FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

# Exploring 3D visualizations for sports data: visualizing the temporal evolution of cross-border player transfers

**Luís Guilherme de Melo Félix Diogo**

U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Mestrado em Engenharia Informática e Computação

Supervisor: Prof. Alexandre Miguel Barbosa Valle de Carvalho

July 11, 2025

# Exploring 3D visualizations for sports data: visualizing the temporal evolution of cross-border player transfers

**Luís Guilherme de Melo Félix Diogo**

Mestrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

President: Prof. Daniel Filipe Martins Tavares Mendes

Examiner: Prof. Paulo Miguel de Jesus Dias

Supervisor: Prof. Alexandre Miguel Barbosa Valle de Carvalho

July 11, 2025

# Resumo

No panorama digital atual, apresentar informação complexa e dinâmica de forma clara e visualmente apelativa é essencial para acomodar diferentes perfis de utilizador e níveis variados de literacia visual. Fundada em 2003, a ZOS, Lda. (ZOS) detém a maior base de dados de futebol do mundo e é um fornecedor de informação de referência para adeptos e profissionais do desporto.

Dada a escala e complexidade dos seus dados, a ZOS enfrenta o desafio de apresentar informação com dimensão espacial, temporal e interconectada, como transferências internacionais de jogadores ou trajetórias de carreira de jogadores, de forma acessível, intuitiva e sem sobrecarregar o utilizador.

Este trabalho estuda e propõe um método para visualizar dados geograficamente distribuídos, em constante evolução temporal e interconectados. A hipótese central é que visualizações 3D podem representar eficazmente este tipo de dados com relações dinâmicas, preservando a componente narrativa da visualização. A solução proposta consiste numa visualização interativa em globo 3D, que apresenta transferências internacionais de jogadores através de arcos animados, permitindo explorar os dados ao longo de vários anos.

Para validar esta abordagem, é desenvolvido um protótipo funcional para a web utilizando Three.js, uma biblioteca JavaScript para gráficos 3D. É realizada uma experiência com participantes de diferentes perfis, que interagem com o protótipo para identificar padrões nos dados, como os principais exportadores e importadores de jogadores, picos temporais de atividade no mercado de transferências, e trajetórias específicas na carreira de jogadores entre países e clubes. Os participantes respondem a questões baseadas em tarefas específicas e preenchem um questionário System Usability Scale (SUS) para avaliar a compreensão, eficácia e usabilidade geral da solução.

Os resultados demonstram que visualizações baseadas em globos 3D são eficazes na representação de dados futebolísticos com dimensão espaciotemporal, oferecendo uma forma envolvente e intuitiva de explorar padrões complexos de transferências internacionais, validando assim a hipótese proposta. No entanto, o estudo revela também áreas importantes a melhorar, nomeadamente a desorganização visual causada pela sobreposição de arcos e a escalabilidade limitada. Desenvolvimentos futuros devem focar-se na melhoria dos mecanismos de filtragem, na responsividade para dispositivos móveis e na integração de elementos 2D complementares, como tabelas, gráficos ou painéis de resumo, para apoiar a interpretação detalhada dos dados e facilitar a transição entre exploração visual e análise.

**Palavras-chave:** Visualização de Informação, Informação Histórica, Infografias 3D, Relações Geoespaciais, Transferências de Futebol

**Classificação ACM:** Computação centrada no ser humano → Visualização → Domínios de aplicação de visualização → Visualização de Informação

# Abstract

In today's digital landscape, presenting complex and dynamic information in a clear and visually engaging way is essential to accommodate diverse user profiles and varying levels of visual literacy. Founded in 2003, ZOS, Lda. (ZOS) hosts the world's largest football database and serves as a key information provider for sports fans and professionals.

Given the scale and intricacy of its data, ZOS faces the challenge of presenting spatial, temporal, and interconnected information, such as cross-border player transfers or player career paths, in a way that is accessible, intuitive, and not overwhelming to users.

This work studies and proposes a method for visualizing geographically distributed, time-evolving, and interconnected data. The central hypothesis is that 3D visualizations can effectively represent geographically distributed, time-evolving data with dynamic relationships, while preserving the narrative within the visualization. A solution is proposed in the form of an interactive 3D globe-based visualization that presents cross-border player transfers using animated arcs, with the ability to explore data across multiple years.

To validate this approach, a web-based functional prototype is developed using Three.js, a JavaScript 3D graphics library. An experiment is conducted involving participants with varied backgrounds, who interact with the prototype to identify patterns in the data, such as the main exporters and importers of football players, temporal peaks in transfer activity, and specific trajectories in a player's career across countries and clubs. Participants complete task-based questions and a System Usability Scale (SUS) questionnaire to evaluate user comprehension, effectiveness, and overall usability.

The results demonstrate that 3D globe-based visualizations are effective for representing spatio-temporal football data, offering a compelling and intuitive way to explore complex international transfer patterns, thus validating the proposed hypothesis. However, the study also reveals key areas for improvement, particularly regarding visual clutter from overlapping arcs and limited scalability. Future developments should focus on improved filtering mechanisms, enhanced mobile responsiveness, and the integration of complementary 2D elements, such as tables, plots, or summary panels, to support detailed data interpretation and facilitate smoother transitions between exploration and analysis.

**Keywords:** Information Visualization, Historical Information, 3D Infographics, Geospatial Relationships, Football Transfers

**ACM Classification:** Human-centered computing → Visualization → Visualization application domains → Information visualization

# UN Sustainable Development Goals

The United Nations Sustainable Development Goals (SDGs) provide a global framework to achieve a better and more sustainable future for all. Comprising 17 goals, they aim to tackle pressing global challenges such as poverty, inequality, climate change, and access to education and innovation.

Although this dissertation does not directly address environmental or humanitarian issues, it contributes to specific SDGs by promoting technological innovation, digital inclusion, and visual literacy. The development of an accessible web-based prototype for visualizing complex spatio-temporal football data supports improved digital infrastructure and equitable access to knowledge and data interpretation tools.

The specific Sustainable Development Goals addressed are as follows:

**SDG 4** Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

**SDG 9** Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

**SDG 10** Reduce inequality within and among countries

| SDG | Target | Contribution | Performance Indicators and Metrics |
|---|---|---|---|
| 4 | 4.4 | Enhances digital literacy and data comprehension by providing interactive and exploratory visualizations accessible to a wide range of users. | Percentage of users reporting improved understanding of football data; Grade B in SUS score. |
| 9 | 9.5 | Fosters innovation in sports data visualization by developing an interactive 3D tool. | Successful development and deployment of a web-based 3D prototype. |
| | 9.C | Expands access to complex data via a responsive, web-friendly interface designed for multiple devices and user skill levels. | Device compatibility; Free to use. |
| 10 | 10.2 | Promotes inclusivity by enabling users with diverse technical backgrounds to intuitively explore and interpret sports data. | Participant diversity across demographics and self-assessed football and data visualization knowledge levels. |

# Agradecimentos

Mais difícil do que aceitar que o meu percurso enquanto estudante chega agora ao fim, é conseguir traduzir por palavras aquilo que estes anos significaram para mim. Foi uma jornada linda, cheia de altos e baixos, mas sempre rodeado de pessoas incríveis às quais agradeço aqui.

Em primeiro lugar, ao meu professor e orientador, Alexandre Valle de Carvalho, um muito obrigado por toda a disponibilidade, ajuda e orientação que me deu. Foi mesmo um prazer poder contar consigo e obrigado também pela confiança que senti que depositou em mim.

Ao Marco, ao Pedro, e a toda a gente com quem tive o prazer de privar nas instalações da ZOS, obrigado por todo o apoio que me deram. Obrigado por me terem acolhido e integrado tão bem no vosso meio. Não havia mesmo melhor ambiente possível para realizar esta dissertação.

Aos meus pais, obrigado pelo vosso total apoio e principalmente pela vossa enorme paciência, especialmente nestes últimos 7 anos. Tenho a certeza que foram bem mais longos para vocês do que para mim.

A todos os meus avós, mas em particular aos dois que me viram crescer mais de perto. Em 2002, escrevia a minha mãe na sua tese de doutoramento: "*Aos meus pais, por tantas horas em que, além de avós, foram pai e mãe do meu Guilherme.*" Já não vão a tempo de ler esta mas espero que soubessem que muito do que sou hoje se deve a vocês, e vou estar para sempre grato por isso.

Aos meus amigos de sempre, por todas as vezes em que precisei de vocês e estavam lá para mim, apenas à distância de uma chamada ou de uma ida ao Garden.

À AEFEUP, à Logística, e a todos os grandes amigos que daí surgiram, por 4 anos fantásticos de associativismo que me deram tanto. Não teria sido a mesma coisa sem vocês.

Por fim, àqueles que me permitiram fazer parte daquela que hoje é a minha casa, a FEUP. Já não bastava a ESIC ter-me dado alguns dos meus melhores amigos, ainda pôs no meu caminho professores fantásticos que me levaram a estar onde estou hoje. Prof. Teresa Carmo, Prof. Alda Roma, Prof. Paulo Piscinas, entre muitos outros... obrigado por tudo o que me ensinaram e, acima de tudo, por me fazerem gostar de aprender. Prof. António Rocha, não me esqueci de si, e é a si a quem devo o maior agradecimento. OBRIGADO por me ter feito apaixonar por Biologia e, consequentemente, acabar por passar 1 ano em Bioengenharia antes de ir parar a Informática. Sem isso, não teria conhecido o grupo mais fantástico de pessoas que a faculdade me podia ter trazido. A estes, não agradeço, mas sem todos eles, sejam mais novos, mais velhos, ou da fantástica casta de 018, tenho a certeza que não estaria aqui hoje a finalizar este meu percurso. Não estaria aqui porque, naquelas vezes em que o caminho mais fácil parecia ser desistir, foi a vocês que me agarrei para ganhar vontade de continuar. Não estaria aqui porque um rapaz de Metal e outro de Bio não se teriam conhecido, mudado os dois para Informática, e feito o seu percurso juntos até ao fim. Mas acima de tudo, não estaria aqui porque eu não seria eu. Seria outra pessoa qualquer, completamente diferente, para quem estes anos não teriam significado o que significaram.

Guilherme Diogo

*"Mamba out."*

Kobe Bryant

# Contents

# List of Figures

# List of Tables

# Listings

# Abbreviations and Symbols

| | |
|---|---|
| 3D | Three-dimensional |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| 2D | Two-dimensional |
| DOSNES | Doubly Stochastic Neighbor Embedding |
| EDM | Equilibrium Distribution Model |
| MISM | Matrix of Isosurface Similarity Maps |
| API | Application Programming Interface |
| JSON | JavaScript Object Notation |
| HTML | HyperText Markup Language |
| CSS | Cascading Style Sheets |
| ID | Identifier |
| URL | Uniform Resource Locator |
| USA | United States of America |
| SUS | System Usability Scale |
| UK | United Kingdom |
| 1-1 | Country-to-Country |

# Chapter 1

# Introduction

Users' attention spans keep getting shorter [1] as the amount of available information increases. This urges the need to present content in a user-friendly way with enough clarity for everyone to understand, regardless of their level of knowledge. Taking this into account, companies end up with the question of how to present information to their users in a compelling and more captivating format. This becomes a daily challenge for ZOS, Lda., a company that manages a substantial volume of complex and historical data that must be presented attractively and understandably. Multiple innovative techniques have been developed worldwide to present data across several visual formats, making it extremely important to choose the most effective one depending on the characteristics of the data and the context where it will be applied.

## 1.1 Context and Motivation

Founded in 2003, ZOS, Lda. (ZOS) is an information provider for sports fans and professionals [2]. Its vast collection includes data from various sports over the years and is constantly being updated and improved, always ensuring accuracy and trustworthiness. In terms of football, ZOS has the world's largest football database [2], containing information about the history of football, national teams, players, and clubs of major, minor, and regional leagues. This level of detailed information can only be obtained through constant collaboration with multiple institutions and the contributions of some enthusiastic supporters, making the company one of the leaders in the industry.

Given the extensive and diverse nature of the collected data, it is essential to present it efficiently to users [3]. This leads ZOS to focus on finding new methods and techniques to deliver its vast sports data effectively. The use of 3D infographics is one of the emerging approaches to do it, offering a dynamic and interactive way of visualizing complex data. These tools provide an immersive experience that makes understanding large volumes of data and their relations easier. By exploring and implementing 3D infographics, ZOS aims to present its extensive historical data in a more captivating and informative format, able to improve its user engagement.

## 1.2 Problem Identification

Sports data's extensive and intricate domain poses a distinctive challenge in creating informative, clear, and comprehensive visual representations. Users have different profiles and comprehension levels and increasingly reduced attention spans [1]. This diversity exacerbates the difficulty of designing an infographic that is both informative and accessible to a broad audience while being captivating [4]. This work focuses on exploring the available tools and technologies for the creation of 3D infographics for the visualization of the temporal evolution of cross-border player transfers. Two main dimensions can be considered for this problem:

- The volume and direction of player transfers between football clubs in different countries vary over time, reflecting shifts in global football dynamics, market behavior, and league composition.
- The number of countries in the world changes over time, either by new countries being formed or existing countries ceasing to exist.

Capturing and presenting inherently dynamic data, such as the temporal evolution of cross-border player transfers, is complex. The challenge lies in accurately reflecting the fluctuations in the volume and direction of player transfers between football clubs in different countries and maintaining visual coherence as the data evolves over time [5]. The dynamic nature of player transfers creates intricate interdependencies between countries, clubs, and players.

This interconnected and time-evolving data must be represented in a way that balances clarity with detail [3], ensuring that the visualizations are informative without overwhelming the user. In this context, exploring advanced 3D visualization techniques capable of representing geographical and temporal dimensions becomes essential, ensuring that the visual representation remains coherent, attractive, and accessible to a broad audience [4]. Balancing the need for detailed information with visual clarity is the core issue addressed in this work.

## 1.3 Hypothesis

Given the previous problem, this work formulates the following hypothesis:

*3D visualizations effectively represent geographically distributed, time-evolving data with dynamic relationships while preserving the narrative within the visualization.*

Three-dimensional visualizations are well-suited for representing geographically distributed and time-evolving data. By incorporating dynamic elements such as animations or interactive layers, these visualizations allow for a more precise depiction of changing relationships between entities over time. The spatial and temporal dimensions are seamlessly integrated, making it easier for users to comprehend complex, dynamic data sets without losing the narrative thread. This method enhances user engagement by offering an intuitive and visually compelling way to explore the evolution of data across different regions and time periods.

## 1.4 Research Questions

The hypothesis defined in the previous section can be divided into the following research questions:

- **Q1**: Which 3D visualization techniques are commonly used to present time-evolving interconnected data?
- **Q2**: Which visualization techniques are commonly used to present time-evolving data?
- **Q3**: Which 3D visualization techniques using a globe are available to represent the relationships between players, countries, clubs, and leagues, and how do these techniques accommodate temporal variations?
- **Q4**: What are the challenges and limitations of using 3D visualizations for representing time-evolving data in a global sports context?

## 1.5 Methodology

To validate the hypothesis presented in Section 1.3, this dissertation follows a structured methodology composed of four main stages, each supporting the design, implementation, and evaluation of the proposed solution.

Initially, a **Systematic Literature Review** is conducted using the PRISMA methodology [6] to identify, filter, and analyze relevant publications on visualization techniques for interconnected and time-evolving datasets. The review focuses on approaches such as 3D representations, globe-based visualizations, and temporal transitions.

Following this, a **State-of-the-Art** is written to complement the literature review. This step aims to synthesize the reviewed works into a cohesive narrative, addressing current trends and innovations in visualizing spatial-temporal-network data. It highlights best practices and design challenges across sports data, 3D visualizations, and time-evolving storytelling, thus laying the groundwork for an informed design process.

Then, based on the insights obtained, a **solution** is defined to address ZOS's need to visually represent cross-border football player transfers in a clear, engaging, and interactive way. This proposed solution is then replicated into a functional **prototype**. The development process is carefully documented, outlining the chosen technologies and explaining how they are applied to implement the visualization.

Finally, an online **experiment** is conducted to evaluate the solution's usability, interpretability, and overall effectiveness. This stage includes defining the experiment protocol and detailing the execution. The results are then analyzed and discussed to extract valuable conclusions.

## 1.6 Document Structure

This dissertation is organized into six chapters in addition to the introduction. This section provides a brief overview of each of those chapters.

Chapter 2 presents a Systematic Literature Review that explores existing visualization techniques for interconnected and time-evolving data.

Chapter 3 analyzes the state-of-the-art in information visualization, focusing on sports data, interconnected data, 3D visualization, and the integration of spatial and temporal dimensions.

Chapter 4 defines the proposed solution and describes the development of the prototype, outlining the design decisions, technologies used, and implementation process.

Chapter 5 details the experiment used to evaluate the prototype, including the participant selection, evaluation metrics, and data collection methods.

Chapter 6 analyzes the results obtained during the experiment, discussing user performance, usability metrics, and qualitative feedback.

Chapter 7 summarizes the key findings of the dissertation, discusses its limitations, and outlines potential directions for future work.

# Chapter 2

# Literature Review

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach was used to conduct a systematic literature review [6, 7]. This methodology provides a complete and meticulous way of reviewing the existing literature and, therefore, offers a good foundation for addressing the previously identified problem.

This systematic literature review used EndNote, a reference management system, to simplify the process. From now on, the following nomenclature will be used:

- **Database -** digital library, source of the extracted data;
- **Search Term -** single term or word used on a search string;
- **Search String -** full expression, made of search terms combined with operators, used to search for records;
- **Search Records -** results of a search string;
- **Papers -** final documents resulting from the literature review process.

## 2.1 Research Domain

Data visualization's constant and rapid evolution creates an urgent demand for new approaches and techniques. Due to the increasing complexity, size, and user demands regarding information, new challenges are faced daily. As a result, research has increasingly focused on creating new visual representations that are both accessible and rich in detail, balancing user needs with the intricacies of the data.

This study aims to find new methods for visualizing highly connected sports data with temporal progression using 3D visualization techniques and to determine their strengths, weaknesses, and possible uses in this area. To explore 3D infographics that accurately depict the dynamics of player movement across time, a methodical examination is performed following the PRISMA methodology. The PRISMA flow diagram in Figure 2.1 provides a comprehensive overview of the systematic methodology used to select the most relevant literature for this study. The process follows four key stages:

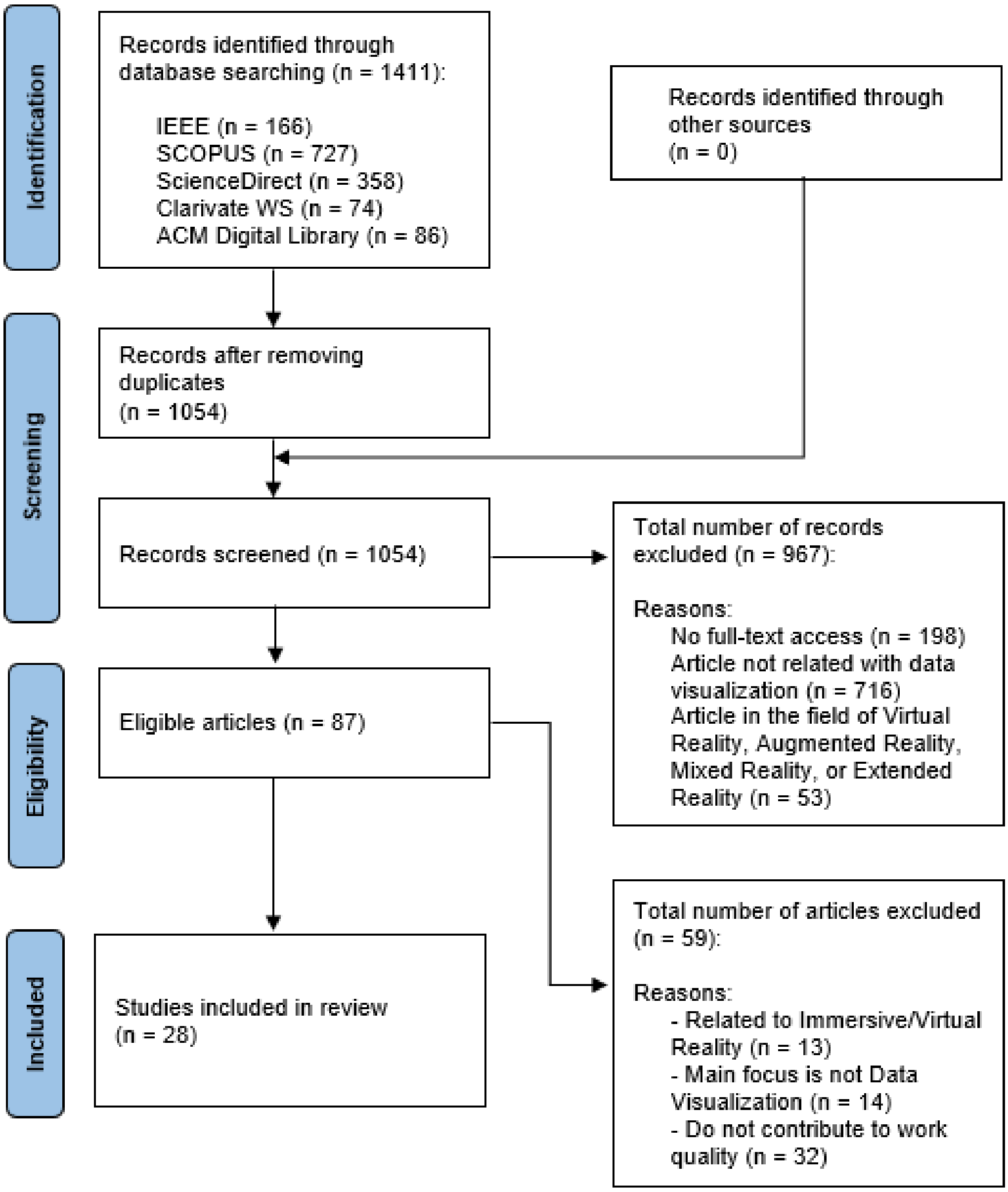



Figure 2.1: PRISMA flowchart.

- **Identification:** The initial step involved searching through five major digital libraries: IEEE Xplore (166 records), SCOPUS (727 records), ScienceDirect (358 records), Clarivate Web of Science (74 records), and ACM Digital Library (86 records). A total of 1,411 records were retrieved through database searching. No additional records were identified from other sources.
- **Screening:** After removing duplicates (357 records), the remaining 1,054 records underwent a thorough screening process. This step focused on titles, abstracts, and keywords to

assess relevance to data visualization. At this stage, 967 records were excluded for various reasons:

  - Lack of full-text access (198 articles).
  - Articles not related to data visualization (716 articles).
  - Articles primarily focused on Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), or Extended Reality (XR) (53 articles).

- **Eligibility:** The 87 relevant articles were further reviewed for full-text eligibility. During this stage, 59 articles were excluded due to:
  - Being related to immersive/virtual reality (13 articles).
  - Its primary focus is not on data visualization (14 articles).
  - Lack of contribution to work quality (32 articles).
- **Inclusion:** Finally, 28 articles were selected as the most relevant and high-quality studies for the research objectives. These studies focus on techniques and methods that align closely with the challenges of visualizing highly connected and time-evolving sports data using 3D approaches.

## 2.2 Identification Phase

A comprehensive search was performed across several scientific digital libraries to gather information on the specified domain. This search encompassed a range of academic papers, including journal articles, conference proceedings, and books.

Five major digital libraries were selected as the primary databases for this research: IEEE Xplore, SCOPUS, ScienceDirect, Clarivate Web of Science, and ACM Digital Library. After selecting these platforms, keywords were carefully defined and refined to ensure relevant results. While initial broad search terms produced a high volume of papers, they were adjusted to incorporate more specific keywords, enabling the identification of the most pertinent studies. Table 2.1 overviews the initial search strings used.

Table 2.1: Initial search strings.

| | |
|---|---|
| SS1 | (geographical OR geospatial) AND (data OR information) AND "3D visuali?ation" |
| SS2 | (temporal AND (data OR information) AND visuali?ation) |
| SS3 | (interconnected OR relational) AND (data OR information) AND visuali?ation |
| SS4 | (sports OR football OR soccer) AND (data OR information) AND visuali?ation |

Using the wildcard operator (?), the search strings consider the different spellings of the word "visualization" in British and American English. Since the words "data", "information", and "visualization" appear in multiple contexts, all four search strings are built to provide more specific results by specifying some of the data characteristics related to this particular context.

### 2.2.1 Refinement Process

The IEEE Xplore library was chosen to refine the search strings in Table 2.1. Each of the four strings was subjected to a similar process, described in the following sections.

#### 2.2.1.1 SS1 Refinement Process

The search string *(geographical OR geospatial) AND (data OR information) AND "3D visuali?ation"* retrieved only 17 search records. In the next version of this search string, *(3D AND (geographical OR geospatial) AND (data OR information) AND visuali?ation)*, the search term "3D" was moved to the beginning of the expression, and the total number went up to 75 search records. Despite the increase in the number of search records, the content of these was out of the desired context. The search terms "geographical" and "geospatial" were the motive for this misalignment, and they ended up being dropped, defining the final search string as *(data OR information) AND "3D visuali?ation"*, retrieving a total of 121 search records.

#### 2.2.1.2 SS2 Refinement Process

The search string *(temporal AND (data OR information) AND visuali?ation)* initially retrieved 8320 search records. This overflowing number of search records may be caused by using the search term "temporal", which is probably too generic to be used. The search string was changed to *(time-evolving OR time-varying) AND (data OR information) AND visuali?ation*, returning 779 search records. As there were still many results, the search string was refined one last time to specify even more of its search. The final search string *(chronological OR temporal) AND (evolution OR progression) AND (data OR information) AND visuali?ation* returned a total of 462 search records (11 after filters were used), resulting in a more relevant set.

#### 2.2.1.3 SS3 Refinement Process

The search string *(interconnected OR relational) AND (data OR information) AND visuali?ation* resulted in a set of 120 search records. Despite being a manageable number of search records, their contents were not what was desired for this search. So the search string was altered to *(interconnected OR relational) AND (data OR information) AND "3D visual?ation"*, to encompass 3D visualizations. This retrieved only 4 search records, which were still unsuited for this study. The search terms "interconnected" and "relational" were substituted by the search terms "transactions", "transfers", and "movement", providing more specificity and resulting in a more adequate search string. The final search string *(transactions OR transfers OR movement) AND (data OR information) AND "3D visuali?ation"* returned a total of 24 relevant search records.

#### 2.2.1.4 SS4 Refinement Process

The search string *(sports OR football OR soccer) AND (data OR information) AND visuali?ation* returned a total of 71 search records. However, most of these search records were unrelated to football/soccer data. The search term "sports" was removed, resulting in the final search string *(football OR soccer) AND (data OR information) AND visuali?ation*, retrieving 10 search records. All were related to football/soccer, fulfilling the initial goal of this search string.

#### 2.2.1.5 Refinement Process Result

Having concluded the refinement process, the final versions of the search strings to be used are the following:

- **SS1 -** (data OR information) AND "3D visuali?ation"
- **SS2 -** (chronological OR temporal) AND (evolution OR progression) AND (data OR information) AND visuali?ation
- **SS3 -** (transactions OR transfers OR movement) AND (data OR information) AND "3D visuali?ation"
- **SS4 -** (football OR soccer) AND (data OR information) AND visuali?ation

Using this set of search strings, a total of 1411 search records were retrieved across all the databases. The results were imported into the reference management system being used (EndNote). Table 2.2 shows the results specified by the number of records each search string gets from each library.

Table 2.2: Search string results per database.

| | SS1 | SS2 | SS3 | SS4 | Total |
|---|---|---|---|---|---|
| IEEE Xplore | 121 | 11 | 24 | 10 | 166 |
| SCOPUS | 228 | 273 | 161 | 65 | 727 |
| ScienceDirect | 100 | 175 | 70 | 13 | 358 |
| Clarivate Web of Science | 44 | 22 | 7 | 1 | 74 |
| ACM Digital Library | 29 | 47 | 2 | 8 | 86 |

### 2.2.2 Search Criteria

Data, information, and visualization concepts are transversal to multiple areas. As a result, the search records are limited to those published in the last 5 years (since 2019) in the area of Data or Information Visualization. Each database has its specific filters and, therefore, each of the following search criteria is adapted to each one of them:

- **IEEE Xplore -** 2019+ journal articles, conference papers, and books published under the topic of “data visualization”;
- **SCOPUS -** 2019+ articles, conference papers, and books, written in English, in the subject area of “Computer Science” or “Engineering”, under the keyword of "Data Visualization", in the final stage of publication and with full-open access;

- **ScienceDirect -** 2019+ research articles in the subject area of “Engineering” or “Computer Science” with full-open access;
- **Clarivate Web of Science -** 2019+ research articles, written in English, in the research area of “Computer Science” or “Engineering” and free to read (search within Topic - title, abstract, and keywords);
- **ACM Digital Library -** 2019+ research articles (search within Title, Abstract).

## 2.3 Screening Phase

The next step in the PRISMA methodology is the screening phase. The goal is to eliminate duplicate search records and those that fail to match some defined inclusion and exclusion criteria so that the most relevant literature is available at the end of the process.

### 2.3.1 Inclusion and Exclusion Criteria

The previous phase ended with a total of 1411 search records. Therefore, defining rigorous inclusion and exclusion criteria is essential to filter the results. Table 2.3 indicates all the exclusion criteria, while Table 2.4 includes all the inclusion criteria. Each search record was subjected to the criteria in a phased approach.

Table 2.3: Exclusion Criteria

| **Phases** | **Exclusion Criteria** |
|---|---|
| 1 | Duplicated, retracted, or work-in-progress |
| 5 | Published with the keywords of Virtual Reality, Augmented Reality, Mixed Reality, or Extended Reality |

Table 2.4: Inclusion Criteria

| **Phases** | **Inclusion Criteria** |
|---|---|
| 2 | Published in the English language |
| 3 | Full-text access |
| 4 | Published with the keyword of Data visualization |

The established criteria are designed to ensure the selected literature is relevant to the research objectives. To further refine the selection, each of the remaining search records will be subjected to a detailed analysis of its components (title, abstract, introduction, etc.) to confirm whether they are useful and relevant.

### 2.3.2 Duplicates Removal

All duplicates are removed using EndNote's built-in function "Find Duplicates", resulting in a set of 1054 unique records and 357 duplicates.

### 2.3.3 Screened Records

Out of the 1054 unique records, 967 failed to match some of the criteria defined in Table 2.4, or matched one or more of the criteria defined in Table 2.3. This resulted in their exclusion, leaving 87 eligible papers remaining.

## 2.4 Eligibility Phase

The following phase of the PRISMA methodology, the eligibility phase, focuses on individually analyzing the remaining 87 papers to determine each paper's relevance to the research objectives. The result of this phase is the final list of papers to consider for this research.

### 2.4.1 Data Extraction

The relevant information to be extracted was carefully determined at this stage to ensure a thorough analysis of the selected papers. Data extraction aims to systematically gather consistent and relevant information from each study, enabling a structured comparison and evaluation. This process ensures that all papers are assessed equally and that no critical aspect of the research objectives is overlooked. Specifically, the extracted data will allow us to:

- Identify the key characteristics of each study, such as its domain, publication type, and context.
- Understand the visualization techniques and tools applied, their applications, and how they align with the identified research gaps.
- Evaluate the contributions of each paper to determine its relevance and applicability to the proposed solution.
- Ensure traceability and replicability of the review process by systematically documenting essential information.

Table 2.5 outlines the specific data extracted from the papers.

Table 2.5: Information extracted from papers.

| Category | Data |
|---|---|
| Basic Information | Title<br>Author(s) |
| Publication Information | Source<br>Year<br>Quartile<br>Type |
| Research Information | Application Domain<br>Research Purpose<br>Used visualization techniques and tools<br>Relevance and Contribution to the proposed solution |

By extracting this information, we ensure a comprehensive understanding of the current state of the art, identify trends, and establish a strong foundation for designing an innovative solution that addresses the gaps and challenges in visualizing interconnected and time-evolving data using 3D techniques.

### 2.4.2 Included Records

In this stage, a complete analysis of the full text of the papers was performed, storing the data described in Table 2.5. After this information was gathered, the relevant papers were grouped by domain. Tables 2.6, 2.7, 2.8, 2.9, and 2.10 summarize the grouped papers. Each table categorizes the records based on their respective domains, highlighting the diversity of approaches and tools relevant to the research goals.

From the initial 1411 papers, only 28 remained after this process and were considered relevant. These remaining studies are grouped into five domains, as discussed below.

The first domain focuses on sports data visualization, emphasizing the use of visual techniques tailored to sports metrics and analyses, as shown in Table 2.6. These papers explore a range of methods, from interactive applications for basketball metrics to comparative approaches in sports data visualization, and methodologies for uncovering insights in virtual football datasets. These studies directly contribute to understanding the unique requirements and challenges of visualizing sports-related information.

Table 2.6: Papers related to Sports Data.

| Domain | Research Purpose | Year | Reference |
|---|---|---|---|
| Sports Data Visualization | To create an interactive application using R-Shiny for visualizing basketball shooting areas and related metrics | 2022 | [8] |
| Sports Data Visualization | To compare and analyze different sports data visualization techniques and approaches | 2022 | [9] |
| Dataset Analysis and Exploration | To uncover insights and patterns that can enhance the understanding of the virtual football world | 2024 | [10] |

The second domain addresses interconnected data, particularly networks and relationships between entities. Table 2.7 includes studies introducing tools like Hermes for exploring economic networks, real-time frameworks for tracking COVID-19 dynamics, and visualizing dissemination patterns. These studies demonstrate the utility of network-based approaches to reveal patterns and relationships in large datasets. These can be adapted to visualize complex interactions in sports data, such as player transfers and the relations between leagues.

Table 2.7: Papers related to Interconnected Data.

| Domain | Research Purpose | Year | Reference |
|---|---|---|---|
| Network Visualization and Exploration Tool | To introduce Hermes, a guidance-enriched Visual Analytics environment designed to explore intricate economic networks | 2020 | [11] |
| Real-time Data Visualization Framework | To develop a real-time visualization framework that captures the dynamics of COVID-19 spread facilitated by human movement | 2021 | [12] |
| COVID-19 Dissemination Tracing and Analysis | To introduce COnVIDa, a web-based tool designed to compile and visualize multidisciplinary data related to the COVID-19 pandemic | 2021 | [13] |
| COVID-19 Dissemination Tracing and Analysis | To understand COVID-19 dissemination patterns and provide insights into patterns and relationships involved in the disease transmission | 2022 | [14] |

The third domain explores 3D data visualization techniques, as shown in Table 2.8. These studies cover topics like open-source tools for antenna visualization, hybrid 2D-3D approaches for air traffic management, and the use of 3D globe models for visualizing global data connections. These techniques provide a foundation for applying 3D visualizations to sports data, particularly for presenting spatial and contextual relationships in an engaging and interactive way.

Table 2.8: Papers related to 3D Data.

| Domain | Research Purpose | Year | Reference |
|---|---|---|---|
| Open-source visualization tools | To evaluate and compare open-source visualization tools available for antenna engineers | 2023 | [15] |
| Marine Geological Data Visualization | To develop a marine geological data resource-sharing platform that enhances the accessibility and utilization of marine geoscience data | 2020 | [16] |
| Hybrid 2D-3D Visualizations | To evaluate the effectiveness of the hybrid 2D-3D radar display in improving air traffic management | 2021 | [17] |
| 3D Data Visualization | To use 3D visualizations to display movie data on web pages through the use of WebGL and Three.js | 2021 | [18] |
| 3D Digital Earth Model | To solve the display of 3D earth models in web browsers | 2023 | [19] |
| High-Dimensional and Graph Data Visualization | To enhance visualization techniques by adapting the doubly stochastic neighbor embedding (DOSNES) method to spherical manifolds | 2019 | [20] |
| 3D Scientific Visualization | To introduce Mayavi, an open-source, versatile 3D scientific visualization package designed to offer accessible and interactive tools for data visualization | 2024 | [21] |

The fourth domain focuses on time-evolving data visualization, with papers summarized in Table 2.9. These studies explore dynamic tools like ALDO, EventsVista, and Circular to address the challenges of visualizing temporal changes and event-based networks. Their relevance extends to visualizing sports data, where player movements and team dynamics are inherently time-dependent.

Table 2.9: Papers related to Time-Evolving Data.

| **Domain** | **Research Purpose** | **Year** | **Reference** |
|---|---|---|---|
| Digital Framework for E-Learning | To introduce ALDO, an advanced digital framework to support integrated facilities for effective and active e-learning | 2020 | [22] |
| Event Visualization and Interpretation | To introduce EventsVista, a data visualization software to display and analyze events based on the assessment of previous data visualization methods | 2023 | [23] |
| Event-based Networks Visualization | To introduce Circular, an interactive exploration environment designed to visualize event-based networks | 2021 | [24] |
| Trajectory Data Visualization | To enhance the understanding of complex trajectory datasets by proposing a framework that interprets attributes in both geographical and abstract spaces | 2019 | [25] |
| Scientific Data Visualization | To develop a scalable and adaptable method for producing narrated documentaries that effectively convey molecular structures and processes | 2023 | [26] |
| Sentiment and Stance Visualization | To explore how uncertainty is represented in visualizations of sentiment and stance data | 2023 | [27] |
| Temporal Data Visualization | To create an intuitive and effective method for visualizing temporal data by introducing a 3D perspective to traditional 2D plots | 2019 | [28] |
| Scientific Data Visualization | To introduce MISM, an intuitive tool to identify significant isosurfaces and compare them across spatial, temporal, and value dimensions | 2019 | [29] |
| Dynamic Data Visualization | To introduce COVERAGE Data Viewer, a web-based tool that enables integrated and dynamic visualization of various satellite and in-situ data types | 2023 | [30] |
| Bipartite Networks | To introduce Dynamic Bipartite Cluster Flows (Dynamic BiCFlows), an interactive visualization method for displaying large, time-dependent bipartite graphs, combining hierarchical aggregation with filtering using linked lists | 2020 | [31] |

Finally, the fifth domain centers on user-centric design in information visualization, summarized in Table 2.10. These studies focus on enhancing user engagement, perceptual clarity, and interaction. They emphasize the importance of designing accessible and intuitive visualizations, particularly for complex datasets like sports data.

Table 2.10: Papers related to User-Centric Design.

| Domain | Research Purpose | Year | Reference |
|---|---|---|---|
| Visual Analytics | To find the trends, major challenges, and future directions of visual analytics | 2019 | [32] |
| Color Selection in Visualizations | To introduce the Equilibrium Distribution Model (EDM), an approach for automatically selecting colors that provide optimal perceptual contrast in visualizations | 2019 | [33] |
| Enhancing Data Visualization | To enhance the visualization of data through automatic generation of modifications, embellishments, and annotations | 2020 | [34] |
| Collections Exploration | To introduce Touch the Collections, a 3D user interface that visually outlines and interactively exhibits communities and collections | 2023 | [35] |

## 2.5 Critical Analysis

Creating a link between the selected papers and the research questions described in Section 1.4 is essential to evaluate their relevance to this study.

The selected papers provide a comprehensive overview of various approaches to interconnected, time-evolving, and 3D data visualization, aligning with this study's research questions and goals. Papers focused on sports data visualization [8, 9, 10] contribute directly to understanding the different types of sports data, understanding the visualization of temporal data (Q2), and addressing the challenges of representing this type of data in a sports context (Q4). For instance, studies exploring basketball shooting metrics [8] and comparative visualization techniques [9] offer insights into temporal performance visualization, while the analysis of virtual football datasets [10] highlights challenges in handling complex sports data.

Papers related to interconnected data [11, 10, 13, 14] center on techniques for mapping interconnected entities (Q1), such as economic networks [11] and COVID-19 dissemination patterns [10, 13, 14], which are adaptable to visualizing global player movements and relationships (Q3). These studies also provide examples of managing dynamic and dense datasets, offering valuable lessons for addressing similar challenges in sports data visualization (Q4).

In the domain of 3D visualization, seven papers [15, 16, 17, 18, 19, 20, 21] explore techniques relevant to interactive 3D sports infographics. These include methods for designing 3D Earth models [18, 19], aligning closely with globally visualizing player movements (Q3). Additionally, practical challenges related to rendering and interactivity in large datasets (Q4) are addressed through studies on hybrid 2D-3D visualizations [17] and open-source tools [15, 21].

Time-evolving data visualization is covered by ten papers [22, 23, 24, 25, 26, 27, 28, 29, 30, 31], which include frameworks such as ALDO [22] and EventsVista [23] that demonstrate interactive temporal data exploration (Q2). Tools like Dynamic BiCFlows [31] and Circular [24] provide approaches to visualizing relational data over time, directly addressing challenges in representing player transfers and other dynamic sports data (Q4).

Lastly, four papers on user interaction and perception [32, 33, 34, 35] highlight the importance of user-friendly visualization design and its implications for understanding data. The introduction of approaches such as the Equilibrium Distribution Model (EDM) for color selection [33] and interactive interfaces [34, 35] demonstrates practical ways to enhance user experience. Additionally, [32] identifies trends and challenges in visual analytics, offering future directions that align with creating impactful and intuitive visualization systems (Q4). These contributions underscore the value of prioritizing perceptual and usability considerations in designing visualizations for complex datasets.

Overall, the selected papers align well with the research questions. Techniques from 3D data papers [18, 19, 20] and hybrid visualization methods [17] address challenges in interconnected 3D visualization (Q1). Temporal visualization tools [23, 24, 25] provide solutions for representing time-evolving data (Q2). Relationships and mapping techniques [11, 12, 14, 18, 19] contribute to player transfer visualizations worldwide (Q3). At the same time, challenges in complexity and interactivity are addressed across sports [10], time-evolving data [31], and user perception studies [33] (Q4).

This critical analysis confirms that the selected records provide meaningful insights into visualizing interconnected and time-evolving sports data while addressing the study's core research objectives.

## 2.6 Gap Analysis

Despite the breadth of research on data visualization techniques, several gaps remain that are particularly relevant to this study's objectives. These gaps highlight the need for novel approaches and underscore the contributions this research aims to make:

- **Integration of 3D, Temporal, and Network Data Visualization:** While there is significant work on 3D visualization techniques [18, 19, 20], temporal data representation [23, 24, 25], and network visualization [11, 12, 14, 18, 19], there is limited research on integrating these three dimensions into a cohesive visualization framework. Existing studies often focus on one or two of these aspects, leaving a gap in understanding how to effectively combine them for complex, interconnected datasets such as those in sports analytics.

- **Dynamic and Interactive Representations:** Many visualizations in the reviewed literature provide static or minimally interactive data representations. For example, trajectory visualizations [25] effectively show movement patterns but lack advanced interaction mechanisms for users to explore relationships, temporal trends, or granular details. This limits their applicability for tasks requiring detailed analysis or real-time insights.

- **User-Centric Design for Domain-Specific Applications:** Although user-centric design principles are explored in papers like [32], there is a lack of focus on adapting these principles to the specific needs of sports data visualization. Tools that align with the cognitive requirements of analysts, coaches, and fans are still underdeveloped.
- **Storytelling in Visualizations:** The importance of storytelling in visualization is acknowledged [26], yet few studies effectively integrate narrative techniques into visualizations of interconnected and time-evolving datasets. This gap represents an opportunity to create visualizations that inform and engage users by emphasizing the narrative aspect of the data.

## 2.7 Additional Papers

In addition to the 28 selected papers through the execution of the PRISMA methodology, it was decided to include an additional paper that did not meet the initial inclusion criteria. Shneiderman's *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations* [36] was later included due to its foundational nature and its lasting relevance in the field of information visualization.

This influential work presents a classification system of visualization techniques, known as the "*Task by Data Type Taxonomy*", which remains valid and widely used today. The taxonomy provides a structured framework to understand how data can be effectively visualized for various tasks, making it an essential reference for researchers. Its inclusion was deemed critical for writing a comprehensive state-of-the-art, as it offers valuable insights into designing informative and engaging visualizations for users.

## 2.8 Summary

This chapter provides a detailed overview of the literature review conducted, establishing a foundation for the state-of-the-art analysis presented in the next chapter.

The Systematic Literature Review explores visualization techniques and tools for interconnected and time-evolving data, specifically focusing on 3D visualizations and their applications in dynamic sports data. The review systematically examines studies across various domains, including sports data visualization, network analysis, 3D visualization, and time-evolving data visualization. By integrating these perspectives, the review highlights the challenges of representing dynamic relationships, maintaining visual clarity, and ensuring interactivity in complex datasets.

Key findings emphasize the role of 3D visualizations in enhancing data comprehension, particularly in visualizing temporal changes and interconnected relationships, such as those seen in sports contexts. The analysis also underscores the importance of user-friendly design and perceptual elements, such as interactivity and visual storytelling, in making complex data accessible and engaging. This literature review provides valuable insights into the current knowledge of data visualization approaches and their relevance to the specific requirements of representing the temporal evolution of cross-border player transfers.

# Chapter 3

# State of the Art

This chapter provides a comprehensive review of state-of-the-art information visualization techniques, focusing on techniques relevant to sports data, network analysis, 3D visualization, and time-evolving data. Section 3.1 introduces the fundamental role of visualization in transforming complex datasets into actionable insights, while Section 3.2 synthesizes these perspectives by addressing the integration of spatial, temporal, and network dimensions in trajectory visualization. These sections establish a foundation for understanding current challenges and innovations in visualizing complex, evolving datasets.

## 3.1 Information Visualization

In the age of digital transformation, the large volume and complexity of data available across various domains have made information visualization an indispensable tool. Organizations and individuals can now access extensive datasets from digital platforms, scientific research, social media, and online databases. While rich in potential insights, these datasets often pose significant challenges due to their scale, variety, and interconnected nature, making extracting meaningful and actionable information a difficult task [33, 32].

Information visualization bridges raw data and actionable understanding, transforming complex datasets into intuitive visual representations [33, 34]. Doing so helps organize and document data and facilitates exploration, pattern recognition, and analysis [23, 10]. Visualization plays a critical role in enhancing cognitive processes, helping users uncover insights that might remain hidden in textual or tabular formats [24, 27]. This dual role of representation and cognitive facilitation underscores the importance of information visualization in modern decision-making processes.

As data grows increasingly interconnected and dynamic, advanced visualization techniques have become essential, including in the sports domain. Studies have explored innovative approaches such as graph-based representations for relational data [11, 31], temporal visualization tools to display time-evolving patterns [23, 25], and 3D visualization techniques for enhancing spatial and contextual understanding [19, 18]. These methods enable users to navigate through large, complex datasets, revealing relationships, trends, and changes over time [17, 21].

### 3.1.1 Sports Data

Sports data visualization became an indispensable tool in modern sports analytics, offering transformative insights that extend beyond the traditional boundaries of sports management and performance analysis. The evolution of sports data visualization has been influenced by the development of innovative visualization technologies and data analysis techniques [9], enabling more engaging and dynamic ways to represent complex datasets.

As stated by Tamboli *et al.* [10], *"In terms of players and spectators, football is the most popular sport in the world"*. European football clubs were estimated to have generated $27 billion in revenue in 2017, making football a global passion and a vital component of the world economy [10]. The immense popularity and financial stakes associated with football underscore the critical role of data in optimizing performance, enhancing fan engagement, and supporting strategic decisions within the sport.

Key types of sports data include box score data, which captures statistical summaries of events in a game, tracking data, representing continuous spatiotemporal in-game motion, and metadata, which encompasses game-related details such as player profiles, transfers, or stadium information [9]. These data sets are essential for understanding player performance, team dynamics, and game outcomes.

For instance, the R-Shiny application for basketball shooting area visualization [8] developed by Bani *et al.* offers an interactive platform to analyze player performance metrics across court zones, as shown in Figure 3.1. Similarly, the comparative studies on sports data visualization approaches developed by Bhatt *et al.* highlight the strengths and limitations of various techniques in making sports data accessible and engaging [9], such as the traditional scatter, box, and line plots, as shown in Figures 3.2, 3.3, 3.4. The exploration of virtual football worlds provides additional insights into the visualization of interconnected datasets within sports [10].

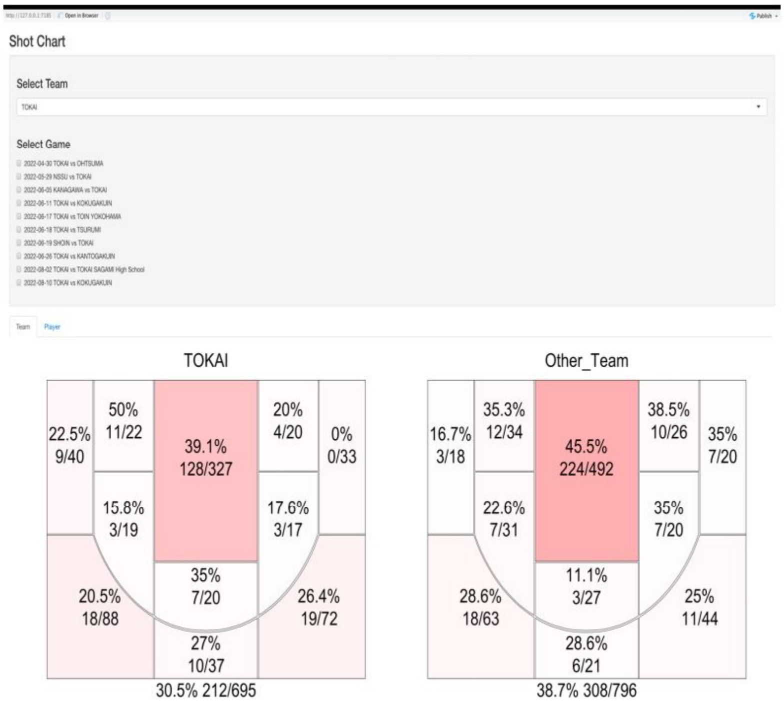



Figure 3.1: Basketball shooting area visualization: The application interface [8].

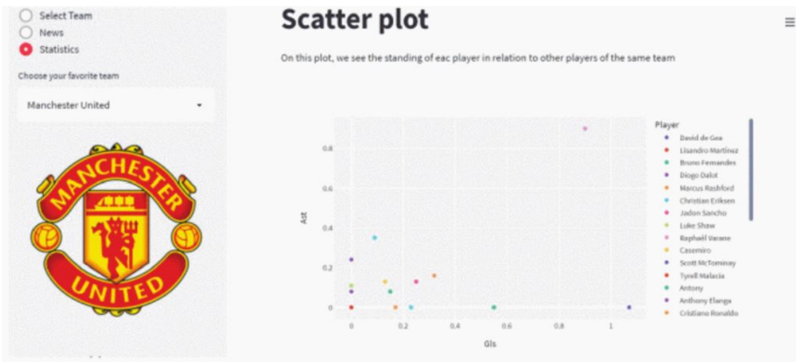



Figure 3.2: Visualization through Scatter Plots [9].

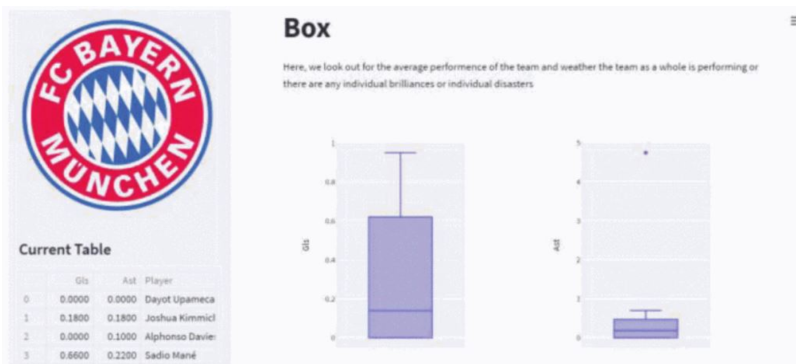



Figure 3.3: Visualization through Box Plots [9].


Figure 3.4: Visualization through Line Plots [9].

Key challenges in sports data visualization include managing the scale and complexity of datasets, providing interactivity tailored to diverse user groups, and effectively representing temporal changes and relationships. These challenges underscore the importance of matching data types and tasks to appropriate visualization techniques, ensuring that users can intuitively interact with and extract meaningful insights from the data [36].

### 3.1.2 Network Data

Network data visualization plays a pivotal role in exploring and understanding complex relationships between interconnected entities. Networks can represent a wide variety of domains, from social and economic interactions to disease dissemination, providing insights into patterns, dependencies, and dynamics. The selected studies highlight different approaches to network visualization, offering valuable lessons for representing interconnected sports data, such as player transfers between different leagues.

Network visualization techniques often focus on representing relationships and patterns within large datasets. In Hermes, developed by Leite *et al.* [11], visualization is achieved through a multi-board system that provides both high-level overviews and detailed network explorations. The "Overview Board", shown in Figure 3.5, incorporates adjacency matrices, parallel coordinates, and map-based visualizations to contextualize regional and sectoral data. Meanwhile, the "Network Board", shown in Figure 3.6, enables detailed pattern-matching tasks using a node-link diagram for query construction and result visualization.

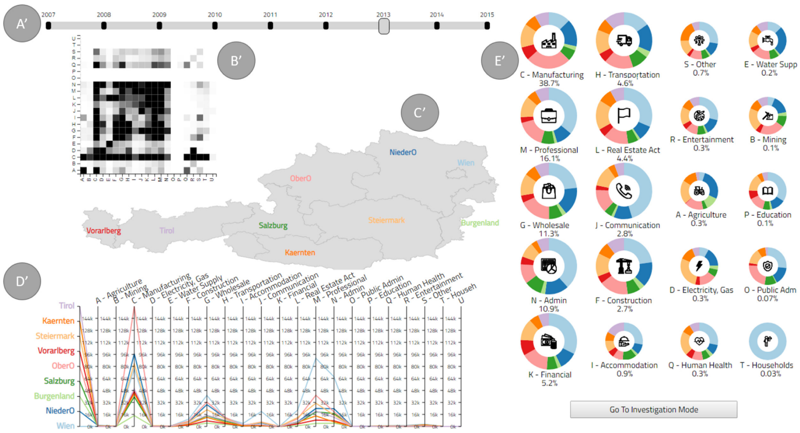


Figure 3.5: Hermes: Overview Board Interface [11].

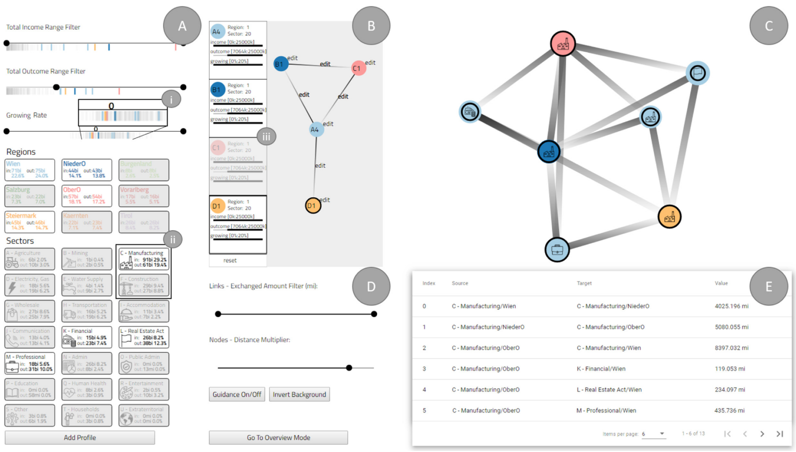


Figure 3.6: Hermes: Network Board Interface [11].

The framework supports tasks such as identifying economic trends, evaluating supply chains, and understanding monetary flows across regions. By integrating interactive and guidance-enriched features, Hermes [11] enables users to navigate complex datasets and extract actionable insights efficiently. These features align closely with Shneiderman's [36] Visual Information-Seeking Mantra: *"overview first, zoom and filter, then details on demand."*, which emphasizes the importance of starting with a more general overview, allowing users to zoom and filter, and then providing details on demand. This approach ensures users can transition seamlessly from macro-level trends to micro-level information.

Additionally, Medeiros *et al.* [14] explore the use of knowledge graph techniques to visualize relationships and dependencies within interconnected systems. Knowledge graphs represent data as nodes and edges, enabling users to gain structural insights into complex datasets. Figure 3.7 illustrates an example of a knowledge graph, where central nodes, clusters, and significant connections are visually highlighted. This approach is particularly effective for emphasizing the structural relationships within the data, allowing users to navigate through dependencies, clusters, and key pathways.

Unlike traditional graphical or temporal visualizations, knowledge graphs abstract the dataset into a network structure, focusing on relational aspects. This abstraction makes them highly effective for analyzing and understanding connections that may otherwise remain hidden in large datasets. For instance, knowledge graphs excel in identifying hubs or communities within networks, which can provide critical insights for decision-making.

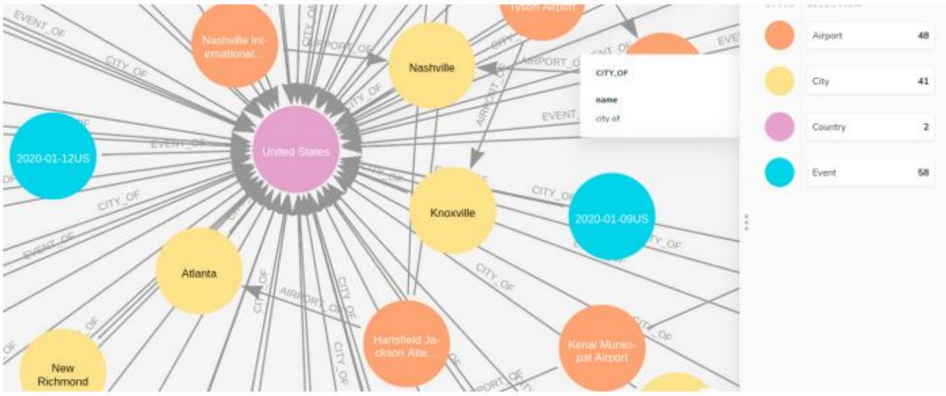



Figure 3.7: Subset of the knowledge graph [14].

However, implementing effective knowledge graph visualizations has a lot of challenges. Key issues include managing the density of datasets or ensuring usability for non-expert users. Addressing these challenges often requires advanced visualization techniques, such as filtering methods to reduce visual clutter, clustering algorithms to group related nodes, and interactive exploration functionalities to allow users to focus on specific areas of interest.

Key challenges in network visualization include managing dense datasets, ensuring scalability across multi-level analyses, and preserving the usability of visual tools for non-expert users. These are addressed by incorporating filtering, clustering, and interactive exploration functionalities [36].

### 3.1.3 3D Visualization

3D visualization plays a crucial role in representing complex, multidimensional data, offering intuitive ways to explore spatial, temporal, and relational information. Its versatility has led to applications across diverse fields such as engineering [15], geosciences [16], air traffic management [17], and scientific research [21]. By incorporating depth, interactivity, and realism, 3D techniques provide richer insights into intricate datasets, making them indispensable for analyzing dynamic and interconnected data.

The effectiveness of 3D visualization tools in engineering is demonstrated through their ability to enhance the interpretation of intricate datasets, as highlighted by Arya *et al.* [15]. Similarly, improvements in data sharing and accessibility in marine geosciences are achieved through 3D visualizations, as shown by Cheng *et al.* [16]. These applications illustrate how tailored 3D solutions address domain-specific challenges effectively.

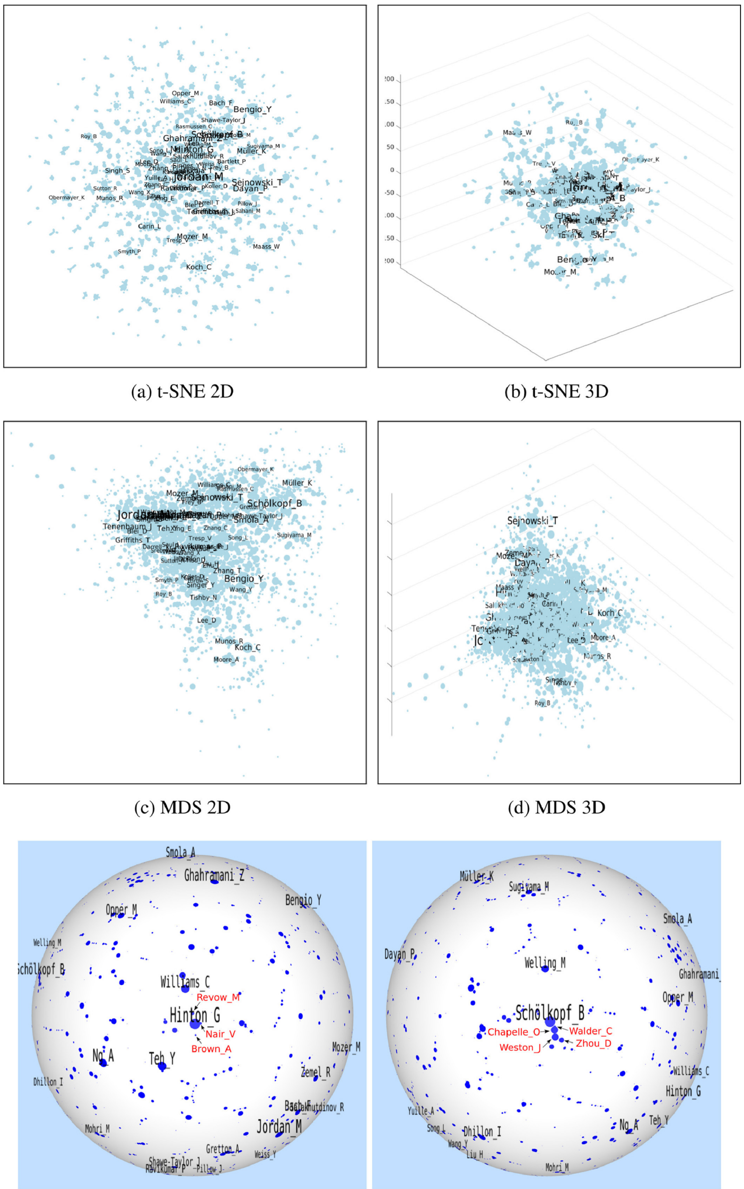
Hybrid 2D-3D visualizations expand the potential of this technology by combining 2D simplicity with 3D depth. Their value in air traffic management, as emphasized by Li *et al.* [17], lies in enabling users to seamlessly switch between high-level overviews and detailed analyses while reducing overlapping problems. This approach aligns with Shneiderman's principle of integrating overview and detail [36], proving particularly effective for applications requiring scalability and precision.

Recent advancements, such as 3D Earth models (Figure 3.8), highlight the continuous evolution of this field. The Digital Earth platform, developed by Li *et al.* [19], visualizes global geospatial data in web browsers and supports dynamic, time-based data displays. This enables users to explore spatio-temporal patterns, offering new opportunities for understanding global relationships and trends.

Figure 3.8: Example of a 3D Earth model [37].

The value of 3D visualization in enriching our capacity to explore, understand, and derive insights is underscored by Trajkovska *et al.* [21]. Tools like Mayavi [21], an open-source scientific visualization platform, exemplify the versatility of 3D visualization for both general and specialized applications. By combining accessibility with interactivity, Mayavi provides an effective resource for exploring large and complex datasets.

For high-dimensional datasets, adapting doubly stochastic neighbor embedding (DOSNES) to spherical manifolds, as proposed by Lu *et al.* [20], provides a method to map data onto spherical representations while maintaining spatial coherence. However, the technique introduces significant visual clutter when representing dense datasets. As shown in Figure 3.9, spherical visualization suffers from overlapping elements, which can obscure relationships and reduce interpretability. This highlights the need for more effective methods to manage clutter and enhance clarity in 3D visualizations.



(a) t-SNE 2D
(b) t-SNE 3D
(c) MDS 2D
(d) MDS 3D
(e) DOSNES (viewpoint 1)
(f) DOSNES (viewpoint 2)

Figure 3.9: Comparison between 2D and 3D views using DOSNES [20].

Key challenges in 3D visualization include scalability and performance, particularly when handling large or dynamic datasets in real time. Visual clutter and occlusion can obscure critical relationships, making it difficult to interpret dense data effectively. Perceptual challenges, such as accurately judging depth and spatial relationships, further complicate user interaction. Additionally, designing intuitive and accessible interfaces for non-expert users remains a persistent issue. Integrating spatial, temporal, and relational dimensions into a cohesive visualization without overwhelming users adds another layer of complexity. Addressing these issues is crucial to fully harnessing the potential of 3D visualization for interconnected and dynamic datasets.

### 3.1.4 Time-evolving Data

Visualizing time-evolving data is essential for understanding dynamic changes and patterns in datasets that evolve over time. By capturing and representing temporal relationships, such visualizations enable users to explore trends, monitor progress, and gain insights into the underlying data processes. The selected studies demonstrate diverse approaches to temporal data visualization, offering valuable perspectives on addressing challenges related to time-dependent datasets.

As stated by Bartolini *et al.* [22], frameworks like ALDO demonstrate how scalable and interactive systems can effectively handle temporal data. ALDO [22] provides a dynamic interface for tracking learning progress in e-learning environments, using advanced data visualization tools like maps, graphs, and timelines to explore temporal and spatial data.

Event-based and historical data visualization approaches have advanced significantly, focusing on uncovering patterns and relationships in time-sensitive datasets. Circular, presented by Filipov *et al.* [24], provides an interactive exploration environment designed for visualizing event-based networks, particularly in historical contexts. Circular enables users to examine relationships between entities, such as individuals, organizations, or events, and observe how these relationships evolve over time. It can highlight shifts in alliances, trends in communication patterns, or the rise and fall of influence within a network. By incorporating interactivity, users can dynamically filter, zoom, or rearrange data to focus on specific time periods or subsets, offering a more detailed understanding and revealing hidden trends and patterns.

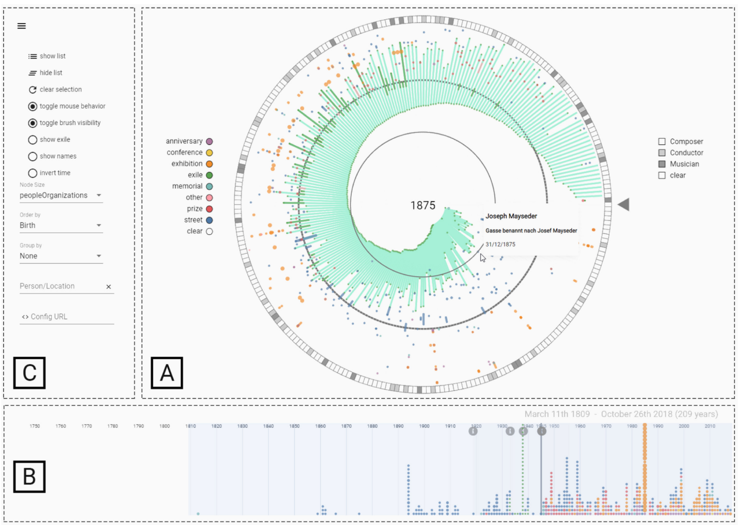



Figure 3.10: Circular: Prototype Overview [24].

Similarly, EventsVista, introduced by Farhat *et al.* [23], focuses on analyzing and interpreting event data through intuitive visual representations, reducing the difficulty of understanding multiple event dimensions, and making it easier to explore complex temporal relationships. In addition to these visual representations, EventsVista [23] provides a *Statistics Page*, as shown in Figure 3.11, to present an extra coverage of all event dimensions.

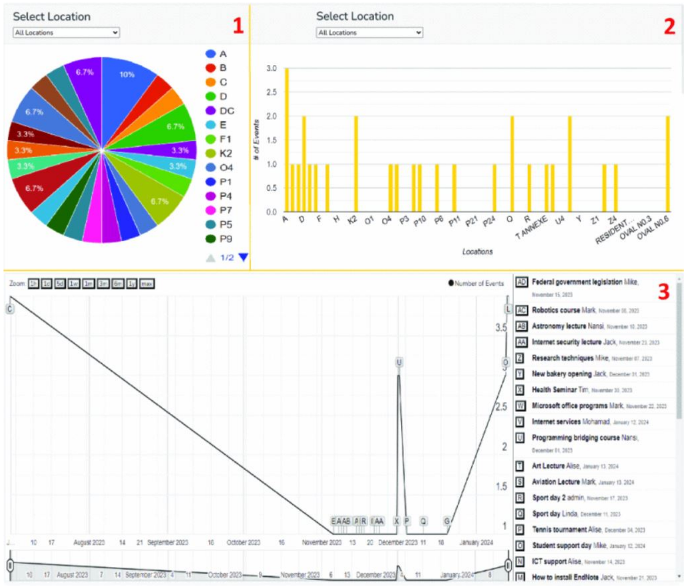



Figure 3.11: EventsVista: Statistics Page [23].

Innovations in 3D visualization have further enhanced temporal data representation. Schmidt *et al.* [28] present Popup-plots, a "2.5D" perspective on traditional 2D plots, offering users an immersive view of temporal changes, allowing them to engage with the data in a single view and enhancing the ability to identify trends and anomalies. This is done by strengthening 2D plots by incorporating temporal information aligned with the current viewing direction. Figure 3.12 shows the differences between this and existing approaches. Similarly, Tao *et al.* [29] present an innovative interface called the matrix of isosurface similarity maps (MISM), which provides tools to analyze significant isosurfaces across spatial, temporal, and value dimensions, illustrating the potential for multidimensional temporal data visualization in scientific research.

| Property | 2D Plot | Popup-plot | 3D Plot, Space-time Cube |
|---|---|---|---|
| Number of views | single | single<br>all aspects of the data can be studied in one view | multiple<br>3D view, additional 2D projections in separate views |
| Coordinate system | Cartesian | ellipsoidal<br>space bent to view time | Cartesian |
| Dimensionality | 2D | 2.5D<br>popup greeting card analogy, 2D frontal view, but increasingly 3D when opening | 3D (2D)<br>3D space-time cube with separate 2D projections along axes |
| Time exploration | projection | rotation | rotation |
| Time encoding | none<br>color or line thickness may be used | distinct (intrinsically)<br>dendrochronology, inner circles resemble past whereas outer ones future | ambiguous<br>axes labels required |
| Transitions | none | continuous<br>ellipsoidal model allows seamless transitions between 2D projections | none<br>static visualization, 2D projection potentially faded in |
| Orientation | distinct<br>viewing and projecting along z-axis | distinct (intrinsically)<br>encoded in orientation of ellipsoids | ambiguous<br>orientation cube required |
| Comprehensibility | good<br>well-known plot | medium<br>requires accomodation time | good<br>z-axis intuitively resembles time |

Figure 3.12: Comparison between Popup-plots and existing approaches [28].

Dynamic data visualization tools like COVERAGE Data Viewer, introduced by Tsontos *et al.* [30], tackle the challenges of large-scale, time-dependent datasets. COVERAGE integrates satellite and *in-situ* data to present a comprehensive view of temporal changes. This tool stands out for its innovative features, shown in Figure 3.13, like tightly synchronized horizontal and vertical views of data evolution over time, along with an interactive "one-stop" capability for data subsetting, and a flexible temporal control interface that supports animations and step-by-step temporal navigation. COVERAGE also addresses data interoperability challenges by filtering datasets by spatial, temporal, and variable criteria. It promotes collaborative usage by enabling users to share interactive visualization configurations via URL links, making it a powerful example of how dynamic visualization frameworks can enhance understanding and engagement with complex, time-evolving datasets.

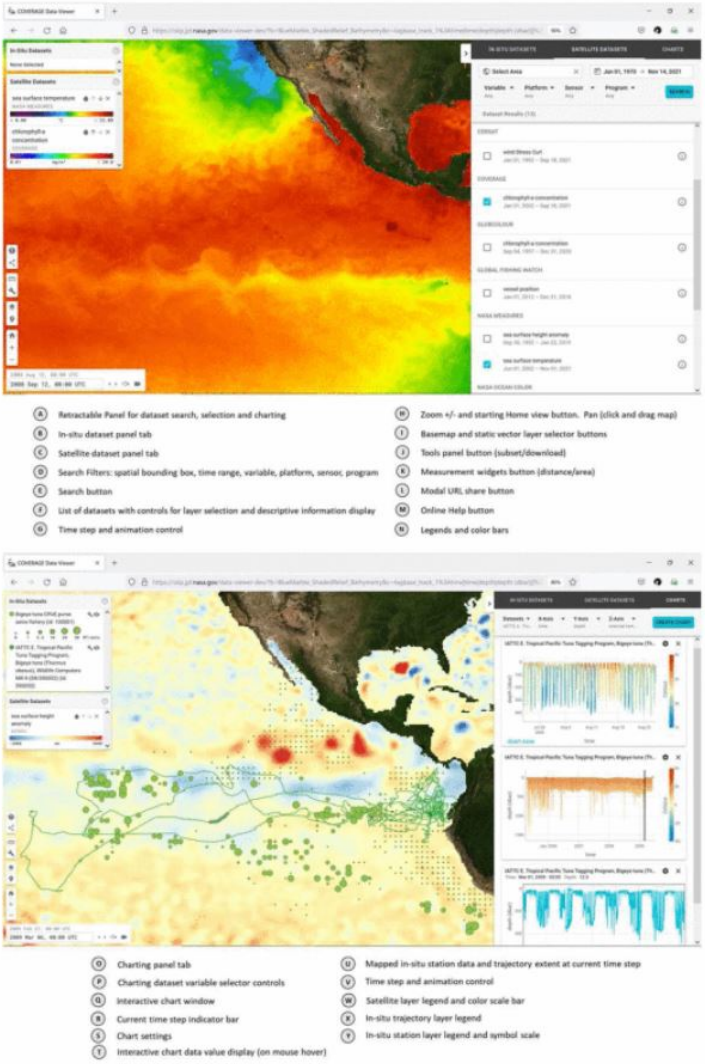



Figure 3.13: COVERAGE: User interface [30].

Waldner *et al.* [31] tackled the challenges of visualizing dynamic data by developing Dynamic BiCFlows, an interactive method tailored for exploring large, time-dependent bipartite graphs. Bipartite graphs, which consist of two distinct sets of nodes connected by edges, often represent complex relationships, such as those between users and items or events and entities. Dynamic BiCFlows innovatively integrates hierarchical aggregation and filtering techniques with linked list structures to manage the complexity of large datasets. These features allow users to explore high-level overviews by aggregating related nodes while also enabling the examination of finer details through filtering and drill-down interactions. For instance, users can track how connections between the two sets of nodes evolve over time, revealing trends, periodic patterns, or anomalies that might otherwise be obscured.

Additionally, visualizations that address uncertainty and user perception, such as sentiment and stance visualizations, as stated by Ramalho *et al.* [27], provide valuable insights into effectively representing ambiguous or complex temporal changes. Figure 3.14 shows an example of using multimodal methods to extract sentiment from a single individual. These tools demonstrate the importance of clarity and interpretability when presenting time-evolving data.

Figure 3.14: Screenshot of the UXMood visualization [27].

Key challenges in visualizing time-evolving data include maintaining scalability while dealing with large datasets that change dynamically. Ensuring interactivity and real-time responsiveness is essential, particularly when tracking complex temporal relationships. Managing visual clutter is another significant issue, as overlapping or dense temporal elements can obscure meaningful patterns. Additionally, balancing clarity with the complexity of temporal data remains a persistent issue, especially when incorporating multiple dimensions such as spatial and relational data. Addressing these challenges is crucial to creating intuitive and effective temporal visualizations that meet the needs of diverse audiences.

#### 3.1.4.1 Storytelling in Information Visualization

Storytelling and narrative techniques are relevant in making data visualizations engaging and comprehensible. Integrating narrative elements enhances the user's ability to extract meaningful insights from complex datasets by providing context and coherence. As highlighted in Kouřil *et al.*'s

Molecumentary [26], storytelling involves creating a structured flow of information, often using visual aids like animations, guided tours, and verbal annotations to explain data relationships and dynamics effectively.

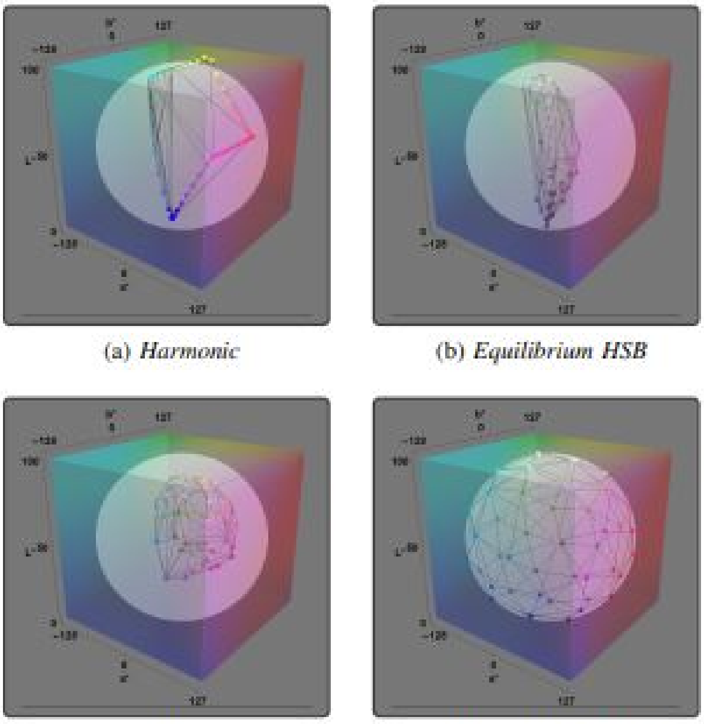
In dynamic and time-evolving datasets, storytelling becomes even more critical. It ensures that the complexity of changes over time is communicated in a structured way, allowing users to follow the progression and understand the data's implications intuitively. Tools that integrate storytelling, such as automated narrative synthesis or text-to-speech functionalities, represent significant advancements in making data visualization both insightful and engaging for diverse audiences.

### 3.1.5 User-Centric Design in Information Visualization

Creating compelling visualizations requires a user-centric approach, prioritizing the audience's needs, goals, and cognitive processes to ensure clarity, accessibility, and meaningful insights. For the message to be passed effectively, the visualization must be intuitive, accessible, and engaging enough so that the user correctly perceives it. Recent studies provide valuable insights into the principles and advancements in user-centric visualization design.

Visual analytics, for instance, bridges the gap between computational analysis and human cognition by combining powerful data processing capabilities with intuitive interfaces. Cui [32], in his study on trends and challenges in this field, underscores the importance of adaptable tools that cater to diverse user profiles, ensuring inclusivity and usability across a wide range of applications.

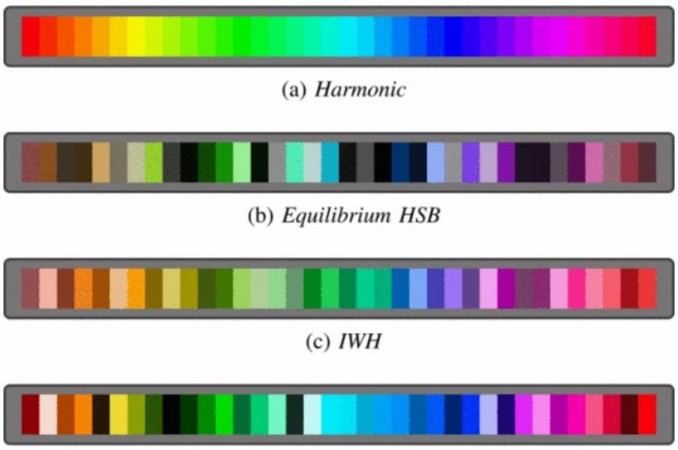
Color selection can significantly impact how users retain valuable insights from a visualization. As stated by Maji *et al.* [33], *"The human visual system tends not to perceive absolute colors of objects but rather the differences between them, i.e., their contrast"*. Their proposed approach, Equilibrium Distribution Model (EDM) [33], introduces an innovative method for selecting colors with optimal perceptual contrast, as shown in Figures 3.15 and 3.16. By addressing the challenge of designing informative and visually appealing visuals, EDM [33] ensures that datasets with numerous features remain comprehensible to users, regardless of their varying levels of visual literacy.



(a) *Harmonic* (b) *Equilibrium HSB*

(c) *IWH* (d) *Equilibrium LAB*

Figure 3.15: Distribution of color points using four different color schemes [33].



(a) *Harmonic*

(b) *Equilibrium HSB*

(c) *IWH*

(d) *Equilibrium LAB*

Figure 3.16: Discrete Color maps of the four approaches [33].

Annotations, modifications, and embellishments further enhance visualization by adding layers of meaning and context to raw data. Research on data visualization enhancement, performed by Suwanworaboon *et al.* [34] highlights the value of automatically generated annotations that emphasize key insights and trends, reducing cognitive load and enabling users to focus on actionable information. These enhancements make it easier for users to interpret complex datasets and draw meaningful conclusions.

Interactive 3D interfaces represent another advancement in user-centric design, immersing users in the data exploration process and fostering deeper engagement. For example, Touch the Collections, presented by Yang *et al.* [35], introduces a 3D user interface that visually outlines and interactively displays communities and collections. This tactile, spatial approach enhances users' ability to explore relationships intuitively and understand complex datasets.

User-centric visualization design faces several challenges, including balancing complexity with simplicity, ensuring accessibility for diverse audiences, and maintaining scalability for large datasets. These advancements underscore the importance of prioritizing user needs in designing visualization tools, ensuring they are functionally robust, engaging, and intuitive for a wide range of users.

## 3.2 Network Data with Spatio-temporal Dimensions

Integrating spatial, temporal, and network data visualizations is crucial for understanding dynamic systems where relationships evolve across time and space. Such visualizations combine multiple dimensions of information to provide an intuitive exploration of complex datasets, enabling users to uncover patterns and interactions that are otherwise difficult to discern.

As stated by He *et al.* [25], line-based visualizations are highly effective for trajectory analysis, offering clear and intuitive representations of movement patterns and flow distributions. They are handy for highlighting origin-destination relationships and effectively communicating associated attributes. Trajectory visualizations are usually shown through 2D maps [25], representing movement or connections between entities across geographical locations, as shown in Figure 3.17. However, as the number of trajectories on the screen increases, cluttering problems can be found, making it hard to visualize data effectively.

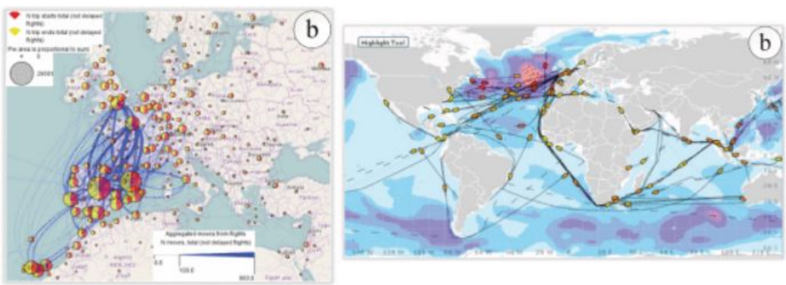



Figure 3.17: Geospatial-based visualizations for trajectory data [25].

Several advanced techniques have been developed to address the cluttering issues in trajectory visualizations, integrating spatial, temporal, and network dimensions to improve clarity and usability. For instance, Lavanya *et al.* [12] illustrates the visualization of COVID-19 dissemination through source-to-destination mapping that visualizes global trajectories across geographic regions.

By using geospatial-based line visualizations, this approach provides an intuitive understanding of movement patterns and connectivity. However, as trajectory data scales up, it still faces challenges in maintaining clarity. The same study also incorporates network visualization techniques, as observed in Figure 3.18, to abstract geographic connections into node-link diagrams, revealing structural relationships and patterns within the data, such as high-traffic nodes or dominant routes.

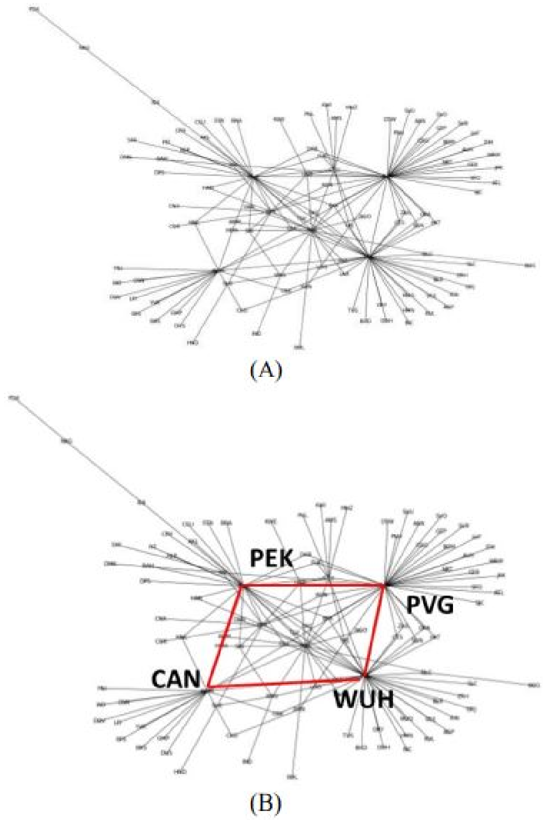



Figure 3.18: Network of Source to Destination Paths [12].

To further mitigate the challenges of clutter and scale, 3D visualization techniques have emerged as powerful tools. Li *et al.* [18] introduces globe-based visualizations, as shown in Figure 3.19, which represent trajectories as arcs on a 3D map. This method provides an immersive perspective, allowing users to explore global movement patterns in a way that reduces overlap and improves clarity compared to 2D representations. 3D visualizations effectively highlight relationships between geographically distant locations by leveraging depth and spatial separation.

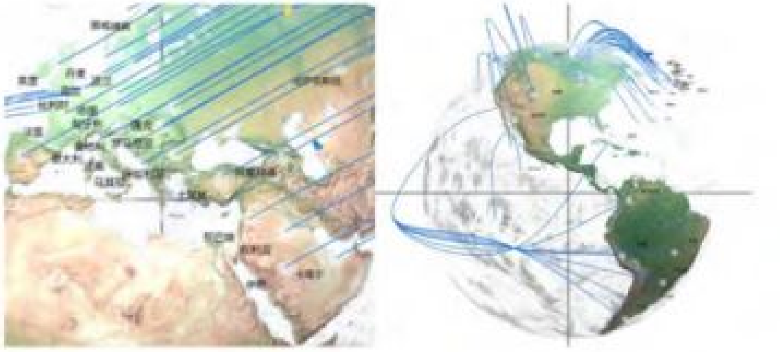


Figure 3.19: Connections between countries on a globe [18].

The final result is an animation, where users can press the play button (see Figure 3.20) and watch the trajectories appear and change over time, accurately representing the temporal evolution to the users.


Figure 3.20: 3D Earth: region in MovieLens [18].

Beyond 2D and 3D mapping, Beltrán *et al.* [13] propose a dual-space representation of trajectory data, combining geographic and abstract spaces to enhance interpretability. Figure 3.21 shows the data selection for display and analysis. This approach visualizes spatial relationships and integrates temporal attributes, enabling a deeper understanding of how trajectories evolve over time. Such temporal dynamics are critical in domains where movement patterns are influenced by changing contexts, such as football player transfers or pandemic spread. By offering interactive features like filtering and zooming [36], this methodology addresses clutter issues by allowing users to focus on specific data subsets.

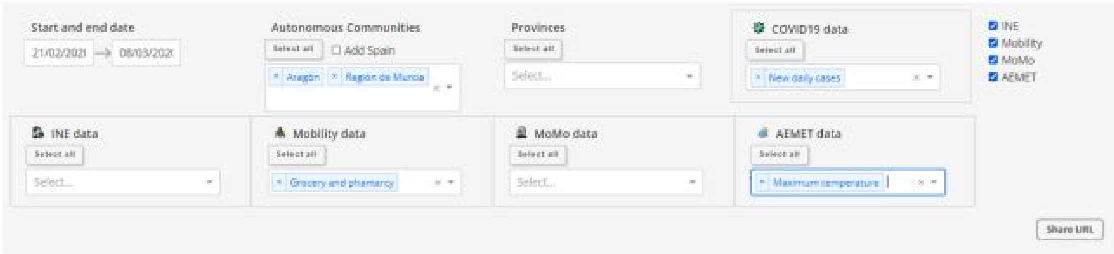



Figure 3.21: Data selection for display and analysis [13].

The integration of these techniques demonstrates the evolution of trajectory visualization from traditional 2D maps to more sophisticated methods that incorporate temporal and spatial dynamics. Line-based visualizations, while foundational, can benefit significantly from the addition of interactive features, dual-space representations, and 3D perspectives.

## 3.3 Summary

This chapter has presented a comprehensive exploration of recent techniques in information visualization, with a specific focus on sports data, network data, 3D visualization, time-evolving data, and the integration of spatial, temporal, and network dimensions. This analysis highlights the diversity of visualization strategies and their critical role in addressing the complexities of interconnected and dynamic datasets.

The discussion began with the importance of information visualization as a tool to bridge the gap between raw data and actionable insights, particularly in domains like sports, where large datasets demand intuitive and interactive solutions. From visualizing basketball metrics [8] to exploring virtual football networks [10], the reviewed studies emphasized the need for techniques that align with user needs and provide clarity among growing data complexity.

Network data visualization techniques, as examined through tools like Hermes [11] and knowledge graphs [14], underscored the potential of relational approaches to capture dependencies and dynamic interactions. These methods have been shown to support multi-level analysis and enhance usability, particularly when applied to interconnected datasets such as player transfers or economic systems.

The exploration of 3D visualization further expanded the potential for representing high-dimensional data. By leveraging depth and spatial separation, 3D techniques ranging from hybrid 2D-3D approaches [17] to globe-based visualizations [19] were demonstrated to address the challenges of clutter and occlusion, providing a more immersive and detailed perspective of data relationships.

Time-evolving data visualization presented unique challenges in capturing temporal dynamics while maintaining clarity. Studies on frameworks like COVERAGE [30] and dynamic bipartite graph visualization [31] illustrated the value of scalable, interactive tools in tracking patterns over time. The integration of storytelling techniques complements these approaches [26], which enhanced user engagement and facilitated the understanding of complex temporal data.

The integration of network, spatial, and temporal dimensions was explored as a critical advancement in trajectory visualization, where techniques such as 3D mapping and dual-space representation offered innovative solutions to longstanding issues of clutter and complexity. The evolution of these methods demonstrated the growing importance of interactivity and multidimensional perspectives in uncovering insights from dynamic datasets.

In conclusion, the reviewed studies underscore the importance of adopting multi-faceted visualization techniques to address the challenges of interconnected, time-evolving datasets. The insights gained from this chapter provide a solid foundation for advancing the visualization of global sports data, particularly in representing the evolution of player movements across leagues and countries. This chapter highlights the potential for further innovation in creating visualization tools that are functionally robust, engaging, and accessible to diverse audiences.

# Chapter 4

# Solution

This chapter introduces the proposed solution for visualizing cross-border football player transfers over time using a 3D globe representation. Section 4.1 explains in detail the proposed solution, providing insights into its planned features. Section 4.2 covers the prototype's design and implementation details, representing the proposed solution. Section 4.3 summarizes what is said in this chapter.

## 4.1 Proposed Solution

The systematic literature review and state of the art conducted in Chapters 2 and 3 revealed several key insights regarding visualizing geospatial, temporal, and relational data. Geospatial visualizations, particularly those employing 3D globes, offer an intuitive representation of global relationships and patterns. When combined with temporal controls and interactive elements, these visualizations can effectively communicate complex, interconnected data that evolves over time.

Traditional approaches to visualizing transfer networks often employ 2D network diagrams or matrix representations [11], which can effectively show relationships but often struggle to convey the geographic context inherent in cross-border player transfers. Alternatively, map-based visualizations can represent geographic information clearly but may become cluttered when displaying numerous connections simultaneously.

The literature emphasizes that effective visualizations should balance complexity with accessibility, particularly for web-based applications intended for diverse audiences [13]. Additionally, interactive elements and animation can significantly enhance user engagement and facilitate deeper data exploration, provided they are implemented thoughtfully to avoid overwhelming users.

Based on these considerations and drawing inspiration from existing globe-based visualizations such as *MovieLens* [18] and *Globe.gl* [37], the proposed solution is a **3D interactive globe visualization with animated transfer arcs**. The primary concept behind this visualization is representing countries as nodes on a 3D globe, with arcs connecting countries to represent player transfers between them, as shown in Figure 4.1.

The core features of the proposed solution include:

Figure 4.1: Proposed Solution - 3D Globe Visualization.

- **Geospatial Representation:** Countries are positioned accurately on a 3D globe, providing an intuitive geographic context for the transfer data.
- **Transfer Arcs:** Curved arcs connect countries, representing player transfers. The thickness of an arc indicates the volume of transfers between countries, while the direction of the arc (from origin to destination) shows the transfer direction.
- **Temporal Evolution:** A time slider interface allows users to navigate through different years, observing how transfer patterns have evolved over time. Animation between years visually conveys the changing patterns.
- **Interactive Filtering:** Users can filter the visualization by specific countries, viewing only incoming or outgoing transfers, or isolating transfers between specific country pairs.
- **Player-Specific Tracking:** A search function enables users to find specific players and visualize their career paths across different countries and years.
- **Contextual Information:** Tool-tips and information panels provide additional details about transfers, countries, and players when users interact with visualization elements.

These features are aligned with the requirements identified in Section 1.2. Priority levels were set for feature implementation due to the difficulty of creating a fully functional 3D visualization: first, the core globe representation with precise country positioning and visualization of transfer arcs; second, the incorporation of timeline controls and animation; and last, the inclusion of search, filtering, and detailed information display capabilities.

## 4.2 Prototype Implementation

The prototype consists of four primary components organized in distinct layers: data acquisition, data management, visualization engine, and user interface. The overall architecture and data flow are illustrated in Figure 4.2.

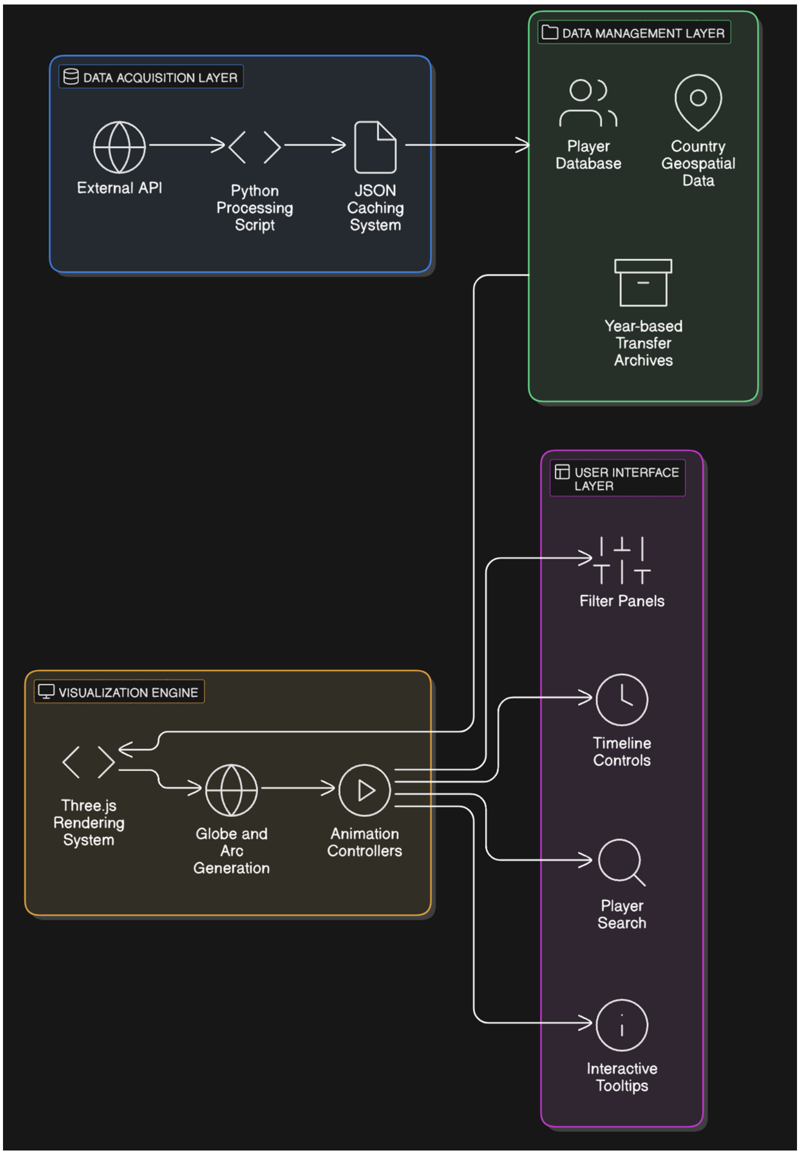



Figure 4.2: Prototype's Architecture and Data Flow.

The **Data Acquisition Layer** forms the system's foundation, where football transfer data is obtained from the zerozero.pt external API. This raw data is processed through a Python script that transforms and structures the information before storing it in a JSON caching system. This caching mechanism significantly reduces loading times and minimizes repeated API requests.

The **Data Management Layer** organizes the processed data into three key repositories: the Player Database, containing comprehensive information about football players and their career paths; Country Geospatial Data, providing geographical coordinates for global mapping; and year-based Transfer Archives, storing the temporal transfer data from 1950 to 2025.

The **Visualization Engine** powers the interactive 3D representation using Three.js as the core rendering system. This engine generates the globe and transfers arcs while managing animations and transitions between different visualization states. The rendering system is optimized to handle complex 3D scenes with numerous transfer arcs while maintaining smooth performance.

The **User Interface Layer** provides multiple interaction mechanisms for exploring the data: Filter Panels for selecting specific countries or transfer types; Timeline Controls for navigating through different years; Player Search functionality for tracking individual player careers; and Interactive Tool-tips providing detailed information about specific transfers.

The visualization is implemented as a web application, making it accessible across different platforms without requiring specialized software. The frontend is developed using HTML, CSS, and JavaScript, with the 3D visualization powered by Three.js [38] and the Three-Globe [39] library extension. The data processing pipeline transforms raw transfer data into optimized formats for visualization, enabling efficient rendering and interaction with complex global transfer patterns spanning multiple decades. Further details on each component's implementation are provided in the subsequent sections.

### 4.2.1 Technology Selection

Several factors were considered when evaluating technologies for implementing the 3D globe visualization, including rendering capabilities, performance, browser compatibility, and ease of integration with web standards. After assessing various options, the following technologies were selected:

- **Three.js:** This JavaScript library provides a high-level API for WebGL [40], making it ideal for creating and displaying 3D content in web browsers. Three.js offers built-in support for essential features such as cameras, lighting, materials, and geometries, while also providing flexibility for custom implementations. Its extensive documentation and active community support further justified its selection.

- **Three-Globe:** Built on top of Three.js, this specialized library offers functionalities explicitly for globe visualizations, including methods for rendering countries, arcs, and points on a 3D globe. It provides convenient APIs for managing geospatial data and creating visually appealing globe-based visualizations.

To validate the suitability of these technologies, initial experiments were conducted to test key aspects of the visualization. Globe Rendering tests confirmed that Three-Globe could efficiently render a detailed globe with country boundaries and customizable appearance. Arc Visualization experiments demonstrated the ability to create visually distinct arcs with controllable properties such as height, thickness, and color. At the same time, Performance Testing under heavy loads confirmed that the selected technologies could maintain acceptable frame rates even with complex visualizations.

These initial explorations validated the technology choices and provided valuable insights for fully implementing the visualization.

### 4.2.2 Data Processing

The visualization relies on comprehensive football player transfer data spanning multiple years. This data required significant processing to transform into a format suitable for efficient visualization.

#### 4.2.2.1 Data Structure

The raw transfer data follows a hierarchical JSON structure that organizes football player transfers by league, season, and participating clubs. The structure begins with league-level metadata, including the competition name and logo URL. Within this top level, the data is segmented by seasons, each containing an array of participating clubs.

For each club, the structure includes:

- Club identifiers (ID, logo URL)
- Geographical metadata (continent association)
- Two key transfer groupings: incoming transfers and outgoing transfers, both organized by country

The country-level groupings contain:

- Country identifiers (flag image URL, continent, and country name)
- A count of transferred players to/from that nation
- Detailed player records featuring:
  - Player ID, full name, position, and date of birth
  - Career movement information via origin/destination club IDs and names
  - Unique transfer identifiers
  - Player profile links

As said above, the raw JSON transfer data requires significant transformation to become suitable for spatial visualization and analysis. This processing stage normalizes the hierarchical source data into two specialized file formats: geospatial arc files for transfer route visualization, divided by year, and a comprehensive player database for career trajectory analysis. The transformation ensures standardized geographic coordinates, consistent country codes, and optimized data structures that facilitate efficient rendering and querying.

The processed output consists of:

- **Arc files** - Contain geospatial transfer routes between countries with associated metadata, as shown in Listing 4.1.

```
{
    "type": "Transfer",
    "arcs": [
        {
            "type": "transfer",
            "from": "POR",
            "to": "ING",
            "startLat": 39.0524,
            "startLong": -8.495397,
            "endLat": 52.288958,
            "endLong": -0.238655,
            "color": "#F76B15",
            "scale": 0.5,
            "count": 6,
            "players": [
                "Sergio Leite",
                "M\u00e1rio Jardel",
                "Silas (1976)",
                "Cristiano Ronaldo",
                "C\u00e2ndido Costa",
                "H\u00e9lder Postiga"
            ],
            "player_ids": [
                "1768",
                "1579",
                "1345",
                "430",
                "1730",
                "1293"
            ]
        },
        // More arcs
    ]
}
```

Listing 4.1: Structure of the arc files.

- **Player database** - Stores complete career paths and attributes for all transferred players, as shown in Listing 4.2.

```
{
  "players": {
    "1579": {
        "id": "1579",
```

```
    "name": "Cristiano Ronaldo",
    "position": "Avan\u00e7ado",
    "birthDate": "1985-02-05",
    "transfers": [
      {
        "year": "2003",
        "from_country": "POR",
        "to_country": "ING",
        "from_club_id": "16",
        "from_club_name": "Sporting",
        "to_club_id": "87",
        "to_club_name": "Manchester United"
      },
      // More transfers
   ],
   "country_flags": {
      "ING": "https://www.zerozero.pt//img/bandeiras/16
        _eng_imgbank_flag.png",
      "POR": "https://www.zerozero.pt//img/bandeiras/1
        _por_imgbank_flag.png",
      // More flags
    },
    "display_name": "Cristiano Ronaldo"
  },
  // More players
 }
}
```

Listing 4.2: Sturcture of the player database.

This dual-file architecture separates spatial transfer patterns from player metadata, enabling efficient visualization rendering while maintaining comprehensive player information. The arc files optimize geospatial operations with previously defined coordinates and aggregated transfer counts, while the player database serves as a centralized reference for detailed career analysis. The transformation process ensures both datasets remain linked through consistent player and country identifiers.

To ensure smooth performance, particularly when animating between years or handling user interactions, several optimization techniques were applied to the data:

- **Deduplication:** In cases where multiple transfers occurred between the same countries in the same year, these were consolidated into a single arc with an updated count and player list. This significantly reduced the number of arcs that needed to be rendered.
- **Coordinate Caching:** Country coordinates were cached to avoid redundant calculations, particularly when generating arcs between the same countries across different years.

- **Progressive Loading:** Data for each year is loaded on demand rather than preloading the entire dataset, reducing initial load times and memory usage.

- **Efficient Lookup and Indexing:** Player and country data were indexed using dictionaries (hash maps) to enable efficient lookup and filtering operations. This approach avoids repeated linear scans, ensuring responsive filtering, search, and animation rendering.

These optimizations were crucial for maintaining responsive performance, especially when users interact with the visualization through filtering, searching, or timeline navigation.

### 4.2.3 Visual Design

The visual design of the 3D globe visualization focuses on clarity, aesthetics, and effective data representation. Various visual elements and properties were carefully considered to create an intuitive and engaging user experience.

#### 4.2.3.1 Globe Appearance

The globe serves as the foundation of the visualization, providing geographic context for the transfer data. Its design balances detail with performance considerations:

- **Surface Appearance:** Instead of textures, the globe surface uses dynamic colors that adapt to light and dark modes, balancing performance and clarity.

- **Country Boundaries:** Subtle borders delineate countries, helping users identify specific nations without creating visual clutter.

- **Atmosphere Effect:** A subtle atmospheric glow surrounds the globe, enhancing the three-dimensional appearance and providing depth cues.

- **Lighting:** Directional and ambient lighting create appropriate shadows and highlights on the globe surface, enhancing the perception of its three-dimensional form while ensuring that all regions remain visible.

#### 4.2.3.2 Transfer Arcs

Transfer arcs are the primary data elements in the visualization, representing player movements between countries:

- **Shape:** Arcs follow a curved path between countries, with the curve's height proportional to the distance between countries. This approach prevents arcs from intersecting with the globe's surface while creating visually distinct paths.

- **Thickness:** The thickness of an arc corresponds to the number of players transferred between countries, with thicker arcs representing higher transfer volumes.

- **Color:** As shown in Figure 4.3, different color schemes are employed to distinguish between various types of transfers:
  - Orange for general or incoming transfers
  - Blue for outgoing transfers when viewing a specific country
  - Green for transfers involving a specific player
- **Animation:** Arcs feature a flowing animation along their length, creating a sense of direction and movement. This animation uses a dash pattern that moves along the arc path, enhancing the perception of transfer direction.

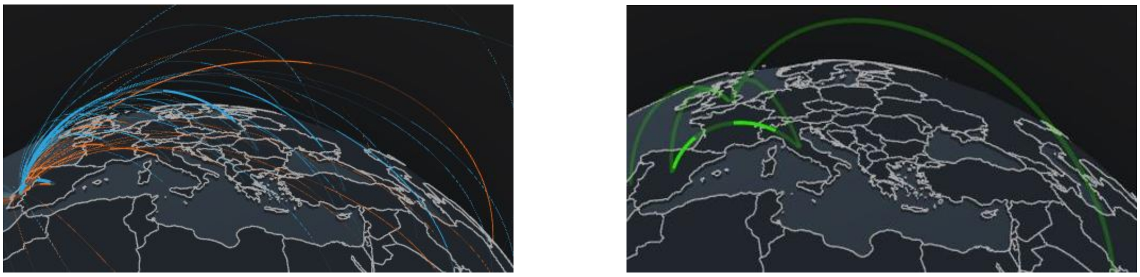


Figure 4.3: Different types of arcs.

#### 4.2.3.3 Color Schemes

Color plays a crucial role in the visualization, both for aesthetic purposes and for encoding data attributes:

- **Theme Colors:** The visualization supports both light and dark themes, with visual elements adapting accordingly, as shown in Figure 4.4.
- **Data-Driven Colors:** Colors are used consistently to encode specific data attributes:
  - Bright arc colors represent transfer types (general, incoming, outgoing, player-specific)
  - Neutral colors are used for interface elements to avoid competing with data visualization
- **Accessibility Considerations:** While our primary color schemes are designed for contrast, future iterations may incorporate additional testing to ensure compatibility with color vision deficiencies.

### 4.2.4 Interaction Mechanisms

The visualization incorporates various interaction mechanisms to enable users to effectively explore and analyze the transfer data. These interactions are designed to be intuitive and responsive, enhancing the overall user experience.

Users can navigate and explore the globe using several intuitive controls:

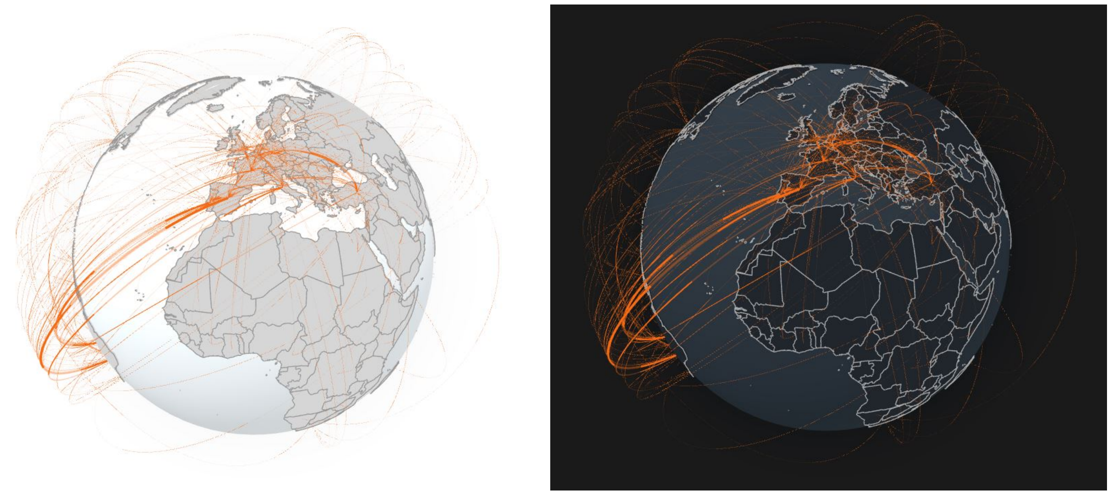


Figure 4.4: Comparison between light and dark modes.

- **Rotation:** Click and drag operations rotate the globe, allowing users to choose the area of the globe they want to see. The rotation is implemented using Three.js' OrbitControls, which provides smooth and responsive movement.
- **Zoom:** Mouse wheel or pinch gestures adjust the zoom level, enabling users to focus on specific regions or view the entire globe. Zoom limits are implemented to prevent users from zooming too far in or out.
- **Hover Interactions:** Hovering over transfer arcs displays tool-tips with relevant information, such as country names, transfer counts, and player lists. This provides immediate context without requiring additional user actions.

In terms of filtering, several options are offered to help users focus on specific aspects of the transfer data:

- **Country Filters:** Users can select specific countries to include or exclude from the visualization, focusing on transfers relevant to their interests.
- **Transfer Direction:** Filters allow users to view only incoming transfers, only outgoing transfers, or both, providing flexibility in analyzing transfer patterns.
- **Country-to-Country Filter:** A specialized filter enables users to isolate transfers between two specific countries, facilitating detailed analysis of bilateral transfer relationships.
- **Player Filter:** Users can filter the visualization to show only transfers involving specific players, making it easy to track individual careers. This player filter has the following functionalities:
  - **Auto-complete:** As users type in the search field, an auto-complete drop-down displays matching player names, as well as their positions and birth dates, facilitating quick selection.

– **Career Visualization:** When a player is selected, the visualization highlights all transfers involving that player across all years, effectively displaying their career path.
– **Player Information:** When a player is selected, additional information about their career is displayed, including the clubs they played for.

These filters can be combined to create highly specific views of the transfer data, enabling detailed analysis of particular aspects of interest.

### 4.2.5 Timeline Integration

The timeline component is a crucial visualization element, enabling users to explore how transfer patterns have evolved over time. Its implementation balances functionality with usability to create an intuitive temporal navigation experience.

The timeline interface is designed to resemble familiar media player controls, making it immediately recognizable and intuitive, as shown in Figure 4.5.


Figure 4.5: Timeline Interface Design.

Key elements of the timeline interface include:

- **Year Slider:** A horizontal slider allows users to navigate directly to specific years by dragging the slider handle. The current year is prominently displayed above the slider.
- **Navigation Buttons:** Forward and backward buttons enable step-by-step navigation through years, providing precise control over temporal exploration.
- **Play/Pause Button:** This button initiates or halts an automatic progression through years, creating an animation showing how transfer patterns have evolved.
- **Speed Control:** A button toggles between normal (1x) and fast (2x) playback speeds, giving users control over the pace of the temporal animation.

## 4.3 Summary

This chapter has introduced the proposed solution for visualizing football player cross-border transfers using a 3D globe representation. The solution leverages the inherent geographic nature of transfers between clubs of different countries, representing countries as nodes on a globe and transfers as arcs between these nodes. Various visual encodings, including arc thickness, color, and height, convey different aspects of the transfer data.

The implementation details covered in this chapter demonstrate how the conceptual design was realized using web technologies, particularly Three.js and related libraries. The visual design balances aesthetic considerations with effective data representation, while the interaction mechanisms enable intuitive exploration of the complex, multidimensional dataset. The timeline integration facilitates temporal analysis, allowing users to observe how transfer patterns have evolved over time.

The next chapter will detail the experimental protocol designed to evaluate the effectiveness, usability, and engagement of this visualization solution.

# Chapter 5

# Experiment

This chapter presents the experiment developed to assess the accuracy, usability, engagement, and understandability of the 3D globe-based visualization representing cross-border football player transfers over time. Section 5.1 describes the protocol to follow. Section 5.2 refers to the conducted pilot test, while Section 5.3 details the actual execution of the experiment. Section 5.4 summarizes this chapter.

## 5.1 Protocol

A protocol was developed to ensure a structured and consistent approach to evaluating the visualization. It defines the experiment's objectives, setup, recruitment strategy, experiment procedure, and assessment methods used to measure user performance, engagement, and perceived usability.

### 5.1.1 Objective

The experiment aims to evaluate the hypothesis that *interactive 3D visualizations* can effectively convey geospatial, temporal, and relational data in the context of cross-border football player transfers.

The central research questions are:

- **Q1:** Can users interpret the spatial and temporal transfer patterns displayed?
- **Q2:** Can users easily interact with the visualization to extract insights?
- **Q3:** Does the visualization maintain user engagement and encourage exploration?

Detailed answers to these questions can help assess the current state of the prototype and guide future refinements, ensuring the tool aligns with user expectations and the needs of the intended audience.

### 5.1.2 Setup

To support this methodology, the following parameters were established:

- **Where and When:** The visualization is hosted online, enabling users to participate remotely and autonomously. The experiment spans approximately two weeks.
- **Equipment and Software:** Participants only required an internet-connected device with a web browser. While compatible with mobile and tablet devices, desktop usage was recommended to ensure optimal interaction with the 3D interface.
- **Participation Format:** The experiment was open to any voluntary participant without restrictions. All sessions were conducted remotely and individually, without the presence of a coordinator, observer, or facilitator. This unsupervised setup allows for scalable participation and simulates the real-world scenario of a user discovering the visualization organically on a website. The lack of assistance ensures an authentic evaluation of the system's intuitiveness and ease of use.
- **Help Resources:** No external instructions, tutorials, or documentation were provided during the experiment. Participants were expected to explore the visualization independently, reinforcing the goal of assessing its clarity and usability in a natural context.

It is important to note that although the experiment remained open for an extended period, the prototype itself remained unchanged throughout. This ensured consistent conditions for all participants and guaranteed that any conclusions drawn from the results were based on the same version of the system. Enhancements or adjustments to the visualization, if required, will only be considered after the completion of the experiment.

### 5.1.3 Participant Recruitment

A diverse and representative sample of participants was targeted to reflect the broad characteristics of ZOS users, considering different levels of expertise, backgrounds, and interests.

Three distinct strategies were employed:

1. **Internal outreach:** Company employees were invited via messages on Slack [41].
2. **Platform promotion:** A news post on zerozero.pt [1] engaged the platform's core audience, which is football enthusiasts and regular users.
3. **Broader outreach:** Friends and colleagues less familiar with football or data visualizations were invited to ensure representation of casual users.

[1] https://www.zerozero.pt/noticias/um-mundo-de-transferencias-a-vista-a-nova-funcionalidade-do-zerozero-/803380

### 5.1.4 Assessment Method

The assessment method consists of a formative questionnaire developed on Google Forms [42], which evaluates demographics, experience, task performance, and usability perception. It is divided into four parts: **Informed Consent**, **Pre-test Section**, **Task-specific Questions**, and **Post-test Questionnaire**. The questionnaire can be observed in its entirety in Annex B.

#### 5.1.4.1 Informed Consent

In the initial section of the questionnaire, participants are presented with a short introduction outlining the goal and purpose of the experiment. They are informed that their participation is entirely voluntary and that they may withdraw at any time without penalty. It is also stated that all data collected will remain confidential and be used exclusively for academic research.

Before proceeding, participants are required to provide informed consent by agreeing to:

- Provide demographic and experience information.
- Allow anonymous recording of your interaction data for research purposes.
- Evaluate the usability and intuitiveness of the visualization.

#### 5.1.4.2 Pre-test Section

The pre-test section of the questionnaire collects essential background information from each participant, including demographics and self-assessed familiarity with relevant domains. All responses are anonymous and used solely for statistical analysis within the scope of this study, with no data shared or used for commercial purposes.

**Demographic Information:** Participants provide their age and indicate their gender by selecting one of the following options: Male, Female, Other, or Prefer not to say. This information contributes to understanding the composition of the participant pool.

**Experience Levels:** To gauge prior familiarity, participants are asked to rate their experience on a 7-point Likert scale (1 = Very unfamiliar, 7 = Very familiar) in relation to three domains:

- Understanding of 3D visualizations
- Expertise in football
- Experience with data visualization tools

These responses help assess how prior domain knowledge might influence participants' interaction with the visualization and their performance in the experiment.

#### 5.1.4.3 Task-specific Questions

The task-specific questions evaluate the ability of the participants to interact with the visualization and extract meaningful insights based on the transfer data it represents. This approach measures whether participants can navigate and interpret different views, filters, and temporal layers effectively, using the interface autonomously.

Each task is structured as a direct question with a predefined correct answer, allowing for objective evaluation of participant responses. The majority of answers were submitted as free-text inputs, encouraging users to explore the visualization independently and express their findings in their own words. One task was presented as a multiple-choice question to reduce cognitive load in a more complex comparison. This mixed-format approach balances user freedom with analytical consistency while minimizing ambiguity and accommodating a diverse participant profile. The questions were carefully designed to test specific features of the visualization, such as filtering, temporal navigation, pattern recognition, and player-specific tracking.

The tasks vary in scope and complexity, progressing from simple filtering queries to more interpretive or multi-step reasoning tasks. They cover both macro-level trends (e.g., aggregate country flows) and micro-level information (e.g., a player's trajectory). The following list presents each question, the functionality it was designed to assess, and the correct answer to be provided during the experiment:

**T1:** Which country received the most outgoing transfers from the USA in 2023? (Correct answer: *Canada*) - Assesses the ability to filter by source country and year, and interpret outgoing arcs geographically.

**T2:** Which country transferred the most players to England in 2002? (Correct answer: *France*) - Assesses the ability to filter by source country and year, and interpret incoming arcs geographically.

**T3:** Which South American country transferred more players to Europe in 1997? (Correct answer: *Brazil*) - Evaluates intercontinental filtering and comparative analysis of transfer volumes between continents.

**T4:** How many countries did Nicolas Anelka play in? (Correct answer: 7) - Examines player search functionality and the ability to trace a career path across time and space.

**T5:** In 2007, which pair of countries transferred the most players between them? (Correct answer: *Portugal/Brazil*) - Focuses on interpreting edge thickness and bilateral transfer filtering within a given year.

**T6:** What's the year of the first transfer from China to a European country? (Correct answer: 1988) - Tests the timeline navigation feature and the identification of temporal transfer events.

**T7:** What's the most common continent from which players transfer to England over time (Correct answer: *Europe*) - Assesses trend identification across multiple years and spatial groupings.

Participants' answers to these tasks offer valuable information about how well they explore the visualization's features and their capacity to interpret dynamic, time-based data. A detailed analysis of the accuracy of these responses is presented in the next chapter, with attention to patterns related to participant characteristics and interaction behavior.

#### 5.1.4.4 Post-test Questionnaire

The System Usability Scale (SUS) [43] was used to evaluate the visualization's overall usability. SUS is a well-known and standardized tool for evaluating the perceived usability of interactive systems through a short 10-item questionnaire (see Annex B, Figures B.5 and B.6), using a 5-point Likert scale (1 = Totally Disagree, 5 = Totally Agree). It is particularly well-suited for usability testing in academic and industrial contexts since it offers a rapid yet accurate indicator of effectiveness, efficiency, and user satisfaction. However, its interpretability can be somewhat limited, as it does not offer diagnostic insights into particular usability issues. To fill this gap, three optional open-feedback questions were added to the end of the questionnaire:

- What aspects of the visualization did you find most intuitive?
- What aspects did you find most confusing?
- Do you have suggestions for improvement?

These open feedback questions were chosen to complement the quantitative results by capturing participants' subjective experiences with the visualization. They aim to identify which features were perceived as intuitive, which elements caused confusion, and to gather constructive suggestions for improvement. This qualitative input provides valuable context to interpret usability scores and task performance, helping to uncover design issues or user needs that may not be evident through structured responses alone.

### 5.1.5 Experiment Procedure

Each participant followed this sequence:

1. Access the landing page presenting the link for the questionnaire and the option to start the visualization.
2. Access the dedicated page presenting the visualization in its initial state (most recent year).
3. Complete the voluntary questionnaire.
4. Engage freely with the visualization during questionnaire completion.

### 5.1.6 Completion and Success Criteria

The experiment is considered complete if the participant submits the full questionnaire after agreeing to the informed consent terms.

Each visualization-related task is evaluated in a binary manner: **correct (1)** or **incorrect (0)** based on pre-established answers. Success criteria focus on task accuracy and overall usability, as task completion time is not directly measured.

## 5.2 Pilot Test

A pilot test was conducted on April 30$^{\text{th}}$ to validate the structure and flow of the experiment, including the clarity of the participant journey, the usability of the visualization, and the overall coherence of the materials provided. This step aimed to ensure that participants would be properly oriented before interacting with the prototype and that the data collected would be reliable and consistent.

The test involved two participants with experience in both football and data visualization, selected for their familiarity with the domain and ability to provide insightful feedback. The procedure consisted of the following steps:

1. Participants accessed the experiment through a dedicated link and were introduced to the project independently.
2. They interacted with the prototype and completed the full questionnaire, including all structured and open-ended sections.
3. Observations were made during the session to detect usability issues or any confusion in the user flow.
4. At the end, informal interviews were conducted to gather qualitative impressions and improvement suggestions.

**Adjustment Made:** Based on the feedback collected during the pilot test, a single improvement was introduced: the creation of a dedicated landing page to guide participants before accessing the visualization. Originally, the experiment opened directly into the interactive globe, which left participants unsure about the objective and the next steps. The new landing page includes:

- A brief explanation of the project's purpose
- Information about the author and the scope of the thesis
- A direct link to the questionnaire
- A clear call-to-action button to access the visualization

This change ensures that participants are properly contextualized before starting the task, improving the structure and user experience of the overall protocol.

No further changes to the prototype or questionnaire were deemed necessary based on the pilot observations.

## 5.3 Execution

The experiment was conducted between May $2^{nd}$ and May $17^{th}$, during which the prototype and questionnaire were publicly accessible.

**Participation:** A total of *54* participants submitted complete questionnaires. Recruitment targeted a diverse audience, including domain experts, casual football fans, and general users, reflecting the platform's real-world user base.

**Engagement:** Participants actively interacted with the visualization, exploring multiple years, filters, and transfer paths. Interaction logs indicate an average session duration of approximately *16 minutes*, with some users spending over 30 minutes engaging with the interface. This suggests a high level of curiosity and exploratory behavior beyond task completion.

**Technical Performance:** No significant issues were reported during the experiment. The visualization loaded reliably across supported browsers and devices, with no crashes or critical usability barriers. Minor performance variation was noted on older mobile devices, particularly when rendering dense transfer data, but it did not impact overall functionality.

**Data Collection:** All questionnaire responses were stored in Google Forms and exported in a structured format for analysis.

The results collected during this phase serve as the basis for the detailed analysis presented in the following chapter.

## 5.4 Summary

This chapter presented the full experimental process designed to evaluate the usability, clarity, and engagement potential of the 3D globe visualization for international football player transfers.

The protocol was structured into clearly defined phases, including informed consent, a pre-test questionnaire, task-specific questions, and a post-test usability evaluation. A pilot test with two experienced participants led to the introduction of a dedicated landing page, ensuring participants received appropriate context before exploring the visualization.

The final experiment was conducted between May $2^{nd}$ and May $17^{th}$, with 54 participants completing all steps. The questionnaire responses were successfully collected.

The combination of a diverse participant sample, task accuracy measurement, SUS scoring, and open feedback provides a solid foundation for evaluating the visualization's effectiveness and for guiding future improvements.

# Chapter 6

# Results Analysis

This chapter presents the results of the experiment introduced in Chapter 5 and its respective analysis. Section 6.1 examines the data collected through the questionnaire, including the outliers identified, participant demographics, task accuracy, System Usability Scale (SUS) scores, and open-ended feedback. Section 6.2 offers an interpretation of the findings, highlighting key patterns and their implications for the usability and effectiveness of the visualization. Finally, Section 6.3 summarizes this chapter.

## 6.1 Formative Questionnaire

The formative questionnaire serves as a way of analyzing participant profiles, evaluating task performance, and assessing the usability of the 3D visualization. It collects both quantitative and qualitative data, offering insight into how the participants interacted with the system and how they perceived its clarity and were able to engage with all its functionalities. Section 6.1.1 explains the process of identifying and removing possible outliers while the following four sections delve into four distinct aspects of this questionnaire: demographics (Section 6.1.2), task performance (Section 6.1.3), usability evaluation (6.1.4), and qualitative feedback (6.1.5).

### 6.1.1 Outliers Identification

As described in Section 5.1.2, the experiment was conducted remotely, with participants engaging autonomously through an online platform. While this approach enables broader participation and allows users to explore the visualization without any external influence, it also introduces potential risks related to the quality of the answers received. Without direct control over how participants interact with the prototype or complete the questionnaire, there is a possibility that some responses do not reflect genuine engagement or proper thinking. Participants might attempt to answer the questions based solely on their existing football knowledge, without actually interacting with the visualization, or even submit random responses without genuine effort, potentially compromising the reliability of the results. To ensure this reliability, submissions were reviewed using two criteria to identify possible outliers:

- **Identical SUS Responses:** submissions in which the participant provided the same value for all 10 (or nearly all) items of the System Usability Scale (SUS), such as all ones or all fives.
- **Extreme SUS Scores:** submissions with a SUS standard score (z-score) with an absolute value greater than two (more than two standard deviations above the mean).

The first criterion targets inattentiveness or mechanical completion of the SUS form, as the mentioned pattern suggests that the participant did not engage meaningfully with the usability evaluation, given that the SUS is structured with alternating positive and negative statements. The second criterion identifies statistical outliers whose usability ratings significantly deviate from the population, indicating either exceptionally biased or unreliable feedback.

The SUS score [43] is calculated by applying a specific transformation to each of the 10 questionnaire items. Odd-numbered items (positive statements) are scored by subtracting 1 from the participant's response, while even-numbered items (negative statements) are scored by subtracting the response from 5. The SUS score is then obtained by adding these adjusted values and multiplying the result by 2.5 to scale it to a 0–100 range. The z-score is calculated for each submission using the following standard formula:

$$z = \frac{x - \mu}{\sigma}$$

Where:

- $x$ is the SUS score of the individual submission
- $\mu$ is the mean SUS score of all submissions
- $\sigma$ is the standard deviation of the SUS scores

The values used for this calculation are summarized in Table 6.1.

Table 6.1: Descriptive statistics of SUS scores before outlier removal.

| | Min. Accepted | Max. Accepted | Mean | Standard Deviation |
|---|---|---|---|---|
| SUS Score | 42.17 | 97.27 | 69.72 | 13.78 |

Submissions meeting either of these two criteria were excluded from the analysis to preserve the integrity of the results. This ensures that all subsequent metrics are based on a consistent and reliable dataset.

Applying these criteria resulted in the exclusion of 6 out of 54 total submissions, leaving 48 valid responses for the subsequent analysis.

### 6.1.2 Demographics

Among the 48 valid submissions that remained after the outlier identification and removal process, 36 participants identified as male (75%) and 12 as female (25%).

Table 6.2 shows the distribution of participants' ages and their self-reported knowledge in the domains of football, 3D visualizations, and data visualization tools, all evaluated on a 7-point Likert scale (1 = Very unfamiliar, 7 = Very familiar).

Table 6.2: Distribution of age and domain knowledge among participants.

| | Min. | 1st Quartile | Median | 3rd Quartile | Max. | Std. Deviation |
|---|---|---|---|---|---|---|
| Age | 19.00 | 22.00 | 26.00 | 38.00 | 55.00 | 11.15 |
| Football Knowledge | 2.00 | 4.75 | 6.00 | 7.00 | 7.00 | 1.66 |
| 3DViz Knowledge | 1.00 | 3.75 | 5.00 | 5.00 | 7.00 | 1.62 |
| DataViz Knowledge | 1.00 | 4.00 | 5.00 | 5.00 | 7.00 | 1.46 |

The demographic data presented in Table 6.2, along with the histogram and bar charts visible in Figures 6.1, 6.2, 6.3, and 6.4, shows that most participants fall within the age range of 20-27, with a median age of 26. Most of them reported high familiarity with football, aligning with the expectations given the platform's sports-oriented audience. In contrast, self-assessed knowledge of 3D visualizations and data visualization tools shows a broader distribution across the scale.

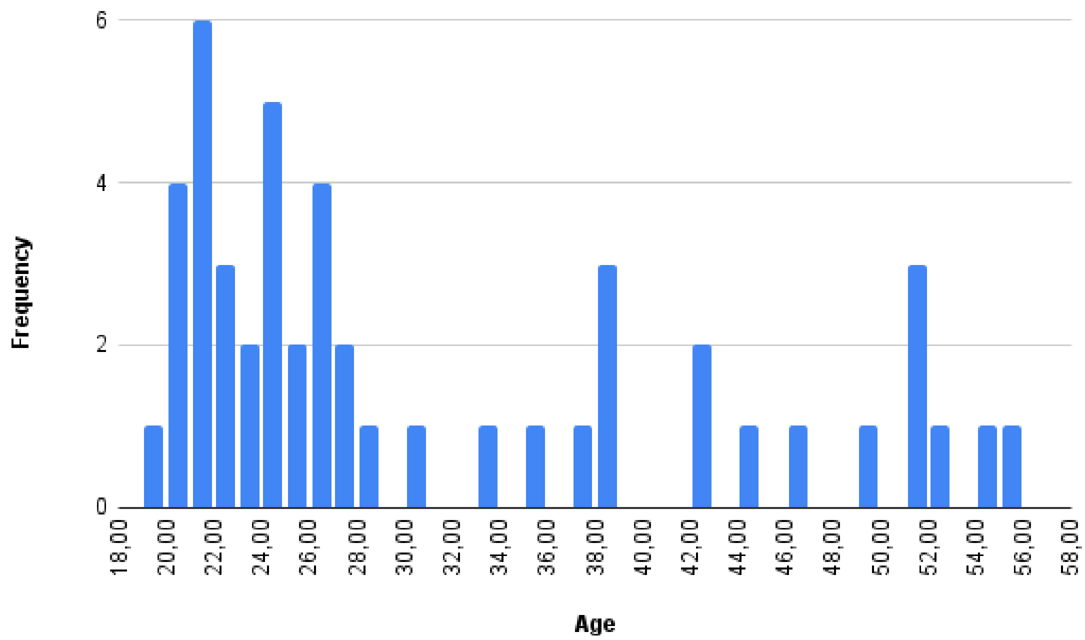



Figure 6.1: Participant distribution by age.

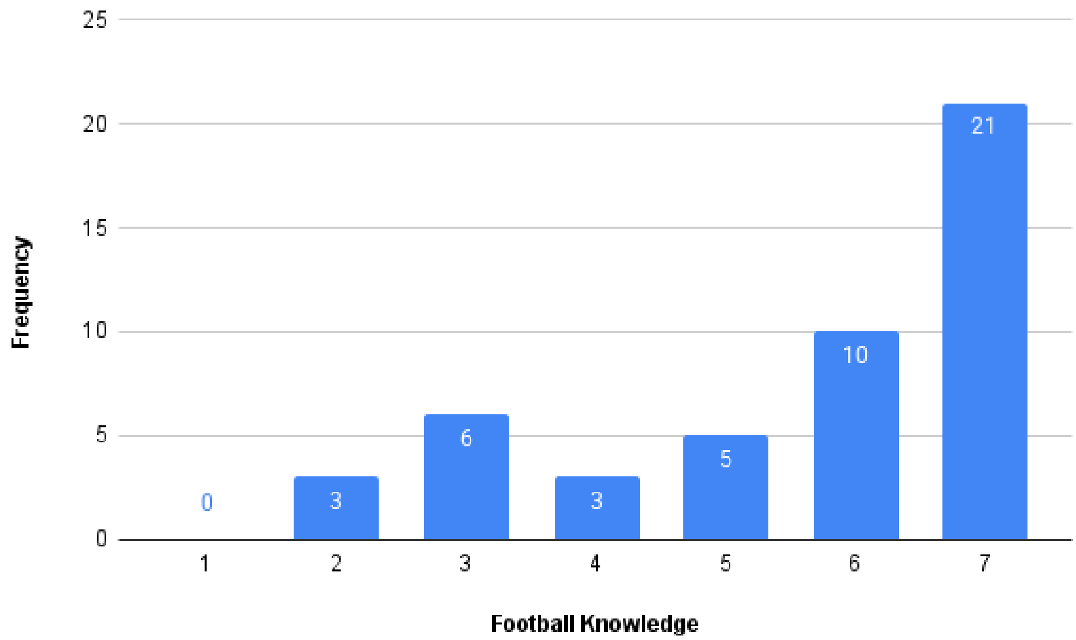



Figure 6.2: Participant distribution by level of football knowledge.

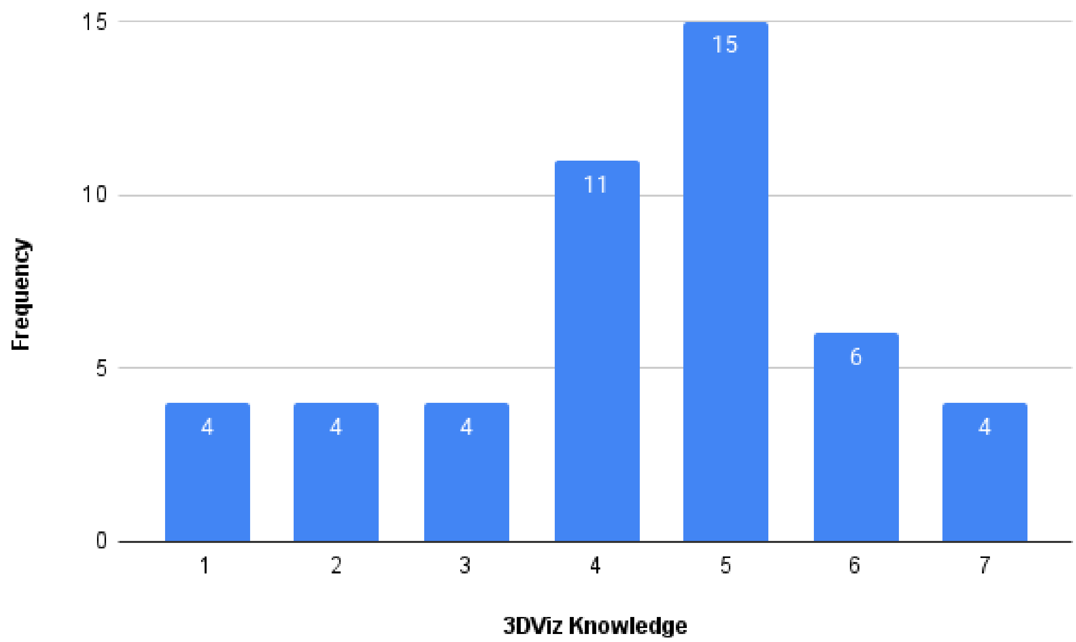



Figure 6.3: Participant distribution by level of 3D visualizations knowledge.

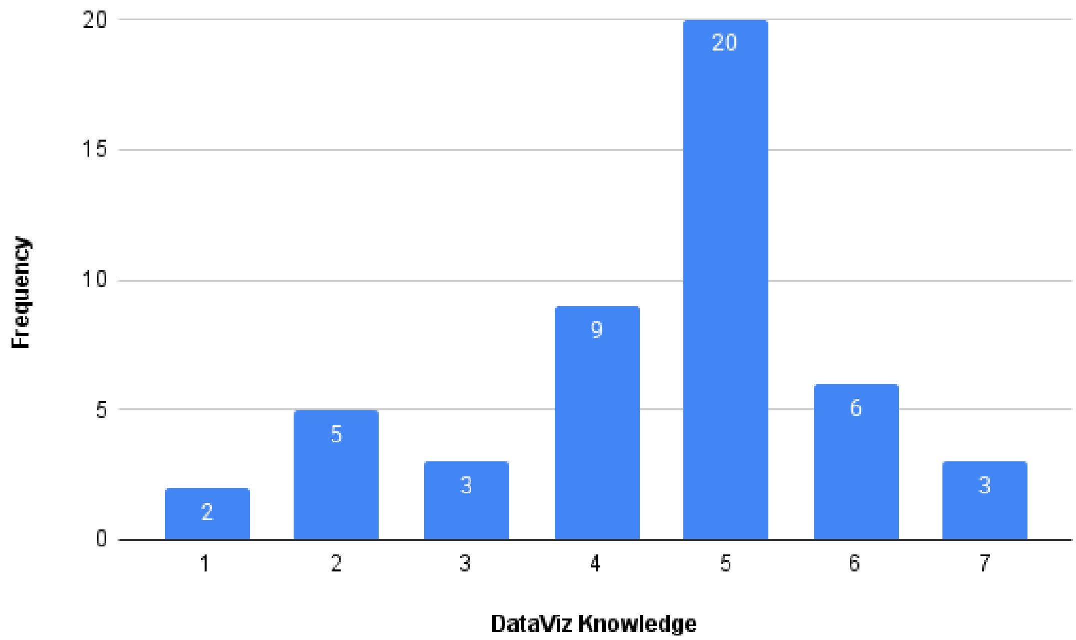



Figure 6.4: Participant distribution by level of data visualization tools knowledge.

This variability indicates a heterogeneous participant pool, which helps evaluate the visualization's effectiveness across users with varying levels of technical background.

### 6.1.3 Task Performance

In order to assess the effectiveness of the visualization and the participants' ability to interact with it meaningfully, task performance was evaluated based on the correctness of each submitted answer. A task was marked as successful if the participant's response matched the predefined correct answer established during the experiment design, as displayed in Section 5.1.4.3. Every question was mandatory, so there were no empty answers. Table 6.3 shows the accuracy rate for each of the seven task-specific questions across the 48 valid submissions.

Table 6.3: Task accuracy across all valid responses.

| Task | Correct Answers | Accuracy |
|---|---|---|
| T1 | 21 / 48 | 43.75% |
| T2 | 32 / 48 | 66.67% |
| T3 | 43 / 48 | 89.58% |
| T4 | 23 / 48 | 47.92% |
| T5 | 38 / 48 | 79.17% |
| T6 | 13 / 48 | 27.08% |
| T7 | 33 / 48 | 68.75% |

From the values presented above in Table 6.3, it can be observed that the number of correct responses varied across the seven tasks, ranging from 13 to 43 correct answers. Tasks T3 and T5 recorded the highest number of correct responses, with accuracies of approximately 80% or above, while T6 had the fewest, with an accuracy of approximately 27%. The detailed distribution of answers per task is described in the following paragraphs.

**Task 1.** 21 out of 48 participants answered correctly (43.75%), identifying *Canada* as the country that received the most outgoing transfers from the USA in 2003. The remaining answers were distributed across multiple countries, most commonly *Mexico* (7 times) and *England/UK* (5 times). There were also 7 other countries selected between 1 and 3 times, as well as one "*I don't know.*" answer. Fig. 6.5 shows the bar chart with the distribution of the answers among participants.

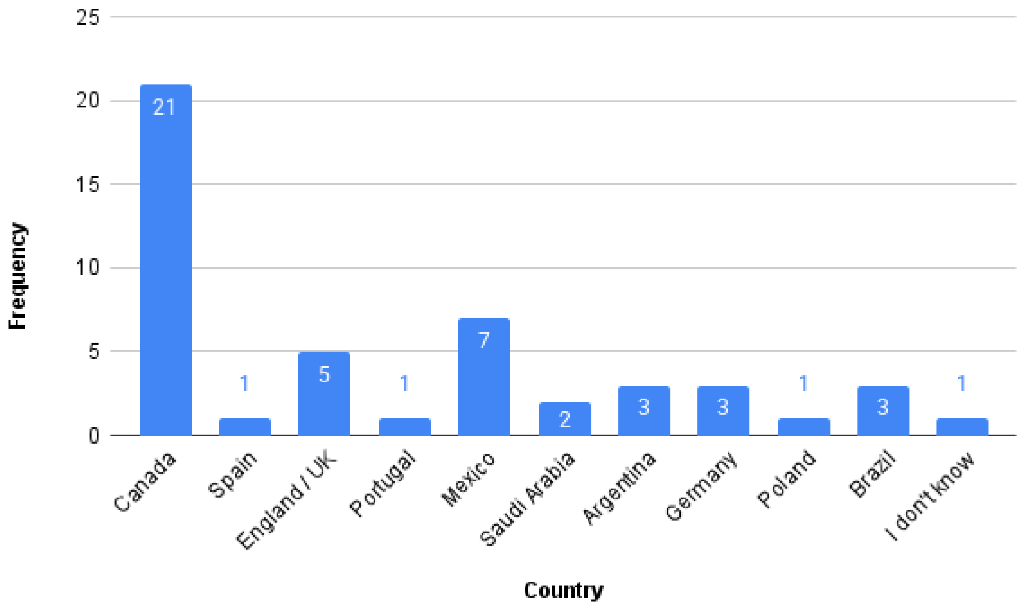



Figure 6.5: Distribution of participants' answers to Task 1.

**Task 2.** 32 out of 48 participants answered correctly (66.67%), identifying *France* as the country that transferred more players to England in 2002. The second and third most given responses were *Spain* (5 times) and *Italy* (3 times), respectively. There were also 5 other countries selected between 1 and 2 times, as well as two "*I don't know.*" answers. Fig. 6.6 shows the bar chart with the distribution of the answers among participants.

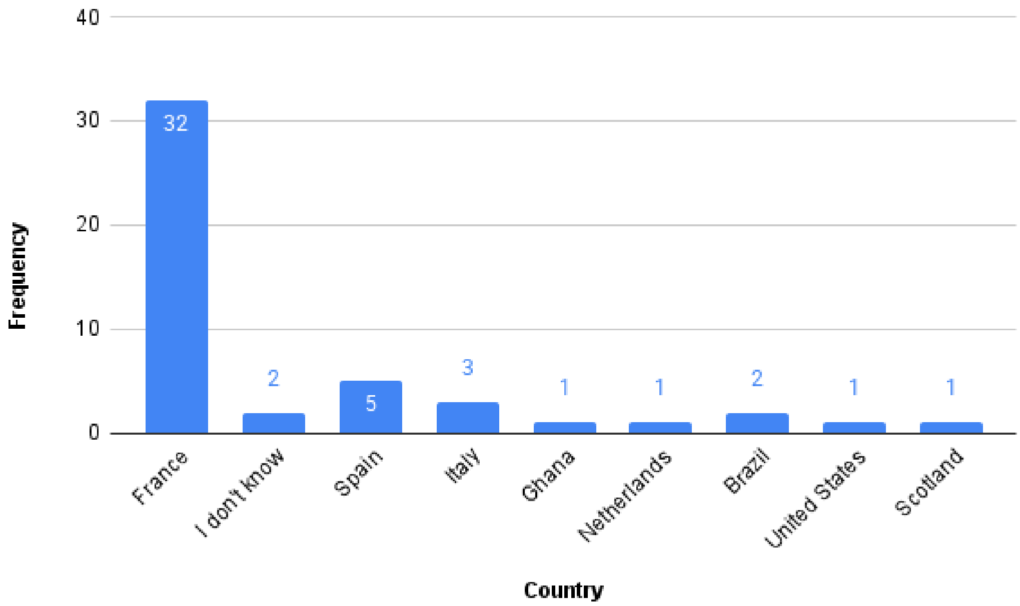



Figure 6.6: Distribution of participants' answers to Task 2.

**Task 3.** 43 out of 48 participants answered correctly (89.58%), identifying *Brazil* as the South-American country that transferred more players to Europe in 1997. The only other countries selected were *Argentina* (2 times) and *Colombia* (1 time). There were also 2 answers stating it was hard to identify that specific country. Fig. 6.7 shows the bar chart with the distribution of the answers among participants.

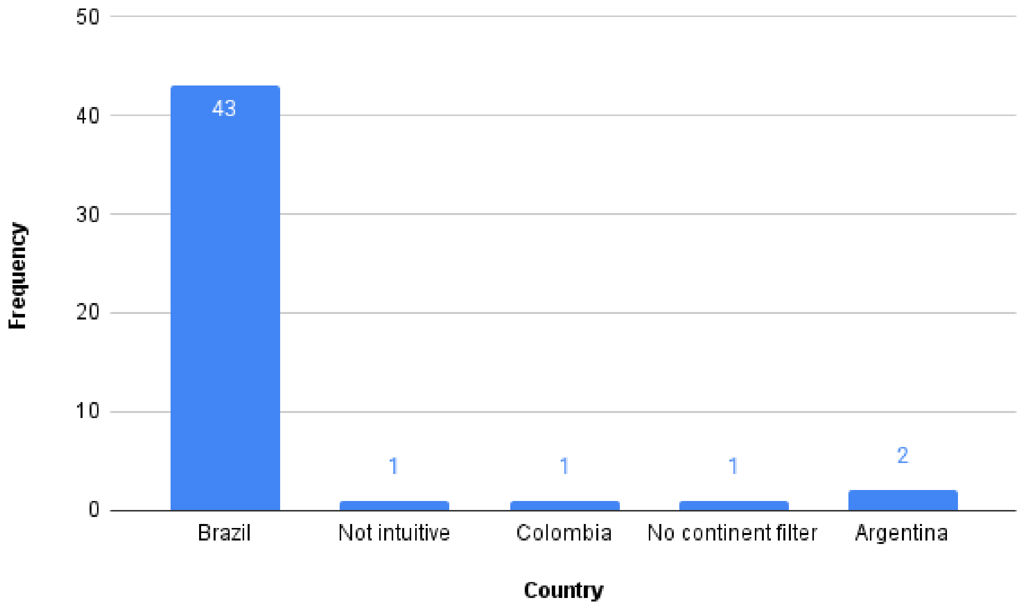



Figure 6.7: Distribution of participants' answers to Task 3.

**Task 4.** 23 out of 48 participants answered correctly (47.92%), identifying 7 different countries where *Nicolas Anelka* played at. The following were the other most common answers: 8 countries (7 times), 5 countries (5 times), 4 countries (4 times), and 6 countries (3 times). There were 6 other different answers, each with a count of one. Fig. 6.8 shows the bar chart with the distribution of the answers among participants.

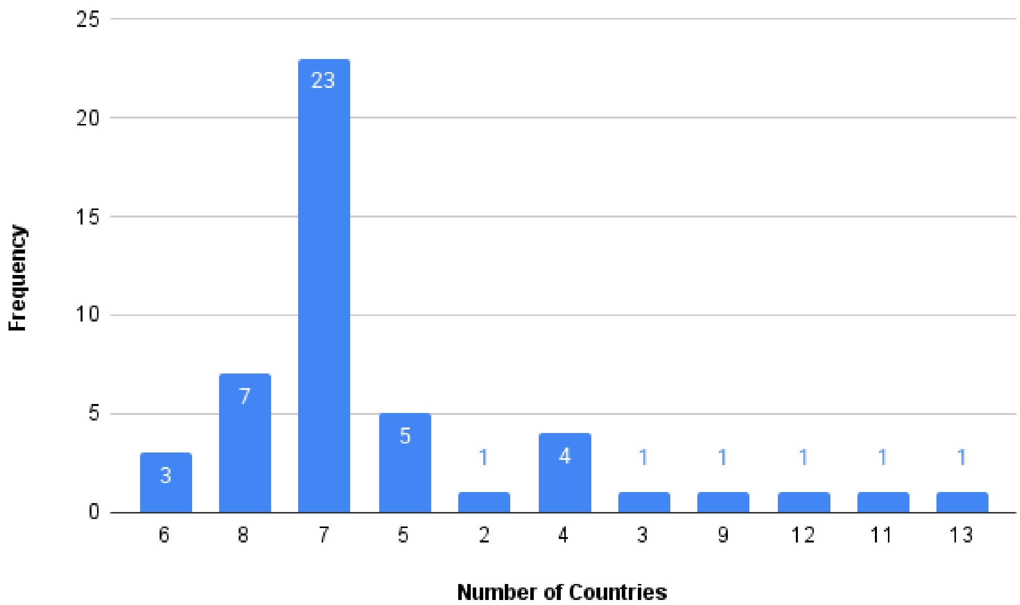



Figure 6.8: Distribution of participants' answers to Task 4.

**Task 5.** 38 out of 48 participants answered correctly (79.17%), identifying *Portugal* and *Brazil* as the pair of countries with the most transfers between each other. Only three different answers were recorded, with them being *France* and *England* (5 times), *Spain* and *Argentina* (4 times), and *France* and *Belgium* (1 time). Fig. 6.9 shows the bar chart with the distribution of the answers among participants.

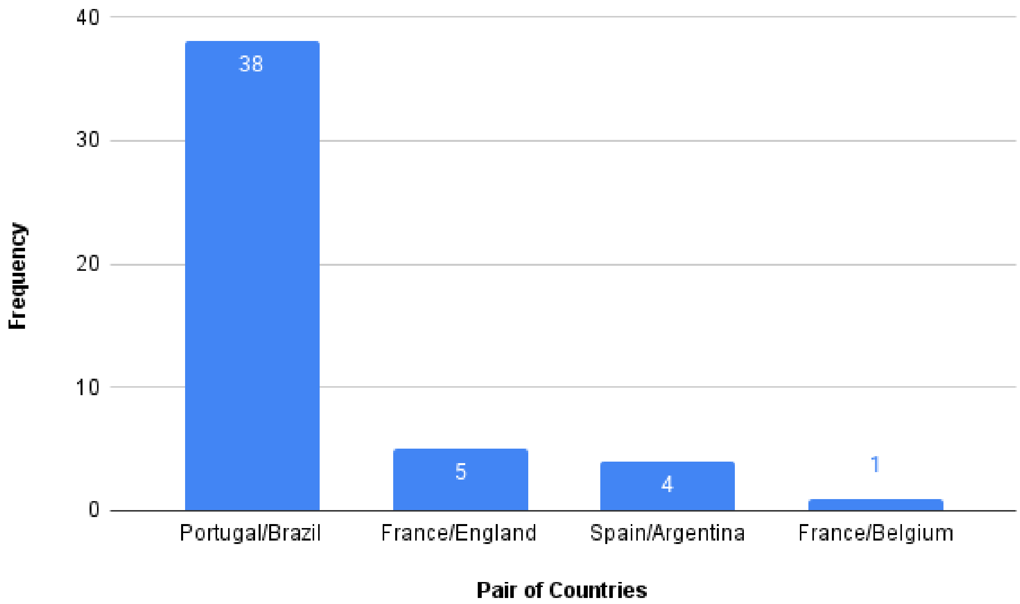



Figure 6.9: Distribution of participants' answers to Task 5.

**Task 6.** 13 out of 48 participants answered correctly (27,08%), identifying 1988 as the first year in which a transfer from *China* to a European country occurred. A close second most given answer was the year of 1995 (12 times). There were 17 other different answers, each being given between 1 and 2 times. Fig. 6.10 shows the bar chart with the distribution of the answers among participants.

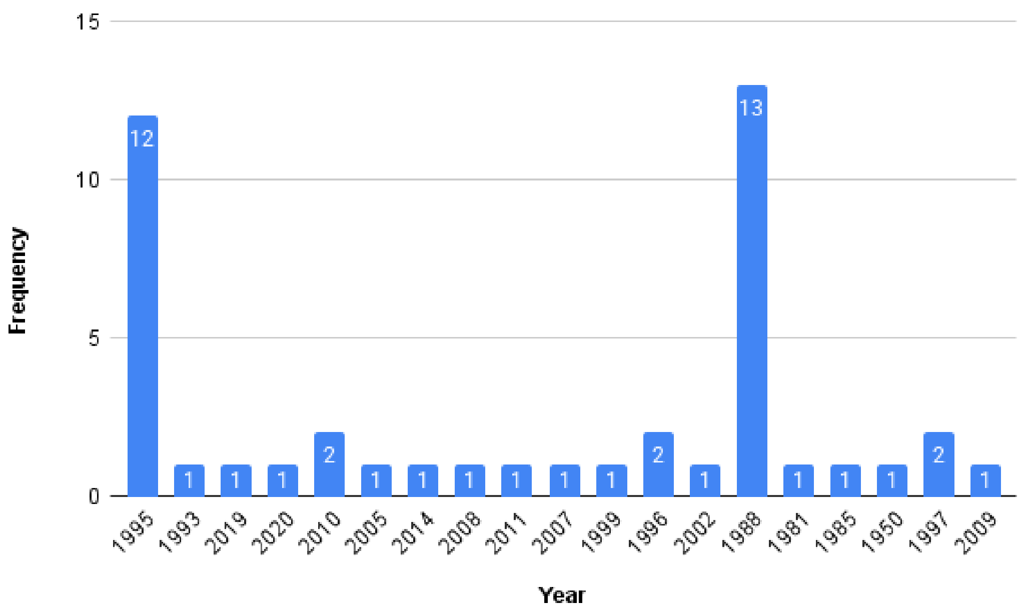



Figure 6.10: Distribution of participants' answers to Task 6.

**Task 7.** 33 out of 48 participants answered correctly (68.75%), identifying *Europe* as the continent that transfers the most players to England over time. Other continents were given as an answer, like *South America / America* (7 times) and *Africa* (1 time). The remaining answers were either single countries like *France* (3 times), *Ireland* (1 time), and *Scotland* (1 time), or answers stating it was hard to identify that specific continent. Fig. 6.11 shows the bar chart with the distribution of the answers among participants.

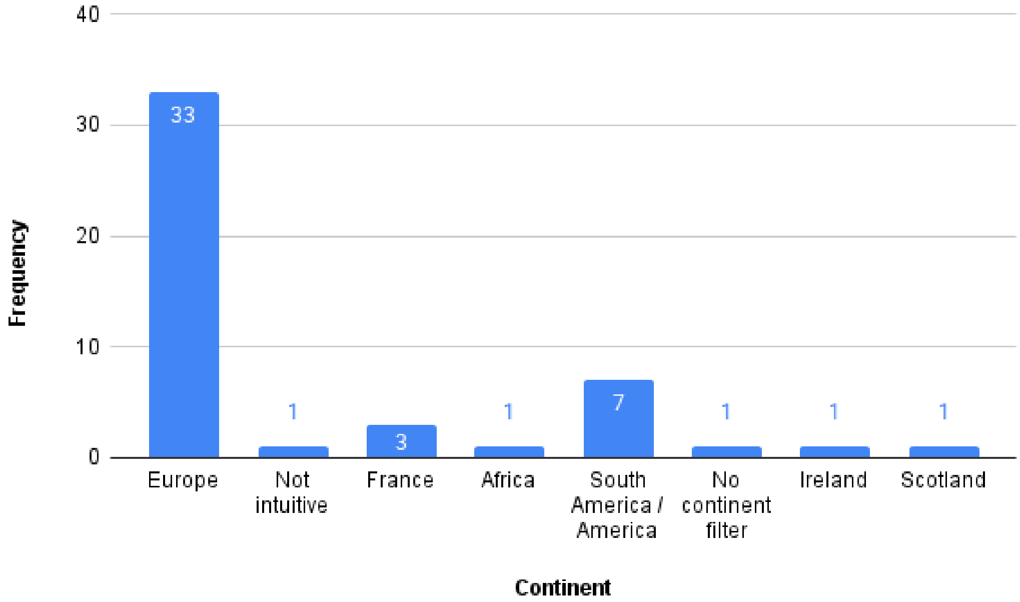



Figure 6.11: Distribution of participants' answers to Task 7.

#### 6.1.3.1 Task Performance per Participant

As observed in Figure 6.12, the number of wrong answers given by each participant varied considerably across all 48 submissions. A total of 6 participants (12.5%) were able to answer all questions correctly, while 7 participants (14.58%) made 1 error and 11 participants (22.92%) made 2 errors, reaching a total of 24 participants (50%) who made 2 or fewer errors. On the opposite side, 9 participants (18.75%) made 5 errors, 2 participants (4.17%) made 6 errors, and no one got all answers wrong, totaling 11 participants (22.92%) who made 5 or more errors. The remaining 13 participants made between 3 and 4 errors, with 5 participants (10.42%) making 3 errors and 8 participants (16.67%) making 4 errors.

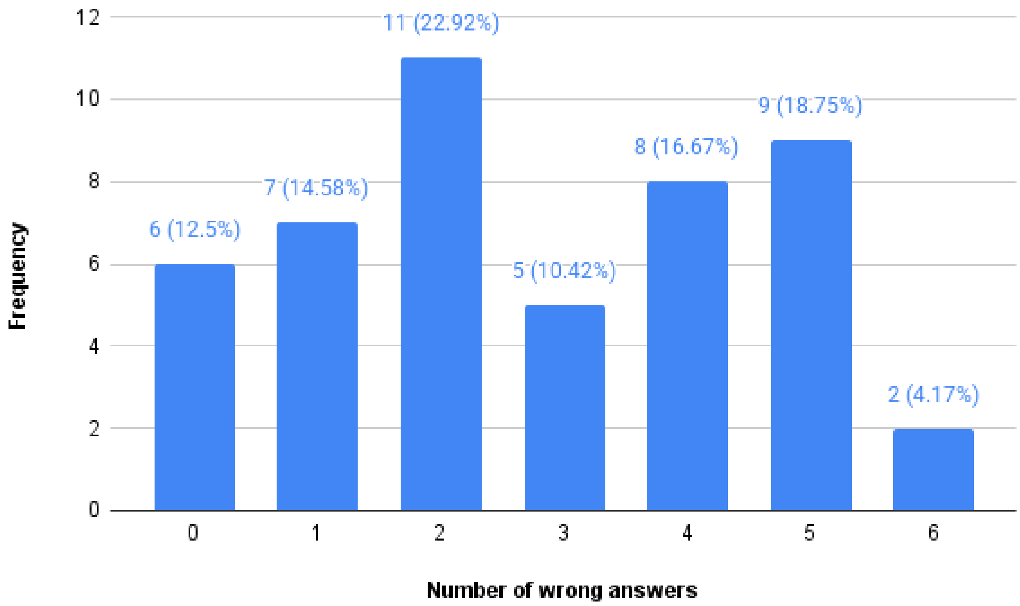



Figure 6.12: Participant distribution per number of wrong answers.

#### 6.1.3.2 Domain Knowledge Influence in Task Performance

In order to understand how prior experience influences task performance, participants self-rated their familiarity with the domains of football, 3D visualizations, and data visualization tools using a 7-point Likert scale. Three different combinations of knowledge levels (two from the three domains at each time) were analyzed using two types of visual tools: bubble charts (Figures 6.13, 6.15, and 6.17) and heatmaps (Figures 6.14, 6.16, and 6.18).

The bubble charts show the distribution of wrong answers among participants with the same knowledge levels, with color representing the number of errors and bubble size indicating frequency. The heatmaps complement this by highlighting the mean number of errors within each pair of domains.

The combination of the domains of **3D Visualizations** and **Data Visualization Tools** (Figures 6.13 and 6.14) shows that participants with higher levels of expertise in both of these domains tend to commit fewer errors. This is backed by the large number and size of blue bubbles (2 or fewer errors) at the upper right corner of the chart and by the low mean errors in the same quadrant of the heatmap.

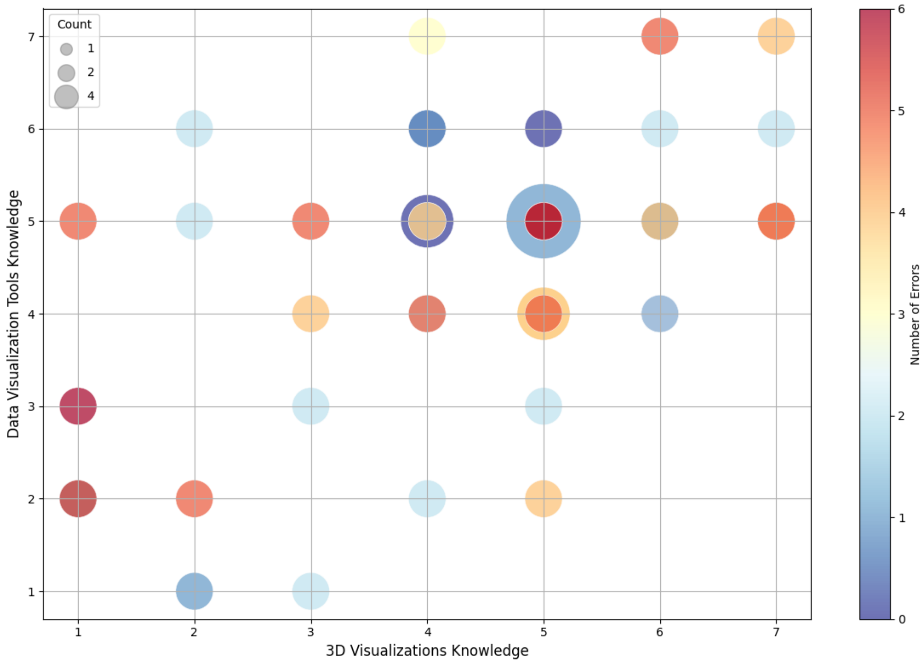



Figure 6.13: Relation between number of wrong answers and domain knowledge levels (3D Visualizations and Data Visualization Tools).

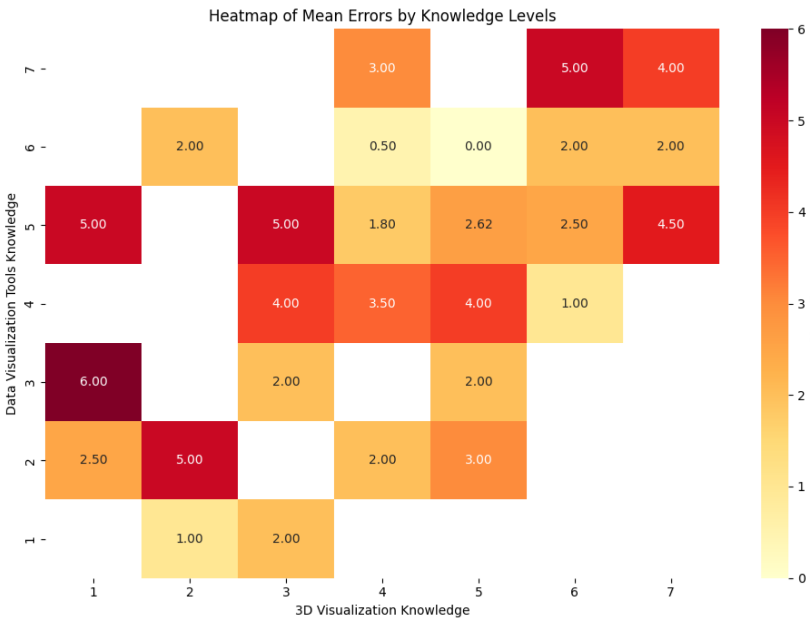



Figure 6.14: Heatmap of the mean of wrong answers for each pair of domain knowledge levels (3D Visualizations and Data Visualization Tools).

In the combination of the domains of **Football** and **Data Visualization Tools** (Figures 6.15 and 6.16), the pattern is less uniform. While some blue bubbles still appear at the upper right corner of the chart (e.g., (6,5) and (7,6)), the distribution is more scattered. The error range of the participants who classify themselves as football experts (7) goes from 0 to 6 errors, suggesting that familiarity with football alone is not enough to obtain a high task accuracy.

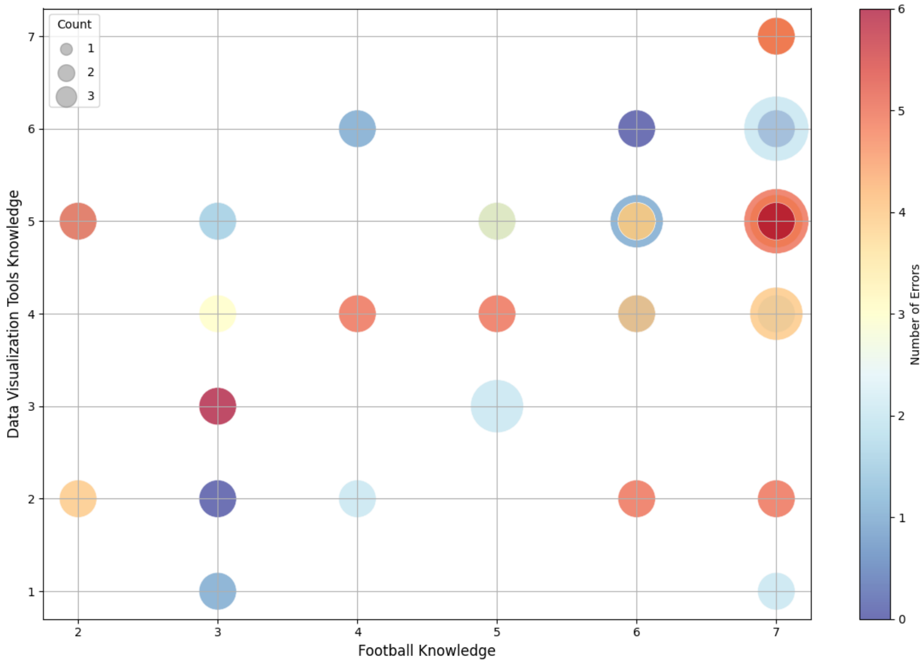



Figure 6.15: Relation between number of wrong answers and domain knowledge levels (Football and Data Visualization Tools).

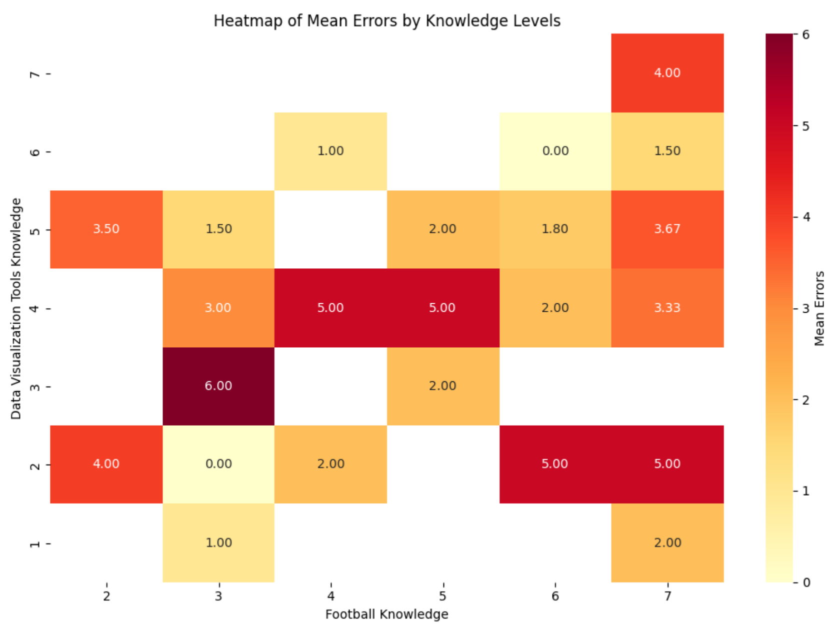



Figure 6.16: Heatmap of the mean of wrong answers for each pair of domain knowledge levels (Football and Data Visualization Tools).

Lastly, the combination of the domains of **Football** and **3D Visualizations** (Figures 6.17 and

6.18) shows something similar to the previous combination. There are also some blue bubbles at the upper right corner of the chart (e.g., (6,4) and (6,5)) but the error range of the participants who classify themselves as football experts (7) also goes from 0 to 6 errors, contributing to the suggestion that familiarity with football alone does not guarantee a good performance.

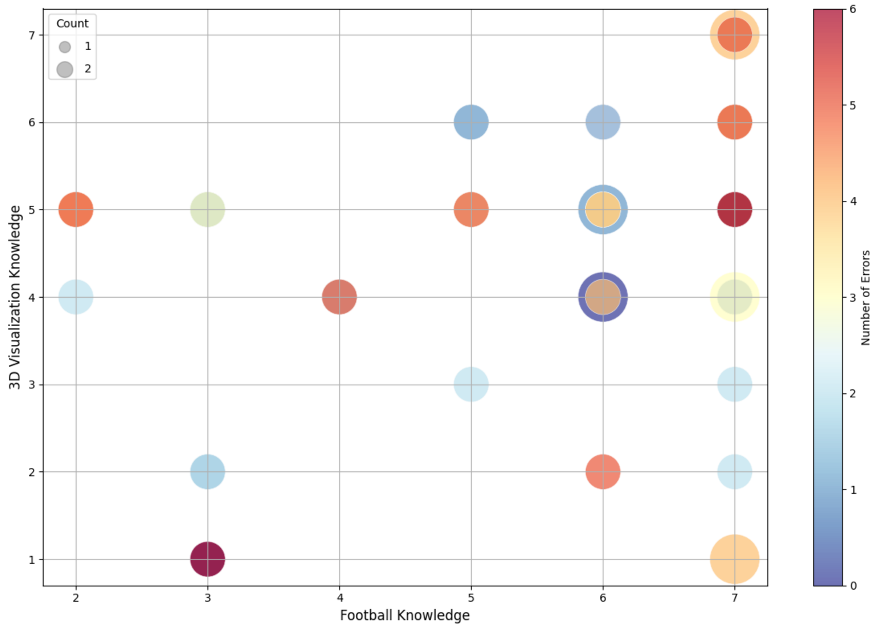



Figure 6.17: Relation between number of wrong answers and domain knowledge levels (Football and 3D Visualizations).

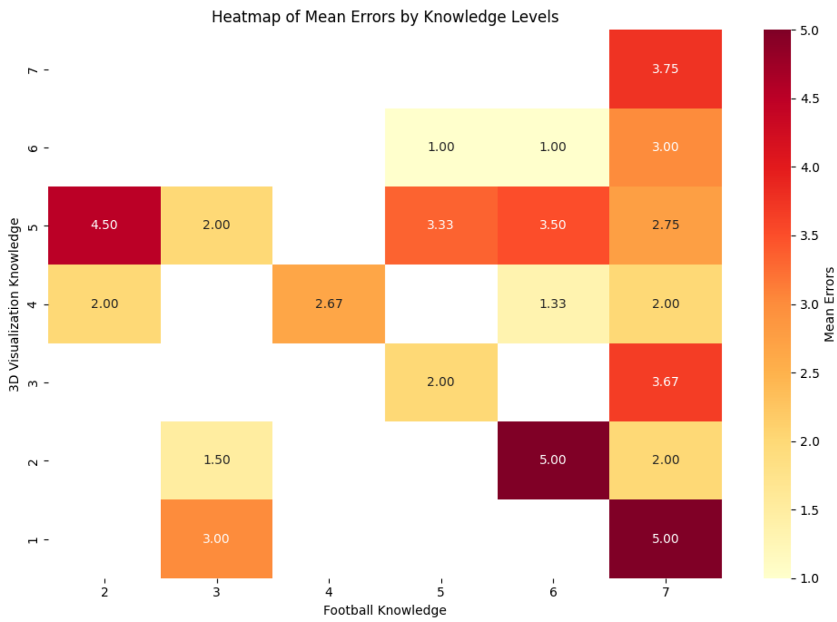



Figure 6.18: Heatmap of the mean of wrong answers for each pair of domain knowledge levels (Football and 3D Visualizations).

To statistically support the visual findings, an ordinal regression model [44] was applied using the participants' self-rated knowledge in the three domains (**Football**, **3D Visualizations**, and **Data Visualization Tools**) as predictors and the number of wrong answers as the ordinal outcome variable. The results can be observed in Table 6.4.

Table 6.4: Ordinal regression coefficients for each domain knowledge.

| **Domain Knowledge** | **Coefficient** | **Interpretation** |
|---|---|---|
| Football | +0.0098 | Weak positive influence |
| 3D Visualizations | -0.0442 | Slight negative influence |
| Data Visualization Tools | -0.0179 | Slight negative influence |

This method estimates the probability of a response falling at or below a given category as a function of predictor variables, assuming that the effect of each predictor is consistent across the response levels. The resulting coefficients indicate the direction and relative strength of the influence of each domain of knowledge on the likelihood of committing more errors. A negative coefficient suggests that higher expertise in that domain is associated with fewer errors, while a positive coefficient suggests the opposite.

In this case, the results align with the visual analysis: the most influential domain appears to be **3D Visualizations**, with a slight negative coefficient, indicating that greater familiarity with 3D concepts may contribute to improved task performance. **Data Visualization Tools** also shows a small negative influence, also indicating the same influence on task performance. Interestingly, **Football** knowledge exhibits a weak positive coefficient, reinforcing earlier observations that football familiarity alone does not predict higher task accuracy.

It is also useful to explore how domain-specific knowledge might have affected participants' performance on individual tasks. The frequency of right and wrong answers for a particular task (T1-T7) across the different self-assessed knowledge levels can be observed in Annex C (Figures C.1 to C.7). Instead of focusing only on overall performance, these bar charts offer a more extended look at how domain familiarity may be associated with greater accuracy in answering specific questions.

While no single domain can accurately predict success in every task, some significant patterns can be seen. Participants with higher self-rated expertise in both **3D Visualizations** and **Data Visualization Tools** performed significantly better in Task 6 (T6) - the question with the lowest overall accuracy - indicating that these types of technical knowledge were helpful to understanding that specific visualization. However, there does not seem to be a direct correlation between task success and football expertise. Indeed, individuals who assessed their own football knowledge at the highest level (7) frequently performed less accurately than those who gave themselves mid-to-high ratings (5 or 6). This could be a sign of an inflated self-evaluation or over-reliance on past football knowledge, which could result in inaccurate assumptions and a lack of attention to detail when analyzing the visual content. It is important to notice that the sample leans toward higher self-assessments (5 to 7), so care should be taken when interpreting this analysis.

### 6.1.4 Usability Evaluation

As mentioned in Section 5.1.4.4, participants used a 5-point Likert scale to rate the 10 items in the SUS questionnaire, in order to express their overall satisfaction with the system usability and effectiveness. The minimum value, 1st quartile, median value, 3rd quartile, maximum value, and standard deviation are observable in Table 6.5 and Figure 6.19.

Table 6.5: Summary statistics for each SUS item.

| | **Min.** | **1st Quartile** | **Median** | **3rd Quartile** | **Max.** | **Std. Deviation** |
|---|---|---|---|---|---|---|
| **I1** (Appreciation) | 1.00 | 3.00 | 3.00 | 4.00 | 5.00 | 1.16 |
| **I2** (Complexity) | 1.00 | 1.75 | 2.00 | 3.00 | 5.00 | 1.06 |
| **I3** (Ease of use) | 1.00 | 3.00 | 4.00 | 4.00 | 5.00 | 0.85 |
| **I4** (Ease of use) | 1.00 | 1.00 | 2.00 | 2.00 | 4.00 | 0.88 |
| **I5** (Consistency) | 2.00 | 3.75 | 4.00 | 4.00 | 5.00 | 0.71 |
| **I6** (Consistency) | 1.00 | 1.00 | 2.00 | 3.00 | 3.00 | 0.80 |
| **I7** (Ease of use) | 1.00 | 3.00 | 4.00 | 4.00 | 5.00 | 1.01 |
| **I8** (Ease of use) | 1.00 | 2.00 | 2.00 | 3.00 | 4.00 | 0.83 |
| **I9** (Confidence) | 2.00 | 3.00 | 4.00 | 4.00 | 5.00 | 0.90 |
| **I10** (Ease of use) | 1.00 | 1.00 | 2.00 | 3.00 | 5.00 | 0.94 |

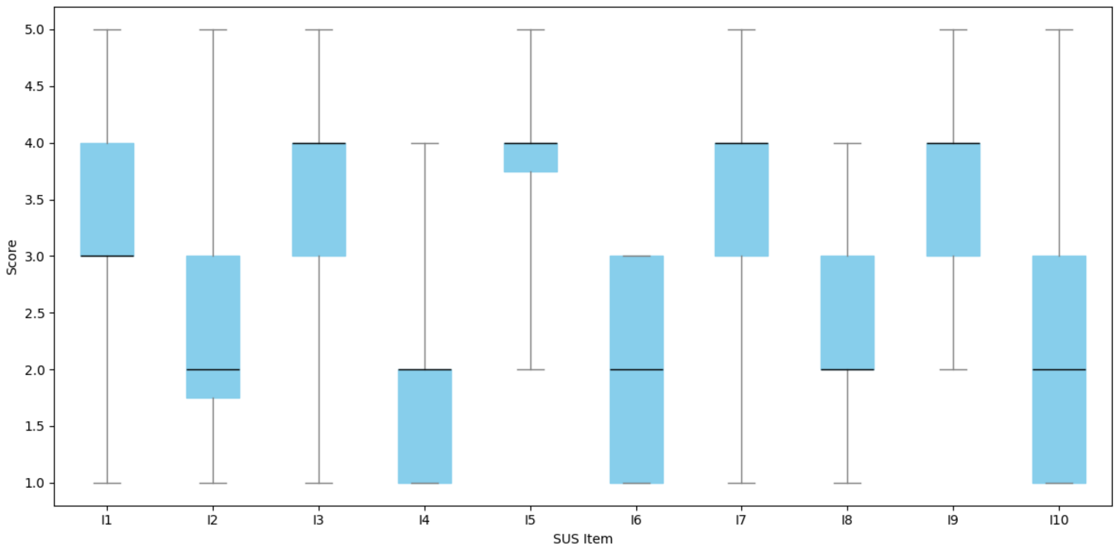



Figure 6.19: Summary statistics for each SUS item.

Since the odd-numbered SUS items (I1, I3, I5, I7, and I9) are positively phrased (higher values represent more favorable evaluations) and the even-numbered SUS items (I2, I4, I6, I8, and I10) are negatively phrased (lower values represent more favorable evaluations), as already mentioned in Section 6.1.1, these results show that participants generally had a positive experience with the visualization.

The 1st quartile values of the odd-numbered items are always equal or greater than 3.00, and their median values are 3.00 for one of them and 4.00 for four of them, which is relatively high. On the other hand, the 3rd quartile values of the even-numbered questions are always equal to 3.00 or lower, and their median values are 2.00, which is also positive. The standard deviation values range between 0.71 and 1.16, showing a reasonably constant range, whereas there is some response variability. Items with exceptionally high quartile values like I5 (" *found the system's functions well integrated.*") reinforce the idea that most users find the system consistent and well integrated.

In order to measure overall user satisfaction, each response's System Usability Scale (SUS) score was calculated by applying a specific transformation (already described in Section 6.1.1) to each of the 10 questionnaire items, scaling the score to a 0-100 range. The average SUS score for all the participants was 68.80, which, according to the table in [43], corresponds to a B grade. This suggests that participants generally found the system to be effective, efficient, and satisfactory to use, indicating a good level of usability. Although there is still room for improvement in some areas, this result supports the presented statistical information and the conclusion that the visualization offers a positive user experience.

### 6.1.5 Qualitative Feedback

As described in Section 5.1.4.4, there were three optional open feedback questions at the end of the questionnaire so users could better express how they felt about the system. Among the 48 participants whose submissions were considered valid, 27 of them decided to provide this additional feedback, highlighting how they viewed and engaged with various functionalities and providing additional context to the quantitative findings. Their feedback was grouped according to the three open questions in the form: most intuitive aspects, most confusing aspects, and suggestions for improvement.

#### Most Intuitive Aspects

- Eight participants found the **filter menu** to be one of the most intuitive features, praising its clarity and usefulness for narrowing down information.
- Five participants highlighted the **overall visual design** as clear and self-explanatory, making the interface easier to interpret.
- Four participants found the **temporal navigation** easy to use and helpful for exploring transfers over time.
- Three participants appreciated the use of **arc thickness** to represent transfer volume, finding this visual aspect easy to interpret.
- Two participants pointed out that the **player filter** was particularly easy to use for locating individual player data.

- Two participants mentioned the **ease of navigation** around the globe and the responsiveness of the system as strengths.
- One participant praised the **interactivity** of the visualization overall.
- One participant highlighted the **1-1 transfer filter** as being especially straightforward.
- One participant commented that the **globe itself** was an intuitive way to represent geographical data.
- One participant found it easy to identify **connections between countries**, without needing extra guidance.

**Most Confusing Aspects**

- Seven participants pointed to the **overlapping of arcs** as the most confusing issue, making it difficult to identify a specific arc.
- Four participants found the **tooltip on arcs** unclear or difficult to interpret when hovering over transfers.
- Three participants mentioned parts of the **filter menu** as confusing or not intuitive, in contradiction with the users who valued this functionality.
- Three participants reported difficulty in interpreting **connections between continents**, suggesting that these were not clearly visualized.
- Two participants were unsure how to use the **timeline navigation**, either due to unfamiliarity with the controls or a lack of clear cues.
- Two participants specifically criticized the **year selection** interface as difficult to use.
- One participant had trouble interpreting **transfer arc visuals**, particularly in crowded areas.
- One participant was confused about how **arc thickness** represented transfer counts.
- One participant expressed frustration with not being able to **view multiple countries simultaneously**.
- One participant wanted a way to **compare transfers between countries** directly and found it hard to do with the current setup.

**Suggestions for Improvement**

- Eight participants suggested adding a **table with detailed transfer information**, enabling users to see numerical data alongside the visual interface.
- Two participants recommended implementing **more zoom functionality**, making it easier to explore specific regions of the globe in detail.

- Two participants suggested adding a **country-to-continent filter** so that those trends can be easily observed.
- Two participants requested that the **selected country change its color on the globe** to improve clarity when focusing on specific areas.
- One participant suggested adding **country flags to the player filter dropdown**, in order to better identify the players.
- One participant asked for an **improved year selection interface** to allow smoother interaction with the timeline.
- One participant asked for an **improved player view**, making it easier to analyze or compare players.
- One participant suggested refining the **tooltip**, making it easier to interpret data.
- One participant asked for a **player nationality filter** to further customize search and analysis.
- One participant suggested adding a **continent filter**, which already exists, and a **player position filter** to also customize search and analysis.
- One participant requested **full globe rotation**, enabling users to reposition the view freely to focus on any region.
- One participant preferred having **data access come first**, using the 3D globe only as a secondary, visual complement to the information.
- One participant suggested **improving the appearance of transfer arcs**, to make them easier to follow visually.
- One participant mentioned that the system would benefit from **more data but a less overwhelming presentation**, implying a need for simplification or improved filtering.
- One participant proposed using **extra colors to distinguish the destination continent** of transfer arcs, increasing visual clarity.

In general, the provided feedback was positive and showed active engagement with the visualization. While a few participants identified minor areas of confusion, many praised the system's user-friendly features and made helpful recommendations for future enhancements.

## 6.2 Discussion

This section covers the findings and implications of the previously presented results, offering insights into the 3D visualization's usability and effectiveness, as well as potential improvements based on participant feedback and task performance.

### 6.2.1 Outlier Identification and Data Reliability

The purpose of the outlier identification technique used in Section 6.1.1 was to guarantee data reliability in a remote and uncontrolled environment where participants carried out the experiment on their own. Submissions that most likely showed disengagement or mechanical answering were filtered out with the aid of this method, which was only based on identifying patterned or statistically extreme SUS responses. It is possible, however, that isolated atypical responses (possibly resulting from actual confusion or misinterpretation) were kept because the criteria focused on complete submissions rather than individual task answers. While the used technique increases the dataset's overall robustness, noise at the level of individual task answers might not be completely removed.

### 6.2.2 Demographics and Recruitment Strategies

As observed in Section 6.1.2, the demographic statistics show a reasonably balanced sample in terms of gender and age distribution, with participants ranging from football fans to users with extensive knowledge of 3D or data visualization tools, successfully brought together by the recruitment strategies employed. This diversity offers a strong basis for evaluating the visualization's effectiveness and usability for various user types. The visualization's sports-oriented focus is supported by the high self-reported football knowledge, and both inexperienced and experienced users tested the system thanks to the wider distribution of technical expertise. Such a diverse participant pool enhances the findings' generalization and strengthens the validity of the drawn conclusions.

### 6.2.3 Task Performance and Domain Knowledge Influence

The results show a considerable variation in task accuracy, with tasks like T3 and T5 reaching over 79% accuracy, while other tasks like T1 (43.75%), T4 (47.92%), and especially T6 (27.08%) had significantly lower success rates. These disparities highlight the impact and influence that task design and cognitive load can have. T3, which asked participants to identify the South American country with the most outgoing transfers to Europe in a specified year, benefited from having a clear and visually salient answer (*Brazil*), depicted with thick and prominent arcs. Similarly, T5 achieved high accuracy in its responses, possibly explained by the straightforward filtering operation necessary for its completion, aligned with the fact that it was a multiple-choice question instead of being an open question like other tasks.

On the other hand, T1 asked participants to identify the country with the most incoming transfers from the United States in a specified year by also visually comparing arc thickness, but achieved less than half of the accuracy of T3. This problem was probably caused by visual clutter and arc overlapping, which made the task more visually demanding, as it did not have a clear and visually salient answer. T4 asked users to identify how many countries Nicolas Anelka played in, and the key to the low accuracy achieved might be the fact that, in his case, he played in the same country on multiple occasions (he was transferred to England 5 times during his career).

This might have caused some misunderstandings, as it required close attention from the user to avoid being misled into the wrong answer.

The worst-performing task (T6) required both navigating the timeline and recognizing subtle arcs on the globe, which was found challenging by most of the participants. The second most given answer to this question (1995), given by 12 participants and only 1 less time than the correct answer, represents the second time that a player was transferred from China to Europe. This leads to the belief that some players missed the arc representative of the first transfer and were only able to identify the one relative to the second transfer, proving that less noticeable visual elements were hard to isolate and identify. One participant even identified 1950 as the correct answer - the year of the first transfer from Macau to Europe. It was considered a wrong answer due to the fact that Macau was a Portuguese territory back then and only returned to Chinese control in 1999.

Domain knowledge also seems to have an impact on performance. Participants with high self-reported familiarity in both 3D visualizations and data visualization tools tended to perform better overall, and particularly on more complex questions, while football knowledge by itself did not increase task accuracy. In fact, football experts were occasionally outperformed by those with moderate knowledge, evidencing the fact that reliance on past knowledge may have resulted in assumptions rather than proper exploration of the visualization. This is further supported by the results of the ordinal regression, which indicated that knowledge of 3D visualization and data visualization tools was negatively correlated with errors, in contrast to football knowledge, which was positively correlated with errors.

### 6.2.4 System Usability and Perceived Experience

As shown in Section 6.1.4, the system obtained a B grade based on its average SUS score of 68.80, suggesting that users generally found the visualization to be useful and satisfactory. The majority of participants believed the system to be reliable, well integrated, and reasonably simple to use, as evidenced by the low values in the negatively phrased items and the high values in the positively phrased items.

However, there was more variation in responses regarding some aspects, including complexity, ease of use, and the need for technical support. The participants' diverse backgrounds and different levels of experience with visualization tools are the probable causes of this variation. Many users found the system easy to use, but some had trouble with certain interactions. This suggests that the system needs better visual guidance, clearer onboarding processes, or tutorials to help less experienced users.

These findings are supported by the qualitative feedback collected from the open-ended questions, displayed in Section 6.1.5. Several participants highlighted features like the filter menu or the visual design as particularly intuitive, which is consistent with the positive SUS results. At the same time, recurrent points of confusion, such as overlapping arcs or difficulties observing the correct tooltips, help to explain the variability mentioned above.

The suggestions' emphasis on functional improvements rather than fundamental issues suggests that users thought the system was good but could be improved. Participants suggested enhancements aimed at increasing efficiency and understanding rather than challenging its usefulness, which implies that they view value in this type of visualization. The feedback highlights a gap between user expectations and system functionality, especially in terms of clarity and ease of interpretation.

Taken together, the SUS scores and qualitative feedback show that the visualization effectively supports exploratory analysis but can be improved, especially in terms of assisting inexperienced users and providing additional and complementary ways to interpret the data.

### 6.2.5 Final Remarks

This last section responds to the research questions presented in Section 5.1.1 and draws conclusions from the findings examined in this chapter.

**Q1: Can users interpret the spatial and temporal transfer patterns displayed?**
Participants were generally able to interpret spatial patterns, according to the task performance analysis, especially in T3 and T5. This was particularly true when there were strong visual cues like thick arcs or limited filtering scopes. Nonetheless, the reduced accuracy in T1, T4, and T6 indicates that in more intricate or visually cluttered situations, the ability to detect spatial and temporal patterns may deteriorate. Misunderstandings of subtle temporal transitions (T6) or transfer arcs overlapping (T1) suggest that some design improvements, such as improved timeline interactions or clearer transfer arcs, might help users to better identify these patterns.

**Q2: Can users easily interact with the visualization to extract insights?**
The distribution of participant responses and the overall SUS score of 68.80 (Grade B) show that the system's usability is generally evaluated as positive. The majority of users thought that the visualization made sense and was simple to use. However, usability issues that could prevent some users from extracting valuable information are highlighted by the variation in individual item responses and the frequent qualitative feedback regarding problems like overlapping arcs or tooltip clarity. This implies that even though the system facilitates interaction effectively overall, contextual assistance and better onboarding would be advantageous for users who are less familiar with data or 3D visualizations.

**Q3: Does the visualization maintain user engagement and encourage exploration?**
Various users interacted with multiple aspects of the interface and provided insightful recommendations for improvements, as supported by the qualitative feedback. They engaged with filters, timeline controls, and spatial elements in diverse ways, and some even recognized subtle problems like the need for additional tables with more detailed information, which suggests an exploratory mindset. The volume and specificity of feedback suggest that the system was successful in encouraging active engagement, even though some users might have relied more on past football knowledge than on actually exploring the visualization. Still, the observed

challenges with subtle interactions offer a chance to improve the flaws and further increase user engagement.

Overall, the results show that the 3D visualization can effectively represent significant patterns in cross-border football player transfers and facilitate exploratory analysis among a diverse user base. However, there is still room for improvement in terms of usability and interpretability, particularly in terms of reducing visual clutter, adding complementary ways of interpreting the data, and providing better assistance for less experienced users.

## 6.3 Summary

This chapter presented the results of the experiment conducted to assess the usability and effectiveness of the 3D globe visualization for depicting cross-border football player transfers. The formative questionnaire provided insights into participant demographics, task performance, perceived usability, and suggestions for improvement. To guarantee the reliability of the data, outlier submissions were identified and subsequently eliminated.

The results show that participants gave a positive SUS evaluation and demonstrated satisfactory task accuracy, indicating that the visualization was generally well-received. Qualitative feedback provided additional insights into the user experience by highlighting both areas of strength and room for improvement. These findings support the idea that 3D visualizations have the capacity to effectively display complex temporal and spatial patterns to a broad audience in an intuitive and engaging way.

# Chapter 7

# Conclusions

This final chapter summarizes the main contributions of this work and reflects on the results obtained by the experimental evaluation of the proposed solution. Section 7.1 revisits the research questions and synthesizes findings from the previous chapters to answer them. Section 7.2 outlines possible directions for future improvements to the system.

## 7.1 Final Remarks

This dissertation is based on the hypothesis that 3D visualizations can effectively represent geographically distributed, time-evolving data with dynamic relationships, while preserving the narrative within the visualization. It tackles a significant challenge faced by ZOS: presenting complex, dynamic, and interconnected data in a clear and engaging way to a broad audience. Given the volume and complexity of cross-border player transfers, the main objective was to study and propose an approach to enable users with varying levels of football knowledge or data visualization experience to understand spatial and temporal patterns. The proposed approach centers on a 3D globe-based visualization that emphasizes both the geographic and chronological dimensions of cross-border player movements. A set of research questions guided the study, aiming to identify the most suitable 3D techniques for representing interconnected and time-evolving data, evaluate the use of globe-based interfaces in similar scenarios, and explore the limitations and challenges of this form of visual storytelling.

A Systematic Literature Review (SLR) was conducted to explore different visualization techniques suited for interconnected and time-evolving data, focusing on 3D representations. It identified common issues like the challenge of maintaining visual clarity in dense networks or the cognitive load involved in interpreting 3D perspectives. Despite this, the results reinforced that 3D visualizations can improve user understanding, particularly when complex relationships and temporal transitions are present. The literature review also illustrated the importance of integrating interactivity, storytelling, and a good spatial representation to enhance the accessibility and engagement of dynamic datasets in the context of global sports data. These findings directly influenced the proposed solution's design.

The State-of-the-Art analysis extended the literature review by examining novel visualization techniques that apply to temporal, spatial, and interconnected data, especially in the sports industry. It addressed tools such as dynamic timelines, 3D globe-based systems, and knowledge graphs, highlighting how they can help to manage complexity, improve interactivity, and support user comprehension. Integrating network, temporal, and spatial dimensions emerged as a key trend, with solutions like hybrid 2D-3D layouts, trajectory mapping, and visual storytelling to increase clarity and user engagement. However, despite their relevance and applicability, not many of these techniques are directly applied to sports data, particularly in the context of global football player transfers. This indicates the existence of a gap that this dissertation aims to fill.

The proposed solution was reflected in a prototype created with Three.js, a robust JavaScript 3D rendering library, chosen for its compatibility with modern browsers and alignment with ZOS's web-based structure. Using animated arcs across a 3D globe, the system seamlessly combines temporal and spatial dimensions to display cross-border football player transfers. Thanks to the implementation of timeline controls and filtering mechanisms, users can examine data from various years, nations, and players. The primary concern was ensuring visual clarity and performance optimization, due to the large dataset's nature and the existence of overlapping elements. The visualization supports multiple interaction modes, such as a one-to-one transfer filter and a dedicated player view.

An experiment was conducted to evaluate how well the 3D visualization could represent temporal and spatial patterns regarding cross-border football player transfers, as well as its usability, interpretability, and user engagement. It was carried out remotely via an online platform, lasted approximately two weeks, and involved a wide range of participants with diverse levels of football and data visualization knowledge. The methodology ensured qualitative and quantitative data collection by combining demographic information, self-rated (from 1 to 7) levels of domain knowledge, task-specific questions, a System Usability Scale (SUS) form, and open-ended feedback questions. The collected data allowed for an extensive evaluation of user comprehension, interaction, and overall satisfaction with the prototype.

The experiment results showed that the visualization was generally well-received and that participants could understand the temporal and spatial patterns of player transfers with relative ease. This is reflected in the high accuracy achieved in some tasks (particularly those involving more noticeable visual cues) and the overall B grade on the System Usability Scale, with an average SUS score of 68.80. Nonetheless, while most users interacted confidently with the system, performance varied depending on the complexity of the tasks and the participants' technical background. While reliance on football knowledge alone occasionally resulted in incorrect assumptions, users who were more familiar with 3D visualizations and data visualization tools performed better. The qualitative feedback complemented these findings by highlighting both the strengths of the interface, such as its filtering mechanisms or its visual design, and areas for improvement, particularly visual clarity regarding arc overlapping.

Based on the work performed and the results achieved, the research questions posed at the beginning of this dissertation can now be addressed:

**Q1: Which 3D visualization techniques are commonly used to present time-evolving interconnected data?**
The literature review identified animated force-directed graphs and globe-based arc visualizations as prominent techniques. Methods such as arc layering, node-link encoding, and temporal animation help preserve relational continuity over time. However, challenges like occlusion and scalability remain significant obstacles.

**Q2: Which visualization techniques are commonly used to present time-evolving data?**
Standard approaches include time series plots, small multiples, and timeline visualizations, often enhanced with animated transitions to reveal changes between time slices. When dealing with large or complex datasets, interactivity and filtering play key roles in mitigating cognitive overload.

**Q3: Which 3D visualization techniques using a globe are available to represent the relationships between players, countries, clubs, and leagues, and how do these techniques accommodate temporal variations?**
Few systems directly apply 3D globe-based visualizations to football data. Still, tools like Globe.gl and custom Three.js implementations use arc animations, hover effects, and time sliders to show geographic and temporal patterns. These solutions are effective for broad exploration, yet often lack the granularity and control needed for detailed user-driven analysis.

**Q4: What are the challenges and limitations of using 3D visualizations for representing time-evolving data in a global sports context?**
Key limitations include visual clutter (especially from overlapping arcs), a steeper learning curve for inexperienced users, and interaction difficulties on smaller devices. These factors were reflected in both task accuracy and user feedback. Careful onboarding, consistent visual encoding, and responsive interaction design are essential to improve usability in such contexts.

Altogether, this dissertation validates the hypothesis that 3D visualizations can effectively represent geographically distributed, time-evolving data with dynamic relationships while preserving the narrative within the visualization. While the prototype already supports exploratory analysis and user engagement, future iterations could further improve clarity, interactivity, and accessibility by refining its interaction logic and visual guidance, helping ZOS present complex data to a global audience more straightforwardly.

## 7.2 Future Work

While the developed 3D globe-based visualization prototype effectively presents cross-border football player transfers and demonstrates the potential of this type of spatio-temporal visual storytelling, several directions remain open for future exploration and improvement.

One of the main limitations is the prototype's reliance on a limited dataset, consisting only of transfers from the top divisions of 15 countries. Expanding the scope to include more leagues and nations would offer a more comprehensive view of global player transfer patterns. However, such expansion would also amplify existing issues like visual clutter and arc overlapping, compromising clarity. To mitigate this, future iterations should enhance filtering mechanisms, potentially requiring users to focus on a single continent, country, or league at a time, to reduce visual overload and improve interpretability.

Data interpretation can also be improved by combining the 3D globe with coordinated 2D views, such as tables, plots, or bar charts. These complementary perspectives offer alternative ways to explore and analyze the data, especially when the spatial encoding alone is insufficient. Connected interactions between 2D and 3D components (e.g., clicking on a country and viewing its transfer breakdown in a table) would strengthen both exploratory and analytical workflows.

Another important step is to add onboarding elements or interactive tutorials to support first-time users. As revealed in the experiment, participants had varying levels of experience with 3D visualizations, which occasionally influenced their ability to fully engage with the system. A more structured introduction, featuring contextual tooltips, onboarding panels, or progressive interaction hints, could significantly improve usability across different user groups.

Further technical improvements could focus on optimizing arc design to reduce clutter, using techniques such as dynamic arc bundling, conditional opacity, or simplified transitions. Ensuring responsive layouts and touch-friendly interactions would also enhance accessibility, especially on smaller screens and mobile devices.

# Bibliography


[1] Veronika S. Novikova. "On the Issue of New Generation's Media Consumption Specifics: Key Motives, Trends and Problematics". In: *2024 Communication Strategies in Digital Society Seminar (ComSDS)*. 2024, pp. 101–104. DOI: 10.1109/ComSDS61892.2024.10502075.

[2] ZOS. http://www.zos.pt.

[3] Siming Chen et al. "Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling". In: *IEEE Transactions on Visualization and Computer Graphics* 26.7 (2020), pp. 2499–2516. DOI: 10.1109/TVCG.2018.2889054.

[4] Christina Stoiber, Margit Pohl, and Wolfgang Aigner. "Design Actions for the Design of Visualization Onboarding Methods". In: *2023 IEEE VIS Workshop on Visualization Education, Literacy, and Activities (EduVis)*. 2023, pp. 1–10. DOI: 10.1109/EduVis60792.2023.00007.

[5] Mohammed Ali et al. "TimeCluster: dimension reduction applied to temporal data for visual analytics". In: *Vis. Comput.* 35.6–8 (June 2019), pp. 1013–1026. ISSN: 0178-2789. DOI: 10.1007/s00371-019-01673-y. URL: https://doi.org/10.1007/s00371-019-01673-y.

[6] PRISMA. *PRISMA - transparent reporting of systematic reviews and meta-analyses*. http://www.prisma-statement.org/.

[7] Murni Fatehah, Vitaliy Mezhuyev, and Mostafa Al-Emran. "A Systematic Review of Metamodelling in Software Engineering". In: June 2020, pp. 3–27. ISBN: 978-3-030-47410-2. DOI: 10.1007/978-3-030-47411-9_1.

[8] R. Bani and Y. Yamamoto. "Development of an R-Shiny-based Shooting Area Visualization Application for Use in Basketball". In: *2022 20th International Conference on ICT and Knowledge Engineering (ICTKE)*. 2022, pp. 1–6. ISBN: 2157-099X. DOI: 10.1109/ICTKE55848.2022.9983260. URL: https://ieeexplore.ieee.org/document/9983260/.

[9] V. Bhatt, U. Aggarwal, and C. N. S. V. Kumar. "Sports Data Visualization and Betting". In: *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*. 2022, pp. 1–6. DOI: 10.1109/SMARTGENCON56628.2022.10083831. URL: https://ieeexplore.ieee.org/document/10083831/.

[10] S. Tamboli and A. Anand. "FIFA Dataset Analysis and Match Prediction". In: *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*. Vol. 1. 2024, pp. 1–5. DOI: 10.1109/ICCPCT61902.2024.10672929. URL: https://ieeexplore.ieee.org/document/10672929/.

[11] Roger A. Leite et al. "Hermes: Guidance-enriched Visual Analytics for economic network exploration". In: *Visual Informatics* 4.4 (2020), pp. 11–22. ISSN: 2468-502X. DOI: https://doi.org/10.1016/j.visinf.2020.09.006. URL: https://www.sciencedirect.com/science/article/pii/S2468502X20300371%20https://www.sciencedirect.com/science/article/pii/S2468502X20300371?via%3Dihub.

[12] A. Lavanya et al. "A Real-Time Human Mobility Visualization of Covid-19 Spread from East Asian Countries". In: *2021 Eighth International Conference on Social Network Analysis, Management and Security (SNAMS)*. 2021, pp. 1–8. DOI: 10.1109/SNAMS53716.2021.9732103. URL: https://ieeexplore.ieee.org/document/9732103/.

[13] Enrique Tomás Martínez Beltrán et al. "COnVIDa: COVID-19 multidisciplinary data collection and dashboard". In: *Journal of Biomedical Informatics* 117 (2021), p. 103760. ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.jbi.2021.103760. URL: https://www.sciencedirect.com/science/article/pii/S1532046421000897%20https://pmc.ncbi.nlm.nih.gov/articles/PMC8007529/pdf/main.pdf.

[14] Gabriel H. A. Medeiros et al. "Tracing and analyzing COVID-19 dissemination using knowledge graphs". In: *Procedia Computer Science* 207 (2022), pp. 2172–2181. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2022.09.277. URL: https://www.sciencedirect.com/science/article/pii/S1877050922011632%20https://pmc.ncbi.nlm.nih.gov/articles/PMC9578937/pdf/main.pdf.

[15] R. K. Arya et al. "Open-Source Visualization Tools for Antenna Engineers". In: *2023 IEEE Microwaves, Antennas, and Propagation Conference (MAPCON)*. 2023, pp. 1–6. DOI: 10.1109/MAPCON58678.2023.10464133. URL: https://ieeexplore.ieee.org/document/10464133/.

[16] Y. Cheng et al. "Design and Implementation of Marine Multiple Geoscience Sharing WebGIS System Based on OpenLayers". In: *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT*. 2020, pp. 164–167. DOI: 10.1109/ICCASIT50869.2020.9368799. URL: https://ieeexplore.ieee.org/document/9368799/.

[17] F. Li et al. "Usability Evaluation of Hybrid 2D-3D Visualization Tools in Basic Air Traffic Control Operations". In: *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2021, pp. 3367–3372. ISBN: 2577-1655. DOI: 10.1109/SMC52423.2021.9658807. URL: https://ieeexplore.ieee.org/document/9658807/.

[18] M. Li, C. Li, and M. Shi. "Movie Data Visualization Based on WebGL". In: *2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter)*. 2021, pp. 74–78. DOI: 10.1109/SNPDWinter52325.2021.00023. URL: https://ieeexplore.ieee.org/document/9403450/.

[19] X. Li et al. "Design of Product Characteristics 3D Model Display System Based on Digital Earth". In: *2023 8th International Conference on Image, Vision and Computing (ICIVC)*. 2023, pp. 660–664. DOI: 10.1109/ICIVC58118.2023.10270452. URL: https://ieeexplore.ieee.org/document/10270452/.

[20] Yao Lu, Jukka Corander, and Zhirong Yang. "Doubly Stochastic Neighbor Embedding on Spheres". In: *Pattern Recognition Letters* 128 (2019), pp. 100–106. ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2019.08.026. URL: https://www.sciencedirect.com/science/article/pii/S0167865518305099%20https://www.sciencedirect.com/science/article/pii/S0167865518305099?via%3Dihub.

[21] A. Trajkovska, A. Bocevska, and B. Ristevski. "3D Scientific Visualization Using Mayavi". In: *2024 59th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*. 2024, pp. 1–4. ISBN: 2603-3267. DOI: 10.1109/ICEST62335.2024.10639745. URL: https://ieeexplore.ieee.org/document/10639745/.

[22] Ilaria Bartolini and Andrea Di Luzio. *ALDO: An Innovative Digital Framework for Active E-Learning*. Conference Paper. 2020. DOI: 10.1145/3369255.3369287. URL: https://doi.org/10.1145/3369255.3369287%20https://dl.acm.org/doi/pdf/10.1145/3369255.3369287.

[23] M. K. Farhat et al. "EventsVista: Enhancing Event Visualization and Interpretation". In: *2023 10th International Conference on Behavioural and Social Computing (BESC)*. 2023, pp. 1–7. DOI: 10.1109/BESC59560.2023.10386648. URL: https://ieeexplore.ieee.org/document/10386648/.

[24] Velitchko Filipov et al. "Gone full circle: A radial approach to visualize event-based networks in digital humanities". In: *Visual Informatics* 5.1 (2021), pp. 45–60. ISSN: 2468-502X. DOI: https://doi.org/10.1016/j.visinf.2021.01.001. URL: https://www.sciencedirect.com/science/article/pii/S2468502X21000012%20https://www.sciencedirect.com/science/article/pii/S2468502X21000012?via%3Dihub.

[25] J. He et al. "Variable-Based Spatiotemporal Trajectory Data Visualization Illustrated". In: *IEEE Access* 7 (2019), pp. 143646–143672. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2942844. URL: https://ieeexplore.ieee.org/ielx7/6287639/8600701/08846208.pdf?tp=&arnumber=8846208&isnumber=8600701&ref=.

[26] D. Kouřil et al. "Molecumentary: Adaptable Narrated Documentaries Using Molecular Visualization". In: *IEEE Transactions on Visualization and Computer Graphics* 29.3 (2023), pp. 1733–1747. ISSN: 1941-0506. DOI: 10.1109/TVCG.2021.3130670. URL: https://ieeexplore.ieee.org/ielx7/2945/10032251/09627526.pdf?tp=&arnumber=9627526&isnumber=10032251&ref=.

[27] Bárbara Ramalho, Joaquim Jorge, and Sandra Gama. "Representing uncertainty through sentiment and stance visualizations: A survey". In: *Graphical Models* 129 (2023), p. 101191. ISSN: 1524-0703. DOI: https://doi.org/10.1016/j.gmod.2023.101191. URL: https://www.sciencedirect.com/science/article/pii/S1524070323000218%20https://www.sciencedirect.com/science/article/pii/S1524070323000218?via%3Dihub.

[28] J. Schmidt et al. "Popup-Plots: Warping Temporal Data Visualization". In: *IEEE Transactions on Visualization and Computer Graphics* 25.7 (2019), pp. 2443–2457. ISSN: 1941-0506. DOI: 10.1109/TVCG.2018.2841385. URL: https://ieeexplore.ieee.org/document/8368315/.

[29] J. Tao et al. "Exploring Time-Varying Multivariate Volume Data Using Matrix of Isosurface Similarity Maps". In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 1236–1245. ISSN: 1941-0506. DOI: 10.1109/TVCG.2018.2864808. URL: https://ieeexplore.ieee.org/document/8440120/.

[30] V. M. Tsontos, F. Platt, and J. T. Roberts. "COVERAGE Data Viewer: A Generalized Web-Based Tool for Integrated Visualization of Oceanographic Satellite and In-Situ Data". In: *OCEANS 2023 - MTS/IEEE U.S. Gulf Coast*. 2023, pp. 1–8. ISBN: 0197-7385. DOI: 10.23919/OCEANS52994.2023.10337364. URL: https://ieeexplore.ieee.org/document/10337364/.

[31] Manuela Waldner, Daniel Steinböck, and Eduard Gröller. "Interactive exploration of large time-dependent bipartite graphs". In: *Journal of Computer Languages* 57 (2020), p. 100959. ISSN: 2590-1184. DOI: https://doi.org/10.1016/j.cola.2020.100959. URL: https://www.sciencedirect.com/science/article/pii/S2590118420300198%20https://www.sciencedirect.com/science/article/pii/S2590118420300198?via%3Dihub.

[32] W. Cui. "Visual Analytics: A Comprehensive Overview". In: *IEEE Access* 7 (2019), pp. 81555–81573. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2923736. URL: https://ieeexplore.ieee.org/ielx7/6287639/8600701/08740868.pdf?tp=&arnumber=8740868&isnumber=8600701&ref=.

[33] S. Maji and J. Dingliana. "Perceptually Optimized Color Maps for Visualizing Large Numbers of Features". In: *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*. 2019, pp. 1–6. DOI: 10.1109/ICACCP.2019.8882980. URL: https://ieeexplore.ieee.org/document/8882980/.

[34] P. Suwanworaboon, S. Lynden, and S. Tuarob. “Enhancing Visualization Applications Using Open Data Sources”. In: *2020 17th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. 2020, pp. 30–35. ISBN: 2642-6579. DOI: 10.1109/JCSSE49651.2020.9268315. URL: https://ieeexplore.ieee.org/document/9268315/.

[35] L. Yang and Z. Zhang. “Touch the Collections: A 3D UI Prototype in DSpace-7”. In: *2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. 2023, pp. 267–268. ISBN: 2575-8152. DOI: 10.1109/JCDL57899.2023.00054. URL: https://ieeexplore.ieee.org/document/10266161/.

[36] Ben Shneiderman. “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations”. In: *The Craft of Information Visualization*. Ed. by BENJAMIN B. BEDERSON and BEN SHNEIDERMAN. Interactive Technologies. San Francisco: Morgan Kaufmann, 2003, pp. 364–371. ISBN: 978-1-55860-915-0. DOI: https://doi.org/10.1016/B978-155860915-0/50046-9. URL: https://www.sciencedirect.com/science/article/pii/B9781558609150500469.

[37] Vasco Asturiano. http://www.globe.gl.

[38] Three.js - JavaScript 3D Library. https://threejs.org/.

[39] ThreeJS WebGL class to represent data visualization layers on a globe. https://www.npmjs.com/package/three-globe.

[40] WebGL: 2D and 3D graphics for the web. https://get.webgl.org/.

[41] Salesforce. *Slack*. https://slack.com/.

[42] Google. *Google Forms*. https://workspace.google.com/products/forms/.

[43] UIUX Trend. *Measuring and Interpreting System Usability Scale (SUS)*. https://uiuxtrend.com/measuring-system-usability-scale-sus/.

[44] P. McCullagh. *Regression models for ordinal data*. 1980.

# Appendix A

# Prototype Code and Resources

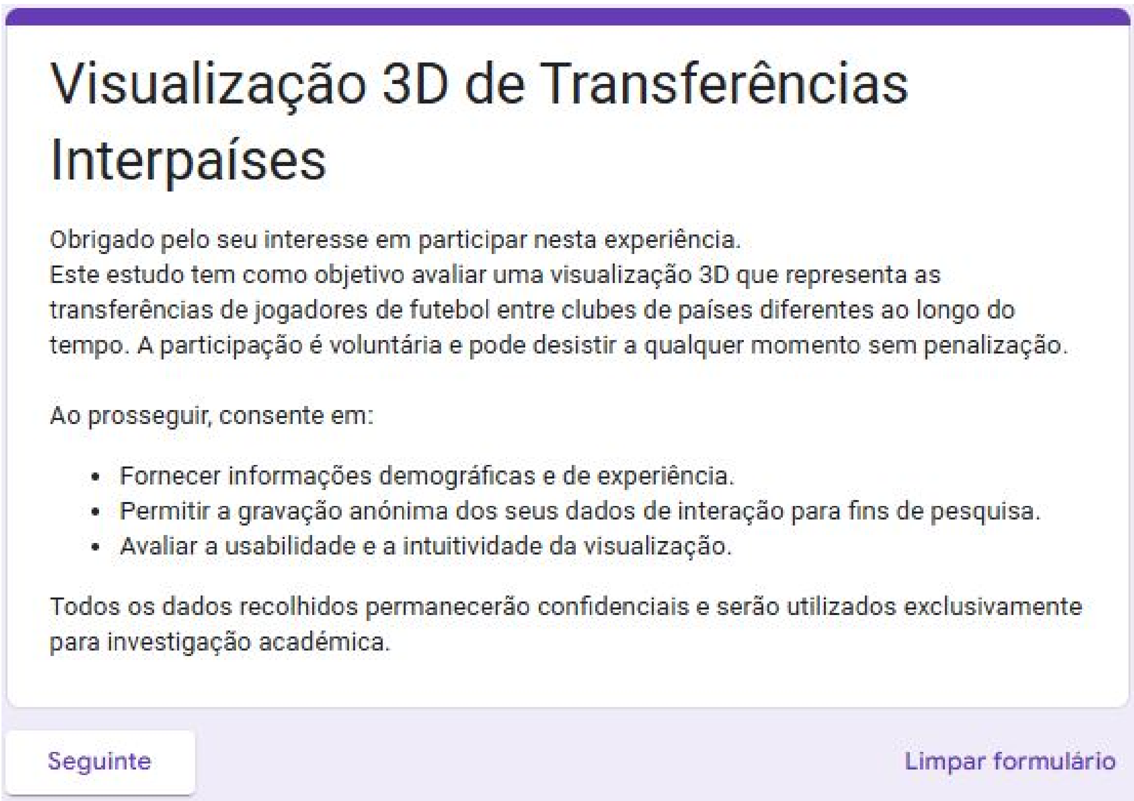
This appendix provides access to the prototype's source code and to the online version that was available to the users.

- **Prototype's Source Code:** https://github.com/Me-Salva/Tese/tree/main/src
- **Online Version:** https://transfers-globe.zerozero.pt/

# Appendix B

# Formative Questionnaire

This appendix shows the figures of the formative questionnaire as it was presented to participants (in Portuguese) during the experimental procedure.



Visualização 3D de Transferências Interpaíses

Obrigado pelo seu interesse em participar nesta experiência.
Este estudo tem como objetivo avaliar uma visualização 3D que representa as transferências de jogadores de futebol entre clubes de países diferentes ao longo do tempo. A participação é voluntária e pode desistir a qualquer momento sem penalização.

Ao prosseguir, consente em:

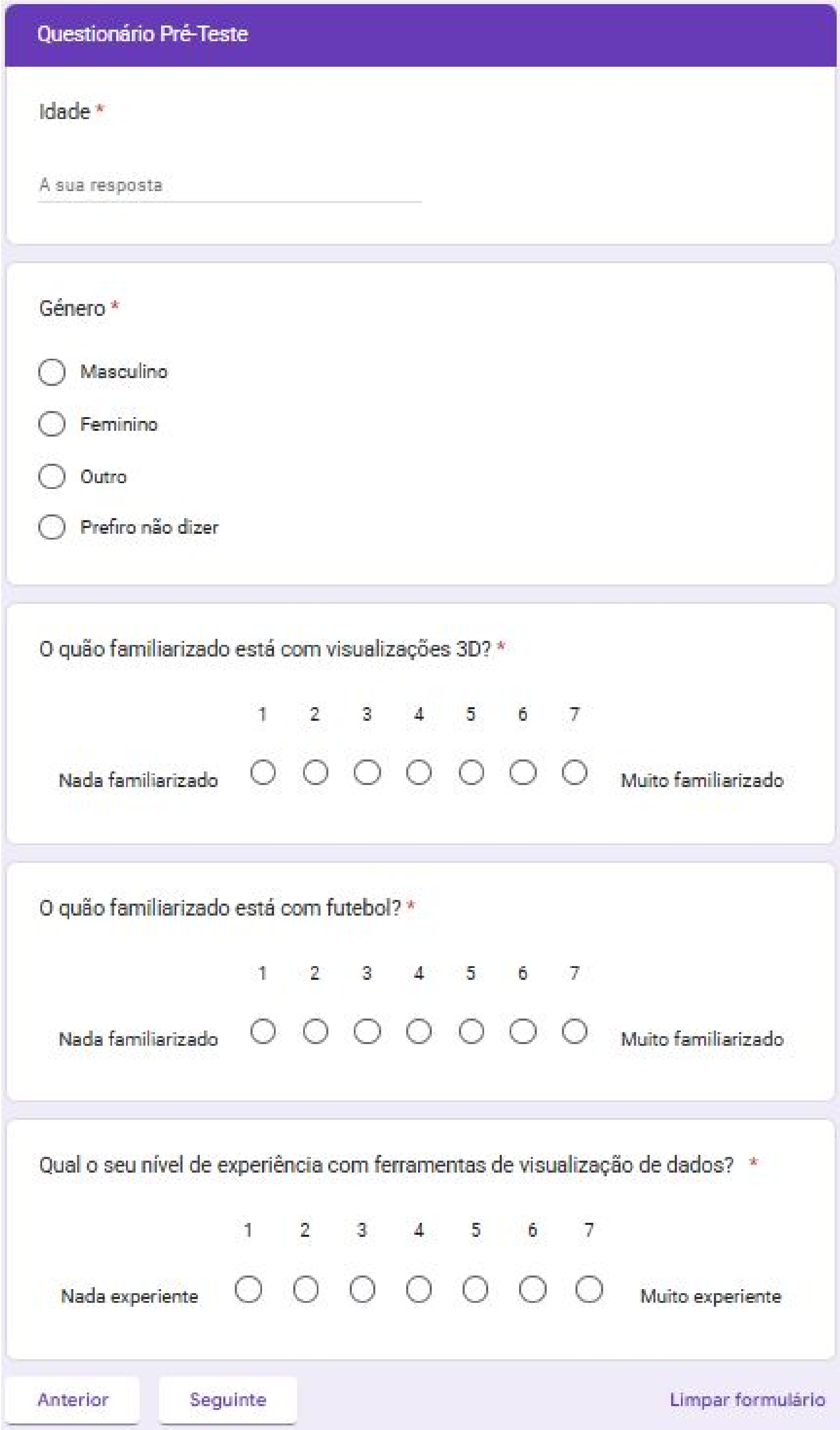
- Fornecer informações demográficas e de experiência.
- Permitir a gravação anónima dos seus dados de interação para fins de pesquisa.
- Avaliar a usabilidade e a intuitividade da visualização.

Todos os dados recolhidos permanecerão confidenciais e serão utilizados exclusivamente para investigação académica.

Seguinte

Limpar formulário

Figure B.1: Evaluation Survey - Informed Consent



Questionário Pré-Teste

Idade *

A sua resposta

Género *

- Masculino
- Feminino
- Outro
- Prefiro não dizer

O quão familiarizado está com visualizações 3D? *

Nada familiarizado 1 2 3 4 5 6 7 Muito familiarizado

O quão familiarizado está com futebol? *

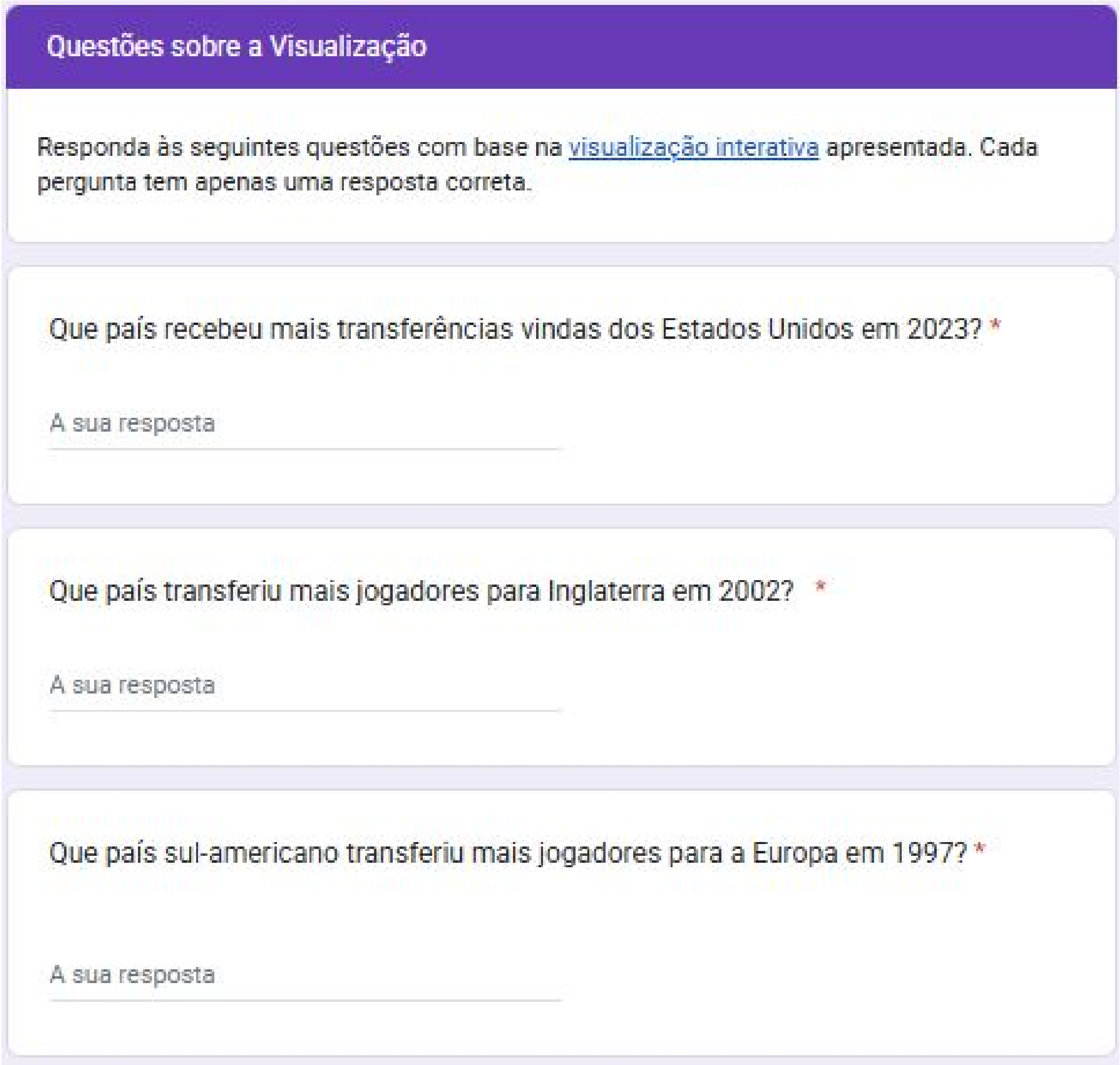
Nada familiarizado 1 2 3 4 5 6 7 Muito familiarizado

Qual o seu nível de experiência com ferramentas de visualização de dados? *

Nada experiente 1 2 3 4 5 6 7 Muito experiente

Anterior | Seguinte | Limpar formulário

Figure B.2: Evaluation Survey - Pre-test Questionnaire



**Questões sobre a Visualização**

Responda às seguintes questões com base na visualização interativa apresentada. Cada pergunta tem apenas uma resposta correta.

Que país recebeu mais transferências vindas dos Estados Unidos em 2023? *

A sua resposta

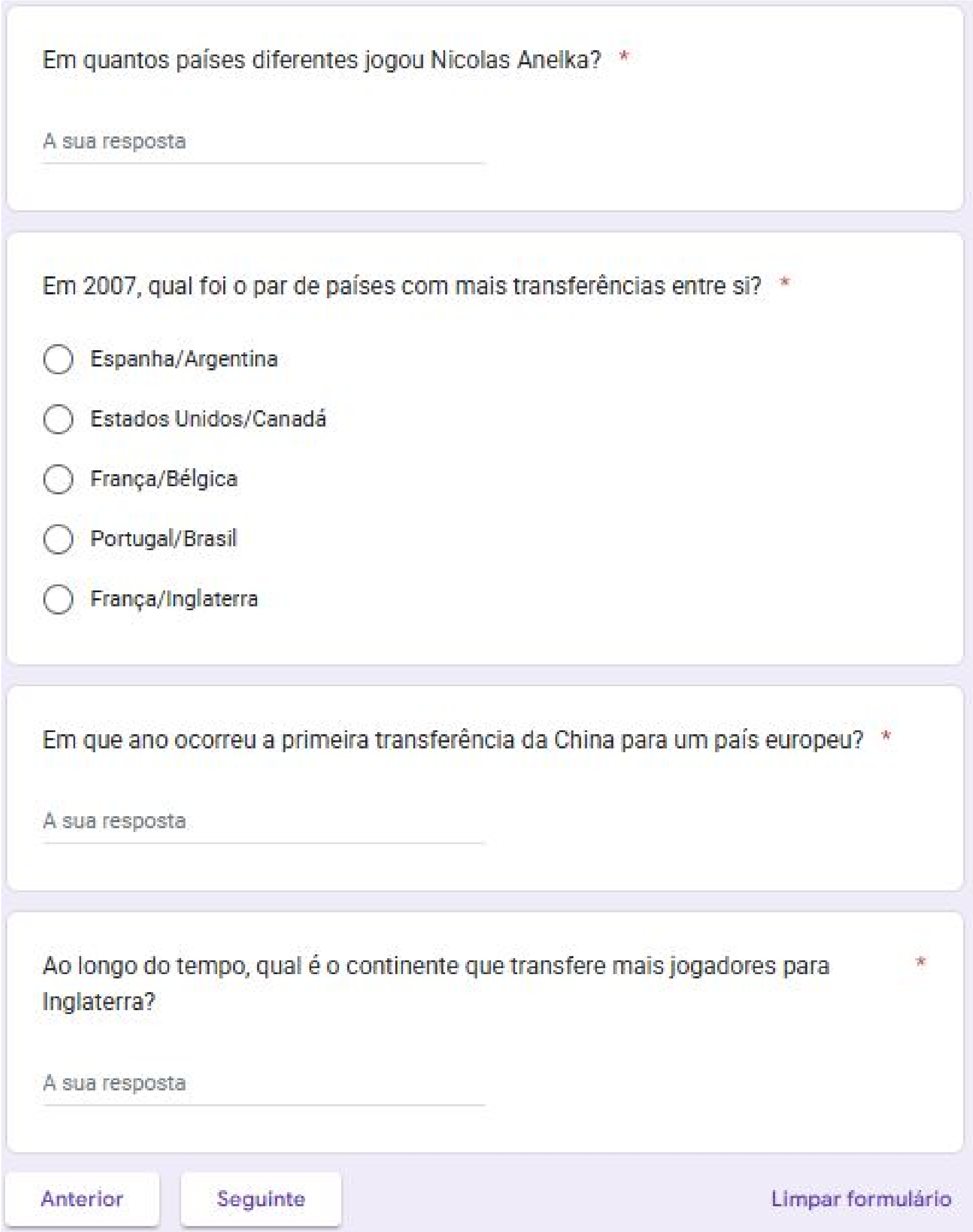
Que país transferiu mais jogadores para Inglaterra em 2002? *

A sua resposta

Que país sul-americano transferiu mais jogadores para a Europa em 1997? *

A sua resposta

Figure B.3: Evaluation Survey - Task-specific Questionnaire (Part 1)



Em quantos países diferentes jogou Nicolas Anelka? *

A sua resposta

Em 2007, qual foi o par de países com mais transferências entre si? *

- Espanha/Argentina
- Estados Unidos/Canadá
- França/Bélgica
- Portugal/Brasil
- França/Inglaterra

Em que ano ocorreu a primeira transferência da China para um país europeu? *

A sua resposta

Ao longo do tempo, qual é o continente que transfere mais jogadores para Inglaterra? *

A sua resposta

Anterior Seguinte Limpar formulário

Figure B.4: Evaluation Survey - Task-specific Questionnaire (Part 2)

**Questionário Pós-Teste**

Gostaria de utilizar este sistema com frequência. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Achei o sistema desnecessariamente complexo. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Achei o sistema fácil de utilizar. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Acredito que necessitaria de apoio técnico para utilizar este sistema. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

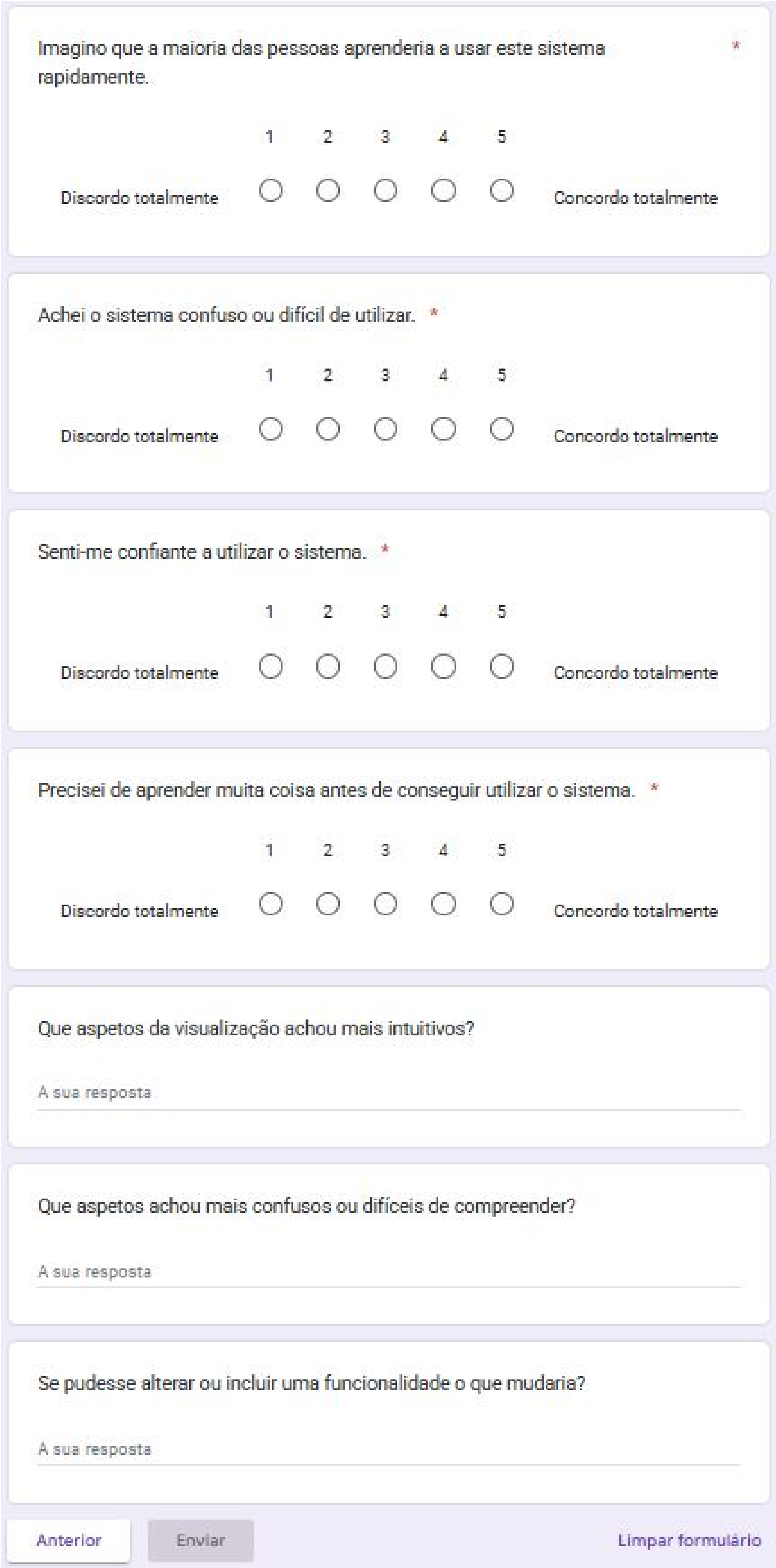
As funcionalidades do sistema parecem bem integradas. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Notei muita inconsistência no sistema. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Figure B.5: Evaluation Survey - Post-test Questionnaire (Part 1)



Imagino que a maioria das pessoas aprenderia a usar este sistema rapidamente. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Achei o sistema confuso ou difícil de utilizar. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Senti-me confiante a utilizar o sistema. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Precisei de aprender muita coisa antes de conseguir utilizar o sistema. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Discordo totalmente | ○ | ○ | ○ | ○ | ○ | Concordo totalmente |

Que aspetos da visualização achou mais intuitivos?

A sua resposta

Que aspetos achou mais confusos ou difíceis de compreender?

A sua resposta

Se pudesse alterar ou incluir uma funcionalidade o que mudaria?

A sua resposta

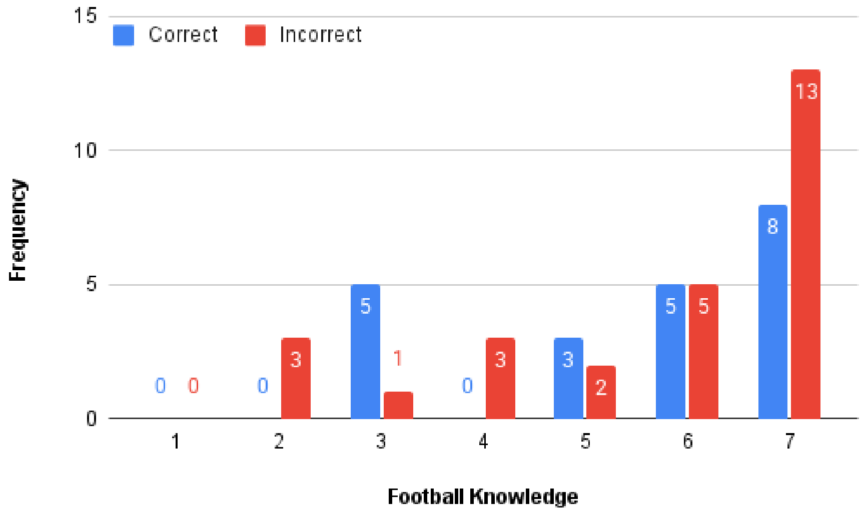
Anterior Enviar Limpar formulário

Figure B.6: Evaluation Survey - Post-test Questionnaire (Part 2)

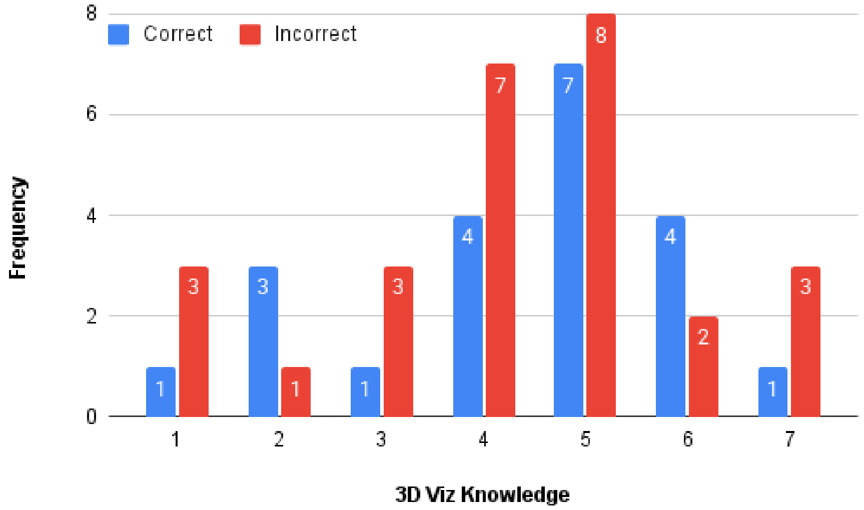
# Appendix C

# Supplementary Results

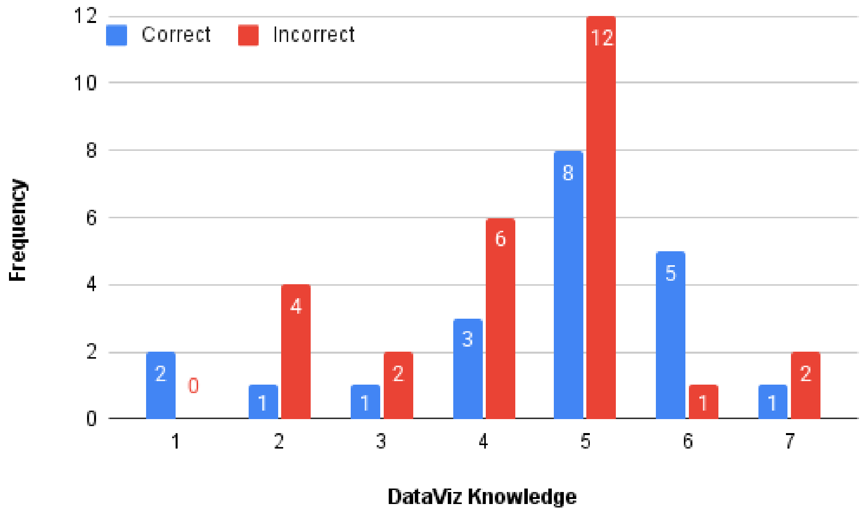
This appendix includes supplementary charts that complement the findings presented in the results analysis section.




(a) Correct and incorrect answers per football knowledge.

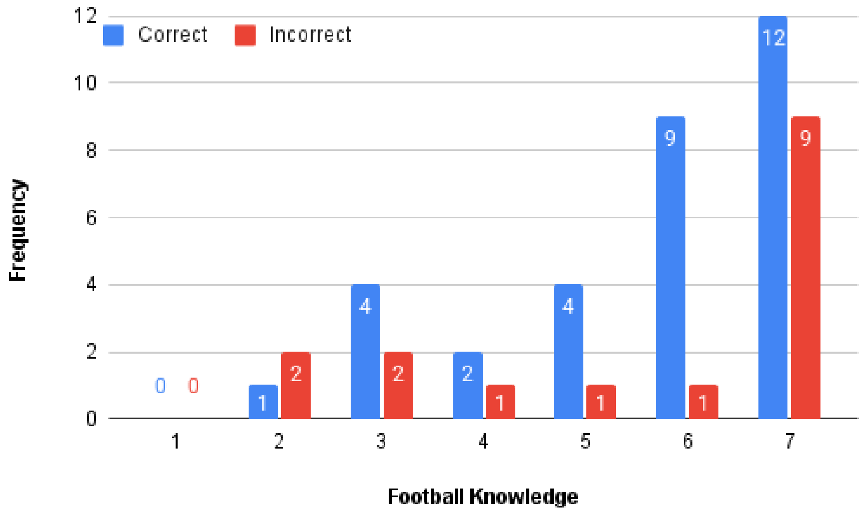



(b) Correct and incorrect answers per 3D visualization knowledge.

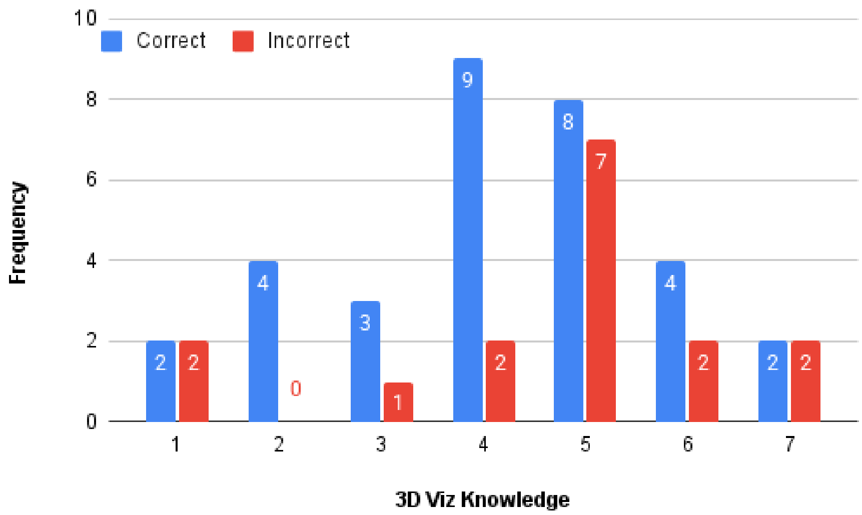



(c) Correct and incorrect answers per data visualization tools knowledge.

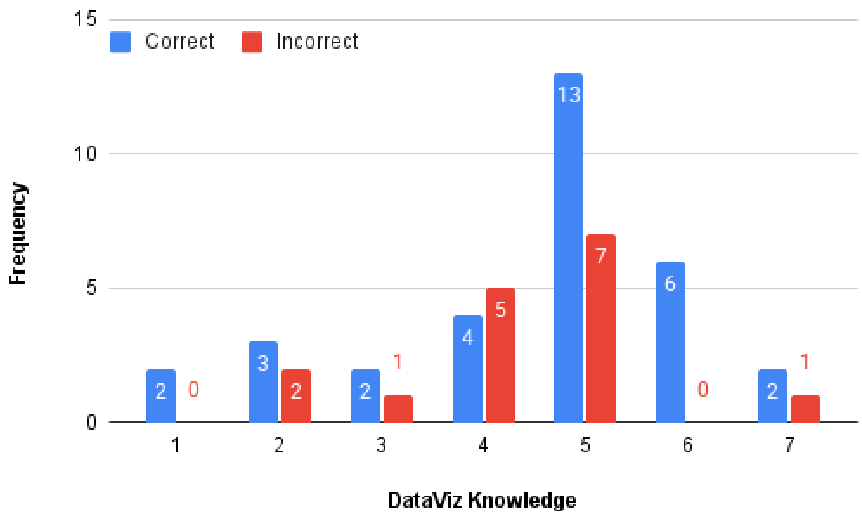
Figure C.1: Distribution of correct and incorrect answers per domain knowledge to question T1.




(a) Correct and incorrect answers per football knowledge.

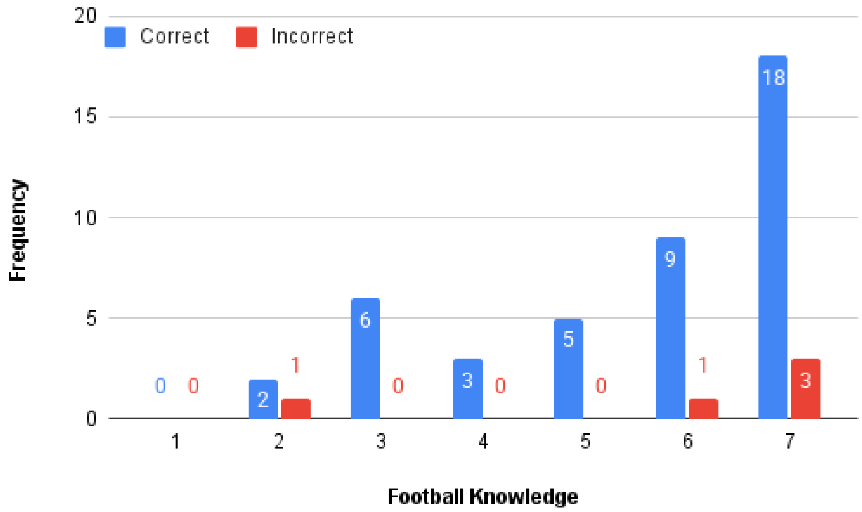



(b) Correct and incorrect answers per 3D visualization knowledge.

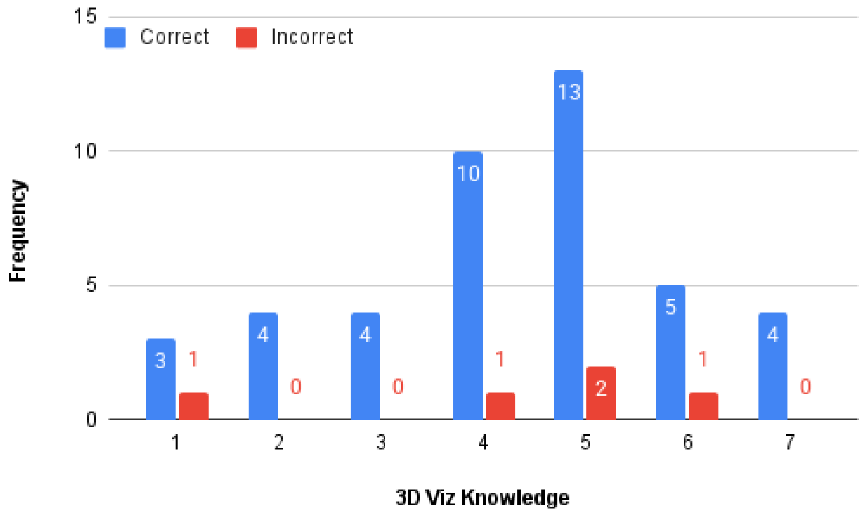



(c) Correct and incorrect answers per data visualization tools knowledge.

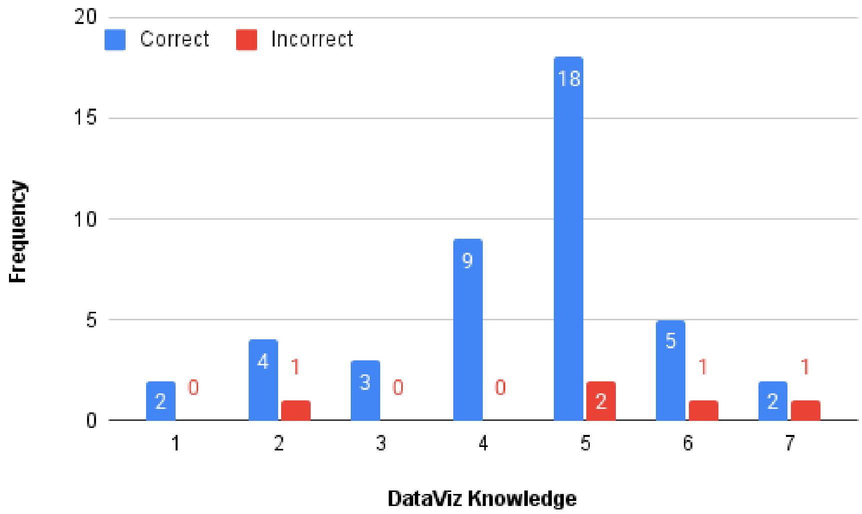
Figure C.2: Distribution of correct and incorrect answers per domain knowledge to question T2.




(a) Correct and incorrect answers per football knowledge.

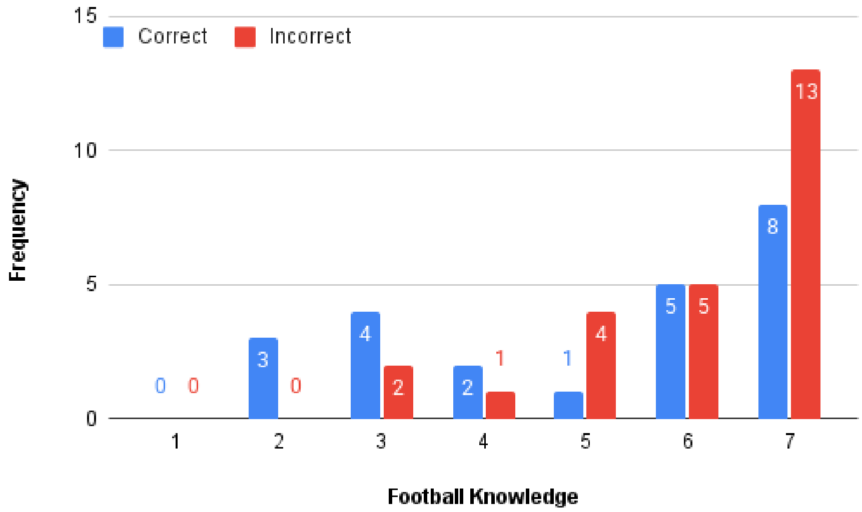



(b) Correct and incorrect answers per 3D visualization knowledge.

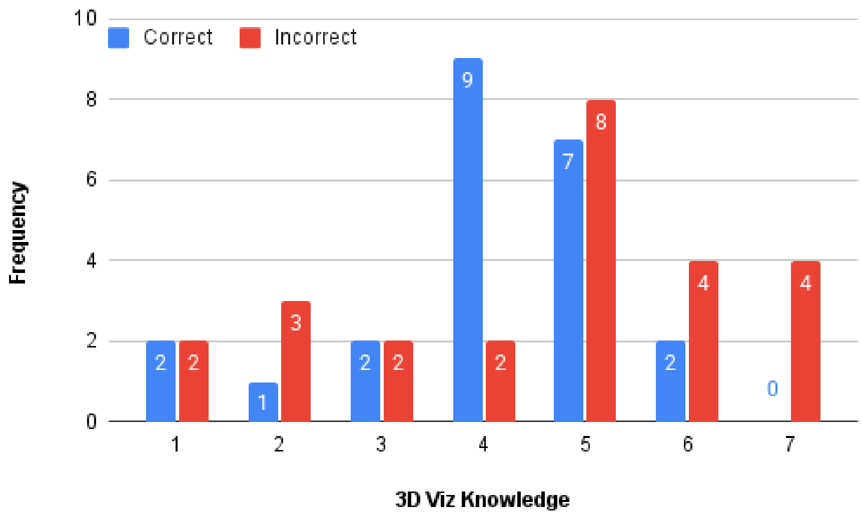



(c) Correct and incorrect answers per data visualization tools knowledge.

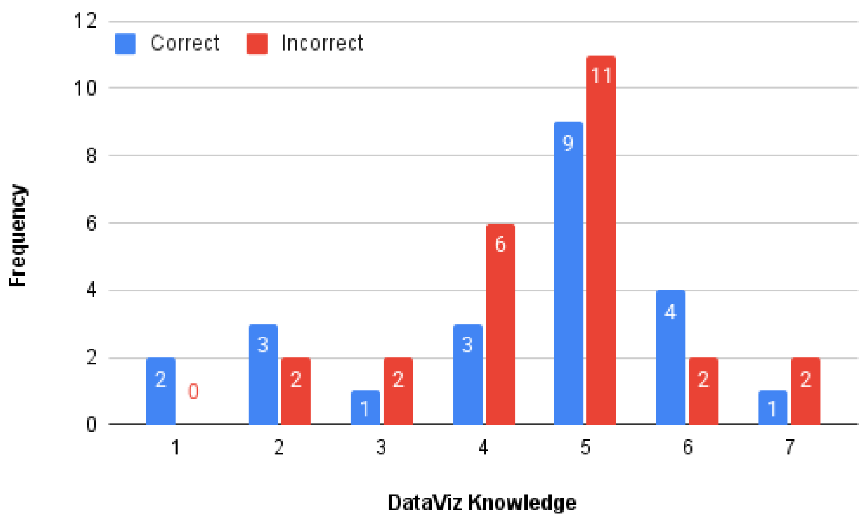
Figure C.3: Distribution of correct and incorrect answers per domain knowledge to question T3.




(a) Correct and incorrect answers per football knowledge.

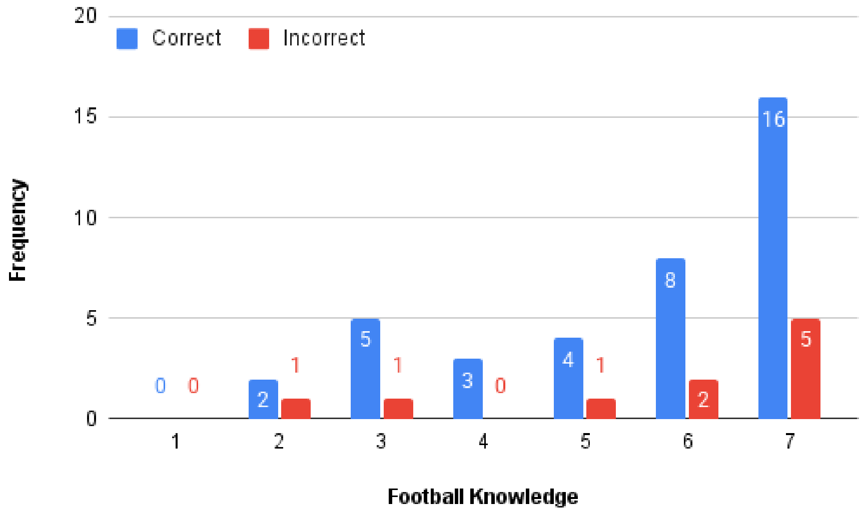



(b) Correct and incorrect answers per 3D visualization knowledge.

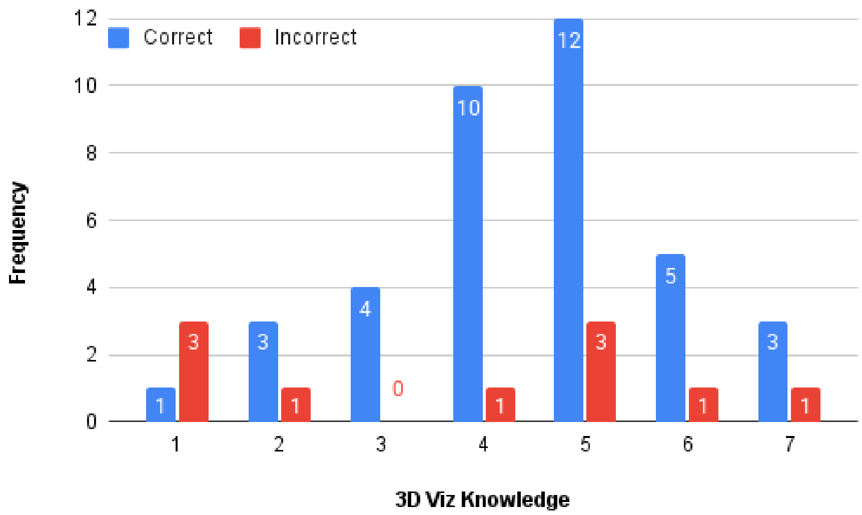



(c) Correct and incorrect answers per data visualization tools knowledge.

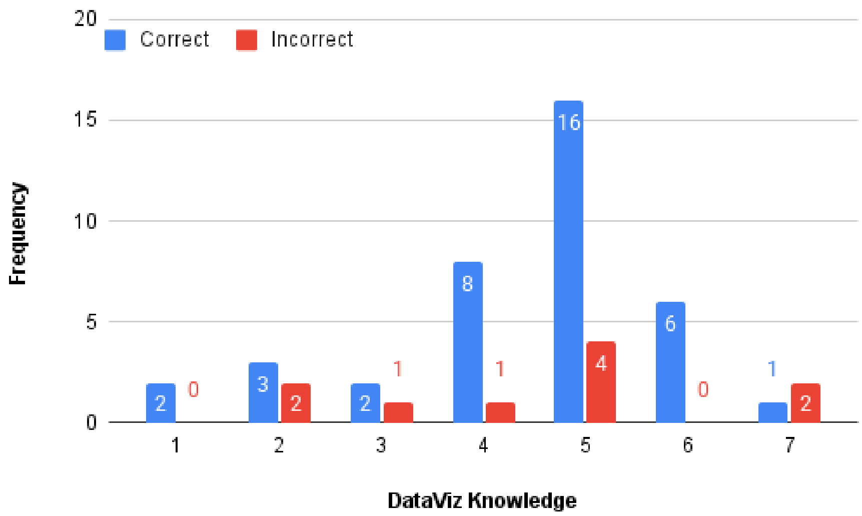
Figure C.4: Distribution of correct and incorrect answers per domain knowledge to question T4.




(a) Correct and incorrect answers per football knowledge.

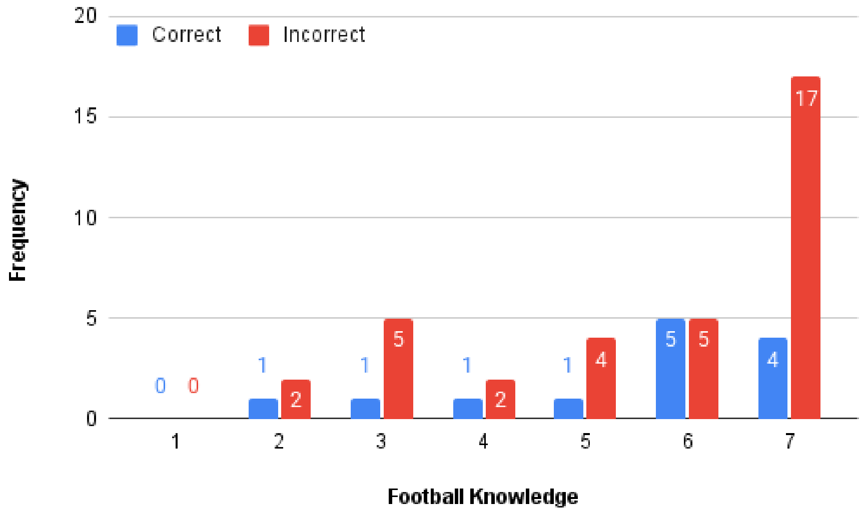



(b) Correct and incorrect answers per 3D visualization knowledge.

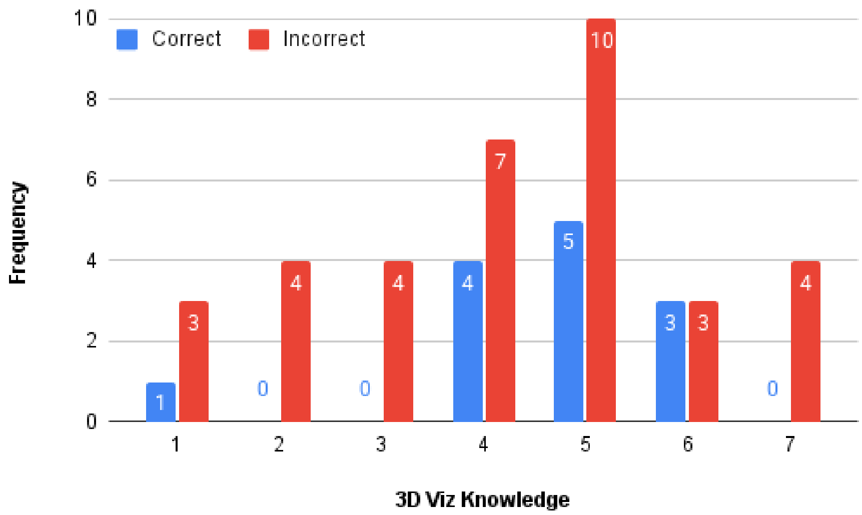



(c) Correct and incorrect answers per data visualization tools knowledge.

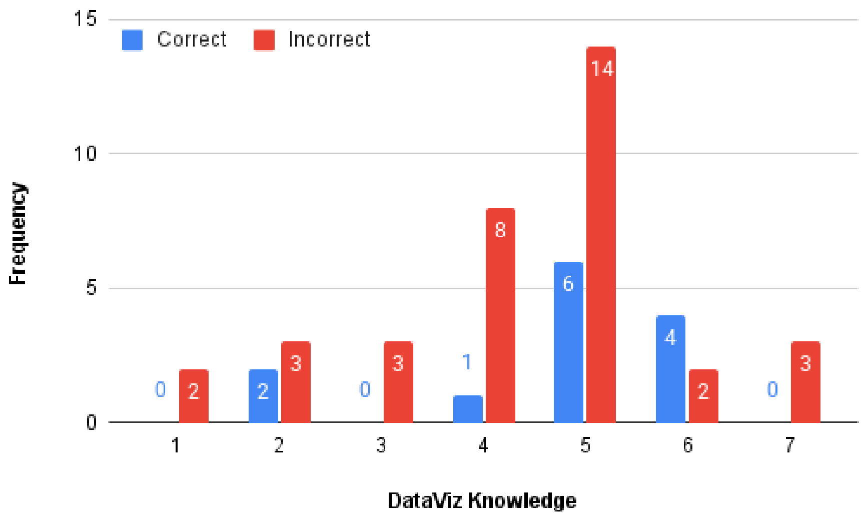
Figure C.5: Distribution of correct and incorrect answers per domain knowledge to question T5.




(a) Correct and incorrect answers per football knowledge.

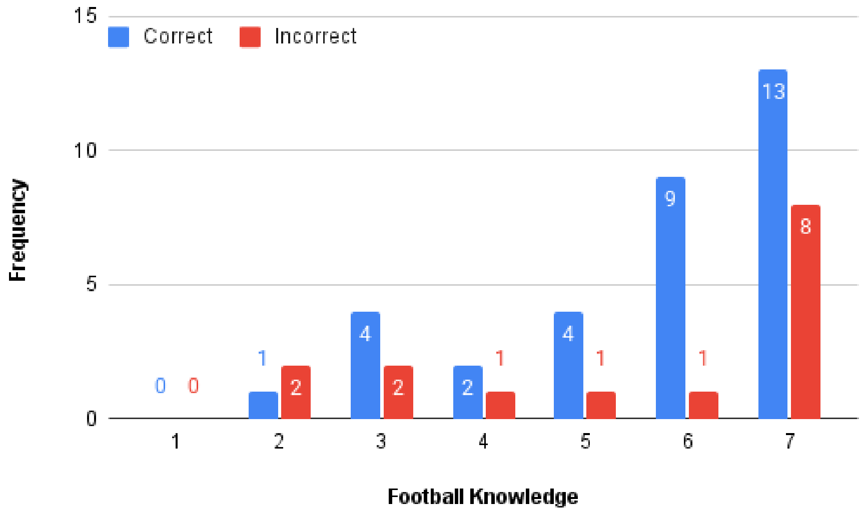



(b) Correct and incorrect answers per 3D visualization knowledge.

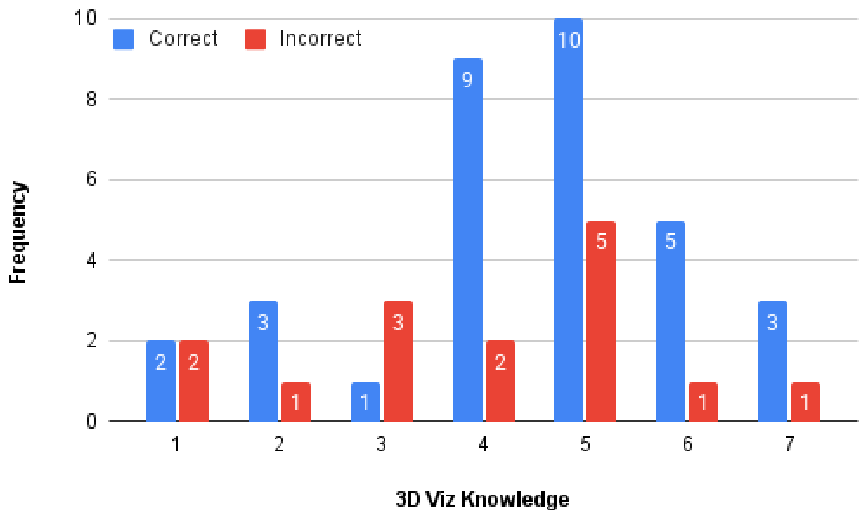



(c) Correct and incorrect answers per data visualization tools knowledge.

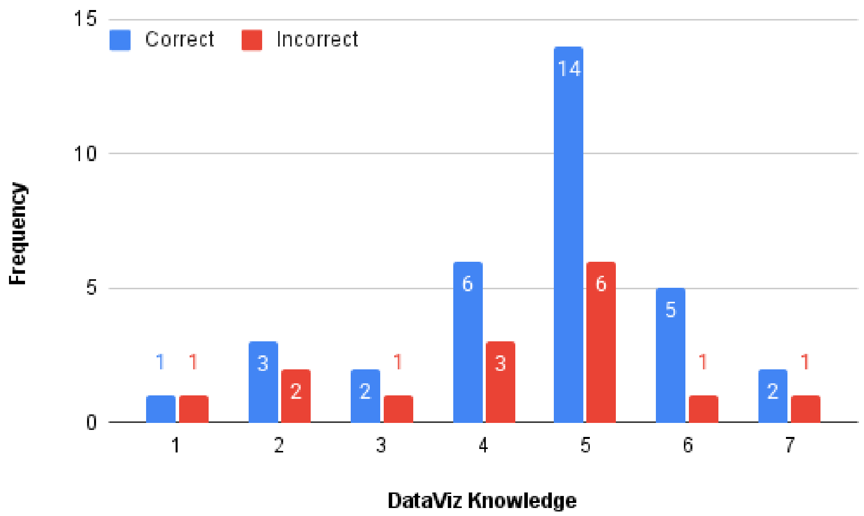
Figure C.6: Distribution of correct and incorrect answers per domain knowledge to question T6.




(a) Correct and incorrect answers per football knowledge.




(b) Correct and incorrect answers per 3D visualization knowledge.




(c) Correct and incorrect answers per data visualization tools knowledge.

Figure C.7: Distribution of correct and incorrect answers per domain knowledge to question T7.