

# Projektbericht zum Modul Information Retrieval und Visualisierung Sommersemester 2021

## Visualisierung von Daten des Videospiels Fifa 19

Johannes Lange

13. August 2021

### 1 Einleitung

#### Tamara Munzner angucken

Die Visualisierung von Daten nimmt mit Hinblick auf Big Data und damit immer unübersichtlicheren Grunddaten an Bedeutung zu. Um aus großen Mengen von Daten neue Informationen zu gewinnen reicht es nicht aus die Daten direkt zu analysieren. Durch die Anwendung der richtigen Visualisierungstechniken können neue Informationen gewonnen werden. In diesem Bericht geht es um die Visualisierung von Daten aus dem Computerspiel Fifa 19. Da sich diese Daten allerdings auf tatsächliche Fussballer beziehen besteht die Hoffnung mittels Visualisierung dieser Daten Rückschlüsse auf die tatsächlichen Sportler schließen zu können.

#### 1.1 Anwendungshintergrund

Die Frage, welche sich bei diesen Daten stellt ist, welchen Nutzen sie außerhalb eines Videospiels haben. Deswegen ist mein Ansatz zu vergleichen wie genau diese Daten die Wirklichkeit widerspiegeln. Ein Beispiel dafür könnte sein ob sich das Rating eines Spielers mit der körperlichen Entwicklung eines Spielers bewegt. Sollte dies zutreffen müssten Spieler im Alter ihres physischen Peaks das höchste Rating haben.

#### 1.2 Zielgruppen

Zielgruppen für Scatterplot:

Parallele Koordinaten:

Baumdiagramm: Hier lässt sich erkennen welche Sportliga Europas im Durchschnitt die besten Spieler hat. Deswegen ist dies für X interessant.

Mögliche Zielgruppen: Videospieler, Teilnehmer einer Fantasy Fußball Liga

### **1.3 Überblick und Beiträge**

Als erste Visualisierung habe ich mich für einen Scatterplot entschieden um mit diesem einen groben Überblick über mögliche Trends in den Daten zu ermöglichen.

Als zweite Technik habe mich für die Technik der parallelen Koordinaten entschieden, mit dieser lassen sich Zusammenhänge zwischen zwei Merkmalen gut erkennen.

Als dritte und letzte Technik habe ich die Baumdarstellung ausgewählt um sichtbar zu machen welche Fussballliga Europas im Durchschnitt die besten Spieler hat. Deswegen ist dies für X interessant.

## 2 Daten

Grundsätzlich eignen sich die Daten gut um die gewünschten Fragestellungen beantworten zu können, jedoch enthält der Grunddatensatz einige Felder, die für die Visualisierung nicht nötig sind, deswegen wurde der Datensatz in der Vorvorarbeitung noch verkleinert (Siehe 2.2 auf Seite 4). Außerdem ist der Datensatz sehr groß, was gerade bei den parallelen Koordinaten zu Problemen führen kann wenn der ganze Datensatz angezeigt wird, deswegen wurde sich bei den parallelen Koordinaten dazu entschieden nach zusätzlichen Dimensionen wie Nationalität zu Filtern um dies so übersichtlicher zu gestalten. Da Datenwerte wie Größe, Alter und Rating der Spieler diskret sind wurde sich dazu entschieden im Scatterplot die Anzahl an Spielern welche in diesem Punkt enthalten sind auszugeben. Weiterhin wurde die Opazität der Punkte verringert, da so zu sehen ist an welchen Stellen sich mehrere Spieler überlagern.

**2.1 Technische Bereitstellung der Daten**

**2.2 Datenvorverarbeitung**

### **3 Visualisierungen**

**3.1 Analyse der Anwendungsaufgaben**

**3.2 Anforderungen an die Visualisierungen**

**3.3 Präsentation der Visualisierungen**

**3.3.1 Visualisierung Eins**

**3.3.2 Visualisierung Zwei**

**3.3.3 Visualisierung Drei**

**3.4 Interaktion**

### **4 Implementierung**

### **5 Anwendungsfälle**

**5.1 Anwendung Visualisierung Eins**

**5.2 Anwendung Visualisierung Zwei**

**5.3 Anwendung Visualisierung Drei**

### **6 Verwandte Arbeiten**

### **7 Zusammenfassung und Ausblick**

**Anhang: Git-Historie**