Data Science Project - Darts

Emilie Terhaar, Sara Rolfs, Tyra Kausch

Christian-Albrechts-Universität zu Kiel

Introduction

This project dives into the game darts using data science techniques. It aims to transform raw data into structured datasets suitable for visualization and analysis. The goal is to answer predefined research questions by gathering data from multiple sources, processing it, and generating analytical datasets.

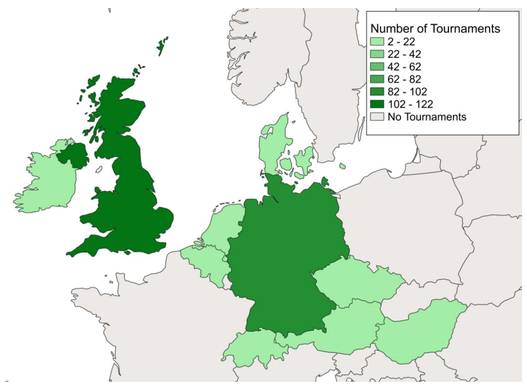
Method

Formulate Research Questions → Web Scraping → Data Cleaning and Storage → Data Visualization and Analysis

Results

How are the tournament's locations distributed?

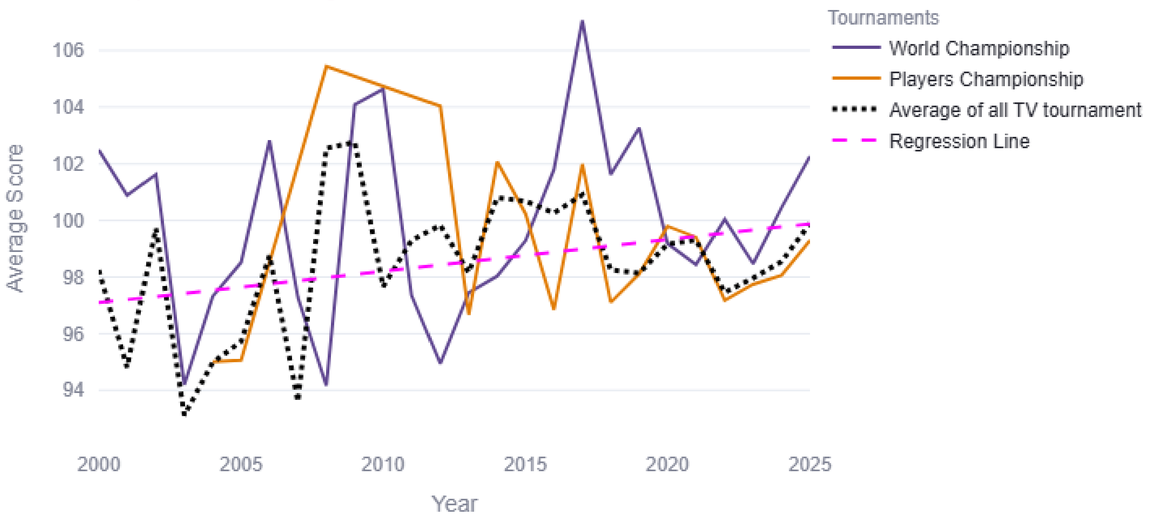
There is a big difference between the number of tournaments held in the UK and Germany and other countries. This shows that there is a much bigger influence on darts in Europe than in the rest of the world.



How do the averages of tournaments vary over time?

For this, it was assumed that the sport has reached a higher maturity level. Meaning the skill of players participating has become better with strategies and coaching. In this diagram, only the winners were looked at.
The World Championship shows greater volatility with peaks around 2010 and 2017/2018 where the average was above 106. All lines generally go up, which can be seen in the upward tilted regression line. The average scores seem to range between 94 and 105. This validates the assumption that the sport has matured.

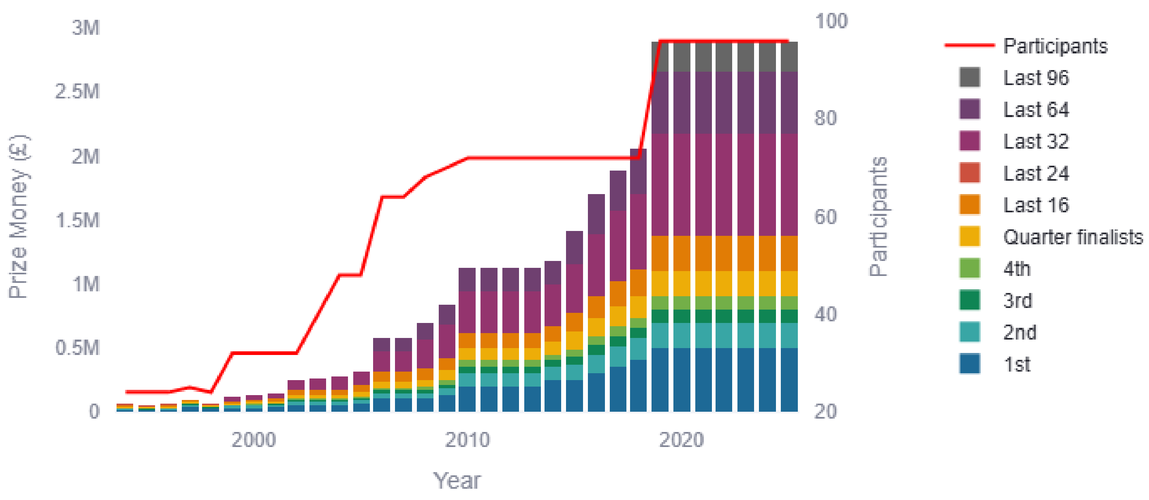
Development of Average Scores over time



How does the price money and number of participants vary over time?

In this graph, the prize money and number of participants are rising. There's a clear correlation: Major increases in participants typically coincide with significant prize money increases. The graphs show that while the total prize pool has increased, the distribution maintains a similar structure.

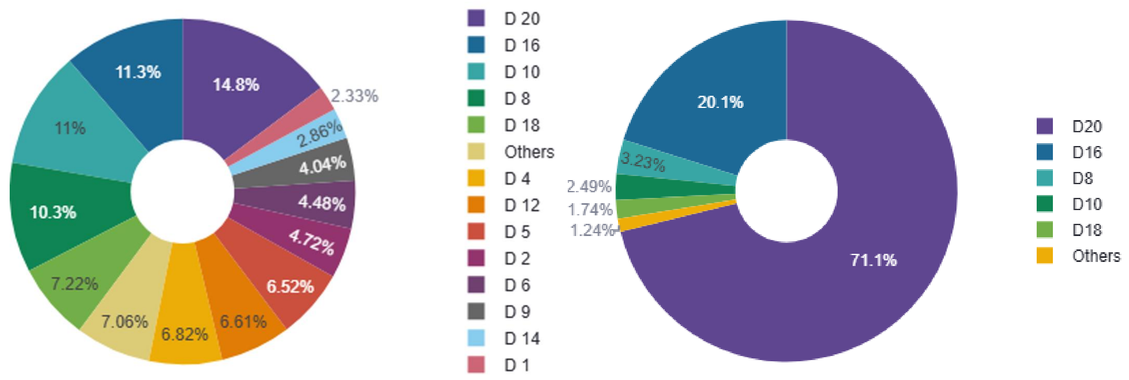
Development of Prize Money and Participants over time



What are most popular double fields and their corresponding check out quotes?

A leg ends by hitting a double field, called a check out. These create the statistic of check out percentages. Other than in the check outs, doubles are usually not targeted.
The left pie chart shows that D20, D16, D10, and D8 are the most popular double fields. This makes sense, since if a throw on D20 or D16 fails, the player can target the double field with half the value to finish the leg. This can also be seen in the right chart, where most player's most frequently thrown double is the D20 with 71.1% and the D16 with 20.1%.

Distribution of Throws and Hits on Double Fields



How does the player's nationality effect the rankings?

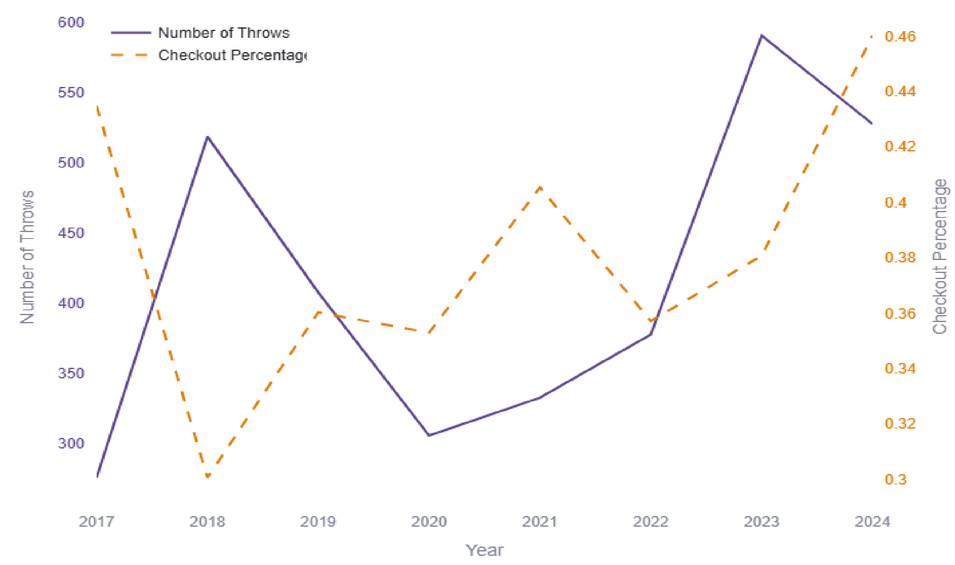
The bigger the bubble, the more players belong to that nationality. As seen, there are a lot of English players who are also very highly ranked, suggesting that the sport might be better developed there.
There is one player in the Netherlands (Michael van Gerwen) who dominated darts for a few years and is now on a high-ranking position. Other countries like Germany, Japan, and Australia have representation, but their players are less frequent and frequent lower ranks.

Nationalities and their Rankings

How does the performance of individual players change over time?

The graph below shows Rob Cross' number of throws and check out percentages for the D1 from 2017 to 2024. This displays the phenomenon, where a check out percentage drops and the number of throws on this double follows in a time-shifted manner. This implies that the player noticed that this double does not work that well for him, so he adjusts his throws.

Number of Throws and Check Out Percentages of D1 – Rob Cross



Bibliography

https://app.dartsorakel.com/
https://mastercaller.com/tournaments/pdc-world-championship/
https://www.dartn.de/Dart-Profis
https://de.wikipedia.org/wiki/Professional_Darts_Corporation
https://de.wikipedia.org/wiki/PDC_World_Darts_Championship
https://www.flashscore.de/