Create a Tableau Story: Players in FIFA 18

Serdar Ozsoy

The dataset contains all the statistics and playing attributes of all the players in the Full version of FIFA 18 is obtained from here:

https://www.kaggle.com/thec03u5/fifa-18-demo-player-dataset/data

After removing duplications and data preparation, final dataset is here.

https://github.com/serdarozsoy/datasets/blob/master/CompleteDataset5.csv

The first version of story:
https://public.tableau.com/profile/serdarozsoy#!/vizhome/first_version/Story2

Final version of story:
https://public.tableau.com/profile/serdarozsoy#!/vizhome/final_story/StoryPlayersinFIFA18

Summary

Football players dataset of FIFA 18 and Tableau are used for creating data visualization that tells player skills varying with age, wage-value relationship and nationality distribution. As a result; Spain is the leading country with highest number of the most valued player, older players are likely worse than youngers in agility, balance, sprint speed and stamina skills, older goalkeepers are likely better than younger ones in goalkeeping skills. Interactive visualizations are common in this project to allow audience to create to further explore particular points especially for age-skills relationship and club insights with Value, Wage, Overall, Nationality, Position and Skills.

Design

My initial design was the same of the first draft. Firstly, I planned to use the most valuable 1000 players with “filled map” by nationality dimension. Also, I aimed to show player names when audience look and select correspondent nationality in filled map. Secondly, I compare clubs with their total paid wages and total player values by using scatter plot with circles. Size of circles changes with overall performance rating and plot can be filtered by Player Position. Third dashboard was similar the previous one, just one-level deeper. It compares players in the selected same team. In that page, I targeted to show player skills when audience select the correspondent player in the plot. Lastly, I select Spanish La Liga football league to investigate and compare numbers of native players. In addition, it is possible to see nationality distribution and line-up of selected team.

Feedbacks

This is collection of verbally feedbacks from my two colleagues during my work.

> *General: Story flow could be more fluent, parts are slightly distinct each other.*

I changed story flow by adding new parts and removing irrelevant parts. I added a new introducing story point which shows all players by nationality on map. In next drafts, I also added indicators for that players of the selected nation. I decided to main point of my story is "age and players", so I investigate relationships with age more deeply in the middle slides by using "Pages". In the last two final story points, I increased interactivity by using actions. I removed Initial Story Point - 4 to avoid losing focus and restricting the audience in narrow investigation by jumping to the Spanish League.

General: There are no trend lines in any plots.

I used trend lines when investigations relationship age with player skills. Polynomial model trends gave result with high R-squared values that means correspondent formulas explain most part of the data.

Initial Story point - 1: Numbers on the countries so distractive and not readable.

I removed text labels for nationalities.

Initial Story point - 2: Layout of page is not well-designed. Middle column is squeeze other columns in.

I used horizontal column to place that informational indicators.

Initial Story point - 2: Annotation may help readers to understand main points.

I created annotation to emphasize that Barcelona and Real Madrid are differ from others.

Extras:

I added annotation for Spain in Final Story Point - 2.

I added age dimension to comparison of clubs by color palette.

I made a bar chart sorted clubs by average overall and average potential value. This is actioned by selection of club from the left plot.

Problematic Parts:

Actions should be designed carefully with source sheet, target sheet and clearing selection. I spent excessive amount of time when researching Aggregation of top n-parameter of nested sorting fields and conditional formatting with calculated fields (dynamic-colored up/down shapes) but results are not appropriate to use.

Datasheet had some duplicated values so I have removed that duplicated IDs. Furthermore, due to Tableau Map and other WMS services did not let me use England, Scotland, Northern Ireland and Wales as country region, I change them as the United Kingdom in dataset.

Since some names are duplicated but they are different players, I used ID as dimension to define unique player instead of Name.

Resources:

https://community.tableau.com/community/forums/

https://www.tableau.com/support/knowledgebase