Week 4 Project- Data Viz

Fifa 18 Database

Project Flow

Extract

- Download Fifa 18 dataset
- Verify that data does not contain overwhelming nulls or inaccuracies

Transform

- Appropriate format data types found with dataset (ex. Country = ‘Geographical’ instead of ‘String’)
- Set data hierarchies to improve organization of data and their relationships

Load

- Use Tableau to connect to provided dataset and transformed data
- Perform various forms of data visualization on selected data

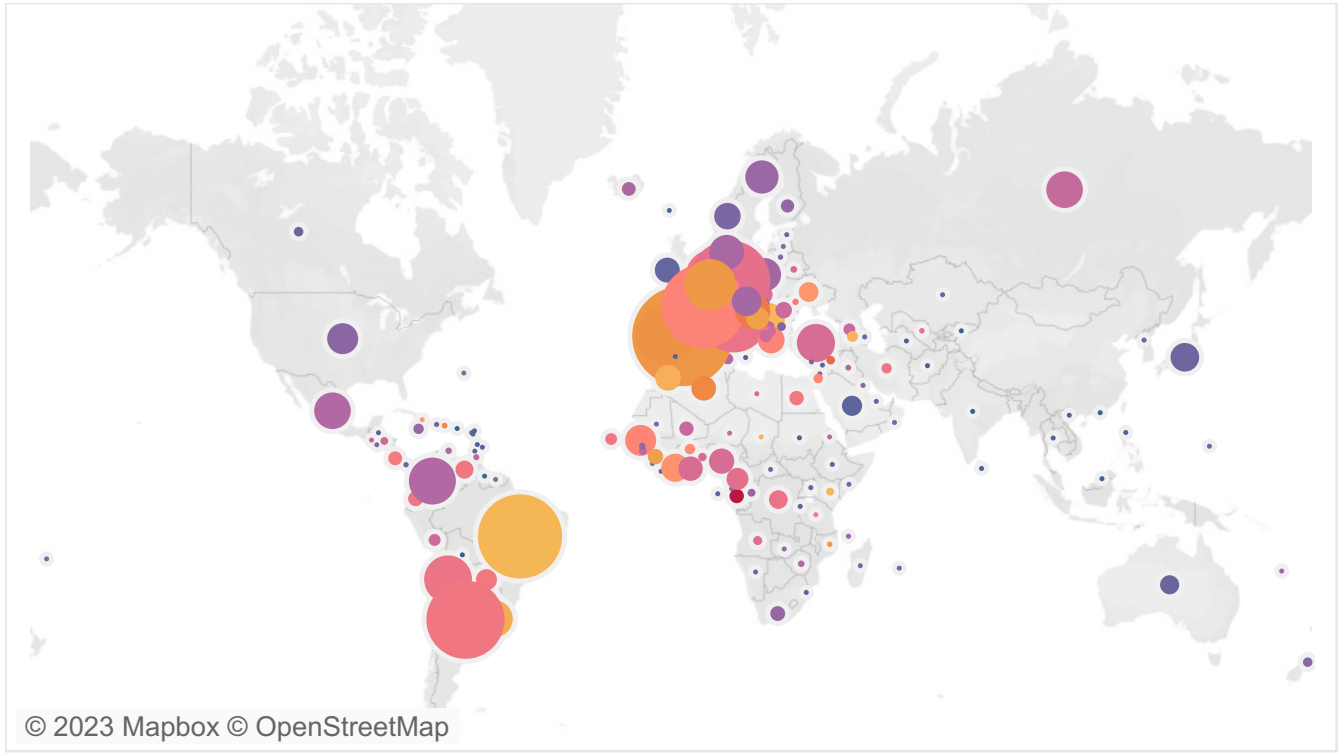
Central Premise/ Question

- Assume the role of the owner of a new football club (ex. Lighthouse F.C.). Which players does your team recruit and why?

Factors

- Financial constraints would lie high on the list of priorities. That being the case, the best players for our club would be determined based on a ratio of price to performance.
- The first inquiry would be made on the entire market of players to see which countries have the biggest accruals of talent and player value on the global scale.
- Next, we narrow the search by clustering player groups by independent variables such as overall, age, and potential.
- As a result, we have grouped players into three groups (developing players, developed players, and players past their prime), of which, developing players seem to be the most lucrative.

Map of Player Market (Wage and Value)



Sum of Value (€)

0
1,000,000,000
2,000,000,000
3,000,000,000
4,000,000,000
4,603,275,000

Avg. Value (€)

0 7,903,333

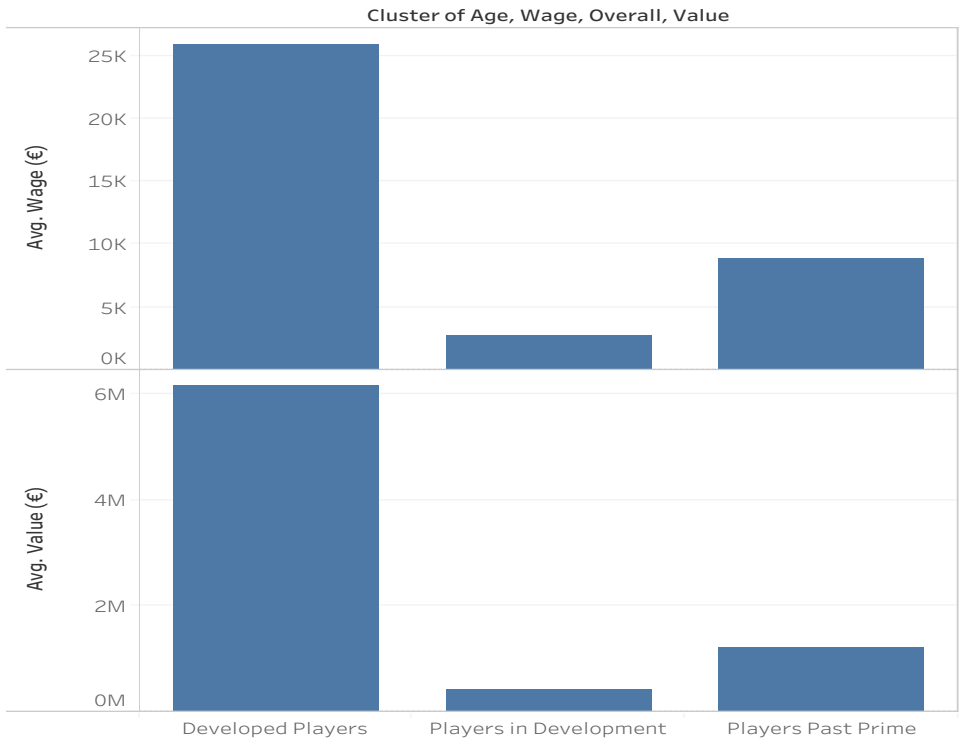
Map based on Longitude (generated) and Latitude (generated). Color shows average of Value (€). Size shows sum of Value (€). Details are shown for Nationality.

Using these bar charts, we ascertain that players in development (young + talented) offer the greatest return on investment in terms of skill.

Avg Total Attribute of Player by Cluster

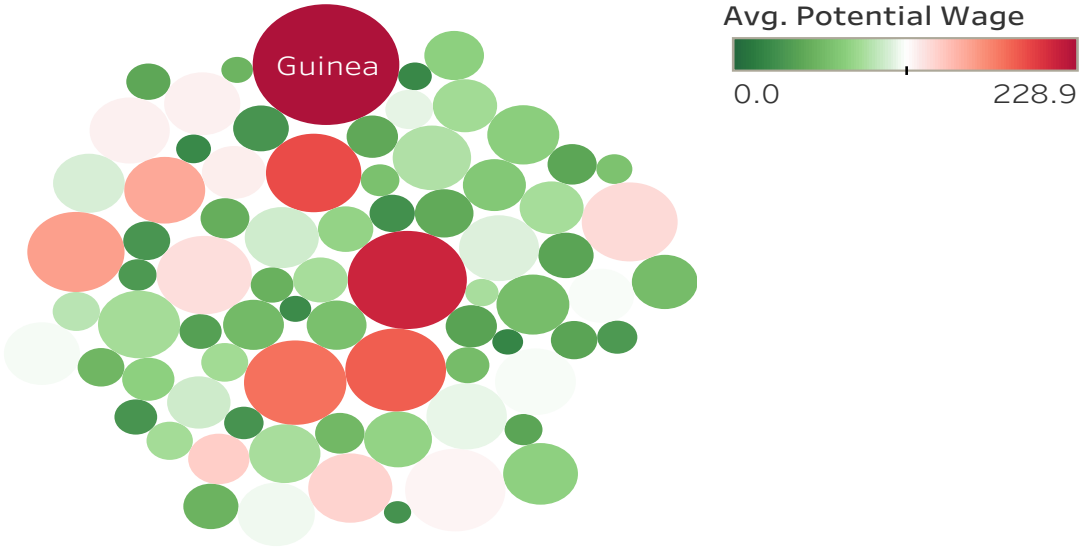
Average of Potential and average of Total Attributes for each Cluster of Age, Wage, Overall, Value.

Avg Wage and Values of Players in Clusters



Average of Wage (€) and average of Value (€) for each Cluster of Age, Wage, Overall, Value.

Potential Market for Developing Players



Nationality. Color shows average of Potential Wage. Size shows average of Potential Cost. The marks are labeled by Nationality. The data is filtered on Age and count of ID. The Age filter keeps 7 members. The count of ID filter ranges from 5 to 753.

Using the bar charts from the previous slide, we narrow our search to players grouped into our in-development cluster. Looking at this cluster further we can see which countries offer more attractive markets for these types of players.

Results

Lastly, we want to focus our search further by looking at the countries that on average offer the lowest valuations and contracted wages on their developing players.

Thus, we can see that countries like Mali, Morocco, Venezuela, and Argentina will offer us the best bang for buck for our investments on potential players.

Challenges

- One challenge that I came across was the lack of time centered data found in the dataset. This limited the ability to forecast future trends in the footballer market that would further enhance our insights and conclusions.
- Another was getting my file corrupted by Tableau after completion, another reminder to push to git more often!