DATA VISUALIZATION USING TABLEAU

SUMMARY

In this project, i worked on baseball players dataset to create visualizations among different relations among different variables in the dataset. There are about 1157 rows of data, each containing details about many players. First i have shown general information regarding the data and then compared different variable to visualize the relations between these variables like height, weight, Home Runs, Average Score and Handedness. Created Visualizations to answer the questions that i want to explore initially.

Initial Story

https://public.tableau.com/profile/mallidi.akhil.reddy#!/vizhome/Story_299/Story1?publish=yes

Final Story

https://public.tableau.com/profile/mallidi.akhil.reddy#!/vizhome/PerformanceofBaseballplayers/PerformanceofBaseballplayers?publish=yes

DESIGN

Initial Story

- In the first plot, to get the general summary of data and the total range's of data present in the dataset, i plotted a plot between the sum of Measure Values.

- In the second plot, i plotted a bar plot to see how many different handedness players present in the dataset, by plotting between handedness(Categorial Value) and number of records.

- In third plot, a bar plot is plotted between handedness and sum of heights. Since it is sum of records, and there are more right handed players the results may not be accurate.

- So fourth plot is plotted between average of heights and handedness to get insights between them.

- In fifth plot, a bar plot is plotted between handedness and sum of weights. Since it is sum of records, and there are more right handed players the results may not be accurate.

- So sixth plot is plotted between average of weights and handedness to get insights between them.

- In seventh plot, scatter plot is plotted between weight and height to know about the relationship between these variable's.

- Outliers are plotted to dig more deeper in the eight plot.

- In the ninth plot, scatter plot is plotted between height and weight after removing these outliers.

- In tenth plot, top ten players are visualized with respect to average score.

- In 11th plot, top ten players are visualized with respect to Home Runs (HR).

- In 12th plot, to get insights about which players(handedness) has good performance(based on average score, plotted against sum of average score).

- In 13th plot, to get insights about which players(handedness) has good performance(based on average score, plotted against median of average score).

- In 14th plot, to get insights about which players(handedness) has good performance(based on HR score, plotted against sum of HR score).

- In 15th plot, to get insights about which players(handedness) has good performance(based on HR score, plotted against median of HR score).

- In 16th and 17th plot, Home Runs (HR) score was plotted against height and weight, to get the top most performer's height and weight.

- In 18th and 19th plot, Average score was plotted against

height and weight, to get the top most performer's height and weight.

Final Story

- The name of the Dashboard1 is renamed to Summary of the data.

- The name of the Story1 is renamed to Performance of Baseball players.

- Logarithmic Trend Line is added to the scatter plot plotted against height and weight.

- Since i am comfortable with these colors, i am skipping this suggestion.

FEEDBACK

- The name of the summary dashboard (named as Dashboard1) should be changed and it should be explanatory about the topic.

- The name of the story (named as should be changed Story1) should be changed and it should be descriptive and explanatory.

- Add a trend line in the scatter plot plotted against height and weight to get more clear understanding.

- Use more Attractive color palette.

CONCLUSION

- Left handed players performance is higher when compared to right handed and both.

- The highest performing player's (based on Average Score) weight is about 170 pounds.

- The highest performing player's (based on Average Score) height is about 72 inches.

- The highest performing player's (based on Home Runs Score) weight is about 195 pounds.
- The highest performing player's (based on Home Runs Score) height is about 72 inches.

RESOURCES

1 - Udacity Nanodegree Videos

2 - https://onlinehelp.tableau.com/current/pro/desktop/en-us/trendlines_add.htm

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

4 - https://community.tableau.com/thread/120707

5 - https://stackoverflow.com/questions/25321106/tableau-filtering-on-duplicates