

Basket Ball Oproject :

SUMMARY : I have selected Basket Ball data file . It is a data set containing 1,157 baseball players including their handedness (right, left handed or both), height (in inches), weight (in pounds), batting average, and home runs. I explored fields like number of Records , finding top players .I have explored features like number of records , handedness of players, height of players , weight of players , average batting score, homeruns made. I have a set of TOP PLAYERS , and two new variables BMI and No. of records = 2. The graphs which I've used in my Explanatory analysis are Bar-chart,Piechart and Scatter-plot and Line Graph. Link of Initial File is : <https://public.tableau.com/profile/kinjal7806#!/vizhome/Basketballinitial1/Story1> Final File link is : <https://public.tableau.com/profile/kinjal7806#!/vizhome/BasketballFinal/Story1>

DESIGN :

1) I have created a set of top players , I have selected 21 players so that the subset was able to include Both handed player , as number of Both handed players is less , selecting number like 10,15 didn't include significant number of both handed players. Top 21 players are selected on the basis of high score of homeruns.

Updated information : Why use a bar graph for Top Players.:

As I sorted the data in descending order , a bar graph was a very good representation to show this skewed data and ,bar-graph will show the best comparison with bars and I've took the set of highest HR's scored in each Handedness category and visual encoded with color.

2) I made piechart of top players having average by handedness.

3) Similarly I made piechart of top players having Homerun by Handedness.

Updated information : Why use a pie chart for above two cases:

Handness is divided in 3 categories i.e left , right and both , and pie chart is an easy representation of how much each category contains . Looking at the pie chart we can approximately guess which category of players contribute more.

4) Then I plotted a Line graph , with average batting score on y axis and Homerun on x axis and plotted them into sections with respect to handedness.

Updated information : Why use a Line graph for handedness , Homerun and Average Relationship:

I have used line graph because it shows easy transiting in the trend of values , also it efficiently conveys the range . For example : For the values for average batting score to homerun , line graph of both handed players as a short range and doesnot show any downtrend, this message is conveyed properly. While plotting the summary of handedness on Dashboard I have taken care of the hoverness and actions made with help of highlights.

5) While working with number of records I observed that somehow players with maximum height showed high Number of record.

Updated information : Why use a Bar graph and Visual Encoding :

For Tall Player:

While plotting the height of player in lbs. on the bar graphs, I have sorted it in descending order , visually encoding the height showed an easy trend in decrease in the value of height of the player, also it is height of player is shown as label, for better clarity , hence double encoding is done.

For maximum number of records:

Bar graph is used on other hand to effectively show that the player has reached no. of records = 2 , distinct values of number of records are on Y axis, also label boolean value , True, for No. of records = 2 is shown for clarity.

6) I created a boolean variable No. of Records = 2 , which is true for Sum(Number of Records) equals to 2 , hence while comparing this boolean value to height , tall players showed better number of records.
relationship of average and handedness , relationship of homerun and handedness .

7) I found that BMIs of 49 and above show maximum average and Homerun.

7, I found that BMI of 15 and above show maximum average and homerun.

Also , I worked on different levels of BMI , and categorized them in Normal , Overweight and Obese.

Updated information : Why use a Scatter plot and Visual Encoding for BMI, Height , Weight:

There were n number of data samples and BMI divided in 3 groups (Normal,Over-weight and Obese) and changes Bmi class from(normal to over-weight) can be clearly noticable with changes in height and weight . Visual Encoding is used to show which data point belongs to which of the 3 groups.

FEEDBACK : I shared and discussed the project file with my udacity friend , he suggested me I should work with height and weight , like creating Body Mass Index, then I tried to explore relationship of BMI, average and Homerun .

Reference : Udacity tutorials.