

DATA ANALYTICS AND VISUALISATION WITH TABLEAU

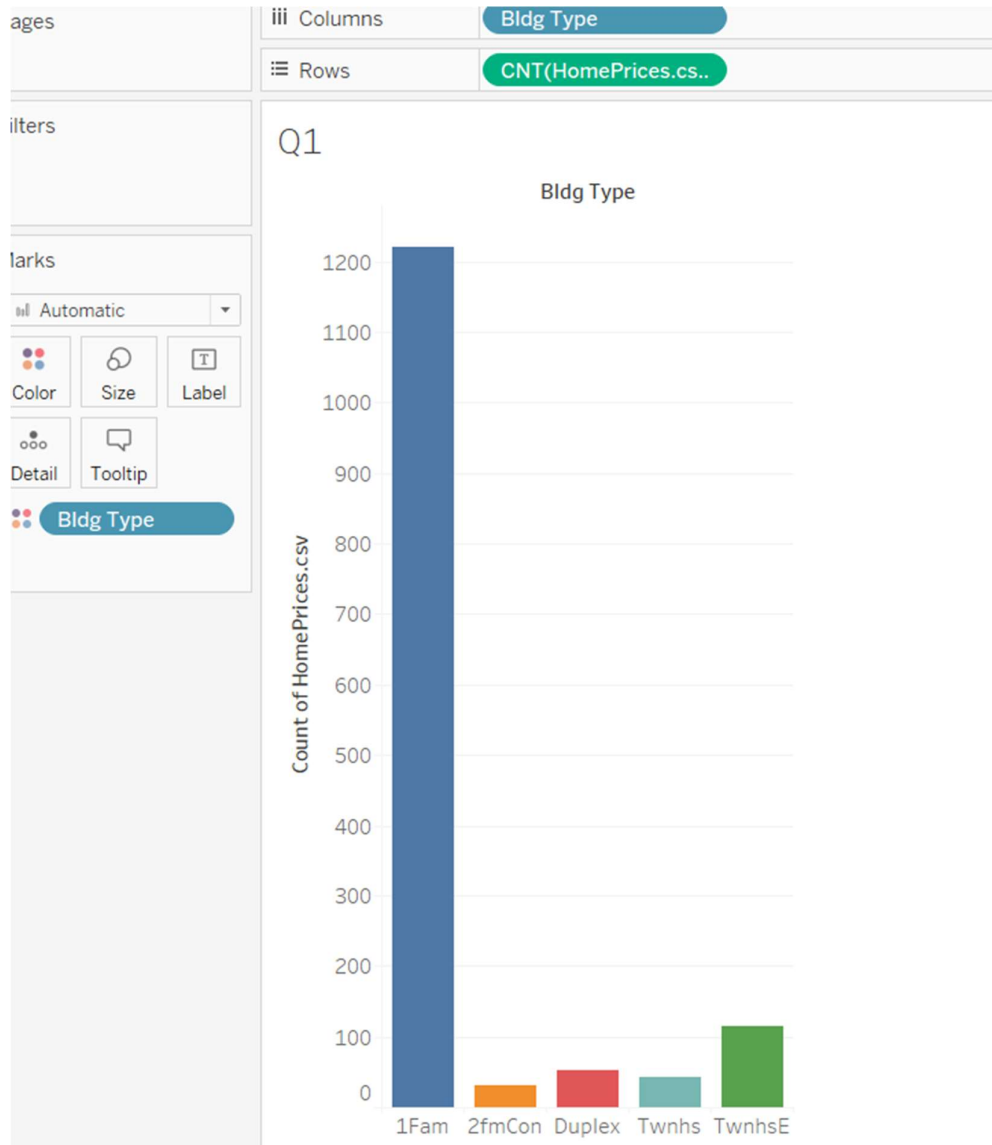
Assignment: Day 3

PES2UG19CS315

Rahul S Bhat

7th semester

1. Create a bar chart showing the distribution of building types:



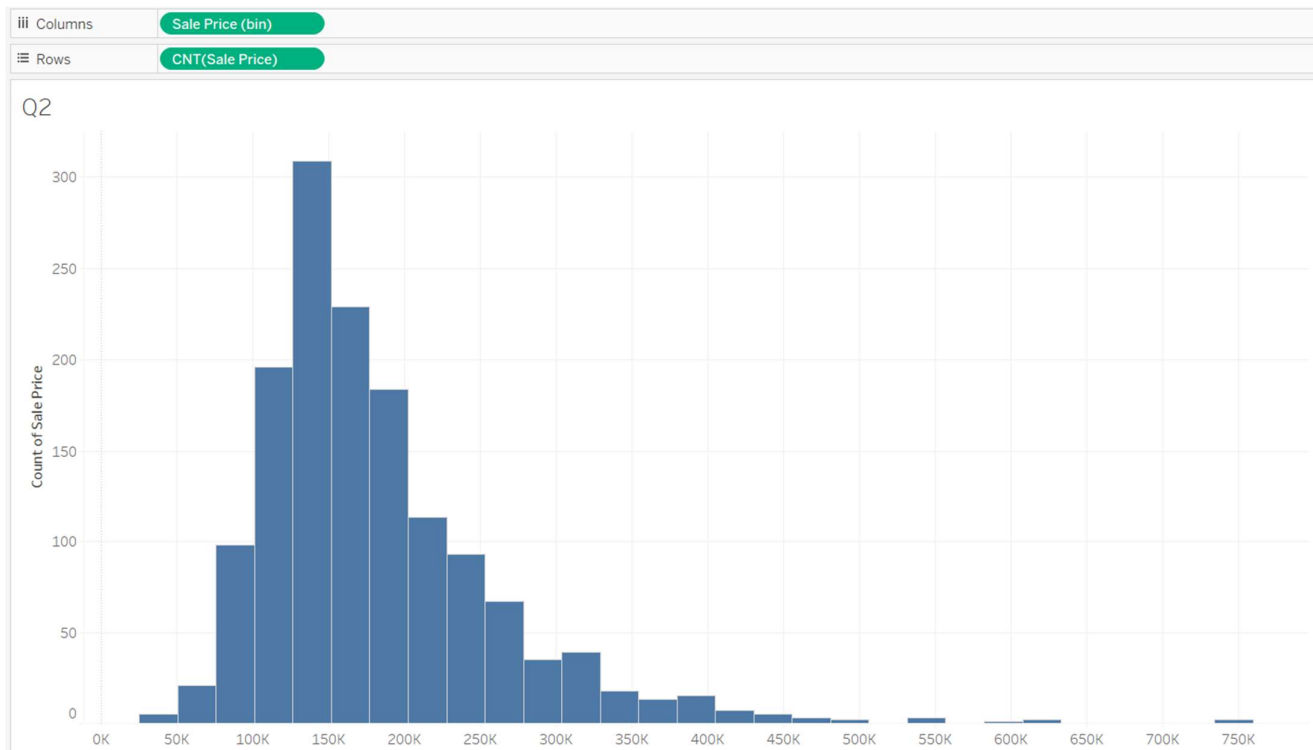
Question 1: What have you learned about the distribution of building types? Which Building Type has the most homes for sale? The least?

The distribution shows the no. of homes for sale for each building type. The distribution is not uniformly spread across different building types.

1Fam has the most homes for sale

2fmCon has the least homes for sale

Create a histogram of Sale Price:



Question 2: What field did Tableau create to make the histogram? Why?

The field created by Tableau is Sale Price(bin)

Histogram depicts the distribution of data across bins and this requires a fixed bin size to be calculated which is automatically calculated by Tableau

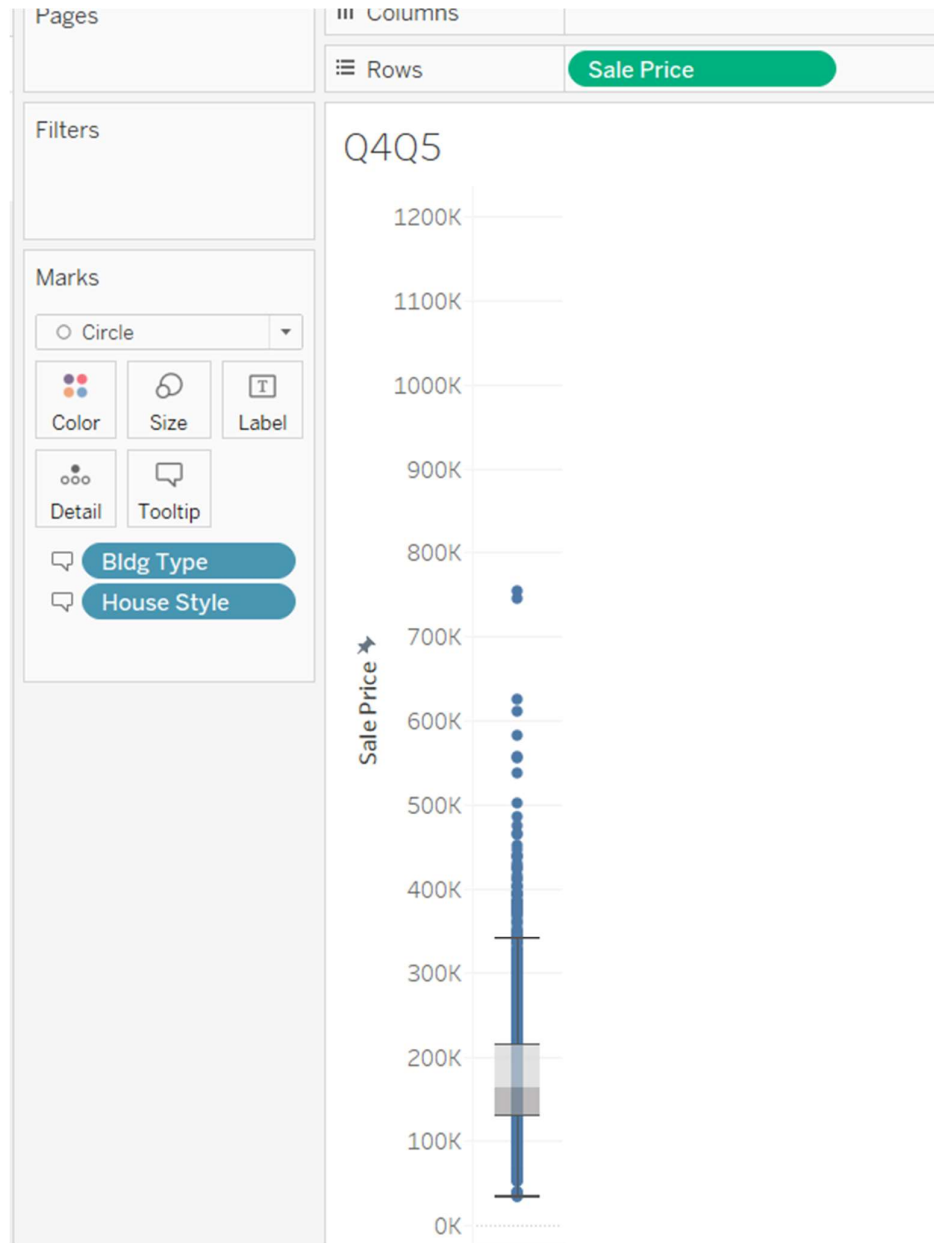
Question 3: What is the shape of the histogram? Will the mean or median have a higher value? Which value (mean or median) would be a better measure for the center of the distribution?

Histogram is left skewed

Mean will have the higher value than the median

As the distribution is skewed, Mean would be a better measure for centre of distribution

Create a boxplot of Sale Price:



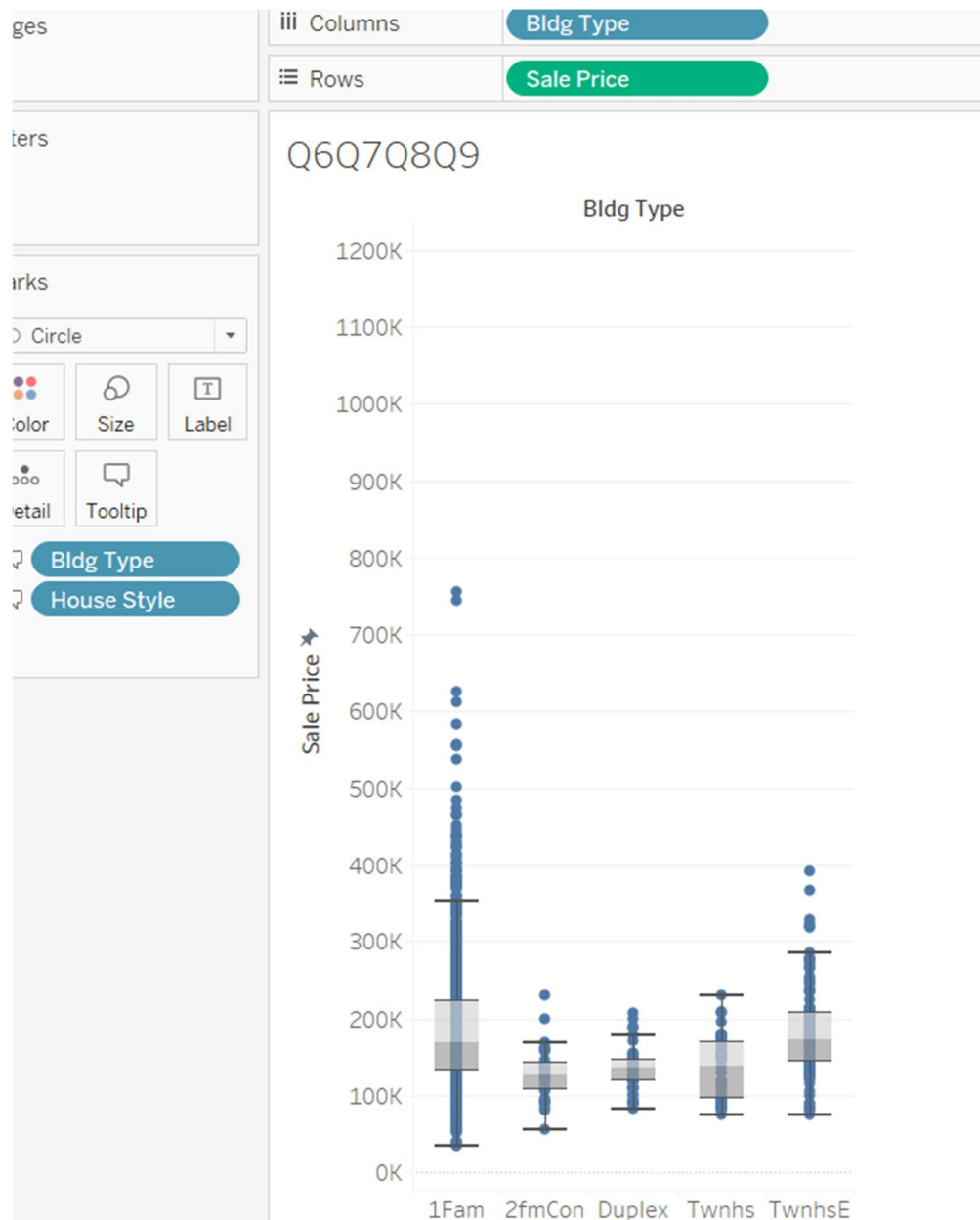
Question 4: Are there outliers present? If yes, are outliers above or below the center of the distribution?

Outliers are present and are above the centre of distribution

Question 5: What house style has the highest sale price?

2 Story house style has highest sale price

Create a set of box plots of Sale Price for Building Type:



Question 6: Which building type has the lowest median Sale Price?

Lowest median sale is for 2fmCon

Question 7: Which building type does not have any outliers in Sale Price?

Twnhs type has no outliers

Question 8: Which distribution has the largest spread? The smallest spread?

Largest spread is for 1Fam. Smallest spread is for Duplex

Question 9: If you were a real estate agent working with a new client, how would this graph be helpful to you? Based upon this view, what other questions do you want to ask? (*hint: Look at the other fields in dimensions for possible ideas. Remember that you can add any dimension to the columns shelf*)



1Fam type has the largest spread while the Duplex has the least spread

Considering the Sale condition, it is seen that the 1Fam Normal has the widest spread amongst all and we can say that people are more likely to opt for 1Fam Normal as there are a lot of options in this home style. And second stands the 1FmPartial.

There are a less options for 2fmCon Partial, Twnhs Family and TwnhsE Family and the prices are low as well. With least no. of options to select from, people are more likely to not to buy it.

Other questions would be the contributing factors like the locality, tenant type, maintainance, good facilities nearby etc which might also have an effect on the sales