

Section 2

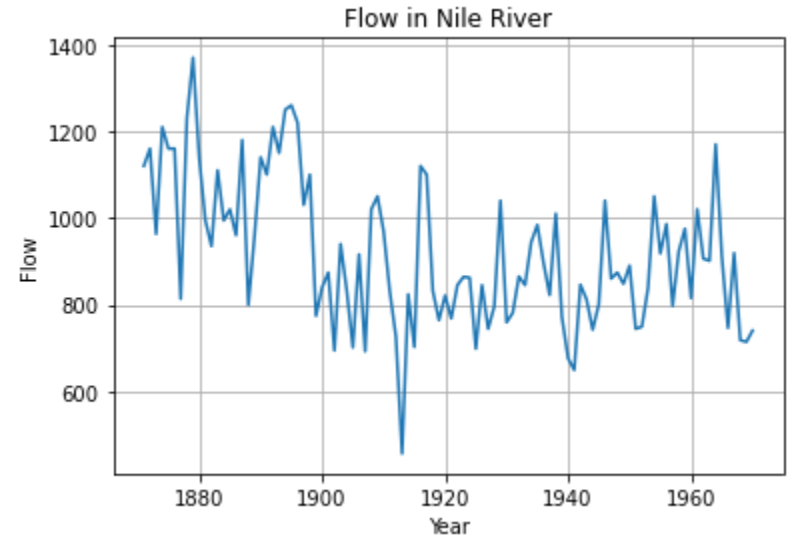
Data Visualization (Summary)

Variable Type	Plots
Time Series	Line Plot
One Discrete	Bar, Pie
One Continuous	Histogram, Density (KDE), Box and Whisker
Two Continuous	Scatter Plot
One Continuous, One Discrete	Box and Whisker (multiple)

Selecting a Plot

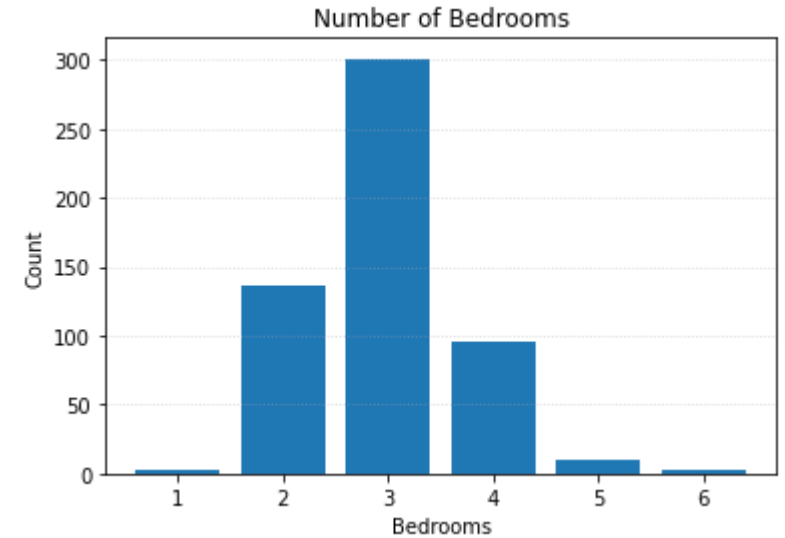
❖ Line plots could be used to show trends.

- ❖ Time on the X axis
- ❖ Variable on the Y axis



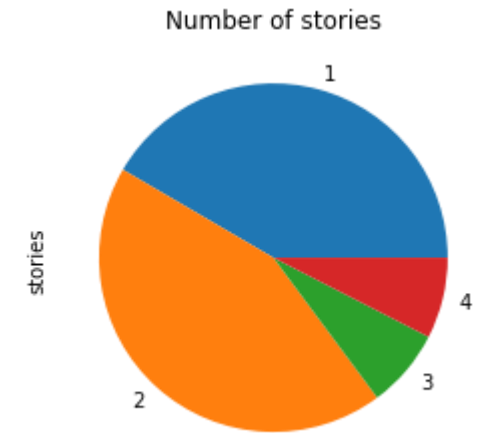
Line Plot

- ❖ Bars of different lengths show the distribution
 - ❖ Category on the X axis
 - ❖ Counts or the proportions on the Y axis
- ❖ If there are too many categories, it could become confusing.
- ❖ Don't show averages per category using bar plot. Use boxplot or violin plot in that case.



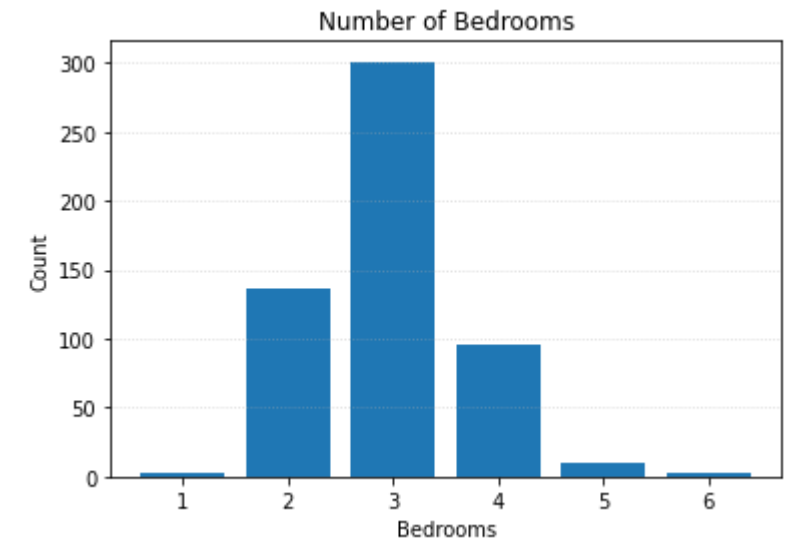
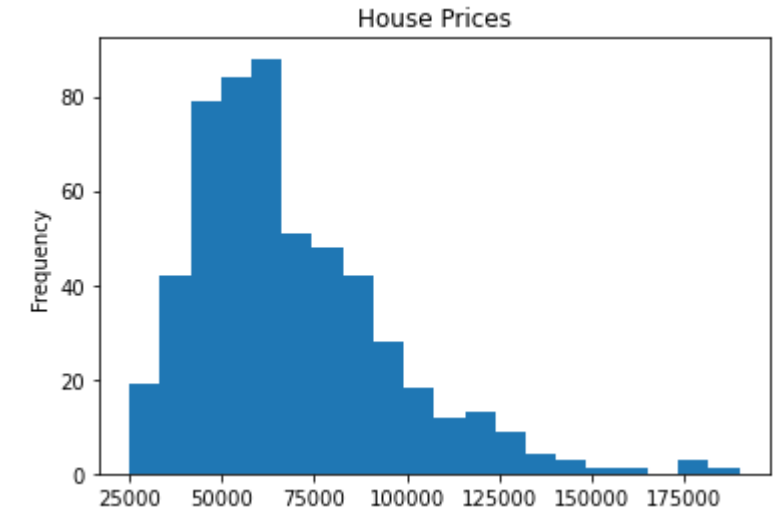
Bar Plot

- ❖ Pies show the distribution
 - ❖ Category on the X axis
 - ❖ Counts or the proportions on the Y axis
- ❖ Be careful to use it when the proportions are almost equal. In those situations, bar chart might be a better choice.



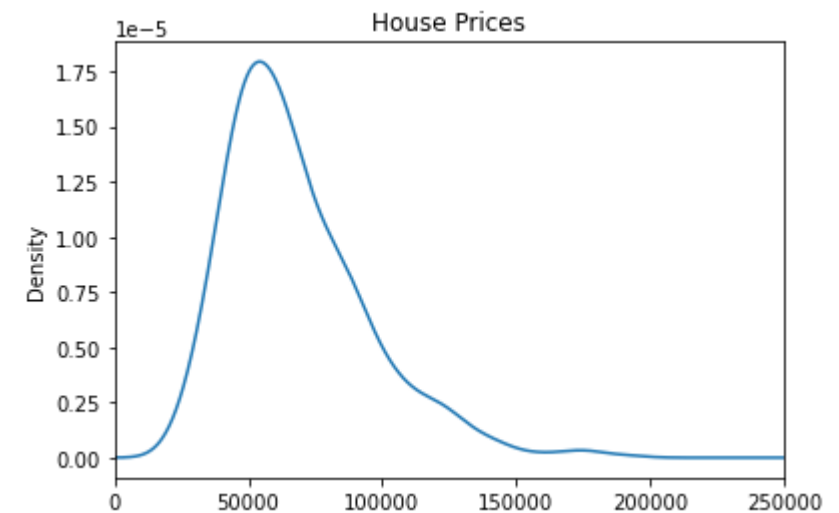
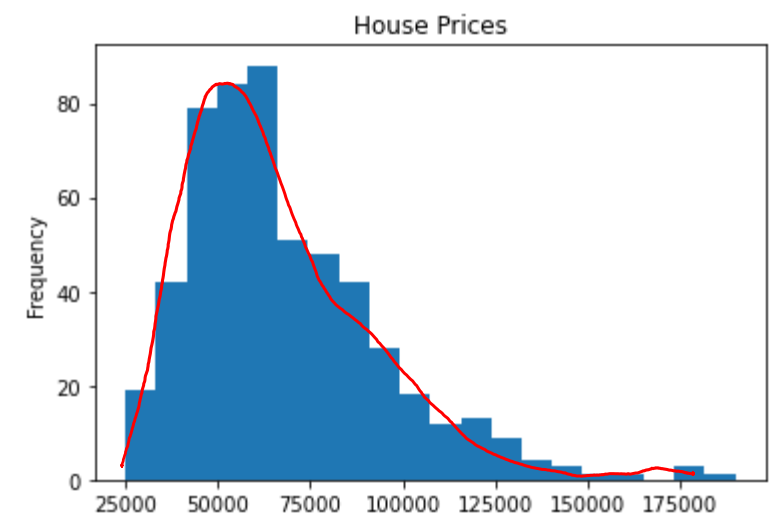
Pie Plot

- ❖ Histograms show the distribution of a continuous variable
 - ❖ Value on the X axis
 - ❖ Frequency on the Y axis
- ❖ Bar Plot has the discrete variable on X axis, and Histogram has a continuous variable.



Histogram

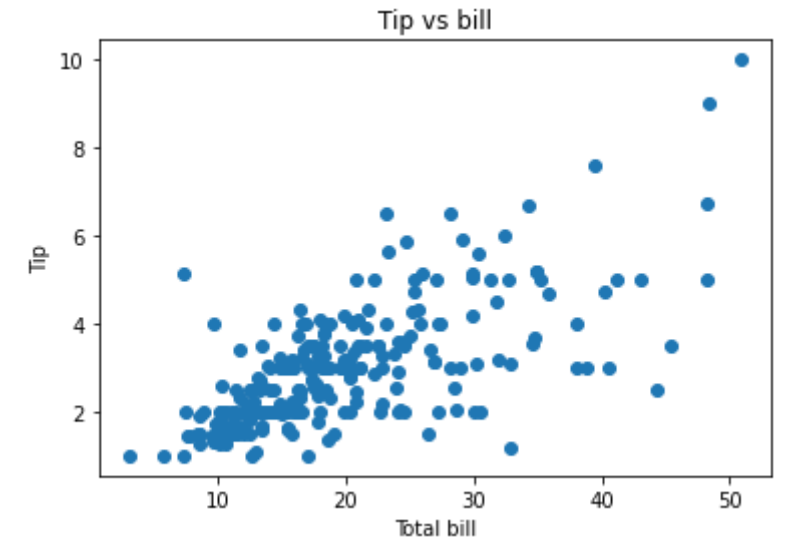
- ❖ KDE or the Density Plot is used to visualize the distribution of observations (like Histogram)



KDE or Density Plot

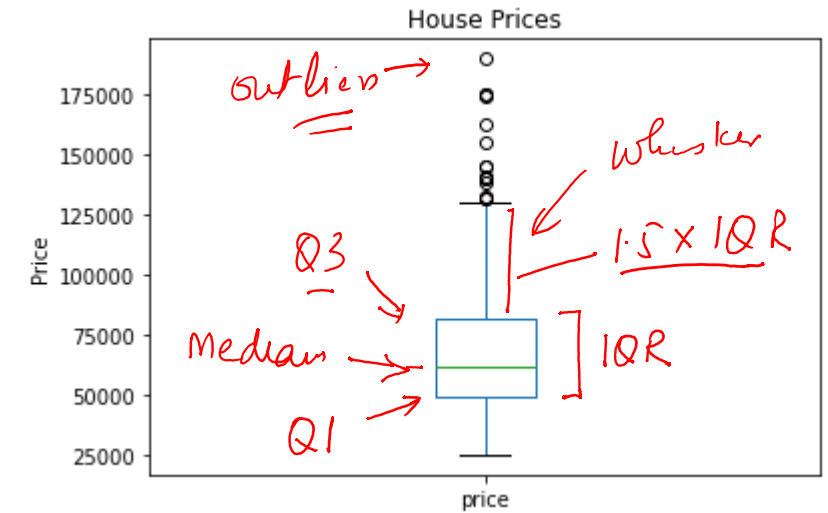
❖ Scatter plot is used to show the relationship between two continuous variables.

- ❖ Predictor on the X axis
- ❖ Response on the Y axis

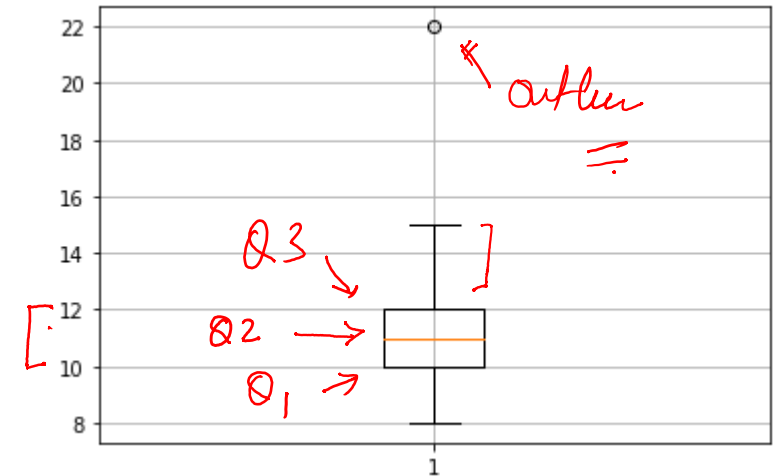
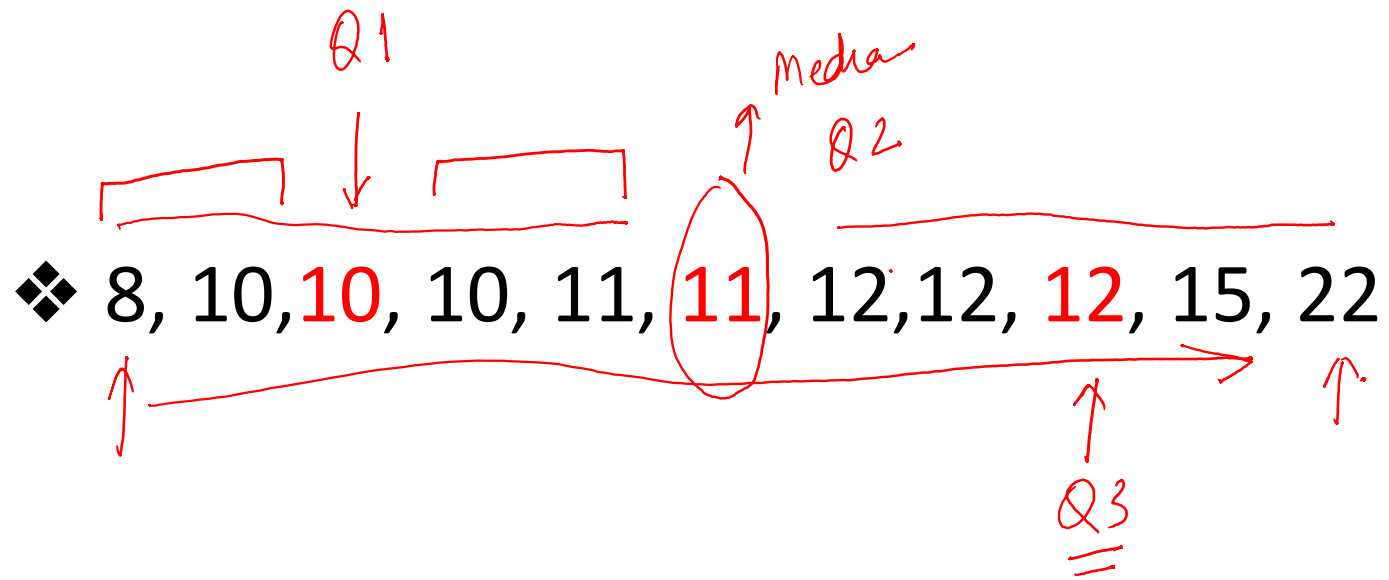


Scatter Plot

- ❖ Box-and-Whisker Plot shows the summary data as three quartiles (Q1, Q2 and Q3) and whiskers.



Box and Whisker Plot



Box and Whisker Plot

- ❖ Multiple Box-and-Whisker Plots show the distribution of a continuous variable(price) by a discrete variable (number of bedrooms).



Box and Whisker Plot (Multiple)