Seaborn

Agenda

Key Takeaways-

- What is Seaborn?
- Plotting All basic plots using Seaborn
- Assessing symmetricity and skewness
- Heatmap
- Some Advanced visualization libraries

Seaborn

- Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

"If Matplotlib "tries to make easy things easy and hard things possible", seaborn tries to make a well-defined set of hard things easy too." – Michael Waskom (Creator of Seaborn)

- Seaborn integrates closely with pandas data structures.
- It requires less plotting effort compare to Matplotlib.

Heatmap

- A heatmap is a two-dimensional graphical representation of data where a matrix values are represented in different shades of colors.
- A heatmap aims to provide a color coded visual summary of data/information.
- Seaborn allows annotated heatmaps as well.

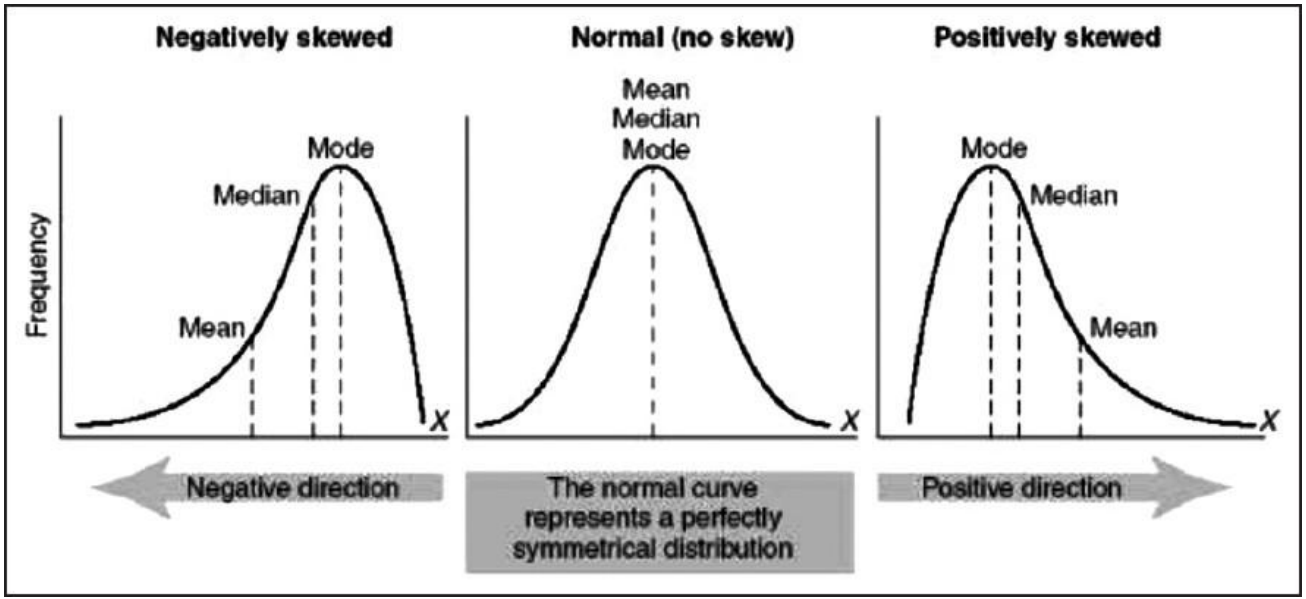
Skewness

- Skewness is a statistical parameter to measure asymmetricity about the mean in a distribution of random variable.
- This parameter value can be positive, negative or undefined.
- A negative value indicates data is left skewed whereas a positive value indicates data is right skewed.

$$Skewness = \frac{3 * (mean - median)}{standard\ deviation}$$

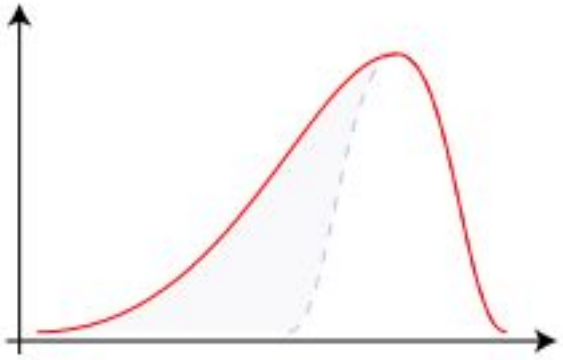
Symmetric and Skewed data

- For a Symmetric data, Mean = Median = Mode
- For a Left skewed data, Mean < Median < Mode
- For a Right skewed data, Mean > Median > Mode



Quiz 2

The following dataset distribution belong to -



A. Right skewed data

B. Left skewed data

C. Mean > Median for this dataset

D. Mean < Median for this dataset