

Python for Visualization

Agenda



- 1. Pop Quiz
- 2. Data Visualization
- 3. Visualization One variable
- 4. Visualization Two variables
- 5. Visualization Multiple variables

Pop Quiz



- 1. What is data visualization?
- 2. Why does data visualization help?
- 3. How to visualize numerical and categorical variables?
- 4. What is correlation?

Data Visualization



- Data visualization is the process of translating data and metrics into charts, graphs and other visuals.
- The resulting visual representation of data makes it easier to identify patterns, trends, and outliers hidden in the data, enabling us to gain better insights.
- We can use different charts/plots to visualize different kinds of data
- Each chart/plot helps us gain insights from a different perspective





Plot	Type of Data	Usage	Example
Histogram	Numerical	Helps us understand data distribution by dividing it into bins and showing the number of observations in each bin via bars	0000 10000 10000 20000 20000 5000 60000 40000
Histogram with density curve	Numerical	Helps us understand data distribution by displaying a distribution curve on top of the histogram bars	00 00 100 100 100 100 100 100 100 100 1
Boxplot	Numerical	Helps us understand data distribution and skewness by displaying the data in the form of a box divided by quartiles is file is meant for personal use by diegoff@gmail.com only.	1500 2000 2500 3000 3500 4000 auto, weight





Plot	Type of Data	Usage	Example
Line Plot	Numerical	Helps us understand the trend or pattern in the data by displaying it as straight lines formed by connecting individual data points	340 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 -
Violin Plot	Numerical	Helps us understand data distribution by plotting a density curve symmetrically around a boxplot	50 100 160 2d0 250 300 terruposeer
Bar Graph	Categorical Th	Helps us understand data distribution by showing the counts of observations in each level (or group) using bars is file is meant for personal use by diegoff@gmail.com only.	80 - 66 - 66 - 70 - 70 - 70 - 70 - 70 - 7

Visualization - Two Variables



Plot	Type of Data	Usage	Example
Scatter Plot	Numerical	Helps us understand potential relationship between two numerical variables	250 - 200 - 200 - 200 - 200 - 200 250 300
Implot	Numerical	Helps us understand and measure the relationship between the two variables quantitatively	250 200 200 200 300 300 400 and
Joint Plot	Numerical	Helps us understand the distribution and relationship between two numerical variables on the same plot. Is file is meant for personal use by diegoff@gmail.com only.	200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 - 200 -





Plot	Type of Data	Usage	Example
Strip Plot	Categorical	Helps us to visualize the distribution of a numerical variable wrt different categories of a categorical variable	20 20 20 20 20 20 20 20 20 20 20 20 20 2
Swarm Plot	Numerical	Helps us to visualize the distribution of a numerical variable wrt different categories of a categorical variable and avoids overlapping of data points	40000 - 40000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000 - 20000





Plot	Type of Data	Usage	Example
Pair Plot	Numerical	Helps us understand the relationship between two or more pairs of numerical variables	
Cat Plot	Numerical	Helps us understand relationship between a numerical variable and one or more categorical variables	1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360 1360
Heatmap		Helps us understand the correlation between pairs of columns in the data s file is meant for personal use by diegoff@gmail.com only.	wheel_base - 1



Happy Learning!

