Data visualization: graphical representation of data to understand trends, outliers, and patterns that exist in the data.

If you’ve ever stared at a massive spreadsheet of data and couldn’t see a trend, you know how much more effective a visualization can be.

make sense of the trillions of rows. tell stories by curating data into a form easier to understand

[making data more understandable](https://www.forbes.com/sites/jeffkauflin/2017/07/20/the-five-most-in-demand-skills-for-data-analysis-jobs/#3e300312c7ce).

1. A line chart is a type of chart used to show information that changes over time.
2. A bar plot is a plot that presents categorical data with rectangular bars. The lengths of the bars are proportional to the values that they represent.
3. A[histogram](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.hist.html#pandas.DataFrame.plot.hist) is a representation of the distribution of data
4. [Box plots](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.plot.box.html#pandas.Series.plot.box) are used for depicting data through their quartiles. A single box plot can convey a lot of information, including details about the interquartile range, median, and outliers

Diagram

Description automatically generated

1. An[area plot](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.plot.area.html#pandas.Series.plot.area) displays quantitative data visually.
2. A[scatter plot](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.scatter.html#pandas.DataFrame.plot.scatter) is used to plot correlations between two variables
3. A[pie plot](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.pie.html#pandas.DataFrame.plot.pie) is a proportional representation of numerical data in a column

[Box plots](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.plot.box.html#pandas.Series.plot.box) are used for depicting data through their quartiles. A single box plot can convey a lot of information, including details about the interquartile range, median, and outliers

bar chart is a plot that presents categorical data with rectangular bars. The lengths of the bars represent the measurement corresponding to each category

line chart is a type of chart used to show information that changes over time.

A[pie plot](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.pie.html#pandas.DataFrame.plot.pie) is a proportional representation of

A[scatter plot](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.plot.scatter.html#pandas.DataFrame.plot.scatter) **graphs pairs of numerical data, with one variable on each axis, to look for a relationship between them**

* Linearity refers to whether a data pattern is linear (straight) or nonlinear (curved).
* Slope refers to the direction of change in variable Y when variable X gets bigger. If variable Y also gets bigger, the slope is positive; but if variable Y gets smaller, the slope is negative.
* Strength refers to the degree of "scatter" in the plot. If the dots are widely spread, the relationship between variables is weak. If the dots are concentrated around a line, the relationship is strong.

Chart, scatter chart

Description automatically generated

8:15 ->8:25