UDACITY

< Return to Classroom

Investigate a Dataset

REVIEW

HISTORY

Meets Specifications

Dear student,

You did a very good job! I'm also sending you two good links regarding communicating your results:

A Data Scientist's Guide to Communicating Results
Communication in Data Science

And two links about data science best practices:

Data Science Best Practices
10 Best Practices for Data Science with Impact

Keep up the good job!
Keep udacious!

Code Functionality

All code is functional and produces no errors when run. The code given is sufficient to reproduce the results described.
Excellent work.

The project uses NumPy arrays and Pandas Series and DataFrames where appropriate rather than Python lists and dictionaries. Where possible, vectorized operations and built-in functions are used instead of loops.

Numpy and Pandas have been used, well done!

The code makes use of functions to avoid repetitive code. The code contains good comments and variable names, making it easy to read.

I strongly encourage you to practice the use of functions to avoid repetitive coding.

Quality of Analysis

The project clearly states one or more questions, then addresses those questions in the rest of the analysis.

The report states clear and relevant questions that are being addressed by the following analysis.

Data Wrangling Phase

The project documents any changes that were made to clean the data, such as merging multiple files, handling missing values, etc.

Exploration Phase

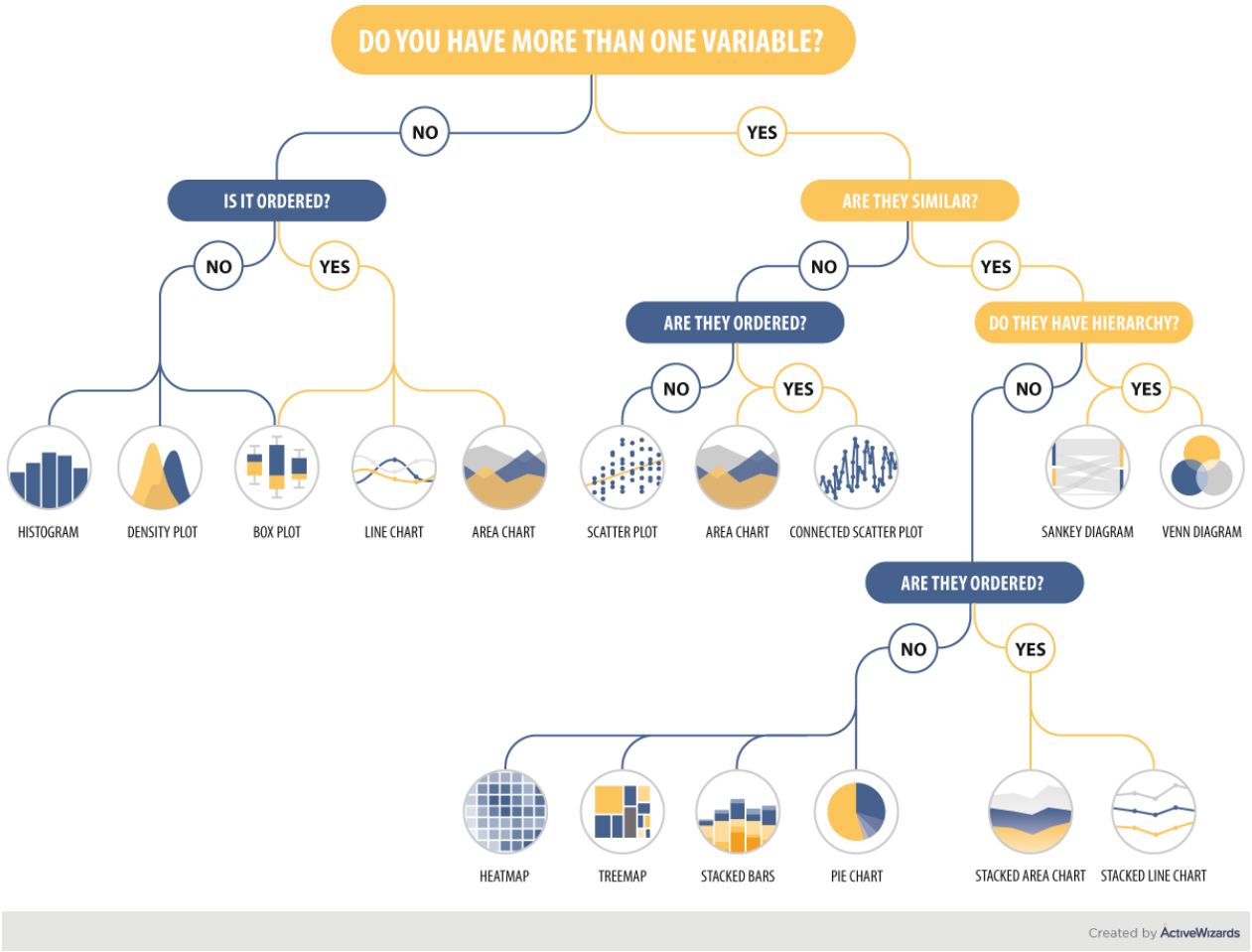
The project investigates the stated question(s) from multiple angles. At least three variables are investigated using both single-variable (1d) and multiple-variable (2d) explorations.

The project's visualizations are varied and show multiple comparisons and trends. Relevant statistics are computed throughout the analysis when an inference is made about the data.

At least two kinds of plots should be created as part of the explorations.

Visualizing data requires a lot of patience and determination because it's not easy selecting the best visualization to match with a given data type. Well enough, the project rightly builds descriptive visualizations.

Here is a nice infographic showing when to use each plot type:



Conclusions Phase

The results of the analysis are presented such that any limitations are clear. The analysis does not state or imply that one change causes another based solely on a correlation.

Communication

Reasoning is provided for each analysis decision, plot, and statistical summary.

Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted.

DOWNLOAD PROJECT

RETURN TO PATH