

[◀ Return to Classroom](#)

Communicate Data Findings

REVIEW

HISTORY

Meets Specifications

Kudos Isaac,

This was a challenging project. You were able to uncover some interesting results. The whole process of evaluating data using analytical or statistical tools to discover useful information can be a delicate one, but this is indispensable to understand what is going around and especially if you intend to build an ML model out of it. Now you have a polished project that you can showcase on your Github. Please share what you learned with your peers.

Notes

- **"Exploratory research"** is what you do to get acquainted with the data. You might start with a theory or query, or you might just really get into the data and figure out what might be important about it. *Exploratory exploration is the method of turning over 100 rocks to reveal 1 or 2 precious gemstones.*
- **"Explanatory review"** is what comes when you have something unique that you want to show the audience-probably those 1 or 2 precious gemstones. *In my blogging and writing, I seem to concentrate mainly on this latter piece, the explanatory analysis, where you've already been through the exploratory analysis, and from this, you've decided something concrete that you want to explain to a given audience: in other words, when you want to say a story with info."*

The choice between descriptive, exploratory, and explanatory research should be made with your research question in mind. What does your question ask? Are you trying to learn the basics about a new area, establish a clear "why" relationship, or define or describe an activity or concept? So, when we work on a Case study (while closely examining the data within a specific context), keep in mind that there are various advantages, in that they present data of real-life situations and they provide better insights into the detailed behaviors of the subjects of interest, however, they can also be criticized for their inability to generalize their results.

Code Quality

All code is functional (i.e. no errors are thrown by the code). Warnings are okay, as long as they are not a result of poor coding practices.

Great work! You took a lot of variables from the dataset and did an excellent job of systematically exploring it and coming up with some interesting findings.

The project uses functions and loops where possible to reduce repetitive code. Comments and docstrings are used as needed to document code functionality.

Well done. Comments were used regularly and were useful where they were made. There were enough Markdown cells to keep track of what was happening, though it might have been good to have used comments to introduce each code cell, rather than waiting for after the plot.

Exploratory Data Analysis

The project (Parts I alone) contains at least 15 visualizations distributed over univariate, bivariate, and multivariate plots to explore many relationships in the data set. Reasoning is used to justify the flow of the exploration.

You are using an excellent mixture of uni-, bi- and multivariate plots to deliver the main points of your analysis, well done! [Here](#) is an interesting article on the differences between univariate and bivariate data.

Questions and observations are placed regularly throughout the report, after **each** plot or set of related plots.

Tip: Use the ""Question-Visualization-Observations"" framework throughout the exploration.

Tip: For the Part I notebook, use *File > Download as... > HTML or PDF* menu option to generate the HTML/PDF.

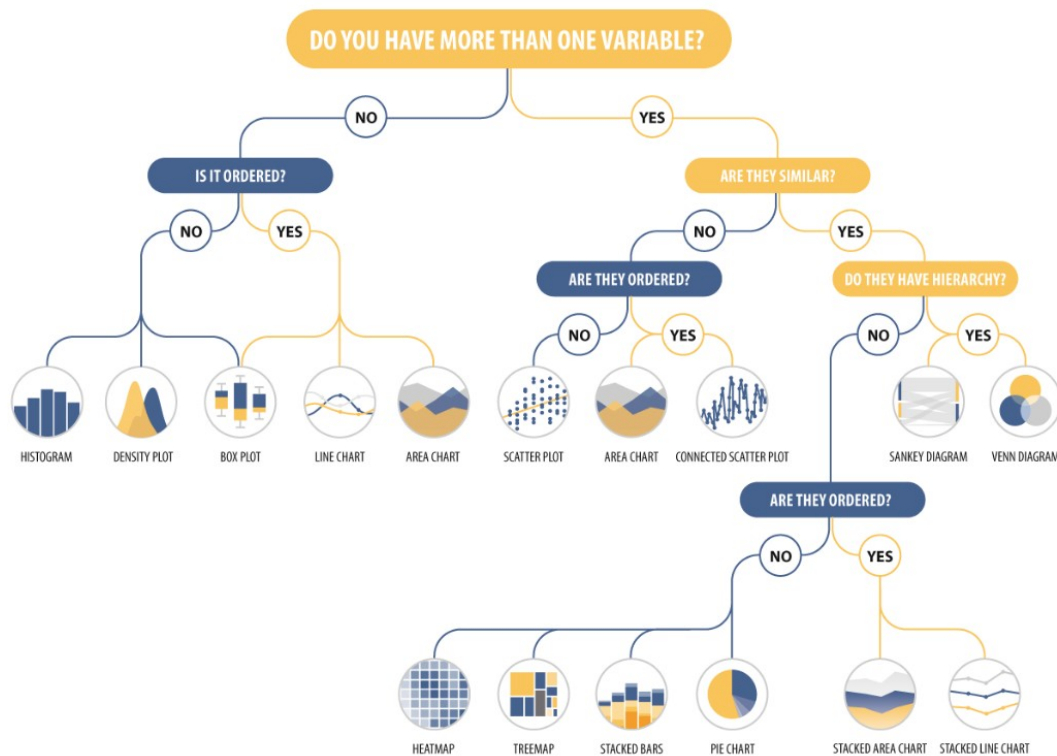
You are guiding the reader nicely through your visualizations by elaborating on them throughout the analysis process, awesome! However, it would have been better to introduce the comment of the visual before the code cell rather than after

"Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted. This includes choice of appropriate plot type, data encodings, transformations, and formatting (title, axis-labels) as needed.

Tip: Do not overplot or incorrectly plot ordinal data."

You are using descriptive names which is awesome and comments in your code explaining what is happening, so overall readability wise a very good job!

Below is a guide to help you choose the correct chart based on the question to answer and the variables under consideration.



Explanatory Data Analysis

The README.md must include a summary of main findings that reflects on the steps taken during the data exploration. It should also describes the key insights that are conveyed by the explanatory presentation.

Tip: The README.md summary is based on the exploration report (Part I notebook) and will guide your explanatory slide deck (Part II notebook) .

Well done, your readme provides a very neat overview of your analysis! To learn more about the importance of readme.md files in repositories, you can read this [article](#)

- A slideshow (HTML file) is provided, with at least 3 visualizations, to convey key insights. Only selective plots are added to the slideshow from the exploratory analysis

selective plots are added to the slideshow from the exploratory analysis.

- The total number of visualizations in the slideshow is less than 50% of the number of visualizations in the exploratory analysis. For example, if the exploratory analysis (Part I) has 18 visualizations, the slideshow can have (3 - 8) visualizations.
- The key insights in the slideshow match those documented in the README.md summary.
- Each visualization in the slideshow is associated with comments that accurately depict their purpose and observation.

Tip: For Part II notebook, use the `jupyter nbconvert` command to generate the HTML slide show.

Well done. Jupyter notebook slides offer a simple, clear layout and are incredibly easy to create. While they do not offer the amount of formatting and design features as other presentation applications, they do a very good job of presenting code and data visualizations for technical audiences

When it comes to most skills, practice usually makes perfect. The more we do it, the better we become at it. Easy, right? Well, with writing, this isn't always so simple. Transcribing your thoughts onto paper quickly and efficiently is pretty challenging for anyone, no matter how often they write.

All plots in the slideshow are appropriate, meaning the plot type, encodings, and transformations are suitable to the underlying data.

All plots in the slideshow are polished, meaning all plots have a title with labeled axes and legends. Labels include units as needed. In other words, each plot must have - chart title, x/y axis label (with units), x/y ticks, and legend.

Great job, The visualizations chosen and the key insights conveyed are well-connected to the findings from the exploration.

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)