



Recap

Recap

In this lesson:

1. You motivated the need for data visualization by showing that summary statistics don't tell the full story. You saw datasets where the summary statistics were the same, but the actual data were very different!
2. You did a review of data types. In general there are quantitative and categorical variables. Quantitative variables can be either discrete or continuous, while categorical variables are either ordinal or nominal.
3. You looked at univariate plots. In most cases a **histogram** should be used for quantitative data, while a **bar chart** should be used for categorical data. There are some cases where you might use one of the other plots.
4. You then looked at bivariate plots, where you were comparing two variables to one another. **Scatter plots** are the most common way to visualize two quantitative variables, while a **line chart** is common for data that you are watching over time. If you are comparing two categorical variables, the best choice is probably a **side-by-side bar chart**.
5. You learned about **correlation coefficients**, which provide the strength and direction of linear relationships. You learned a rule of thumb for determining whether the relationship between two quantitative variables is **strong**, **moderate**, or **weak**.
6. You then looked at cases where we had more than two variables. You learned that using these plots effectively is about building the plot that helps you see the insight that answers the question you have.
7. You gained some insight into visual encodings and data dashboards, which will be a part of the next lessons!



Recap
