PROCEDURE (STEP BY STEP)

We start by cleaning our data. Tasks during this section include:

- Drop NaN values from DataFrame

- Removing rows based on a condition

- Change the type of columns (to_numeric, to_datetime, astype)

Once we have cleaned up our data a bit, we move to the data exploration section. In this section we explore 5 business questions related to our data:

- What was the best month for sales? How much was earned that month?

- What city sold the most product?

- What time should we display advertisemens to maximize the likelihood of customer's buying product?

- What products are most often sold together?

- What product sold the most? Why do you think it sold the most?

To answer these questions we walk through many different pandas & matplotlib methods. They include:

- Concatenating multiple csvs together to create a new DataFrame (pd.concat)

- Adding columns

- Parsing cells as strings to make new columns (.str)

- Using the .apply() method

- Using groupby to perform aggregate analysis

- Plotting bar charts and lines graphs to visualize our results

- Labeling our graphs