

- There are two datasets in the data folder.
 - Data source #1: **Denver Hotels Performance**. This CSV provides Denver hotels' performance data, such as Popularity, Response rate, Room score, etc.
 - Data source #2: **Denver Hotels Info.csv**. This CSV provided Denver hotels' information such as Location, Brand, Room rate, Price Type, etc.
- Before diving deep into analysis, please note that Data source #1 and Data source #2 are provided by two independent data vendor companies. In other words, some hotels from data source #1 may or may not have an identical name in data source #2, and vice versa. To give a clear example, the same hotel in can be listed as '*Ac Hotel New Orleans Bourbon/French Quarter Area*' on Expedia.com, but '*Ac Hotel by Marriott New Orleans French Quarter*' on TripAdvisor. This is a very common issue when dealing with hospitality data set, and a simple join on strings will miss exclude many hotels' information. **Be careful when you merge two CSVs.**
- Both CSVs have manual typos and missing values. Try to identify/correct those errors first and then execute your EDA.
- Please run/save/submit your code on **Jupyter notebook** or **Google Colab**. You may use any available Python libraries to help you finish this test, such as Numpy, Pandas, Matplotlib, Seaborn, Scikit Learn, etc.
- Potential directions:
 - Is there any relationship between Hotel Overall Score and Location?
 - Is there any relationship between Popularity and Price?
 - How do you rank all the unique hotels in different segments? Any reasoning to back the ranking algorithm you are using?
- Things to consider:
 - What correlation techniques are you using and why?
 - Which features are the most/least meaningful when explaining a property's performance data?
 - How can you visualize your findings in a meaningful way?
- Please report any hypotheses/assumptions/insights/conclusions you make or find.