DTSC-580: Data Manipulation

Assignment: Lego

Directions

In this assignment, you will load various tables that contain information about Lego from the Rebrickable website and perform analysis on this data. First, download the notebook and `csv` files, ensuring that the `csv` files are in the same folder as your notebook.

Your first task will be to go through the data, and review the schema as shown in the notebook to understand how the tables all fit together. There is no data dictionary that I am aware of, but you should be able to understand this data by reviewing the files and schema and visiting Rebrickable's Download page and API documentation. Note that you must use the files downloaded from Brightspace for this assignment and not download new files from the Rebrickable website.

When you have answered all the questions in the notebook, save your notebook as `lego.ipynb` and submit it to CodeGrade for automatic grading.

Show Work

Remember that you must show your work. Students submissions are spot checked manually throughout the term to verify that they are not hard coding the answer from looking only in the file or in CodeGrade's expected output. If this is seen, the student's answer will be manually marked wrong and their grade will be changed to reflect this.

For example, if the answer to Q1, the mean of a specific column, is 22:

```
# correct way

Q1 = df['column_name'].mean()

# incorrect way

Q1 = 22
```

Below is a simple rubric indicating how you will be graded:

Correct Answers 100%
Are your answers to the questions correct? Did you set up the requested functions correctly and they return correct outputs?