**Extract**: The first set of data came from an existing pandas file. I exported the dataframe to a csv, imported that csv into another pandas file, then clipped off the select columns and rows that I wanted to compare to the new data. The new set of data came from data.world. It was a list of Superbowl winners, their regular season win percentage, their opponents, and the opponents regular season win percentage, and it was a csv file. Data.world happened to have a built in SQL query capability, so I simply isolated the years I wanted to keep using the built in query tool, then downloaded the edited csv to my local machine. Once it was on the local device, I pulled it into my pandas file and began changing the data around.

**Transform**: The simplest transformation process was to cut off the columns that were unnecessary for this analysis. After that, some renaming of column values took place to try to standardize the two dataframes to make it easier to reference one another. Specifically, one dataframe had team names listed as abbreviations while the other had the full, written out team names. We used a replace function to make those changes in pandas.

**Load**: The final report consists of python dataframes that were exported to SQL for querying. The goal was to create a query that could be applied to each year, find the Superbowl Champion’s cover percentage against the spread, and use that information to try to find a correlation (i.e. a team that had a winning record but had a winning record against the spread might be an indication that the team was undervalued by oddsmakers and, thus, would have yielded a high payout to win the Superbowl)