ETL Project Report

French Tourism 2017:

* Extracted :
  + data from site into excel
* Transformed:
  + Read excel file into jupyter notebook
  + Created six dataframes from excel sheets
  + Cleaning:
    - Renaming columns, new indexes, dropping rows
  + Created tables in pgAdmin
  + Exported dataframe to pgAdmin tables
  + Created clean csv files
* Load:
  + Created database in pgAdmin
  + loaded data from dataframes into pdAgmin tables
* Challenges:
  + Translating from French to English
  + Finding public data
  + Calling different sheets from excel
  + Matching column names in pgAdmin table to DF table
  + Matching primary key to id column in dataframe