**ETL Project Proposal: Federal Budget/Deficit vs Marginal Income Tax Rate**

Team: Erin Mandell, Susan Roedl

**Data set 1**: US Federal Budget Deficit / Surplus by Year since 1940

<https://www.whitehouse.gov/omb/historical-tables/>

[Table 1.3—Summary of Receipts, Outlays, and Surpluses or Deficits (-) in Current Dollars, Constant (FY 2012) Dollars, and as Percentages of GDP: 1940–2024](https://www.whitehouse.gov/wp-content/uploads/2019/03/hist01z3-fy2020.xlsx)

**Data set 2:** IRS / Historical Lowest and Highest Income Tax Brackets

<https://www.irs.gov/statistics/soi-tax-stats-historical-table-23>

Load both excel spreadsheets directly into the Jupyter notebooks.

Drop several columns from each excel file, then strip extraneous notations off of the Tax Rate table.

Load both tables into a SQL database, and set up the schemas in pgAdmin.

Join tables on the year.