**Group 5 - ETL Project Proposal**

**Peter Anderson, Casandra Gormsen, Jennifer Gabrysiak, Zach Liebler**

**Wages vs. Gross Rent in the United States**

**Overview:** We’d like to perform the ETL process on two datasets, combining Wage data and Gross Rent data in the United States, by zipcode.

**Process:**

* Use Pandas in Jupyter Notebook to load in CSV, remove unnecessary information, and create copy
* Use SQLAlchemy’s Create\_Engine Function to load cleaned DataFrames into database
* Use Pgadmin4 to create database for loading
* Create SQL query to join both tables

**Datasets:** Sourced from Kaggle

* [United States Wages Dataset](https://www.kaggle.com/pavansanagapati/us-wages-via-zipcode)
* [United States Gross Rent Dataset](https://www.kaggle.com/goldenoakresearch/acs-gross-rent-us-statistics?select=Gross+Rent+Database+Kaggle+Document.pdf)