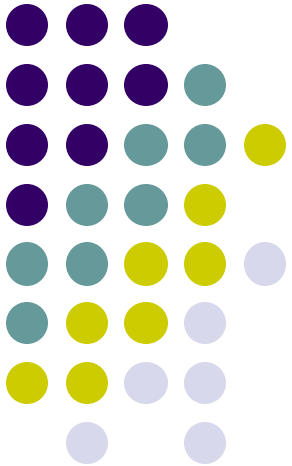
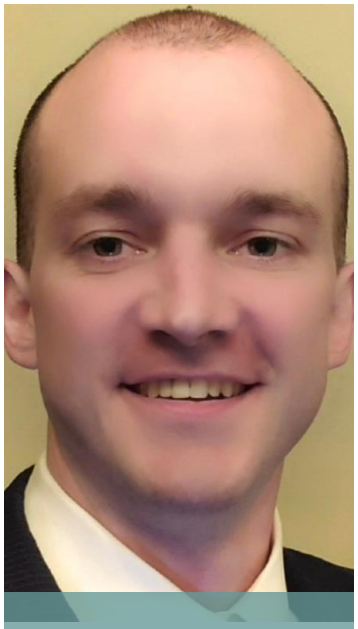
Analyzing Retail Sales through Product and Industry Classifications

Group 6



The Team



Nathaniel Van Schyndel

Dev-10 Associate

James Miller

Dev-10 Associate

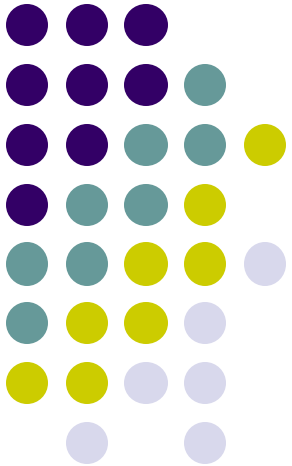
Gavan VanOver

Dev-10 Associate

Cole Ballard

Dev-10 Associate

Introduction



Retail Services

Retail trade “includes establishments engaged in selling merchandise in small quantities to the general public, without transformation, and rendering services incidental to the sale of merchandise.” (U.S. Census Bureau)

Initial Questions

- How have sales changed over time?
  - What kind of trends can be observed over time?

- Are there relationships between macroeconomic principles and retail sales?
  - Specifically, Consumer Price Index (CPI), RPI, the number of retail employees (USTRADE), and the number of wholesale employees (USWTRADE).

Datasets

Monthly Retail Trade Report (accessed on August 2, 2022)

- Monthly Census data from 1992 to 2020 listing the total retail sales broken down by NAICS codes

Consumer Price Index (CPI) United States (accessed on June 29, 2022)

- Provides the monthly Consumer Price Index for the US since the beginning of 1960

Macroeconomics US (accessed on June 29, 2022)

- The original dataset provides over 258 different macroeconomic measures for each month since the beginning of 1960

All Sectors: Products by Industry for the U.S. (accessed on June 29, 2022)

- 2017 Census data broken down by both product codes and industry codes. Provides different measures such as the number of establishments and the amount of sales revenue

What We Want to Know

- Can we use historical retail sales data to predict future trends in retail sales?

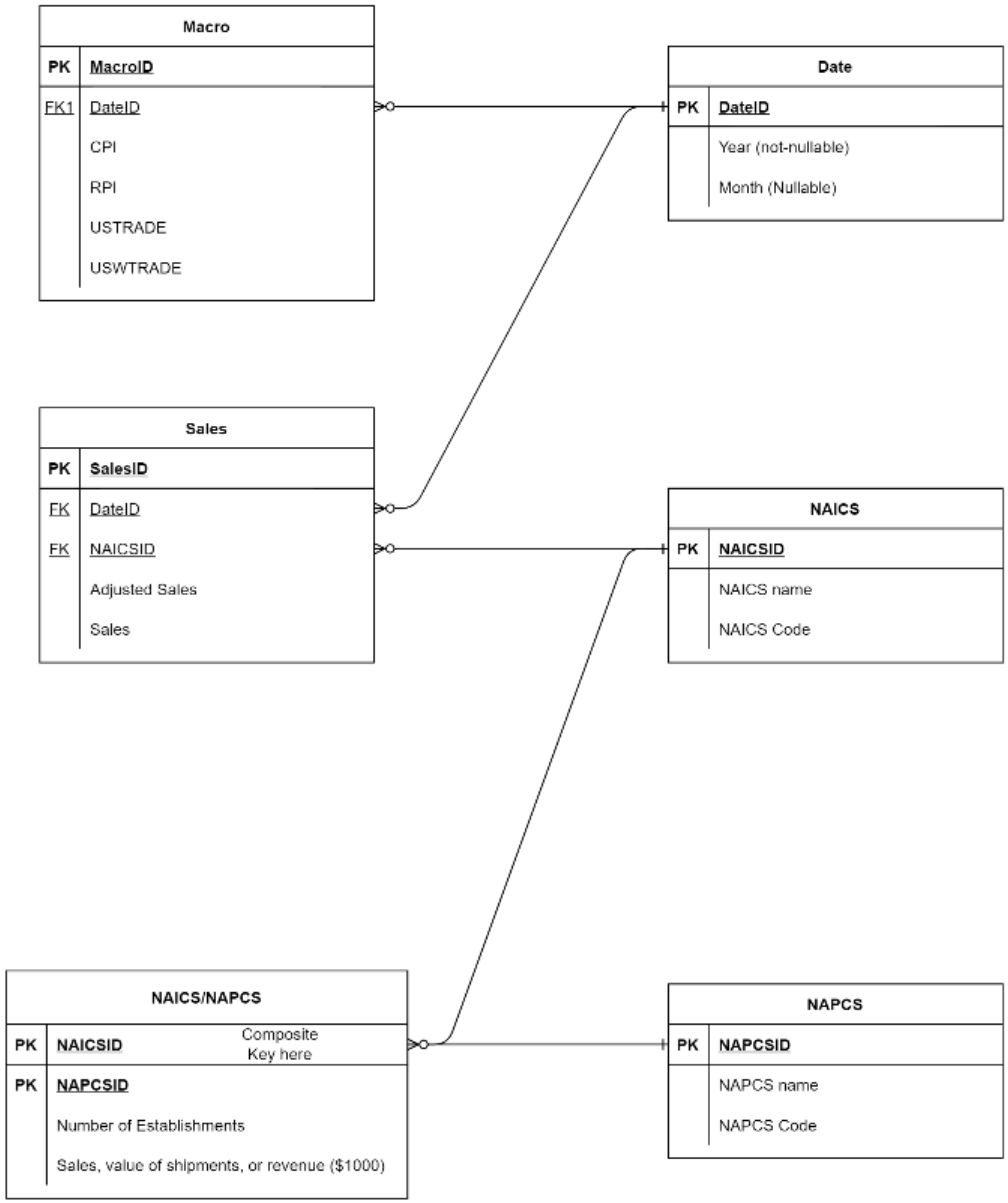
- What industries have the highest number of sales?
  - Which products are sold the most within those industries?

Next Steps

Extract, Transform, and Load desired data from individual datasets into an SQL database

- Create a pipeline in the cloud to automate ETL for non-static data.
  - non-static datasets
- For static data, complete an initial ETL.
  - static datasets

Extract and Transform



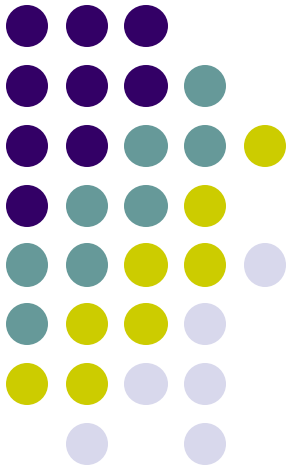
All the data was coming from Excel workbooks or CSV's.

We fit the datasets into 6 tables in 2nd normal form and used several foreign keys to establish relationships between data.

Load

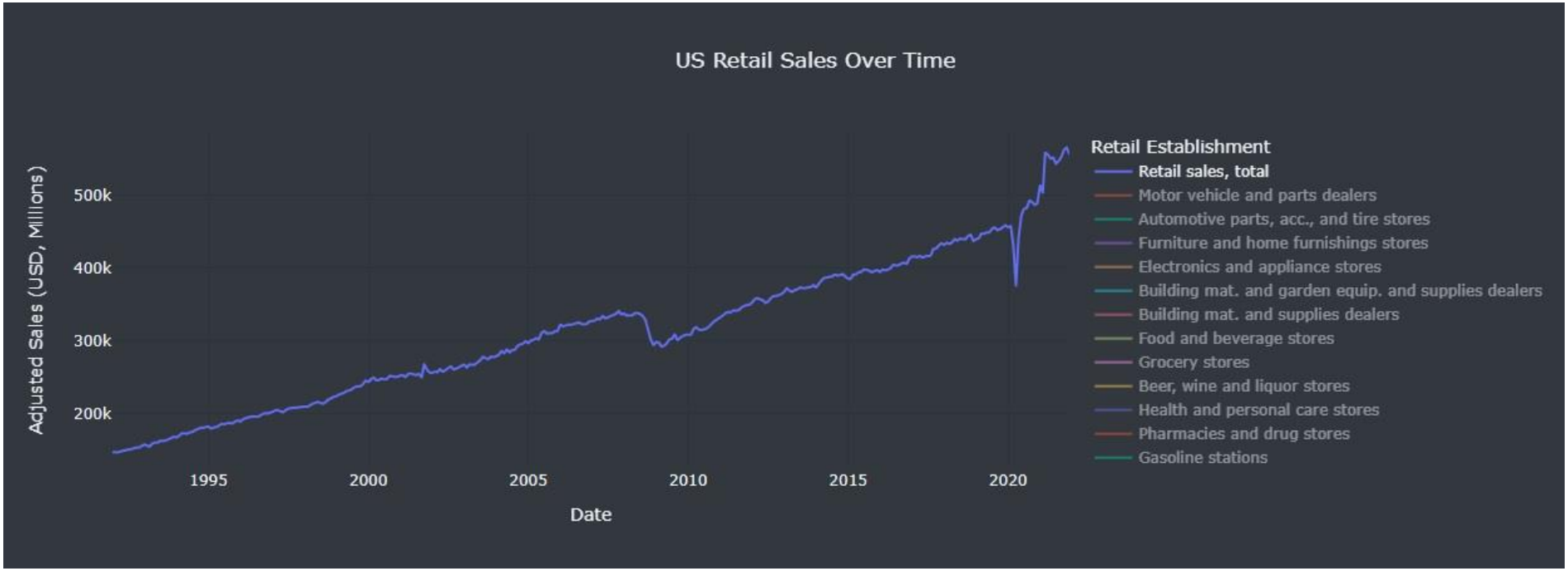
We then pushed the tables into a MSSQL database where it could be accessed by our dashboard.

Results

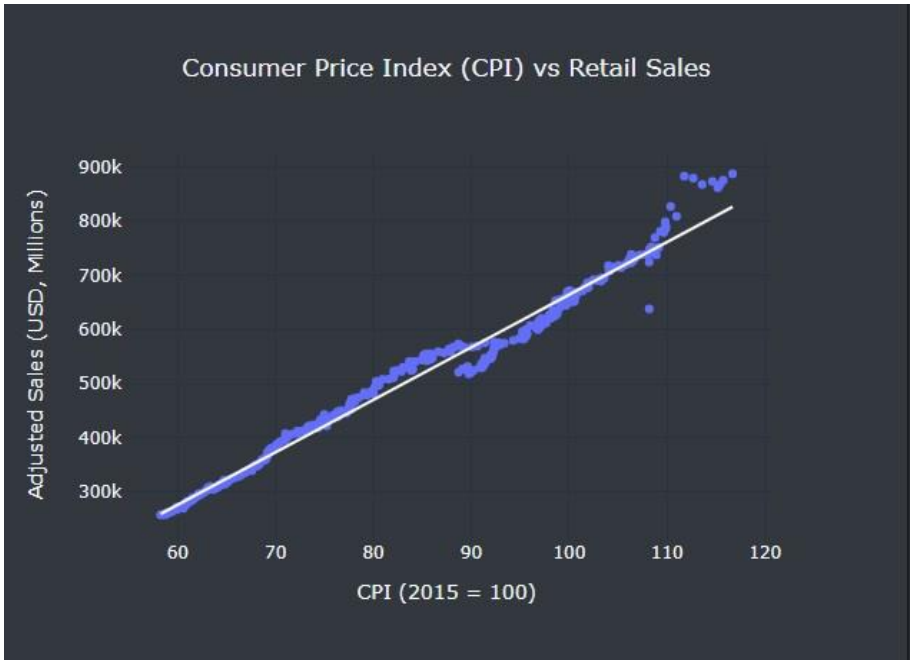


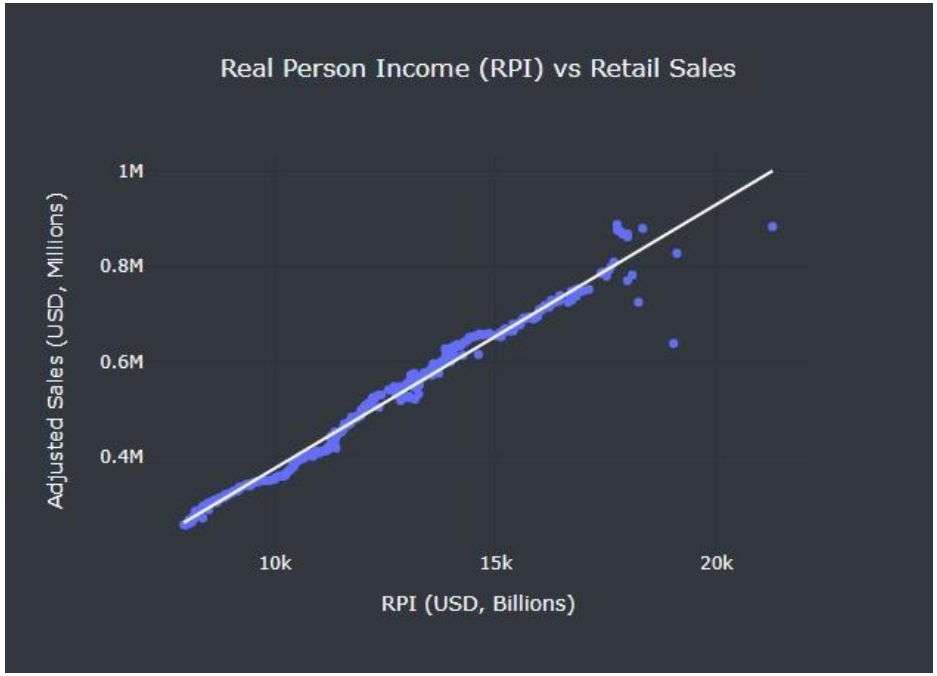
Dashboard: https://us-retail-sales-dashboard.herokuapp.com/

Retail Sales Over Time



CPI & RPI vs Retail Sales



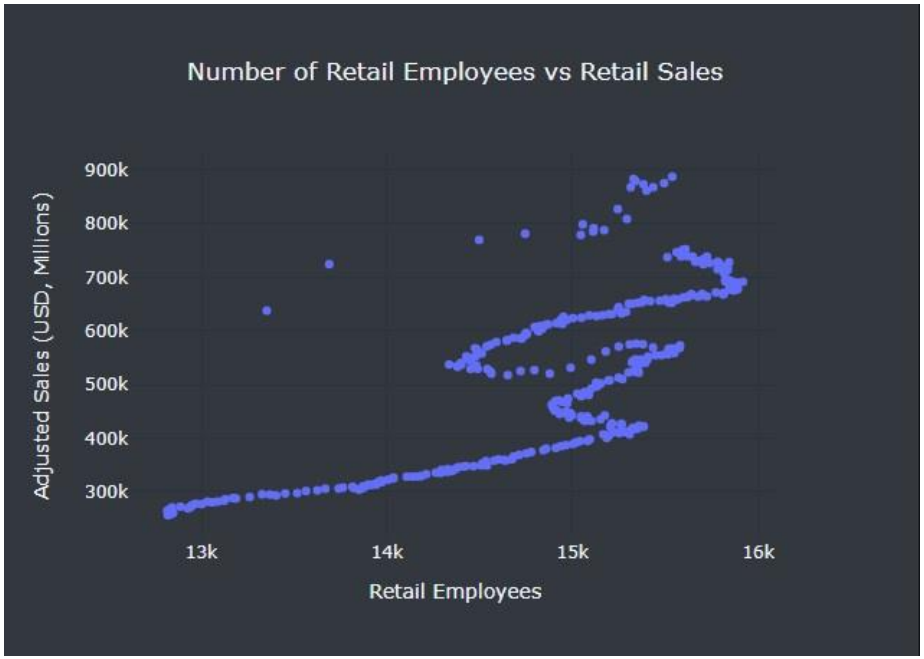


Correlation Between Sales, CPI, and RPI

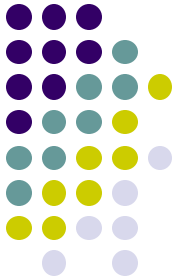
Confirming the positive linear relationships

	Retail sales, total	CPI	RPI
Retail sales, total	1		
CPI	0.980778512	1	
RPI	0.987666331	0.982462	1

Number of Employees vs Retail Sales

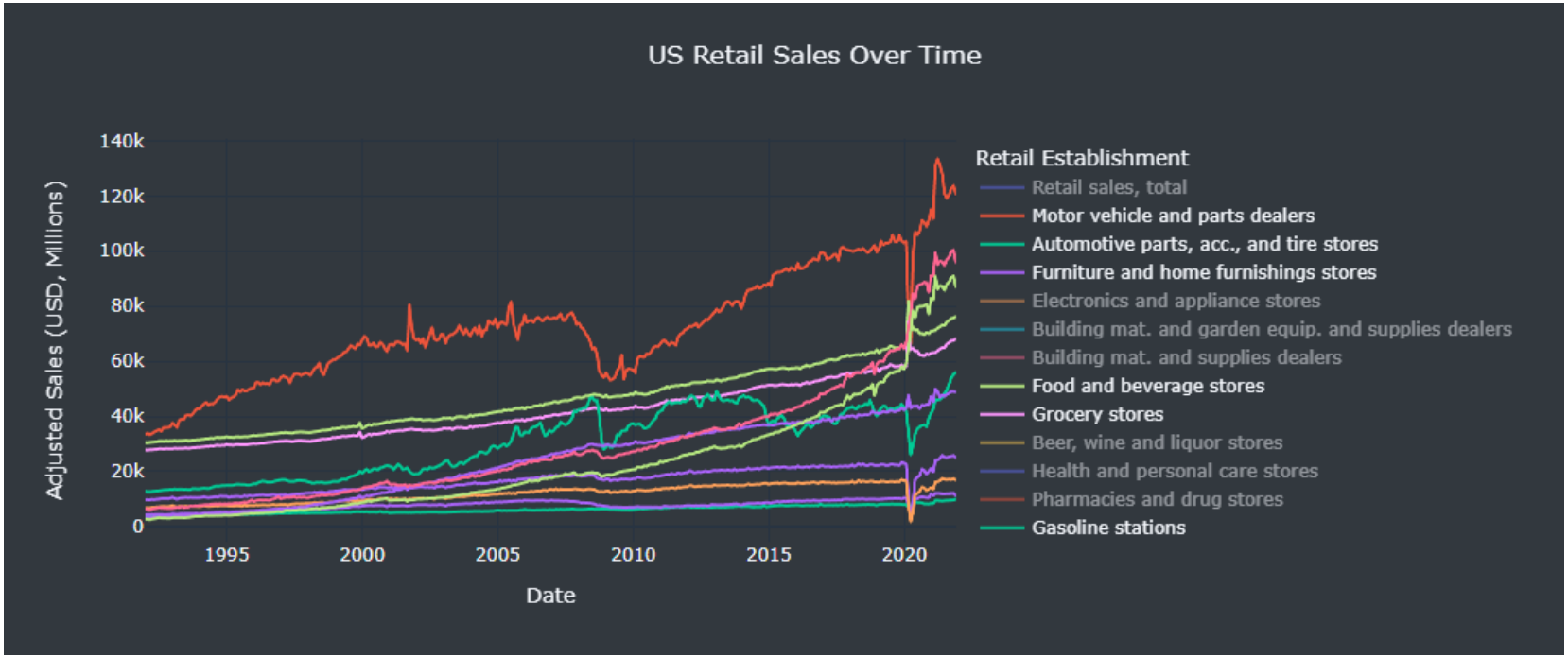


Correlation Between Sales, USTRADE, and USWTRADE



	Retail sales, total	USWTRADE	USTRADE
Retail sales, total	1		
USWTRADE	0.546972217	1	
USTRADE	0.740087702	0.910832715	1

How Do Sales Compare by industry?



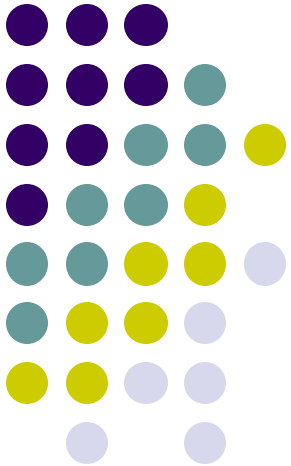
Nonstore retailers

Electronic shopping and mail order houses

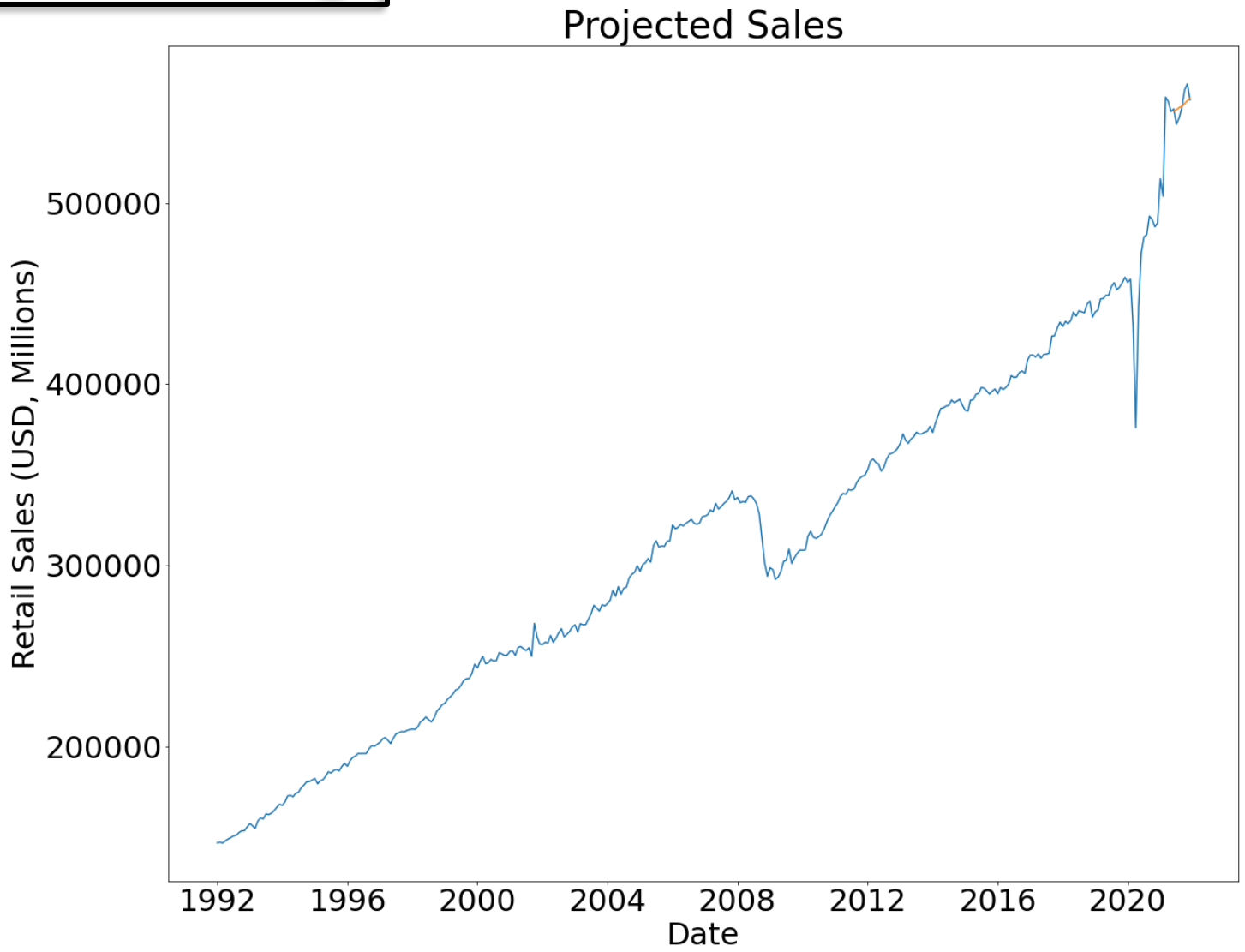
Warehouse clubs and superstores

Top 3 NAPCS in Top 3 NAICS (2017)

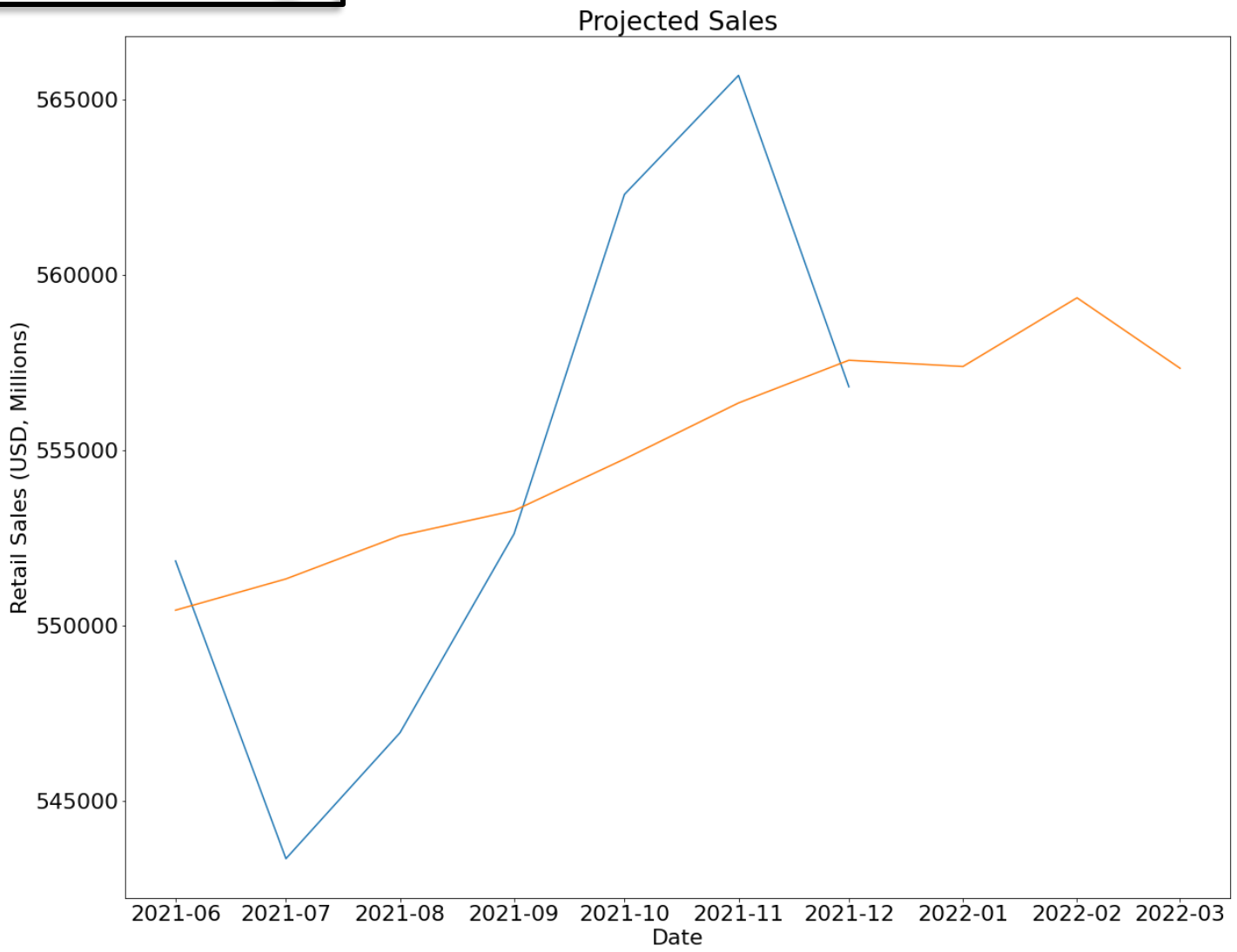
Forecasting



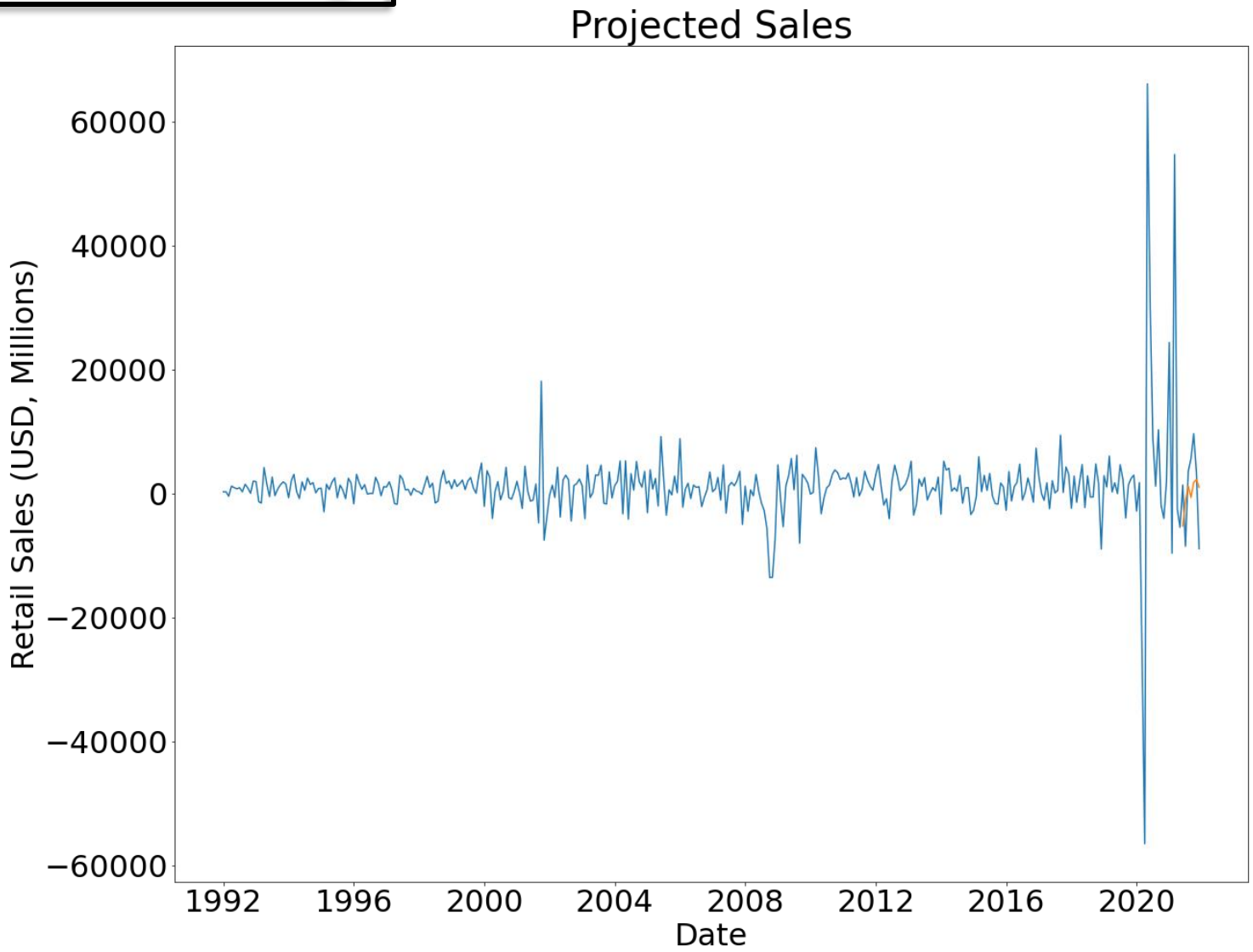
Forecasting



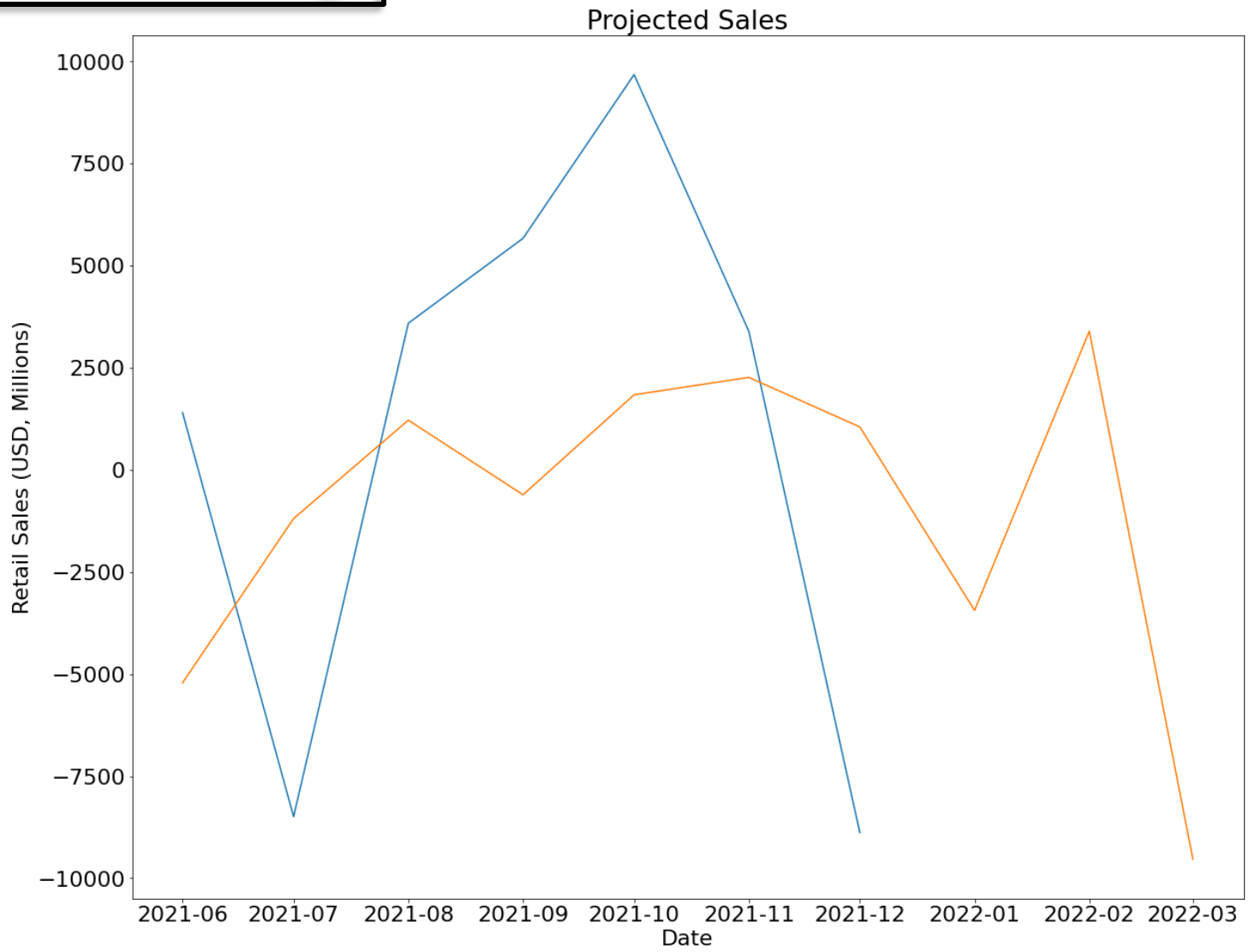
Forecasting



Forecasting



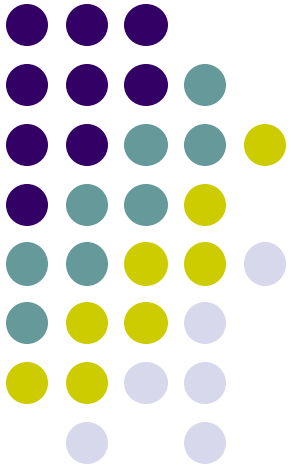
Forecasting



Forecasting

- Due to the volatile and unpredictable nature of retail sales in the United States, it would be irresponsible to forecast more than just a few months out.
- This forecast should continue on a rolling basis, updating as new data becomes available.

Summary



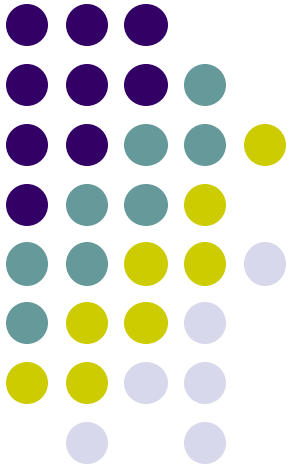
Summary

- CPI and RPI have strong, positive linear relationships with retail sales
- There are observable trends in US retail sales data over time
  - Overall, sales are increasing over time
- The plots depict the volatile nature of retail sales
  - Many peaks and troughs over time

Summary

- Out of the many different industries:
  - Motor vehicles and parts dealers is the largest single industry by a wide margin
  - Non-store retailers and electronic retailers have risen drastically in the last several years to become the 2nd and 3rd largest industries
- Our time series model is moderately useful for predicting the immediate future of retail sales but will need constant updates to remain relevant.

Questions?



Sources

- Bureau, U. S. C. (2017). *All Sectors: Products by Industry for the U.S.: 2017*. Explore census data. Retrieved August 3, 2022, from https://data.census.gov/cedsci/table?q=ECNNAPCSPRD2017.EC1700NAPCSPRDIND&n=N0600.44&tid=ECNNAPCSPRD2017.EC1700NAPCSPRDIND&hidePreview=true
- Avigan, A. (2020, February 16). *Consumer price index (CPI)*. Kaggle. Retrieved August 4, 2022, from https://www.kaggle.com/datasets/aavigan/consumer-price-index-usa-all-items
- Dchaen. (2022, May 11). *Macroeconomics us*. Kaggle. Retrieved August 4, 2022, from https://www.kaggle.com/datasets/denychaen/us-macro?select=US_MACRO110522.csv
- Bureau, U. S. C. (2022, July 15). *Monthly Retail Trade Report*. United States Census Bureau. Retrieved August 10, 2022, from https://www.census.gov/retail/index.html