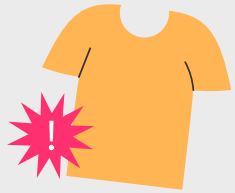
U.S. Department Stores Time Series Analysis

Ayako Homma | May 2023

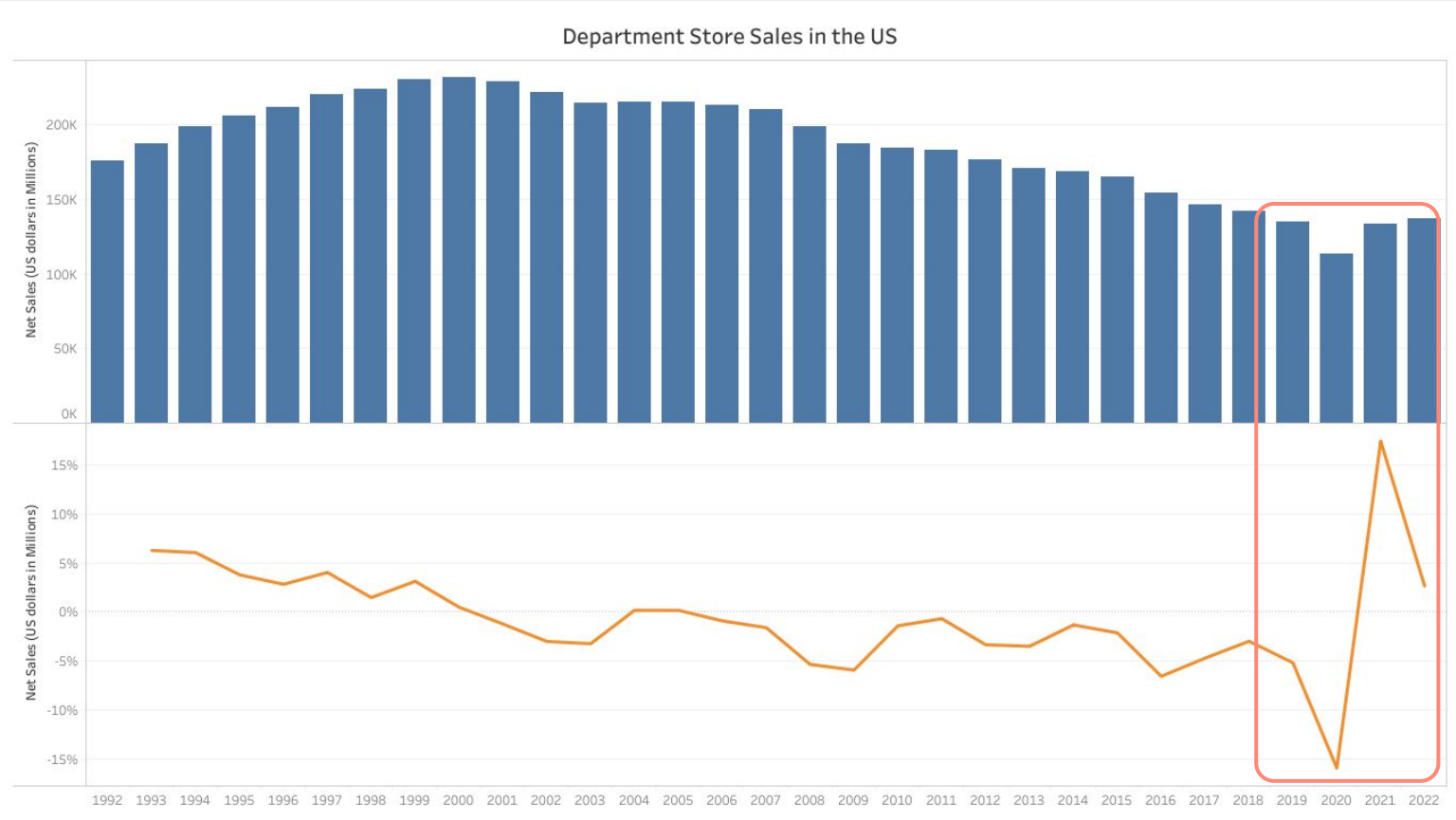
Problem Statement



The objective of this project is to develop a time series analysis and forecasting model that accurately predicts the overall retail channel sales for department stores in the US, using the US Census Monthly Retail Sales data. The model should take into account factors such as seasonality, trends, and holidays, and should be trained and tested on historical sales data to optimize its performance.

The outcome of this project will provide valuable insights into the underlying patterns and trends in the data, allowing department stores to make informed decisions about their sales strategies, inventory management, pricing, and marketing efforts. By accurately forecasting overall retail channel sales, department stores can enhance their profitability and competitiveness in the market.

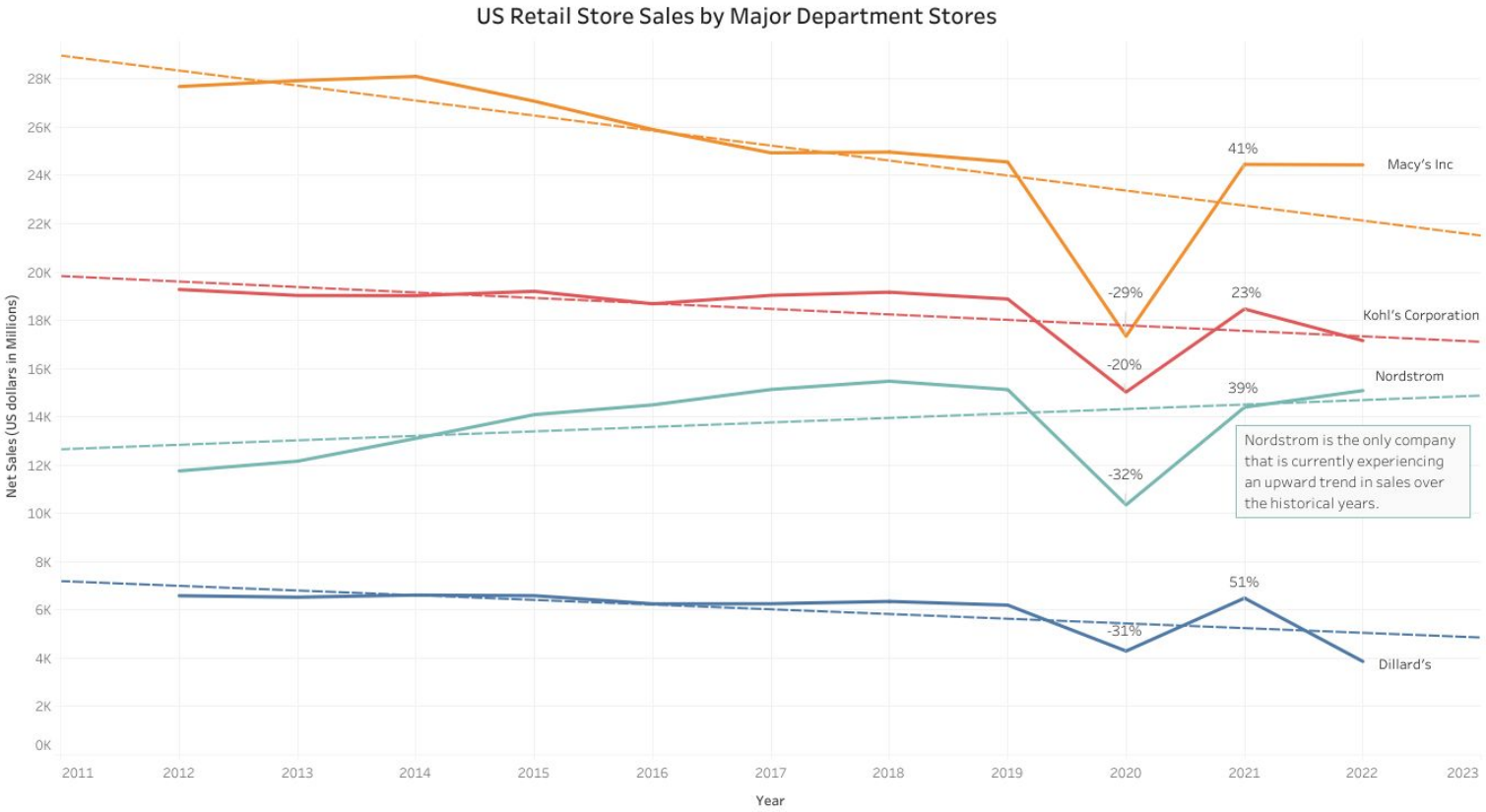
US Department Store Retail Performance



Department store sales saw a decline by 16% in 2020 due to the COVID-19 pandemic. However, after a year, sales quickly recovered and returned to pre-pandemic levels.

Major US Department Stores

Major companies such as Macy's, Kohl's, Nordstrom, and Dillard's saw a double-digit decline in sales in 2020 due to the pandemic. Among these companies, only Nordstrom is currently experiencing an upward trend in sales over the years 2011-2022.



Key Steps in the Research Process

Step 1

Data Collection & Cleaning

Step 2

Time Series Analysis

Step 3

Data Modeling

Step 4

Conclusion & Recommendations

01

Data Collection & Cleaning

Data Collection

US Census

Monthly Retail Trade and Food Services (Survey)

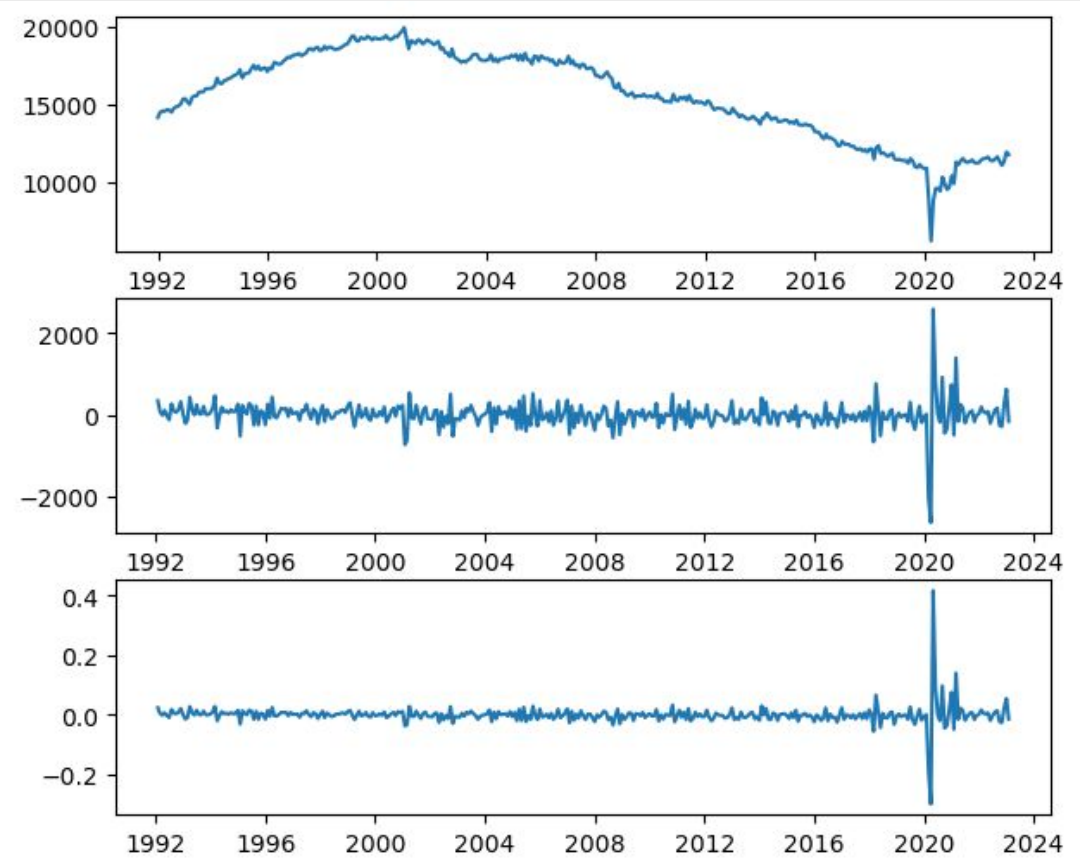
- **Industry or category**: 4521: Department Stores
- **Date range**: 1992.01 - 2023.03*, monthly data
- Seasonally adjusted sales [millions of US dollars]

*For the Monthly Retail Trade survey, all current month estimates are preliminary estimates, which will be superseded in following months by revised estimates.

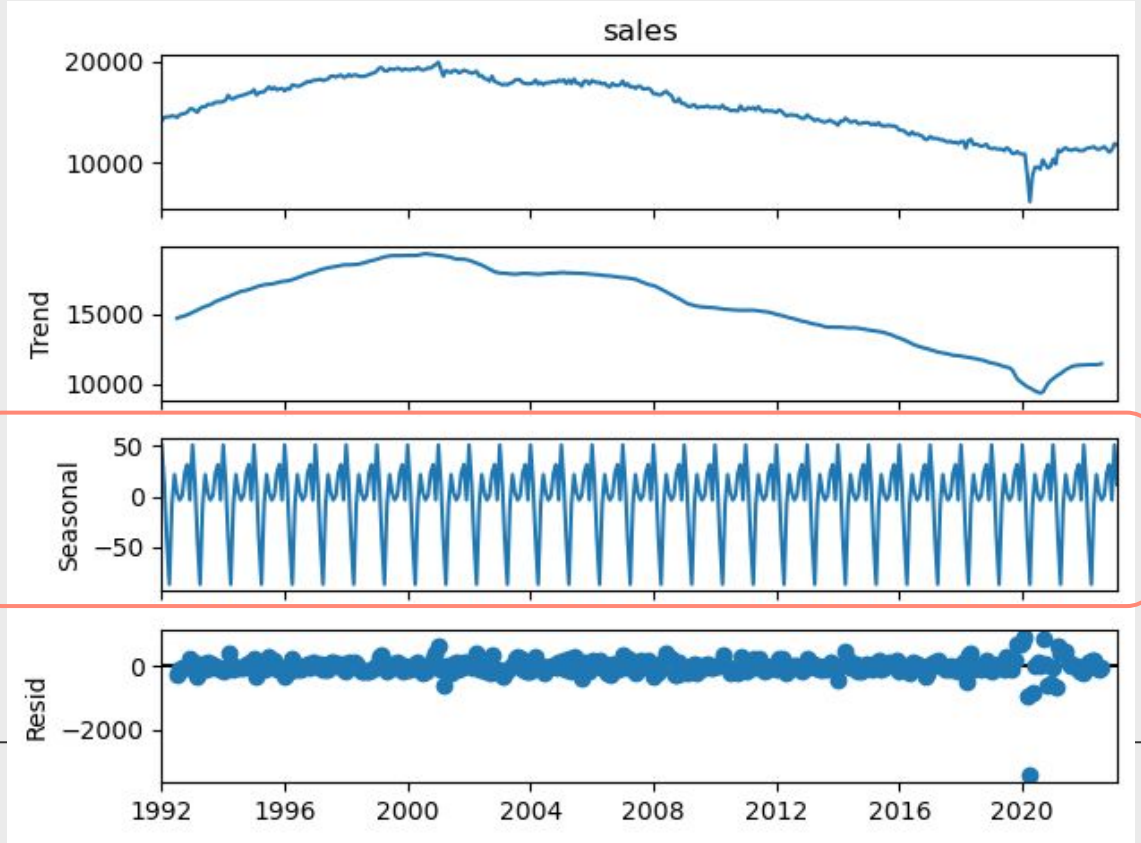
02

Time Series Analysis

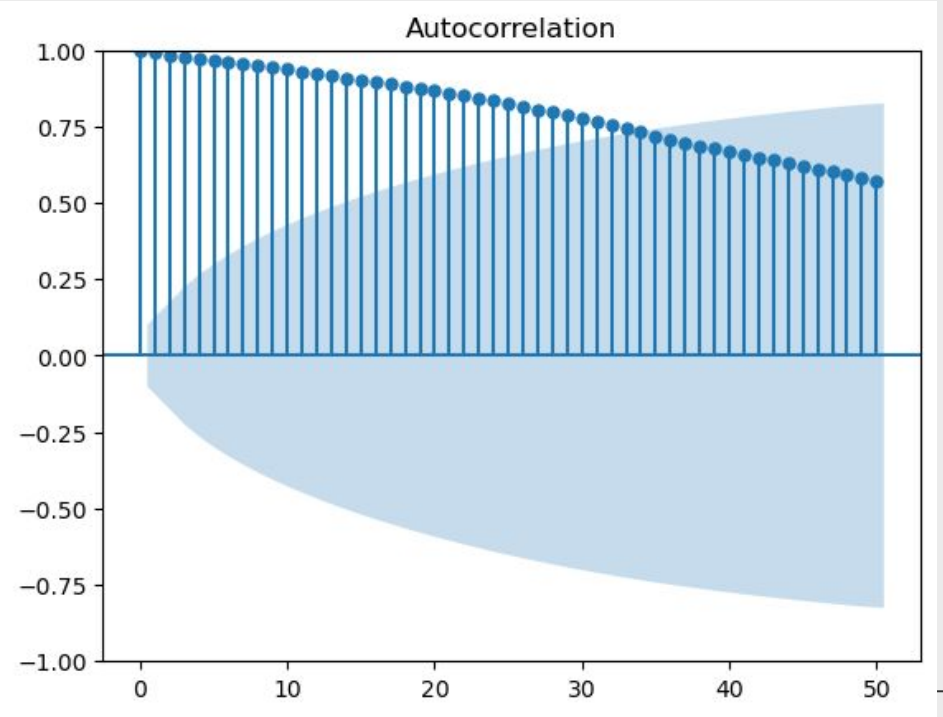
Trends

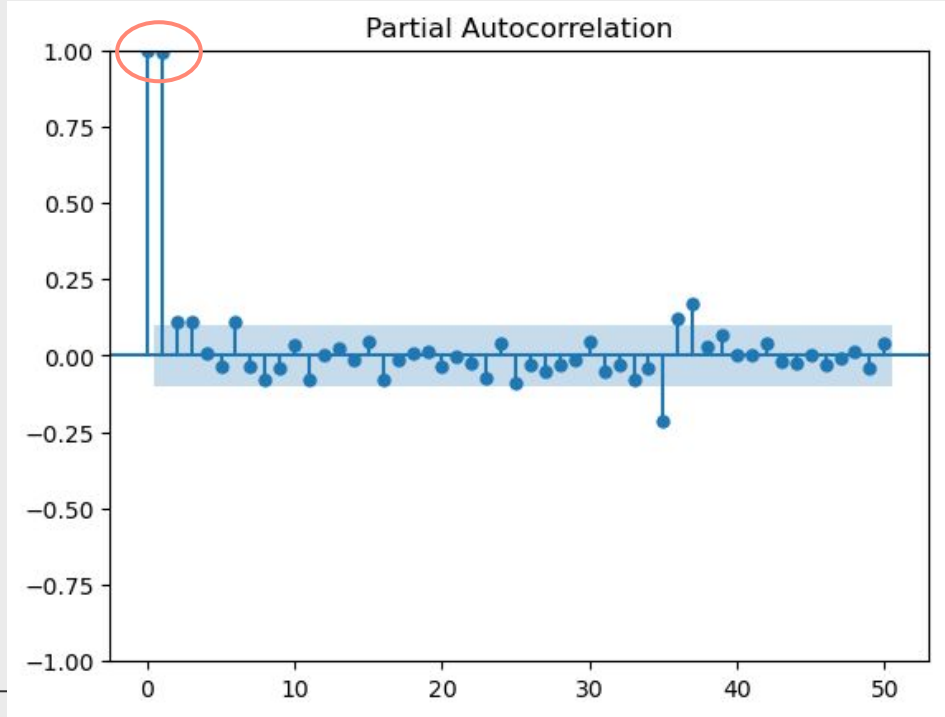

20000
15000
10000
1992
1996
2000
2004
2008
2012
2016
2020
2024
2000
0
−2000
1992
1996
2000
2004
2008
2012
2016
2020
2024
0.4
0.2
0.0
−0.2
1992
1996
2000
2004
2008
2012
2016
2020
2024

Seasonality



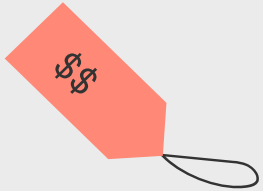
Autocorrelation





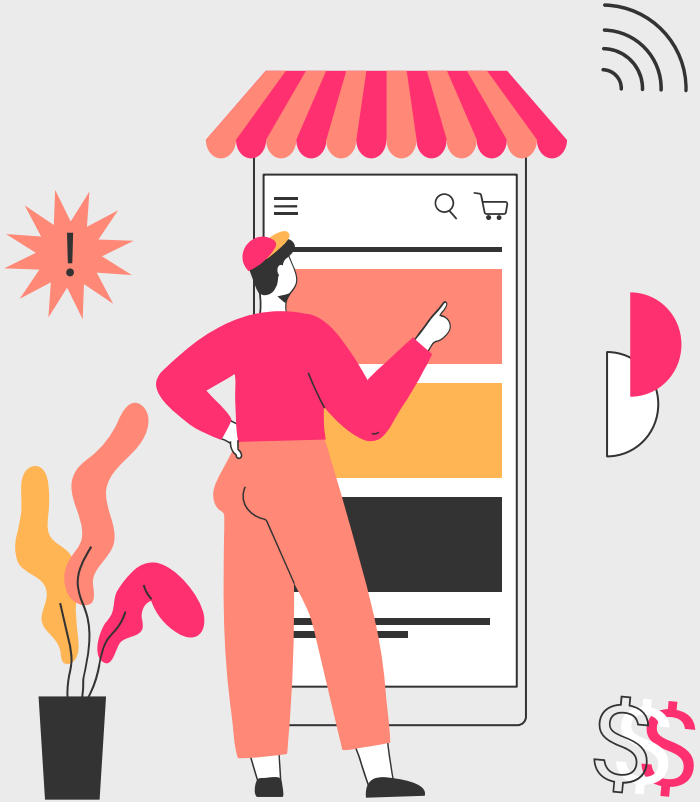
03

Data Modeling



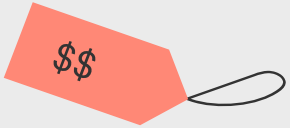
Model Performance & Evaluation

Models	R2	RMSE	MAE	AIC
Linear	0.675	581.987	306.025	966.967
ARIMA (4,1,3)	-0.342	1181.795	966.768	3981.306
SARIMAX (4,1,3)	-1.593	1642.587	1165.682	3876.310



04

Conclusion & Recommendations



Conclusion & Recommendations

The linear model seems to be performing the best out of the three models. The linear model has the highest R2 score and lowest RMSE value and AIC value. ARIMA and SARIMAX had negative R2 scores, meaning that they are performing worse than a model that simply predicts the mean of the target variable. They also have much higher RMSE and AIC values compared to the linear model.

It is recommended to use the linear model for forecasting sales for department stores in the US. However, further analysis should be conducted to ensure the robustness and accuracy of the model. For example, the model could be tested on more recent data to validate its performance and adjust it accordingly. Additionally, other models and techniques could be explored to improve the accuracy of the forecast and provide additional insights for department stores to optimize their sales strategies.

Thanks!

Do you have any questions?