Northeastern University

Final DS 2500 Project:

Covid-19's Impact on Retail Sales

Christopher David and Noah Burra

Contact: burra.no@northeastern.edu and david.c@northeastern.edu

Professor: Rush Sanghrajka

Due Date: 12/3/24

Problem Statement and Background:

COVID-19 took the world by storm in 2019 and peaked in 2020 affecting every individual regardless of status or location. Although serious health and ethical problems occurred, the problem we chose to analyze for this project was the effect of COVID-19 on retail trade sales. What drove us to choose this project topic was the fact that we both witnessed and were part of the drastic shopping habits during as well as after the pandemic. Moreover, we also observed a rise in e-commerce sales and wanted to compare the difference between before and after as it relates to our current majors surrounding economics and behavior. Furthermore, by obtaining results from this project, we believe that it will provoke immense interest in entrepreneurs and business holders as most are still trying to navigate in the post-pandemic world and as well as trying to adapt to ever-changing consumer behaviors. Understanding such changes is crucial not only for the survival of businesses but also for their ability to thrive in a rapidly changing marketplace. Retailers that can change quickly can hold on to customer loyalty and continue to make money whereas companies that fail to innovate often are at a disadvantage. The topic is important to economists and policymakers because retail sales are an important metric in the measurement of economic recovery and stability. Further, the rise of e-commerce sales and its impacts on logistics, supply chain management, and urban infrastructure are wide-ranging. The implications of our findings could better position firms with the resources needed to enhance demand forecasting, improve inventory management, and advance resilience to future disruptions, with greater repercussions for economic resilience and growth.

Hypothesis:

Essential products retail sectors experienced a small decrease in sales during the pandemic compared to non-essential sectors. E-commerce retail sectors remained resilient during the pandemic, experiencing minimal impact, while traditional retail trade sectors faced significant declines in sales.

Introduction to our Data:

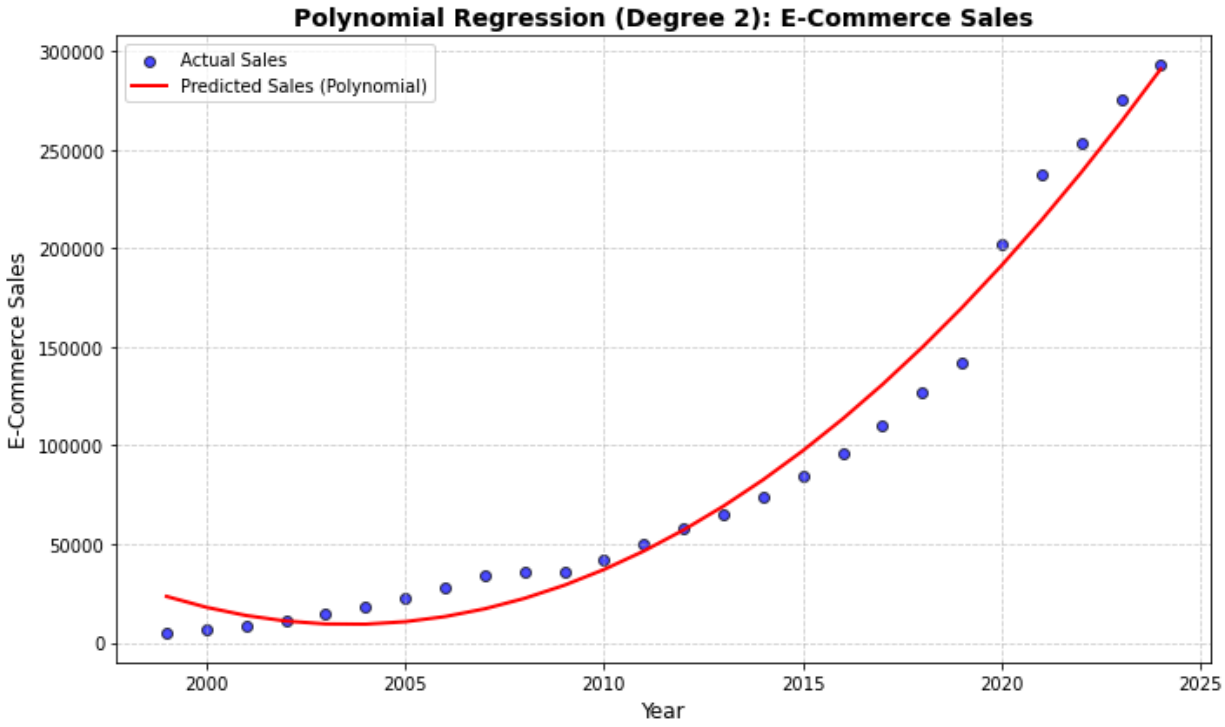
The dataset used in this work brings together several sources that capture the economic and behavioral impacts of COVID-19 on retail trade sales within the United States. Key components involved data from the U.S. Census Bureau covering retail trade sales, including total sales and e-commerce-specific sales, which were based on a survey of retail establishments. These data provided reliable historical and current trends in retail industries. Although these data sources are publicly available and are gathered based on accepted practices, there are several privacy-related ethical considerations, particularly about disaggregated data that inadvertently reveal information about specific businesses or consumers. Furthermore, sources of bias may come from the limitations of the collection itself, biases such as underrepresentation by smaller or unconventional retailers, which could then lead to gaps in understanding the broader impacts on business types.

Data Science Approaches:

In our analysis, we integrated several data science algorithms and techniques to extract insightful information from the dataset. We used Pandas, which cleaned and merged our data in such a manner that the datasets were all well-structured and compatible for analysis. We went ahead to use Matplotlib, combined with NumPy, to create visualizations for our data that would greatly help portray trends and relationships. In addition, we fitted and trained a polynomial regression model using sklearn.linear_model, trying to find a way to make good predictions of outcomes from established patterns. We also did Pearson's correlation to compute the coefficients for the strength and direction of relationships between key variables. These methods provided a structured approach toward the analysis of the dataset to elicit actionable insights.

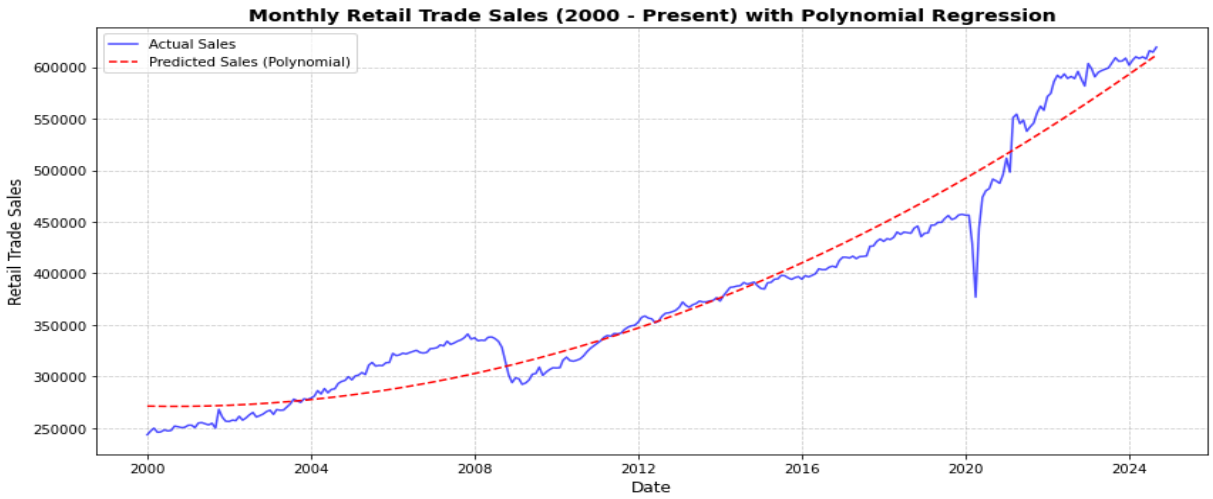
Results/ Analysis + Visualizations:

To either disprove or prove our hypothesis, we decided that we first wanted to conduct a polynomial regression on E-commerce sales during the period between 2000 and 2025. As pasted below:

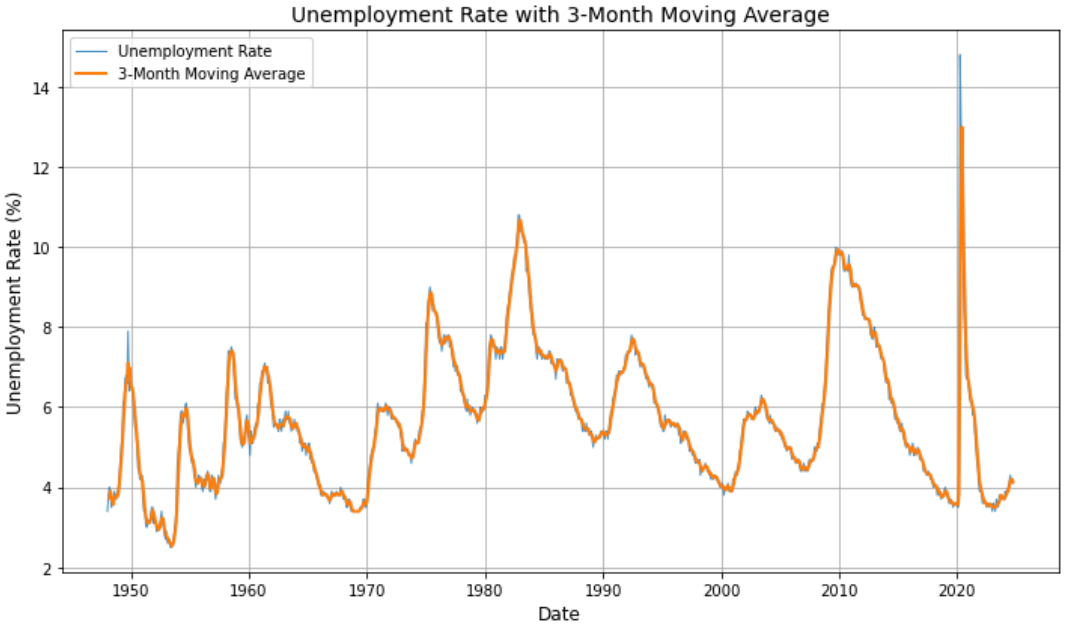


The graph shows us the exponential growth in e-commerce sales, showing no evident decline regardless of the pandemic as well as the huge increase in sales between 2019 and 2021. This provides reasoning in support of our hypothesis that sectors involved in online products experienced rapid recovery and sustained growth compared to other retail sectors during and after the pandemic. Moreover, we can see that up till 2015 the increase in e-commerce was linear. We wanted to research more into this reason and found that it was due to the availability of broadband internet. Additionally, with this new availability, we saw the creation of websites like AMAZON and EBAY that provided an online marketplace for browsers, streamlining shopping and improving efficiency.

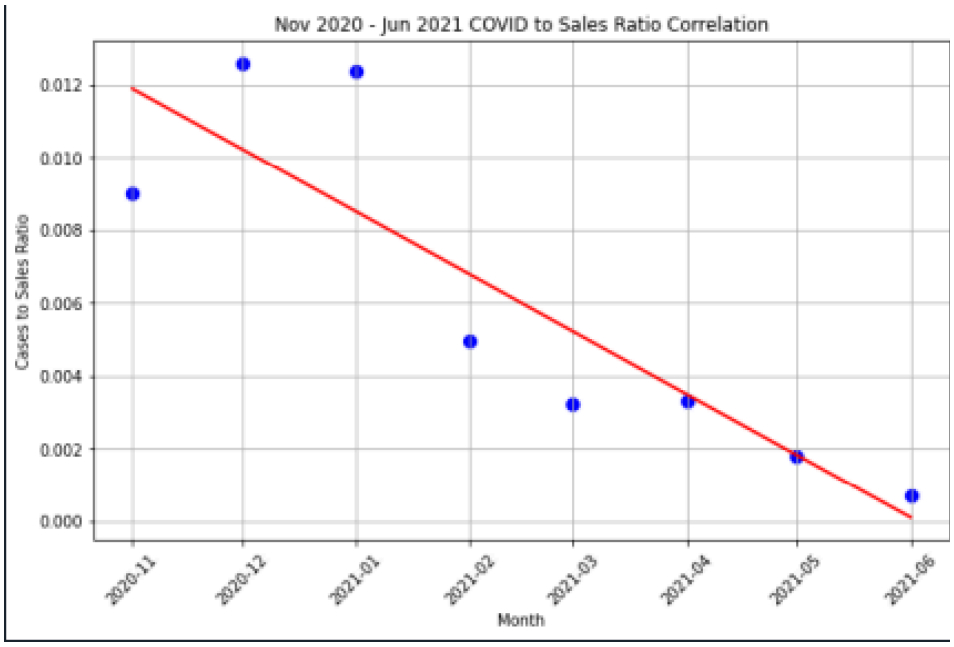
To truly see the effect of COVID-19 on sales we decided to also do a polynomial regression on Retail Trade sales. Shown below:



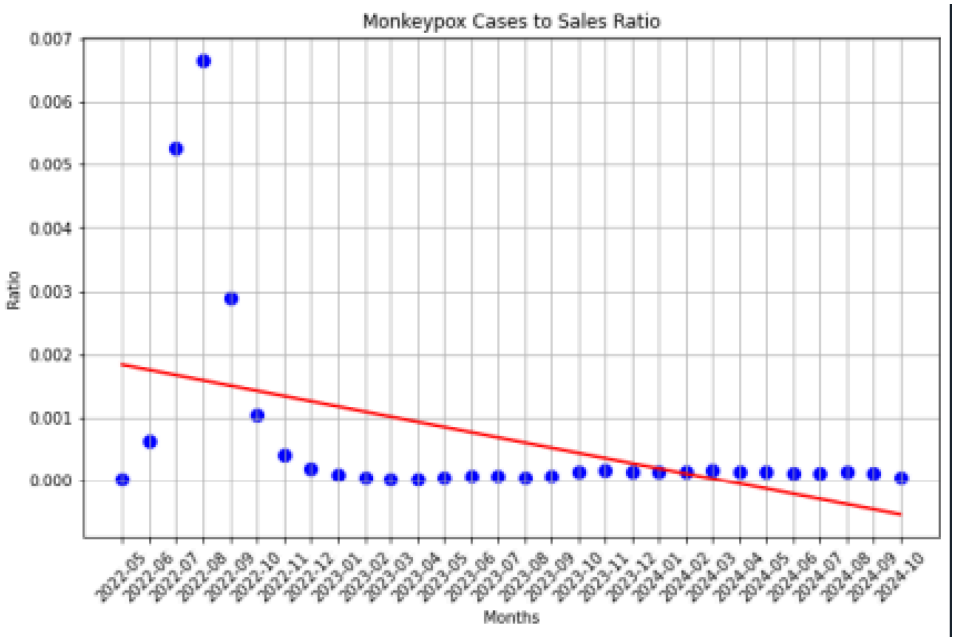
We see that compared to the predicted sales, there was a huge decline in 2020, the peak of the pandemic when Retail Trade sales fell drastically. Regardless of the decline we see that soon after the pandemic Retail trade sales returned to positive and rather grew drastically. At this point, we started to believe that our hypothesis would be disproven and that although ecommerce did gain immense popularity during the pandemic, Retail Trade sales were quick to bounce back and now are higher than they have ever been. After that note, we dove into a deeper topic of consumerism and how human consumer behaviors changed after the pandemic as people were eager to leave their homes and go out and buy items. We also realized that we didn't factor retail sales such as bikes, boats, cars, etc… that grew popular during that period due to social distancing. Despite this oversight, we continued our visualizations and decided to look at unemployment rates and see how they correlated with this data, pasted below:



This data shows a declining unemployment rate until the beginning of the pandemic in 2019. In 2019, the rate rose exponentially from approximately 3% to over 14%. This can be attributed to the increased lockdown restrictions that businesses had to take. However, the main factor is decreased consumer demand, especially in non-e-commerce sales. This caused people to spend less on non-essential items which left many people jobless.



This graph shows a correlation between COVID-19 cases and retail sales. It plots the months as an x and cases to sales ratio as the y. This data shows that there is a steep negative regression line during the peak months of COVID as shown in the graphs from November 2020 to June 2021. What this means is that COVID-19 decreased in personal commerce as people were in lockdown and did not spend money on nonessential goods as much. A similar trend was shown in a more recent pandemic such as the monkeypox virus. The graph for that is shown below:



This graph shows a steep negative regression between the start and present times of the monkeypox virus. Here COVID can be compared to the monkeypox data as it shows that pandemics do affect commerce sales. It also shows that COVID had a larger impact on the retail sales.

Conclusion:

This project studied the effects of COVID-19 on retail sales and considered the resilience and continued growth of e-commerce sectors in contrast to the massive decline and gradual recovery of traditional retail trade. Our analysis underlines the critical role e-commerce has played in maintaining economic activity during disruptions and shows the adaptability of consumer behavior in response to unprecedented challenges. Future work will seek to determine the longer-term impacts of e-commerce growth on brick-and-mortar stores, especially small businesses that may still be recovering. More detailed analysis of particular retail sectors—luxury goods, outdoor equipment, or automotive sales, for example—may yield more insight into which industries have thrived and why. Thirdly, the study might be extended to include data from other countries, thus providing a comparative view of how different regions coped with retail disruptions. Lastly, research on the evolution of post-pandemic customer behavior, especially concerning hybrid shopping models, would provide businesses with helpful insights into innovating and staying competitive in the dynamically changing retail environment.