Adidas Sales Data Marketing Analysis

By Randy Koliha & Isaiah Sarria

Abstract

The point of our research is to find how total sales and sales per capita are influenced by specific demographic features. We tested demographic features like gender, gender preferences, region, states, etc. The findings show that there are gender differences in the volume of sales and seasonality of gender specific sales. Our research showed that there are states that have more sales than other states. When accounting for sales per capita there also were apparent differences in our geological mapping. The differences showed that certain states had significantly more sales per capita than other states. The aim of this project is to show stakeholders of Adidas what they need to know in order to better market to specific demographics, the visualization provided gives an understanding of seasonal trends and preferences. To implement our research, we gathered retailer of Adidas from Kaggle as well as population data from the US Census. After we gathered the data, we put that data into Tableau and simply used drag and drop methods to output visualizations of our choice that best illustrate how the demographic features impact total sales and sales per capita. We also used R to gather some summary statistics to understand broadly what the nature of the data was. The last tool we used for the research is Excel in order to view the data and also create the sales per capita column to make certain visualizations. Despite our interesting findings the data did come with some issues. In the 2020 portion of the data that we gathered; states were missing at high levels. Even in the 2021 there were 4 states missing from the data. However, in the 2021 the states that are present are complete.

Introduction

The Adidas analysis aims to show possible Adidas stakeholders how they should market their products given certain parameters. Specifically, characteristics like region and product times as well as gender and many more. In the project, we took our Adidas data from Kaggle, which was sourced from multiple firms that sell Adidas, and made graphs using Tableau that highlight how different demographic factors influence sales in dollars and sales in units. The graphs we outputted mostly highlighted how these factors affected sales over time. With the contribution of both team members, the outcomes are generally an understanding of how total sales are distributed geographically, how sales are distributed among retailers, and the performance of different product types, male and female time series of total sales, product type time series of total sales, region time series of total sales, sales method time series of total sales, retailer time series of total sales.

1 Goals

1.1 Motivations

We wanted to do this project to find out how different attributes affect sales of particular items throughout the years, and upon finding the dataset this report is on, we found the data to be a good fit for this kind of analysis.

The ultimate goal of this project is to help Adidas, and potentially similar clothing brands, of what kinds of demographic segmentations affect sales, specifically region, gender, purchase type, sales method, retailer preference, and the preference of the kind of apparel that is purchased. With these potential insights, retailers can better understand U.S consumers and their preferences dependent on those mentioned demographic variables to increase company performance. Our mission is to better serve customers with the

clothes they want when they want it by visualizing their purchases.

1.2 Project Target

The key beneficiaries of the results of this project are Adidas's many stakeholders. The firms that sell Adidas may want to know where to market Adidas or when for their customers. Adidas corporate may want to know when and where they should provide their products wholesale. Also alternatively, the brand's competition may want to know how Adidas is doing within certain demographics, etc.

The specific goal of this research is to have a better marketing understanding of what potential demographic factors influence sales. Additionally, we want to learn how to visualize this in a format that the stakeholders and peers can understand simply without data management and analysis knowledge. We also wanted to learn how to visualize this using tableau as our primary

2 Analysis

2.1 Data Source

The data was found from a publicly available website called Kaggle. There is no mentioning of this data being dummy data, as in its only purpose is to be fake data for educational purposes, and the data looks to be very specific, so we assume it is a real dataset that records different Adidas purchases throughout the U.S.

2.2 Summary Statistics

Figure 10 simply explains what the summary statistics are for each feature and whether they are in type character or type numeric.

2.3 Methods

For the research, we utilized drag and drop of features in Tableau to craft particular graphs to answer specific research questions about how demographic features affect total sales. We also utilized R to gather the summary statics and data types. Additionally, we used Excel to add a particular column that calculates state sales per capita. For Tableau, the implementation of the work was dragging the features we wanted to see when comparing different demographics. For instance, we wanted to know how gender played a role in sales over time, so we dragged sales and date in the column and row section, made gender a filter, and selected the line chart graph. We also made sure to put in the date in the filter box to exclude 2020 from the graph because much of the data in that year is incomplete. For the Excel portion, we imported census data for each state's population and did a simple division calculation on sales and population. In doing the R portion, we ran a summary function and gathered the output of the summary statistics. In previous projects, we have utilized R and Python to make visualizations. In this research, Tableau did most of the heavy lifting regarding visualization, whereas R was more code-intensive in gathering data output.

2.4 Results

Figures 1 and 2, both display a map that indicates some sales activity in each state. Figure 2 shows an absolute number of total sales per state, but Figure 1 shows the sales of each state per capita to make more relative comparisons. This has indicated that most of the sales occur in California, but there is a significant market for Adidas products in Wyoming, Alaska, and Vermont, proportionately speaking.

Figures 3 and 4 are helpful for comparing the sales of different men's products and the different women's products and what retailers the two different genders of consumer prefer to shop at. Men seem to be mostly interested in men's street footwear whereas women tend to mostly enjoy women's apparel. Men have a stronger preference for buying shoes at Foot Locker than women and avoid purchasing Adidas goods at Walmart just as much as women. In general, their retailer preferences are very similar.

Figures 5 and 6 dive more into the total dollar sales of men's versus women's products in 2021. Their demand for Adidas in general is the same throughout the year, where there is a spike around July which slows down until it begins to spike back up again during November. It ended at barely an ultimate high in December to then decrease to its lowest point of demand in March. That trend follows for all the product categories, and the only difference is that we can once again distinguish the product types men and women prefer as mentioned before.

Figure 7 shows the trends of sales per region during 2021, and most regions roughly follow the general trend as stated with two exceptions: the

South and the West. The South experiences a much more significant peak around May, while the West experiences a significant peak around September as opposed to the peak around December or July we would normally expect.

Figure 8 shows the trends of sales for the country for each sales method. In-store and online sales follow the expected trend, with the exception of the outlet sales. There seems to be a considerable decrease of shopping at outlets, but interestingly enough the two equal peaks of outlet sales occur before and after the expected July peak (also present in the other two sales methods), and enter the lowest trough during the July peak.

Figure 9 displays the trends for the sales for each retailer. The top three retailers consumers shop at are Foot Locker, Sports Direct, and West Gear, and they follow the same expected trend. Kohl's has a decent following between April and May, after that it falls and remains flat like Amazon and Walmart. Walmart and Amazon are quite flat during the whole year with peaks and troughs that are not even worth mentioning.

3 Conclusion

3.1 Limitations and Future Study

One limitation of the data we gathered was that the 2020 portion of the data was incomplete as far as months recorded. Also, most states did not have data at all this year. We leveraged the 2021 data for the research which had more complete data as far as months of the year. However, in the 2021 portion there were 4 missing states.

In the future we could collect more data on later years or potentially earlier years to see how they compare. We could also collect other demographic information about the people in each state and see if there are differing preferences. Examples of other data we could collect are race, occupation, income levels, cultural preferences, etc.

References

Bureau, US Census. 2023. "State Population Totals and Components of Change: 2020-2023." Census. www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html.

Heemali Chaudhari. 2022. "Adidas Sales Dataset." Kaggle, www.kaggle.com/datasets/heemalichaudhari/adidas-sales-dataset.

A Appendices

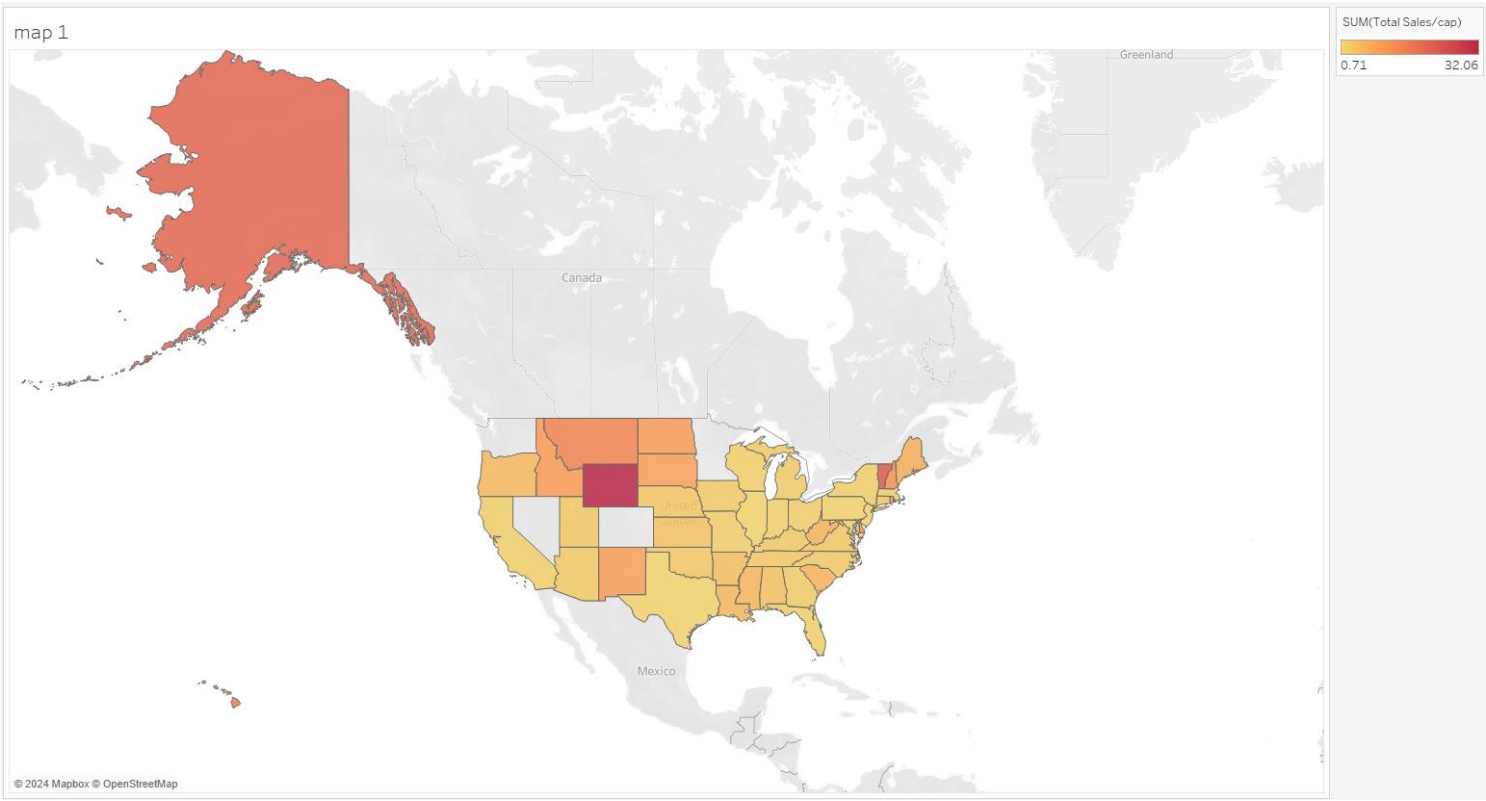

Figure 1: Total sales per capita for each state.

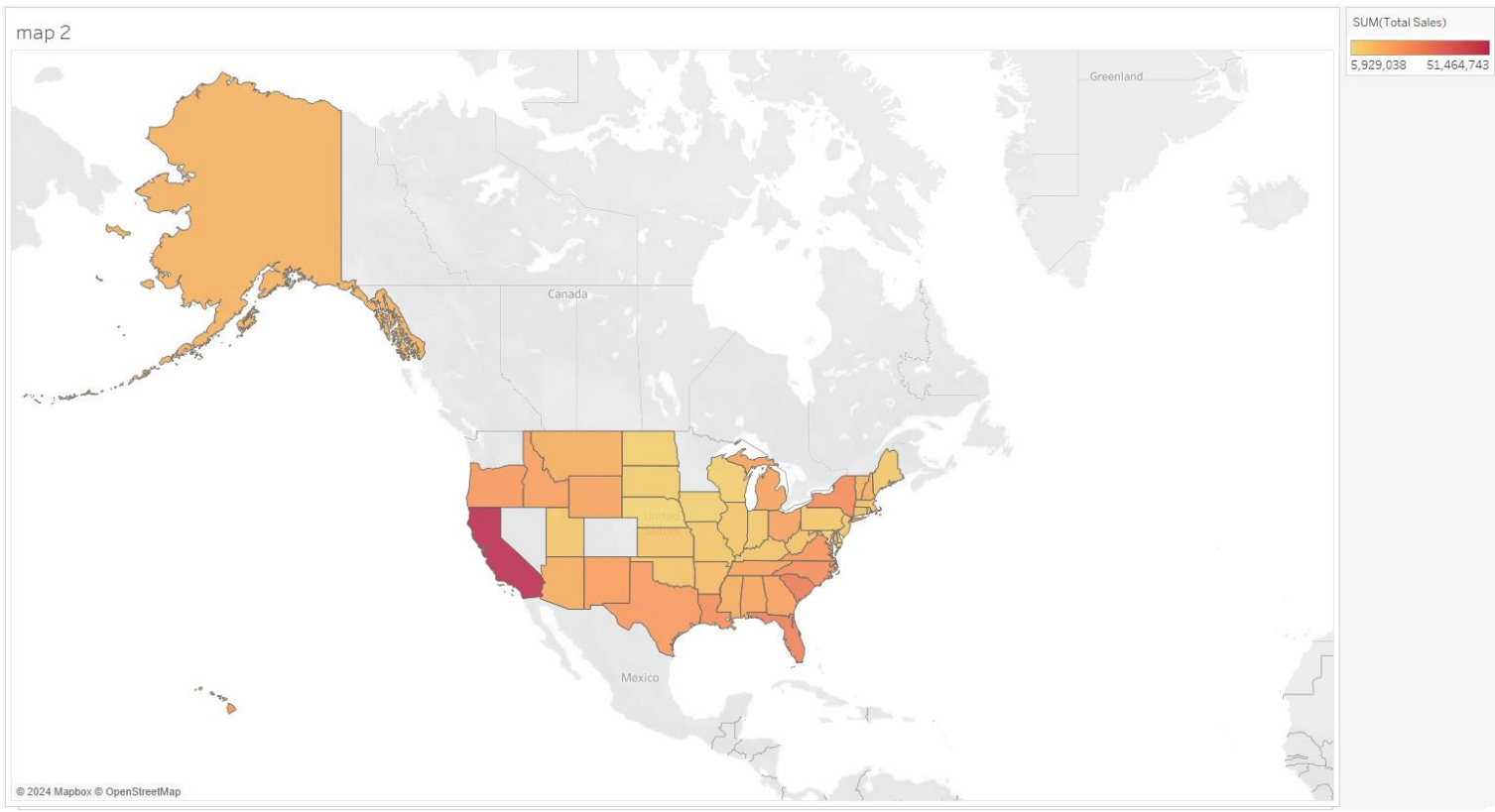

Figure 2: Total sales for each state.

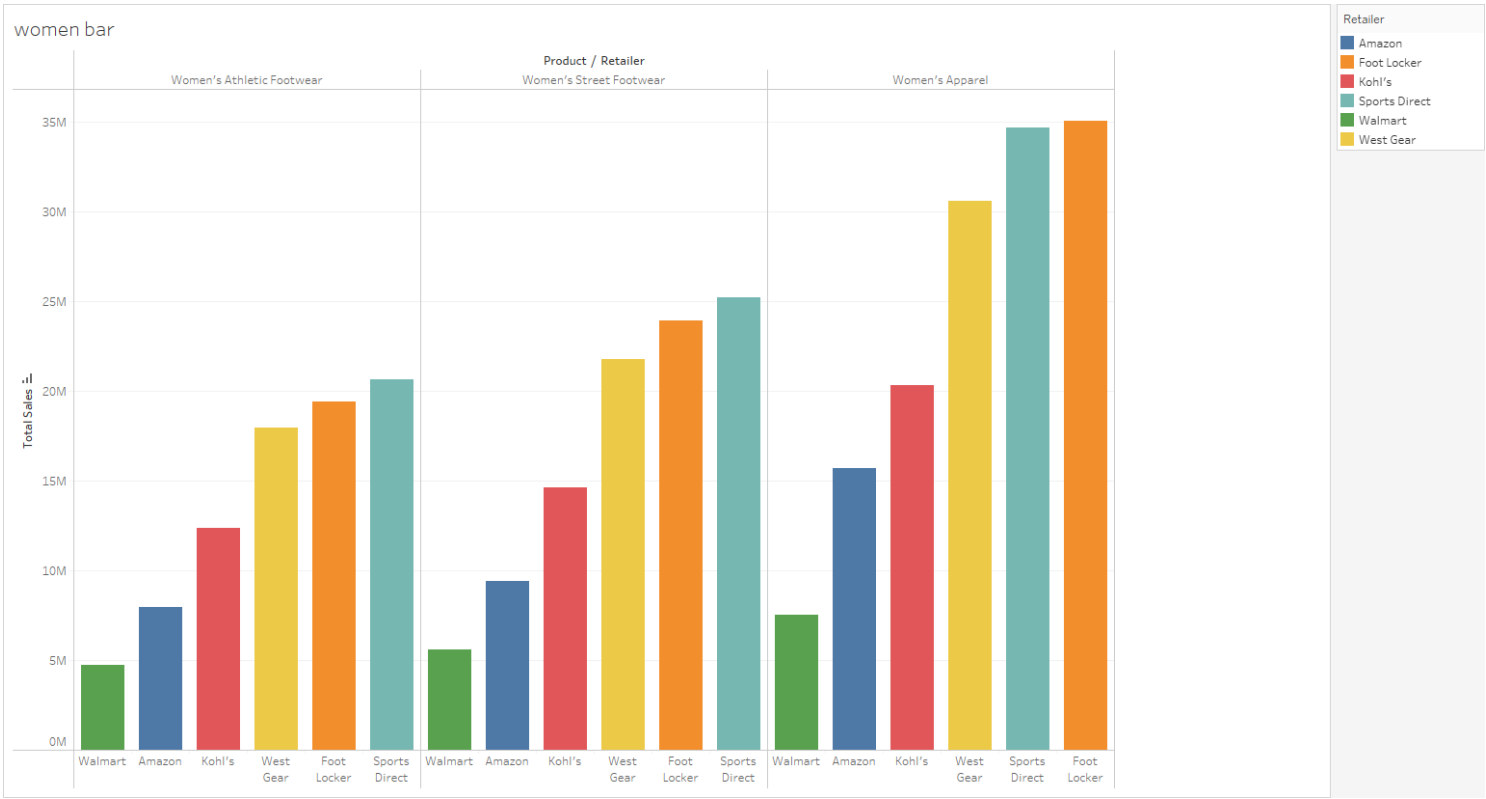

Figure 3: Total sales of each type of women's product and the retailer of those purchases.

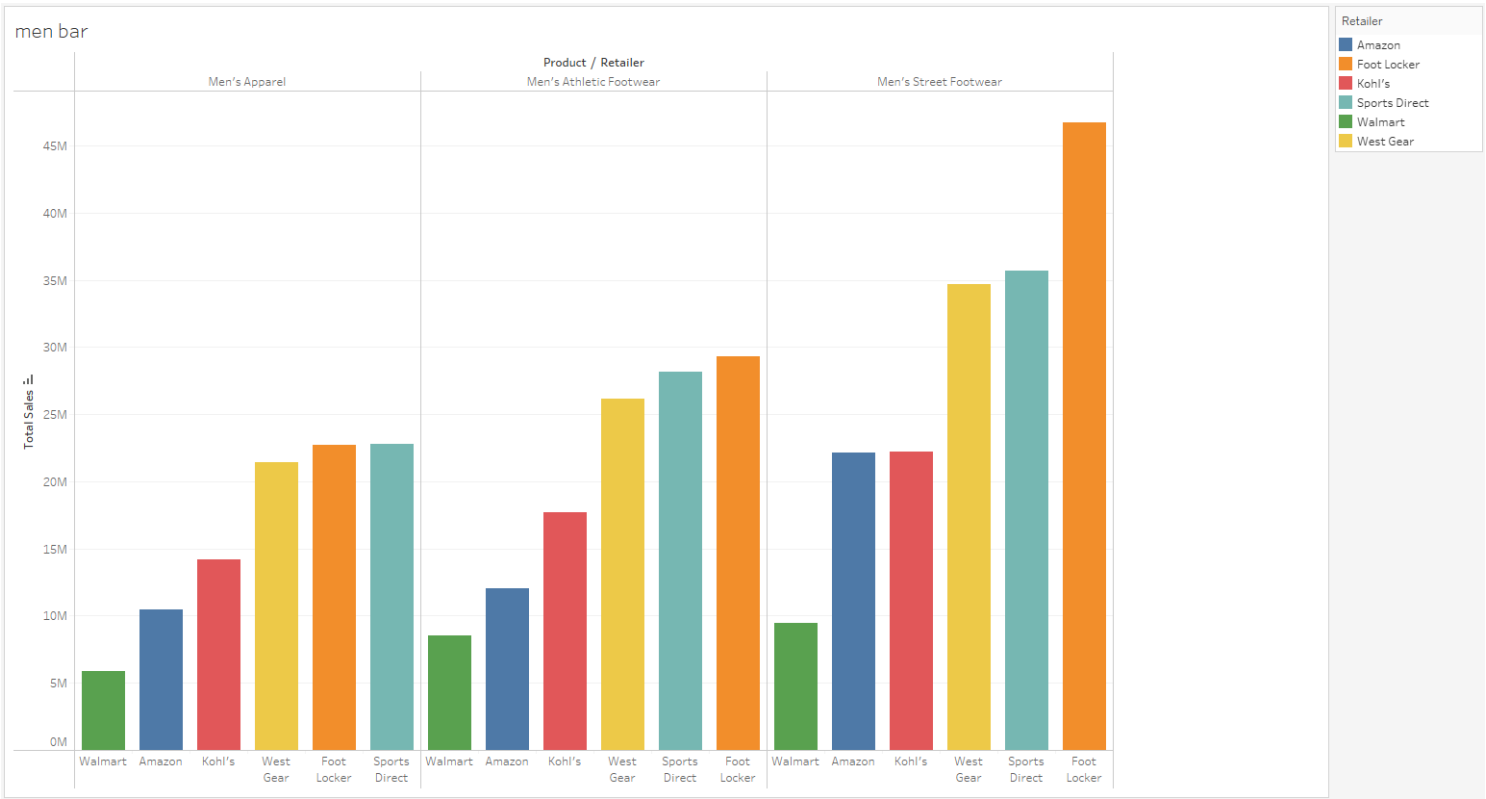

Figure 4: Total sales of each type of men's product and the retailer of those purchases.

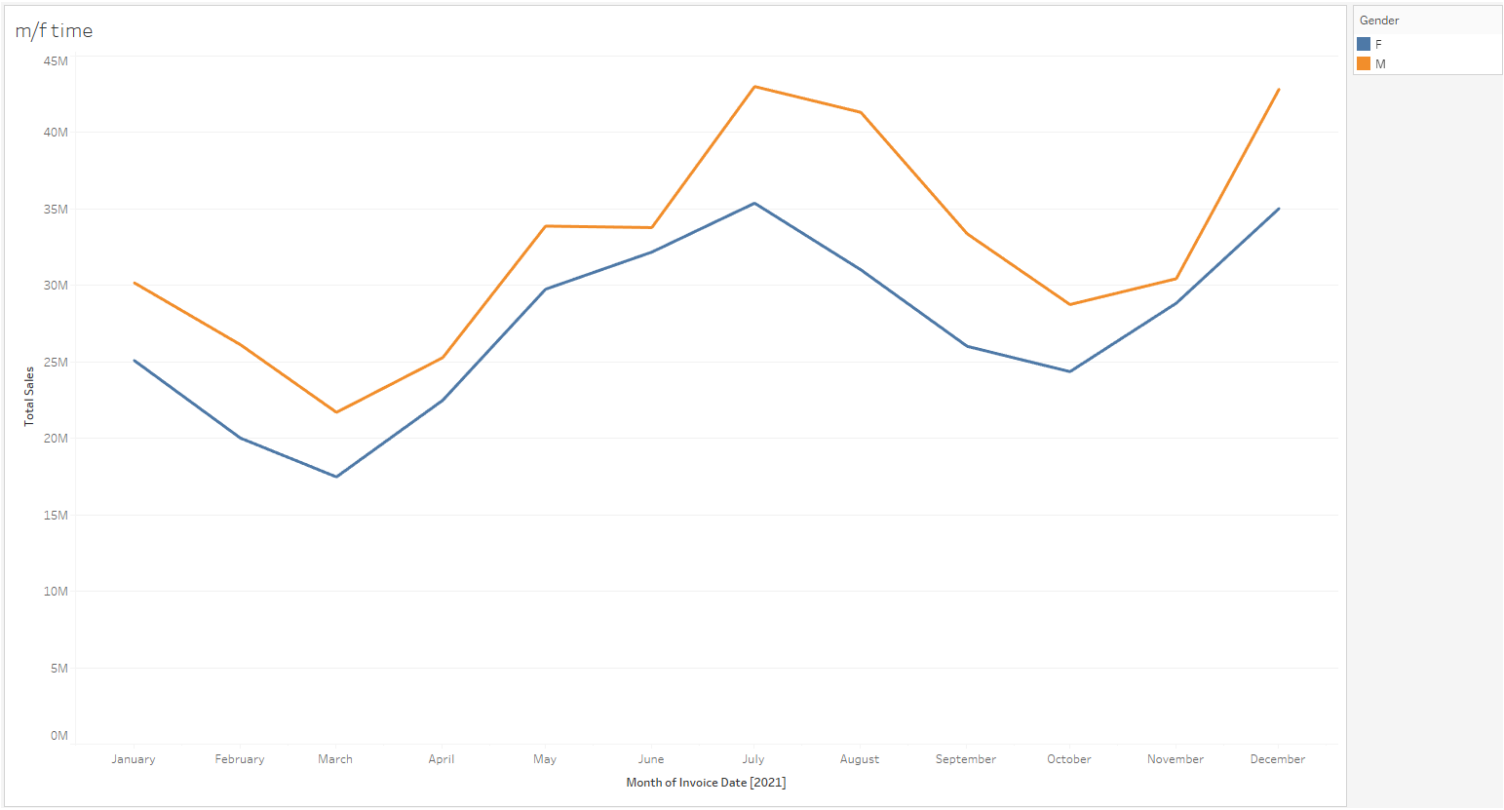

Figure 5: Total sales of men vs. women's clothes over 2021.

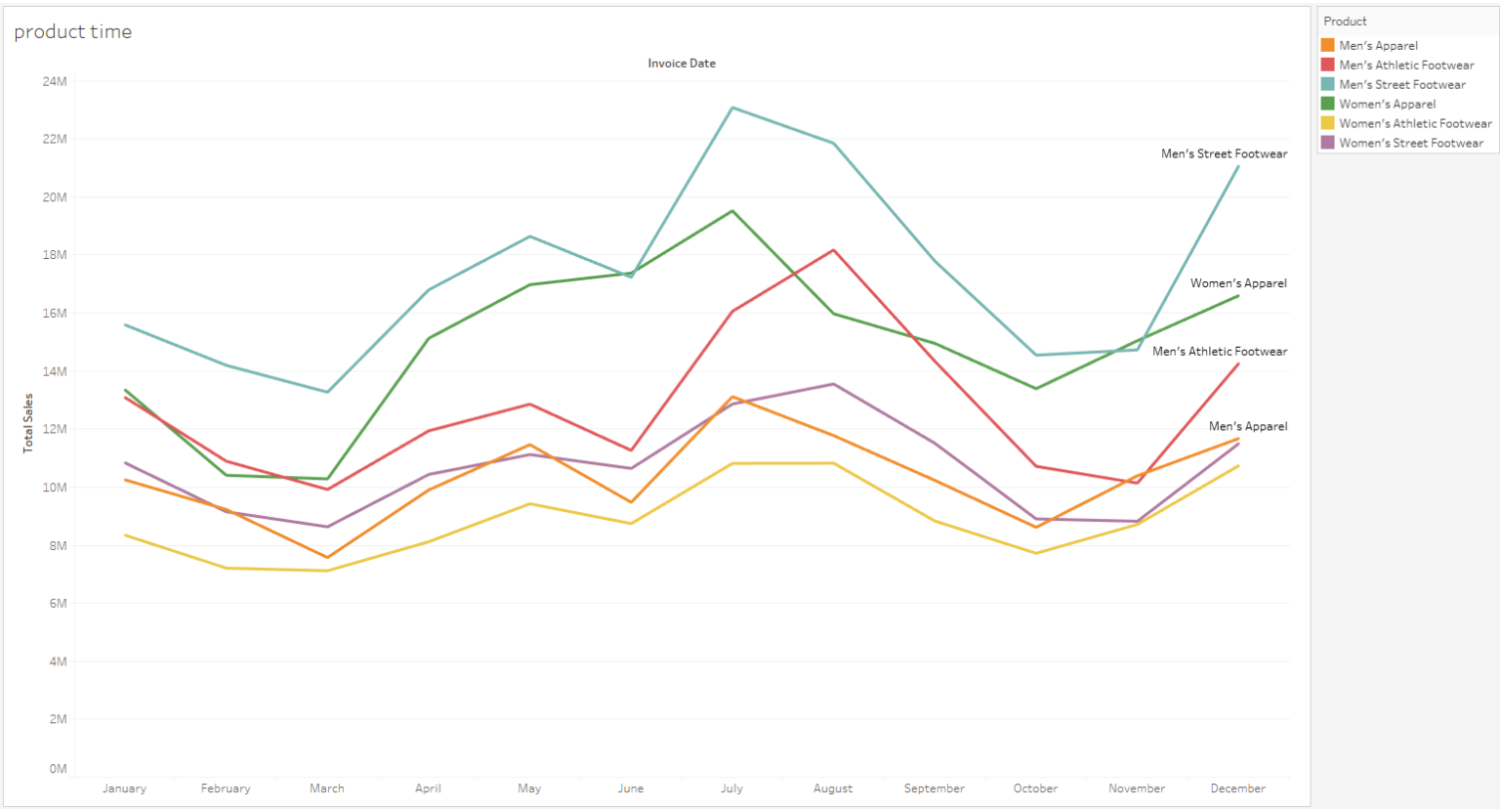

Figure 6: Total sales of each product type in 2021.

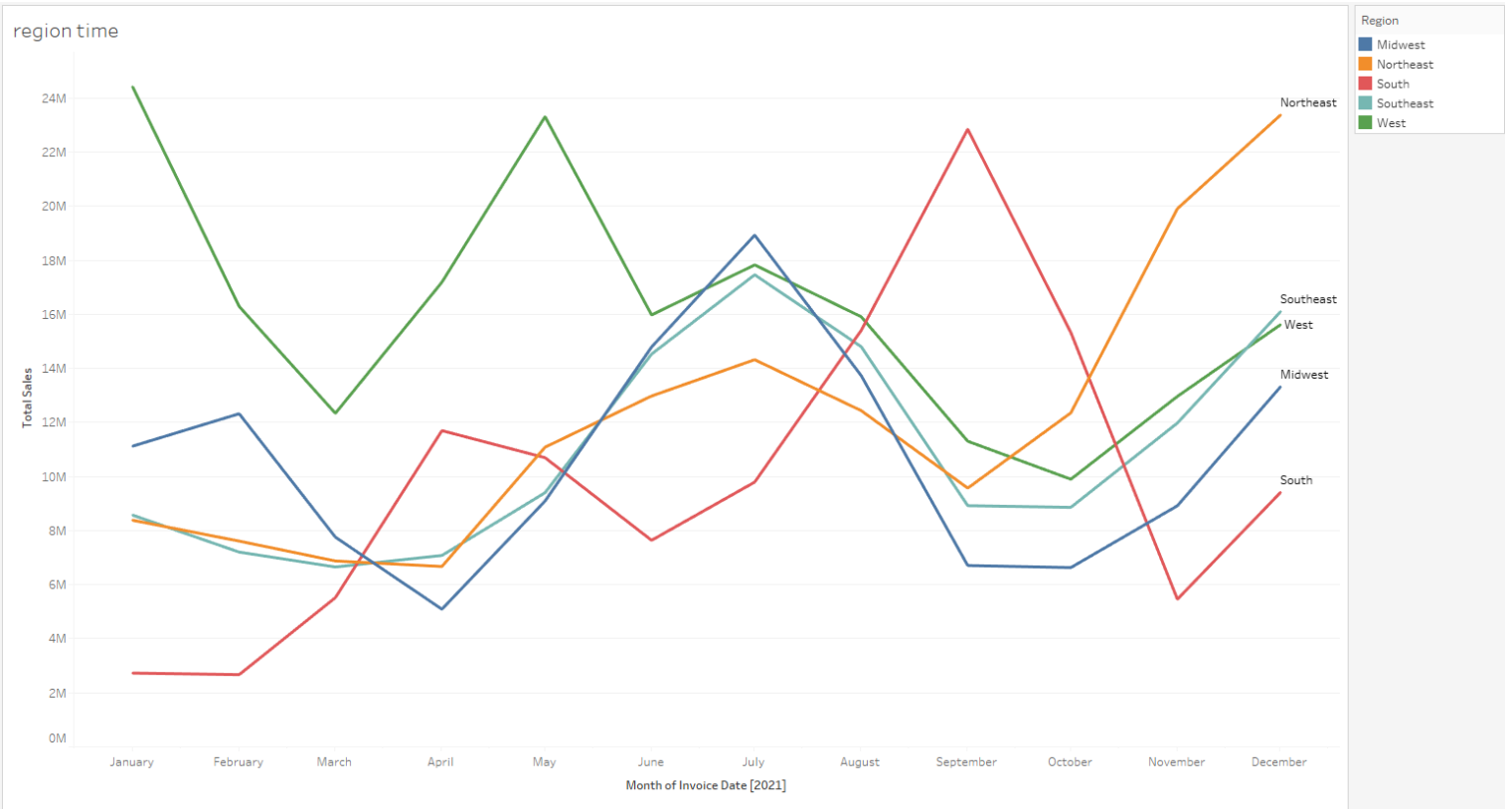

Figure 7: Total sales per region over 2021.

Figure 8: Performance of different sales methods during 2021.

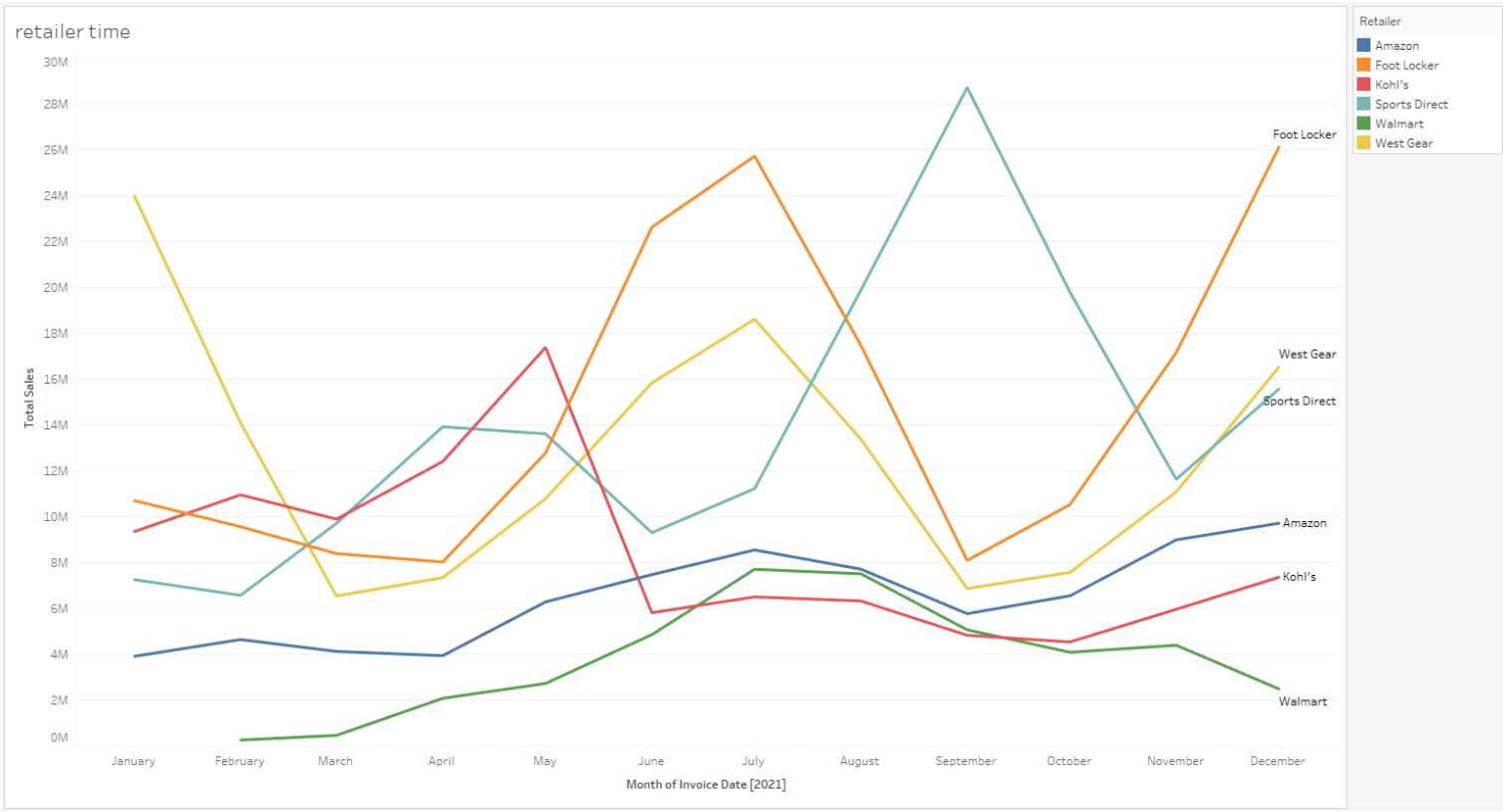

Figure 9: Performance of different retailers during 2021.

```{r}
df <- read.csv("adtest.csv")
```

```{r}
summary(df)
```

```
   Retailer          Retailer.ID       Invoice.Date      Region             State              City             Product            Gender         
 Length:9648        Min.   :1128299   Min.   :43831   Length:9648        Length:9648        Length:9648        Length:9648        Length:9648       
 Class :character   1st Qu.:1185732   1st Qu.:44244   Class :character   Class :character   Class :character   Class :character   Class :character  
 Mode  :character   Median :1185732   Median :44351   Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character  
                    Mean   :1173850   Mean   :44327                                                                                                  
                    3rd Qu.:1185732   3rd Qu.:44455                                                                                                  
                    Max.   :1197831   Max.   :44561                                                                                                  

 Price.per.Unit     Units.Sold       Total.Sales      Operating.Profit Operating.Margin Sales.Method         X2020.pop          X2021.pop       
 Min.   :  7.00   Min.   :   0.0   Min.   :     0   Min.   :     0   Min.   :0.100    Length:9648        Min.   :  577664   Min.   :  579548  
 1st Qu.: 35.00   1st Qu.: 106.0   1st Qu.:  4254   1st Qu.:  1922   1st Qu.:0.350    Class :character   1st Qu.: 2118488   1st Qu.: 2116950  
 Median : 45.00   Median : 176.0   Median :  9576   Median :  4372   Median :0.410    Mode  :character   Median : 5031864   Median : 5050380  
 Mean   : 45.22   Mean   : 256.9   Mean   : 93273   Mean   : 34425   Mean   :0.423                       Mean   : 8628032   Mean   : 8637886  
 3rd Qu.: 55.00   3rd Qu.: 350.0   3rd Qu.:150000   3rd Qu.: 52063   3rd Qu.:0.490                       3rd Qu.:10453812   3rd Qu.:10567100  
 Max.   :110.00   Max.   :1275.0   Max.   :825000   Max.   :390000   Max.   :0.800                       Max.   :39503200   Max.   :39145060  

 Total.Sales.cap     Units.sold.cap  
 Min.   :0.0000000   Min.   :0.000  
 1st Qu.:0.0006625   1st Qu.:0.000  
 Median :0.0029848   Median :0.000  
 Mean   :0.0294473   Mean   :0.001  
 3rd Qu.:0.0239240   3rd Qu.:0.001  
 Max.   :0.7592124   Max.   :0.002  
                     NA's   :9504   
```

Figure 10: Summary statistics of analyzed data.