Executive Summary

U.S / TJIX Retail and E-commerce sales Data analyzation

Overview

Throughout the company analysis, various attributes were related to each other and conclusions were drawn from doing so. Analysis of the general E-commerce and retail industry was also taken into consideration and related heavily to TJIX sales. In observing and analyzing the data, I have concluded that TJIX should very well submerge themselves deeper into the E-commerce world as well as increase their advertisement spending. In return if we do so we can expect to see great returns in TJIX total net sales each year.

Problem Summary

Why does TJIX have a total percentage of retail sales for E-commerce of less than 3%, whereas a competitor such as Macy's E-commerce sales as percentage of total sales is greater than 30%! Can we use E-commerce to increase our total net sales per year and can we use advertisement to increase the total net sales along with E-commerce sales?

Solution Summary

Knowing what we know now, I can confidently say that yes, we can effectively increase both categories of sales by implementing these strategies. We know that TJIX E-commerce sales are directly correlated with TJIX total retail sales, and we have also concluded after running linear regression that TJIX total retail sales is dependent on TJIX E-commerce sales, TJIX advertisement spending, U.S Total retail sales, as well as Year or quarter. With all of this being true and U.S total retail sales growing steadily with the change in Year or quarter we can conclude that any increase in any of these independent variables will cause an increase in TJIX total retail sales.

Value Proposition

The value behind the proposition lies with an increase in TJIX total net sales for each year in the future with an increase in E-commerce sales. An increase in advertising will likely entice this to happen because our linear regression shows us that TJIX total retail sales is dependent on an increase in advertising. With every $1million dollar increase in advertisement spending, there will be about a $12.7 million dollar increase in TJIX total net sales for the year. This was done at a 90% confidence interval with a correlation of 97% and all variables included being significant .

Key Next Steps & Summary

We can quickly begin to implement more advertising to cause an increase in Total retail sales for the company. Along with more advertising we can design a better or more well-versed website and online shopping options to encourage people to shop online as well. With E-commerce clearly on the rise along with retail sales itself this will bring in a whole new crowd of customers and with increased online E-commerce sales, comes increased Total retail sales.

Prior Knowledge

I was able to source Data pertaining to Total U.S retail sales and Total U.S E-commerce sales from 1999-2022. I was also able to source data pertaining to TJIX E-commerce spending for the years 2021-2022 and total retail sales worldwide for the years 2013-2022. Along with this was also Macy's E-commerce spending for the years 2021 and 2022 and total retail sales worldwide from 2013-2022. I was also able to obtain data related to each companies Advertising spending during the years of 2013-2022. Since 1999 E-commerce sales have grown a whopping 5,748% and Retail sales have grown only 246%. This is just another indication that E-commerce is growing faster than retail sales itself. E-commerce as a percent of Total retail sales in the U.S. has also grown by about 14% in the last 23 years! From this data we can ask ourselves why is TJIX E-commerce sales so low if it is growing so steadily in the US and what can TJIX do to improve these E-commerce sales? This may very well have to do with low advertising spending and low E-commerce marketing.

Preparation

Going through both sets of data there were attributes that I needed to create, and estimate based on moving averages and averages. Data pertaining to Total U.S retail sales was transformed to include attributes showcasing whether Retail or E-commerce performed better during a given quarter or a given year. Percentages relating to both categories were analyzed and differentiated to give us numerical data showcasing how large or small the difference between performance of Retail and E-commerce were. Quarters were also split up and recorded as binomial data types to test for seasonal patterns. With the data pertaining to TJIX advertising spending and E-commerce sales during the years of 2013-2022 we needed to estimate E-commerce sales for TJIX for the years 2013-2020 by using moving averages of the given attributes. We were then left with data showing TJIX advertising spending, E-commerce sales, Total US E-commerce sales, and Total US Retail sales for the years 2013-2022.

Modeling

After my data was cleaned and transformed correctly to use for analyzation, I began to dissect, analyze, and model it in different manners. Moving averages and conditional formulas in excel were used to create simple hypotheses and new data pertaining to TJIX business spending and sales. One of my first tasks at hand was running Linear regression with the data. Similarly, correlation matrices were made to prove this point further. I wanted to allow myself to see if any of our attributes in our data had a significant relationship with another or other attributes. If there is a significant relationship between any two attributes, we can then use our produced linear regression equation to predict our dependent attribute based on changes in our

independent attributes. This data was also used to predict future outcomes using decision trees. This is a more simplified analyzation of the data to showcase whether E-commerce is expected to grow more in the future or if Retail is expected to grow more in the future. Seasonal patterns were analyzed to conclude whether each separate quarter (q1-4) had a pattern throughout the 23 years. Data pertaining to advertising expenses was also analyzed by using linear regression as well and with that we can predict exactly how much an increase in advertising spending would advantageously affect TJIX E-commerce and total retail sales.

Application

Using data pertaining to U.S. Retail sales and E-commerce sales that I collected from the years 1999-2022 I was able to use a decision tree within rapid miner using a target variable of "E-commerce" or "Retail". The target variable was made by using conditional formulas in excel. If the quarterly percent change is greater for E-commerce that quarter, it will show "E-commerce" and vis versa. Using this decision tree, I was able to predict future quarters by using moving averages for my input attributes in a scoring dataset and more than 80% of my target variables were predicted as "E-commerce" having a greater % change than Retail sales in the next 10 quarters. The same thing was done with another target variable being whether retail or e-commerce had a greater % increase from a year ago. This was also held consistent with the predictions having E-commerce for more than 80% of the target variables. We can also look at our current data and see that E-commerce has increased by 308% more than Retail since 1999 in each quarter consecutively and 1105% more than retail in each year consecutively.

I used the data pertaining to each companies E-commerce sales during the years of 2013-2022. I used this information and the percentage of each Stores E-commerce sales to the company's total net sales for that year. TJIX had an average of 2.6% of their total sales being E-commerce and Macy's had an average of 30.6% of their total sales being E-commerce. I used these averages to predict the previous year's (Missing data) E-commerce sales for Macy's and TJIX.

I ran a correlation matrix between the attributes 'Year', 'total retail sales', and TJIX company net sales worldwide. Year or quarter had a 93% correlation to 'US Total retail sales. Year also had a 69% correlation to 'TJIX companies net sales worldwide'. I also ran a correlation matrix between 'E-commerce total sales' and 'Total retail sales', and these two attributes of course had a correlation of 96%. I then ran a correlation matrix between 'TJIX net sales' and 'TJIX advertisement spending' (I normalized each attribute first). The results left me with a 60% correlation between the two variables.

A time series with the U.S. retail and E-commerce sales was created and I have concluded that although sales are dependent on year or quarter, there is no seasonal pattern within the data pertaining to quarters of the year.

Another approach I took was running linear regression at a 95% confidence interval, having total retail sales as the dependent variable. In conclusion I observed that Total retail sales in the US directly relates to Total E-commerce sales in the U.S. and having Total TJIX sales as a

dependent variable, I concluded that it of course as hypothesized directly relates to TJIX E-commerce sales. Again, having E-commerce and Retail total sales in the U.S as dependent variables and 'quarter' as independent variable, with every quarter increase there will be an increase in E-commerce of about 2292 ($million) and in retail for every quarter increase there will be an increase of about 8947 ($millions) in Total retail sales. This clearly tells us that as time goes on both Total retail sales as well as total E-commerce sales will increase steadily. From this we can know that since TJIX total retail sales is directly related to U.S. Total retail sales and US total E-commerce sales. All these regressions have correlations above 89% or R-squares above .89.

Regression was also run with TJIX Net sales being the dependent variable and 'Year', 'US total retail sales', and 'TJIX companies advertising spending' being independent variables. With a correlation of 97% and a confidence interval of 90% all attributes were significant.

Knowledge

As I observed, E-commerce is growing steadily and at a faster rate than Retail sales itself. We can calculate that if TJIX were to have similar distribution of 30% of their sales being E-commerce, they would increase their Avg total net sales per year by about $9009 million dollars per year. It is clear to us that as years or quarters advance so will US total retail sales and so will TJIX net sales worldwide. In addition, E-commerce will grow larger as US total retail sales increases. We can also conclude that there is a good chance that TJIX advertising spending has an incredible impact on TJIX net sales. . TJIX total retail sales is expected to increase with the increase in quarter there will be an increase in TJIX retail sales. With every 1 ($millions) increase in TJIX advertisement spending, there will be about 12.7 ($millions) increase in TJIX net sales. This is extremely important because it shows us just how important advertisement spending is for TJIX increasing their net sales. It is obvious that TJIX E-commerce sales have a direct impact on TJIX net sales. In conclusion, with any increase in the independent variables of Total US retail sales, Total US E-commerce sales, Year/Quarter, or TJIX advertising spending will have a significant positive effect on TJIX total net sales and TJIX E-commerce sales.

Reflection

I was able to use charts and mainly pivot tables and Tableau to visualize a lot of this data. These visualizations made it easy for me to see quality information such as how E-commerce is growing faster than Retail sales itself or how TJIX's E-commerce sales were so much lower than its competitions even though their total retail sales were very similar. These visualizations could be used in a presentation but served me as a great tool to look at the data from an easier perspective, allowing me to draw various conclusions from it. In conclusion I think that my initial analysis went well when I first was dealing with only the dataset pertaining to U.S. Total retail sales and E-commerce sales filtered by each quarter. However, when I started diving into the data, I realized that some attributes such as advertisement spending and the company's data pertaining to their sales will be useful in determining if the company should apply more effort toward advertising and E-commerce to increase their net sales. Once I was dealing with two different sets of data, 3 if you count Macy's as another one for reference, it became a lot

more complex to run classifications, regressions, and other analysis. It did take some time, but I got it done after cleaning and transforming some data to fit into the models I was using and parameter settings that were needed. I think in the future when I am to do something like this again, it may be easier to finish analyzing one set of data before involving another one. I believe if I were to break up the datasets and analyze them individually before drawing connections then I would have felt more organized!