Multivariate Analysis of Pepsico’s Revenue

May 3, 2022

BA222 A1

Professor Carlos Casso Dominguez

Jacob Rose U29999110

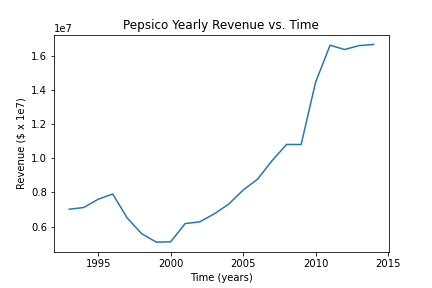
Dana Winer U13501316

Gabriel Moncau U83571123

Andres Tamayo U45018174

Company and Data Description

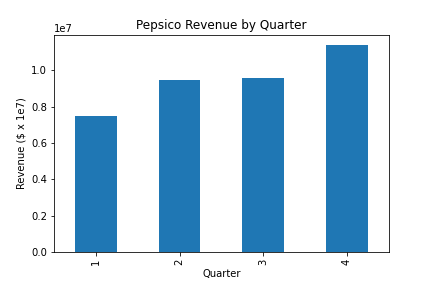
PepsiCo is a leading food and beverage company that manufactures and distributes its products in more than 200 countries. PepsiCo’s products include a wide range of products such as Pepsi, Gatorade, Frito-Lay, and Quaker Oats, among others. PepsiCo is part of the Food and Beverage Industry, including all companies involved in transforming raw agricultural goods into food products. Pepsi is the largest food and beverage company in North America and the second-largest globally, behind only Nestlé.

1

The time period for the data is from 1993 to 2015. In 1993, Mountain Dew's volume doubled, bringing in $2.2 billion in revenue. PepsiCo purchased the orange juice company Tropicana Products in 1998, and merged with Quaker Oats Company in 2001, adding with it the Gatorade sports drink line and other Quaker Oats. In 2003, PepsiCo had “false” allegations in India, so there was a decrease in revenue worldwide. Between 2005 and 2007 PepsiCo revenue started to increase due to new products and unique advertising, which led to PepsiCo be #1 in the beverage industry, surpassing CocaCola. Between 2009 and 2015 there was a high increase in revenue due to buying their own subsidiary and expanding around the world.

Identifying Time Trend, Seasonality and Outliers

Revenues of PepsiCo show a clear upward trend over the last 21 and a half years. However, there is a seasonal trend in the data that can clearly be observed in Figure 1. Although revenues showed a rising trend overall, there was a considerable downward trend between the years of 1996 and 2000. After this, Pepsico noticeably bounced back and have continued on a positive trend since.

2

Seasonal variation is a common feature of revenue data whether the trend is upward or downward. The data is recorded quarterly, and there is a clear variation in quarterly revenue as shown in Figure 2. The average revenue figure is the lowest in the 1st quarter whereas the highest in the 4th quarter. The average revenue in the second and third quarters are almost the same. One of the salient features of the quarterly variation is that the seasonal fluctuations of revenue rise as the year goes on, taking a dip at the start of the new year. Figure 1 of the fitted value shows a clear up and down trend to visualize this, which is included in the previous section. As a result, one of the assumptions of simple linear regression may not be satisfied.

We’ve additionally located two outliers, at times 13 and 63, within the second between regression (reg\_bet1) and the final regression (reg\_aft) by observing the residuals plots of previous regression models, to ensure accurate predictions.

Regression Model

Keeping in mind the four different time trends, and the seasonality of Pepsico’s revenue, we were able to decipher the most optimal regression model for their revenue progression. This was the regression made for time above 67, or 2010 Quarter 1 and beyond because we utilized the last regression to forecast 2015’s revenues, which will be explained in the later section. This means that we only had 19 observations in this regression as the others had shown different time period patterns. This means that our model may be inconsistent and inaccurate as we do not have the usual minimum 30 observations for a three variable regression. We kept the revenue as our dependent variable and using time, C(quarter), and accounting for outliers, we were able to achieve an Adjusted R-squared of 0.964, attributing 96.4% of the effect on revenue to the factors mentioned previously. By accounting for outliers and observing seasonality, we were able to limit biases, such as omitted variable bias. Additionally, the time trend held a 0.001 p-value, and a 4.282 t-value, making it very statistically significant. We are able to make these statistical inferences as usual due to an absence of multicollinearity, which is the estimation of imprecise regression coefficients due to high correlation among the explanatory variables. An increase in each independent variable leads to a significant increase in the revenue, with the partial exception of time, as its effect is close to 100 times less than other independent variables.

Forecast

We used our model to predict the revenue for 2015 and got the results listed in the table below, next to the actual values. Our predictions were generally around ~1 million USD higher than the actual amount, meaning somewhat inaccurate. Looking at the 95% confidence interval suggests we had even less accuracy, as the actual values did not fall within the interval for any of the predicted quarters. Q1 came the closest with the interval’s lower value being ~12.6m, about ~380k, or 3.16%, higher than the actual revenue. While we are unsure what caused our model’s inaccuracy, it is clear that despite having an r-squared of 0.974 our predictions were not equally impressive. It is likely that there were more variables affecting Pepsi’s revenue that were unaccounted for and, thus, decreased our accuracy.

| Time Value (Quarters) | 2015 Q1 | 2015 Q2 | 2015 Q3 | 2015 Q4 |
| --- | --- | --- | --- | --- |
| Predicted Revenue (USD) | 13,322,710 | 17,473,960 | 17,596,760 | 20,783,160 |
| Actual Revenue (USD) | 12,217,000 | 15,823,000 | 16,331,000 | 18,585,000 |

3