Business Insights Report

Forecasting & Predicting the Future Using Data – DAT - 3531 – NYC1

Catarina Jesus, Student ID: 26810706

In the vibrant city of Seattle, where the weather can be very unstable and can vary dramatically from one day to another, businesses can face challenges related to climate conditions in their operations. This report dives deep into an analysis of Seattle´s weather data, exploring variables such as precipitation, maximum temperature, minimum temperature, and wind to predict weather conditions and draw meaningful business insights. By leveraging predictive and forecasting modeling, it is possible to provide valuable insights to empower businesses, optimize operations and enhance their strengths in the face of weather-driven uncertainties.

To build a strong base for the analysis, it was necessary to start examining the historical weather data, from January 1st, 2012, to December 31st, 2015. By exanimating the patterns and trends in the variables of precipitation, temperature extremes, and wind, it was possible to gain a deeper understanding of what makes the weather sunnier. A sunny day will allow better operation of the companies and provide in the long-term, higher chance of profitability.

In the analysis of the Seattle weather dataset, it was employed predictive modeling techniques to predict weather conditions and understand the variables with higher influence on the dependent variable, weather. The techniques used were logistic regression and Gini decision tree regression, which proved to be highly effective in capturing the relationships between the independent variables and predicting outcomes.

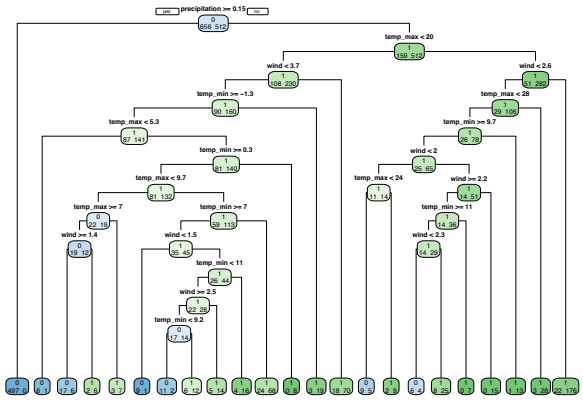

Figure 1- Gini Decision tree

The results revealed that the decision tree model exhibited a higher accuracy rate and better confusion matrix. The decision tree presented above provides a visual representation of the decision rules and their hierarchical structure. The decision tree analysis highlights precipitation as a critical variable affecting the probability of good or bad weather. This precipitation impacts various aspects of business operations and consumer behavior. Unfavorable weather conditions associated with higher precipitation, whether in the form of rain, snow, drizzle, or fog, which means greater and equal to 0.15, can lead to decreased foot traffic and slow down transportation. For instance, this may discourage customers from pursuing purchasing behaviors, thereby impacting business revenues negatively. Moreover, will slow the supply chain and operations of the business due to bad weather conditions.

Furthermore, the maximum temperature is the second variable with a significant influence on the weather conditions in Seattle. Temperature plays a crucial role in customer behavior and the opportunity for the business to play its operations. Warmer temperatures often attract more people to outdoor activities which can benefit businesses. On the other hand, extremely cold temperatures deter customers which will automatically reduce sales.

Lastly, the decision tree analysis also identifies wind as the third variable with a notable influence on weather conditions. Strong gusts can disrupt operations and damage infrastructures which will lead to potential business interruptions and financial losses. In addition, it can influence consumer behavior by affecting comfort experiences outside.

In addition to analyzing historical weather data, the report also focused on forecasting future weather conditions in Seattle to help the businesses and customers to gain valuable insights into potential weather conditions and help them to make decisions to optimize their operations and behaviors and proactively plan for weather uncertainties and enhance their profitability.

Before doing the forecast, it was important to see the behavior of the variables in the analysis all over the years per month. Figure 2 shows the average precipitation per month in the different 4 years, which is possible to observe certain trends. For example, the average precipitation tends to be higher during the winter months, which may be useful for the business to have into consideration and be prepared for potential impacts in their operations and make strategies to overcome these difficulties. Figure 3 exhibits the maximum temperature per month in the 4 years in analysis. By exanimating the plot, it is noticeable that the maximum temperature tends to be higher during the summer months, which indicates higher profitability and foot traffic during that period. Businesses should focus on marketing and product offerings for summer-related products or services. Lastly, figure 4 demonstrates the variations in the wind throughout the months and years. One more time, it is possible to visualize

that wind speed is higher in the fall and winter months. Businesses should adjust their schedules and operations accordingly to these behaviors.

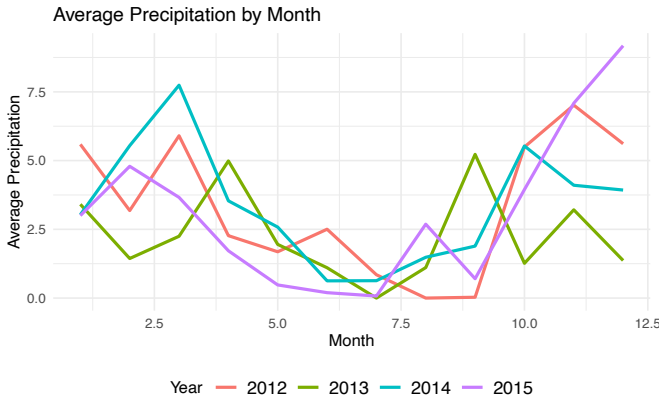

Figure 2- Average Precipitation by Month per Year

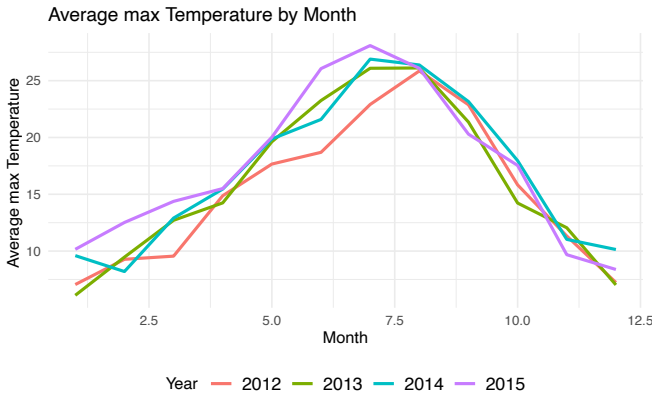

Figure 3- Average Maximum Temperature by Month per Year

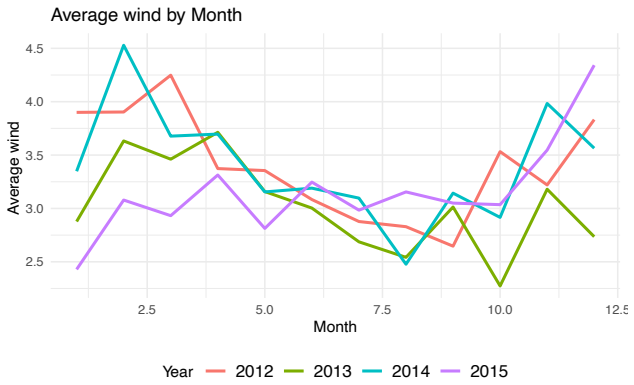

Figure 4- Average Wind by Month per Year

As mentioned before, precipitation plays an important role in weather conditions, by analyzing the forecast for the next 12 months, it was possible to identify the following numbers 9.012280, 9.112194, 9.051743, 9.088318, 9.066189, 9.079577, 9.071477 9.076378, 9.073413, 9.075207, 9.074121, 9.074778, having into consideration that the forecast present a high level of uncertainty between 221% to 724%. Moreover, this variable presents a high percentage of forecast volatilities around 224%. Furthermore, after forecasting the variable maximum temperature for the next 12 months, the numbers found were the succeeding ones 2.002874, 2.584622, 2.944451, 2.963718, 2.988545, 3.260744, 3.613116, 3.954603, 4.086378, 19.513835, 15.547872, 11.674406, 9.540338, keeping in mind that the forecast displays a high level of uncertainty between 200% and 427%. Although this forecast looks more

accurate compared with the forecast of the variable precipitation, once the higher numbers are related to summer months, this forecast disposes of a high percentage of forecast volatilities, about 617%. Lastly, the forecast of the variable wind demonstrates the following numbers 3.984740, 4.121233, 4.049976, 4.087176, 4.067755, 4.077894, 4.072601, 4.075364, 4.073922, 4.074675, 4.074282, 4.074487, consider a level of uncertainty between 49% and 121%. Additionally, this variable, one more time, shows high forecast volatility, around 56%. Overall, high volatility in the forecast of the three variables creates challenges and uncertainties for business operations.

In conclusion, after analyzing the weather in Seattle, and performing the prediction modeling, it was possible to conclude that precipitation, maximum temperature, and wind have a higher impact on the dependent variable, weather. This information is highly correlated with the performance of the business operations and consumers' behaviors so forecasting these weather variables will help businesses to have higher odds to profit. Unfortunately, forecasting these future weather variables is difficult due to the high level of uncertainty and volatility, so businesses should always be aware of seasonal factors and have a backup of unpredicted events.