

Guided Capstone Project Report

The aim of the project:

In this project first I determined the problem statement to conduct my project based on this statement (i.e., proposal). The aim of this project is to maximize the benefits of Big Mountain Resort.

Data analysis:

First I imported the necessary libraries (pandas, numpy etc). I created folders for my data, figures, and models. Then, I looked at the first few rows and columns of the dataset to understand the dataset that I am working on. I performed the descriptive statistics to see the minimum and maximum values, counts (if there is any missing case), and standard deviations.

Then I plotted the variables to understand the distribution and features of the variables. The bar charts revealed that two variables which are state and region are identical. Since this situation leads to collinearity in my model, I dropped the region variable.

I looked at the correlation matrix among the variables. There were some missing cases in the dataset. I handle missing cases by utilizing different approaches like using average values. With the clean data I split the dataset in two parts train and test to conduct linear regression model. In the model, I looked at the effects of other variables on adultweekend; to estimate the adultweekend price. To measure the accuracy of the model, I tested explained variance score and mean absolute error. Finally, I draw scatter plots to visualize my results.



