# Guided Capstone Project Report

## Problem Identification Overview

### Problem:

What opportunities exist for Big Mountain Resort to recoup the $1,540,000 of

additional operating costs from the new chair lift in this season without risking the

consistent yearly 9.2% profit margin and what can be the resulting revenue this season

from the implementation of these opportunities?

One way to answer the above question is to see if Adult Weekend price can be adjusted to increase revenue in the upcoming season to account for the increased operating cost from the addition of new chair lift. Therefore, this project will focus on predicting the expected Adult Weekend price. Our modeling response therefore will be AdultWeekend from the data set.

### Data:

Database manager, Alesha Eisen, has provided the CSV file that contains characteristics and price information for 330 resorts including Big Mountain Resort.

### Model:

Linear regression model was employed because the response and the predictor variables are numeric.

### Deliverables:

* Jupyter notebook outlining modeling process from data exploration to best model results
* Project Report
* Executive Summary

## 

## Data Preprocessing steps

* Imputed missing values with Median, Mean and Zeros as appropriate
* Developed clusters
* Created dummy features for categorical variables
* Standardized the magnitude of numeric features
* Split the data set into training and test sets by using K-Fold cross-validation method

## Model Description & Performance

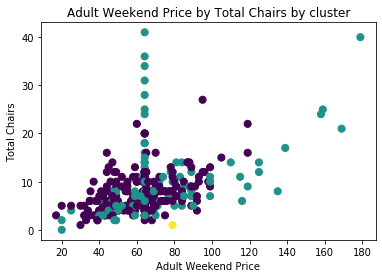
Input data size is 330 rows with 27 variables. Data was split into train and test sets using K-fold cross-validation with 20 folds (k=20) then linear regression was fit to the training set. The following table shows the top 10 features with their coefficients along with the explained variance and mean absolute error as to the model performance.

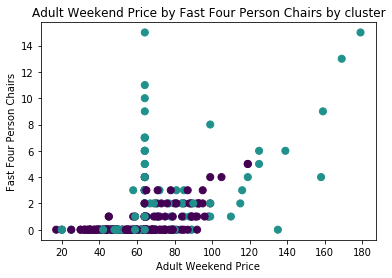
|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Coefficient** | **Explained Variance** | **Mean Absolute Error** |
| total\_chairs | 1.330612e+14 |  | |
| fastQuads | 5.044378e+13 |
| surface | 4.726204e+13 |
| double | 4.164906e+13 |
| triple | 3.715382e+13 |
| quad | 3.011180e+13 |
| fastSixes | 1.495409e+13 |
| trams | 1.284897e+13 |
| fastEight | 1.263179e+12 |
| AdultWeekday | 2.008679e+01 |
|  | | 0.92 | 5.54 |

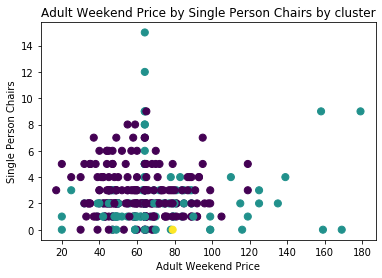
### 

## Model Findings

The model predicted the adult weekend ticket price for Big Mountain Resort to be $89.8 which suggests an increase of $8.8 to the current price of $81. This prediction is based on the features identified from the other 329 resort data fitting the model to Big Mountain resort data.In addition, the following charts demonstrate the relationship between several of the predictors and the criterion by clusters with color code.









### Next Steps

* Assess residual plots and improve model performance by handling biased coefficient estimates and predictors
* Assess if predictors are statistically significant or not (if R-squared value is low)
* Run F-test to confirm whether the relationship between the model and the response variable is statistically significant or not
* Source additional data on the revenue, costs and number of customers for previous seasons to improve the model and build similar model to predict Adult Weekday prices