**Guided Capstone Project Report**

Identifying how the Big Mountain Resort can recoup   the increased operating costs from the new chair this season.

**Summary of Findings:**

Based on the model I chose for identifying relation between features in the dataset and given the characteristics of the resort in comparison to other ski resorts and their unique characteristics, the predicted vs actual AdultWeekend price of Big Mountain resort is:

|  |  |
| --- | --- |
| **Actual AdultWeekend Price** | **Predicted AdultWeekend Price** |
| $81.00 | $64.06 |

Similarly, for all response variables, the predicted vs actual prices:

|  |  |  |
| --- | --- | --- |
|  | **Predicted Price** | **Actual Price** |
| **AdultWeekday** | 58.05 | 81 |
| **daysOpenLastYear** | 115.53 | 123 |
| **projectedDaysOpen** | 120.51 | 123 |

**Top Features that affect the** *AdultWeekend, AdultWeekend, daysOpenLastYear, projectedDaysOpen* **of Big Mountain Resort are:**

Vertical change in elevation from summit to base, Snow making machinery area covered, Count of regular speed four person chairlifts, Count of regular speed three person chairlifts, Number of trams, Number of fast four person chairlifts, Number of fast eight person chairlifts, Count of regular speed single person chairlift, Sum of all the chairlifts at the resort and Number of Terrain Parks at the resort.

**Details:**

In the data pre-processing step, I identified that there is a strong correlation between summit-elev and base\_elev features (0.98). Having scaled the data in Step 4 and 5, I applied the linear regression model on the third model that I chose after Step 3. I **dropped** the following columns:

**'Name'**: being of no relevance for modelling data

**'state':** cannot be treated as an actionable trait associated with ticket prices

**'AdultWeekend'**: this is the response variable

**'summit\_elev','base\_elev'**: these columns have been highly correlated. Also, these features cannot be treated as an actionable trait by the management of Big Mountain.

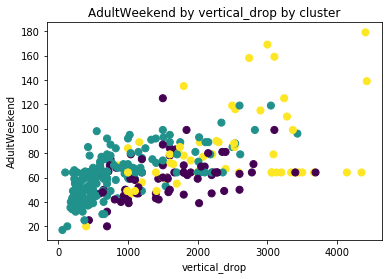
The Linear regression model provided the following coefficients of features with respect to the AdultWeekend prices:

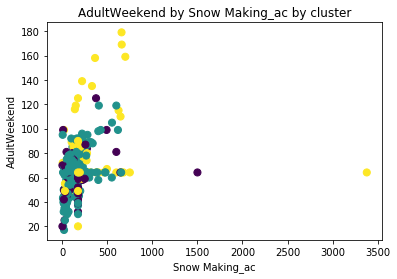


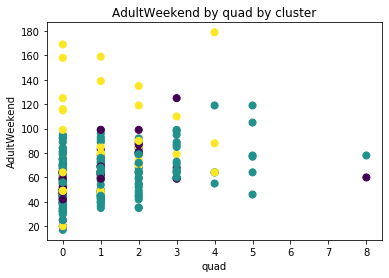
Identified the top features that have the highest correlation coefficients w.r.t the AdultWeekend price, and made scatter plots using some of these to see how they affect the AdultWeekend price individually.

**Here are a few of them:**

1. vertical\_drop is the most strongly correlated with Adult Weekend price with the highest coefficient of ~8.65, which means that this feature is the most important for any further predictions about ticket pricing of Big Mountain Resort

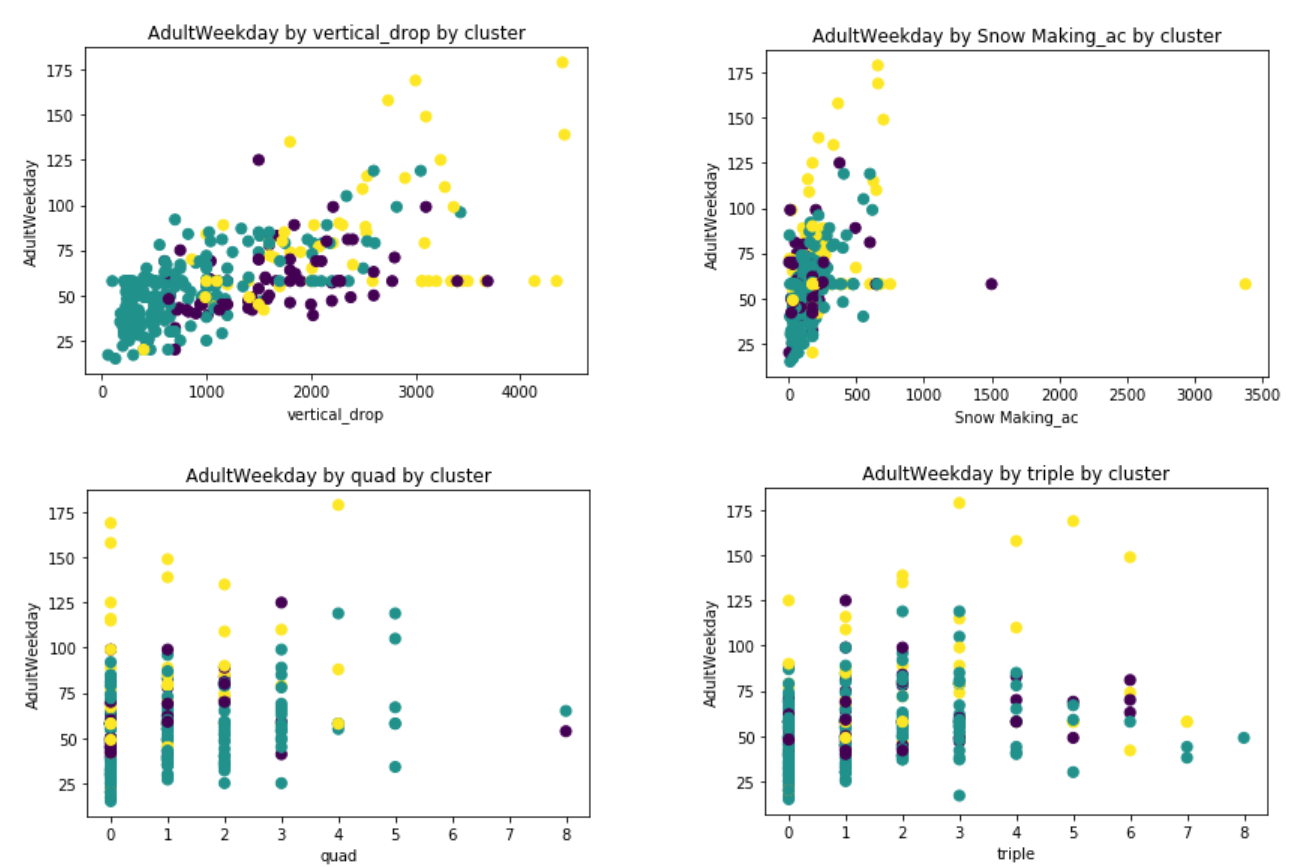


1. SnowMaking\_ac is the second most strongly correlated feature with Adult Weekend price with coefficient of ~5.70
2. Count of regular speed four person chairlifts(quad) also impacts the Adult Weekend price with an increase of ~4.35 times for every unit increase in the the number of such chairlifts.

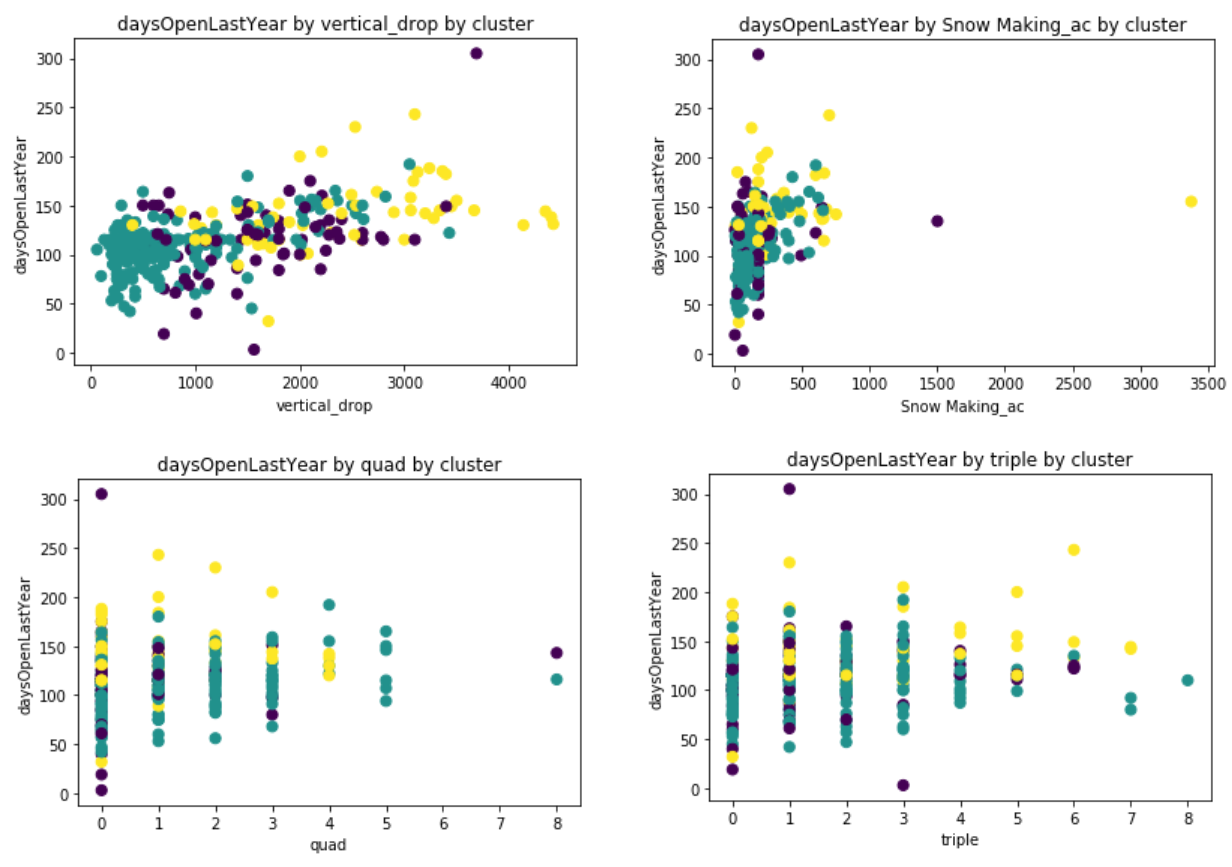


1. Count of regular speed three person chairlifts also impacts the Adult Weekend price with an increase of ~3.38 times for every unit increase in the number of such chairlifts



**AdultWeekday**

**daysOpenLastYear**



**projectedDaysOpen**

