# Big Mountain Resort

Executive Pitch

# Problem Identification

**Problem Statement**

- Big Mountain Resort has recently installed an additional chair lift to increase distribution of visitors across the mountain.
- This has resulted in the operating costs increasing by $1,540,000 this season.
- Need a solution to increase this year's annual revenue to offset the increasing operating costs.

**Problem Identification**

- What strategies Big Mountain Resort needs to resort to by doing a comparative analysis of all the 330 competitor resorts across the US in order to increase this year's annual revenue and keep at ~9.2%?

**Scope of Solution**

- **Aim**: Determine the best strategy to achieve the desired revenue targets

- Scope of solution would be to target the 330 competitors that would be not just limited to local Montana region but even across other states across the US.
- Focus on renewed marketing strategies, customer segment groups, ticket pricing, increased capacity offering to attract more footfalls.

# Recommendations and Key findings

**Key Findings**

- One important factor chosen for the scope of this project was Average Weekend ticket price.
- It was found that Big Mountain Resort fares pretty well as compared to competitors in terms of the Summit elevation and Base elevation it possesses.
- The investment in additional chair lift will definitely provided a boost to the visits even if the ticket price is increased as compared to its current price.
- There are other variables that could add up to the increased revenue (but not accounted for the scope of this project) are:
    - Average Weekday ticket price
    - Number of days to be open

**Recommendations**

- The actual Weekend ticket price for 'The Big Mountain Resort' is **$81**, however, going by the predictions using the modeling, the expected ticket price should be targeted at **$91.14** to stay in-line with the other resorts depending on varied features Big Mountain offers vs. them.
- These increased ticket prices should help the Management achieve their goal of 9.2% of profit margin despite increased CAPEX.

# Modeling Results

## Models Applied

***Linear Regression***

Linear Regression was applied since there was a strong co-relation observed based on heat map created for the dataframe. The objective was clear that there was a need to determine which variable had highest impact on the price of Weekend ticket price in order to achieve the objective of increased revenues.

**Performance Metrics**

In order to choose the best model using the available dataset, Explained Variance and Mean Absolute Error was calculated (see Fig a.)

## Important considerations before building model

- Duplicate and null values were handled appropriately as a part of data cleaning/ wrangling.
- A lot of outliers were found, but considering the size of the dataset which is small, they were not removed to avoid removing relevant data points that contribute to the model findings.
- Multi collinearity was observed for certain variables after building a heatmap indicating co-relation between all variables in the dataset. Those variables were removed.
- Linear relationship was observed between variables and considering the scope of project, **'Adult Weekend price'** was chosen as the dependent variable, considering its high impact on revenues.

# Modeling Results (Contd)

| Model | Explained Variance | Mean Absolute Error | Features Dropped |
|---|---|---|---|
| Model 1. | 0.92 | 5.46 | - |
| Model 2. | 0.91 | 5.77 | 'state' |
| Model 3. | 0.92 | 5.70 | 'state','summit_elev','base_elev |

**Fig a.**

## Chosen Model

- Going by the statistical findings, Model 1 is the best model since due to **highest explained variance** and **lowest mean absolute error**.
- Considering different characteristics, it is important for Big Mountain management to be able to determine a model that is able to predict with highest confidence what variables/ factors would
    - Drive ski resort lift ticket prices
    - Minimize error rate
- If you note, all 3 models are almost leading to same % confidence and the error is also somewhat in the same range. But going by the metrics as indicated in above table, Model 1 could be chosen.
- But you will see that once training data is created and model evaluated on test dataset for 'Big Mountain Resort', we won't need 'State' as a variable any more and it would be okay to choose **Model 2.**

# Modeling Results

*Predictions on the test dataset*

| Variables | Coefficients |
|---|---|
| AdultWeekday | 20.290758 |
| base_elev | 5.069323 |
| summit_elev | 4.405048 |
| averageSnowfall | 1.471846 |
| quad | 1.255350 |
| surface | 1.000788 |
| fastQuads | 0.963656 |
| triple | 0.904347 |
| NightSkiing_ac | 0.696030 |
| total_chairs | 0.609234 |
| Runs | 0.567704 |
| daysOpenLastYear | 0.555854 |
| projectedDaysOpen | 0.477440 |
| LongestRun_mi | 0.468372 |
| vertical_drop | 0.380968 |
| SkiableTerrain_ac | 0.343201 |
| fastSixes | 0.325714 |
| yearsOpen | 0.240124 |
| trams | 0.234007 |
| Snow Making_ac | 0.233699 |
| double | 0.156957 |
| TerrainParks | 0.112444 |
| fastEight | 0.037744 |

**Fig b.**

### Impact on Weekend ticket price

For **'Big Mountain Resort'**, with appropriate model application, training data development and evaluation on test dataset, it was found using the coefficients of co-relations determination that

- **Summit Elevation** - Elevation in feet of the summit mountain at the resort
- **Base Elevation** – Vertical change in elevation from the summit to the base in feet

are the variables that have highest impact on **Weekend Ticket price** (see Fig b.)

# Modeling Results

*Predictions on the test dataset*

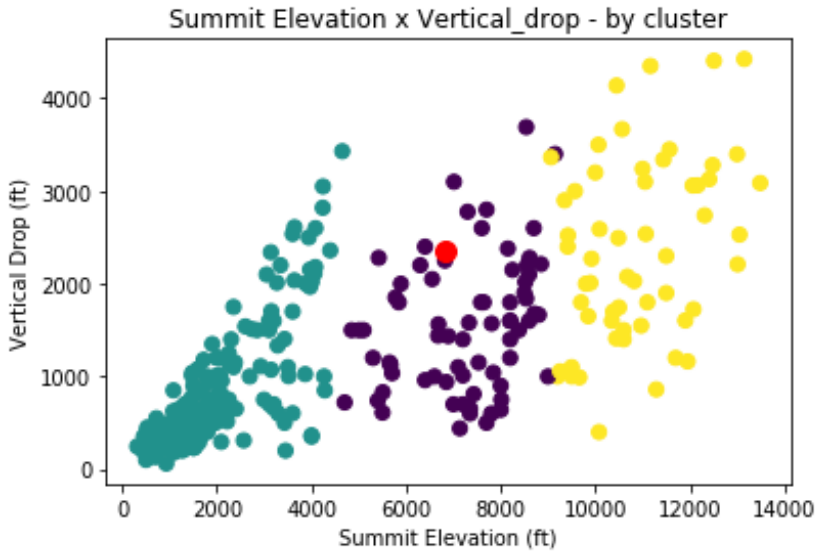


**Fig c.**

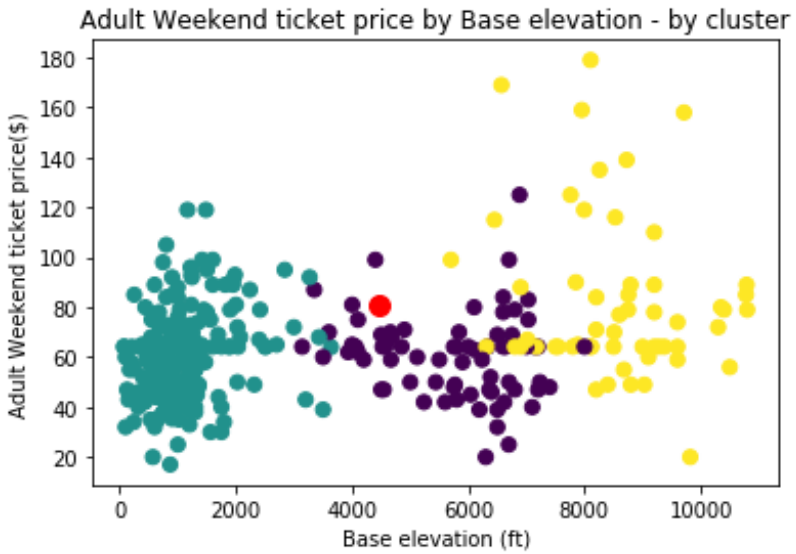


**Fig d.**

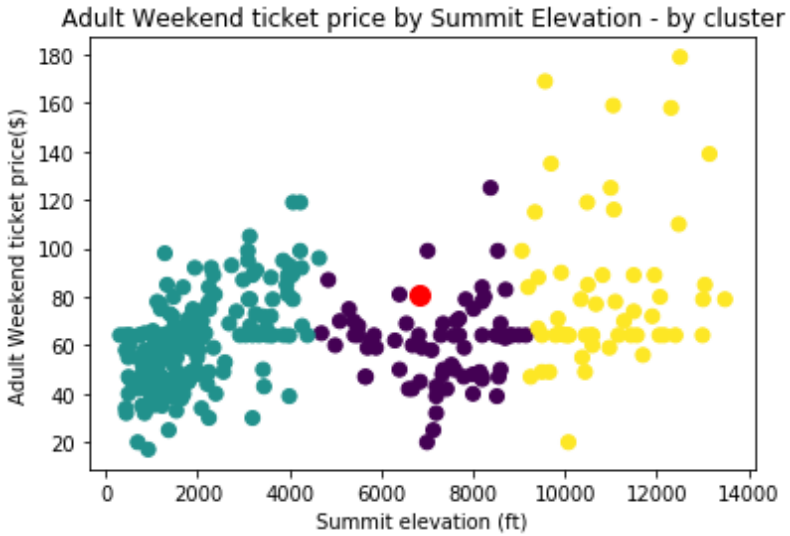


**Fig e.**

# Modeling Analysis

**Scatter plot findings**

The most highly co-related variables to Weekend ticket price $ are

- Adult Weekday ticket price (which is little obvious, hence there is no plotting done for it)
- Base elevation
- Summit elevation

Going by plotting scatter plots for above variables, it can be seen that

- 'Summit elevation and Base elevation' are linearly co-related indicating that resorts with high elevation are more likely to see a deep base elevation too (possibly related to terrain area (not indicated in the graph) , 'Big Mountain Resort' as indicated by the 'red' dot in Fig c is:
  - On the higher side in terms of vertical drop (as compared to the more dense clusters in green and purple)
  - Midway in terms of base elevation
- Both graphs (Fig d and e) are not necessarily linearly placed
  - Big Mountain Resort (as you see in the 'red' dot) falls in the middle cluster (in purple)
  - Price is moderately placed for the ski activities/ features it offers (Both Base and Summit elevation contribute heavily to those activities)

**Graph Conclusion** – As indicated in the 'Modeling Results' slide earlier, it would be wise to increase the weekend ticket price by $10, since there are fewer resorts that offer such great Summit and Base elevation and yet moderately priced

# Summary and Conclusion

**Summary and Conclusion**

- Data Cleaning is a very important factor to ensure any unwanted data is cleaned for data consistencies Ex: There were a lot of null values that had to be handled to avoid NaN errors.
- There could be duplicate data and variables which need to be cleansed in order to avoid any unnecessary bias and multicollinearity that could cause precision compromise to model performance.
- Choosing the right model is important to improve model performance for better predictions.
- 'Big Mountain Resort' was not present in the original data frame possibly due to type errors/ manual data entry, hence it is important to perform necessary data manipulation using data in hand.
- Heatmap and histogram plots were important in the Exploratory Data Analysis phase to identify the type of datasets, relationship between variables, outliers and decisioning in terms of what % of dataset should be retained.
- Building Scatter plots and applying clustering using methods like K-means provides a great visual on where the 'Big Mountain Resort' dataset fared against other resorts w.r.t Average ticket pricing when plotted against the highly co-related features like summit elevation and drop elevation.
- Similar prediction and model findings can be done for the other (following) variables; but is not part of the scope of this project
    - Average Weekday Ticket price
    - Projected Days to be opened
        - All these variables together would contribute to 9.2% profit margin goal which can't be accurately calculated (but can be definitely predicted if there is more information on revenues)