

Guided Capstone Project Report

In [14]:

```
from IPython.display import Image
from IPython.core.display import HTML
Image(url= "https://github.com/p5149247263/1/blob/main/bm.jpg?raw=true" )
```

Out[14]:



☆ Problem Statement:

What opportunities exist for Big Mountain Resort to effectively develop and implement a new pricing strategy that can maximize capitalization in their facilities investments to offset their recent additional operating cost by \$1,540,000 this season.

☆ Resort Story:

Big Mountain Resort offers spectacular views of Glacier National Park and Flathead National Forest, with access to 105 trails. Every year about 350,000 people ski or snowboard at Big Mountain. The business expressed a desire for some guidance on how to select a better value for their ticket price. by considering number of changes hoping to reduce cost without reducing ticket price or increasing ticket price.

☆ Current Resort's Pricing Strategy:

For Big Mountain Resort to base their pricing mainly on just the market average won't be enough to maximize their capitalization investment and can't be sustainable to gain an edge over the competition.

☆ Modeling scenarios:

Big Mountain Resort has been reviewing potential scenarios for either cutting costs or increasing revenue (from ticket prices). Ticket price is not determined by any set of parameters; the resort is free to set whatever price it likes. However, the resort operates within a market where people pay more for certain facilities, and less for others. Being able to sense how facilities support a given ticket price is valuable business intelligence. This is where the utility of our model comes in.

The business has shortlisted some options:

- Permanently closing down up to 10 of the least used runs. This doesn't impact any other resort statistics.
- Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage
- Same as number 2, but adding 2 acres of snow making cover
- Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres

The expected number of visitors over the season is 350,000 and, on average, visitors ski for five days. Assume the provided data includes the additional lift that Big Mountain recently installed.

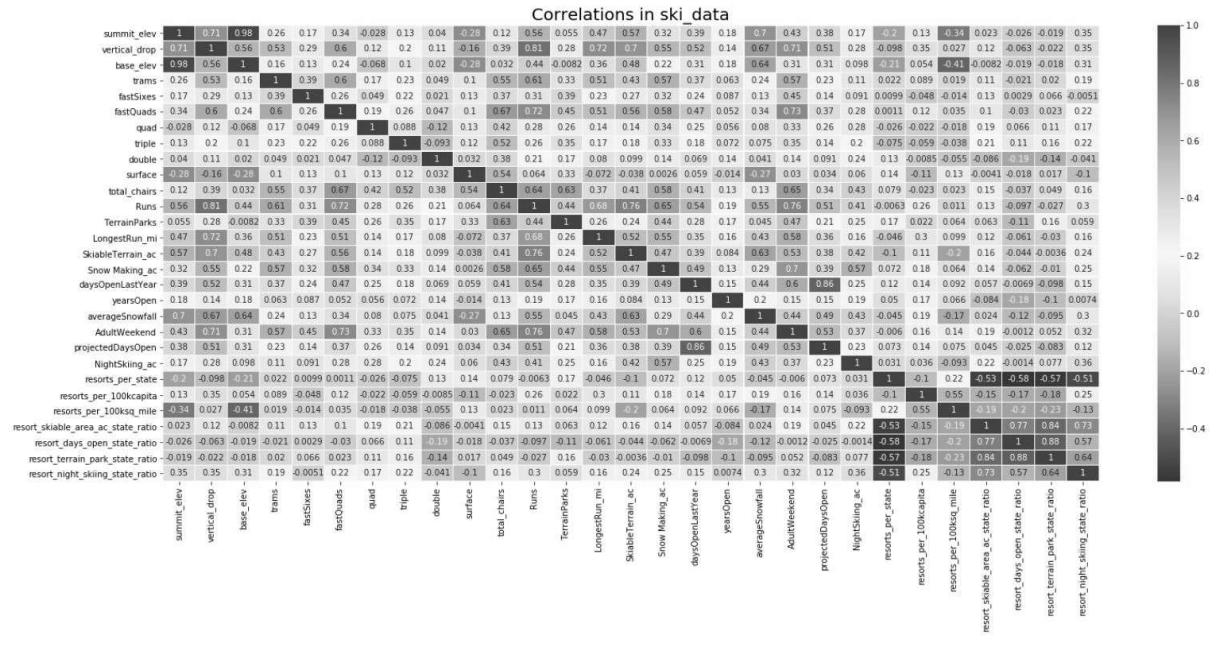
☆ Feature Engineering:

- **Feature Correlation heatmap:** a great way to gain a high level view and identify patterns of relationships amongst the features:
 - summit and base elevation are quite highly correlated. This isn't a surprise.
 - AdultWeekend ticket price, we see quite a few reasonable correlations. fastQuads stands out, along with Runs and Snow Making_ac.
 - The last one is interesting. Visitors would seem to value more guaranteed snow, which would cost in terms of snow making equipment, which would drive prices and costs up.
 - As well as Runs, total_chairs is quite well correlated with ticket price. This is plausible; the more runs you have, the more chairs you'd need to ferry people to them! Interestingly, they may count for more than the total skiable terrain area. For sure, the total skiable terrain area is not as useful as the area with snow making. People seem to put more value in guaranteed snow cover rather than more variable terrain area.

In [15]:

```
Image(url= "https://github.com/p5149247263/1/blob/main/corr.jpg?raw=true")
```

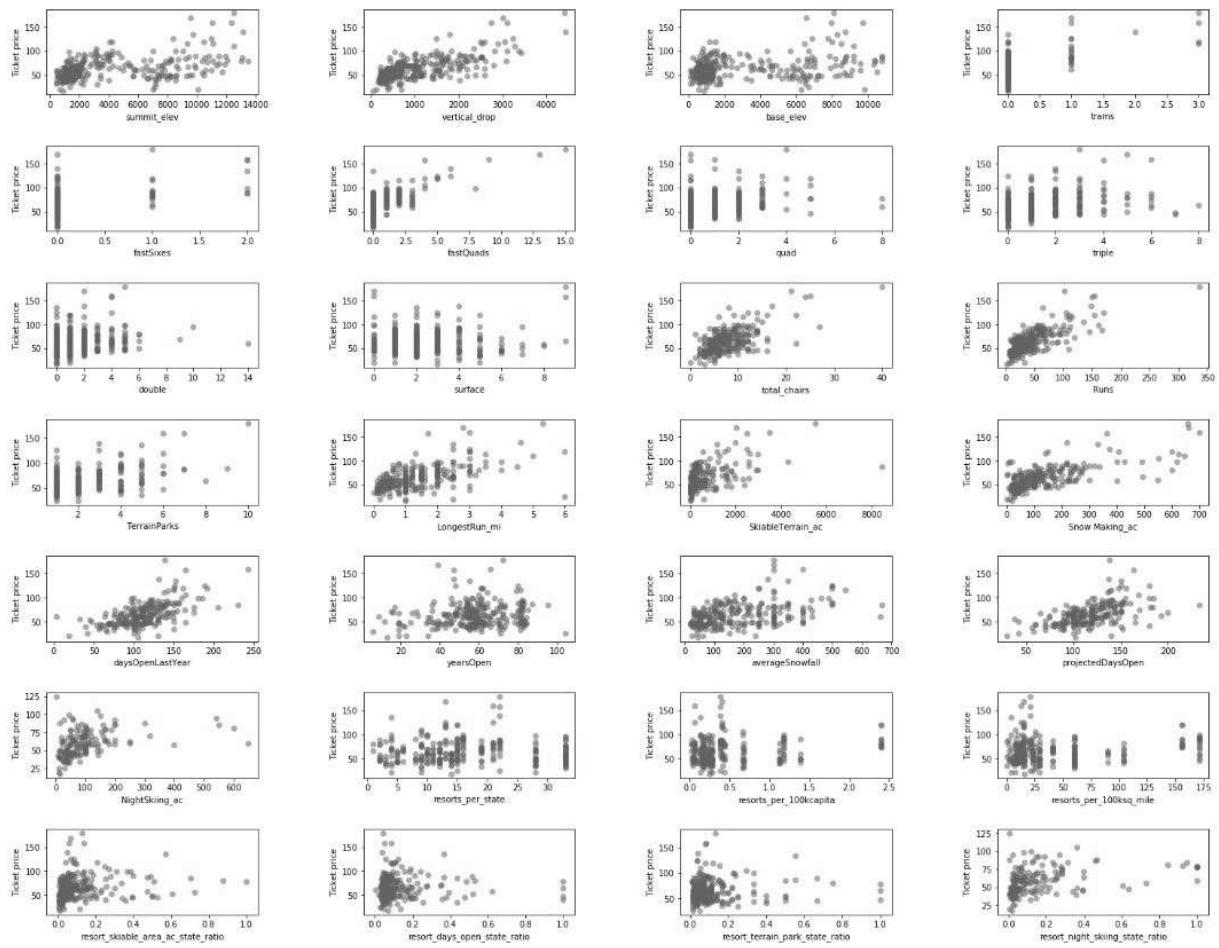
Out[15]:



- Feature Correlation Scatterplots:** Correlations, particularly viewing them together as a heatmap, can be a great first pass at identifying patterns. But correlation can mask relationships between two variables
 - There's a strong positive correlation with vertical_drop.
 - fastQuads seems very useful. Runs and total_chairs appear quite similar and also useful.

```
In [16]: Image(url= "https://github.com/p5149247263/1/blob/main/pairplot.jpg?raw=true")
```

Out[16]:



☆ Pre-Processing and Training Data:

- **Linear Model:** In the process of building the linear model, missing values were imputed with the median and mean values. If ticket prices were predicted using the linear model, they would be off by about \$9. However, the initial linear model was overfitting and needed to be adjusted by the number of features. Through cross-validation, the value of k was set to eight features to focus on: vertical_drop, Snow Making_ac, total_chairs, fastQuads, Runs, LongestRun_mi, trams, and SkiableTerrain_ac. These features fit our initial assumptions from EDA.
- **Random Forest Model:** Was developed. Like the linear model, missing values were imputed with the median and mean values. While imputing the median was helpful, it was not helpful to scale the features. The random forest model revealed that the top four features to consider are fastQuads, Runs, Snow Making_ac, and vertical_drop.
- **Final Model Selection:** After testing both the linear model and random forest model, the project will be moving forward with the forest regression model. Comparison of the two demonstrated that performance on the test set was consistent cross-validation results. Additionally, the cross-validation mean absolute error was lower using the random forest regressor.

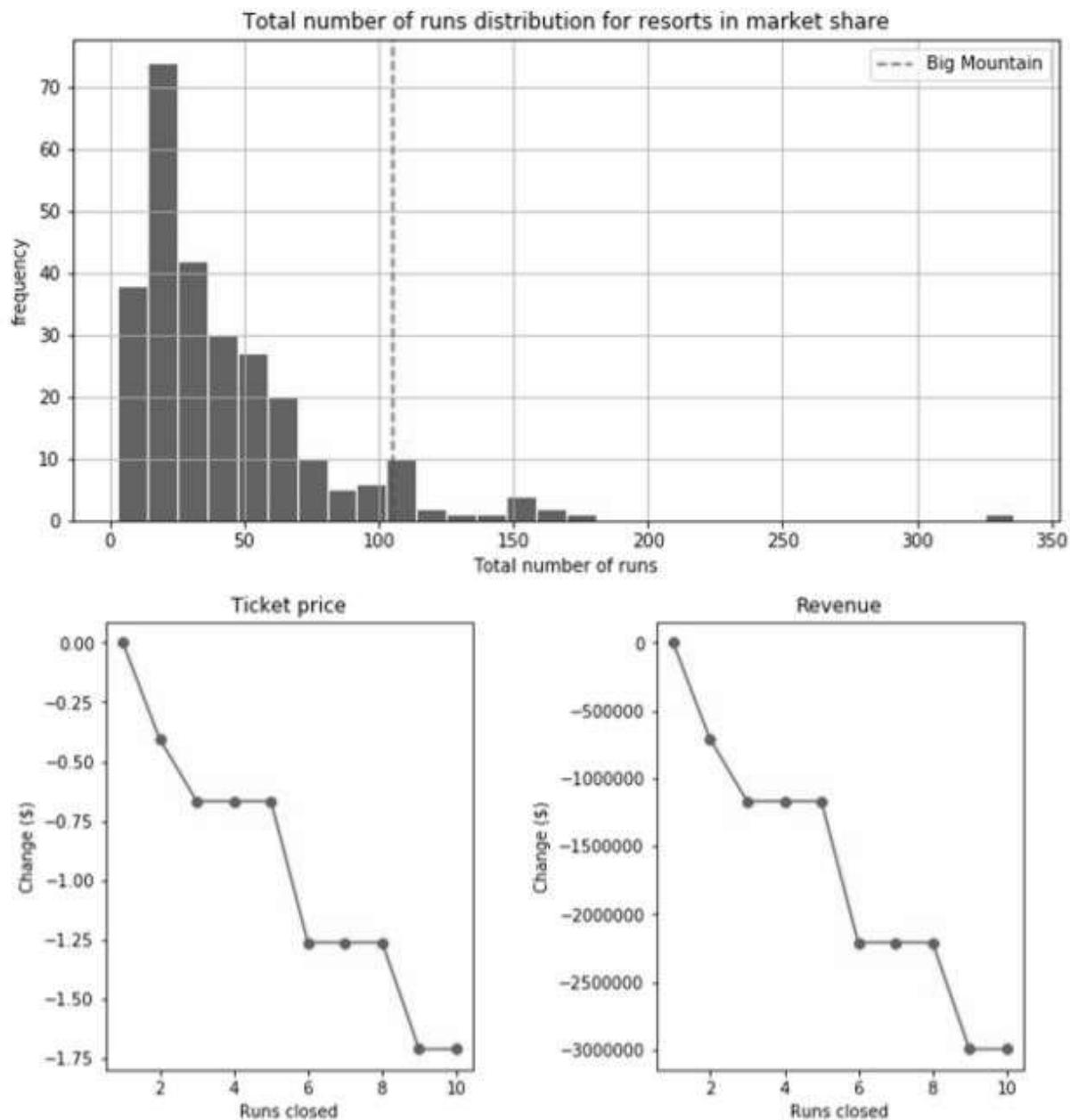
☆ Recommendations:

- Our Model suggests that Mountain Resort's ticket price is lower than the predicted model by 16.31%, and the resort have many potential scenarios for either cutting costs by closing runs or increasing ticket price by increasing vertical drop, adding acres snow making or increasing the longest run.
- Increasing the vertical drop by 150 ft would increase the ticket price by 10.44% from 81 to 89.46, resulting in revenue increase by \$14,811,594.
- Adding 2 acres of snow making would increase the ticket price by 12% from 81 to 90.75, resulting in revenue increase by \$17,068,841.
- When it comes to closing up to 10 used Runs, our Model predicted the following:
 - Closing one run will have no impact on Ticket price or revenue.
 - Closing 2 runs reduce support for ticket price and so revenue by 0.4 and 750,000 respectively.
 - Closing down 3 runs, it seems they may as well close down 4 or 5 as there's same loss in ticket price and revenue by 0.67 and 1.250M respectively.
 - Closing 10 runs reduce support for ticket price and so revenue by 1.71 and 3M respectively.
 - Because we don't know the operating cost per used run, we can't determine how much cost saving will be offset the loss in revenue after closing more than one run.

In [21]:

```
Image(url= "https://github.com/p5149247263/1/blob/main/analysis.jpg?raw=true")
```

Out[21]:



☆ Conclusion:

After applying our Model for ski resort ticket price and leverage it to explore Big Mountain Resort's potential scenarios for increasing revenue, we can conclude that:

- The best scenario where we managed to gain the highest revenue increase possible was by increasing the vertical drop by 150 ft, adding one Chair Lift, adding one run and adding 2 acres of snow making cover. This scenario has increased ticket price by 12% from 81 to 90.75, resulting in a bottom-line increase by 15,528,841 (*After deducting operating costs = 1.54M*).
- Due to lack of data in regards of operating cost per used run and weekdays ticket price, our model cannot recommend closing down used runs or implementing a dynamic ticket pricing.

In [22]:

```
Image(url= "https://github.com/p5149247263/1/blob/main/conclusion.jpg?raw=true")
```

Out[22]:

