**Problem Statement**

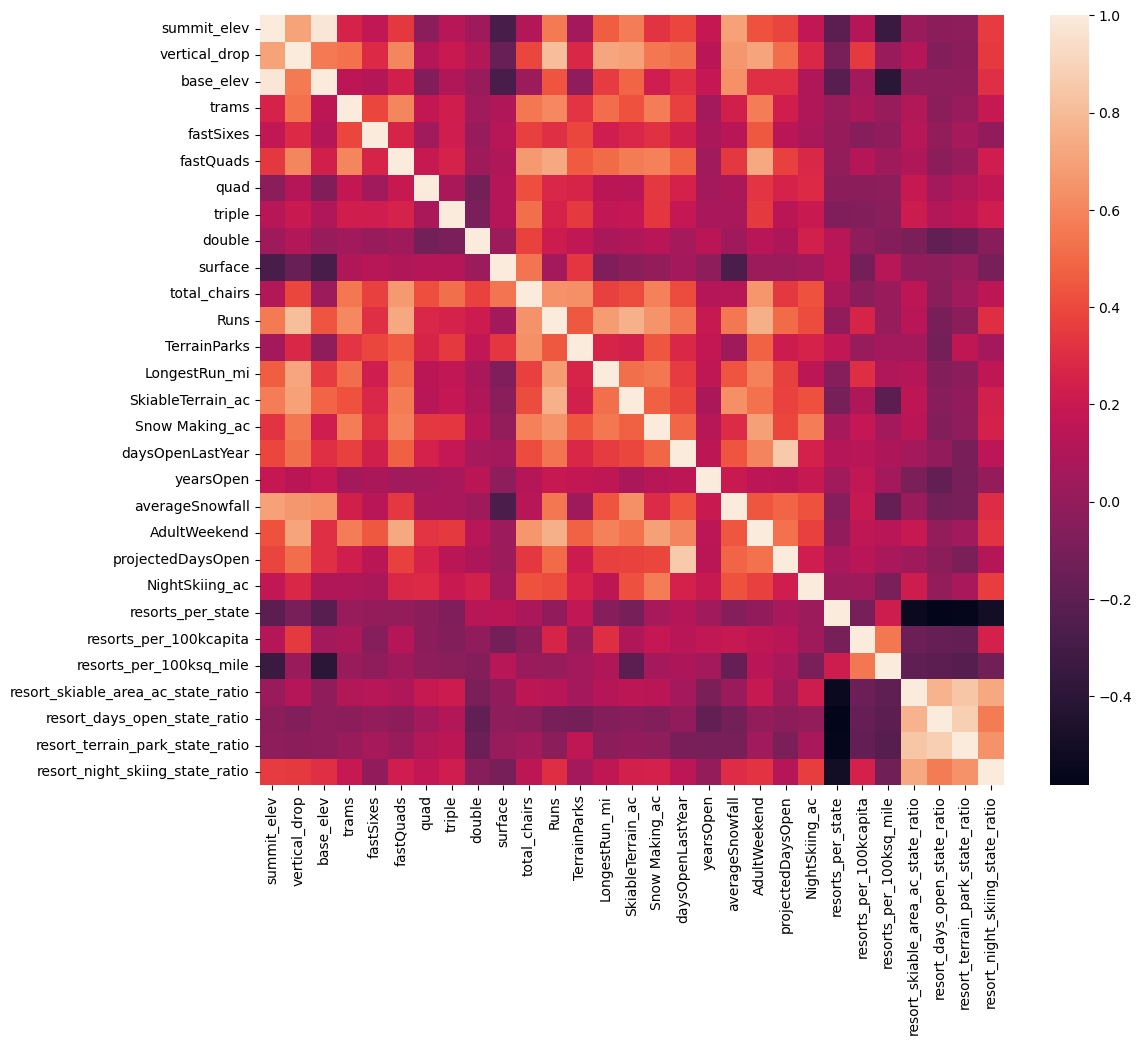
Big Resort Mountain wants to find out how to maximize its profits and cut costs through finding the key resort facilities and utilize more data-driven pricing strategies with proper facility modifications.

**Data Wrangling**

The dataset provided for analysis includes 23 features for evaluation with the ticket price as the feature of interest and also includes data from around 330 different resorts. With this comes some irrelevant and missing data that was handled appropriately.

**Exploratory Data Analysis**

After cleaning and sorting the data, EDA is performed to find relevant characteristics of the data and help understand the data. During EDA, the data was visualized using states as the metric, various features of interest were scrutinized, relevant new features were added, and the data was scaled and presented in a high dimensional format to find relationships between the features. PCA transformation is utilized to derive patterns in the data. The following is the feature correlation heatmap containing a high level view of relationships amongst the features:



The key features of interest are fastQuads, Runs, Snow Making\_ac, and total\_chairs.

**Model Preprocessing**

When beginning to train the data, a train/test split was established of 70/30 respectively. Different models were then tested and compared to the base “mean” model. Metrics such as R-squared, MAE, and MSE are used to measure each models’ performance. The Linear Regression and Random Forest models chosen to be evaluated.

**Algorithms Used**

A pipeline with various methods are used to streamline a model. For example, in the Linear Regression model, the SimpleImputer algorithm is used to form a base model, the StandardScaler is used to scale the data, SelectKBest finds the k best features, and then the Linear Regression model is applied. After this cross validation can be applied to all of the data to ensure coverage of the dataset.

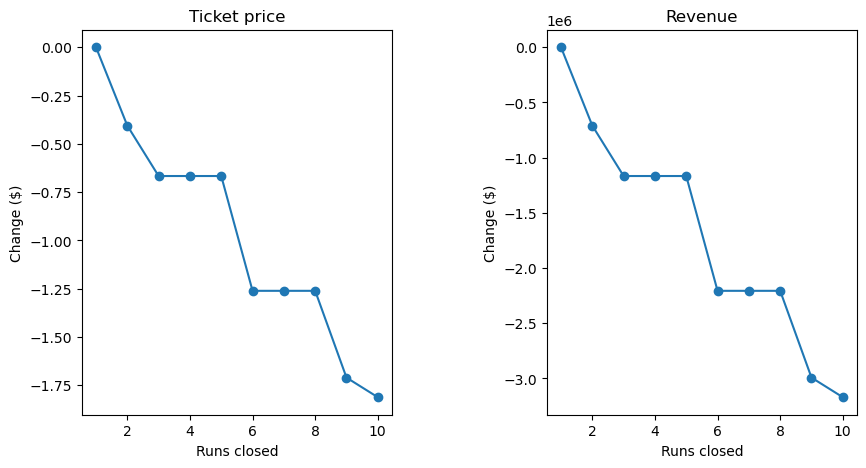
**Winning model**

The winning model is the Random Forest model with an MAE of 9.54 vs the Linear Regression models’ MAE of 11.79.

**Pricing recommendation**

Big Mountain currently charges $81, but our model predicts that the suggested price should be $95.87 given its facilities. Even with the mean absolute error, it's undoubtedly the best course of action to increase the ticket price assuming other resorts are following good pricing strategies.

**Conclusion**

It's worth noting that it may be worth it to close runs. Depending on costs saved from the closures, it may be worth investigating if run closures are worth it. The following shows the projected loss in ticket pricing support depending on how many runs are closed:  


On the basis that each visitor is buying on average 5 day tickets with the current 350,000 annual visitors, adding an additional run that would increase the vertical drop by 150 feet and requiring an additional chair lift would support a ticket price increase of $2 and would expect to amount to an increase of revenue of $3,474,638. Given that the cost of adding an additional chair only increases operation costs by $1,540,000, it is recommended that Big Resort Mountain proceeds with this facility upgrade.

**Future scope of work**

This model heavily relies on the fact that other resorts have proper pricing strategies. Some other relevant data necessary is pricing on additional snow making, chair construction costs, etc. It's also relevant to find how the number of visitors would change if a price increase were to occur. Big Mountain is already fairly high in pricing, but the modeled price is even higher because of Big Mountain’s exceptional performance on most key features.