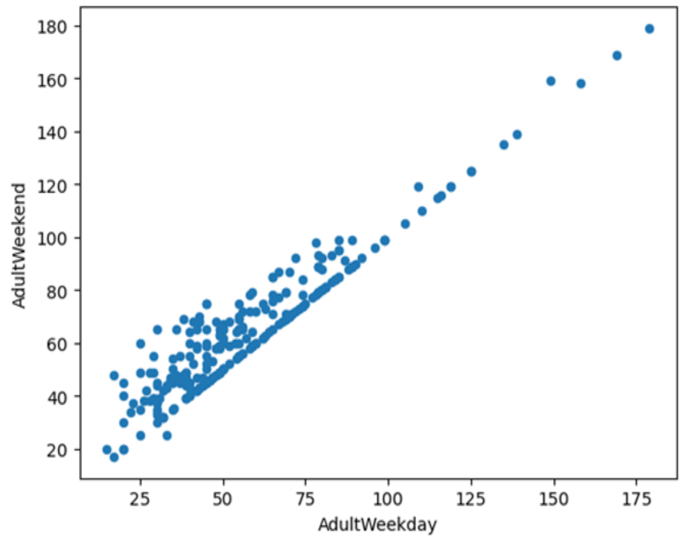
Guided Capstone Project Report [Big Mountain Resort]

Problem Statement

What opportunities exist for Big Mountain to optimize ticket pricing and/or to cut costs based on important facilities within the resort before the start of next ski season to increase the business' revenue by 'above average price' in its market segment?



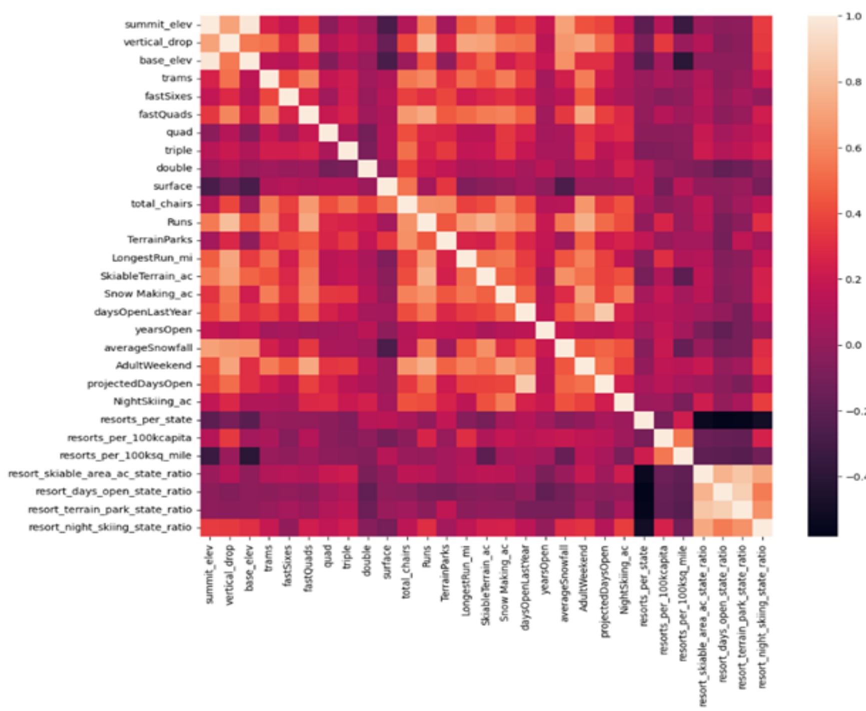
Data Preprocessing and Data Wrangling

The Data Science Method was used in the project and during the processing 82% of the resorts have no missing ticket prices, 3% are missing either weekend or weekday price, and 14% are missing both. 14% of the rows with no price data are dropped because price is our target. Additional US state population and size data with which to augment the dataset.

The relationship between weekday and weekend prices indicated a clear line where weekend and weekday prices are equal. However, in sub $100 resorts, weekend prices are higher than weekday prices (see scatterplot). Hence, I decided to model the adult weekend prices.

Data Analysis

A heatmap was applied to indicate the correlations between features. As such there was significant correlation between Adult Weekend ticket prices and total_chairs, Runs, fastQuads, Snow Making_ac, trams, Longest Run_mi, and vertical_drop.



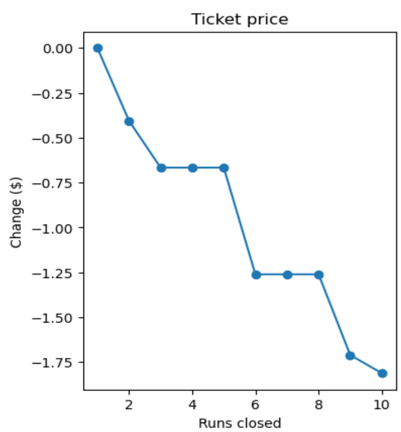
Price Modelling and Recommendations

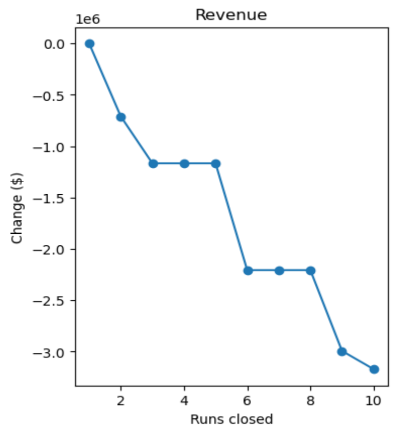
The following function was written to build the winning pricing model.

```python
def predict_increase(features, deltas):
    """Increase in modelled ticket price by applying delta to feature.

    Arguments:
    features - list, names of the features in the ski_data dataframe to change
    deltas - list, the amounts by which to increase the values of the features

    Outputs:
    Amount of increase in the predicted ticket price
    """

    bm2 = X_bm.copy()
    for f, d in zip(features, deltas):
        bm2[f] += d
    return model.predict(bm2).item() - model.predict(X_bm).item()
```

Four scenarios for potential price modelling were identified. The best model was scenario 2 which recommends adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift, without additional snow making coverage. This would increase support for ticket price by $1.99 and bring in an estimated amount of $3,474,638 over the skiing season.

Scenario 1: Permanently closing at most 10 of the least used runs. This does not impact any other resort statistics.





The model says closing one run makes no difference. Closing 2 and 3 successively reduces support for ticket price and so revenue. If Big Mountain closes 3 runs, it seems they may as well close 4 or 5 as there's no further loss in ticket price. Increasing the closures to 6 or more leads to a large drop.

Scenario 2: Adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift, without additional snow making coverage will increase support for ticket price by $1.99 and bring an estimated revenue of $3,474,638 over the ski season.

Scenario 3: Adding a run, increasing the vertical drop by 150 feet, installing an additional chair lift, and an additional 2 acres of snow making coverage. The impact of this scenario yields the same results as scenario 2.

Scenario 4: Increasing the longest run by 0.2 miles and guaranteeing its snow coverage by adding 4 acres of snow making capability. The predicted results indicate that there is no difference whatsoever. Although the longest run feature was used in the linear model, the random forest model (the one we chose because of its better performance) only has longest runway down in the feature importance list.

Conclusion and Future Work: Data quantity was sufficient but with limitations. For example, it is not possible to tell, from the present data, why the other resorts could undercharge or overcharge to use their facilities. Hence, additional data such as operating costs of the resorts would be relevant. However, based on current data, the business leaders in Big Mountain resort should elect scenario 2 because it presents the best model. Further, I recommend following a model deployment process. This way the model will be easier for business analysts to test new parameters in the future.