Guided Capstone Project Report

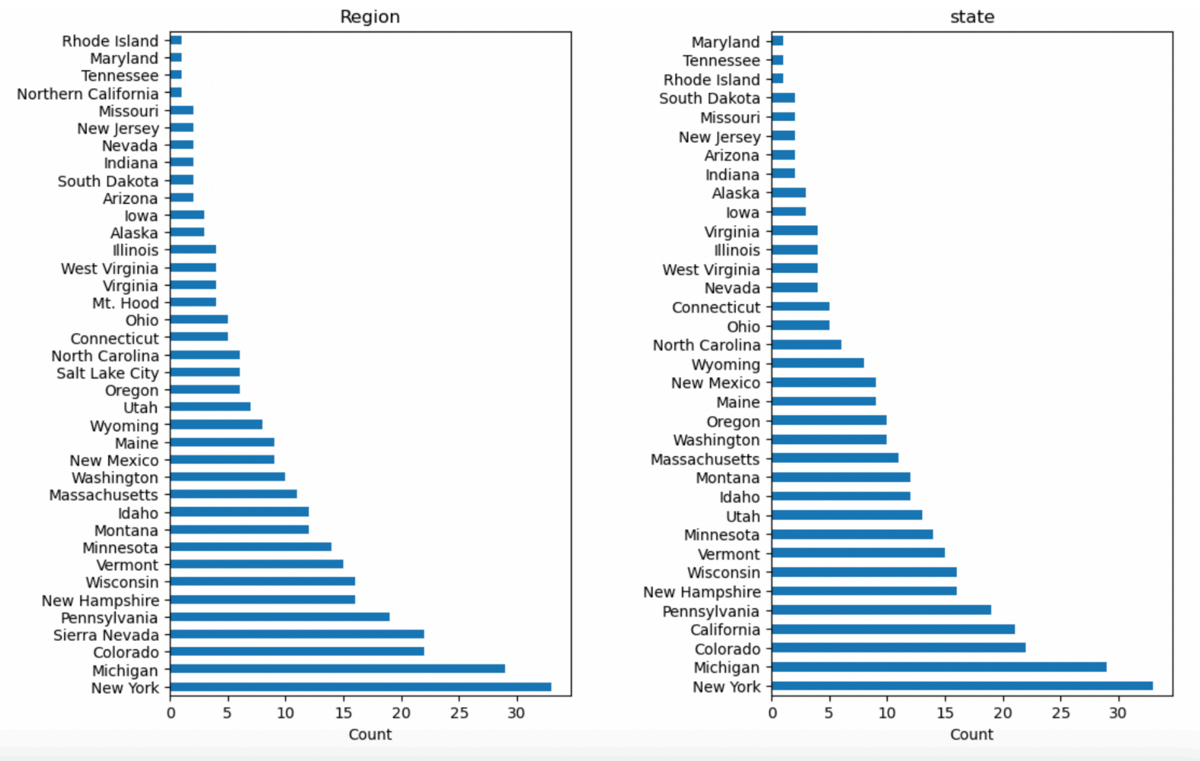
Problem Statement: How should Big Mountain Resort determine the ticket prices of this season and arrange the costs to be able to cover the expenses of new chair lift which increases operation cost by $1,540,000?

Context:

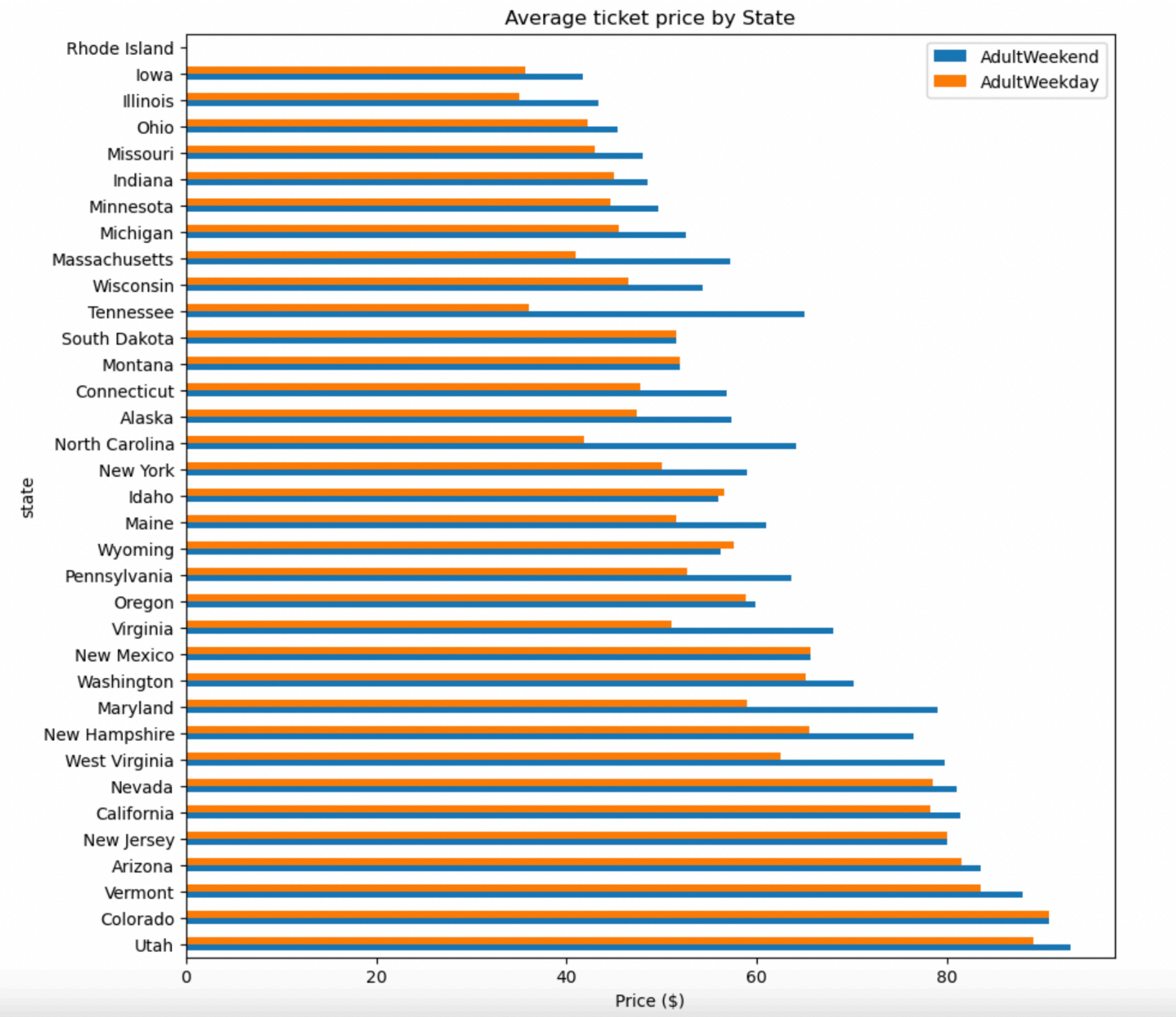
Big Mountain Resort installed a new chair lift to be able to increase the distribution of visitors across the mountain. As a result, they have additional operating cost of $1,540,000 this season. They need to increase their profit in a way that can cover this additional cost. Their strategy is either increase the ticket price or lower the costs.

Summary of Data Wrangling and Exploratory Data Analysis:

In the data set, there are 330 unique resorts with 27 features including information regarding our resort which is "Big Mountain Resort". The data set includes 38 unique regions and 35 unique states, meaning a few states are split across multiple named regions.



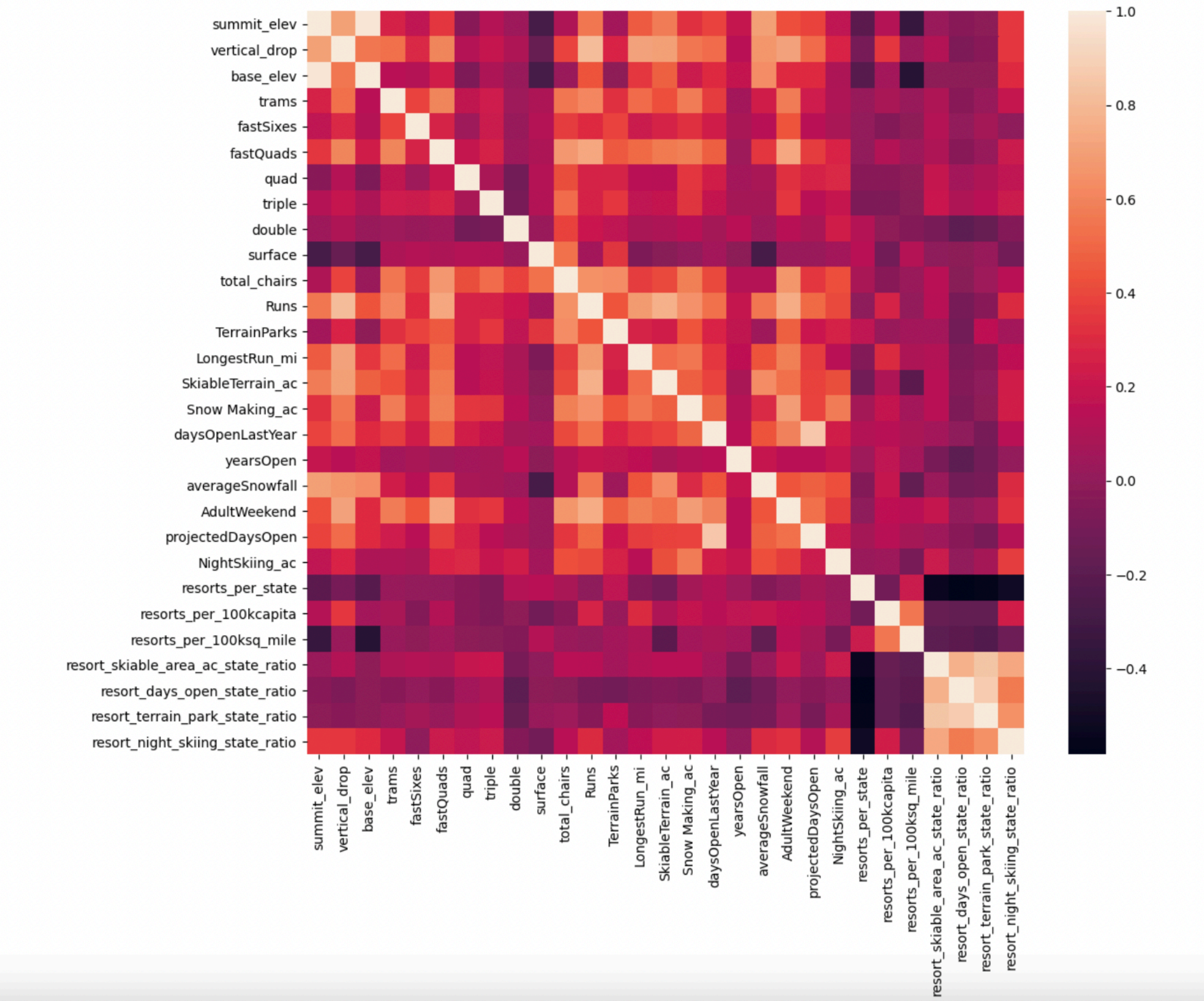
According to the plot, the majority of resorts are in New York. Our resort is in Montana, which comes in at 13th place.



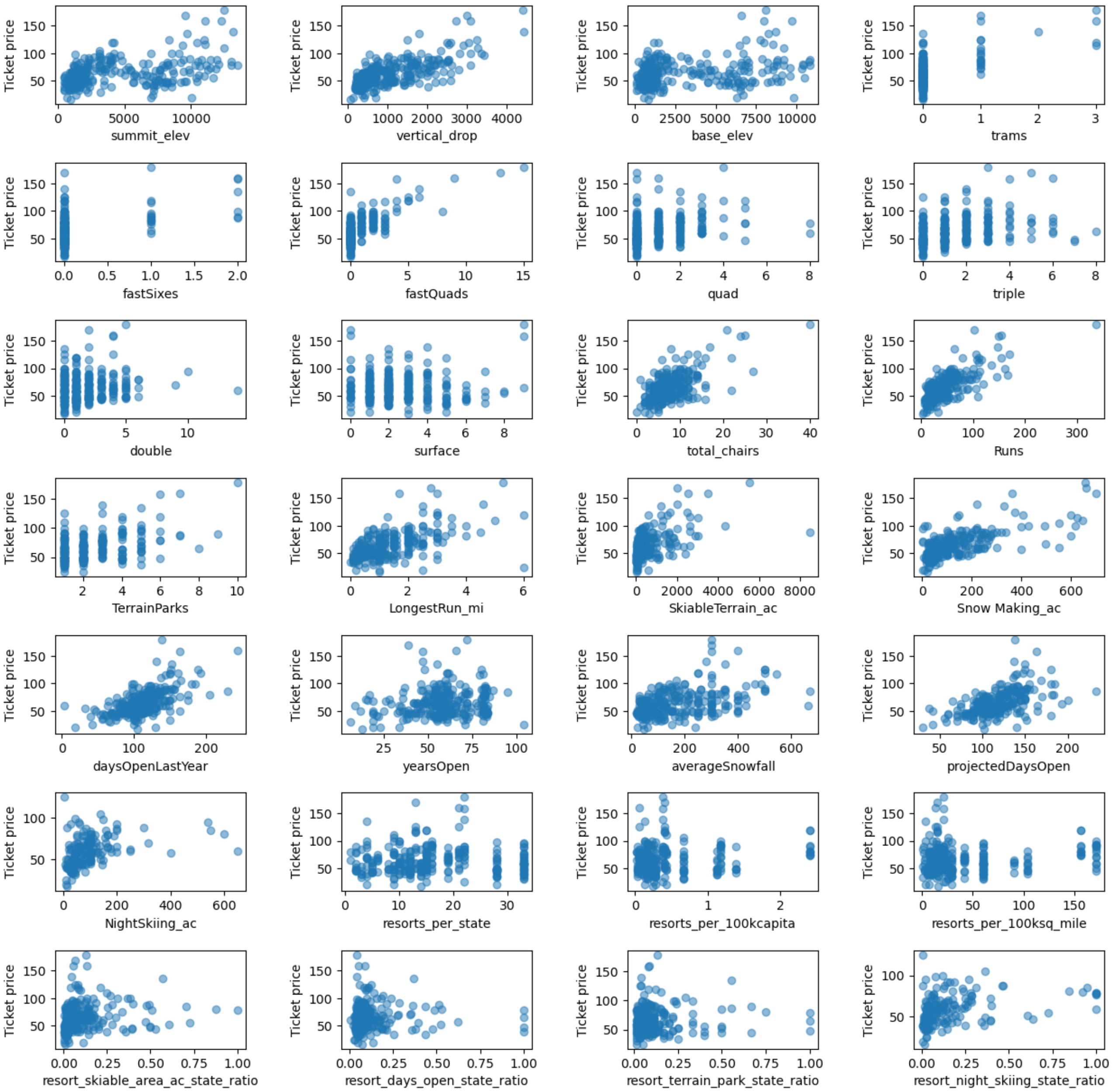
According to average ticket price by state plot, Utah is the one with the most expensive ticket prices.

The dataset provided for analysis contained both numerical and categorical features. The numerical features included variables like the number of resorts per state, total skiable area per state, total days open per state, total terrain parks per state, total night skiing area per state, state population, and state area in square miles. The categorical feature was the "state" variable. In addition to those, resort density metrics, such as resorts per 100,000 capita and resorts per 100,000 square miles, were calculated to provide additional insights.

Principle Component Analysis (PCA) was used to disentangle the interconnected web of relationships. A heat map was created to gain high level view of relationship amongst the feature.



According to this heat map, summit and base elevation are quit highly correlated. The interesting inference might be the positive correlation between the ratio of night skiing area with the number of resort per capita which means that when resort are more densely located with population, more night skiing is provided. For the ticket price, there are some reasonable correlations, such as the number of fast quads, runs, and chairs. Those are quite plausible, since the more runs you have, the more chairs you would need to ferry people to them. The interesting correlation is with snow making area, since visitors would seem to value more guaranteed snow.



Scatterplots were created to visualize relationships between states based on features and average ticket prices. In the scatterplots, there are some high correlations that are clearly picking up on. There's a strong positive correlation with vertical drop, number of fast quads, number of runs, and total chairs with ticket price. These are quite predictable outcomes. The interesting outcome might be with the resort density per 100k capita. When the value is low, there is quite a variability in ticket price, although it's capable of going quite high. Ticket price may drop a little before then climbing upwards as the number of resorts per capita increases. Ticket price could climb with the number of resorts serving a population because it indicates a popular area for skiing with plenty of demand.

The lower ticket price when fewer resorts serve a population may similarly be because it's a less popular state for skiing. The high price for some resorts when resorts are rare (relative to the population size) may indicate areas where a small number of resorts can benefit from a monopoly effect. It's not a clear picture, although we have some interesting signs.

Summary of Modeling:

The modeling phase involved building and evaluating predictive models for ticket prices. The following steps were taken:

- Data was split into training and test sets (70% train, 30% test).
- A baseline model using the mean ticket price was established.
- Various metrics, including R-squared, mean absolute error (MAE), and mean squared error (MSE), were used to evaluate model performance.
- Missing values were imputed using mean and median values, and feature scaling was performed.
- Feature selection was conducted using the k best features approach.
- Linear regression and random forest regression models were trained and evaluated using cross-validation.
- Hyperparameter tuning for the random forest model was performed using GridSearchCV.

<u>The key findings from the modeling phase include:</u>

- The baseline model using the mean ticket price showed that Big Mountain Resort might be overcharging, as the model predicted a higher price.
- Feature importance analysis indicated that vertical drop, fastQuads, runs, snowmaking area, and total chairs were important features in predicting ticket prices.
- Random forest regression outperformed linear regression in terms of mean absolute error and variability.
- It was determined that the dataset size was sufficient for the analysis, and there was no need for additional data.

Summary of Scenario Analysis:

Several scenarios were analyzed to explore potential cost-cutting or revenue-increasing strategies for Big Mountain Resort. The scenarios included closing runs, increasing vertical drop, and expanding snowmaking coverage. The key findings from the scenario analysis include:

- Closing a small number of runs had little impact on ticket prices, but closing a significant number could lead to a substantial drop.
- Increasing vertical drop was found to have a positive impact on ticket prices.
- Expanding snowmaking coverage had no significant impact on ticket prices.
- Big Mountain Resort was found to be competitively positioned in terms of various features, despite its high ticket prices.

Recommendations:

Based on the analysis, the following recommendations are made:

- **Pricing Adjustment**: Big Mountain Resort could consider increasing its ticket prices slightly, as the modeling suggested potential room for an increase. However, any price increase should be carefully implemented to avoid deterring visitors.

- **Scenario Analysis**: The resort should carefully evaluate the scenario of increasing vertical drop, as it was shown to have a positive impact on ticket prices. This decision should be balanced with the cost of installing an additional chairlift.

- **Feature Enhancement**: Continue to focus on maintaining and enhancing key features that contribute to the resort's competitiveness, such as vertical drop, fastQuads, runs, and snowmaking capabilities.

- **Regular Data Analysis**: Regularly analyze visitor data, cost data, and competitor information to make informed pricing and operational decisions.

- **User-Friendly Model Access**: Develop user-friendly interfaces or dashboards for business analysts and decision-makers to explore different scenarios without the need for constant model retraining.

- **Data Collection**: Collect additional data on operating costs, visitor demographics, and preferences to improve the accuracy of future modeling and decision-making.

- **Monitor Performance**: Continuously monitor the performance of the resort and the impact of any changes in pricing or operations on revenue and profitability.

- **Business Leaders' Involvement**: Engage business executives in the decision-making process and ensure that they are aware of the modeling results and recommendations.

In conclusion, the analysis provides insights into potential pricing adjustments and operational strategies for Big Mountain Resort. The resort should carefully consider these recommendations in making important business decisions to enhance its competitiveness and profitability in the ski resort industry.