Summary of recommendations for ***Big Mountain Resort***

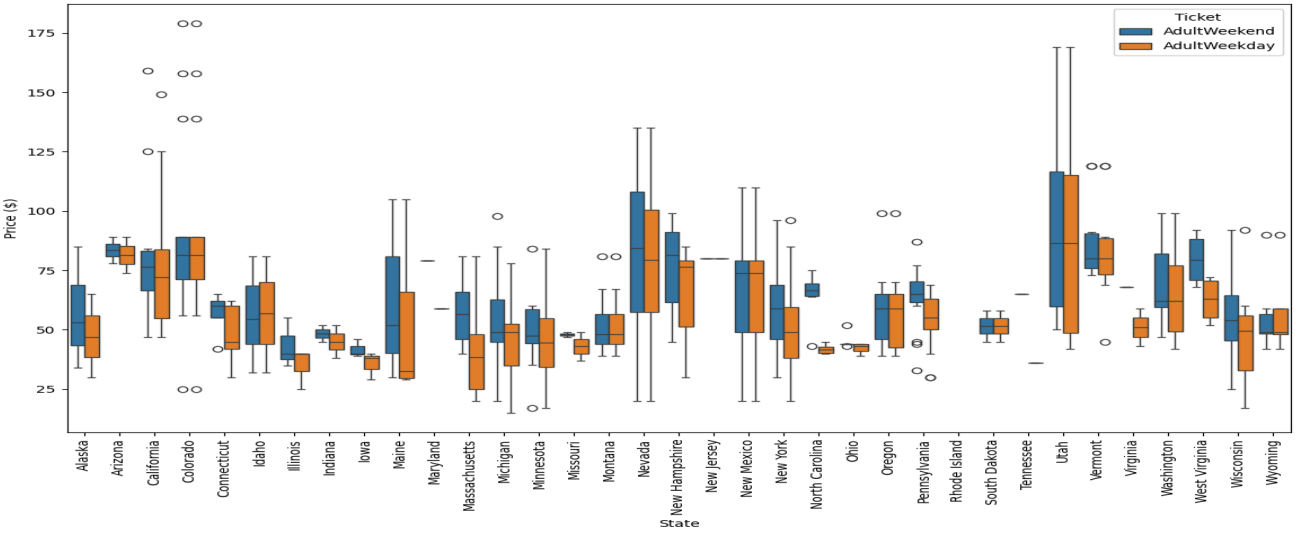
Problem Statement:

Big Mountain Resort wanted to alter it's pricing as the operating cost is increased due to additional chairs lift added in this season.They would like to change the price in this season, so would like to alter the price in 2 weeks to maintain their profit.

The Business given data had **330 Rows and 27 Columns** and 'Big Mountain Resort' was part of this ***ski_data*** data. In the ***ski_data***, there were many missing values such as null,NaN, inappropriate values for that Column etc.These feature columns which didn't have appropriate values were treated either by finding its original value if it's available are treated in the appropriate way as part of ***Data Wrangling*** process.

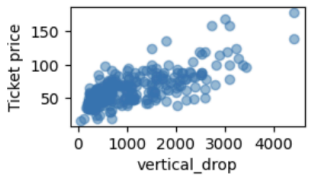
Certain additional data are needed further for the analysis or to better understand about the Skiing Resort location, the state in which it's present,to understand the dynamics of the ***ski_data***. The '**Population and area data for the US states**' was used to cross validate and update the Name/State/Region accordingly. So during this process Data Wrangling, Validation was done so that Data to be Analyzed is valid and useful and this Data was saved and used for Future Exploration and Analysis.

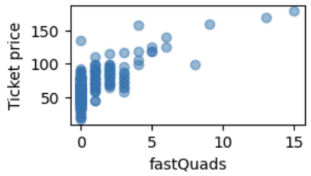
While doing Data Exploration, '*resorts_per_100kcapita'* which is based on per state State Resort and population and *'resorts_per_100ksq_mile'* which is based on per state Resort and state_area/sq_miles 2 additional columns were added to find the importance of the Resort within the State and within 100k radius miles. This helps us get the value/importance of the Resort on that particular location and helps us find a number of competitors within the radius.



While performing the BoxPlot, to see if there are any special case behaviors and to see the median pricing of each state to see what's the Ticket Pricing Trends. We found that ***Big Mountain Resort,*** which belongs to the **Montana** State, has AdultWeekendPrice and AdultWeekdayPrice to be the same, so it would be better to consider the **AdultWeekend** as a better pricing.

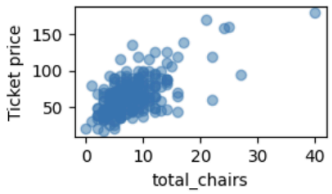
Scatter Plot was done to visualize, to find the Correlation between each feature with Price(Target) to find those Features which Positively impacted the Price and those which negatively impacted and those which did not make any effect.

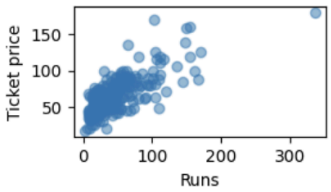




Out of which we were able to figure out that **Vertical_drop** , **fast_squad**, number of **runs**, **Total_chairs** have an impact on the Pricing. We can visually witness that as the value of these increases, the **Ticket Price** Increases.

So these features have **Positive Correlation** with **Ticket Price**. So these have a positive impact for the Price to increase, if any features of these are given more special attention our Ticket Prices value can be increased, Similarly if we remove these it will impact the Price and will cause the price to drop.





For a ML model, to predict perfectly, we treat the data pre-processed data which had handled all the missing values and other Data Validation, imputation either using Statistical Mean or Median and Scaling the numerical features so that all are within the range and is not having any large difference within the feature.

Though the initial data given by the Business was **330 Rows and 26 Columns**, the data is now reshaped to **277 Rows and 36 Columns** after the merge of ***ski_data*** with **state_summary**.

Typically in a ML model we Train with 70% and Test the model with 30% data which is not used for Training to check how the model predicts better with Training. We use Cross validation by doing multiple times of Training the 70% data to make sure we train well so that our predictions are better with minimal errors. Various models such as Linear Regression and RandomForest Regressor were used to Predict the ***Big Mountain Resort*** Pricing. The predicted Ticket Price performance is validated using various metrics such as - R2_Score, MAE. On comparing various the Prediction with the actual values, using RandomForest Regressor the error was less compared to others.

1. The **Big Mountain** Resort currently charge ***AdultWeekend*** = **81.00**
2. The model suggests **Big Mountain** Resort can be priced at **95.87**, whereas the actual price is **81.00**.
3. But based on the Best Features which are more important for a Skiing Resort,and the state which it belongs to, the Price **81** seems to be in the tail end(higher price). Though the skiable area and the number of runs are more, that is not accounting majorly for the price to increase.
4. The business has shortlisted some options: a. Permanently closing down up to 10 of the least used runs. Closing 1 run doesn't affect the cost, closing 2 to 5 runs affects the Ticketing Cost by -0.70, which would make the ticketing price to be around 80, and so on.

b. 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 - *This can increases support for ticket price by 1.99, Over the season, this could be expected to amount to 3474638*

c. Same as number 2, but adding 2 acres of snow making cover - *This can increase support for ticket price by 1.99, Over the season, this could be expected to amount to $3474638. But finally doesn't make much difference for increasing snow making cover for a very less area of about 2 acres makes difference*

d. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres - *Doesn't contribute to the increase in ticket price factor.*

5. **CONCLUSION** : Based on the above 4 scenarios, found to be that *installation of an additional chair lift to bring skiers back up, without additional snow making coverage* contributed approximately **3,474,638** increase in the revenue when the ticket price is increased for **1.99** per person. The Operational cost for the new chair lift addition is **1,540,000** so still there is a Gross Profit of approximately **1934638** for an average of **350,000** visitors for **five days' skiing**.