**Guided Capstone Project Report**

**The problem statement:**

A Big Mountain Resort in Montana is spacious for 350,000 people to ski or snowboard every year. Recently they installed an additional chair lift, so their operating costs have increased to $ 1, 540,000. Now, they need to increase the company’s bottom line and reduce expenses by 20% within the next financial year. So, they are looking for some guidance to select a better value for their ticket price.

**Data Wrangling:**

In this step I collected the data from the library. I imported libraries, like Panda, Seaborn, and OS, to organize the data. Then I loaded data in a Ski\_data data frame. I used some panda methods to explore the data. Here I got two target features to predict the ticket price. **The target features are AdultWeakend** and **AdultWeakday**. Then here, I removed some rows which has no price data.

The next step is **Exploratory & Analysis** data, at this point, I got an idea of what is the data science problem and can the data can help to solve it? Here I used kinds of plots and heatmaps to visualize the data. Then cleaned and defined the data for the next step Preprocessing and Training.

**Training the Models:**

Here, I started to build machine learning models by considering how useful the mean value is as a predictor. In this notebook Preprocessing and Training, I defined some functions using formulas. The regression models have some common steps, so I attributed missing values to the training data sets using different strategies, such as mean and median. Some other steps are scaling the features, training the models, and calculating the performance.

**Evaluation Metric**

**R2=1−∑(yi​−yˉ​)2∑(yi​−y^​i​)2​**

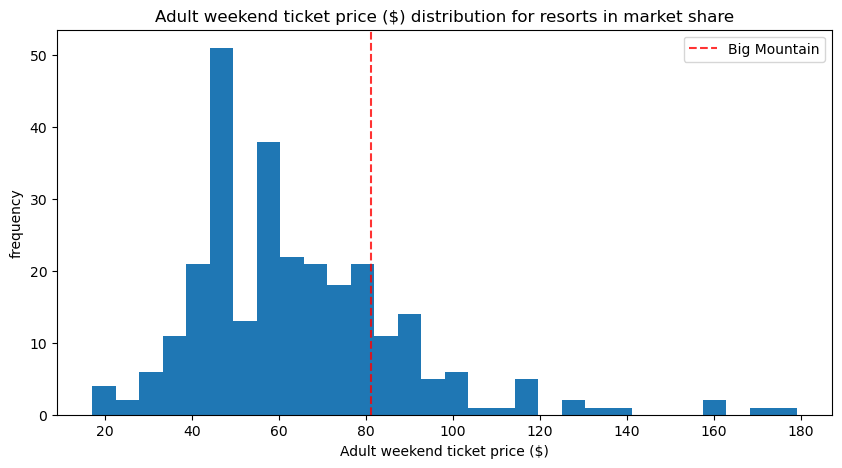
**Mean Absolute Error : MAE=n1​i=1∑n​∣yi​−y^​i​∣**

**Mean Squared Error : MSE=n1​i=1∑n​(yi​−y^​i​)2**

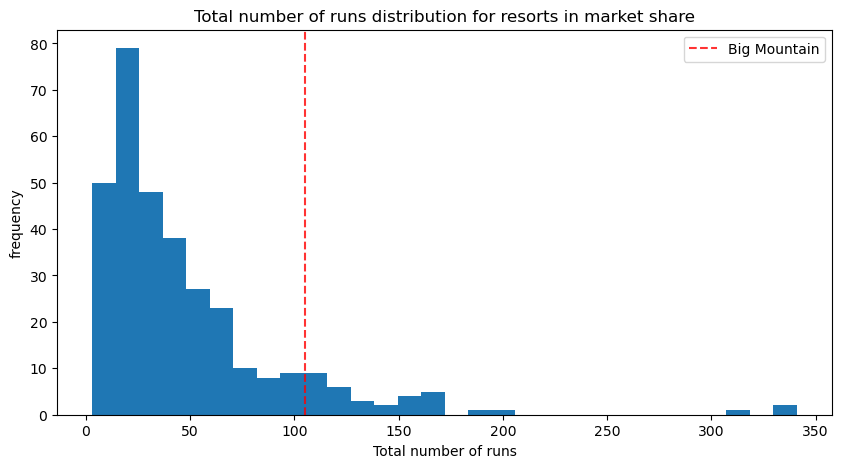
Both the linear regression model and the random forest model have the best step which is cross-validation. Here I got the mea is 5.978 using the random forest regression model performance. The Final step in this project is **Modelling.** Here I used the model to gain insight into what Big Mountain's ideal ticket price could/should be, and how that might change under various scenarios.

**Pricing:**

Big Mountain Resort currently charging 81.00, but the modelled price is 90.55.



The utility of our model comes in, 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. And, increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres.



**Model and Scenario:**

Here, I will decided to use and go forward with the **Random Forest model, because I would say it has a better prediction.**I would recommend second scenario, which is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift. The model says closing one run makes no difference.

**Conclusion:**

In this project, I developed and evaluated a predictive model to assist Big Mountain Resort in determining an optimal ticket price while accounting for increased operating costs and revenue goals. After performing data wrangling, exploratory data analysis, and building several machine learning models, the Random Forest model emerged as the most reliable predictor of the ticket price. Using this approach, I will be able to provide the resort with actionable data-driven insights to adjust their pricing strategy and make informed decisions about infrastructure improvements, ensuring long-term profitability and customer satisfaction.