Deliverables

Part 1 - Problem Statement

1. Hypothesis → How can Big Mountain Resort create a ticket pricing strategy that will (1) better leverage the client's facilities and (2) either cut costs or give more confidence in increasing ticket prices without undermining ticket sales in the long run?
2. Criteria for success → Providing a pricing strategy that will better leverage the client's facilities and either cut costs (without undermining the ticket price) or give more confidence in increasing ticket prices.

Part 2 - Data Wrangling

1. The original number of rows was 330. fastEight and AdultWeekday were removed from consideration because of missing values and significant zero-values. Other issues included (1) missing values in AdultWeekend column, (2) an unreasonable Silverton Mountain skiable terrain error and (3) There was an unrealistic value of 2019 in the yearsOpen column. This was corrected as needed. The target feature was confirmed to be AdultWeekend prices. After initial wrangling, there were 277 rows left.

Part 3 - Exploratory Data Analysis

1. There was no pattern suggested of a relationship between state and ticket price. Consequently, all states were treated equally with regards to modelling.
2. Features ***AdultWeekend ticket price, fastQuads, Runs, Snow Making_ac, resort_night_skiing_state_ratio, vertical drop and total_chairs*** showed good correlations with ticket price. Feature ***resorts_per_100kcapita*** showed variation depending on value. Features ***total_chairs_runs_ratio, total_chairs_skiable_ratio, fastQuads_runs_ratio, fastQuads_skiable_ratio*** showed counterintuitive correlations.
3. The number of visitors per year is unavailable but could have been useful. Additionally, having no ***fast quads*** may limit the ticket price, but if a resort covers a wide area then getting a small number of fast quads may be beneficial to ticket price.

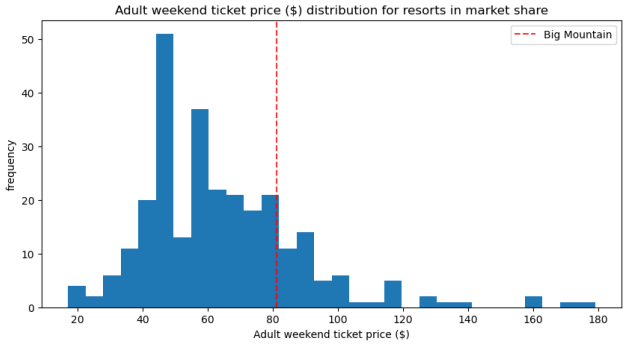
Part 4 - Pre-Processing

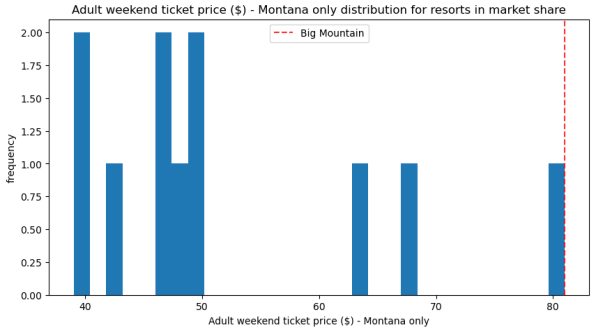
1. I built a linear model using the data. This suggested that ***with median and mean***, you'd expect to estimate a ticket price within ~$9 of the real price. The estimate performance (using r^2 score and mean absolute error) on the test data was consistent with that of the training data.
2. A random forest regressor was performed. The best pre-processing steps were simple imputing using the median and standard scaling. The cross-validation performance on the test data was approximately consistent with that of the training data.
3. I opted for the random forest model. I did this because the random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produced a performance consistent with the cross-validation results.

Part 5 - Modelling

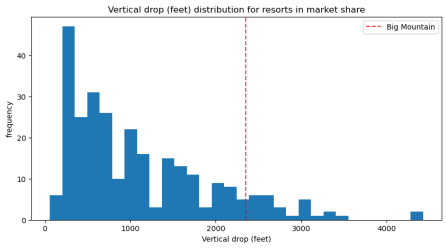
1. Big mountain currently charges $81 for AdultWeekend tickets (ADT). Our model suggests you should charge $95.87. The additional operating cost of the new chair lift per ticket (on the basis of each visitor on average buying 5 day tickets) should prompt a $1.99 increase to $97.86.
2. Useful data that was missing from the dataset include: (1) number of visitors per year, (2) additional data about typical length of stays at these resorts, (3) number of day tickets sold, (4) number of weekly passes sold etc.
3. To test a new combination of parameters in a scenario, (1) this model will be made available for business analysts to use and explore, (2) simple linear, piecewise equations can be provided so that the company can make small adjustments to price (up tro stated limits) without another consultation or (3) a simple web-based app can be developed using the insights from this model/analysis.
4. See Figures 1 to 2 for frequency distribution of ADT for the entire market and for Montana respectively. See Figures 3 to 10 for the frequency distributions for Vertical drop, Area covered by snow makers, Total number of chairs, Number of fast quads, Runs, Longest run, Trams and Skiable terrain area repectively. See Figure 11 for Changes in Revenue and Ticket prices per runs closed.

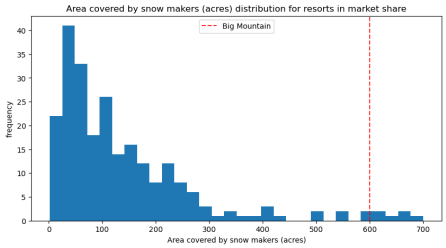
Figure 1 and 2

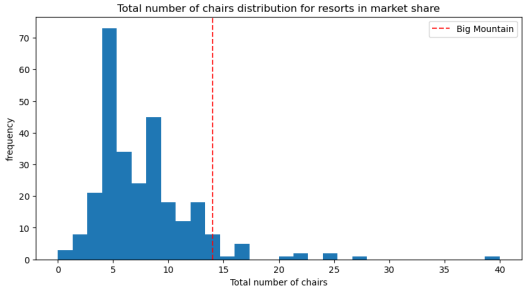




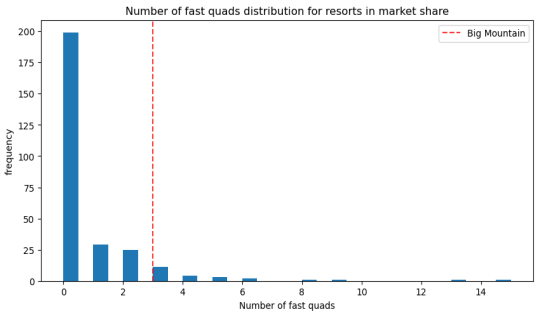
Figures 3, 4 and 5

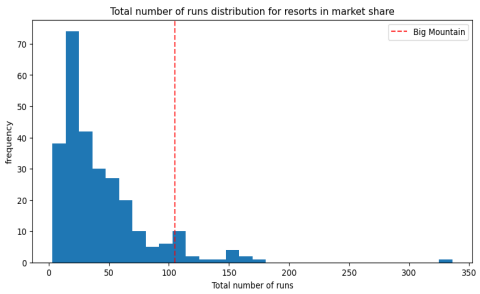






Figures 6, 7 and 8





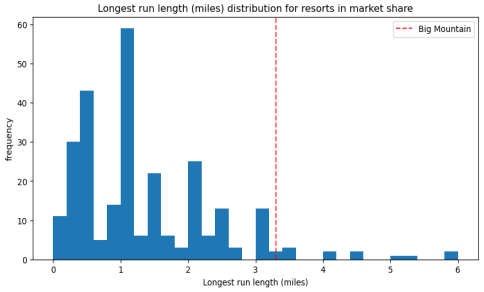


Figure 9 and 10

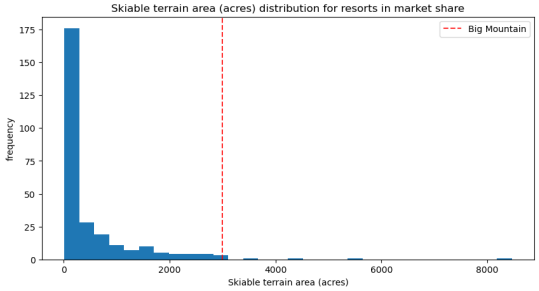


Figure 11 - Changes in Revenue and Ticket prices per runs closed