Capstone Project1 – Final Report
Big Mountain Resort

Objective and Goal:

Big Mountain Resort has recently installed an additional chair lift to better serve visitors. The additional installation increases operating costs by $1,540,000, which negatively impacts the profit. Company needs to set the best value for ticket price to make sure return supports the future investment plans.

Question: Which facilities matter most to visitors? particularly which features or facilities do they most likely pay more for?

Description of the Data:

The dataset used for this project called 'ski_resort_data' that contained data about the 330 resorts in the US. I also used a dataset on information about states. Analyzing and fitting Simple Linear Regression and Random Forest Regressor provided some insightful information about the market and facilities that matters for visitors and impact ticket price.

List of the Original Features and Features Engineered

Name	yearsOpen
Region	averageSnowfall
state	AdultWeekend
summit_elev	projectedDaysOpen
vertical_drop	NightSkiing_ac
base_elev	resorts_per_state
trams	resorts_per_100kcapita
fastSixes	resorts_per_100ksq_mile
fastQuads	resort_skiable_area_ac_state_ratio
quad	resort_days_open_state_ratio
triple	resort_terrain_park_state_ratio
double	TerrainParks
surface	LongestRun_mi
total_chairs	resort_night_skiing_state_ratio
Runs	total_chairs_runs_ratio
SkiableTerrain_ac	total_chairs_skiable_ratio
Snow Making_ac	fastQuads_runs_ratio
daysOpenLastYear	fastQuads_skiable_ratio

Description of original features are available in Fig 1.

Some of the most important findings and recommendations are as follows:

1) Distribution of resorts by states (Fig 2) showed that New York accounting for the majority of resorts. Big Mountain in Montana came in at 12th place.
2) There was a clear line where weekend and weekday prices were equal (Fig 3). Weekend prices being higher than weekday prices seemed restricted to sub $100 resorts. Besides their relationship, less missing values for weekend price resulted into choosing weekend price as a target feature in modeling.
3) Runs and vertical_drop were the most important features in setting price and consequently expected revenue. Ticket price was very sensitive to closing down more than 6 Runs (Fig 4).
4) Features like Snow Making_ac, fastQuades, total_chairs, LongestRun_mi, trams and SkiableTerrain_ac were other important features that the ticket price was sensitive to their changes but not small changes (Fig 5).
5) There were other features like Resort night skiing state ratio (ratio of resort night skiing area to total state night skiing area) seemed the most correlated with ticket price (Fig 5). Perhaps seizing a greater share of night skiing capacity is positive for the price a resort can charge. Such features were correlated with the ticket price but they had relative importance in setting price.

Based on the result of learning curve function, we had plenty of data. In other words, Big Mountain doesn't need to collect more data, since increasing the size of training sample (more than 60) in Cross Validation didn't result better scores (Fig 6).

Fig 1. Original Features and Description

Column	Description
Name	The name of the ski resort.
Region	The region within the United States where the resort is located.
state	The state name where the resort is located.
summit_elev	Elevation in feet of the summit mountain at the resort.
vertical_drop	Vertical change in elevation from the summit to the base in feet.
base_elev	Elevation in feet at the base of the resort.
trams	The number of trams.
fastEight	The number of fast eight person chairs.
fastSixes	The number of fast six person chairs.
fastQuads	The number of fast four person chairs.
quad	Count of regular speed four person chairlifts.
triple	Count of regular speed three person chairlifts.
double	Count of regular speed two person chairlifts.
surface	Count of regular speed single person chairlifts.
total_chairs	Sum of all the chairlifts at the resort.
Runs	Count of the number of runs on the resort.
TerrainParks	Count of the number of terrain parks at the resort.
LongestRun_mi	Length of the longest run in the resort in miles.
SkiableTerrain_ac	Total skiable area in square acres.
Snow Making_ac	Total area covered by snow making machines in acres.
daysOpenLastYear	Total number of days open last year.
yearsOpen	Total number of years the resort has been open.
averageSnowfall	Average annual snowfall at the resort in inches.
AdultWeekday	Cost of an adult weekday chairlift ticket.
AdultWeekend	Cost of an adult weekend chairlift ticket.
projectedDaysOpen	Projected days open in the upcoming season.
NightSkiing_ac	Total skiable area covered in lights for night skiing.

Fig 2. Distribution of Resorts by State

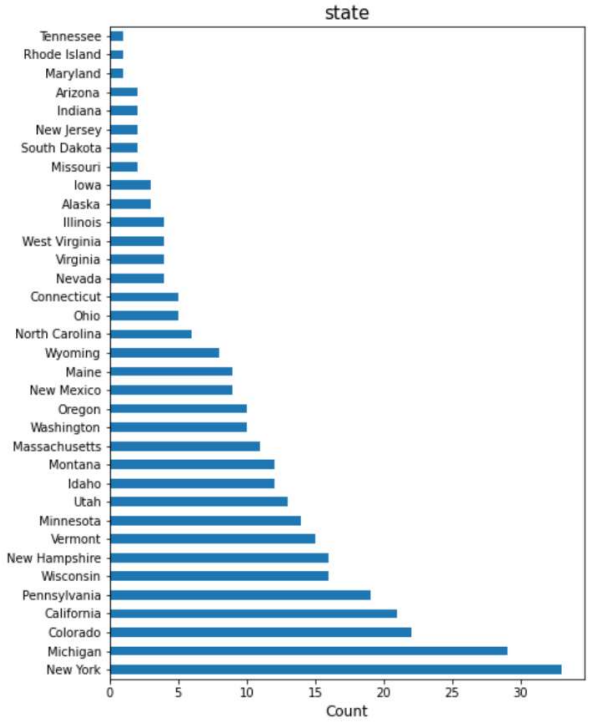


Fig 3. Relationship between Weekend and Weekday Prices

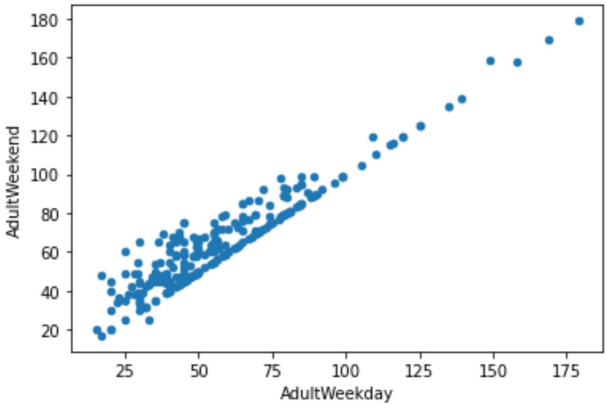
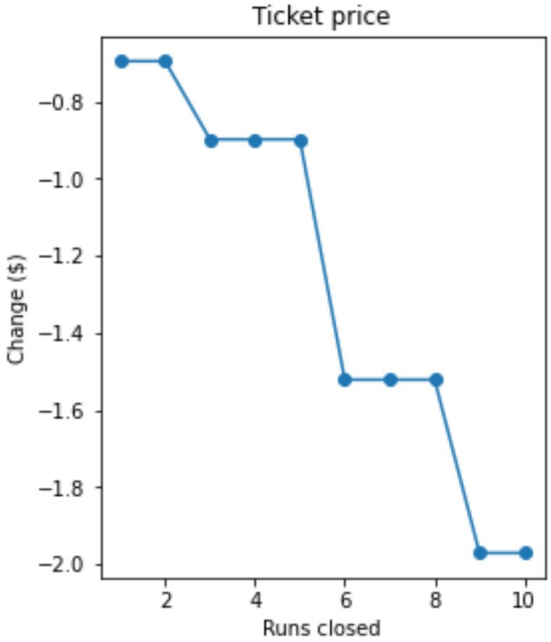


Fig 4. Sensitivity of Ticket Price to Number of Run Closures



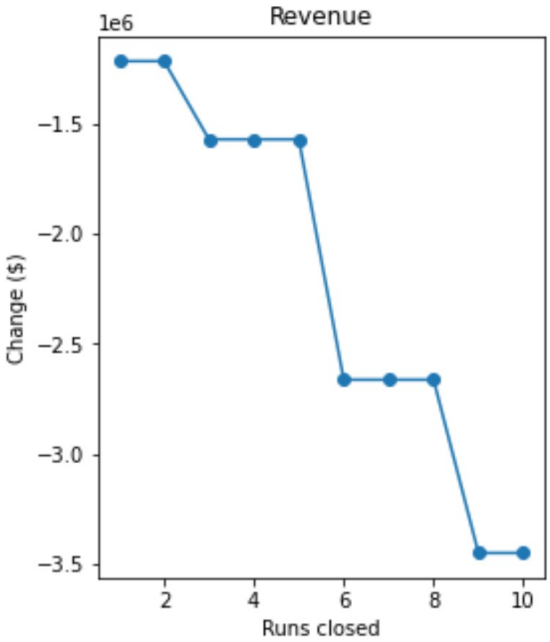


Fig 5. Most Important Features on Ticket Pricing

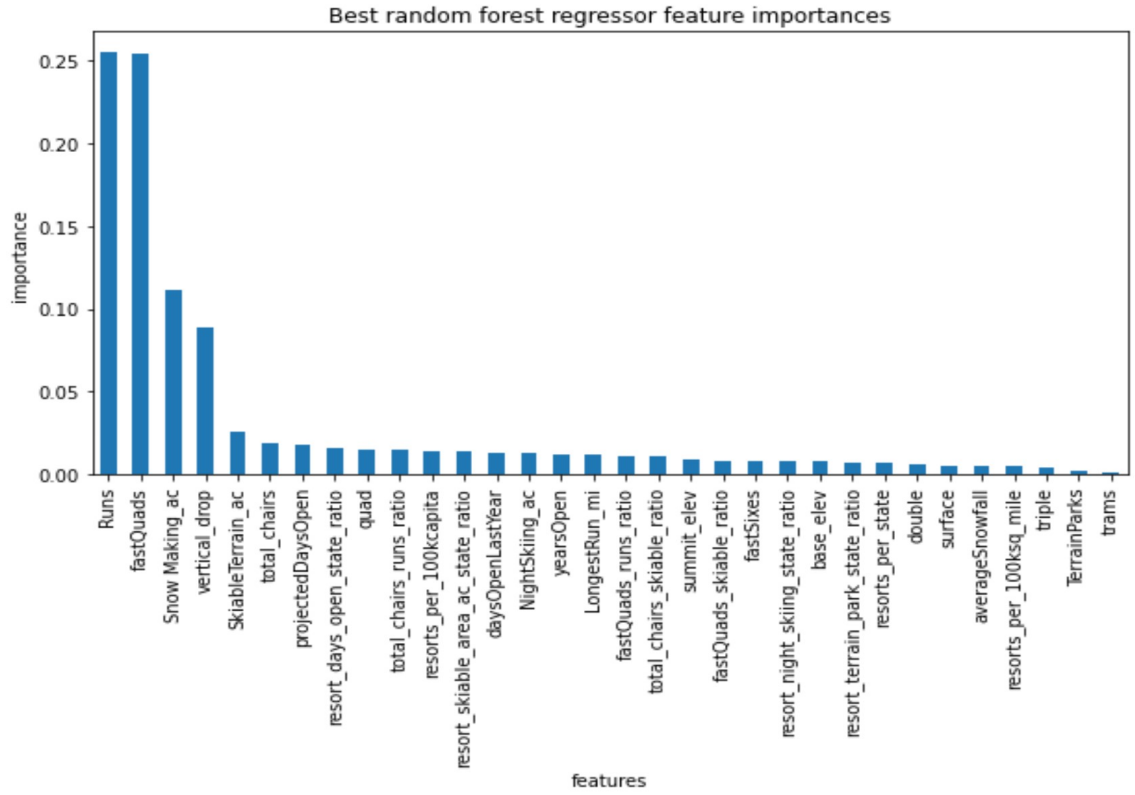


Fig5. Data quantity assessment