**Guided Capstone Project Report**

**Amit Giri**

This guided capstone project is about the client ‘Big Mountain resort’, a ski resort located in Montana. Every year about 350,000 people ski or snowboard at Nig mountain, Big Mountain resort has recently installed an additional chair lift to help increase the distribution of visitors across the mountain but has increased the operational cost by 1.52 million this season. The business wants some guidance on how to select a better value for their ticket price.

**Applying the Data Science**

We applied data science method to get appropriate results. Criteria for success is making better use of the facility and working on ticket pricing or cutting down the other cost without affection the ticket price.

I used python language on Jupyter notebook platform to use the given data to find out the result we wanted. One of the initial step of data science method in Data Wrangling where I collect, organize, define and clean a relevant dataset. About 14% of the rows have no price data. As the price is our target, these rows are of no use, and we dropped these rows. Original data provided by client had 330 rows and 27 columns, after cleaning it came down to 277 rows and 25 columns.

Next step was exploratory data analysis which was to understand the relationship between data and features. In this step we found the positive correlation between the ratio of the night skiing area and resort per capita. When Resort was densely located more night skiing were provided. Night skiing was also positively corelated to the ticket price. 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.

Next step was Pre-processing and training data development, where we build the machine learning model to train our data. I used 70/30 train/test split partition sizes to train and test the data. Which means we use 70% of our data to train and 30% of our data to test using machine learning. As per this step Vertical drop and area covered by snow making equipment had the most positive feature for increasing the ticket price.

Chart, line chart

Description automatically generated

Modeling is the fifth step in Data science method. In this step we cleaned and process the data to make predictive insights. In this step we ran 3 scenarios to see which scenario supports increasing the price which result in more revenue. 1st was to close the 10 of the least used runs but closing different least run had different impact. Closing one run makes no difference while closing 2 and 3 successively reduce support for ticket price and revenue. If Big Mountain closes 3 runs, it seems they may as well close down 4 or 5 as there's no further loss in ticket price. Increasing the closures down to 6 or more leads to a large drop. In 2nd scenario, Big Mountain is adding a run, increasing the vertical drop by 150 feet and installing an additional chair lift. This scenario increase support for ticket price by $1.99 over the season which could be amount to $3474638. In scenario 3 we add 2 acres of snow making on scenario 2 but such a small increase in the snow making area makes no difference. Suggested ticket price increase and revenue remains the same.