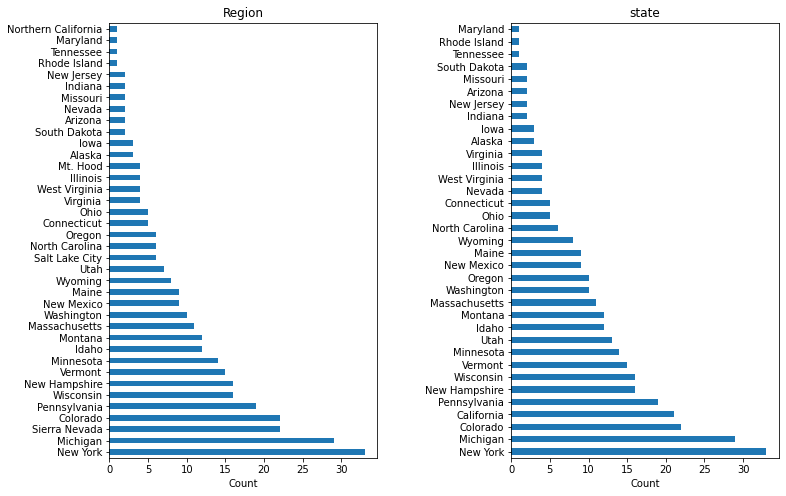
**The Problem:**

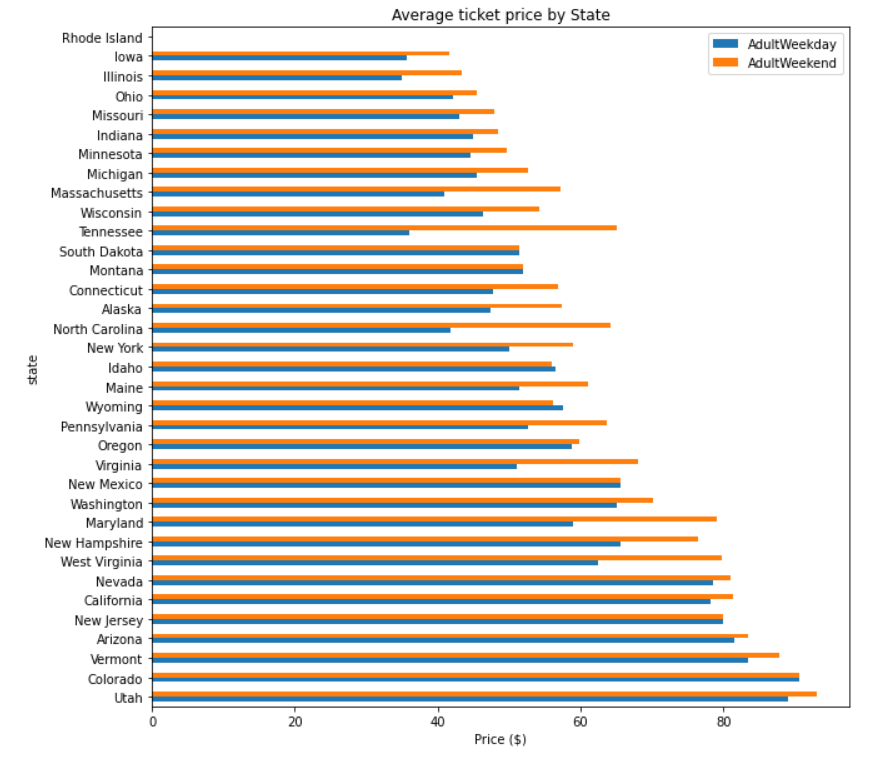
The main purpose of this data science project is to come up with a pricing model for ski resort tickets. The business team has the understanding that the returns are getting low, relative to its position in the market. This model will guide them for future investment plans.

**Data Wrangling:**

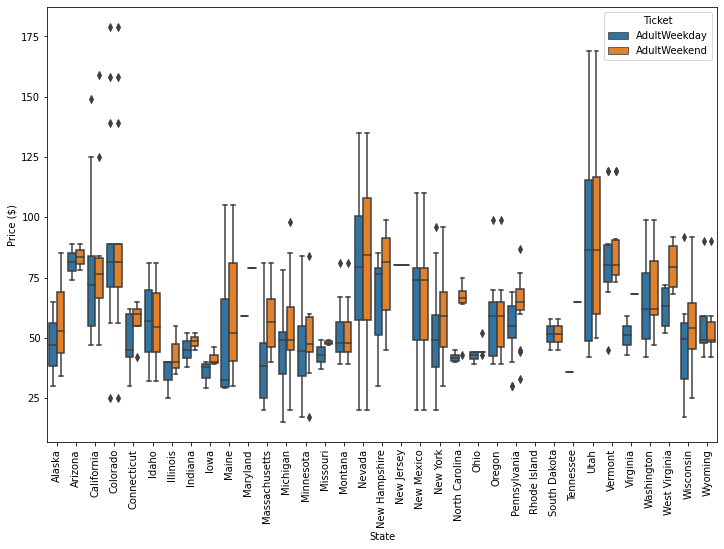
This is the first step of our project focusing on collecting data, organizing it, and ensuring it is well-defined. At first, we find the missing values in the data. The plots below show the resort count in each state and region.



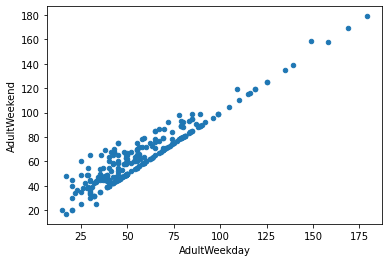
The plot below depicts the average price comparison on weekends versus weekdays. Weekend prices are slightly higher.



The box plot below shows price distribution based on the state. California, Colorado, and Utah have higher prices with the broad range of $25.00 to over $100 respectively.

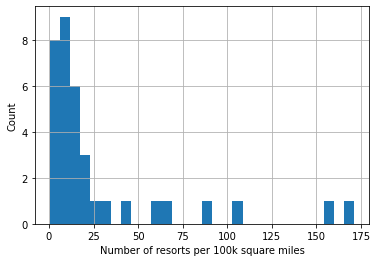
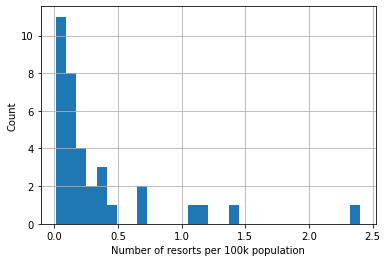


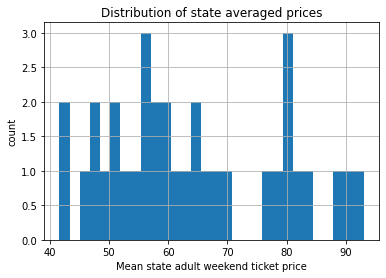
One thing to note here is, we simply lack the required data due to missing values, so we have no other option but to drop those records which account for 15% of the resort's data.



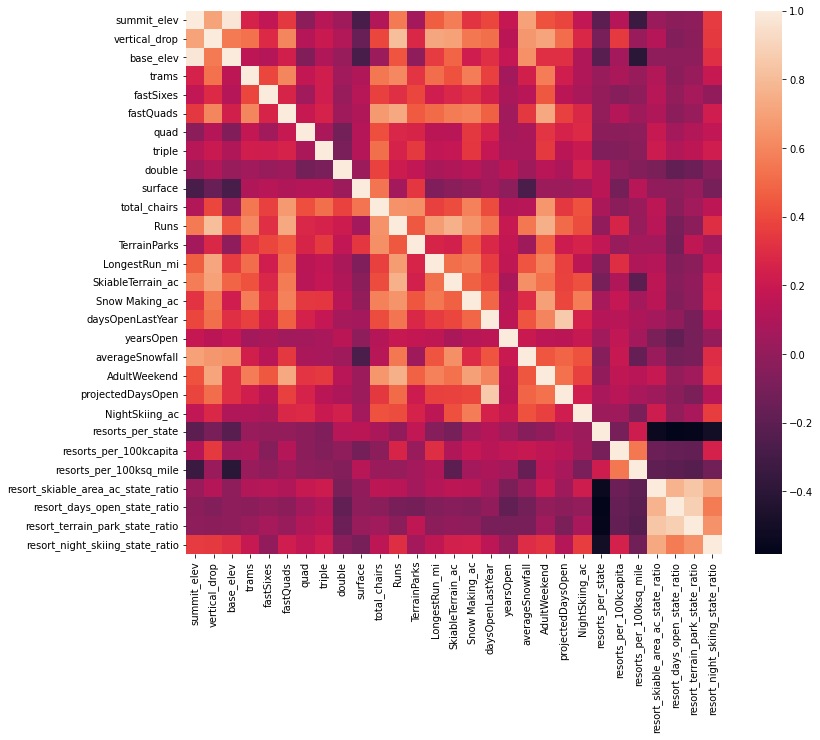
**Exploratory Data Analysis:**

After cleaning and organizing the data, we are going to further analyze the refined data set. In this part we explore top states, total state area, resort per state and resort density. The big states are not necessarily the most populous states as shown below,





The heatmap shows the relationship amongst the features in data.



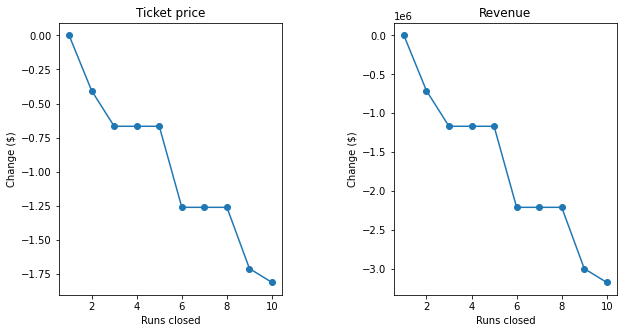
**Pre-processing and Training Data:**

We treated ski resort data as single entity. In machine learning, when you train your model on all the data, you end up with no data aside to evaluate model performance. So, we partition the data into training and testing (70/30), which is very useful final check on the expected performance.

**Modeling:**

Resort ticket price modeling relies on the assumption that other resort in big mountain competitors setting the price based on what people value at certain facilities. The resort’s actual prices were around $81 and $95 with the mean absolute error of $10.39, suggest that their is a room for price increase. Model also suggests that the number of visitors remain the same over the years. The resort also looking to cut some facilities if they have to or increase the ticket price to sustain the reasonable profit. Resort is also free of setting the prices.

The resort wants to look at four different scenarios. First scenario is to close permanently 10 of the least used runs.



The model says closing one run makes no difference. Closing 2 and 3 successively reduces support for ticket price and so revenue. If Big Mountain closes down 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.

Second scenario, increase the vertical drop by 150 feet and install an additional chair lift. This scenario increases support for ticket price by $8.61 with the expected revenue to be $15065471. Adding two acres of snow making could increase the revenue by $17322717 if price is set to $9.90. It makes so such difference, this is scenario three.

Last scenario, if the longest run increase by 0.2 miles, guaranteeing its snow covering by adding four acres of snow making capability, will not going to make any difference respectively.

**Recommendations:**

The lack of visitors per season or year numbers was inhibiting our analysis. It also would have been helpful to see a breakdown of costs, like how many employees are needed per run, equipment cost or per other variables. Our model told us that Big Mountain was undercharging, but maybe it has a bad reputation for some reason, or other information affecting the situation that was not in the data. It would need to be a conversation with the executives, and it would be good to know if there was a previous model that gave the $81 price. Furthermore, we can play with other parameters to see how they affect the model.