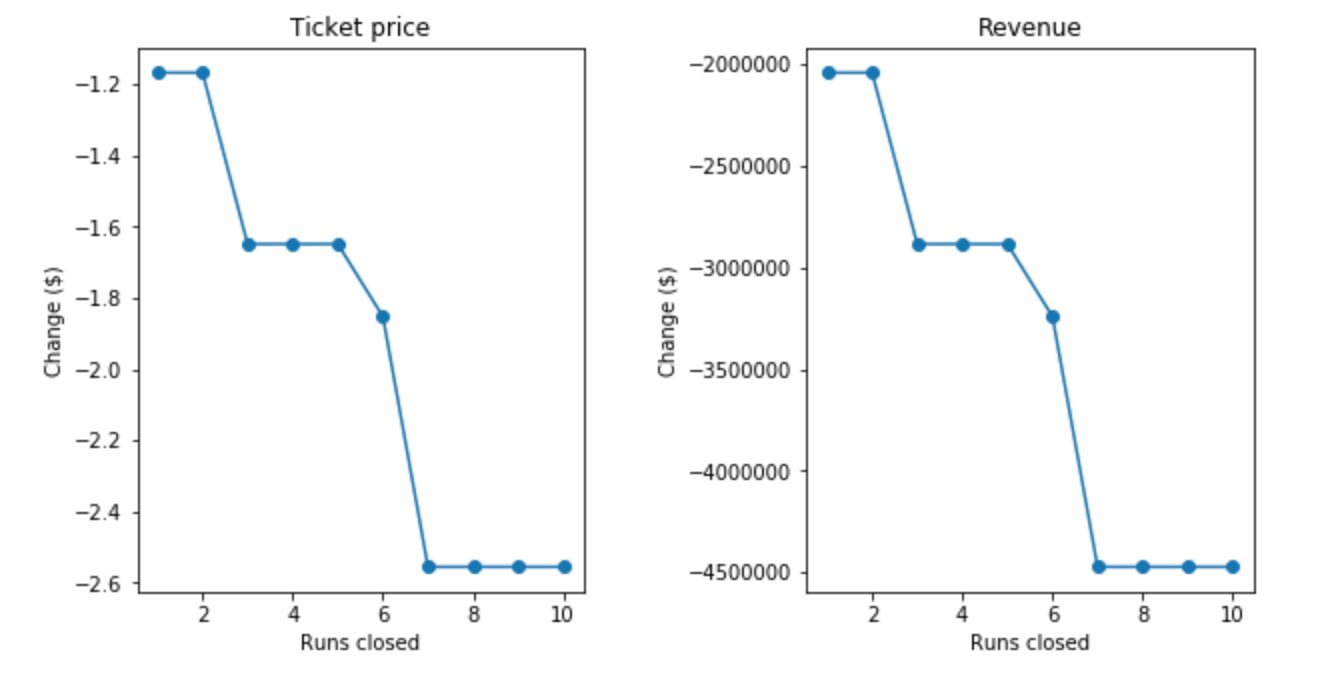
**Final recommendations:**

* Without changing any facilities, Big Mountain Resort modelled price is USD 92.80, actual price is USD 81.00. Even with the expected mean absolute error of $10.41, this suggests there is room for an increase. However, would suggest business to observe impact on number of visitors with piloted increased prices, before finalising it.
* Moreover, suggest dropping one run in scenario 1 (Permanently closing down up to 10 of the least used runs), as the model suggests no change in price/ revenue. For further closes, the suggestion is to start with reducing 4 or 5 and observing the effect on prices/ revenue, as the model predicts almost the same drop for 2 through 4/5 run drops. Do not recommend further run deduction.



*Fig 1: runs closed vs ticket price/ revenue*

* Further, can consider scenario 2 (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), as it predicts an increase in USD 1.33 per ticket, translating to USD 2,333,333 increase in revenue, assuming 350,000 visitors over the season, with an average of 5 tickets per visitor. Increased operating cost/ ticket for the additional chair is: 1,540,000/(5∗ 350,000) = $0.88

**Data and exploratory insights:**

***Raw data overview***

1. 330 rows, including our own resort
2. 27 columns, from which 2 (AdultWeekday and AdultWeekend) could be target features

***Data issues and wrangling steps***

1. Silverton Mountain resort had a wrong entry for SkiableTerrain\_ac of 26819 (only value > 10k sq mi), replaced it with the correct one as per Wiki: 1819
2. Fasteight column was dropped as it has ~50% missing values and rest are 0

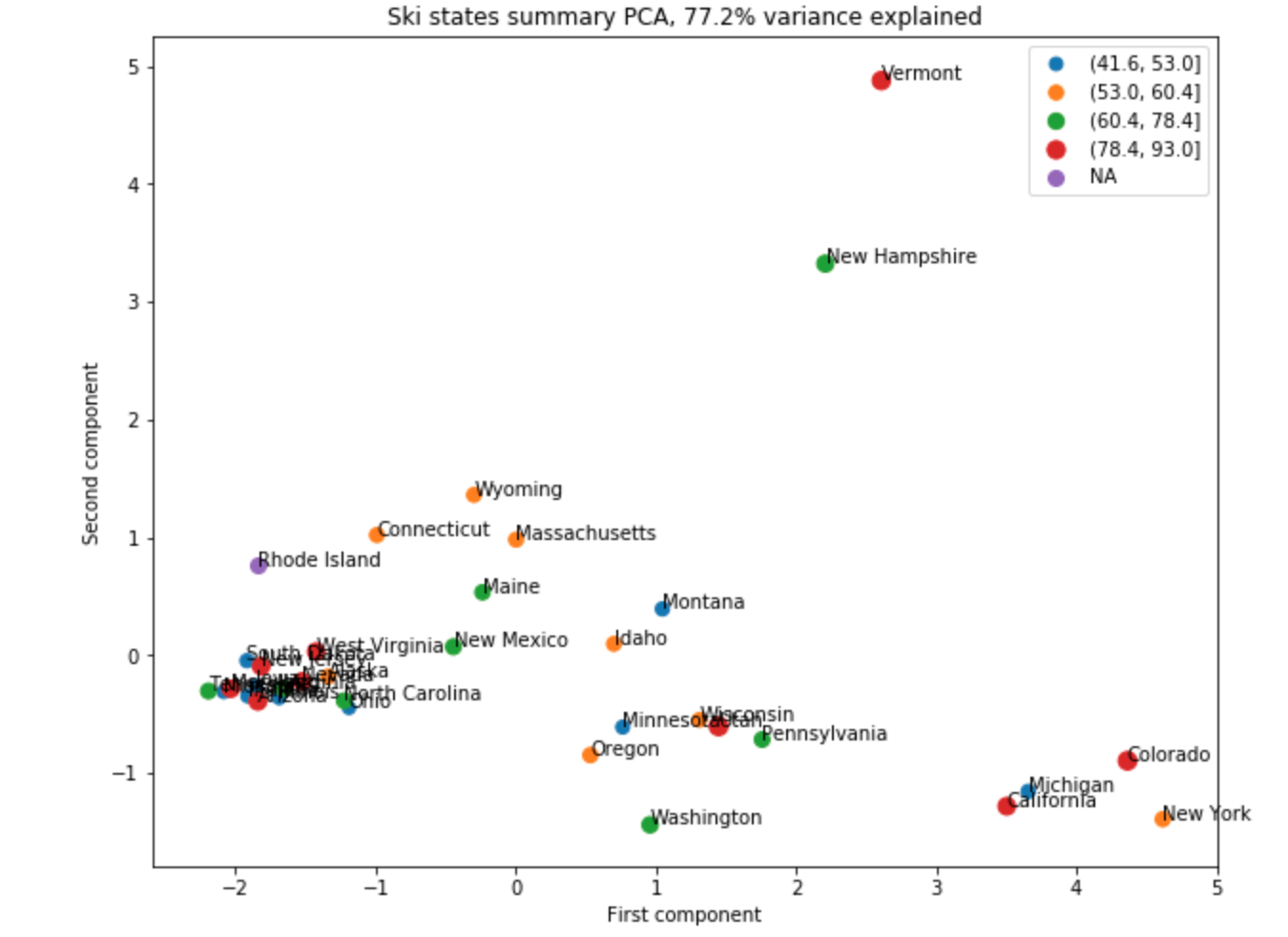
***Other steps***

1. Aggregated State wise data and added population and state size data to create a new csv file, as the prices as well as other features looked state dependent. This csv might. Be useful in next steps
2. Considered Weekend prices as the target feature as the values for weekend and weekday are almost equal for ticket prices >100 dollars (Including our resort), and it has less missing values.
3. First Dropped the 14% rows with both prices missing, then the 3 rows missing only Weekend prices

**Influence of state on ticket prices:**

On performing PCA based analysis of state summary, it's evident that there are extremes in prices for few states: namely Vermont and New Hampshire, that have high no. of resorts per 100 sq m. We also say that other features feature in the PCA 0 and 1, which explain 77% of the variance in the state summary data.

However, there are no easy to recognise patterns in state data to be used directly, so state data was ported to the per resort data to include the 'state competition' features: 'resorts\_per\_state', 'resorts\_per\_100kcapita', 'resorts\_per\_100ksq\_mile', 'resort\_skiable\_area\_ac\_state\_ratio', 'resort\_days\_open\_state\_ratio', 'resort\_terrain\_park\_state\_ratio', 'resort\_night\_skiing\_state\_ratio', 'total\_chairs\_runs\_ratio', 'total\_chairs\_skiable\_ratio', 'fastQuads\_runs\_ratio', 'fastQuads\_skiable\_ratio'



**Model details:**

Linear regression and random forest were the two models tried, both had R square >0 (better than average).

From there, random forest was chosen as it performed better on cross validation. It has a lower cross-validation mean absolute error by almost $1.