GUIDED CAPSTONE PROJECT PRESENTATION

NAZILA GUNDOGDU

PROBLEM STATEMENT

- With $1,540,000 cost for an additional chair, we need to strategize to recoup this investment. Two avenues to explore are:
  - Increasing our ticket prices for the season
  - Introducing some changes to cut costs elsewhere

RECOMMENDATION AND KEY FINDINGS

- A price increase of **$1.99** (from current $81 to $82.99)
- Adding an additional run
- Increasing the vertical drop by 150 feet

MODELING RESULTS AND ANALYSIS

- **Raw Data sources:**
  - Resort data provided by Alesha Eisen, our database manager (size: 330 by 27)
  - Population data obtained from open source
- **After Cleaning:**
  - Dataset I
    - 7 features including state population and area _ later replaced by resort density per 100K capita and 100K square miles
    - Used for comparative cross-state study
  - Dataset II
    - Size: 277 resorts with 29 features
    - Used for modeling

MODELING RESULTS AND ANALYSIS

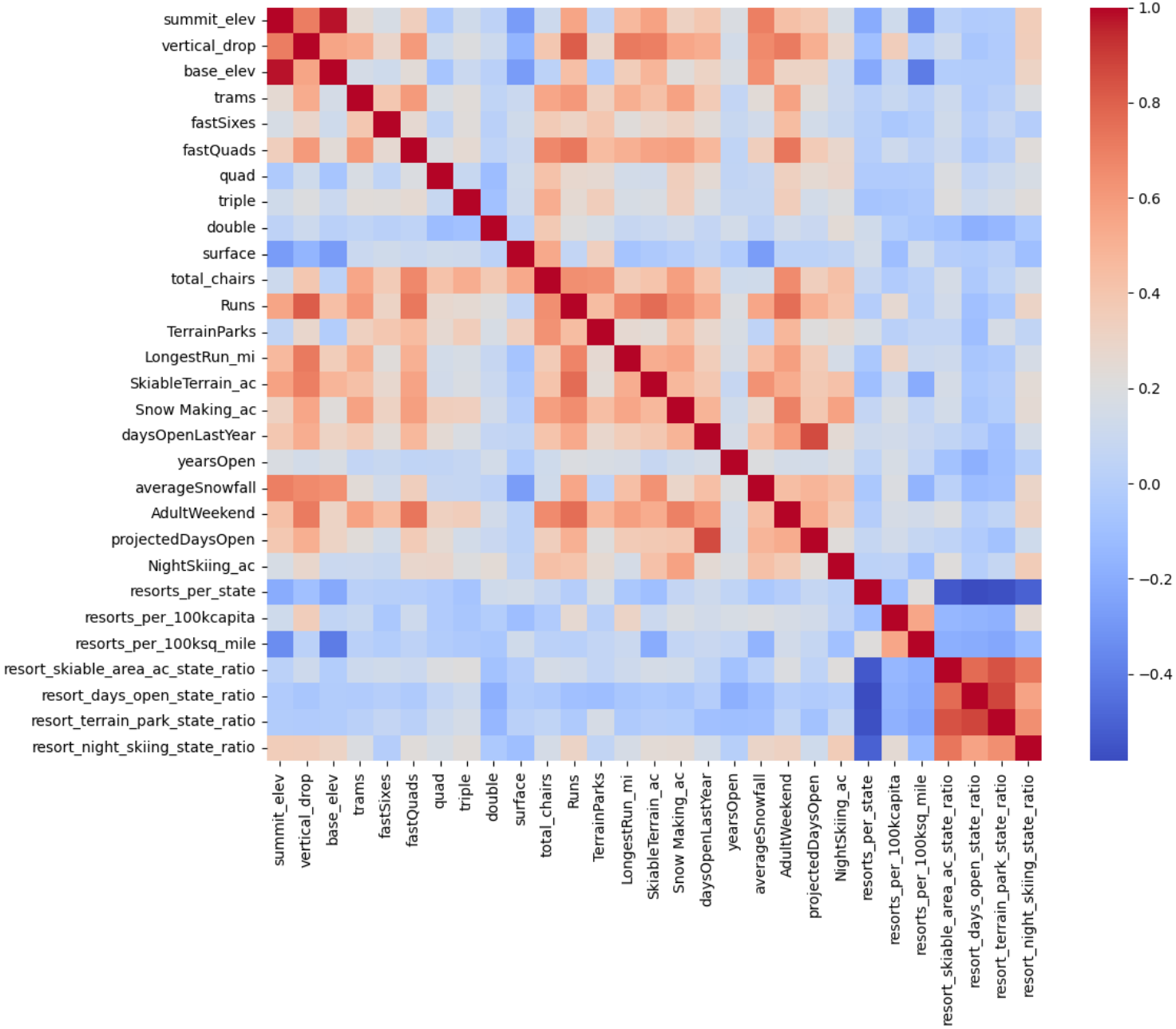
- Step 1: Comparative Cross-State Study → Inconclusive as no clear pattern emerged

Feature	Montana's Status
State Area	3rd
# of Resorts/State	12th
Resort density per 100K capita	4th
Resort density per 100K sq miles	>20th

- Step 2: Custom Dataset → Inconclusive
  - New dataset including resort densities per state, merged with *TerrainParks*, *SkiableTerrain_ac*, *daysOpenLastYear*, and *NightSkiing_ac* from our resort information dataset
  - PCA and dimensionality reduction → same result as Step 1 with no clear pattern emerging

MODELING RESULTS AND ANALYSIS

- Step 3: Back to original dataset + feature engineering → First Direction
  - Added the following 4 features to the data:
    - ratio of resort skiable area to total state skiable area
    - ratio of resort days open to total state days open
    - ratio of resort terrain park count to total state terrain park count, and
    - ratio of resort night skiing area to total state night skiing area
  - Ran a state-wide comparative study
  - Heatmap (next slide) revealed correlation between features



Strong correlation between:

1. *vertical_drop*
2. *total_chairs*
3. *longest_run*

MODELING RESULTS AND ANALYSIS

- Step 4: Modeling
  - Baseline Model:
    - Mean → Mean Absolute Error (MAE) = $19
    - Median → MAE = $9 → The winner method!
  - Model I: Linear Regression
    - Showed no improvement on baseline model using median
    - Cross-validation: 8 is optimal number of features; top 4 being *vertical_drop*, *Snow Making_ac*, *total_chairs* and *fastQuads*
  - Model II: Random Forests → The winner model!
    - Top 4 features: *fastQuads*, *Runs*, *Snow Making_ac*, and *vertical_drop*
    - MAE= $1 less than Linear Regression model

MODELING RESULTS AND ANALYSIS

- Step 5: Scenarios and Prediction
  - Model prediction: room for increasing the ticket price over $15 with an error margin of $10.39
  - 4 scenarios explored
    - Closing runs: showed negative effect
    - Adding Runs, Increasing Vertical Drop by 150 feet and Adding a Chair Lift: supports price increase of $1.99 → The winner scneraio!
    - Previous option + Adding 2 Acres of Snow Making: no difference from previous option
    - Increasing the Longest Run by .2 miles + adding 4 acres of snow making capability: shows no effect

SUMMARY AND CONCLUSION

- **What we did**
  - Cross-state comparison did not show clear patterns in terms of area, population, number of resorts, etc. We could not use this data as is to make price predictions
  - State-wide comparison of various resorts, their prices and the services they offer turned out to be more fruitful
  - Linear Regression and Random Forests were compared and the latter was picked due to lower MAE
  - Following cross-validation, 8 features turned out to be the most relevant to ticket prices
  - Using these 8 features and modeling some plausible scenarios, the following is concluded
- **Conclusion and Recommendation**
  - A price increase of $1.99
  - 1 additional run
  - Increasing the vertical drop by 150 feet.