Ski Weather Conditions Final Report

By Sarah Klute, Shani Dor, Ihunaya Eluwa, and Chiara Blake

Abstract

Whether you are a novice or an experienced skier, various factors can significantly impact the quality of a skiing experience. Among these factors, weather conditions play a crucial role, with snowpack (the depth of snow on a given mountain) being of utmost importance. Utilizing historical weather data, including temperature, snowfall, and rainfall, our objective was to develop a model capable of predicting the snow depth at a given ski resort for an upcoming season and to assist in identifying the optimal resort for skiing. We utilized random forest regression, feature importance (Exhibit 7), multiple and polynomial regression models, and methods of cross validation to investigate this problem. The results indicate that the multiple regression model is the most suitable for predicting snow depth (Exhibit 9). However, despite our efforts, this model exhibits poor performance in accurately forecasting snow depth, resulting in low accuracy when predicting future ski conditions (Discussion).

Introduction

Every year, as the temperature drops and we enter the winter months, many look forward to the long-anticipated ski season. The beginnings of snowfall and temperatures dropping below 32 degrees signal that it's time to bring out your skis or snowboard and get ready for the ski resorts to open. Ski bums (someone that skis frequently) and ski culture have become a widespread phenomenon enjoyed all around the nation. However, with imminent developments in the state of climate change and the rise in global temperatures, ski seasons are beginning to look drastically different as the years go by. A Times article (Vaughan) and BBC article (Gerretsen) both further highlight the grim future of skiing and how climate change threatens this beloved pastime.

For resorts, a lot goes into the preparation for ski seasons in order to maximize profits. For ski bums, knowing where and when there will be good ski conditions allows them to make the best out of their ski season. Additionally, it is important for ski patrollers to have a good understanding of the daily snow conditions, in order to keep people safe. There are many weather-related factors that affect a typical ski season. Major factors include snowfall, snowpack, temperature, rain, and wind. Analysis of these factors can provide us with key insights into the upcoming ski season. In this report we aim to answer the following questions:

1. Can we predict the snowfall conditions for a given resort and time period based on historical data?
2. Which ski resorts should skiers visit for optimal ski conditions?

Once the data is cleaned and initially visualized, we will use five ski resorts' historical weather data to help predict snow depth conditions, to help skiers decide which resorts to visit during the season, and when. Using the historical weather data, we built and trained regression models in order to predict the daily snow depth.

Data Description

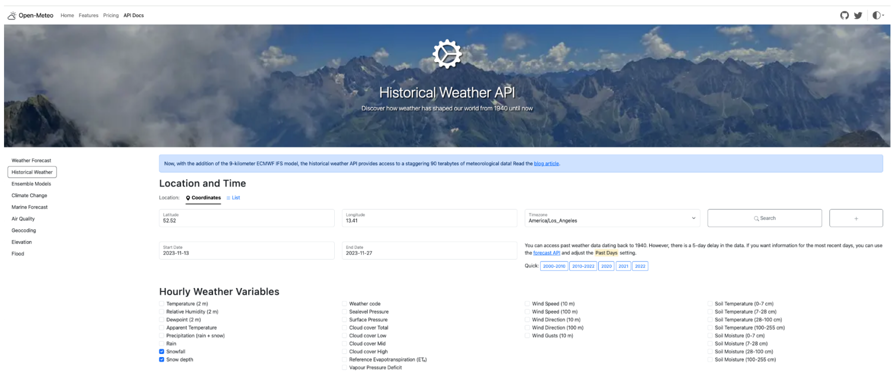

Exhibit 1. Weather API

To answer these questions we will gather data on these factors from the [Open Meteo Historical Weather API](). The maximum number of calls per day to this API is 10,000, this will be kept in mind as we proceed. We will source the data from 5 public ski resorts (Crystal Mountain, Washington; Jackson Hole, Wyoming; Vail, Colorado; Whistler, Canada; and Park City, Utah) for the 2018-2020 ski seasons. These resorts were randomly chosen by our group in an effort to gather data from different areas. The data we will scrape is as follows:

Resort (name of resort)
Location (latitude and longitude of resort)
Start Date:
End Date:
Day Date:
Daily Max Temp (Maximum temperature of that day)
Daily Min Temp (Minimum temperature of that day)
Daily Mean Temp (Mean temperature of that day)
Daily Snowfall Sum (Total snowfall of that day)
Daily Max Wind Speed (Maximum wind speed of that day)
Daily Rain Sum (Total rainfall of that day)
Daily Snow Depth (Depth of snowfall that day)

Pipeline Overview:

To prep and clean our data taken from the API, we defined multiple functions. These functions work together in the pipeline to develop a clean data frame.

1. Run an API weather pull for a specific location/resort.
    a. get_weather(): given latitude, longitude, and date pulls from the Historical Weather API and returns a weather dictionary.

2. Save ski season dates for resorts within a given year
    a. get_dates(): creates a dictionary of resort names, dates, and locations for the ski season of interest for analysis
3. Save hourly data only
    a. hr_to_daily(): converts the hourly data into daily sums
4. Save weather data for each resort into a data frame
    a. season_stats(): takes in resort dictionaries (name, date, and location) and populates the dictionary with weather information utilizing get_weather(). Then compiles a final data frame with daily resort data

Below is a picture of the first few rows of our df_ski Dataframe.

```
df_ski = season_stats(start_dates, leng)
df_ski.head()
```

Table Raw Visualize Statistics

	resort	location	start_date	end_date	day_date	daily_max_...	daily_n
0	crystal mountain	(46.9282, -121.5045)	2018-11-20	2019-04-01	2018-11-20	38.7	
1	crystal mountain	(46.9282, -121.5045)	2018-11-20	2019-04-01	2018-11-21	34.2	
2	crystal mountain	(46.9282, -121.5045)	2018-11-20	2019-04-01	2018-11-22	26.5	
3	crystal mountain	(46.9282, -121.5045)	2018-11-20	2019-04-01	2018-11-23	22.8	
4	crystal mountain	(46.9282, -121.5045)	2018-11-20	2019-04-01	2018-11-24	25.5	

5 rows x 13 columns Jump to top Jump to bottom

Exhibit 2. Ski Dataframe

Visualizations

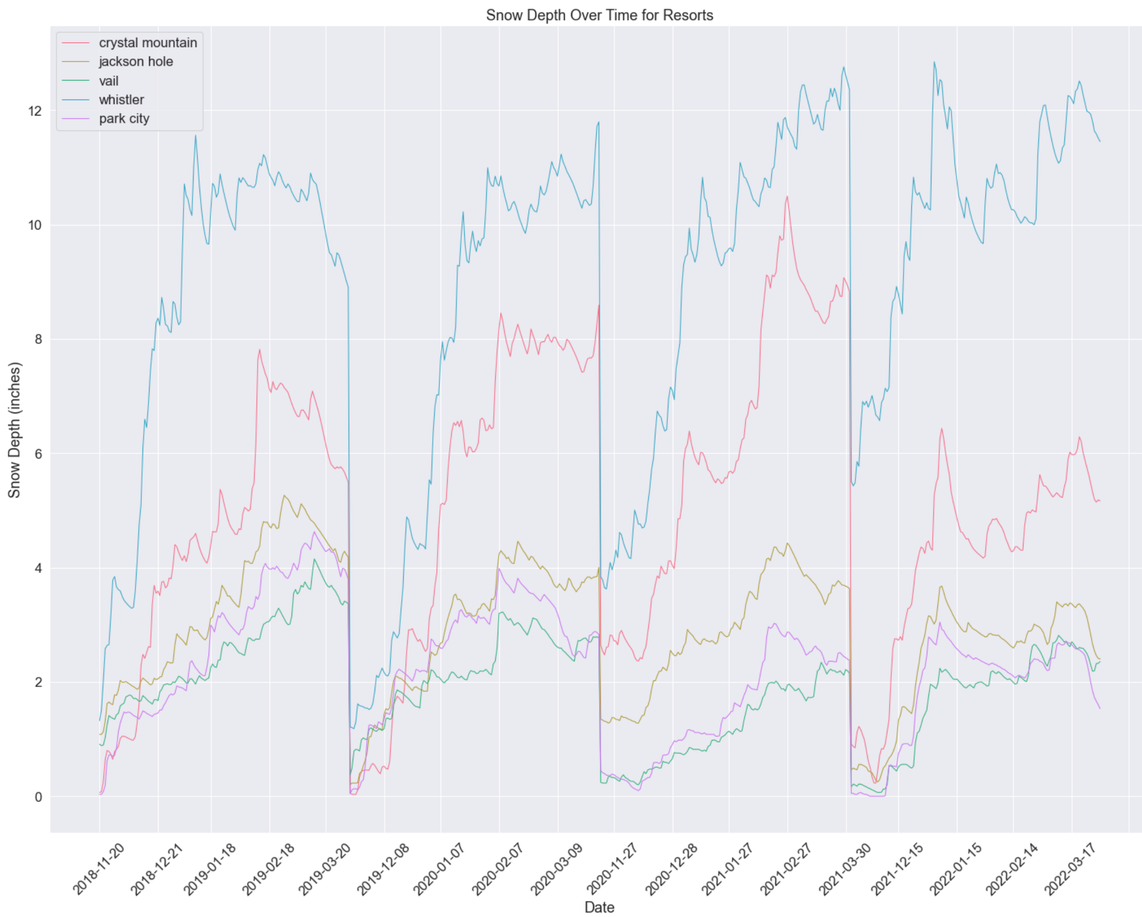

Exhibit 3. Snow Depth Over Time

Exhibit 3 compares the daily snow depths at each of the 5 resorts. It gives insights into which resorts consistently have a greater snow depth revealing Whistler and Crystal Mountains superiority comparatively. Additionally, over the 4-year period, the graph shows which resorts have experienced significantly decreasing snow depths from years previously. Crystal in 2022, compared to the previous season, has seen diminishing levels. Finally, we can also see when snow depth levels peaked for each resort each ski season. In 2022, all resorts, but specifically Whistler, saw snow depth levels peaking much earlier in the season (between December and January) compared to years previously where peaks occurred later into the season.

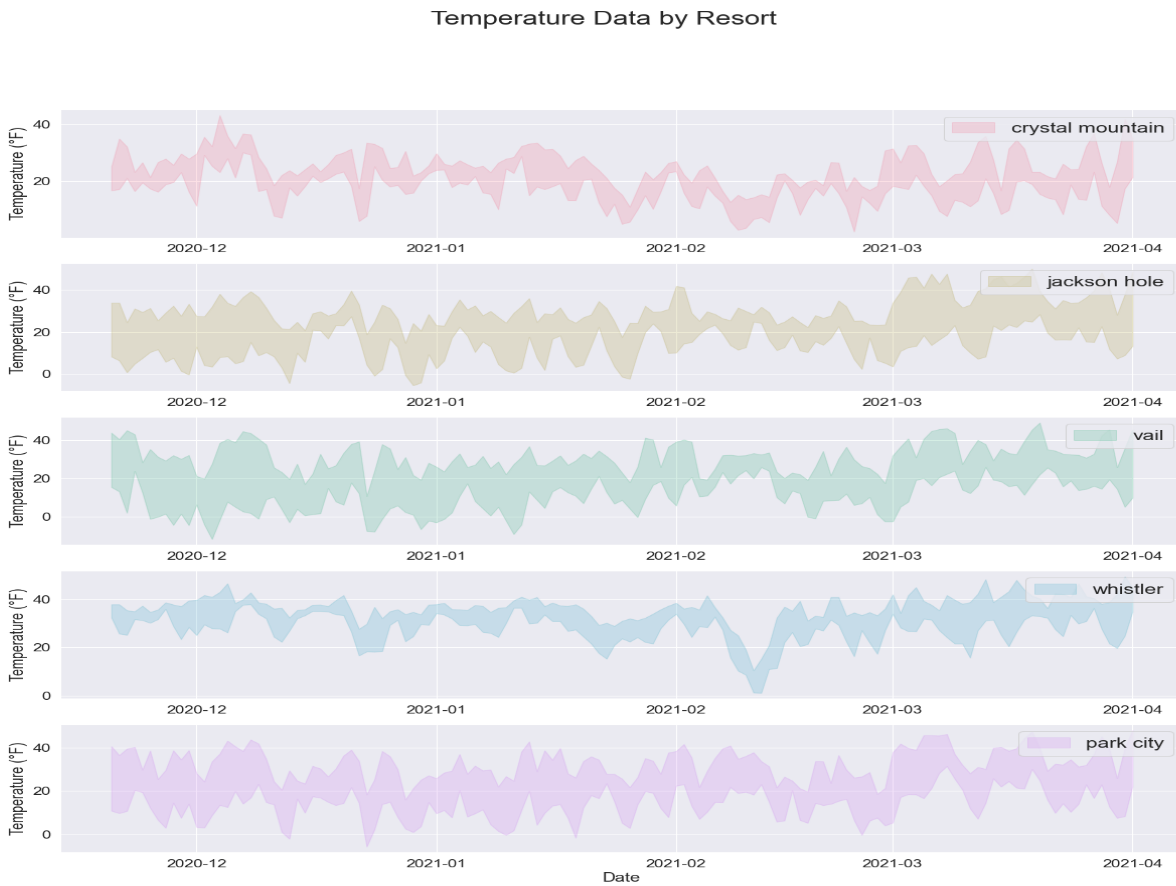

Exhibit 4. Temperature Data by Resort

Exhibit 4 shows the daily temperature data by resort for the 2020-2021 ski season. Daily minimum and maximum temperatures are helpful in determining if it is a good day for skiing. Temperature affects the snowpack and the feeling of the snow (icy/powdery). If daily temperatures are typically optimal during a month (or any time period) we can use that to predict a month with ideal skiing conditions. For example, Park City sees temperatures with a low below 20 during the month of February and a high above 20 during the month of March, if this trend is observed over multiple years we can deduce that February is a better month for skiing than March because the temperatures are lower.

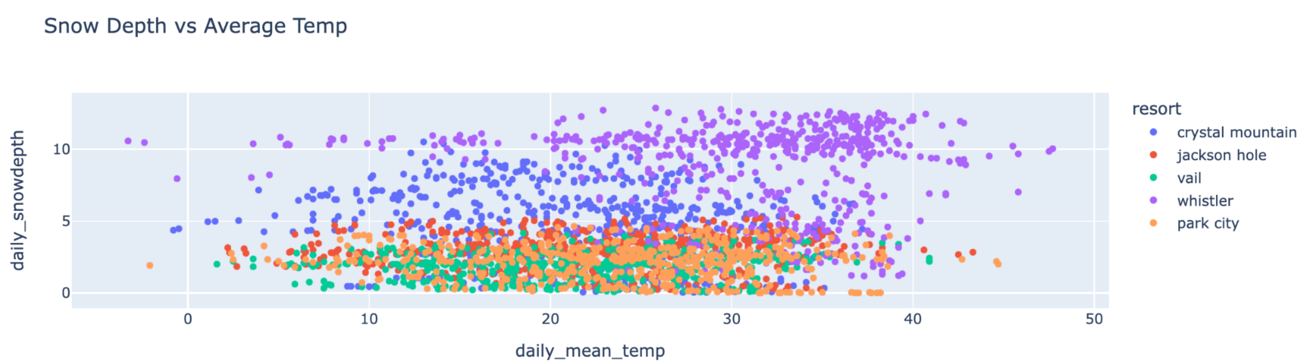

Exhibit 5. Snow Depth vs Average Temperature

Exhibit 5 illustrates that Whistler tends to have the highest snow depth in general. The data seems to be stratified and ordered in relationship to the amount of snow depth at various temperatures. Whistler appears to consistently have the most depth followed by Crystal Mountain, with the rest being about the same (Jackson Hole, Vail, and Park City). This can provide us insight in regards to how we recommend the optimal ski resort for the conclusion of the project. The graph shows that as the daily temperature increases, there is no specific movement as it relates to daily snow depth. Thus, although this graph doesn't exactly illustrate a relationship between daily temperatures and daily snow depth, it is helpful for us to see the differences between resorts.

Exhibit 6. Correlation Matrix

The correlation matrix gives us an initial sense as to which features might have a relationship. At first glance, the minimum and average temp has a strong positive correlation, which is logical as they are somewhat dependent features because the mean temp uses the minimum temp in its calculation. Another positive correlation is seen between the minimum temp and the snow depth. While the correlation is not strong, 0.39 indicates there might be a slight relationship there worth exploring.

Method

The machine learning tools we will use to analyze and answer our analysis questions are random forest regression, multiple linear regression, and polynomial regression. We will utilize historical weather metrics (temp, windspeed, snowfall, rainfall) to train our models and create predictions for future snow levels at each of the five resorts. Since we plan on testing multiple weather features with differing units, we will apply a scaled data frame to models with multiple numeric features. From our ski knowledge, we expect to see features like snowfall heavily impact the snow depth though, from our preliminary analysis of the correlations, it is possible that is not the most important feature. We will use a random forest regressor to weigh the impact of the different features on snow depth to determine which are more impactful in reducing variability. Once we identify which features are most important, those will be the features we include in our multiple regression model to predict snow depth at each resort. With our initial visualizations not indicating a clear linear or polynomial curve in the snow depth data, we will test a polynomial regression model as well to ensure that we create the most accurate model.

The methods that we have chosen to analyze our data allow us to compare the benefits of two machine learning models, polynomial and multiple regression, side by side. The advantages of this method are that we will be using the data from one API pull for all our analyses to keep consistency in the numbers. However, some pitfalls of this method are that we are only testing two models out of a large range of possibilities. There might be different machine learning techniques that can produce better predictions.

Additionally, the method we are using takes data from one data source with a limited number of resorts. A more accurate model might be possible by combining data from multiple sources or including a larger sample of resorts. Our analysis will be a foundation for future models to be created using more resorts or data. We created functions in our code to allow other data scientists to input more information and experiment with more complicated models.

The methods used in this analysis are suitable for our data given that our initial visualizations do not reveal a clear pattern, suggesting the absence of a specific model type as the optimal choice. Additionally, the majority of our weather data is numerical, making regression a very beneficial tool for prediction. Drawing on our understanding of the various factors influencing ski seasons, we recognize the multitude of elements affecting snow depth and the overall skiing experience. Consequently, employing complex, not simple, prediction models becomes most appropriate. Given our current knowledge of machine learning tools, the chosen methods represent the most effective approach for constructing a prediction model without unnecessarily over-complicating the models.

Results

When applying our machine learning tools, we began by running a feature importance test to identify which features most heavily impact snow depth. Our goal is to increase the accuracy of our models by including the features that are most impactful in reducing variability while not over complicating the model. From the graph below, it is evident that certain features are more impactful than others. The daily max temp, daily min temp, daily mean temp, daily snowfall sum, and daily max wind speed account for about 97.6% of the variability of snow depth. Thus, we will use these features in our multiple regression model to minimize some complexity.

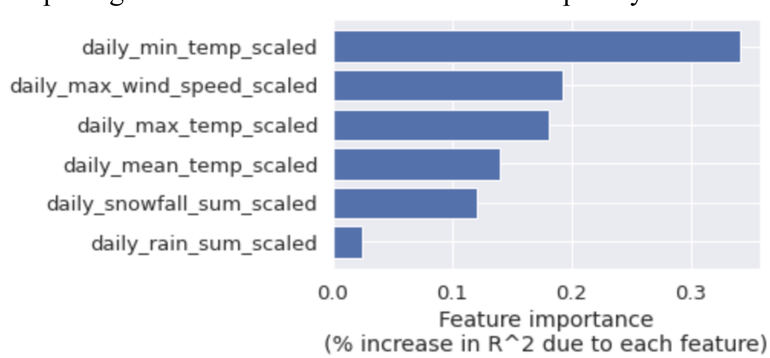

Exhibit 7. Feature Importance

After identifying the important features, we ran a multiple regression model on all historical weather metrics: daily max temp, daily min temp, daily mean temp, daily snowfall sum, daily max wind speed, and daily rain. We used single-fold cross-validation to train our model on a subset of the data and ran our model on the test dataset using a 70-30 split. The cross-validated mean squared error of the multiple regression model using all features was 0.775, and the r^2 score was 0.193. This r^2 score indicates that using a multiple regression model with daily max temp, daily min temp, daily mean temp, daily snowfall sum, daily max wind speed, and daily rainfall explains 20% of the differences in snow depth.

According to our feature importance plot, daily rainfall is the least impactful in accounting for the variability in snow depth. We re-ran our multiple regression model excluding daily rainfall and saw a slight increase in r^2 to 0.237. This highlights that we can achieve a slightly better model for predicting snow depth by not including rainfall as a feature. As mentioned in our method section, we want to be sure we are implementing the best model for predicting snow depth. Since our initial visualizations did not clearly look linear or polynomial, we will test both before running a model on our full dataset.

Similar steps were followed for our polynomial regression model. We started by creating the x and y arrays based on the feature importance graph and non-scaled data frame. We used minimum temperature as the x-feature because it is most important in eliminating snow depth variability as seen in Exhibit 7. We tested the r^2 terms from each polynomial degree.

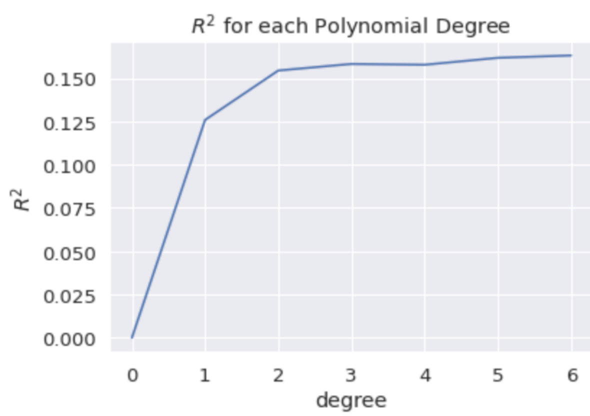

Exhibit 8. r^2 for Varying Polynomial Degrees

The graph above has an elbow at degree 2, but continues to slightly increase through the other degrees. In the interest of creating the most accurate prediction model with the least amount of complexity, we elected to create a polynomial of degree 2. In performing single-fold cross-validation on our data, we trained our model on the training data and ran it on the test data. The cross-validated mean squared error of the polynomial regression model was 9.20 and the r^2 was 0.14. This indicates that the polynomial model of degree 2 using minimum temperature accounts for only 17% of the variability in snow depth. In comparing both the MSE and r^2 from our multiple and polynomial regression models, we found that the multiple regression model has both a lower MSE and higher cross-validated r^2 percentage, so it is the better model.

The analysis of cross-validated r^2 shows that a multiple regression model is best for predicting snow depth with our data, so we proceeded with that model. We ran our multiple regression model on the full data set and got a MSE of 0.778 and r^2 of 0.222. This r^2 indicates that 22% of differences in snow depth are accounted for by the features on our important features list. This is not a good r^2 value as it indicates that less than ¼ of the variability in snow depth is accounted for in our model. Therefore, this model does not explain the variance of the data to an extent that results in usable predictions. This highlights to us that predicting weather is a complicated task and there must be other factors that our model is not taking into account.

Discussion

Multiple Regression Model

$$\hat{y} = 2.11 - 0.07x_1 + 0.77x_2 - 0.44x_3 + 0.11x_4 - 0.23x_5$$

where

- y: `snow depth`
- x_1: `daily max temp scaled`, maximum temperature of that day
- x_2: `daily min temp scaled`, minimum temperature of that day
- x_3: `daily mean temp scaled`, mean temperature of that day
- x_4: `daily snowfall sum scaled`, total snowfall of that day
- x_5: `daily max wind speed scaled`, maximum wind speed of that day

Exhibit 9. Final Multiple Regression Model

Above is our final model in the context of our analysis. Our results from the multiple regression model showed an r^2 score of 0.222, revealing that only 22% of the variance of daily snow depth is attributed to our X features of daily min, max, mean temp, daily snowfall sum and wind speed. Therefore, in the context of predicting snow depth in order to better inform skiers and resorts to maximize potential for the season, our model does not perform well and would leave a lot up to chance. Our results should therefore not be taken at face value as the model assumptions were not sufficiently met.

Although the r^2 is not good, we still checked the multiple regression model assumptions to see if they were met and if not, we discussed avenues to explore in creating a generalizable model. Our independence plot displays a pattern shown amongst the residuals from our model. Our data set is representative of four snow seasons from 2018, 2019, 2020, and 2021. This explains the four groups within the graph that look really similar to each other. Within each year section there are grouped clusters that are representative of each resort. The last resort seems to be routinely more spread out which may indicate a dependency on snow depth based on the resort. We could stratify by resort, then stratify by time and determine if a similar pattern is observed. We would say that this assumption is not met.

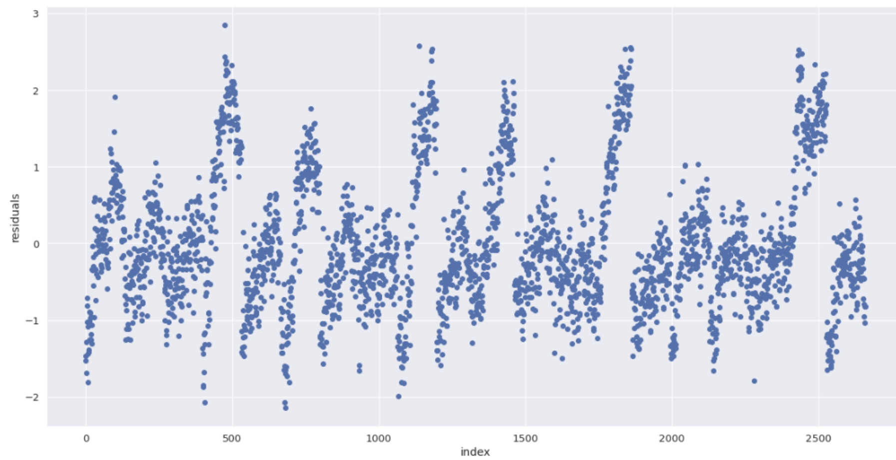

Exhibit 10. Checking Independence Plot

We then went on to check consistent variability among our five features. Max temp is the singular feature that appears to have constant variability across various temperatures. The remaining four features all have some aspect of funneling for various values of the respective x feature as shown in the visuals below. For example, take the graph for snowfall sum. As you move right towards greater snowfall sums, there is less variability within the residuals. This implies that our model is worse at predicting smaller values of snow fall sums than larger ones and does not meet the constant variance assumption. Next steps to try to alleviate this could be transforming these features to fix the issues of funneling and clustering.

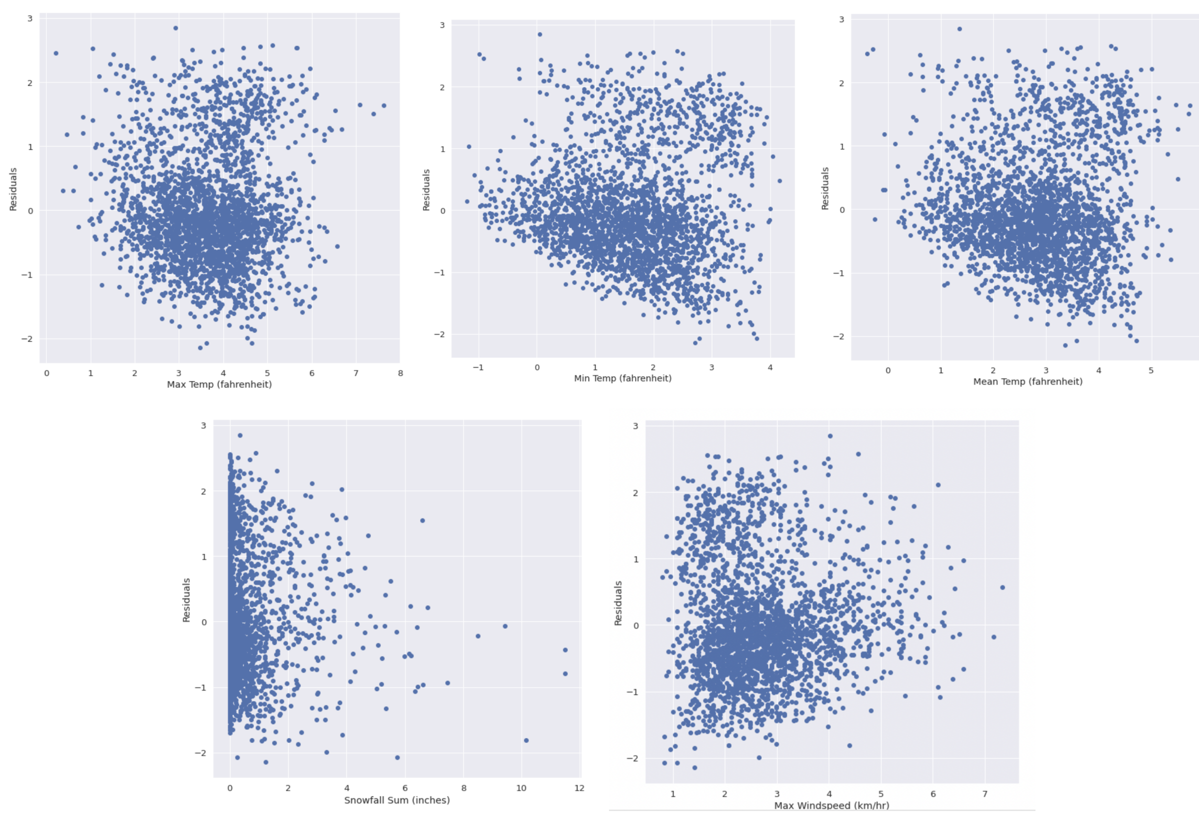

Exhibit 11. Checking Constant Variance Plots

In checking the assumption of normality, we found that a lot of the errors do not lay close to the 45 degree line. The residuals appear to consistently deviate from the line in a curve-like pattern. From this, we can conclude that this assumption is not met. Since neither of the three assumptions we met, we can conclude that our Multiple Linear Regression model was not successful in predicting our target variable and is not generalizable.

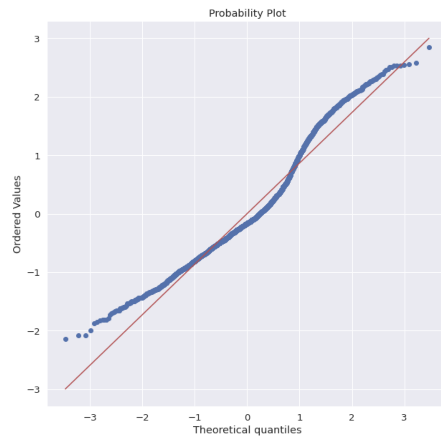

Exhibit 12: Checking Normality Plot

Below, our final visualization uses a distribution plot to show the actual target variable of snow depth in relation to how our model predicted snow depth. As concluded from our rather low r-squared, our model does not do the best in predicting resort snow depth using the features of ax temp, mean temp, min temp, and wind speed. This visualization leads us to the same conclusion as our model goes from under predicting to over predicting back to under predicting again— very inconsistent.

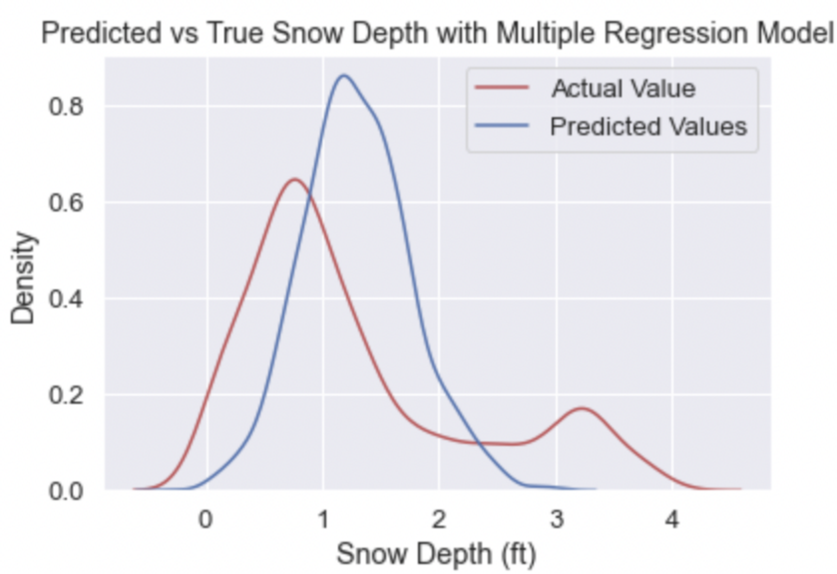

Exhibit 13: Predicted vs True Snow Depths

We were able to formulate some initial answers to our data analysis questions based on our analysis, but feel there is further exploration needed to generate more accurate results. Regarding predicting future snow depth and weather conditions, we were able to make predictions, but the accuracy was very limited. Therefore, exploring different models and potentially incorporating additional features might lead to more accurate results. Since our model was not very accurate, it will not be helpful in determining the most optimal resort for skiing. However, based on our initial visualizations, we can see that Whistler consistently has a greater snow depth than any of the other resorts throughout the ski season. Since all the resorts had fairly similar average temperatures, we can conclude that Whistler resort will be the most optimal for skiing. This conclusion mirrors real life, as out of the resorts that we examined, Whistler is by far the most popular and is known to be one of the top-rated ski resorts in North America.

The unanticipated questions that arose from our analysis and our Multiple Regression Model were discovered by checking our assumptions in the graphs above. The daily snowfall sum in inches is greatly skewed by its outliers, and therefore, the errors over daily snowfall sum are not randomly scattered and do not vary constantly. This was an unanticipated result, although it should have been logically assumed that outliers of heavy snowfall over a single day would cause such deviations. However, in the context of ski seasons and the desire for deep powder days for optimal conditions, including these heavy snowfall days is crucial in visualizing the quality of each resort. To adjust the model in future work or analysis, we could choose to focus solely on these outliers, and therefore, resorts that have the greatest number of powder days. Shifting the analysis to define predictions around these high snow levels in a short amount of time could lead to more interesting and anticipated results. Another alternative that we can explore is specializing our dataset to focus only on one resort over a longer period of time. This approach might yield better results in our model as it would be based on data from a more extended timeline and predict our target variable based on a set of weather conditions from one place rather than four.

Sources

Gerretsen, Isabelle. “Are Ski Trips Coming to an End?” BBC News, BBC, 14 Apr. 2023, www.bbc.com/future/article/20230124-how-climate-change-threatens-to-close-ski-resorts

Vaughan, Adam, et al. “Does Skiing Have a Future?” The Times & The Sunday Times: Breaking News & Today’s Latest Headlines, The Times, 10 Jan. 2023, www.thetimes.co.uk/article/does-skiing-have-a-future-ssxbwdtvj

“Historical Weather API.” *Open-Meteo*, https://open-meteo.com/en/docs/historical-weather-api#hourly=snowfall,snow_depth&daily=temperature_2m_max,temperature_2m_min,temperature_2m_mean,snowfall_sum&temperature_unit=fahrenheit&precipitation_unit=inch&timezone=America%2FLos_Angeles