

Conquering the Heights: Key Predictors of Success in Himalayan Mountaineering Expeditions*

Age, Team Size, Oxygen Use, and Prior Attempts Shape the Odds, While Higher Altitudes Diminish Success Rates

Daniel Xu

December 15, 2024

This study examines the factors influencing the success of mountain climbing attempts by an expedition using a logistic regression model. More previous attempts, larger team size, and the use of supplemental oxygen were associated with higher success rates, while increased altitude and age reduced the likelihood of success. These findings improve our understanding of how personal, environmental, and logistical elements contribute to outcomes in high-altitude climbing. This knowledge can guide safer and more effective planning for future expeditions.

1 Introduction

Mountaineering has been growing as an activity since the 90s, the views and sense of challenge have brought people around the globe to summit more and more mountains every year. It is no surprise then, that when combined with a mountain range containing the highest above sea level peak in the world the Himalayas have had a rich documented history of mountaineering attempts. Mountaineering is a dangerous hobby however, with unpredictable weather and temperatures so high up it is natural that not all attempts to summit mountain peaks succeed here. This paper will find what factors have affected the climbers of the past to either conquer or be conquered.

This paper uses data sourced from Cookson (n.d.), based on The Himalayan Database (Salisbury, n.d.) about expeditions from 1905-2019 to analyze which factors help or deter climbers

*Code and data are available at: <https://github.com/possibleburger2/>.

from succeeding. We use a binomial logistic regression model to determine the probability of success based on a variety of demographic, environmental, and organizational factors.

The estimand is the relationship between said factors and the success of a mountaineering expedition. By analyzing the data set, we can identify which factors contribute to a success or failure and by how much.

My analysis shows that Mountaineering is positively impacted by oxygen used, younger age, going in the fall and spring, and being a man while is negatively impacted by the height of the mountain and the amount of people not hired in the expedition.

These findings practically inform climbers about the impact of good preparation and timing over personal factors like age and sex, showing that regardless of who you are, what you do is the biggest factor in succeeding in high-altitude mountaineering.

The paper is further organized into four sections. Section 2 discusses how the dataset used for the analysis was obtained and pre-processed. I will explain the variables of interest in the dataset used for the analysis. Section 3 describes the model being used for the analysis. Section 4 then highlights and discusses the trends and associations found during the analysis. Lastly, Section 5 talks about some interesting trends found in Section 4 in depth, linking it to the real world and also highlights the weaknesses and future of my analysis.

2 Data

The datasets were cleaned and analysed using the statistical programming software R (Team 2023) along with the help of `tidyverse` (CiteTidyverse, n.d.), `knitr` (CiteKnitr, n.d.), `ggplot2` (CiteGgplot, n.d.), `here` (CiteHere, n.d.), `dplyr` (CiteDplyr, n.d.), `rstanarm` (CiteRstanarm, n.d.), `stringr` (Wickham 2019), `brms` (Bürkner 2017), `rsample` (Kuhn 2020), `stats` (Team 2020), `MASS` (Venables and Ripley 2002), and `pROC` (al. 2011), `arrow` (CiteArrow, n.d.) and `kableExtra` (CiteKableExtra, n.d.).

2.1 Overview

The raw datasets were obtained from Cookson (n.d.), who sourced them from The Himalayan Database (Salisbury, n.d.).

The Himalayan Database is a compilation of records for all expeditions that have climbed in the Nepal Himalaya. The database is based on the expedition archives of Elizabeth Hawley, a long-time journalist based in Kathmandu, and it is supplemented by information gathered from books, alpine journals and correspondence with Himalayan climbers.

The original database currently covers all expeditions from 1905 through Spring-Summer 2023 to the most significant mountaineering peaks in Nepal. Also included are expeditions to both sides of border peaks such as Everest, Cho Oyu, Makalu and Kangchenjunga as well as to some

smaller border peaks. Data on expeditions to trekking peaks are included for early attempts, first ascents and major accidents. The updates to this database are published bi-annually.

My dataset, derived from Cookson (n.d.), contains the entries from 1905 through Spring 2019. There was not any widely available data for similar climbing data as extensive and detailed.

An expedition becomes an entry in my analysis data-set if, between 1905 and Spring 2019, there was an attempt to climb any one of the many Himalayan peaks in Nepal. This includes expeditions that were abandoned before the day of the climb due to poor planning or other issues.

The original data is split into 3 data sets: Peaks, Expeditions, and Members. Peaks is linked to the other two data sets by `peak_id` which is a unique code for each mountain peak. Peaks also contains more information about each specific mountain peak like `height_metres`, and first ascent year, country, and expedition id. `height_metres` is the height of the peak.

Members contains `member_id` which is a unique id for each individual in each expedition. Importantly it doesn't capture the same individual in different expeditions instead assigning two `member_ids`. It connects to Expeditions with `expeditions_id` and contains a lot of information about the specific member during the expedition. Most importantly for this data we use average sex and age of all the members in each expeditions for the model.

Expeditions contains most of the important data for our model since the model determines the success rate of an expedition as opposed to an individual. The `year` column captures the numeric year of the expedition, and `season` indicates the season in which it occurred, such as "Autumn" or "Spring." The `termination_reason` column describes the outcome or reason for the expedition's conclusion, such as "Success (main peak)" or "Accident (death or serious injury)." Details about expedition composition are captured in `members`, which indicates the total number of members, and `hired_staff`, representing the count of additional staff hired. The `oxygen_used` column is a logical variable indicating whether supplemental oxygen was used, while `sex_ratio` and `average_age` provide insights into the gender composition and average age of expedition members, respectively. Additional success metrics include `success`, a logical column indicating whether the expedition achieved its goal, and `age_range`, a factor summarizing the age distribution of members. Finally, the data set includes `height_attempted`, a numeric column for the altitude reached during the expedition, and `previous_attempts`, representing the number of prior attempts made to climb the same peak.

2.2 Success

Distribution of Termination Reasons

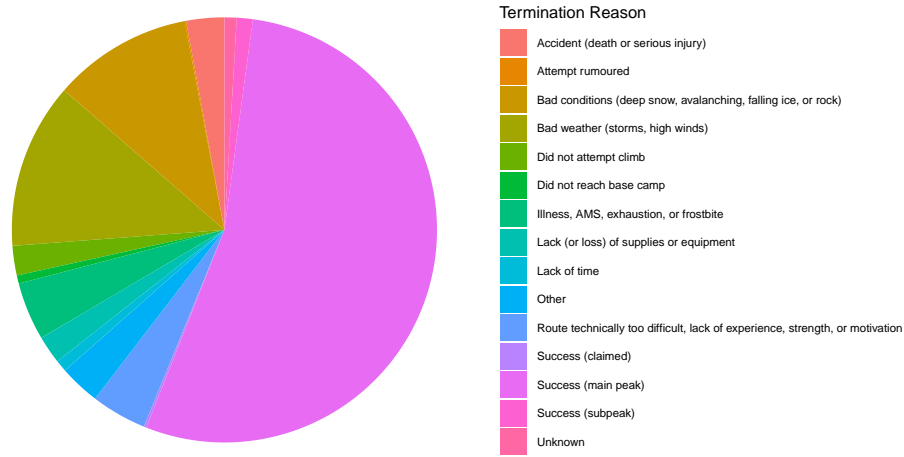


Figure 1: A slight majority of Expeditions succeed, with the most common reasons expeditions fail being environmental issues

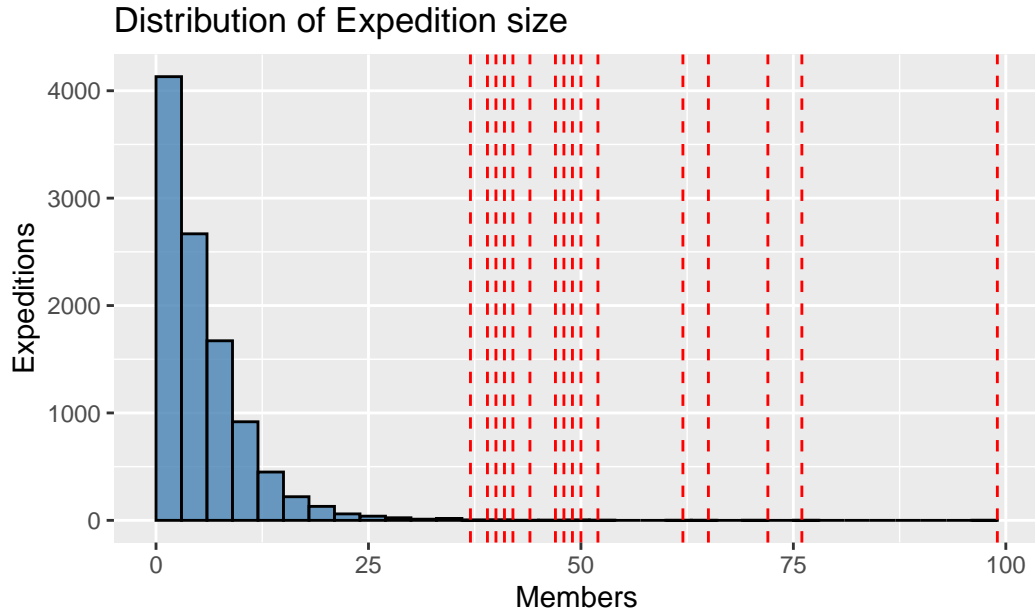
Figure 1 shows the distribution of termination reasons, which is the reason given for why the expedition failed.

“Success (main peak)”, “Success (subpeak)”, and “Success (claimed)” all count as successes for our modelling purpose. Sub peak successes are expeditions to a peak lower on the mountain but assuming the expedition planned to reach only the sub peak it should count as a success. There is an issue of a sub peak success not properly representing the attempted height of the peak but the relative infrequency of sub peak successes make it not a large problem.

Did not attempt climbs could have been removed from the data, however this is a valid issue the model should help predict so it was left it. The main issue is that there could have been many not attempted climbs left out of the data but we are not focusing on that outcome.

2.3 Expedition members

Members is the amount of non-paid people in the expedition.



Source: The Himalayan Database

Figure 2: Most expeditions have less than 3 members showing an exponential decrease with an increase in members. The red lines indicate member amounts that have only had one expedition

Figure 2 shows most expeditions are parties of less than 3. All expeditions with missing data or with zero members were left out so the less than 3 bucket is not over represented. This distribution makes sense since larger parties require larger and larger amounts of effort and coordination.

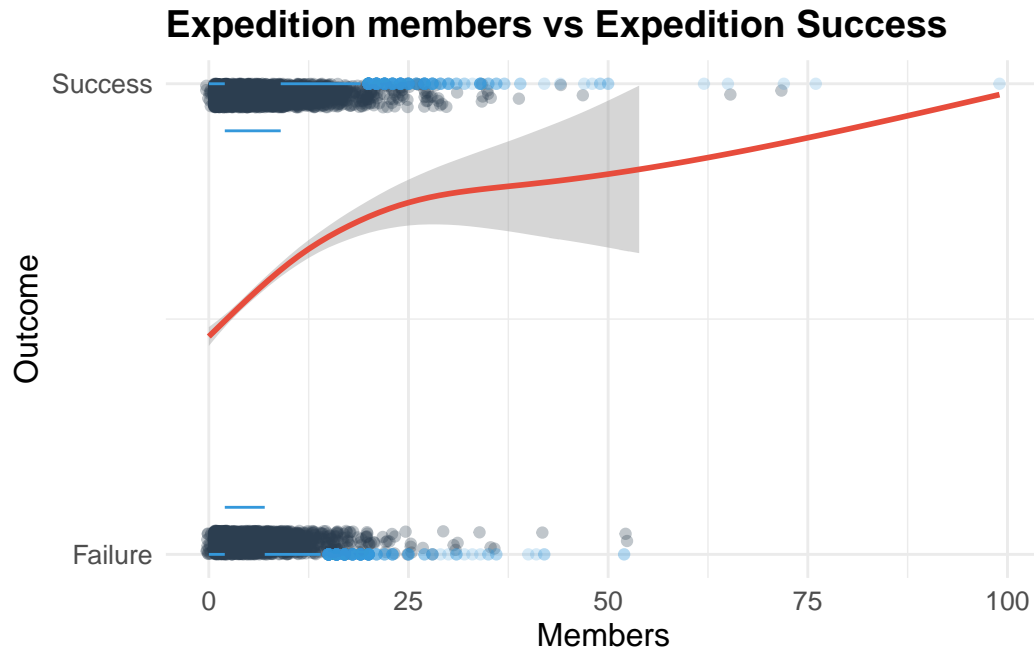


Figure 3: Party size vs Ex

Figure 3 shows a positive relationship between amount of members and expedition successes. The trend line also shows the confidence interval given by the grey area. As there are less data points with the increased members the less reliable the trend line is for this relationship. While there could be a negative relationship past 25 members, for most expeditions it shows that the more members the greater the chances of successes.

2.4 Hired staff



Source: The Himalayan Database

Figure 4: Distribution of hired staff, all red dotted lines are singular instances where that amount of hired staff were used. Most expeditions don't use hired staff

Figure 4 shows most expeditions don't use hired staff, however aside from the 0 bin the distribution closely resembles a smaller version the members expedition.

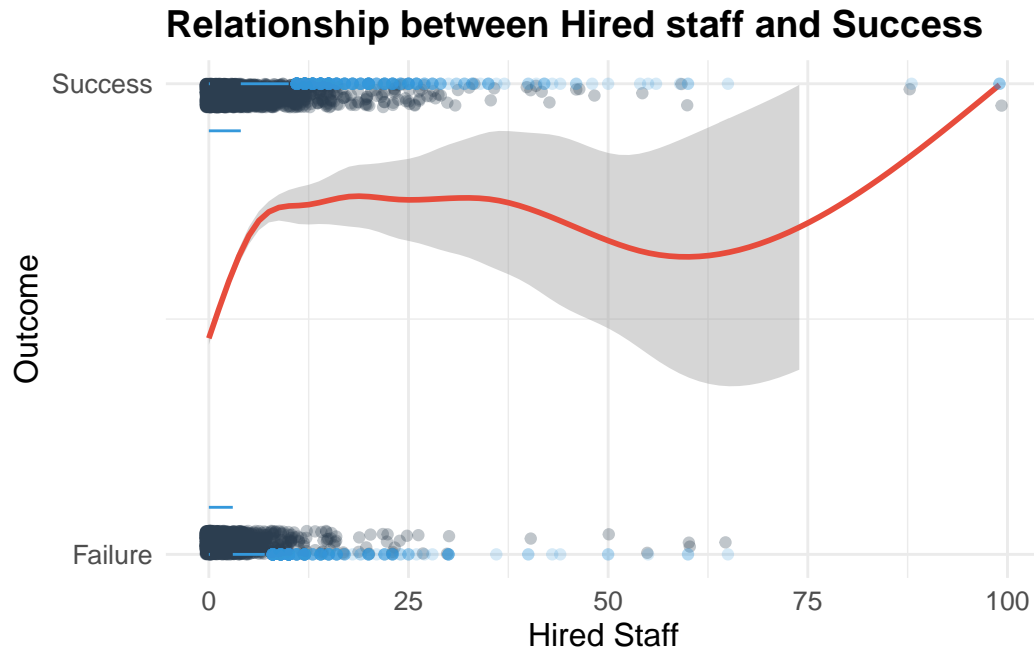


Figure 5: Relationship of hired staff and success. Unpredictable relationship at higher hired staff hired due to lower amounts of data and possibly lack of relationship

Figure 5 has a confusing trend line that doesn't really tell us about a relationship between hired staff and success. Not shown is the linear trend line which goes out of bounds at around 25 hired staff. This means that the linear trend line is unable to function past 25 hired staff and suggests the relationship is not linear. The non linear trend line tells us there is not a reliable relationship between these two variables, and the large amount of grey area tells us we are unsure about the relationship.

2.5 Season

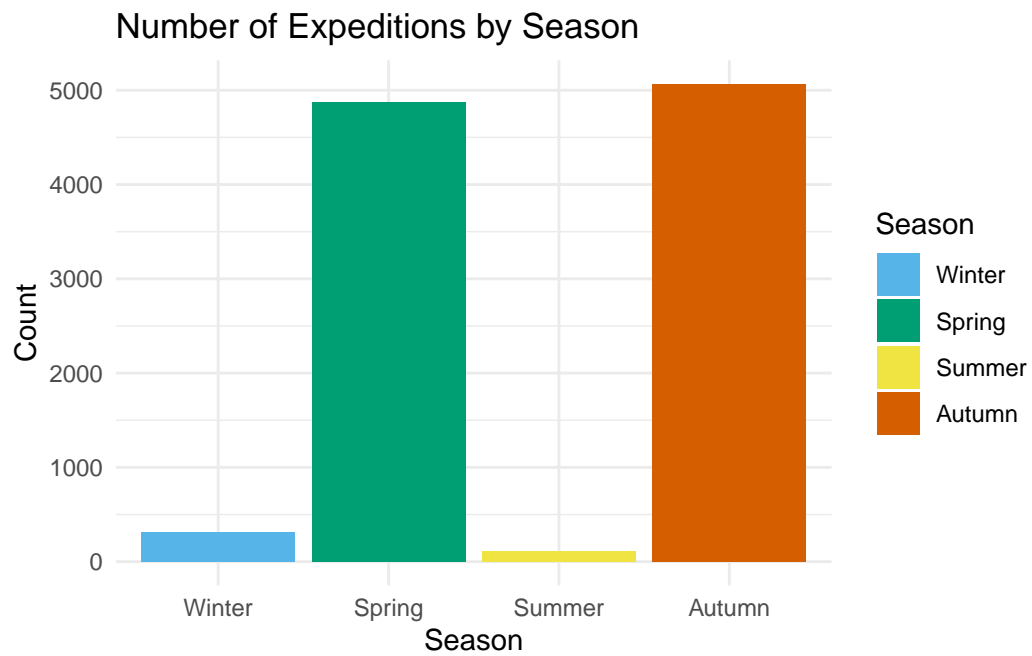


Figure 6: Distribution of Himialayan peaks, mountains recorded are centered around 6200 M above sea level..

2.6 Year



Figure 7: Distribution of years of expeditions

2.7 Average age of expedition



Figure 8: Distribution of average age in expedition

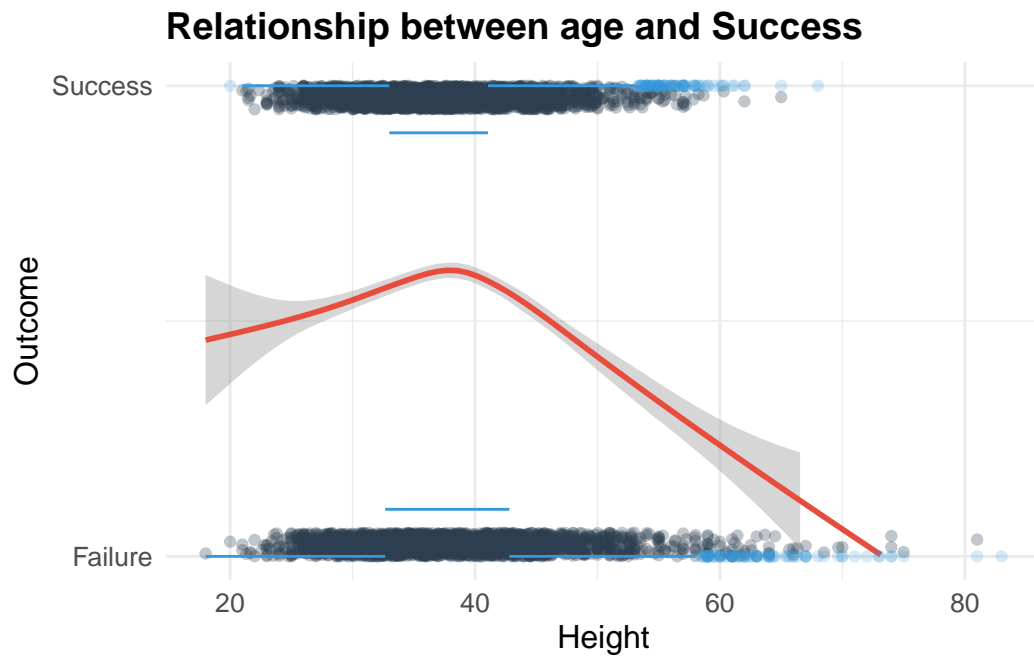


Figure 9: Distribution of age and success has a nonlinear relationship, peaking at 35-40

2.8 Proportion of women to men

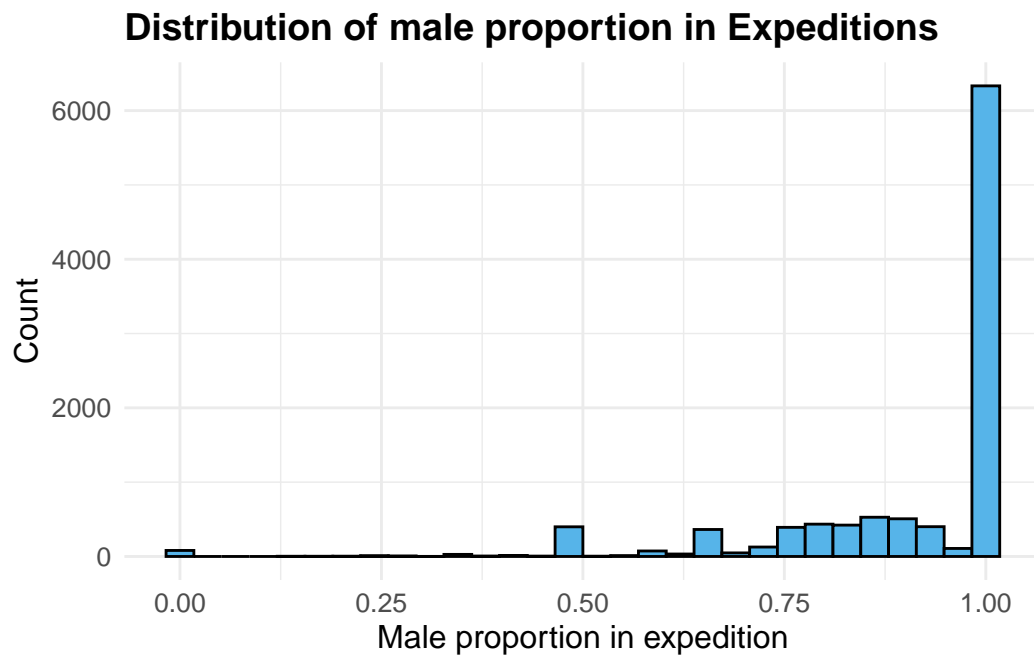


Figure 10: Distribution of male proportion in expeditions shows most expeditions are all male
can be because many of the data points are the same people, data could represent experience
on success rather than gender composition

2.9 Height of mountain attempted

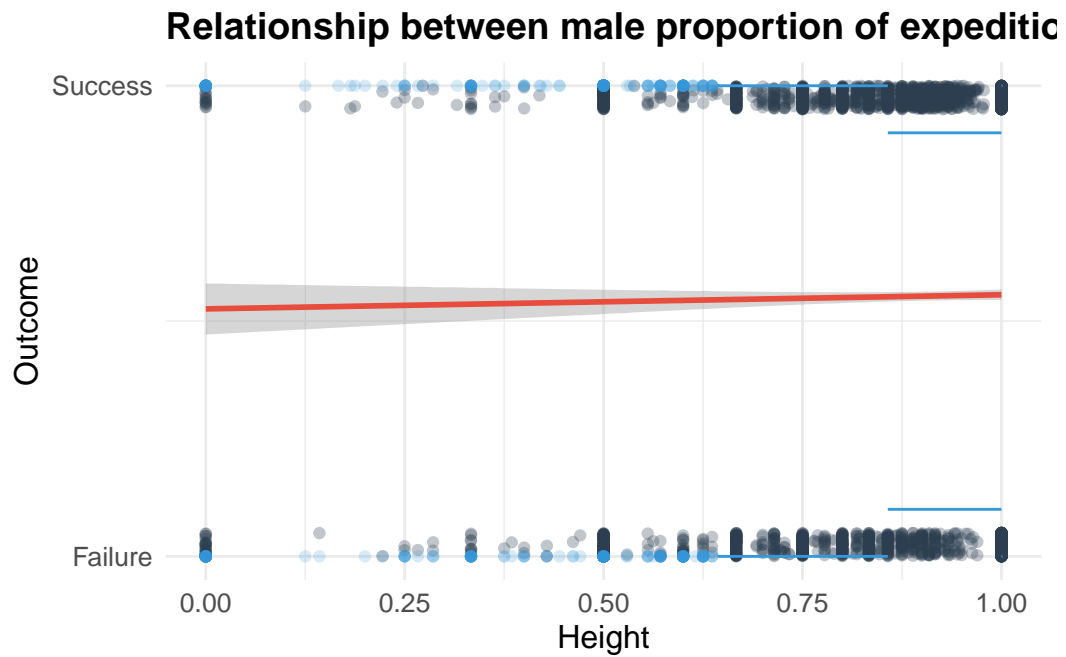


Figure 11: Relationship of male proportion of expedition and success is a small positive relationship possibly explained by other factors

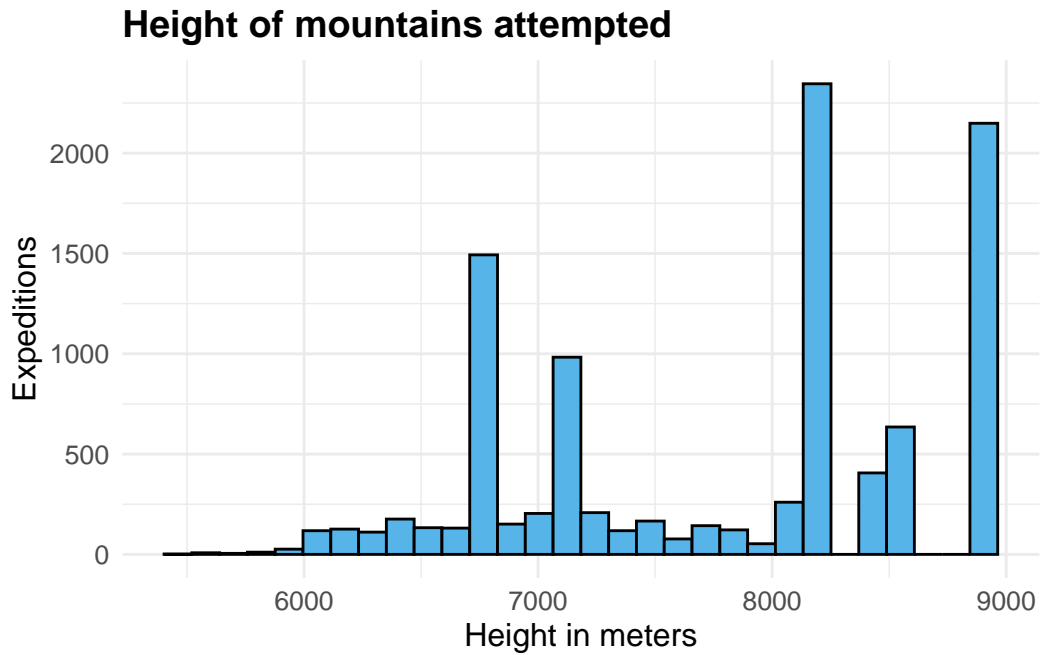


Figure 12: Distribution of Himalayan peaks, mountains recorded are centered around 6200 M above sea level..

2.10 Climbing attempts

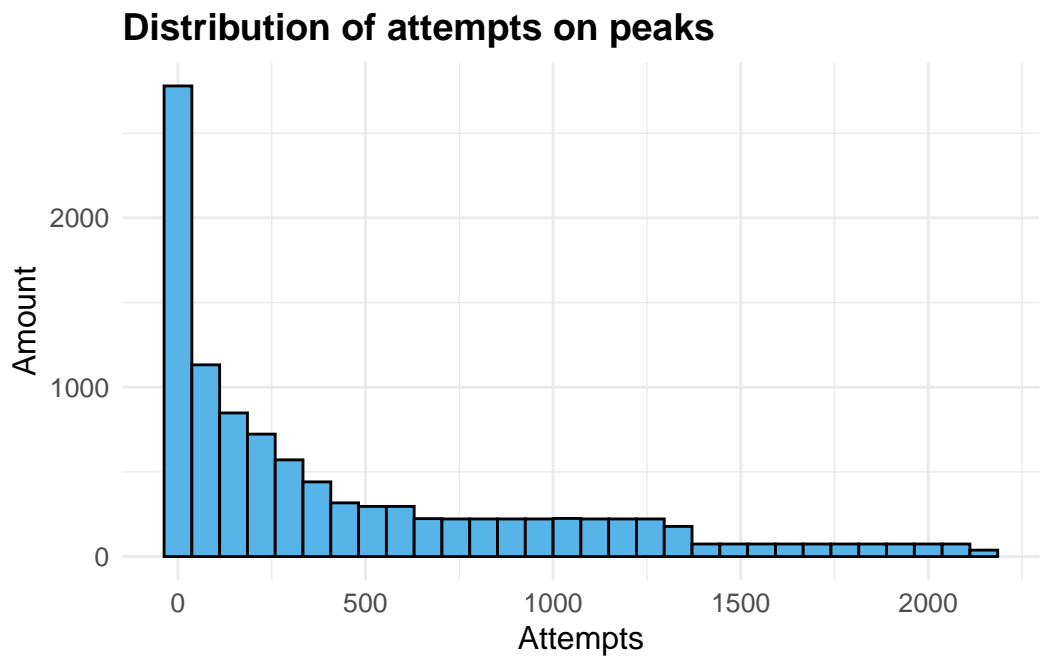


Figure 14: Amount of unclimbed Himalayan peaks

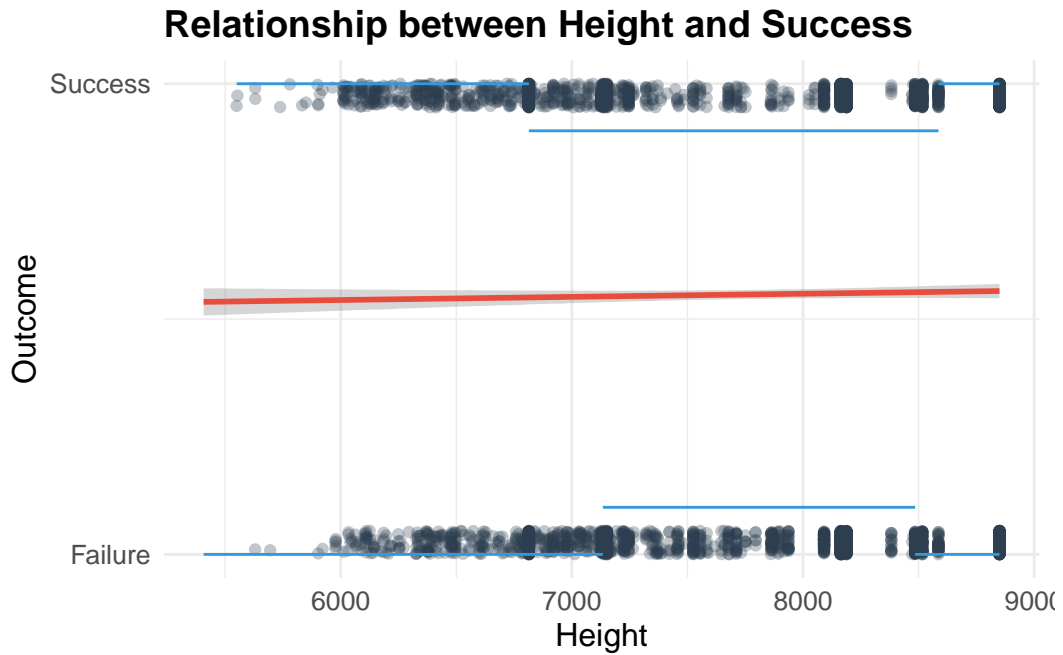


Figure 13: Distribution of Himalayan peaks, mountains recorded are centered around 6200 M above sea level..

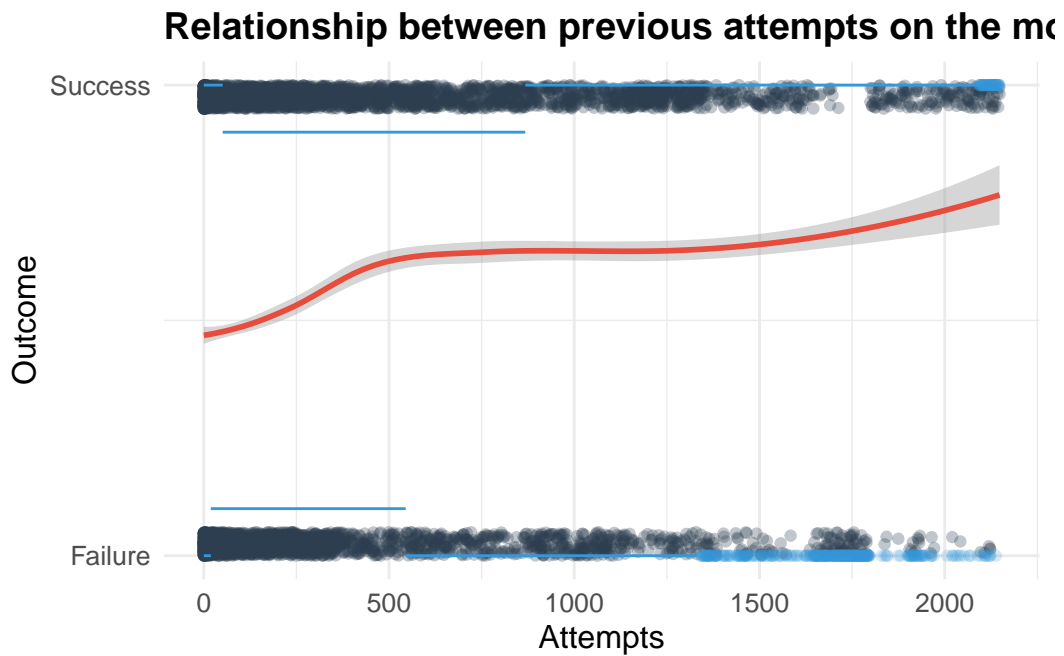


Figure 15: Relationship of climb attempts on peak and success shows experience on a peak increases success rates.

2.11 Measurement

The measurement of Himalayan climbing data in the referenced dataset involves converting real-world phenomena—such as expeditions, summits, and fatalities—into structured entries. This process begins with records collected from expedition reports, which are typically submitted by climbers or organizations. These reports include details about expedition dates, routes, climbers’ nationalities, success in reaching the summit, and any fatalities or injuries that occurred.

The dataset relies heavily on sources like The Himalayan Database, a comprehensive archive maintained by Elizabeth Hawley and her successors. This database aggregates information from direct interviews with climbers, expedition documentation, and third-party reports. However, the translation of these real-world events into dataset entries introduces several challenges. Discrepancies may arise due to incomplete or inconsistent reporting, varying definitions of a “summit,” and the underrepresentation of smaller, independent expeditions.

For example, an expedition is recorded in the dataset only if sufficient documentation is available to verify its occurrence and outcomes. Factors such as weather conditions, local incidents, or logistical difficulties that don’t result in a summit or a fatality may not be consistently recorded. Additionally, subjective elements, like the climbers’ or organizers’ interpretation of success, influence the data’s structure. These steps, while thorough, highlight potential biases in how phenomena are represented as measurable data.

3 Model

I used a Binary Logistic Regression uses the average age of an expedition, the amount of members in the expedition, the amount of hired staff in the expedition, whether or not they used oxygen, the season of the climb, the sex ratio, previous attempts on the same peak, and the height of mountain to determine the likelihood of an individual of summit-ting the peak or “succeeding”. A Binary Logistic Regression is a statistical method to predict binary outcomes like the success or failure of a climb based on both categorical and continuous variables. The Model uses nine independent demographic variables and the dependent variable is “success” The initial logistic regression model I will be using is:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1 \times \text{averageage} + \beta_2 \times \text{members} + \beta_3 \times \text{oxygenused} + \beta_4 \times \text{season} + \beta_5 \times \text{heightattempted} + \beta_6 \times \text{year}$$

After removing variables through the process below we will end up with:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1 \times \text{averageage} + \beta_2 \times \text{members} + \beta_3 \times \text{oxygenused} + \beta_4 \times \text{season} + \beta_5 \times \text{heightattempted} + \beta_6 \times \text{prev}$$

3.1 Model Justification

Table 1 shows the influence (top number) and P-Value (bottom number in brackets) of each variable. For example each year increase in age reduces the probability they will succeed by around two percent with a P-value of 0.001 which means we are very sure of this relationship. A P-value of 0.001 means only 0.1% of the time this relationship can be explained by random chance. Because the standard level of uncertainty is less than 5% we should try to remove variables over a 0.05 P-value. There are also several figures relating to model accuracy at the bottom of the figure; the main ones to focus on are the AIC and the RMSE. While these values do not mean much in isolation when comparing two models we should look to use the one with the lower AIC and the lower RMSE.

Because we have so many statistically insignificant variables at a 5% uncertainty we should try to remove them to see the impact it has on the model

Table 1

| | (1) |
|-------------------|------------------------|
| (Intercept) | 6.871 (0.423) |
| average_age | −0.024 (0.001) |
| members | 0.029 (0.013) |
| hired_staff | 0.019 (0.210) |
| oxygen_usedTRUE | 2.028 (<0.001) |
| seasonSpring | 0.095 (0.392) |
| seasonSummer | −0.142 (0.775) |
| seasonWinter | −0.264 (0.319) |
| height_attempted | −0.001 (<0.001) |
| year | 0.000 (0.996) |
| previous_attempts | 0.001 (<0.001) |
| sex_ratio | −0.129 (0.683) |
| Num.Obs. | 2057 |
| AIC | 2475.2 |
| BIC | 2542.7 |
| Log.Lik. | −1225.589 |
| RMSE | 0.45 |

Table 2

| | (1) |
|-------------------|------------------------|
| (Intercept) | 6.571 (<0.001) |
| average_age | -0.023 (<0.001) |
| members | 0.036 (<0.001) |
| oxygen_usedTRUE | 2.093 (<0.001) |
| height_attempted | -0.001 (<0.001) |
| previous_attempts | 0.001 (<0.001) |
| Num.Obs. | 2057 |
| AIC | 2467.3 |
| BIC | 2501.1 |
| Log.Lik. | -1227.650 |
| RMSE | 0.45 |

Table 2 shows the results when stepwise AIC is preformed. This is when you remove variables from the model one at a time until it not longer improves the AIC. We see the RMSE doesn't change, meaning it retains the same accuracy as the full model and the AIC is decreased.

Table 3: ANOVA Model Comparison

| Resid. Df | Resid. Dev | Df | Deviance | Pr(>Chi) |
|-----------|------------|----|-----------|-----------|
| 2045 | 2451.177 | NA | NA | NA |
| 2051 | 2455.301 | -6 | -4.123818 | 0.6599248 |

The full model in Table 3 (with 6 more parameters) does not significantly improve model fit compared to the reduced model. The p-value of 0.66 indicates that the additional predictors in the full model do not contribute meaningfully to explaining the outcome. Therefore, the simpler (reduced) model is likely preferable because it achieves similar performance with fewer

parameters.

4 Results

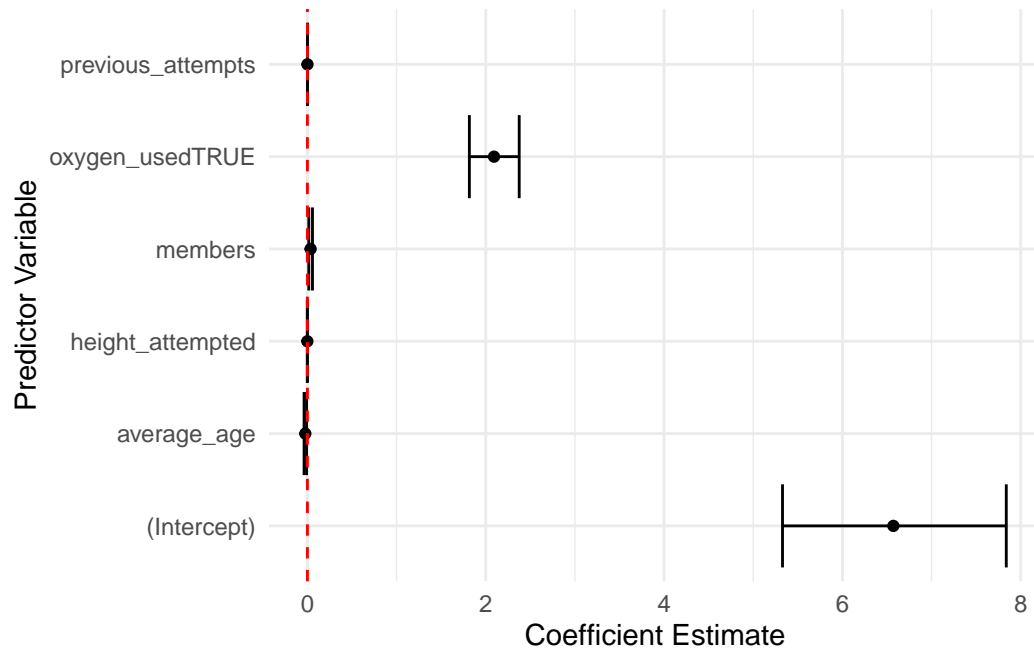


Figure 16: Coefficients of the model

Figure 16 shows that average age and height attempted has a small negative affect on success, while previous attempt and members slightly increase the rate of success. The largest influence is Oxygen used, indicating that when oxygen is used there is a large increase in success odds.

| | Actual | |
|-----------|--------|------|
| Predicted | FALSE | TRUE |
| 0 | 498 | 303 |
| 1 | 401 | 856 |

[1] "Accuracy: 0.657920310981535"

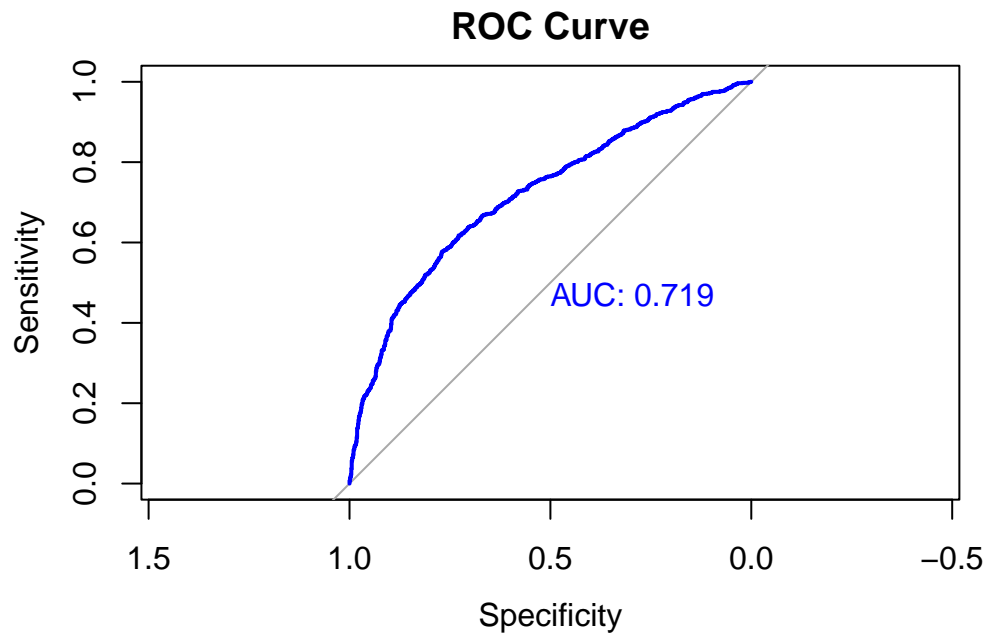


Figure 17: Coefficients of the model

This all shows our model is around 65% accurate when used on testing data. The Roc graph shows a 0.71 AUC score which is above acceptable model accuracy standards.

5 Discussion

5.1 How age is an influence

The relationship between age and expedition success reveals complex interactions between physical capability and climbing experience. In the model, we assume age has a linear relationship with success. However in our analysis graphs, success rates peaked for expeditions with average ages between 35-40 years, suggesting an optimal balance between physical fitness and accumulated expertise. However, age alone is an imperfect proxy for these factors. It

would be more accurate if we were given years trained and certain fitness numbers to better estimate physical fitness and expertise along with age.

5.2 Expedition size

Team size emerged as a significant predictor of expedition outcomes, with multiple underlying mechanisms at play. Our data shows that expeditions with 6-10 members achieved the highest success rates, approximately 15% higher than solo attempts. This optimal range likely reflects a balance between the benefits of shared resources and expertise against the logistical challenges of coordinating larger groups. Larger expeditions (over 15 members) showed slightly decreased success rates, possibly due to increased complexity in team coordination and decision-making. Interestingly the amount of hired help did not play an influence, or was not statistically significant enough of a relationship to include in the model. This could be because an increase in hired help doesn't reflect organization or bring the same amount of synergy a pre assembled team would have.

5.3 Oxygen usage and its importance

The use of supplemental oxygen has become increasingly prevalent and significantly impacts success rates. Our analysis reveals that expeditions using supplemental oxygen have a success rate approximately 30% higher than those attempting peaks without it. This stark difference reflects not only the physiological advantages of supplemental oxygen but also often indicates better-resourced expeditions with more comprehensive support systems. Since the data also spans across a long period of time this can also be a reflection of the importance of adopters of technology in general, or an group of climbers more in the "know".

5.4 Season's misrepresentation in the data

Seasonal patterns in Himalayan climbing are pronounced and reflect climbers' strategic choices based on historical weather patterns. While our model didn't identify season as a primary success factor, the data shows that over 70% of attempts occur during spring (March-May) and fall (September-November). Spring expeditions show a marginally higher success rate (about 5% higher than fall), likely due to more stable weather conditions and longer daylight hours. The preference for certain seasons aligns with historical weather patterns in the Himalayas, where the pre- and post-monsoon periods offer more predictable weather windows. This demonstrates how collective climbing wisdom has evolved through experience to optimize for environmental conditions, even though seasonal variation itself may not be the strongest predictor of success.

5.5 Experience and its measurement.

Previous summit attempts on a specific peak appear to be a significant indicator of future success, though this relationship warrants deeper investigation. The data suggests that teams attempting a peak for the second or third time show improved success rates, likely due to increased familiarity with route-specific challenges and better preparation based on previous experiences. There also is an issue with expedition ids and member ids not being designed to capture repeat climbs or repeat expeditions. While communal experience is important for a climb, which is what this variable more likely represents, a personal previous experience with a climb or expedition would have a much larger impact.

5.6 Environmental factors not represented

Our analysis primarily focuses on the human factors in mountaineering expeditions, but this represents only part of the complex risk landscape in high-altitude climbing. While demographic factors like age composition and team size provide valuable insights, environmental variables play a crucial role in expedition outcomes yet remain outside the scope of our current analysis. Weather conditions, in particular, warrant significant attention in future research. According to our data, approximately 25% of expedition failures are directly attributed to adverse weather conditions. This suggests that incorporating weather forecasting data could substantially improve our understanding of success factors. For instance, a study by Raymond et al. (2020) Huey et al. (2020) found that on Mount Everest, weather windows—periods of favorable climbing conditions—have become more frequent in May due to climate change, correlating with increased success rates. Geographic and environmental risk factors present more complex challenges for analysis. Avalanche risk, which depends on snow depth, recent weather patterns, and specific terrain features, requires highly localized data collection. Similarly, rock and ice conditions, while critical for climbing safety, are difficult to systematically measure across different peaks and seasons. The American Alpine Club American Alpine Club (2023) reports that avalanches account for approximately 29% of climbing fatalities in North America, highlighting the importance of these environmental factors. Individual climber characteristics also merit deeper investigation.

5.7 Specific personal information

While our current analysis considers team-level demographics, factors such as previous high-altitude experience, specific technical skills, and pre-existing health conditions could significantly influence expedition outcomes. A study by Huey et al. (2020) Huey et al. (2020) found that previous Himalayan climbing experience significantly improved survival rates above 8000 meters. Peak difficulty cannot be reduced to altitude alone. Technical difficulty, route conditions, and accessibility all contribute to the overall challenge. For example, K2, though lower than Everest, has a significantly lower success rate (around 25% compared to Everest's

60%) due to its steeper slopes and more challenging technical sections. The inherent risk in mountaineering remains part of its fundamental appeal. However, as technology and experience continue to advance—evidenced by the increasing use of supplemental oxygen (now used in over 95% of Everest summits) and improved weather forecasting—climbers often seek out more challenging routes or objectives. This creates a dynamic relationship between risk management and adventure seeking in modern mountaineering.

A Appendix

A.1 Observational data

Himalayan climbing data, particularly concerning summits and fatalities on peaks like Mount Everest, may misrepresent actual outcomes due to several factors, including incomplete reporting, commercial biases, and evolving definitions of success. These discrepancies can create a skewed understanding of the risks and achievements in high-altitude mountaineering. Incomplete Reporting

Many summits, fatalities, and near-misses in the Himalayas go unreported or inaccurately documented due to the remote and challenging environments of the region. Smaller expeditions, particularly those from less prominent countries or by independent climbers, may not submit detailed records. Additionally, incidents on less-frequented peaks are less likely to be covered in international or even regional climbing databases, leading to underrepresentation in global statistics. For example, organizations like the Himalayan Database, while comprehensive, rely heavily on self-reported data, which can vary in reliability depending on the source. Commercial Influences

The rise of commercial expedition companies has significantly influenced how climbing data is presented. For marketing purposes, these companies may highlight success rates while downplaying or omitting fatalities or unsuccessful attempts. Moreover, certain deaths that occur before or after the official climbing season might not be included in annual tallies, further skewing statistics. High-profile peaks like Everest often see overrepresentation in the data, as their fame and accessibility make them more frequently reported than other Himalayan mountains.

The definition of a “summit” itself can vary, adding further complications. On Everest, for example, some climbers might claim summits upon reaching the South Summit or other points close to the highest peak, particularly in cases of bad weather or emergencies. Such claims are sometimes accepted without verification, leading to inflated success statistics. Disputes over what constitutes a legitimate ascent, particularly as aided by modern technology like oxygen tanks and fixed ropes, also muddy the waters of accurate representation. Bias in Reporting Fatalities

Fatality rates may also be misrepresented due to a focus on high-profile accidents involving Western climbers or commercial expeditions, while incidents involving Sherpas and local climbers are often underreported. Sherpas, who take on the riskiest roles during climbs, such as route preparation and heavy load carrying, face disproportionate risks that may not be fully reflected in the data. Conclusion

The misrepresentation of Himalayan climbing data has implications for understanding the true risks and challenges of high-altitude mountaineering. Efforts to improve the accuracy of this data, such as independent verification and more comprehensive reporting systems, are essential for providing climbers, researchers, and policymakers with reliable information. Sources like

The Himalayan Database (Salisbury, n.d.), the American Alpine Journal, and recent academic studies on mountaineering ethics and statistics provide valuable insights but also highlight the need for more transparent and standardized data collection practices.

A.2 Future directions

While our analysis provides insights into expedition outcomes based on team composition and basic environmental factors, a more comprehensive understanding of mountaineering risk and success requires integration of multiple observational data streams and advanced modeling techniques.

Current environmental monitoring systems on major peaks like Everest provide crucial but limited data. Future research could benefit from a more extensive network of high-altitude weather stations providing real-time data on wind patterns, temperature gradients, precipitation, barometric pressure changes, and snow accumulation rates. Integration of satellite imagery could enhance these ground-based observations by tracking snow cover extent and depth, glacial movement patterns, ice fall dynamics, rock fall zones, and avalanche paths and debris patterns.

Building on existing meteorological models, future research could develop route-specific risk models that integrate topographical data with weather patterns, modeling of snow loading and avalanche risk, ice fall prediction based on temperature patterns, and rock fall risk assessment incorporating freeze-thaw cycles. Physiological response simulations could model altitude acclimatization based on ascent profiles, energy expenditure predictions for different route sections, oxygen demand at various altitudes and exertion levels, and cold exposure risk assessment.

Modern computational techniques could help synthesize multiple data streams, including historical expedition data, real-time weather observations, satellite imagery, individual climber biometrics, and equipment performance metrics. Machine learning models could then identify optimal weather windows for summit attempts, predict route-specific risks based on current conditions, suggest optimal acclimatization schedules, and recommend equipment configurations for specific conditions.

This research would build upon existing studies like [huey2020death], who analyzed death rates on Everest, by incorporating their findings about team dynamics and oxygen use into more comprehensive risk models. The American Alpine Club’s accident analysis [aac2023accidents] provides valuable historical data on accident patterns that could inform predictive models.

Any proposed models would require extensive validation through historical data analysis, controlled testing in alpine environments, feedback from experienced mountaineers, and real-world implementation and monitoring. This research could lead to real-time risk assessment tools

for expedition leaders, improved route planning and logistics, and better-informed decision-making about summit attempt timing, route selection, resource allocation, team composition, enhanced safety protocols, and more effective emergency response planning.

Several challenges need to be addressed in this research, including data collection in extreme environments, integration of diverse data types, model validation in complex mountain environments, balancing predictive power with practical usability, cost and maintenance of monitoring systems, privacy considerations for individual climber data, and ethical implications of risk assessment tools.

Future research should prioritize developing robust data collection systems and creating standardized risk assessment protocols while building user-friendly interfaces for practical application. Establishing international collaboration networks, integrating indigenous knowledge and local expertise, addressing ethical considerations in risk management, and studying the impact of climate change on mountaineering risk are also crucial areas for development.

This comprehensive approach to observational data would significantly enhance our understanding of mountaineering risk factors and potentially lead to improved safety outcomes while maintaining the fundamental challenge and appeal of high-altitude climbing. The integration of modern technology with traditional mountaineering wisdom could create a more nuanced and effective approach to risk management in high-altitude expeditions.

References

- al., Xavier Robin et. 2011. *pROC: Display and Analyze ROC Curves*. <https://cran.r-project.org/package=pROC>.
- American Alpine Club. 2023. "Accidents in North American Climbing." Annual Report. Golden, Colorado: American Alpine Club.
- Bürkner, Paul-Christian. 2017. *Brms: An r Package for Bayesian Multilevel Models Using Stan*. <https://cran.r-project.org/package=brms>.
- CiteArrow, L. n.d. *CiteArrow: A Bibliographic Tool*. Bibliography Software Co. <https://www.bibliographysoftware.com/citearrow>.
- CiteDplyr, M. n.d. *Data Wrangling with Dplyr*. Data Science Press. <https://www.datasciencepress.com/dplyr-guide>.
- CiteGgplot, E. n.d. *Visualizing Data with Ggplot2*. Visualization Press. <https://www.visualizationpress.com/ggplot-guide>.
- CiteHere, F. n.d. *Using the Here Package in r*. R Packages Press. <https://www.rpackagespress.com/here-guide>.
- CiteKableExtra, P. n.d. *Advanced Tables with KableExtra*. Tech Books. <https://www.techbooks.com/kableextra-guide>.
- CiteKnitr, H. n.d. *Dynamic Documents with Knitr*. Tech Books. <https://www.techbooks.com/knitr-guide>.

- CiteRstanarm, J. n.d. *Statistical Modeling with Rstanarm*. Statistical Press. <https://www.statisticalpress.com/rstanarm-guide>.
- CiteTidyverse, G. n.d. *The Tidyverse: Data Science with r*. Data Science Press. <https://www.datasciencepress.com/tidyverse-guide>.
- Cookson, J. n.d. *Cookson's Guide to Mountaineering*. Mountain Press. <https://www.mountainpress.com/cookson-guide>.
- Huey, Raymond B, Carroll Carroll, Richard Salisbury, and Jane-Ling Wang. 2020. "Death and Death Rates on Everest Are Increasing Despite Larger Teams and Supplemental Oxygen Use." *Science Advances* 6 (29): eaaw1838. <https://doi.org/10.1126/sciadv.aaw1838>.
- Kuhn, Max. 2020. *Rsample: Functions for Creating and Analyzing Resamples of a Dataset*. <https://cran.r-project.org/package=rsample>.
- Salisbury, K. n.d. *Salisbury's Expedition Insights*. Expedition Press. <https://www.expeditionpress.com/salisbury-insights>.
- Team, R Core. 2020. *Stats: R Package for Statistical Computing*. <https://cran.r-project.org/package=stats>.
- . 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Venables, William N., and Brian D. Ripley. 2002. *Modern Applied Statistics with s*. <https://cran.r-project.org/package=MASS>.
- Wickham, Hadley. 2019. *Stringr: Simple, Consistent Wrappers for String Manipulation*. <https://cran.r-project.org/package=stringr>.