Understanding Factors Contributing to a Successful Himalayan Expedition*

Analysis using data from Himalayan expeditions from 1905 through Spring 2019 leveraging Bayesian Logistic Regression

Kaavya Kalani

April 19, 2024

This study analyses the relationship between various demographic, environmental and geographic factors and the probability of success in attempting to summit a peak. Data from Himalayan expeditions in Nepal from 1905 through Spring 2019 is used and a Bayesian Logistic Regression model is utilised to analyse the trends and factors influencing a successful summit. Young age, being male and embarking on the expedition in spring or autumn are some factors that increase one's chances of a successful summit. The insights from this study aim to assist in future expedition planning, risk management, and safety protocols.

Table of contents

*Code and data supporting this analysis is available at: https://github.com/kaavyakalani26/himalayan-expeditions-analysis

1 Introduction

Mountaineering, with its blend of adventure and challenge, has captivated explorers for generations. It represents a pinnacle of human endeavor, testing physical strength, mental fortitude, and strategic planning against some of the most formidable natural landscapes on Earth. Amidst the allure of conquering majestic peaks, a critical question arises: what factors contribute to the likelihood of a successful summit attempt? This paper tries to unravel precisely this.

Numerous mountain ranges worldwide attract expedition enthusiasts, with the Himalayas, home to the highest peak, Mt. Everest, being one of the most renowned. This paper focuses on data from expeditions to the Himalayan mountain ranges in Nepal.

My estimand is the relationship between various demographic, environmental and geographic factors (such as the height of the peak, sex, age, season of the expedition and if it was a solo ascent) and the success of an expedition. By analysing the dataset, my goal is to identify trends and factors influencing successful expeditions and determine which factors contribute to a higher likelihood of success.

Data sourced from Cookson (2020), based on The Himalayan Expedition records (Salisbury 2023), is utilized to understand these factors and trends. A Bayesian Logistic Regression model is employed to predict the probability of a successful attempt based on demographic, environmental and geographic factors.

My analysis reveals that mountaineering is a predominantly male-dominated activity, with factors such as young age, male gender, and expeditions in spring or autumn increasing the chances of having a successful summit.

The findings of this research have practical implications for expedition planning, risk management, and safety protocol aiming to enhance decision-making in high-altitude mountaineering, making it safer and more informed in one of the world's most challenging environments.

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

For this analysis, I have combined three datasets into one, which is used for analysis. The datasets were cleaned and analysed using the statistical programming software `R` (R Core Team 2023) along with the help of `tidyverse` (Wickham et al. 2019), `knitr` (Xie 2014), `ggplot2` (Wickham 2016), `here` (Müller 2020), `dplyr` (Wickham et al. 2023), `rstanarm` (Goodrich et al. 2024), `arrow` (Richardson et al. 2024), `broom.mixed` (Bolker and Robinson 2022) and `kableExtra` (Zhu 2024).

2.1 Analysis Dataset

The raw datasets were obtained from Cookson (2020), who sourced them from The Himalayan Database (Salisbury 2023).

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 longtime 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's, contains the entries from 1905 through Spring 2019.

The three datasets I considered included information about all peaks in Nepal, all expeditions on those peaks and all members on those expeditions. I then combines the data from these three datasets to form the main analysis dataset.

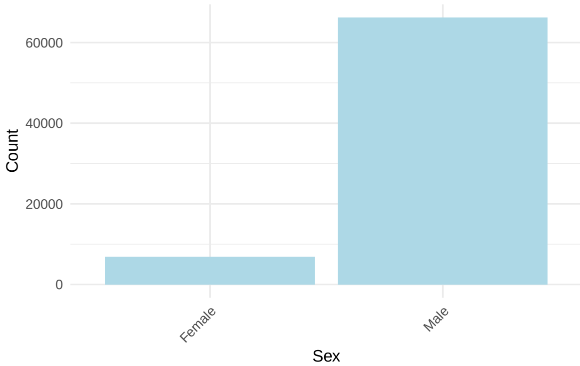
A person becomes an entry in my analysis dataset if, between 1905 and Spring 2019, they attempted to climb any one of the many Himalayan peaks in Nepal. It also included expeditions to both sides of border peaks as mentioned before.

Among the overall range of variables available, I chose the following to be included in the analysis dataset.

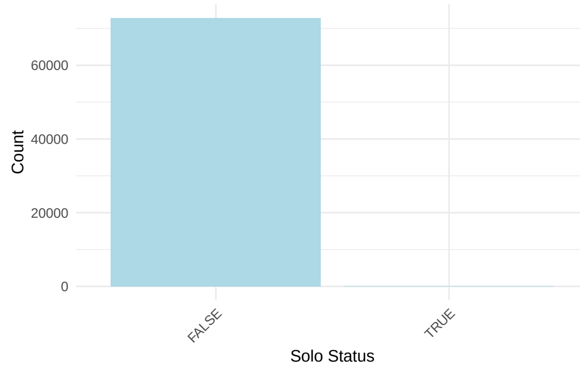
- `Height` is the height range in which the peak's height in metres falls. This is for the peak the person in the current entry is on an expedition for. The categories for this are 5400 - 5749, 5750 - 6099, 6100 - 6449, 6450 - 6799, 6800 - 7149, 7150 - 7499, 7500 - 7849, 7850 - 8199, 8200 - 8549 and 8550 - 8900.
- `Seasons` is the season the expedition is embarked on. This takes on either of the four values: autumn, spring, winter, summer.
- `Sex` is the sex reported by the expedition member and it is either male or female.
- `Age` is the age group in which the expedition member fell at the time of the expedition. Depending on the best available data, this could be as of the summit date, the date of death, or the date of arrival at basecamp. The different categories for this are Under 18, 19-30, 31-40, 41-50, 51-60, 61-70, 71-80 and 81-90.
- `Success` indicates if the person's expedition resulted in a successful summit.
- `Solo` indicates if the person attempted a solo ascent.
- `Died` indicates if the person died during the expedition.

Figure 1 displays the counts for various variables I will be modelling. The plot shows there are significantly higher men than women. This shows the existence of a gender imbalance skewed towards men in mountaineering in the Himalayas. Additionally, most expedition members opt not to attempt solo ascents and fall within the middle-aged range. Lastly, it can be seen that the peak expedition seasons are the more pleasant autumn and spring over the extreme seasons like summer and winter. These count disparities should be considered when analysing success proportions.

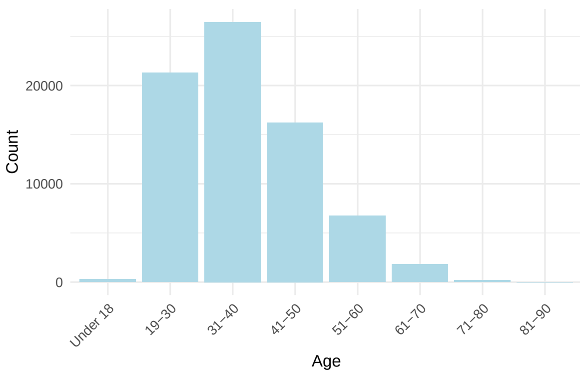
Figure 2 shows the proportion of expedition members who had a successful outcome or died during their expedition. For sex, it is observed that more men had successful expeditions compared to women. Moreover, individuals who embark on solo expeditions have a higher mortality rate, potentially explaining the difference in the number of people attempting solo ascent observed earlier. Some other trends that can be observed are that success rates decline with age, and winter contributes to the most deaths. The success and death proportions based on the height of the peak are very varied and do not follow a specific trend. This goes to show that just the height of the peak alone doesn't define the probability of having a successful summit but there are other factors that contribute to it.



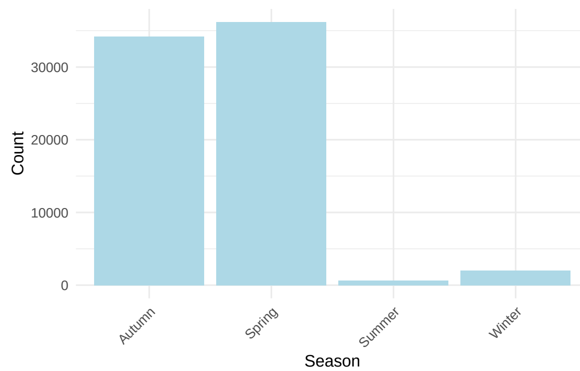
(a) Sex



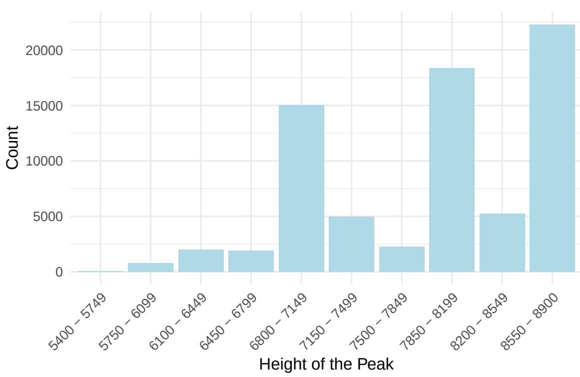
(b) Solo Status



(c) Age

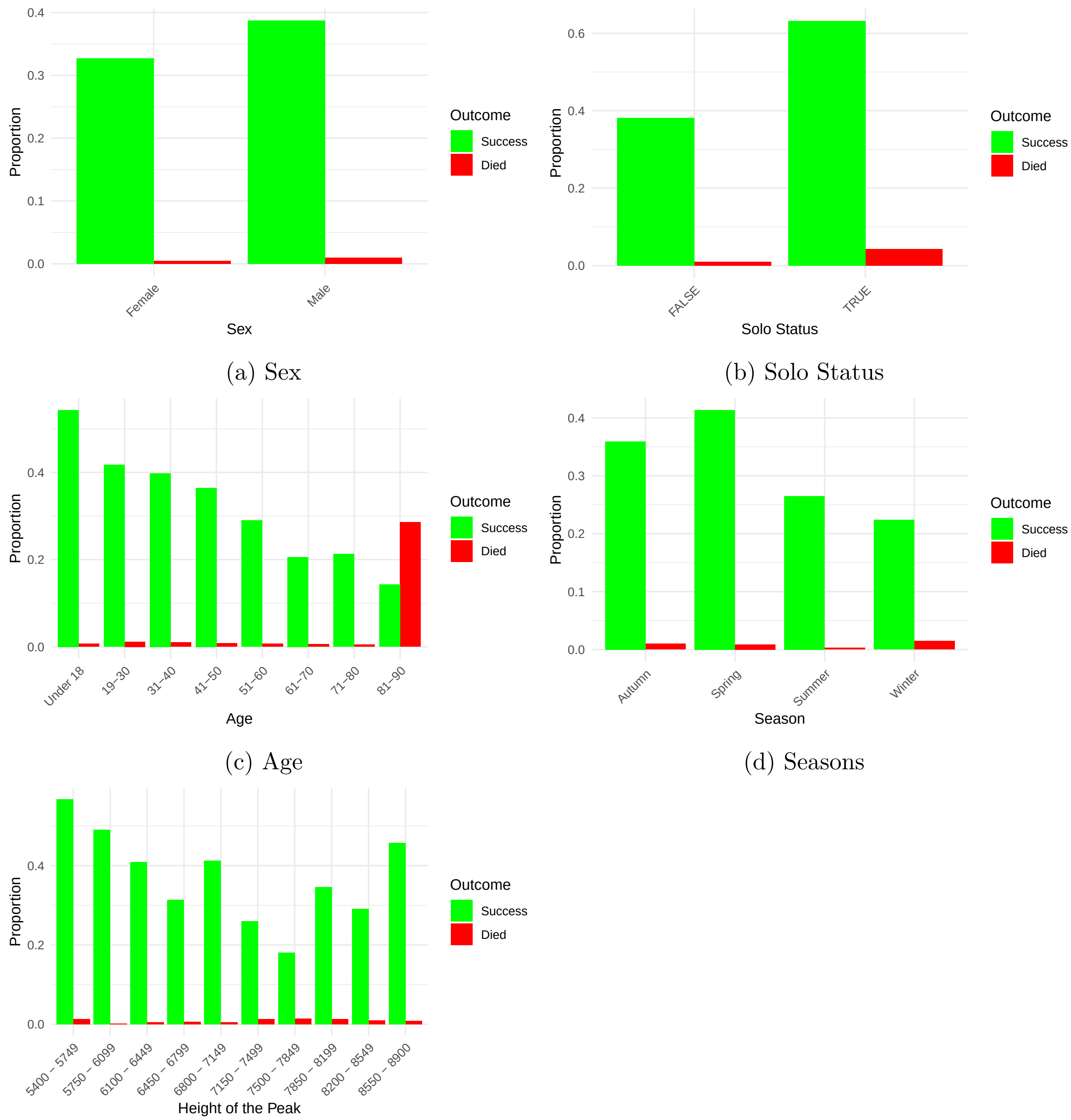


(d) Seasons



(e) Height of the Peak

Figure 1: Counts for the variables of interest



(a) Sex

(b) Solo Status

(c) Age

(d) Seasons

(e) Height of the Peak

Figure 2: Proportions for the variables of interest and the outcome of their expedition

3 Model

I used a Bayesian Logistic Regression model to determine the likelihood of an individual successfully reaching the summit of the Himalayan peak they are attempting to climb during the expedition. Logistic regression is a statistical technique used for binary classification to predict the probability of a categorical dependent variable.

My model will be based on five independent demographic variables: `height`, `sex`, `age`, `seasons` and `solo` and the dependent variable will be `success`.

The logistic regression model I will be using is:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1 \times \text{height} + \beta_2 \times \text{sex} + \beta_3 \times \text{age} + \beta_4 \times \text{seasons} + \beta_5 \times \text{solo} \tag{1}$$

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 2.5) \\
\beta_1 &\sim \text{Normal}(0, 2.5) \\
\beta_2 &\sim \text{Normal}(0, 2.5) \\
\beta_3 &\sim \text{Normal}(0, 2.5) \\
\beta_4 &\sim \text{Normal}(0, 2.5) \\
\beta_5 &\sim \text{Normal}(0, 2.5)
\end{aligned}$$

where,

- $\hat{p}$ represents the probability that someone will successfully summit the peak they are on the expedition for.
- β_0 represents the intercept term of this logistical regression. It is the probability that someone will successfully summit the peak they are on the expedition for if the predictors' values are zero
- β_1 is the coefficient corresponding to height of the peak
- β_2 is the coefficient corresponding to sex of the person
- β_3 is the coefficients corresponding to age of the person
- β_4 is the coefficients corresponding to seasons of the person
- β_5 is the coefficients corresponding to if the person is attempting the summit alone

In my model, normal priors with a mean of 0 and a standard deviation of 2.5 are used for both the coefficients and the intercept. Setting the mean of the priors to 0 implies that there is no expectation of a particular direction or magnitude for the coefficients or intercept. I chose this as I have no expectation of the same. The standard deviation of 2.5 reflects the uncertainty or variability in the prior beliefs. I chose a moderately wide prior to allow for a reasonable amount of uncertainty.

The chosen priors allow the data to largely determine the posterior distribution as they are relatively non-informative. They don't heavily influence the results unless the data provide strong evidence to the contrary.

The use of moderately wide priors can also help regularize the model, preventing overfitting and providing more stable estimates, particularly when dealing with limited data.

Model coefficients with their standard errors, posterior checks for the model and convergence checks for the model can be found in Section 6

4 Results

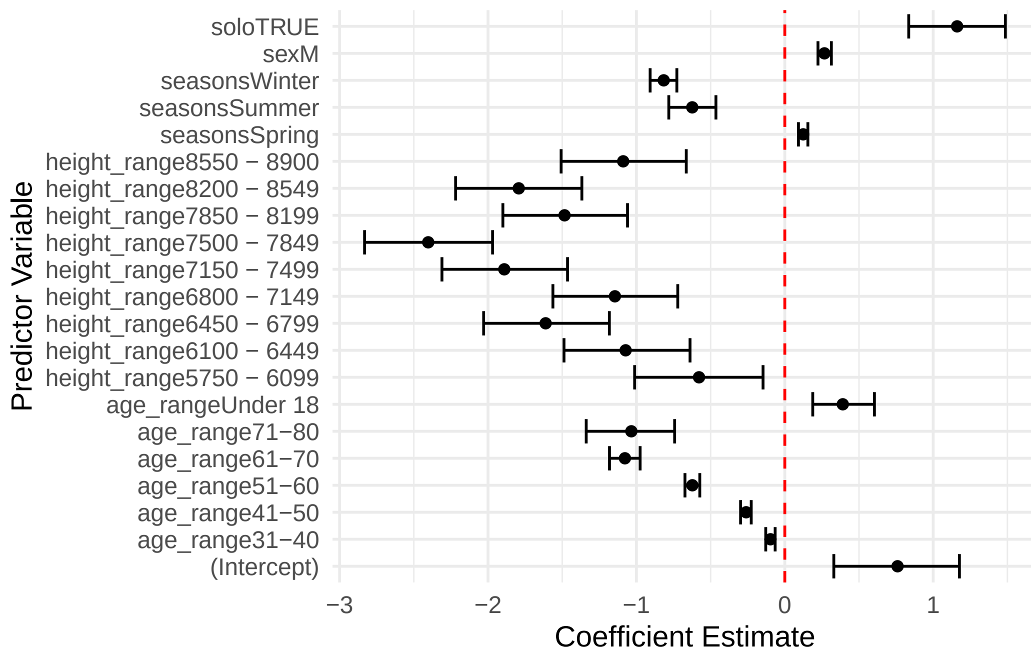


Figure 3: Coefficients of the model

Figure 3 illustrates the coefficients and their corresponding 95% credible intervals for the predictor variables in the Bayesian model. A 95% credible interval means that there is a 95% probability that the true parameter lies within the interval, given the observed data and the model assumptions. Each point represents the estimated coefficient for a predictor variable, with the horizontal lines indicating the credible interval around the estimate. Variables with coefficients to the right of zero indicate a positive association with the outcome variable, suggesting that an increase in the predictor variable corresponds to an increase in the outcome variable. Conversely, coefficients to the left of zero indicate a negative association, implying that an increase in the predictor variable is associated with a decrease in the outcome variable.

These insights can help in understanding the direction and magnitude of the relationships between predictor variables and the outcome in the Bayesian model. The coefficient for the age_range81-90 variable was removed due to its small magnitude, which was affecting the interpretability of other coefficients in the plot.

The analysis shows that males have a slightly higher success probability compared to females. Furthermore, expeditions in summer and winter show lower success probability than autumn, while spring shows a higher success probability. Additionally, the success probabilities decrease with increase in age.

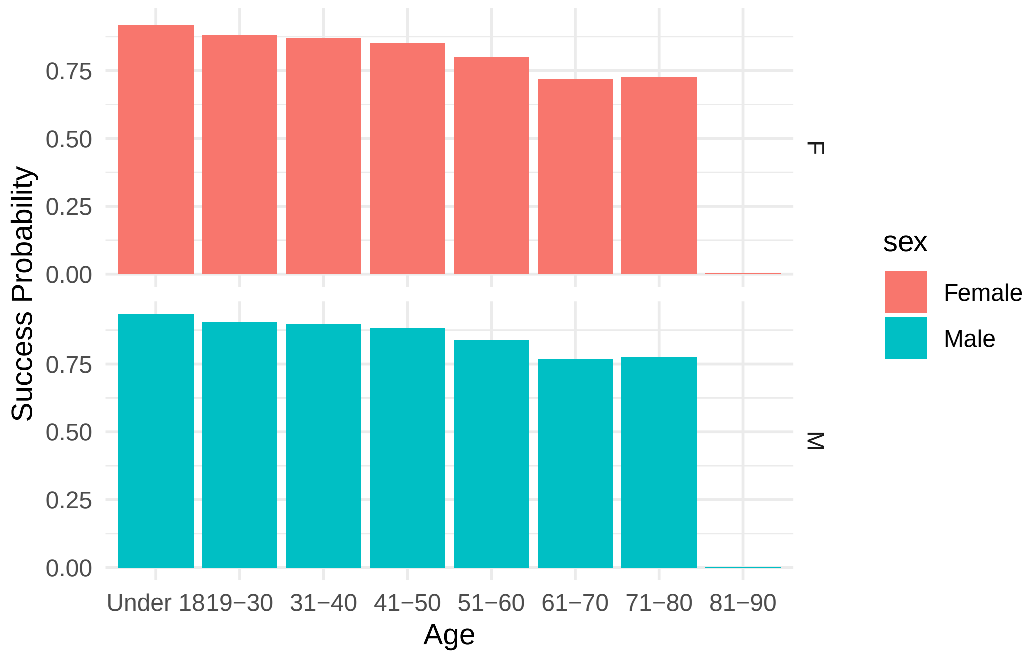


Figure 4: Predicted Success Probability by Age and Sex

Figure 4 shows the predicted success probability based on the age and sex of expedition members. It is evident that success probabilities are slightly higher for males across all ages, consistent with the findings in Figure 3. However, when age is factored in, these differences vary. For individuals under 50, the success probabilities for males and females are similar, but for those aged 51 and above, the gap widens significantly. Success probabilities are notably high, exceeding 0.8 for younger individuals but decrease with age. The sharp decline in the 81-90 age category may be attributed to the smaller number of expedition members in that age group, resulting in less reliable outcomes.

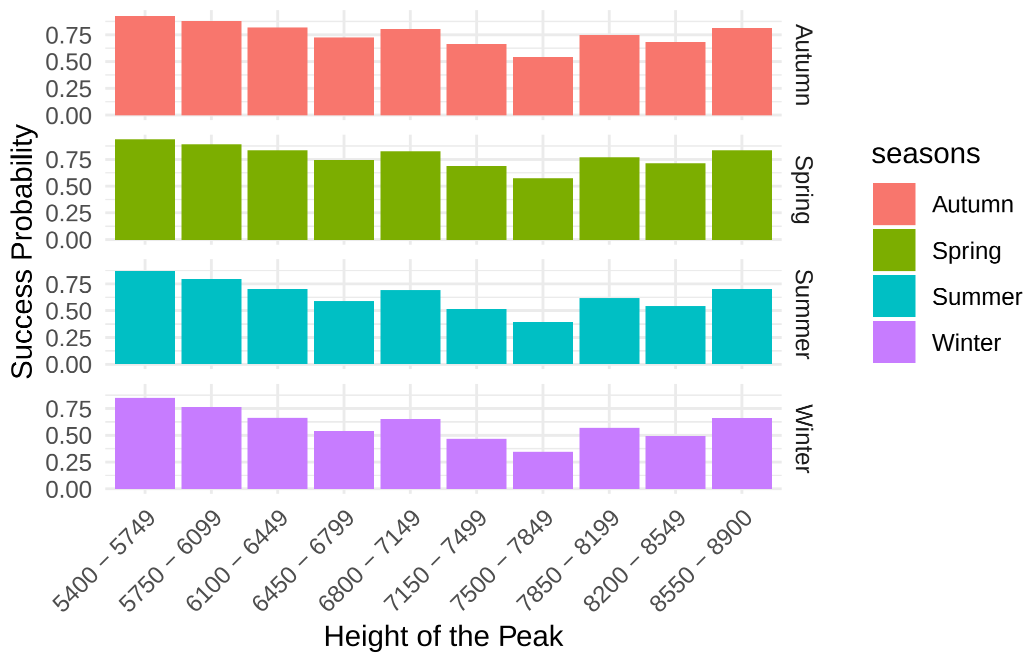


Figure 5: Predicted Success Probability by Height of the Peak and Seasons

Figure 5 shows the predicted success probability based on the height of the peak and the season of the expedition. Success probabilities are higher for shorter mountain ranges and decrease with increase in height. This might have to do with the fact that less higher peaks could be comparatively easier to summit. This, however, cannot be established as the trend as it is not consistent for peaks above 6800 meters.

Winter expeditions have lower success probabilities, probably owning to the extreme cold conditions at higher altitudes. The interaction of these two factors provide an interesting insight. It is observed that the success probabilities during different seasons for peaks between 5400 metres and 5749 metres are all comparatively closer to each other than compared to peaks between 7500 metres and 7849 metres in height. All seasons' success probabilities for peaks between 5400 metres and 5749 metres fall between 0.85 to 0.9 whereas for for peaks between 7500 metres and 7848 metres, the success probabilities go from around 0.35 in winter to around 0.6 in autumn or spring.

5 Discussion

5.1 Male dominance in mountaineering

Mountaineering, as evident by the data, presents a significant gender gap, with men outnumbering women by nearly sixfold in expedition participation. Figure 1 shows the counts of the independent variables and it is observed that there are almost six times more male expedition members. This gender gap underscores the sport's male-dominated nature.

The physical demands of mountaineering, requiring exceptional strength and endurance (Whittaker 2023), align with biological differences between sexes, as men typically have greater muscle mass and strength (ACSM 2023). This may explain why men are more drawn to mountaineering due to their physical advantages. To have the same physical strength, a woman would need to work much harder.

Despite the numerical difference, an interesting observation emerges from the analysis of success probabilities. Figure 3 and Figure 4 show how the success probabilities of men and women are not extremely different. While men have a slightly higher likelihood of reaching the summit, women show commendable success rates despite their lower representation in expeditions.

This suggests that factors beyond sheer physical strength contribute to summit success, and one of those factors can be will power. As mentioned, women often face the reality of needing to exert greater effort due to inherent physical differences, coupled with the underrepresentation within mountaineering circles. The notable success rates among women despite these challenges highlight not only their physical strength but also underscore their exceptional mental resilience and inner motivation.

These findings support the need for initiatives to promote gender diversity in mountaineering. Implementing programs and incentives tailored specifically for women can help overcome barriers, encourage more participation, and bring diverse perspectives and experiences to the mountaineering community.

5.2 Spring and Autumn are the best time for expeditions

Autumn and spring, as transitional seasons, provide favorable conditions for mountaineering expeditions. The periods offer moderate temperatures and more stable weather patterns compared to the extremes of summer and winter. Figure 3 and Figure 5 show the difference in probabilities of success by seasons and highlight higher chances of success in spring and autumn. Searle (2023) provides tips for summit success when climbing Mt. Everest and suggest that the best time to climb it is in the Autumn/ Fall season. This is in-line with our findings.

Conversely, summer and winter present significant challenges for mountaineers. Summer brings the risks of intense heat and thunderstorms, increasing the likelihood of rockfall and other hazards while winter conditions are characterized by extreme cold, high winds, heavy snowfall, and increased avalanche danger, making ascent and descent more difficult and hazardous. As a result, winter is generally considered the least desirable season for mountaineering expeditions.

In conclusion, autumn and spring are the optimal seasons for mountaineering expeditions, and season selection is an important factor in expedition planning.

5.3 Chances of success decline with increase in age

Age is a key factor in mountaineering success, with success rates typically decreasing with increase in age.

Figure 3 and Figure 4 show the impact of age on summit success. The impact of age on summit success is evident in the declining success rates with increase in age, influenced by factors such as reduced muscle mass, strength, and aerobic capacity associated with aging. Older climbers may also face challenges related to agility, flexibility, and recovery capacity, which can affect their performance on challenging terrain and recovery from strenuous exertion.

Moreover, age-related health issues like cardiovascular problems, joint stiffness, and decreased bone density can further hinder older climbers' ability to meet the physical demands of mountaineering. These factors contribute to a decline in overall performance and increase the likelihood of encountering difficulties during expeditions.

Reaburn and Dascombe (2008) highlight some of the same issues and factors when analysing endurance performance in masters athletes.

While age-related decline in mountaineering performance is noticeable, it is essential to recognize that success in mountaineering is not solely determined by physical capabilities. Experience, skill, preparation, and mindset also play an important roles in summit success. Older climbers can compensate for physiological limitations through greater experience, strategic planning, and mental resilience, enabling them to pursue their mountaineering goals despite age-related challenges.

In summary, age-related impacts on mountaineering success are significant, with success rates generally decreasing as climbers age. However, older climbers can overcome some of these challenges through experience, skill, strength training, and a strong mindset, allowing them to continue enjoying mountaineering well into their later years.

5.4 Weaknesses and Limitations

While my analysis offers valuable insights into the factors influencing a successful summit, it is important to acknowledge the weaknesses and limitations.

Firstly, some success probabilities should be viewed with skepticism due to potential inaccuracies. Discrepancies in the counts of different categories may lead to imprecise success probabilities. For instance, the number of individuals attempting a solo ascent was significantly lower than those in group ascents. However, my model coefficients indicate for solo ascents was higher. This could create a misconception that solo ascents have better chances of success, potentially influencing more people to choose this option without considering other contributing factors. These probabilities may be influenced by the limited sample size of solo ascents, which lacks diverse data to account for confounding variables. There might be other factors which play a role like will power. Individuals opting for solo expeditions likely possess high levels of determination which can be a confounding factor playing into their success.

Additionally, while I focused on focused on a few factors, there are numerous other variables that could impact the likelihood of a successful summit. Nationality and group size were two factors which were available in the raw data and I believed were factors which could have played a significant difference. These were later removed due to the lack of consistency in those variable's. For instance, the way the number of members were reported differed from record to record. As the dataset also included the people summiting the other side of border peaks, some countries reported this number differently than others. For expeditions in Nepal, this was usually the number of foreigners listed on the expedition permit whereas in China, this was usually the number of non-hired members.

5.5 Future Directions

Moving forward, additional research is necessary to address the limitations discussed and improve the dataset to enhance the understanding of factors contributing to a successful summit. Including more diverse information can help uncover additional factors influencing a successful summit. Some of these factors that can be explored would be nationality of the person, attempt number, a metric for that could measure their level of preparedness and group size.

Nationality can play a role as the cultural background and national resources can play a significant role in shaping an individual's climbing experience. Different countries may have different levels of access to training facilities, equipment, and support networks for climbers. Additionally, genetic differences between individuals may also play a role.

Analysing the number of attempts made by climbers before a successful summit can reveal patterns related to perseverance, learning from past experiences, and adapting strategies. It can help determine whether repeated attempts increase success rates or reach a point of diminishing returns.

Developing a metric to measure climbers' preparedness can quantify the various aspects of readiness, including physical fitness, technical skills, mental resilience, and logistical planning. Understanding the relationship between preparedness levels and summit success can guide training and preparation efforts.

Group dynamics play an important role in mountaineering, affecting decision-making, safety, resource allocation, and overall team morale. Studying how group size influences summit success can provide insights into optimal team composition, leadership structures, and communication strategies. It can also highlight the trade-offs between the benefits of larger teams and the challenges of coordinating and managing them effectively.

6 Appendix

6.1 Cleaning

For the analysis data, the cleaning steps I took were:

1. Initial merging: The raw data from the `expeditions` and `members` datasets is merged based on a common identifier, `expedition_id`, consolidating information about expedition participants.
2. Column selection and renaming: Irrelevant columns are removed from the merged dataset, and the remaining columns (`peak_id.x`, `season.x`, `sex`, `age`, `success`, `solo`, `died`) are selected for further analysis. Additionally, column `peak_id.x` is renamed to `peak_id`.
3. Secondary merging: The cleaned `expeditions` dataset is merged with the `peaks` dataset based on a common identifier, `peak_id`, to incorporate information about the height of each peak climbed during expeditions.
4. Filtering out incomplete data: Rows with missing values for key variables such as `sex` or `age` are filtered out to ensure the integrity of the dataset.
5. New ranges: Height range categories are created based on predefined ranges in meters: 5400-5749, 5750-6099, 6100-6449, 6450-6799, 6800-7149, 7150-7499, 7500-7849, 7850-8199, 8200-8549, and 8550-8900. Age range categories are defined as follows: Under 18, 19-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, and 91 or older.
6. Final dataset creation: `season.x` is renamed to `seasons`. and then the resulting dataset is further refined to include only relevant columns (`height_range`, `seasons`, `sex`, `age_range`, `success`, `solo`, `died`).

6.2 Analysis dataset

Table 1: Analysis dataset

height_range	seasons	sex	age_range	success	solo	died
5750 - 6099	Autumn	M	19-30	TRUE	FALSE	FALSE
5750 - 6099	Autumn	M	19-30	TRUE	FALSE	FALSE
5750 - 6099	Autumn	M	19-30	TRUE	FALSE	FALSE
5750 - 6099	Autumn	M	19-30	TRUE	FALSE	FALSE
5750 - 6099	Autumn	M	19-30	TRUE	FALSE	FALSE
5750 - 6099	Autumn	M	51-60	TRUE	FALSE	FALSE

Table 1 provides a glimpse of the dataset used for analysis.

6.3 Model summary

Table 2: Summary of the model

term	estimate	std.error	conf.low	conf.high
(Intercept)	0.76	0.25	0.33	1.18
height_range5750 - 6099	-0.58	0.26	-1.01	-0.15
height_range6100 - 6449	-1.07	0.25	-1.49	-0.64
height_range6450 - 6799	-1.61	0.25	-2.03	-1.18
height_range6800 - 7149	-1.14	0.25	-1.56	-0.72
height_range7150 - 7499	-1.89	0.25	-2.31	-1.46
height_range7500 - 7849	-2.40	0.25	-2.83	-1.97
height_range7850 - 8199	-1.48	0.25	-1.90	-1.06
height_range8200 - 8549	-1.79	0.25	-2.22	-1.37
height_range8550 - 8900	-1.09	0.25	-1.51	-0.66
sexM	0.27	0.03	0.22	0.31
soloTRUE	1.16	0.19	0.84	1.49
age_range31-40	-0.10	0.02	-0.13	-0.06
age_range41-50	-0.26	0.02	-0.30	-0.23
age_range51-60	-0.62	0.03	-0.67	-0.57
age_range61-70	-1.08	0.06	-1.18	-0.97
age_range71-80	-1.03	0.18	-1.34	-0.74
age_range81-90	-209.28	182.58	-597.95	-21.95
age_rangeUnder 18	0.39	0.13	0.19	0.60
seasonsSpring	0.12	0.02	0.09	0.16
seasonsSummer	-0.62	0.09	-0.78	-0.46

term	estimate	std.error	conf.low	conf.high
seasonsWinter	-0.82	0.06	-0.91	-0.73

Table 2 shows the coefficients for my Bayesian model along with the standard error and the 95% credible interval. The standard error (SE) is a measure of the precision with which a sample statistic estimates a population parameter. It quantifies the variability of sample statistics around the population parameter.

6.4 Posterior predictive check

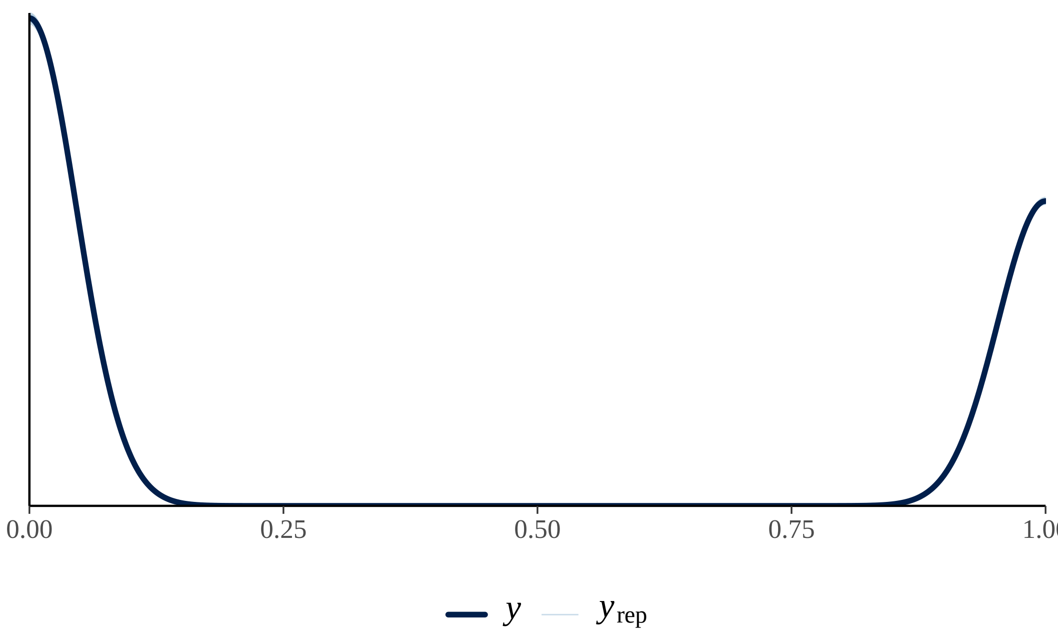


Figure 6: Posterior distribution for logistic regression model

Figure 6 and Figure 7 are used to understand assess the quality of my model.

Figure 6 is the result of a posterior predictive check and is used to compare the actual outcome variable with simulations from the posterior distribution. Figure 7 is the result of comparing the posterior with the prior to see how much the estimates change once data are taken into account. Both these results support the model doing a good job of fitting the data.

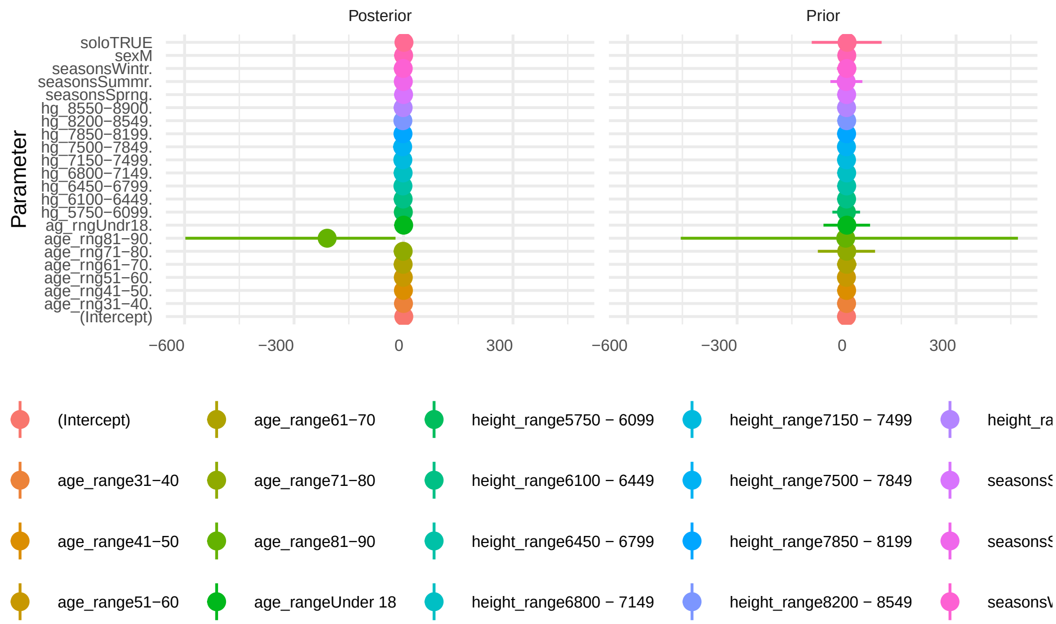
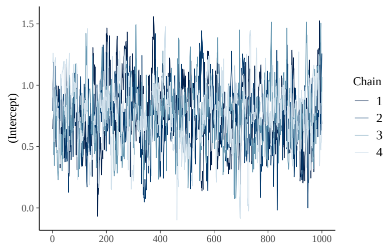


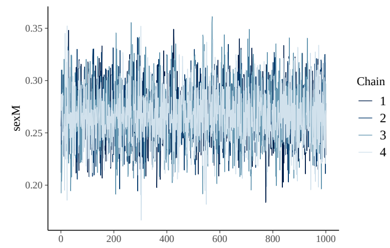
Figure 7: Comparing prior distribution with posterior distribution

6.5 Markov chain Monte Carlo (MCMC) Convergence Check

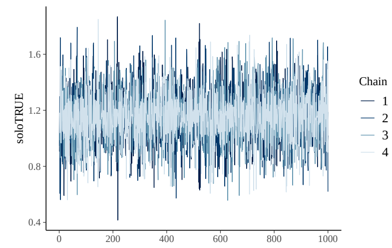
My model, modelled using `rstanarm` (Goodrich et al. 2024) uses a sampling algorithm called Markov chain Monte Carlo (MCMC) to obtain samples from the posterior distributions of interest. Figure 8, Figure 9 and Figure 10 are used to check for the existence of signs that the algorithm ran into issues. Figure 8 and Figure 9 does not suggest anything out of the ordinary and everything in Figure 10 is close to 1 suggesting no problem as well.



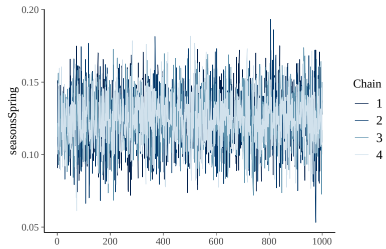
(a) Intercept



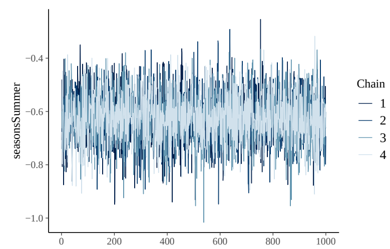
(b) sexM



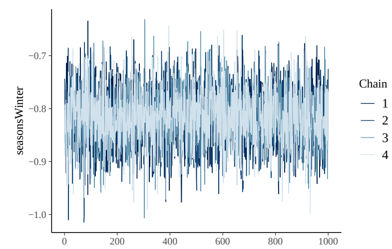
(c) soloTRUE



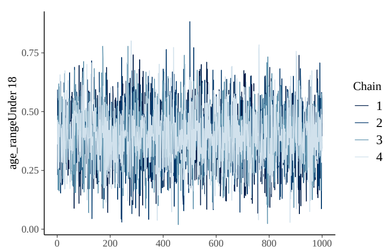
(d) seasonsSpring



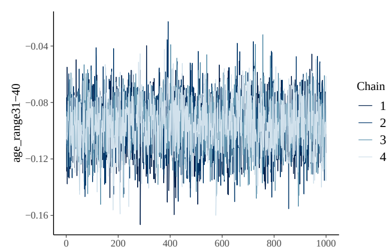
(e) seasonsSummer



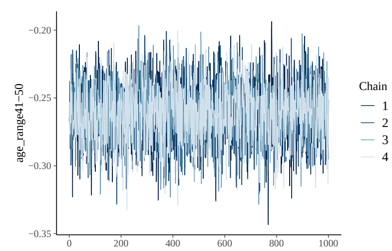
(f) seasonsWinter



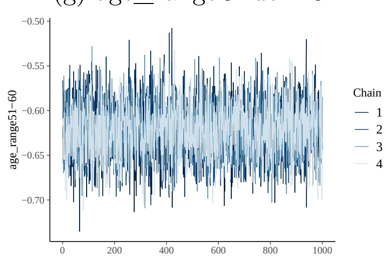
(g) age_rangeUnder 18



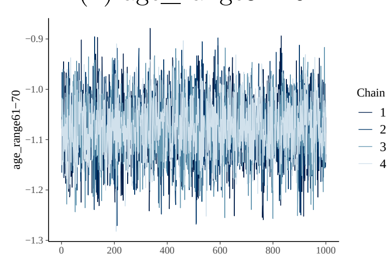
(h) age_range31-40



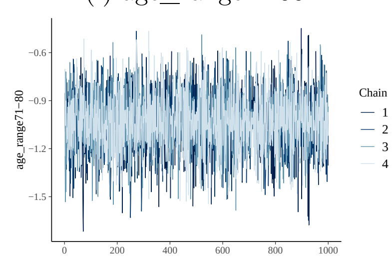
(i) age_range41-50



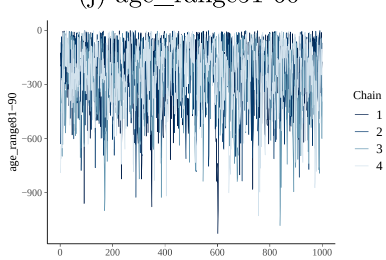
(j) age_range51-60



(k) age_range61-70

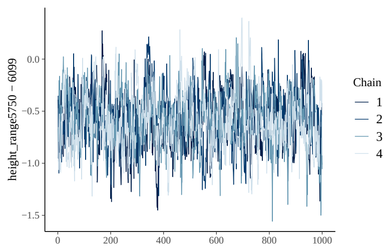


(l) age_range71-80

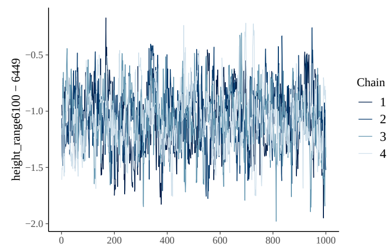


(m) age_range81-90

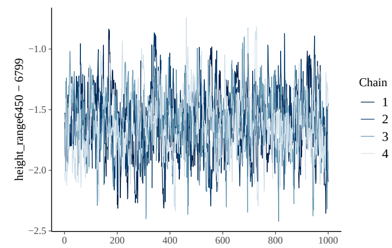
Figure 8: Trace plot



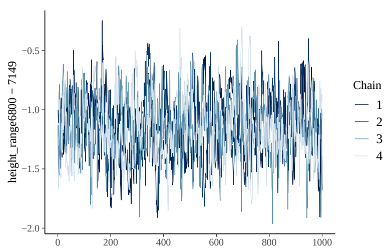
(a) height_range5750 - 6099



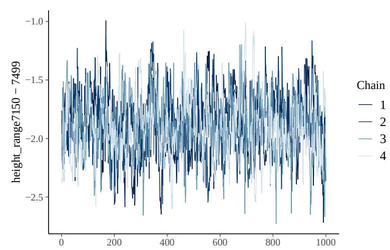
(b) height_range6100 - 6449



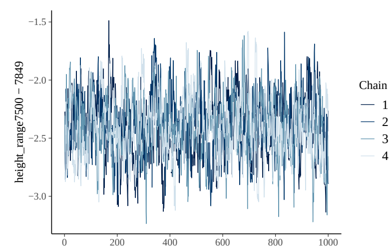
(c) height_range6450 - 6799



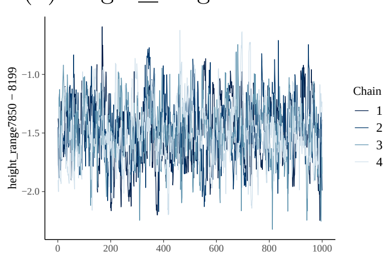
(d) height_range6800 - 7149



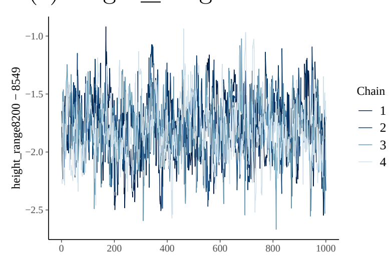
(e) height_range7150 - 7499



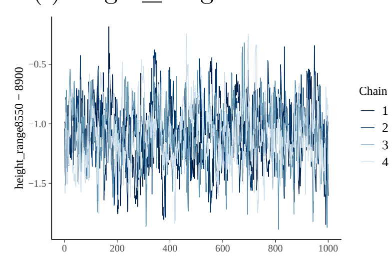
(f) height_range7500 - 7849



(g) height_range7850 - 8199



(h) height_range8200 - 8549



(i) height_range8550 - 8900

Figure 9: Trace plot

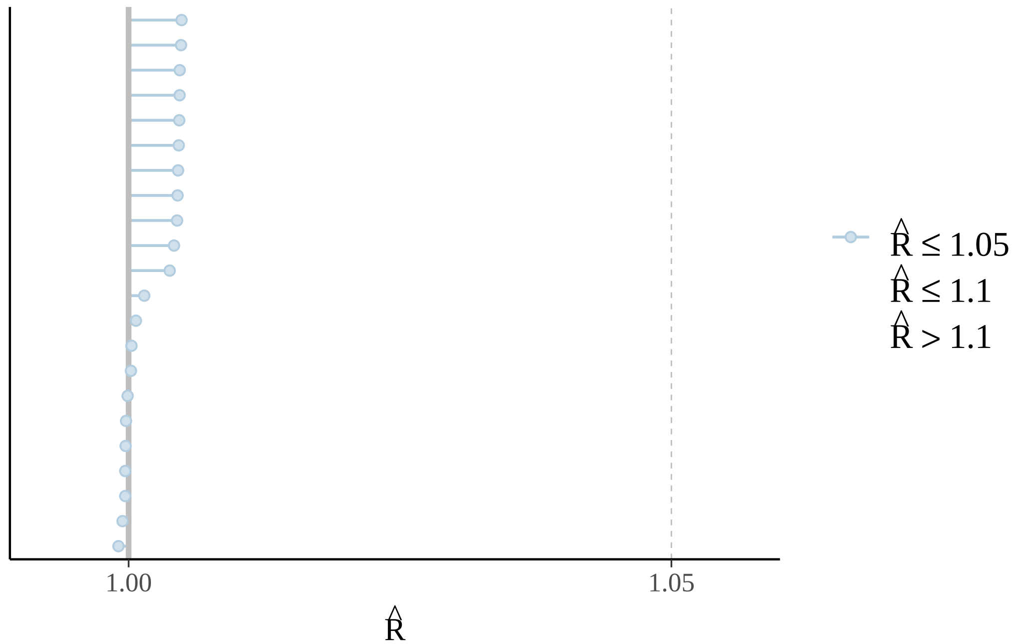


Figure 10: Rhat plot

References

ACSM. 2023. "ACSM Releases Expert Consensus Statement: The Biological Basis of Sex Differences in Athletic Performance." *ACSM_CMS*. https://www.acsm.org/news-detail/2023/09/29/acsm-releases-expert-consensus-statement-the-biological-basis-of-sex-differences-in-athletic-performance.

Bolker, Ben, and David Robinson. 2022. *Broom.mixed: Tidying Methods for Mixed Models*. https://CRAN.R-project.org/package=broom.mixed.

Cookson, Alex. 2020. "Data/Himalayan-Expeditions at Master · Tacookson/Data." *GitHub*. https://github.com/tacookson/data/tree/master/himalayan-expeditions.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reaburn, Peter, and Ben Dascombe. 2008. "Endurance Performance in Masters Athletes - European Review of Aging and Physical Activity." *BioMed Central*. Springer-Verlag. https://eurapa.biomedcentral.com/articles/10.1007/s11556-008-0029-2.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/packag e=arrow.

Salisbury, Richard. 2023. “The Himalayan Database, the Expedition Archives of Elizabeth Hawley.” *The Himalayan Database, The Expedition Archives of Elizabeth Hawley*. https://www.himalayandatabase.com/.

Searle, Hugo. 2023. “Climbing Mount Everest: Key Considerations for Summit Success.” *Mountain IQ*. https://www.mountainiq.com/climbing-mt-everest/.

Whittaker, Leif. 2023. “Endurance Training for Mountaineering.” *The Mountaineers*. https://www.mountaineers.org/blog/mountaineering-training-basics.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with ’Kable’ and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.