Otzi's Adventures

Introduction

5000 years ago, tragedy struck in the Tyrol when a man now known to the world as Otzi the Iceman was killed 3000 meters up in the mountains. There he remained for centuries, his body naturally mummified, until its discovery in 1991. Otzi was not the first - and would not be the last - to perish so high above the Earth, as humankind remains fascinated by mountains and surmounting their many dangers. Perhaps none capture the imagination so much as the "8000ers", the 14 mountains on Earth stretching above 8000 meters. It at this point (the "death zone", as it's called) that oxygen levels in the air become too low to sustain life, and many climbers elect to use supplemental oxygen to assist their efforts. They are the ultimate test in high-altitude mountaineering, and climbing any one of them is a serious undertaking with considerable risks. Yet hundreds attempt them every year, and there is a growing industry of companies offering services to assist less experienced climbers in reaching the summits. Now, one of Otzi's distant descendants, intrepid business man and mountaineering enthusiast Bo L. Zano, wants to begin his own Himalayan expedition company, *Otzi's Adventures*, to help others conquer those masses of rock and ice that looked with such indifference upon his ancestor so many years ago. Having already made his initial fortune in the portable oxygen production game, Bo hopes to form a symbiotic relationship between his two ventures, whereby he sells oxygen to his climbing clients and guides them up and down the mountain safely to climb (and buy) again. Bo has tasked us with analyzing years of data on attempted climbs on the eight 8000ers in the Nepal Himalaya to assess how we can give our clients the best chance of summiting, reduce the risks inherent in climbing, and to explain the many benefits of using supplemental oxygen.
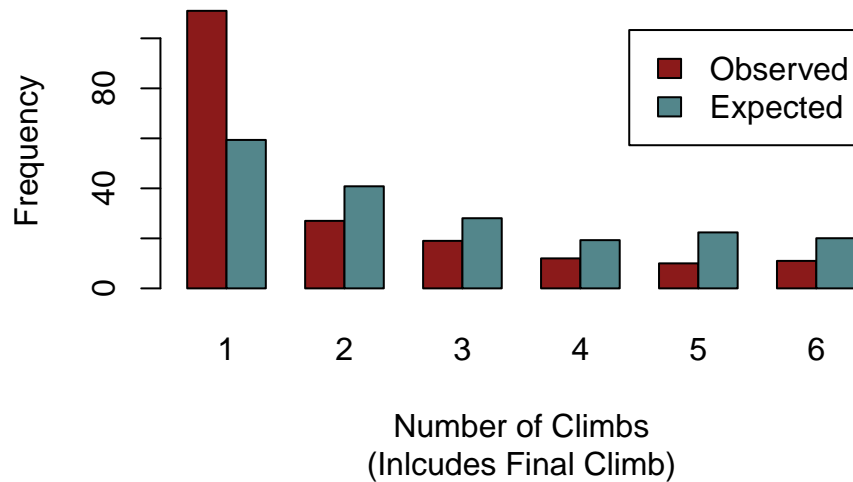
Fitting Distributions

Climbs Before Death - Geometric

Is there a certain number of climbs one can attempt before an unfortunate end is expected? Given our knowledge of the geometric distribution, it makes sense to attempt to model the number of expeditions a climber goes on before dying geometrically. We have discrete, positive data, and want to count the number of trials (climbs) until a condition is met (death). Note that it is possible (preferable, even) that a climber does not die while attempting a peak. In this setup, we merely assign this outcome a probability of 0.

We aggregate the data for each climb into a tally of number of climbs undertaken by each unique climber, and filter to look at only those who ultimately died. Using R's *fitdist* function , we estimate the parameter for a geometric distribution based on the data to be $p = 0.31$, which implies a climber who dies can expect to attempt about three climbs before dying. However, when we compare our observed data with values from the theoretical distribution by binning and running a Chi-Square test, the results ultimately indicate that this data is not consistent with a geometric distribution. The discrepancy between the observed and expected values is clear to see in the joint barplot, where the expected value for one climb is much lower than observed, while for two or more climbs, the expected value is consistently higher.

## Number of Climbs Before Death

Summits Before Death - Geometric

With no luck fitting a distribution on the number of climbs one person attempts before dying, we perform the same test on the number of successful summits one climber achieves before dying. Again, a geometric distribution makes the most sense in this case, for the same reasons as above.

Using R's *fitdist* function again, we estimate the parameter based on this data to be $p = 0.54$, for an expected value of almost two successful summit attempts before death. We again bin our observed and expected values and run a Chi-Square test, which this time indicates the data is, in fact consistent with a geometric distribution. Looking at the joint barplot of observed and expected, we do see quite close agreement throughout the distribution. So, climbers who ultimately die can expect to summit about two mountains, with a variance of about 1.6.

## Number of Succesful Summits Before Death
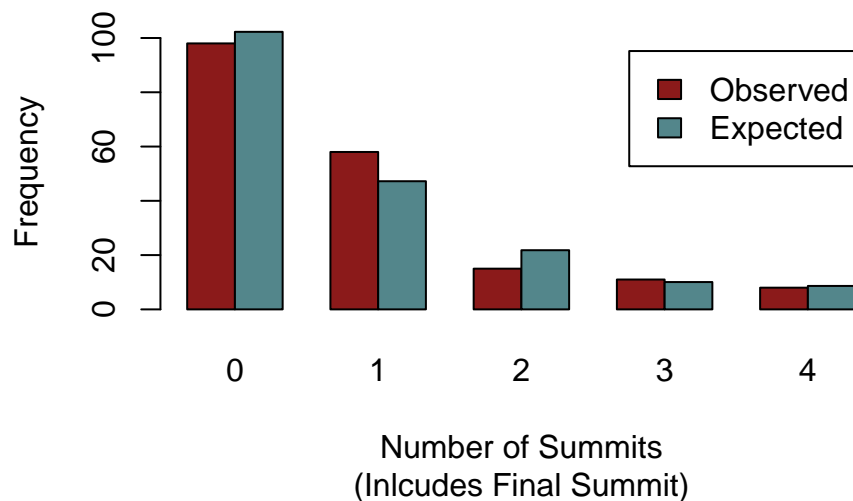
Summit Times - Gamma

Reaching the summit and returning safely are the ultimate goals of any mountaineering expedition, and the time the summit is reached can have substantial implications for the safety of the descent. Most climbers seek to time their summit push to arrive by late morning, in order to avoid descending in darkness or during the common late afternoon storms. We will explore this relationship between summit times and safety in greater depth later, but for now we attempt to fit a distribution to the summiting times for climbers on each peak individually and across all eight combined. In theory, modeling this data with a gamma distribution makes the most sense. We have all non-negative values, and we expect the data to be skewed right with a peak that rises somewhat rapidly (as the majority of climbers summit in the mid-to-late morning) before falling off exponentially (as summiting later in the day is more dangerous and therefore less common). Indeed, looking at a histogram of summit time for each climb across all eight peaks, we observe this general shape.

Histogram of Summit Times for All Peaks
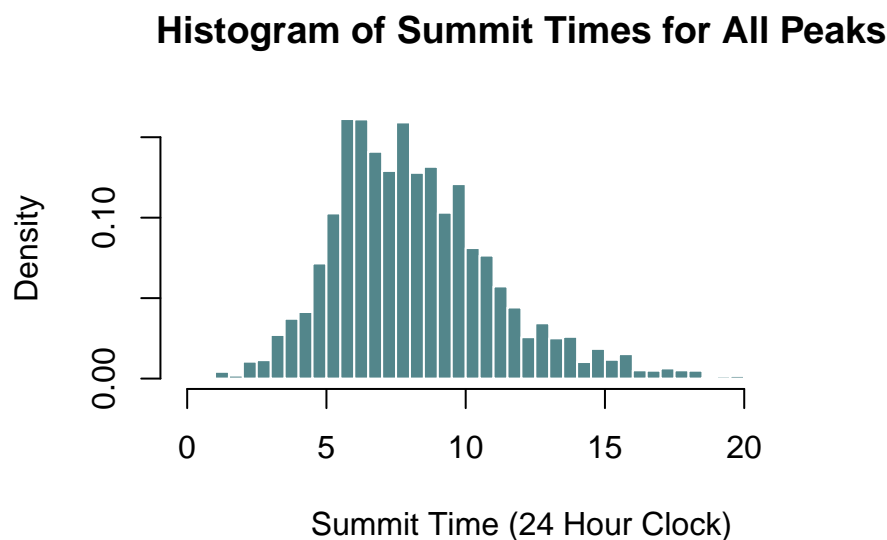
For each peak, we estimate the shape and rate parameters using *fitdist*, and check the fit of the observed data to the expected values given our estimated parameters using a comparison by deciles approach. The results show that in general, summit times cannot be reliably modeled by a gamma distribution. Only one peak, Lhotse, returned a p-value large enough to suggest that the data may come from a gamma distribution. Given the estimated parameters for Lhotse (shape 8.1, rate 0.82), our expected value for summit time is just a little before 10am, with a 12 hour variance.

Histogram of Summit Times for Lhotse
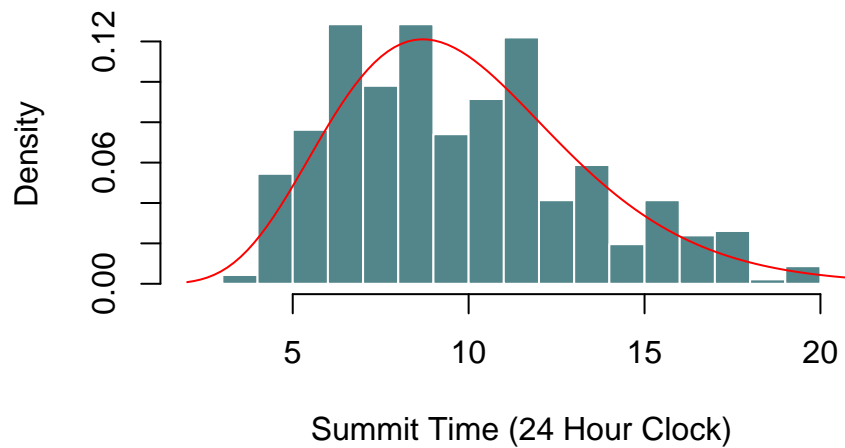
Proportion of Male Expedition Members - Beta

Beta distributions can be applied to modeling continuous random variables whose values range from 0 to 1. Although the initial climber dataset does not contain any variables for which the beta distribution would be appropriate, we create a number of expedition-level variables representing proportions. One of these variables is the proportion of climbers on an expedition who are male. The beta distribution that we estimate from the data generally matches the distribution we observe. However, a chi-sq goodness of fit test shows that the estimated beta distribution does not accurately model the proportion of males climbers in an expedition. In the Q-Q plot, we see that the failure of the beta distribution is likely due to the large portion of our observations (~36%) at the upper bound of 1.

Histogram

Proportion of Expedition Members who are Male

Q–Q Plot
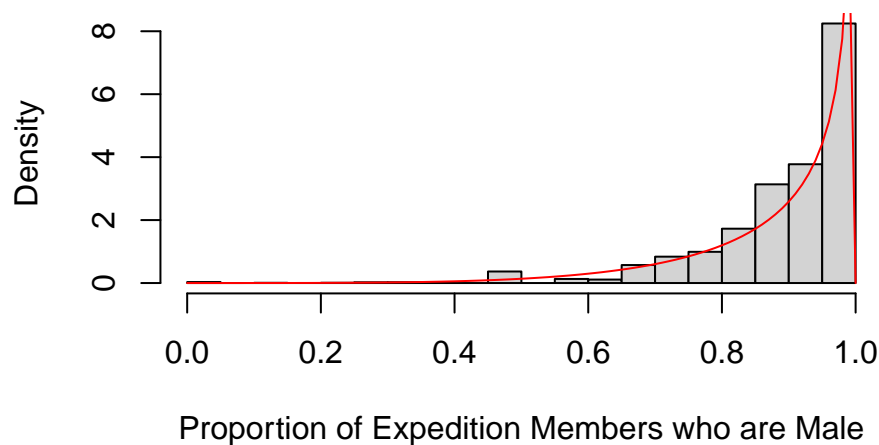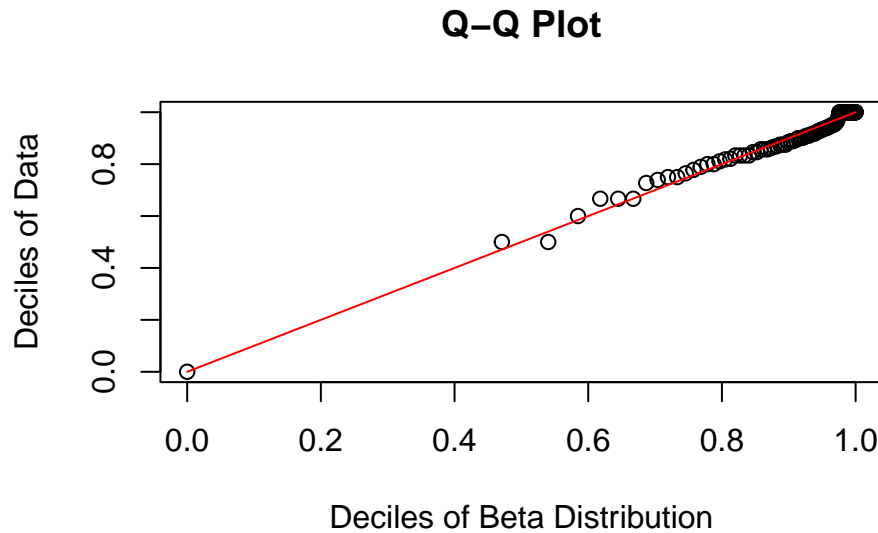
## Miscellaneous Analysis

Survival Plots

Survival plots, common in epidemiology, display the probability of survival past a point in time. Interestingly, The survival function itself is equal to (1 - CDF); so, the survival plots are complementary to our well-known CDF plots. Rather than using time, we adapt the survival plot to show the probability that a climber reaches at least a certain height on Mt. Everest. For climbers who have recorded personal high point, approximately 75% reach the peak of Mt. Everest at 8849m. The probability of reaching point decreases rather steadily from base camp to the summit of Everest, indicating that climbers don't face a universal obstacle on the climb up. We see that male climbers are more likely to reach higher evaluations than female climbers, with the average male climber reaching ~8180m compared to the average female climber at ~7680m. Unsurprising for a physical feat like mountaineering, we find that younger climbers have greater success reaching higher elevations and the peak. Our cohort of climbers who are under 18 summit over 80% of the time; this is assuredly a consequence of selection, as the few children who climb Everest must be heavily vetted to ensure their fitness for the task. On the other hand, climbers over 65 only reach the summit about 25% of the time. We observe climbers designated as "expedition leaders" reach heights above 7000m at lower rates than those without the designation. This is an unexpected finding, as one might think that leaders would be most capable of climbing. We suspect that there is a practical explanation for this, such as leaders going on multiple expeditions and not needing to summit every time or needing to stay behind with climbers who decide not to continue climbing. We find that probability of reaching a given height is similar across the years of our data, with one major exception. In 2015, no climber reaches the peak of Everest and almost no one goes above 7000m. With a bit of investigation, we learn that there was a major earthquake that set off a series of avalanches that resulted in the premature end of the climbing season. Across all years, the popular spring climbing season has the best outcomes, while those who climb in the winter have 0% probability of reaching the summit.
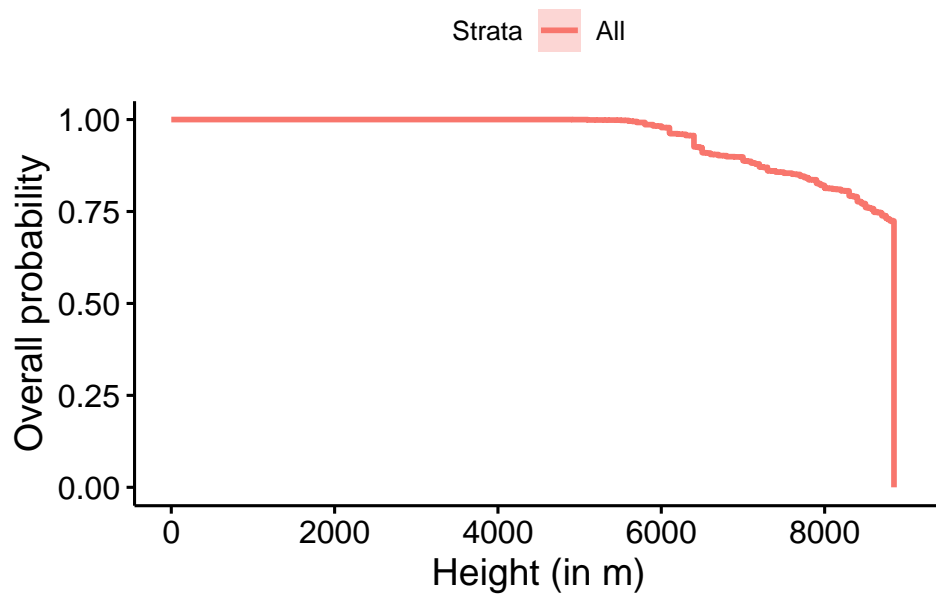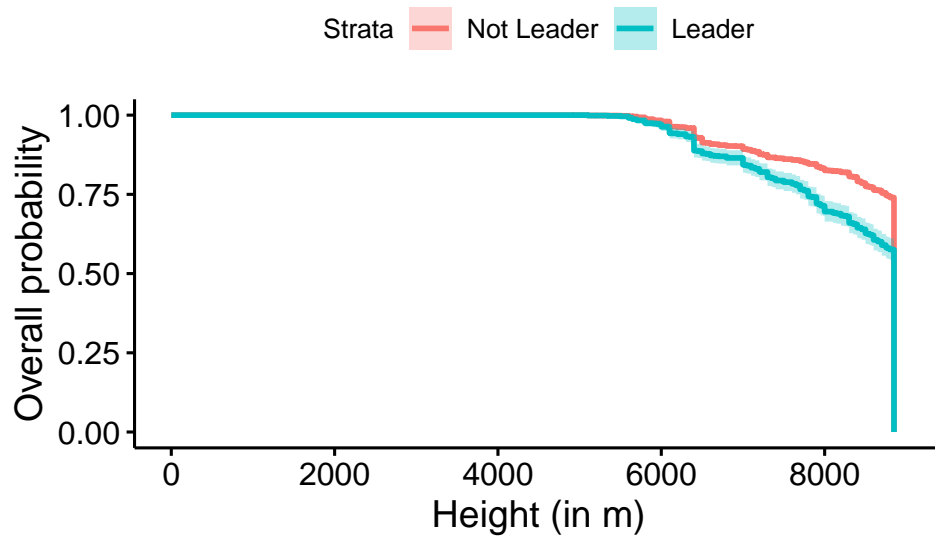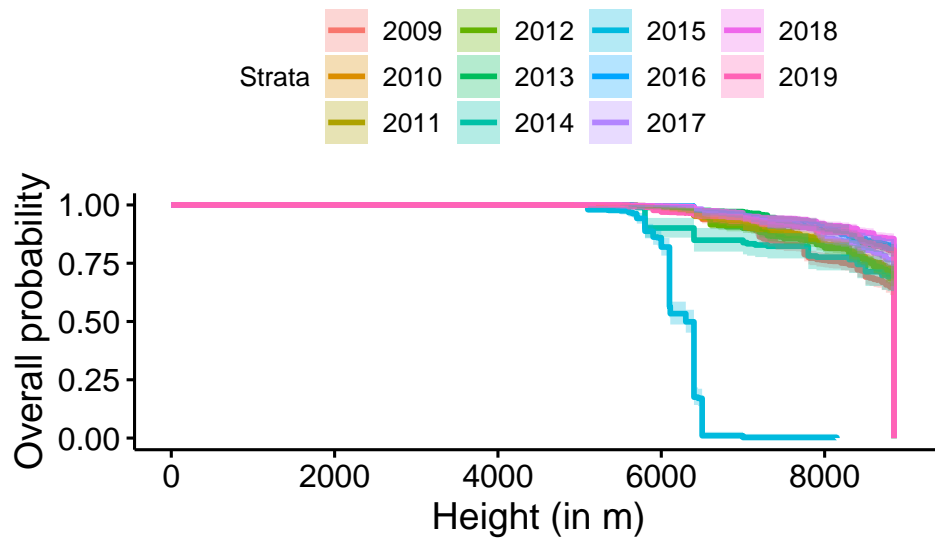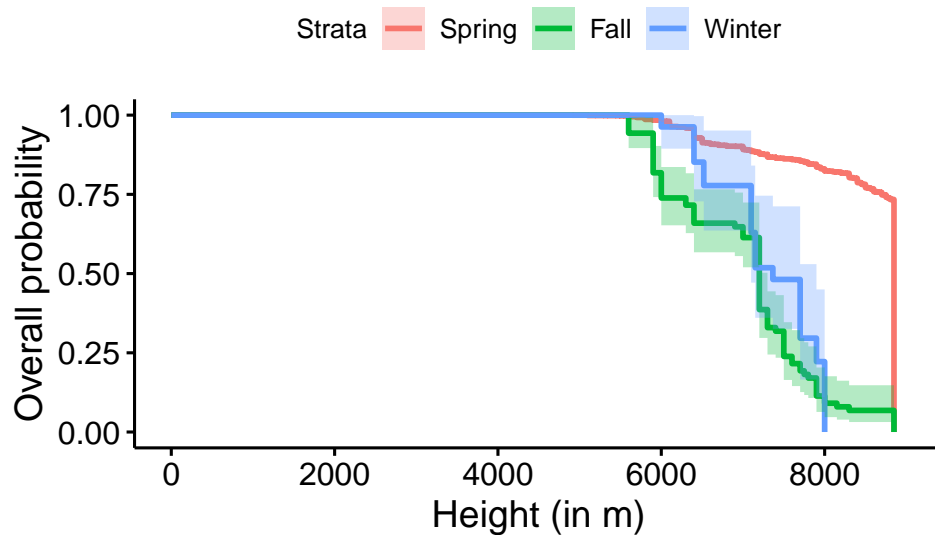
Climbing Everest
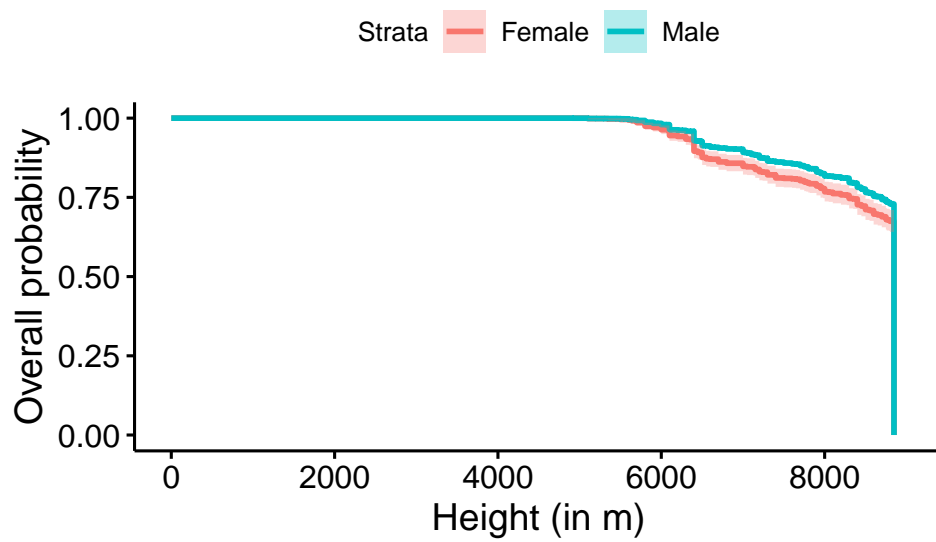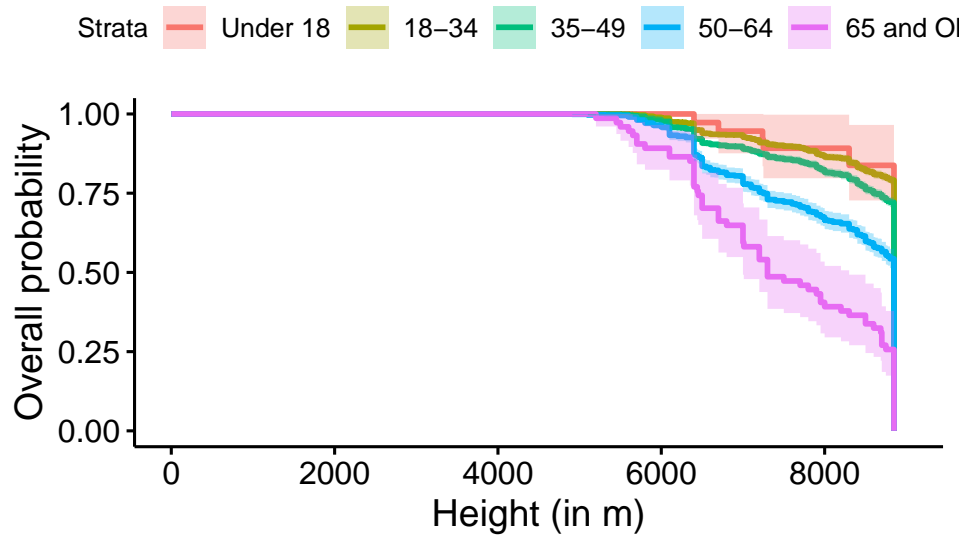
Climbing Everest

Climbing Everest

Climbing Everest

Climbing Everest

Summit Time and Mortality

Do people that die on their climbs summit later than those who survive? We expect the answer to be yes. In mountaineering, teams will usually decide on a "turnaround time" for when they make their final push to the summit from their highest camp. As the name suggests, this is the time at which climbers should turn around, regardless of how close to the top they are. These times can vary by mountain, season, and expected weather conditions, but they are generally set in the early afternoon so that climbers will have enough time to get back down to their high camp before nightfall or afternoon storms set in. Ignoring these times is considered quite dangerous, and there have been some famously disastrous examples of why that is throughout history. To avoid this danger, most climbers try to reach the summit by mid-to-late morning.

Using data for distinct climbs, we can graphically see the disparity between the summit times of those who live and those who die on a particular climb with a simple boxplot, which gives us an initial indication that there is likely a significant difference in summiting time between these groups. The median climber who survives summits around 8am, while those who die are closer to noon or 1pm. The interquartile ranges only overlap slightly, and only beginning of the outliers amongst the survivors summit around the same time as the 75th percentile for those who die.

Summit Times By Mortality
(All Peaks)
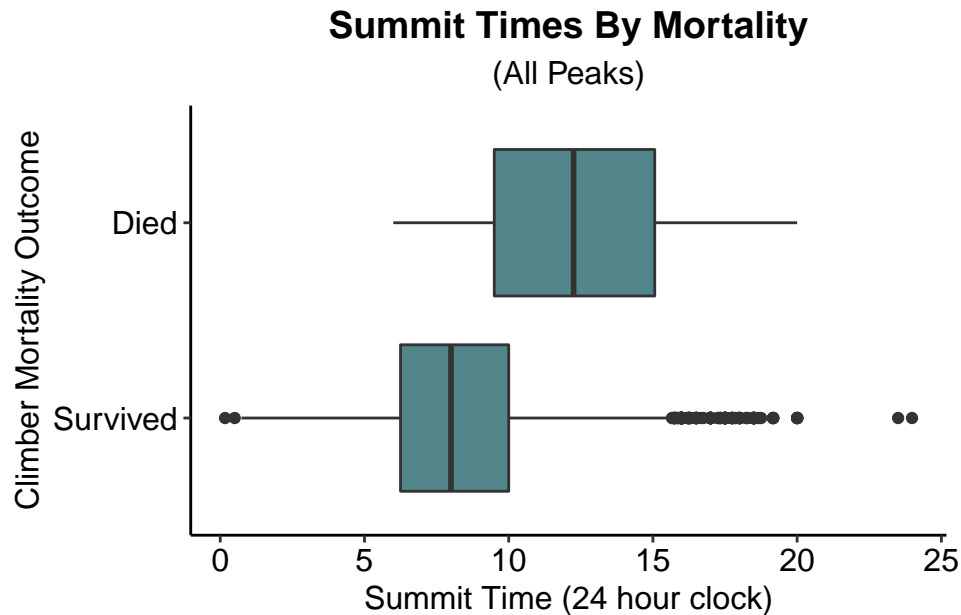
Next, we conduct a permutation to test whether those who die summit significantly later on average than those who survive. The mean summiting times amongst climbers who die on that particular climb is around 12:30pm, while the average for survivors is around 8:30am. This four-hour difference seems quite large, and when we run the permutation test, we can see graphically how far out on the histogram of differences in mean the observed value is. Our p-value indicates there is only a 0.001 probability that we would observe such a large difference due to chance, so we conclude that those who die summit significantly later on average than those who survive. This further highlights the importance of early summits and turnaround times for climber safety.

Difference of Means
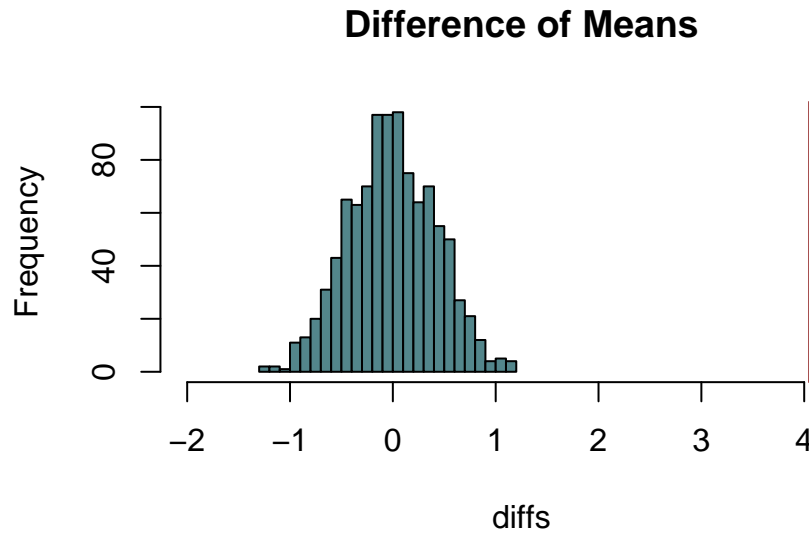
Alpinist vs Businesspeople Summiting Rates

Expeditions up these "8000ers" are no small tasks, and climbers might have any number of motivations. There are professional climbers or avid mountaineers, looking to add another climb; there are scientists, hoping to conduct research along the way; there are novices who have the resources to pursue a Himalayan adventure for the bragging rights or to cross an item off their "bucket-list." We believe that the summit success rates likely varies across different types of climbers. To investigate this, we compare climbers whose occupation is listed as a variant of "alpinist," who we view as likely to be a professional climber, to climbers whose occupation is listed as a variant of "business," who we view as likely to be novices. (Unfortunately we do not have a large enough sample for "professors" or "mathematicians.") We find that business people have a summit success rate of 41.7%, while alpinists have a summit success rate of 47.4%. We use a permutation test to see if 5.7 percentage point lower success rate is statistically significant. At a 5% significance level, we reject the null hypothesis and conclude that business people and alpinist have different summit success rates. We confirm this result by running a chi-sq test on the contingency table of occupation and summit success. Our chi-sq test statistic and p-value match those of our permutation test and again lead us to reject the null hypothesis that business people and alpinists have the same summit success rates.

| | Summit Success | |
	FALSE	TRUE
Alpinists	853	770
Business people	243	174

Histogram of Simulated Rate Differences

Differnce in Summit Rate for Business People vs Alpinists
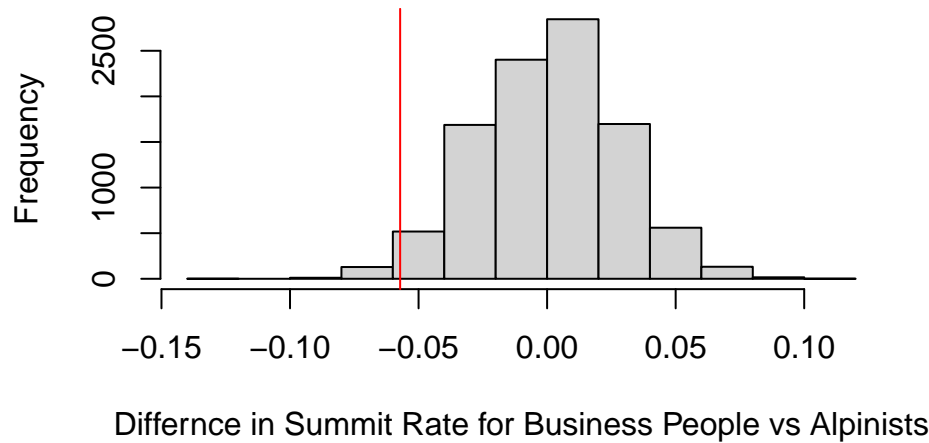
Oxygen Analysis

Altitudes above 8000m are referred to as the "death zone" in mountaineering, because the concentration of oxygen in the air is not high enough to sustain life for extended periods of time. This severe lack of oxygen leads to impaired decision making, deteriorating bodily functions, and puts climbers at greater risk of deadly pulmonary or cerebral edema. And as the name implies, too long in the death zone ultimately has only one outcome. Some climbers choose to brave these dangers head on, relying on incredible fitness, speed, and climbing experience to get them up to the summit and down out of the death zone quickly. However, supplemental oxygen has been used for decades by high altitude mountaineers to combat the effects of high altitude, extend the amount of time they can spend in the death zone, and improve their chances of summiting.

Oxygen Use Impact on Number of Succesful Summits (Linear Regression)

First, we perform a simple regression, using data for individual climbers, aggregating over all of their attempted climbs. We regress the total number of summits a climber achieves on the total number of climbs they attempt, while conditioning on their oxygen use ratio (number of climbs using oxygen over total number of climbs). While the number of climbs is positively correlated with the number of summits achieved in general, the slope of the regression line varies greatly depending on oxygen use.

The model predicts that those who always use oxygen (oxygen ratio of 1) can expect to summit about nine mountains for every 10 climbs they attempt, while those who never use oxygen (oxygen ratio of 0) can only expect to summit between three and four mountains for every 10 climbs. For climbers who use oxygen some of the time (oxygen ratio between 0 and 1) we predict about seven summits per 10 climbs, though there is certainly variation depending on where in the ratio range a particular climber falls. A graph of the regression lines for each group nicely displays the disparities in summit success by oxygen use.

Table 1: Summit Success by Climbs Attempted, Conditional on Oxygen Use

| | Number of Summits | | |
	(O2 Always Used)	(O2 Never Used)	(O2 Sometimes Used)
Number of Climbs	0.907***	0.356***	0.714***
	(0.007)	(0.009)	(0.009)
Constant	−0.093***	−0.274***	−0.662***
	(0.013)	(0.013)	(0.058)
Observations	3,481	3,624	1,744
R^2	0.819	0.317	0.791
Adjusted R^2	0.819	0.316	0.791
Residual Std. Error	0.450 (df = 3479)	0.447 (df = 3622)	1.459 (df = 1742)

Note: *p<0.1; **p<0.05; ***p<0.01

Summit Success by Oxygen Use Over Multiple Climbs
(All Peaks)
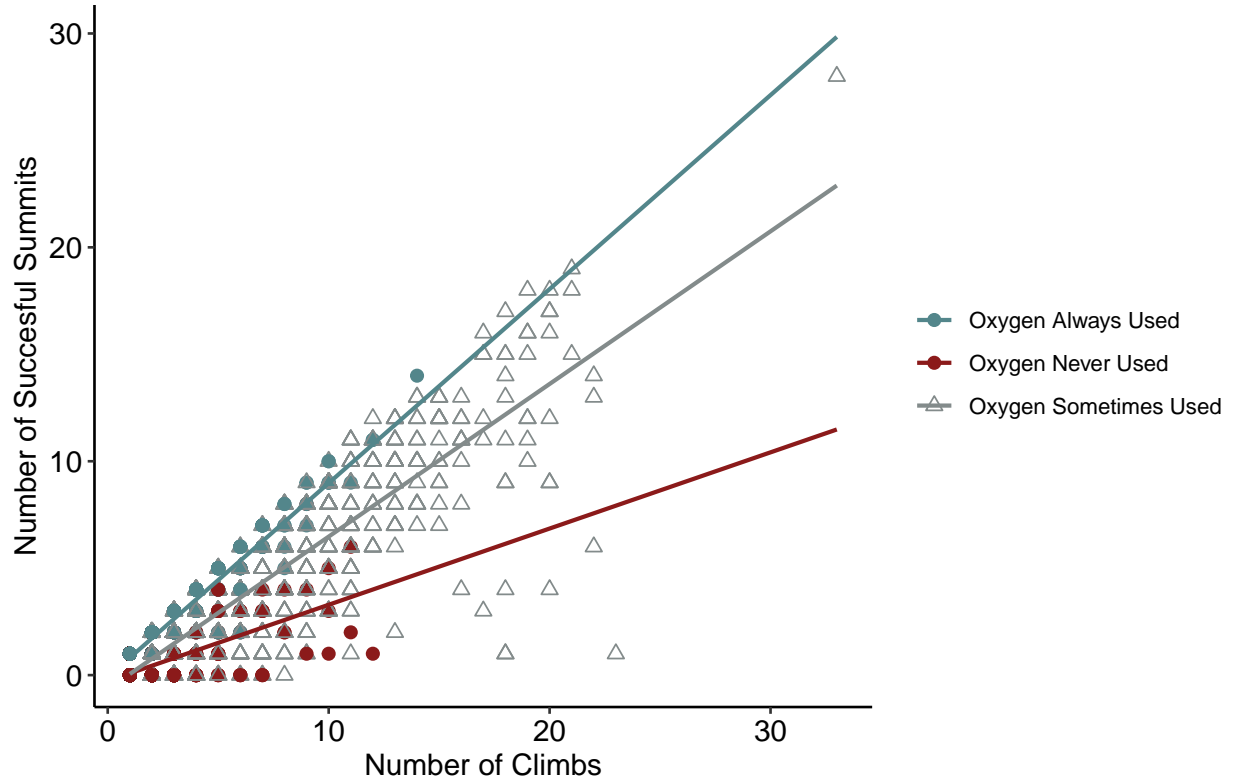
However, to truly capture the effect that a climber's oxygen use ratio has on their number of summits reached, we should control for other factors in a larger model. We estimate the following linear regression:

$$nsuccess_i = \alpha + \beta_1 o2ratio_i + \beta_2 nclimb_i + \beta_3 died_i + \beta_4 calcage_i + \beta_5 sex_i$$

Where *nsuccess* is the number of successful summits, *o2ratio* is the oxygen use ratio, *nclimb* is the total

number of climbs someone attempted, *died* is a dummy variable for whether or not they died (1-yes, 0-no), *calcage* is the age of the climber, and *sex* is a dummy variable for the sex of the climber (1-male, 0-female).

Table 2: Linear Regression Predicion of Summit Success

	Number of Summits
Oxygen Use Ratio	0.986***
	(0.019)
Number of Climbs	0.651***
	(0.003)
Died (1=Yes)	−0.188***
	(0.058)
Age	−0.009***
	(0.001)
Sex (1=Male)	0.084***
	(0.026)
Constant	−0.464***
	(0.041)
Observations	8,575
R^2	0.841
Adjusted R^2	0.840
Residual Std. Error	0.776 (df = 8569)

Note: *p<0.1; **p<0.05; ***p<0.01

The results show that all of our variables are significant predictors of number of successful summits, with oxygen ratio having the largest positive impact, with an increase of about one summit per unit increase in the oxygen ratio. That is to say, someone who always uses oxygen will summit one more mountain than someone who never uses oxygen. Number of climbs enters positively as well, as expected from our earlier regression. An increase in age is associated with a slight decrease in the number of summits, which makes sense given the physical requirements of climbing. A climber who dies is, of course, not going to have the opportunity to summit any more mountains, so we see *died* enter negatively as well. And finally, men summit slightly more peaks than women.

Oxygen Use Impact on Odds of Summiting (Logistic Regression)

However, to best understand oxygen's impact on summit success, it may be best to look at data on a per-climb basis, rather than aggregates over individual climbers' entire history on the mountains. For this, we turn to logistic regression to measure the impact that oxygen use has on the odds of summiting. We estimate the following simple logistic regression:

$$logit(msuccess_i) = \alpha + \beta mo2used_i$$

Where *msuccess* is a binary variable for whether a climber summited on a particular climb (1-yes, 0-no), and *mo2used* is a binary variable for whether or not they used oxygen on that climb (1-yes, 0-no). From the graph, we see that people who use oxygen are at increased odds of summiting, and estimating the model provides striking results. Oxygen use enters significantly in the model, and when we calculate the odds ratio

of summiting based on oxygen use, we get $e^{3.42} = 30.5$, indicating that those who use oxygen are at over 30 times the odds of summiting as those who do not.
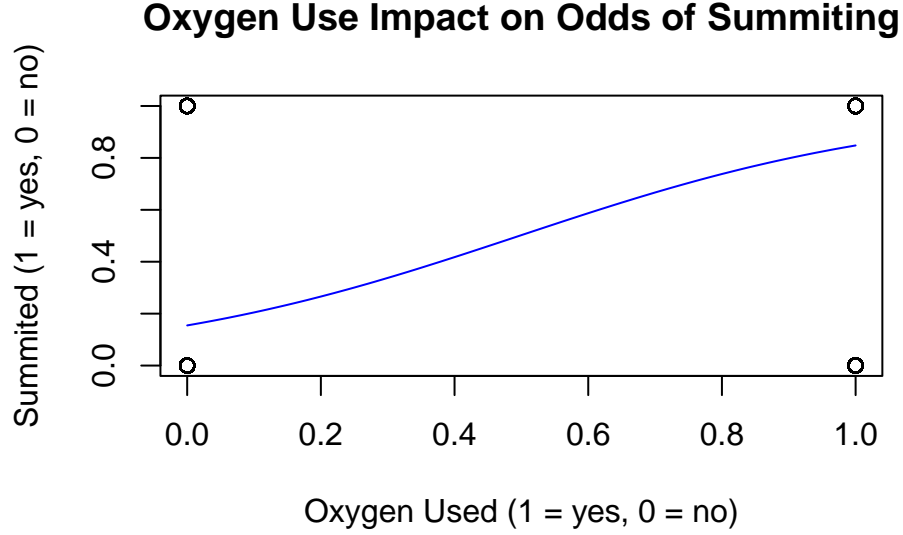
Oxygen Use Impact on Odds of Summiting

Table 3: Oxygen Use Impact on Odds of Summiting - Logistic Regression

	Log-Odds of Summiting
Oxygen Used (1=yes)	3.418***
	(0.041)
Constant	−1.699***
	(0.031)
Observations	18,662
Log Likelihood	−7,989.585
Akaike Inf. Crit.	15,983.170
Note:	*p<0.1; **p<0.05; ***p<0.01

Of course, as in the linear regression model, we should control for other variables to get a fuller picture of what impacts the odds of summiting. We now estimate the following model:

$$logit(msuccess_i) = \alpha + \beta_1 mo2used_i + \beta_2 mainseason_i + \beta_3 sex_i + \beta_4 hired_i + \beta_5 calcage_i$$

Where *mainseason* indicates whether a climb took place in the spring (1) or not (0), since spring is the most common time of year to climb. *sex* controls for the sex of the climber (1-male, 0-female), *hired* controls for whether the climber is hired crew, such as a Sherpa or high-altitude porter (1-yes, 0-no), and *calcage* is the age of the climber. Again, we see that using oxygen significantly increases the odds of summiting, with those using oxygen at about 20 times the odds of summiting as those who don't. Even when controlling for other factors, oxygen has by far the largest impact on odds of summiting. Hired Sherpas and porters are also at slightly increased (1.66 times) odds of summiting. This makes sense given their experience and common role as guides on commercial expeditions, so if any portion of an expedition summits, it is likely some

Sherpas will be amongst the summiters, guiding clients to the top. Each year increase in age leads to very slightly decreased odds of summiting, as expected given the physical nature of mountaineering. A somewhat unexpected result is the decreased (0.55 times) odds of summiting when climbing in the main spring season. We'd expect there to be increased odds given the spring having the most favorable weather conditions in the Himalaya, though perhaps the negative result is due to the larger portion of inexperienced climbers on commercial expeditions outweighing this. Finally, men are at slightly decreased odds of summiting, but the result is not significant.

Table 4: Oxygen Use Impact on Odds of Summiting - Logistic Regression

	Log-Odds of Summiting
Oxygen Used (1=yes)	3.004***
	(0.047)
Main Season (1=Spring, 0=Other)	−0.568***
	(0.050)
Sex (1=Male)	−0.013
	(0.072)
Hired (1=Yes)	0.508***
	(0.050)
Age	−0.020***
	(0.002)
Constant	−0.246**
	(0.108)
Observations	15,936
Log Likelihood	−7,136.121
Akaike Inf. Crit.	14,284.240
Note:	*p<0.1; **p<0.05; ***p<0.01

Weather Analysis

Models to Predict Barometric Pressue

We are interested in the utilization of supplemental oxygen to support climbers on their expeditions. One factor that can impact the demand for supplemental oxygen is weather conditions. As barometric pressure drops, the oxygen saturation drops, thus increasing the likelihood of needing supplemental oxygen. We look at weather conditions at Mt. Everest Basecamp during the spring climbing seasons of 2020 and 2021 to see if there are other weather factors to predict barometric pressure. Looking at basic correlation, we decide to first estimate a uni-variate model that uses temperature to predict barometric pressure.

$$Barometric\ Pressure = \beta_1 Temperature + \epsilon \tag{1}$$

We find a statistically coefficient on temperature of 0.347, meaning that for every 1 degree Celsius increase in temperature we expect barometric pressure to increase by 0.347 hPa. And, 34% of the variation in barometric pressure can be explained by the variation in temperature.

Table 5: Barometric Pressure and Temperature

	Barometric pressure (hPa)
Temperature (degrees C)	0.347***
	(0.007)
Constant	534.027***
	(0.048)
Observations	4,414
R^2	0.336
Adjusted R^2	0.335
Residual Std. Error	1.926 (df = 4412)

Note: *p<0.1; **p<0.05; ***p<0.01

Spring Temperatures and Barometric Pressure at Everes

Residual Plot for Pressure ~ Temperature
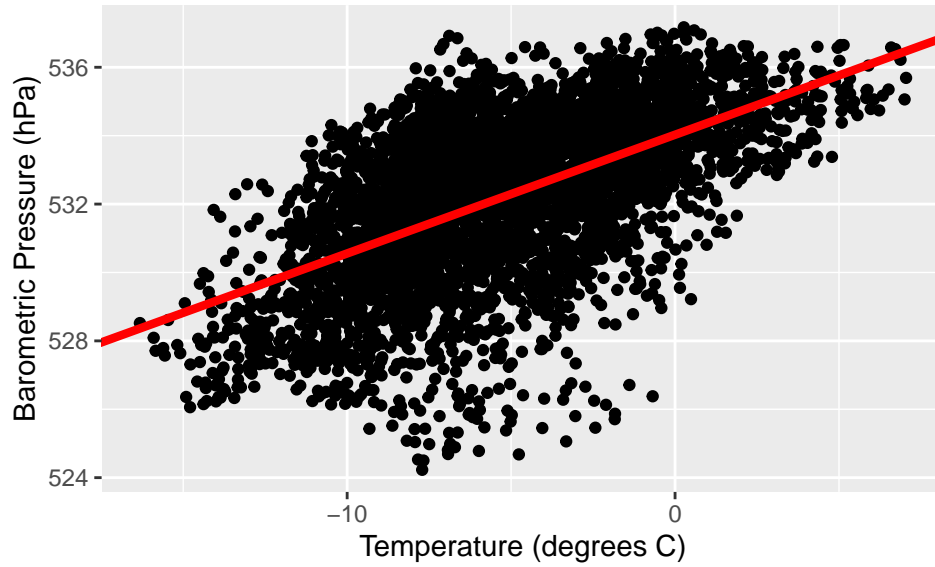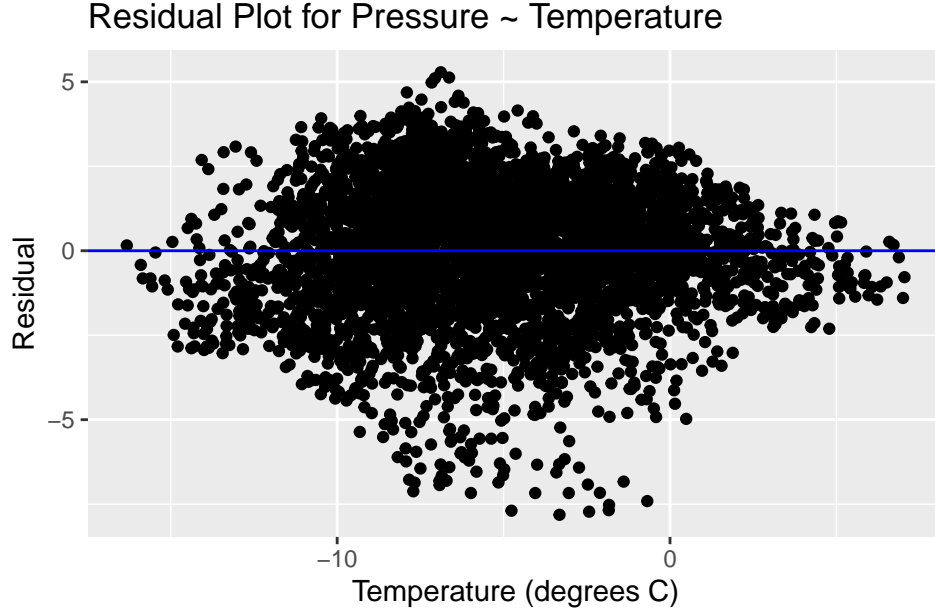
Then, we estimate a bi-variate model that uses temperature and binary variable of high humidity. Ex-ante, we would not suspect that high humidity is a statistically significant predictor as we see low correlation between high humidity and our outcome of barometric pressure, which our plot adding a third dimension of high humidity as color supports.

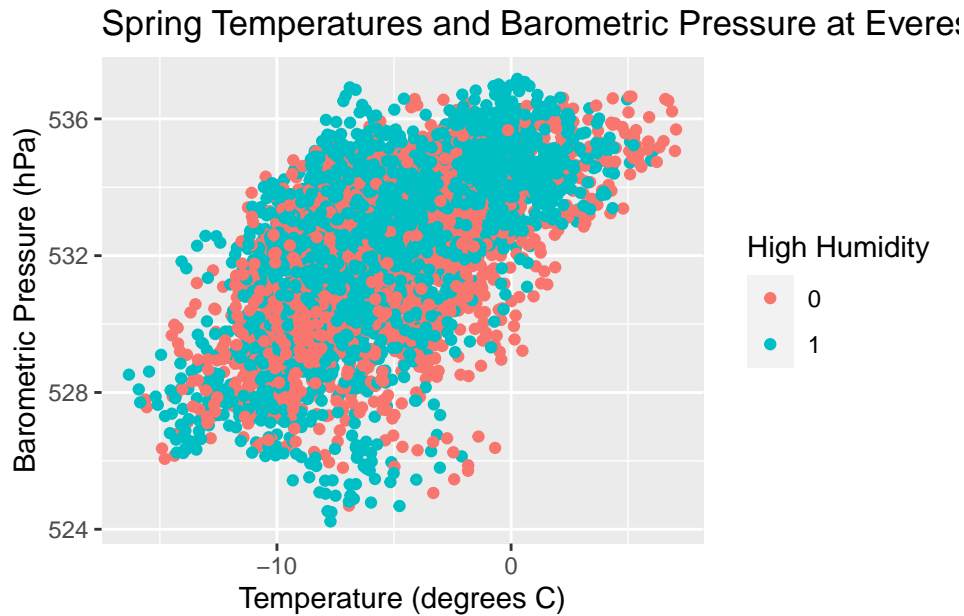$$Barometric\,Pressure = \beta_1 Temperature + \beta_2 High\,Humidity + \epsilon \tag{2}$$

Table 6: Barometric Pressure by Temperature and High Humidity

|	Barometric pressure (hPa)
Temperature (degrees C)	0.346***
	(0.007)
High Humidity	0.647***
	(0.058)
Constant	533.665***
	(0.057)
Observations	4,414
R^2	0.354
Adjusted R^2	0.354
Residual Std. Error	1.899 (df = 4411)
Note:	*p<0.1; **p<0.05; ***p<0.01

We unexpectedly find that the coefficient on high humidity of 0.647 is statistically significant. This means we expect barometric pressure to be 0.647 hPa higher on days with high humidity, holding temperature constant. However, our R-squared coefficient of determination only increase 0.35, so adding the second predictor only increases about power to explain variation in barometric pressure by 1 percentage point. We

17

ultimately conclude that adding the high humidity predictor isn't advantageous, and we should utilizing our uni-variate model.

Average and Seasonal Variation in Temperature

We now have an indirect pathway for predicting the need for supplemental oxygen: at lower temperatures, the barometric pressure is lower, and the oxygen saturation is lower. We only expect to hold expeditions during the spring hiking season (March - May). Therefore, we will look at our average temperature readings for the spring season to ascertain the average temperature at Everest basecamp and seasonal variation in the temperature. We take a random samples of 50 temperature readings to construct 95% confidence intervals for the average temperature. We very reliably contain the true population mean of -8.48 degree Celsius in our confidence intervals. But, we find that the underlying temperatures do not follow a normal distribution so we end up overestimating more often then underestimating. We apply Fourier analysis to our temperature data. Using the first three coefficients (that correspond to our 3 months of spring data), we are able to capture the overall seasonal trend. Temperature is lower in March into the beginning of April and reaches season highs in mid-May. Yet, our models seems to have issues with estimating temperatures at the beginning and end of the season. When we increase the number of coefficients to 13 (to correspond to the 13 weeks of data), we better approximate our data but without getting into the daily variation of temperature. We would likely want to use a model that incorporates between 3 and 13 Fourier coefficients to predict temperature at any point in the spring.

Average temperatures in Spring


Average temperatures in Spring
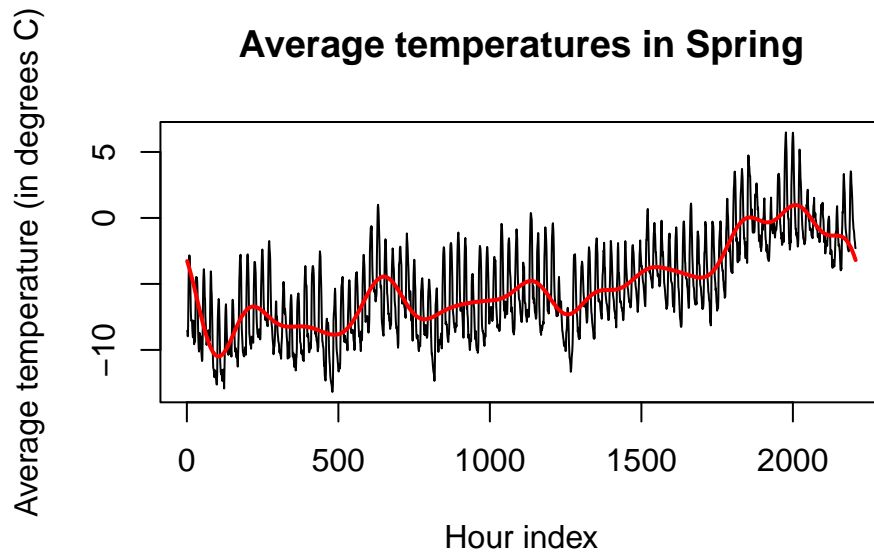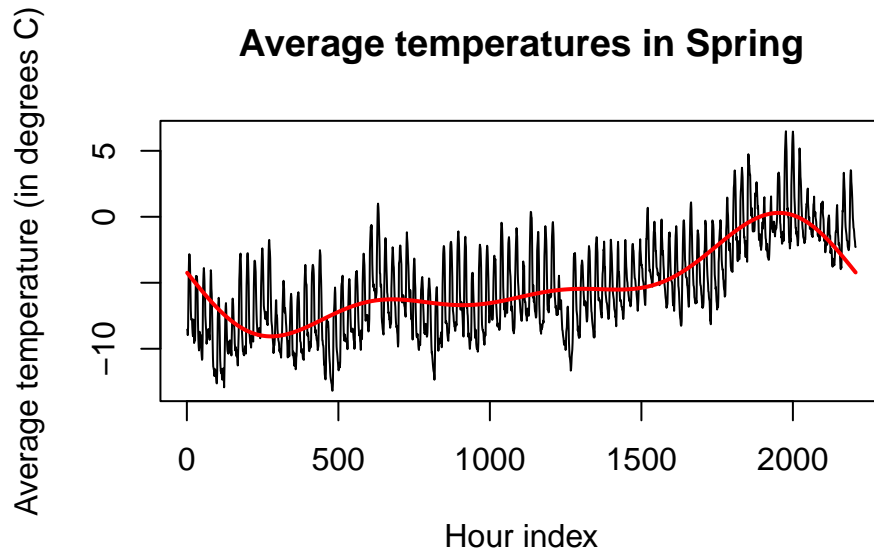
Power Analysis

In the final section of our analysis, we design a randomized controlled trial that could provide an estimate of the causal effect of supplemental oxygen use on mortality. We are interested specifically in whether the risk of death decreases when using oxygen. No one likes dying, so for reasons of ethics and liability, we would like to minimize the risk of it happening on our trips. Moreover, here at Otzi's Adventures, we sell supplemental oxygen, so solid evidence that it reduces mortality risk could be a great marketing asset in our vertically integrated company.

Our study design assumes that we randomize entire expeditions to treatment and control groups, not individual climbers. Otherwise, climbers within the same expedition would have differing treatment statuses,

which might create confusion and logistical problems. Moreover, our treatment condition is the requirement to use oxygen, in contrast to the control condition of being able to choose oxygen or not (prohibiting anyone from using oxygen would be unethical). If some climbers on an expedition were required to use oxygen, perhaps this would encourage all climbers to use oxygen, eliminating the identifying variation in oxygen use that we need to estimate a causal effect. Our primary design challenge then is to identify the number of expeditions that we must randomize to achieve 80% statistical power (hopefully making us eligible for NIH funding), assuming 50:50 allocation between treatment and control.
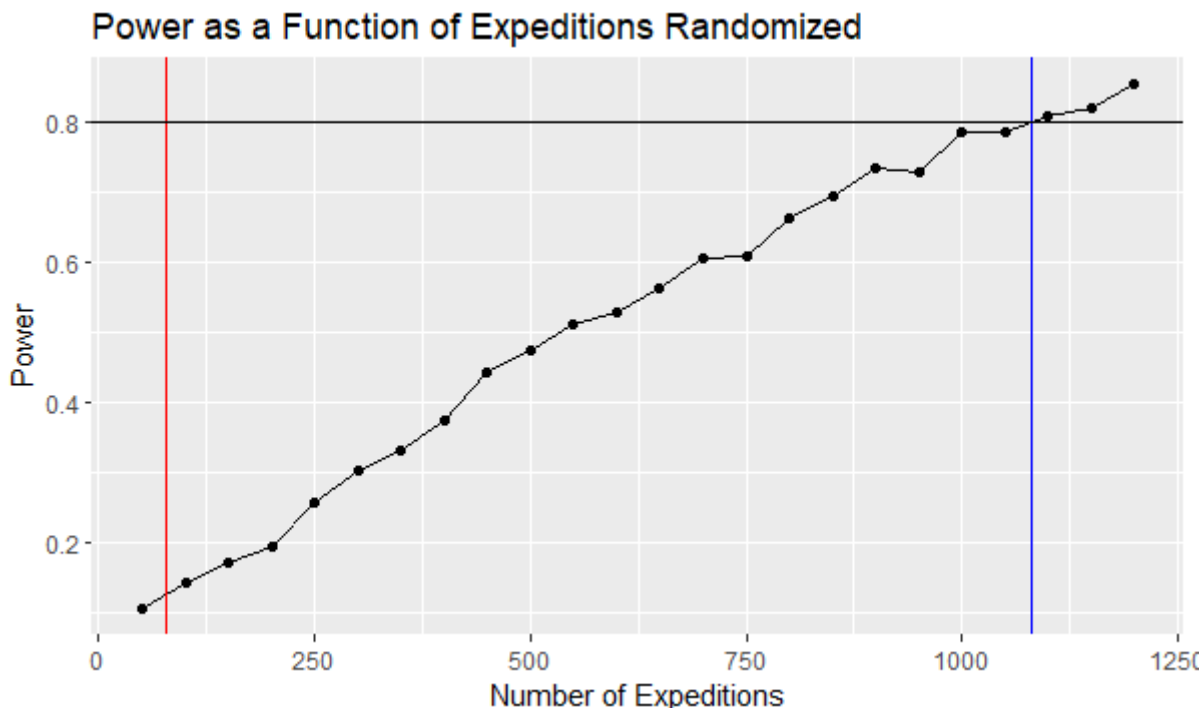
To estimate required sample, we first run a multilevel linear probability model on past data, with death as the outcome. We allow each expedition to have a separate intercept and separate coefficient on supplemental oxygen use that each vary randomly around some population mean. Next, to get a sample size number using a classical approach, we exploit the fact that if you have a true effect 2.8 standard deviations away from the assumed effect under a null hypothesis (in our case, this is 0, or no effect), then you will have 80% power, assuming an alpha value of 5% and a two-tailed test. We take the standard error formula for a treatment effect in a cluster-randomized trial, re-arrange to solve for number of clusters (i.e., expeditions), and substitute in the population variance in the coefficient on oxygen use along with variance due to estimation in a "typical" expedition of mean size and mean oxygen use rate. This approach yields a sample size requirement of ~80 expeditions.

As a check, we estimate sample needs using a second, simulation-based approach. Here, we use the results from the multilevel model to parameterize a multivariate normal distribution of expedition intercepts and slopes on oxygen use. We repeatedly sample from this distribution. Next, we use each expedition's parameters to parameterize a probability value p for a Bernoulli distribution from which to sample climber-level observations. Lastly, we estimate a model, to see whether we detect the true treatment effect we baked into the multivariate normal distribution. The proportion of iterations that result in a statistically significant treatment estimate is the power for that sample size. Finally, we plot power as a function of sample size in a graph. As the reader will note, the simulations suggest we need a sample of over 1000 expeditions for 80% power, in contrast to the 80 expeditions suggested by the classical approach, plotted as a red line. We suspect that this is due to functional form: the linear probability model does not bound probability predictions at 0. We had to truncate to 0 when the predicted value was negative during the simulations,

effectively reducing our assumed treatment effect, and thereby raising sample size requirements for fixed power. The classical approach is ignorant of this constraint because it does not require actually re-sampling data, hence the classical results are likely misleading.

Note: The code that generated the power graph is not included in this .rmd file. The full code takes approximately two hours to run, so we decided to just leave it in the main Otzi's Adventures.R script created for our analysis.