CKME 136 - Mount Rainier Climbing Dataset

Nicole Smith

```

#first, load the dataset  
rainier<-read.csv("C:/Users/nicol/OneDrive/Desktop/rainier.csv")  
summary(rainier)

## Date Route Attempted   
## 7/11/2014: 51 Disappointment Cleaver:2728 Min. : 1.000   
## 7/25/2014: 47 Emmons-Winthrop : 632 1st Qu.: 2.000   
## 5/31/2014: 46 Kautz Glacier : 182 Median : 3.000   
## 7/3/2015 : 46 Fuhrers Finger : 92 Mean : 5.139   
## 6/21/2014: 45 Gibralter Ledges : 76 3rd Qu.: 8.000   
## 6/7/2014 : 43 Little Tahoma : 70 Max. :26.000   
## (Other) :3799 (Other) : 297   
## Succeeded Success\_Percentage Battery\_V Temp   
## Min. : 0.000 Min. : 0.0000 Min. :12.99 Min. : 6.835   
## 1st Qu.: 0.000 1st Qu.: 0.0000 1st Qu.:13.45 1st Qu.:33.377   
## Median : 2.000 Median : 0.5833 Median :13.49 Median :40.049   
## Mean : 2.711 Mean : 0.4990 Mean :13.50 Mean :40.711   
## 3rd Qu.: 4.000 3rd Qu.: 1.0000 3rd Qu.:13.55 3rd Qu.:48.629   
## Max. :71.000 Max. :14.2000 Max. :13.79 Max. :56.154   
## NA's :2182 NA's :2182   
## Humidity Wind\_Speed Wind\_Direction Solare\_Radiation   
## Min. : 12.36 Min. : 0.7535 Min. : 13.49 Min. : 0.0331   
## 1st Qu.: 30.06 1st Qu.: 4.4974 1st Qu.: 55.17 1st Qu.:232.4764   
## Median : 46.14 Median : 7.3960 Median :171.63 Median :308.7490   
## Mean : 47.83 Mean :10.5092 Mean :155.34 Mean :270.7466   
## 3rd Qu.: 64.75 3rd Qu.:14.6579 3rd Qu.:235.50 3rd Qu.:352.2212   
## Max. :100.00 Max. :65.1383 Max. :280.38 Max. :368.0561   
## NA's :2182 NA's :2183 NA's :2182 NA's :2246

#remove rows with more than 4 nas.  
nrow(rainier)

## [1] 4077

keep<-rowSums(is.na(rainier)) <4  
rainierclean<-rainier[keep, ]  
  
nrow(rainierclean)

## [1] 1895

#count how many nas remain  
sum(is.na(rainierclean))

## [1] 65

table(is.na(rainierclean))

##   
## FALSE TRUE   
## 20780 65

sum(is.na(rainierclean$Solar\_Radiation))

## [1] 0

#now remove all remaining nas  
rainierclean<- rainierclean[complete.cases(rainierclean), ]  
  
nrow(rainierclean)

## [1] 1830

#now we need to remove any rows where the success percentage is greater than 1, since that is not logical.  
  
rainier.clean<-rainierclean[(rainierclean$Success\_Percentage<=1), ]  
  
nrow(rainier.clean)

## [1] 1824

summary(rainier.clean)

## Date Route Attempted   
## 7/3/2015 : 46 Disappointment Cleaver :1263 Min. : 1.000   
## 6/6/2015 : 40 Emmons-Winthrop : 293 1st Qu.: 2.000   
## 6/26/2015: 39 Kautz Glacier : 97 Median : 4.000   
## 5/30/2015: 35 Little Tahoma : 36 Mean : 5.384   
## 6/20/2015: 34 Liberty RIngraham Directge: 29 3rd Qu.: 9.000   
## 6/27/2015: 34 Fuhrers Finger : 26 Max. :12.000   
## (Other) :1596 (Other) : 80   
## Succeeded Success\_Percentage Battery\_V Temp   
## Min. : 0.000 Min. :0.0000 Min. :13.04 Min. : 8.221   
## 1st Qu.: 0.000 1st Qu.:0.0000 1st Qu.:13.45 1st Qu.:34.462   
## Median : 2.000 Median :0.5000 Median :13.49 Median :40.171   
## Mean : 2.641 Mean :0.4616 Mean :13.50 Mean :41.337   
## 3rd Qu.: 4.000 3rd Qu.:1.0000 3rd Qu.:13.55 3rd Qu.:48.766   
## Max. :12.000 Max. :1.0000 Max. :13.75 Max. :56.154   
##   
## Humidity Wind\_Speed Wind\_Direction Solare\_Radiation   
## Min. : 12.36 Min. : 0.7535 Min. : 13.49 Min. : 0.0331   
## 1st Qu.: 30.06 1st Qu.: 4.4103 1st Qu.: 49.57 1st Qu.:231.4682   
## Median : 46.01 Median : 7.1409 Median :164.19 Median :308.7490   
## Mean : 47.44 Mean : 9.9872 Mean :153.20 Mean :270.7424   
## 3rd Qu.: 64.33 3rd Qu.:13.6151 3rd Qu.:229.85 3rd Qu.:352.2212   
## Max. :100.00 Max. :65.1383 Max. :279.78 Max. :368.0561   
##

#now, we have a clean dataset.

#correlation  
  
  
cor(rainier.clean$Success\_Percentage, rainier.clean$Battery\_V)

## [1] -0.01747313

cor(rainier.clean$Success\_Percentage, rainier.clean$Temp)

## [1] 0.08782178

cor(rainier.clean$Success\_Percentage, rainier.clean$Humidity)

## [1] -0.05852576

cor(rainier.clean$Success\_Percentage, rainier.clean$Wind\_Speed)

## [1] -0.08966552

cor(rainier.clean$Success\_Percentage, rainier.clean$Wind\_Direction)

## [1] -0.06774289

#wind speed and temp have the highest effect on success rates, and battery voltage has the lest. for the linear model i will use temp, humidity, wind speed and wind direction.

#regression model  
regmodel<-lm(rainier.clean$Success\_Percentage ~ rainier.clean$Temp + rainier.clean$Humidity + rainier.clean$Wind\_Speed + rainier.clean$Wind\_Direction)  
  
regmodel

##   
## Call:  
## lm(formula = rainier.clean$Success\_Percentage ~ rainier.clean$Temp +   
## rainier.clean$Humidity + rainier.clean$Wind\_Speed + rainier.clean$Wind\_Direction)  
##   
## Coefficients:  
## (Intercept) rainier.clean$Temp   
## 0.2921594 0.0053524   
## rainier.clean$Humidity rainier.clean$Wind\_Speed   
## 0.0007881 -0.0044135   
## rainier.clean$Wind\_Direction   
## -0.0002947

#success rate=0.2921594+0.0007881(humidity)-0.0002947(wind direction)+0.0053524(temperature)-0.0044135(wind speed)

#look at which route has the best probability of success  
agg1<-aggregate(Success\_Percentage ~ Route, rainier.clean, mean)  
agg1

## Route Success\_Percentage  
## 1 Curtis RIngraham Directge 0.5000000  
## 2 Disappointment Cleaver 0.4696233  
## 3 Emmons-Winthrop 0.5145982  
## 4 Fuhrers Finger 0.1538462  
## 5 Gibralter Ledges 0.0000000  
## 6 glacier only - no summit attempt 0.0800000  
## 7 Ingraham Direct 0.1538462  
## 8 Kautz Cleaver 0.2857143  
## 9 Kautz Glacier 0.5046756  
## 10 Liberty RIngraham Directge 0.4827586  
## 11 Little Tahoma 0.4270833  
## 12 Mowich Face 0.3333333  
## 13 Nisqually Glacier 0.0000000  
## 14 Ptarmigan RIngraham Directge 0.5000000  
## 15 Success Cleaver 0.5000000  
## 16 Sunset RIngraham Directge 0.0000000  
## 17 Tahoma Cleaver 1.0000000  
## 18 Tahoma Glacier 0.4166667  
## 19 Wilson Headwall 0.0000000

agg2<-aggregate(Success\_Percentage ~ Route, rainier.clean, sum)  
agg2

## Route Success\_Percentage  
## 1 Curtis RIngraham Directge 1.000000  
## 2 Disappointment Cleaver 593.134271  
## 3 Emmons-Winthrop 150.777273  
## 4 Fuhrers Finger 4.000000  
## 5 Gibralter Ledges 0.000000  
## 6 glacier only - no summit attempt 2.000000  
## 7 Ingraham Direct 2.000000  
## 8 Kautz Cleaver 2.000000  
## 9 Kautz Glacier 48.953535  
## 10 Liberty RIngraham Directge 14.000000  
## 11 Little Tahoma 15.375000  
## 12 Mowich Face 1.000000  
## 13 Nisqually Glacier 0.000000  
## 14 Ptarmigan RIngraham Directge 4.000000  
## 15 Success Cleaver 1.000000  
## 16 Sunset RIngraham Directge 0.000000  
## 17 Tahoma Cleaver 1.000000  
## 18 Tahoma Glacier 1.666667  
## 19 Wilson Headwall 0.000000

#although the tahoma cleaver has the highest average success percentage, of 100%, there is only 1 attempt, so we do not have enough evidence to conclude that this route has the highest success percentage.

#export to excel for sql analysis  
#write.xlsx(rainier.clean, "C:/Users/nicol/OneDrive/Desktop/rainierclean.xlsx")