

# Data Profiling Report

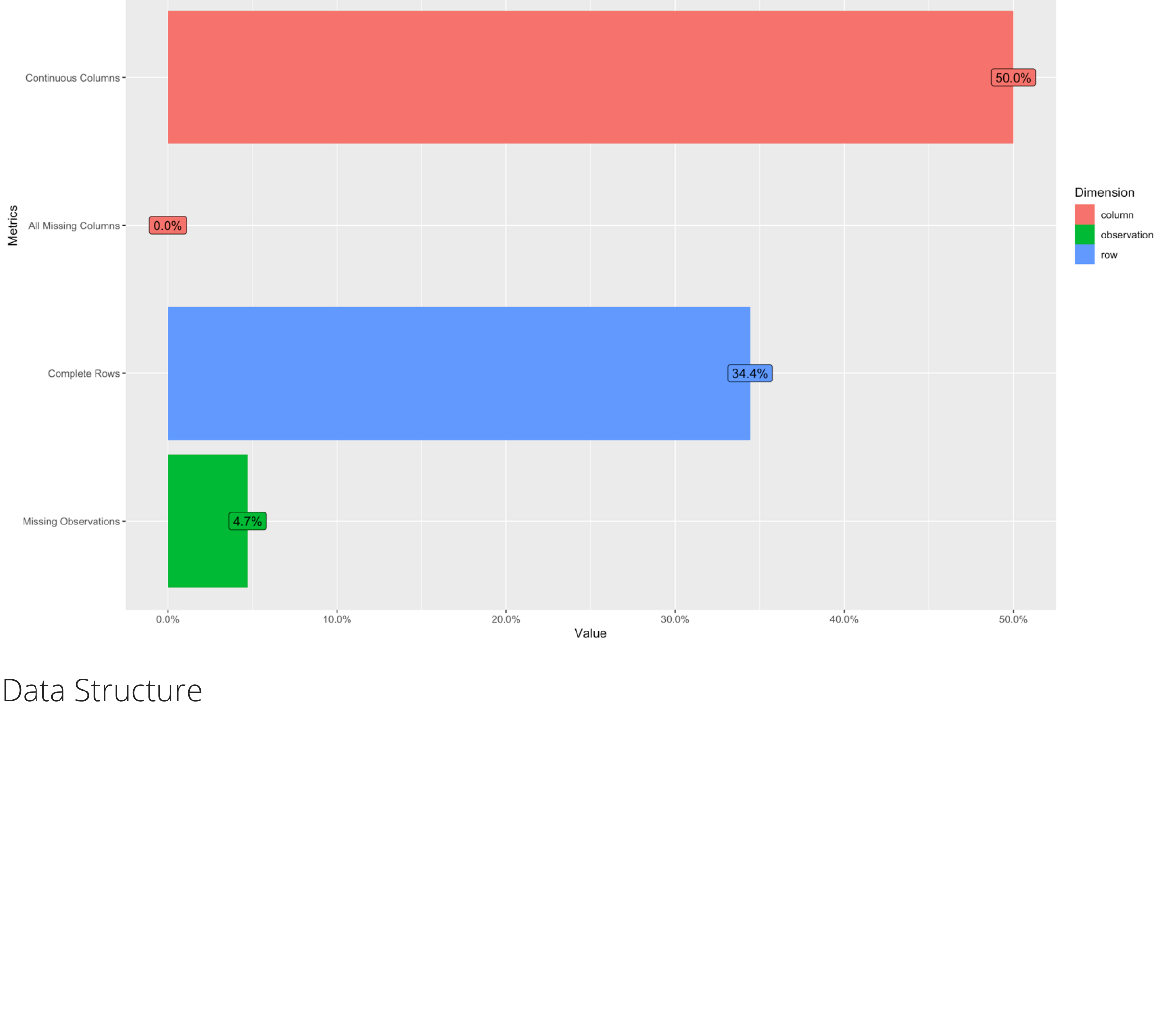
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (with frequency)
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

## Basic Statistics

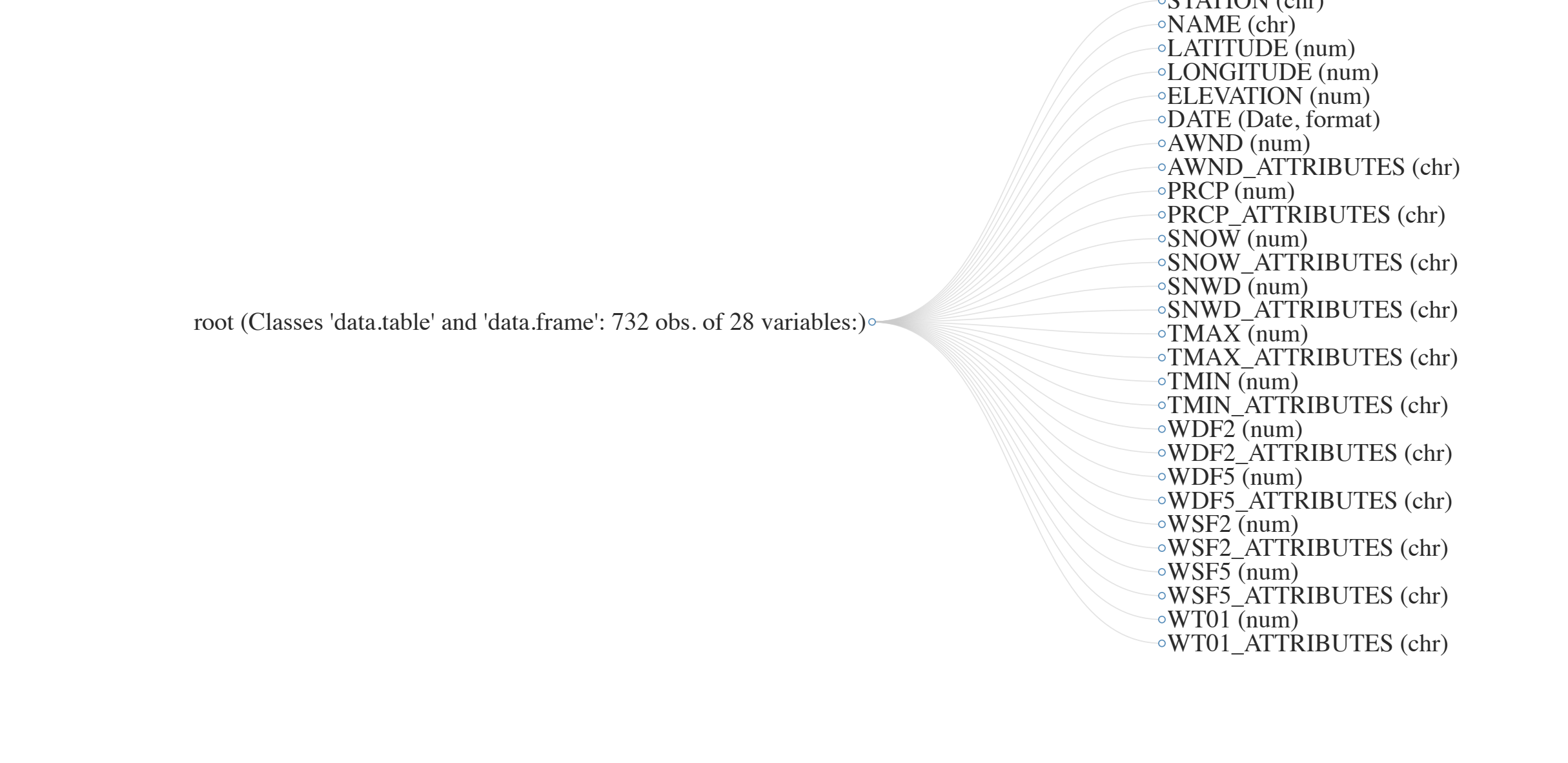
### Raw Counts

Name	Value
Rows	732
Columns	28
Discrete columns	14
Continuous columns	14
All missing columns	0
Missing observations	968
Complete Rows	252
Total observations	20,496
Memory allocation	167.5 Kb

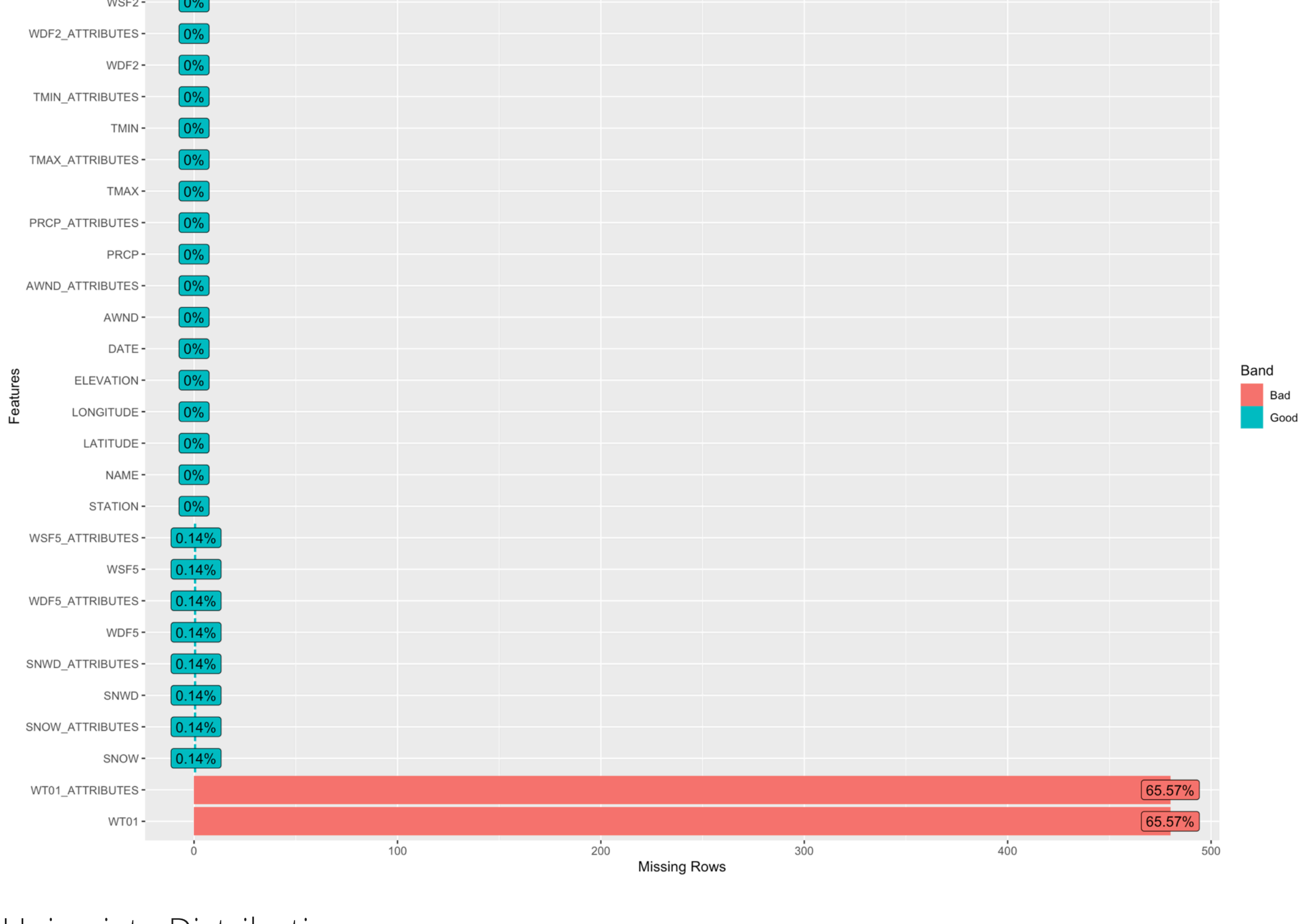
### Percentages



## Data Structure

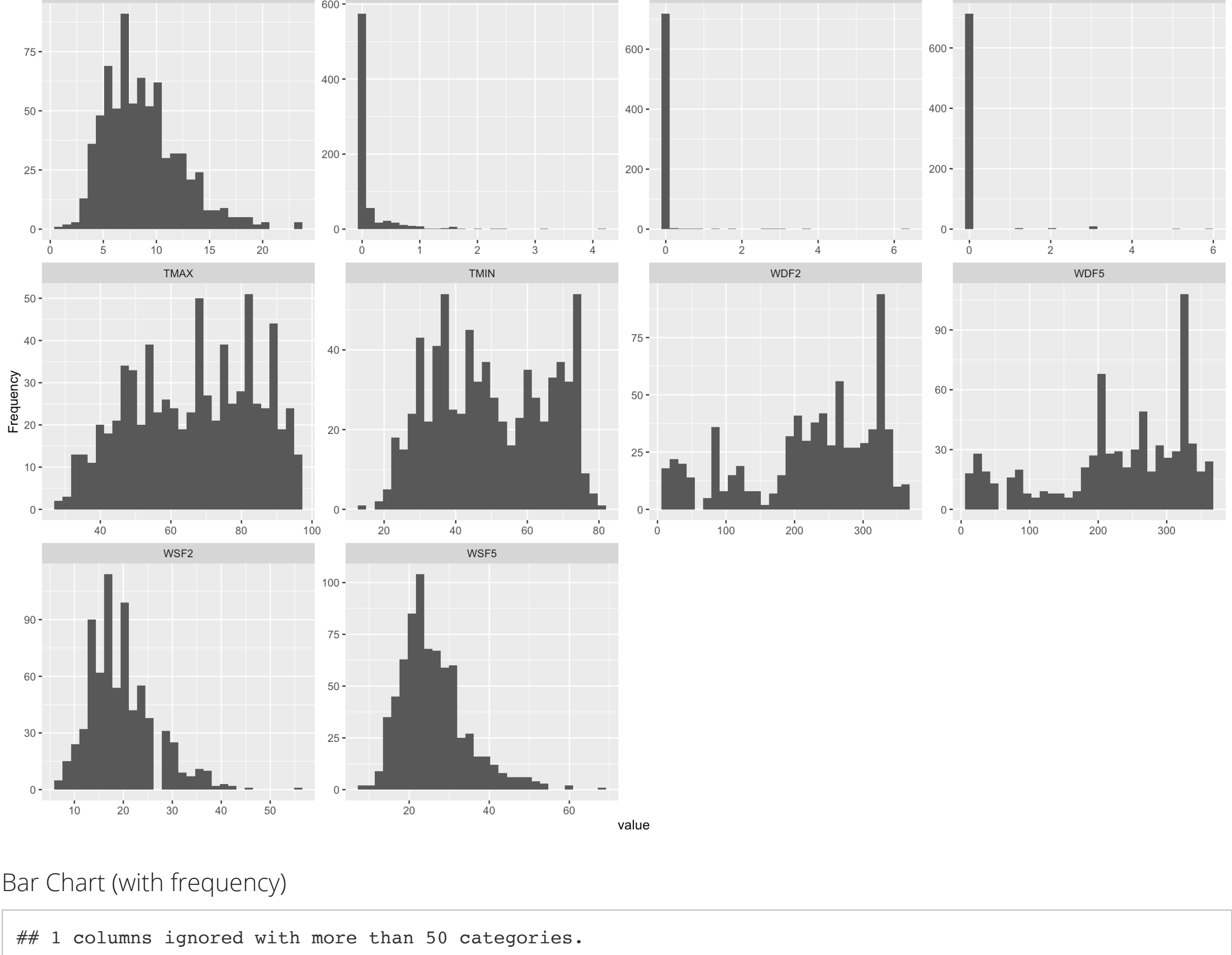


## Missing Data Profile



## Univariate Distribution

### Histogram



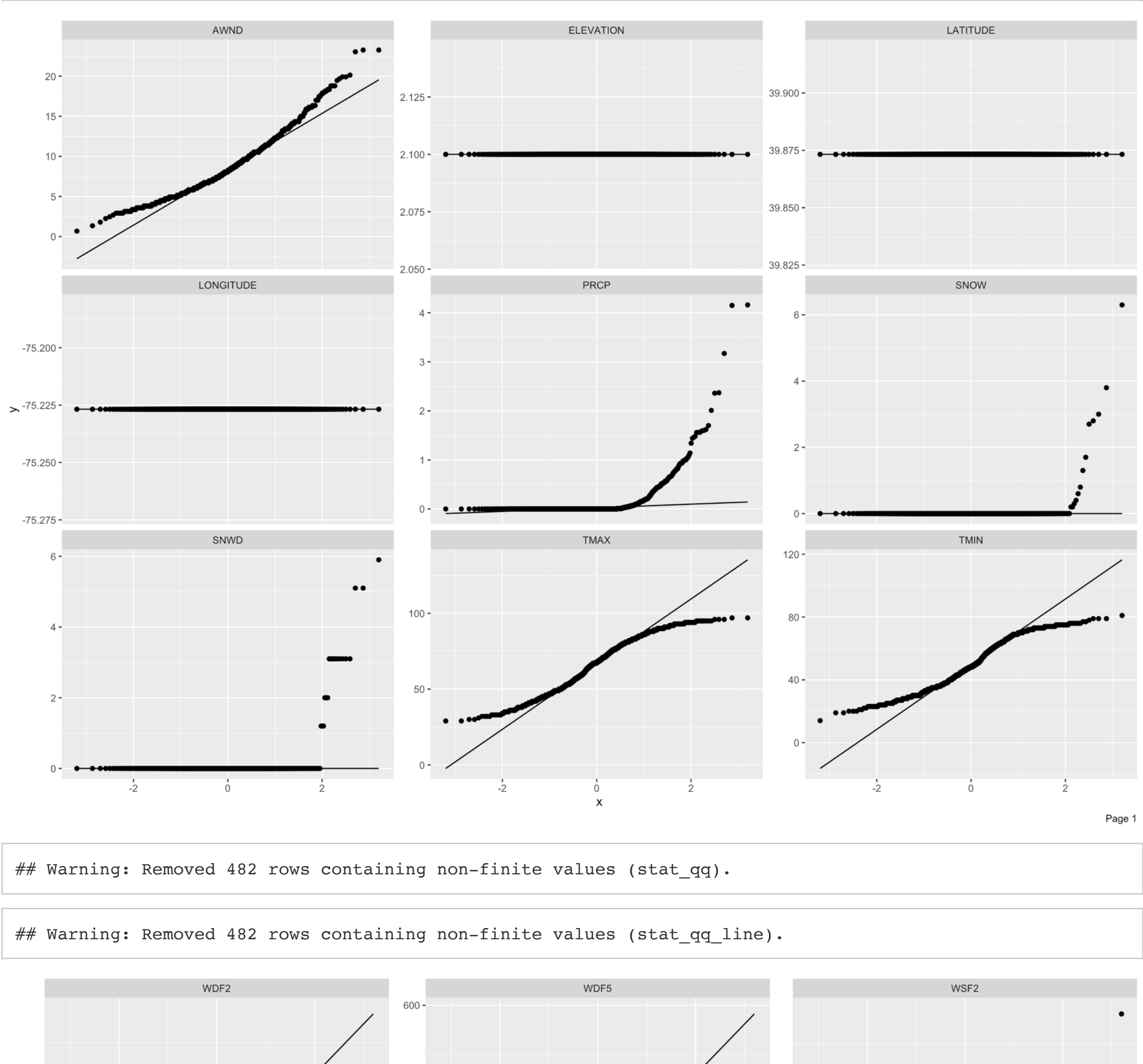
### Bar Chart (with frequency)

## 1 columns ignored with more than 50 categories.  
## DATE: 732 categories

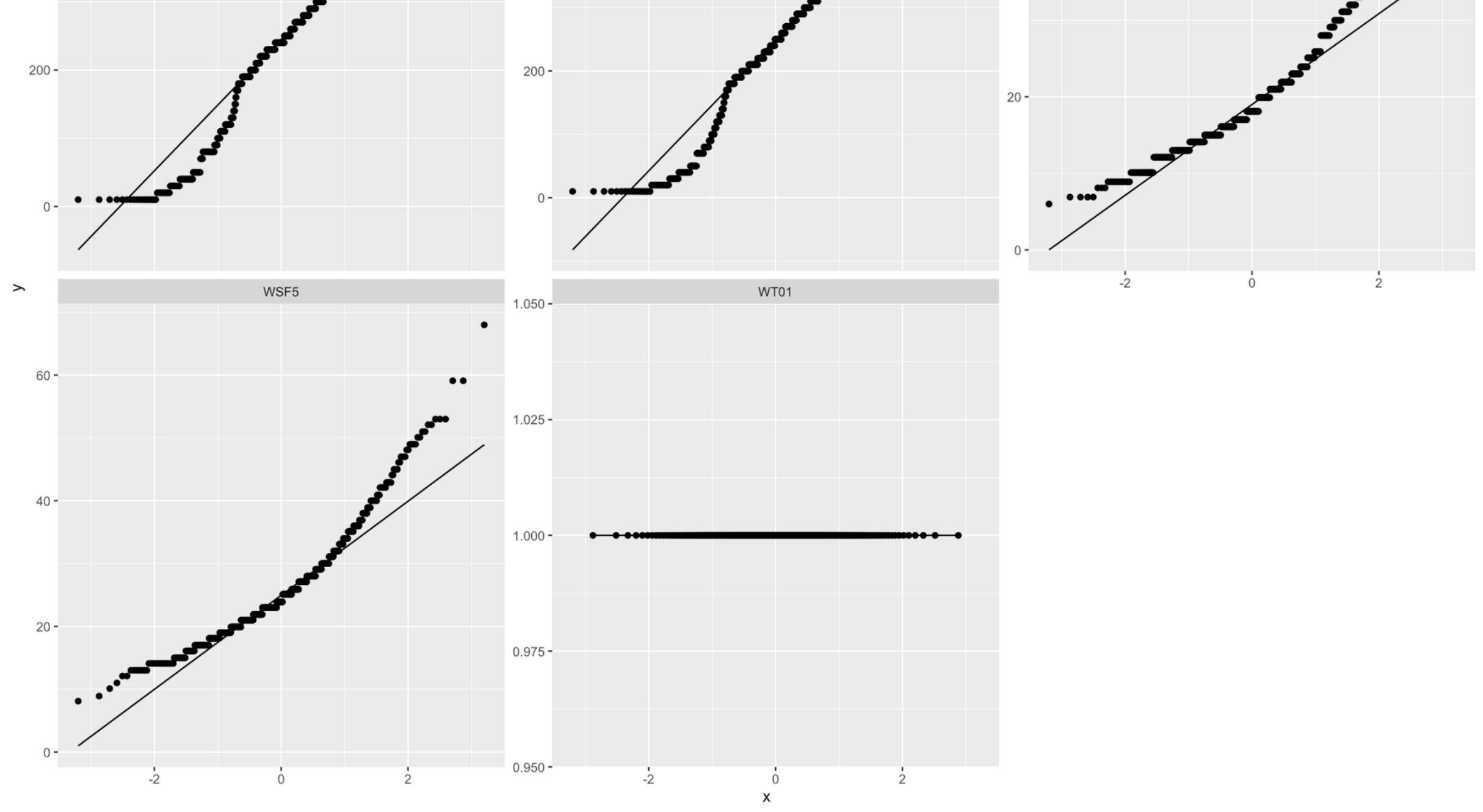


### QQ Plot

## Warning: Removed 2 rows containing non-finite values (stat\_qq).  
## Warning: Removed 2 rows containing non-finite values (stat\_qq\_line).



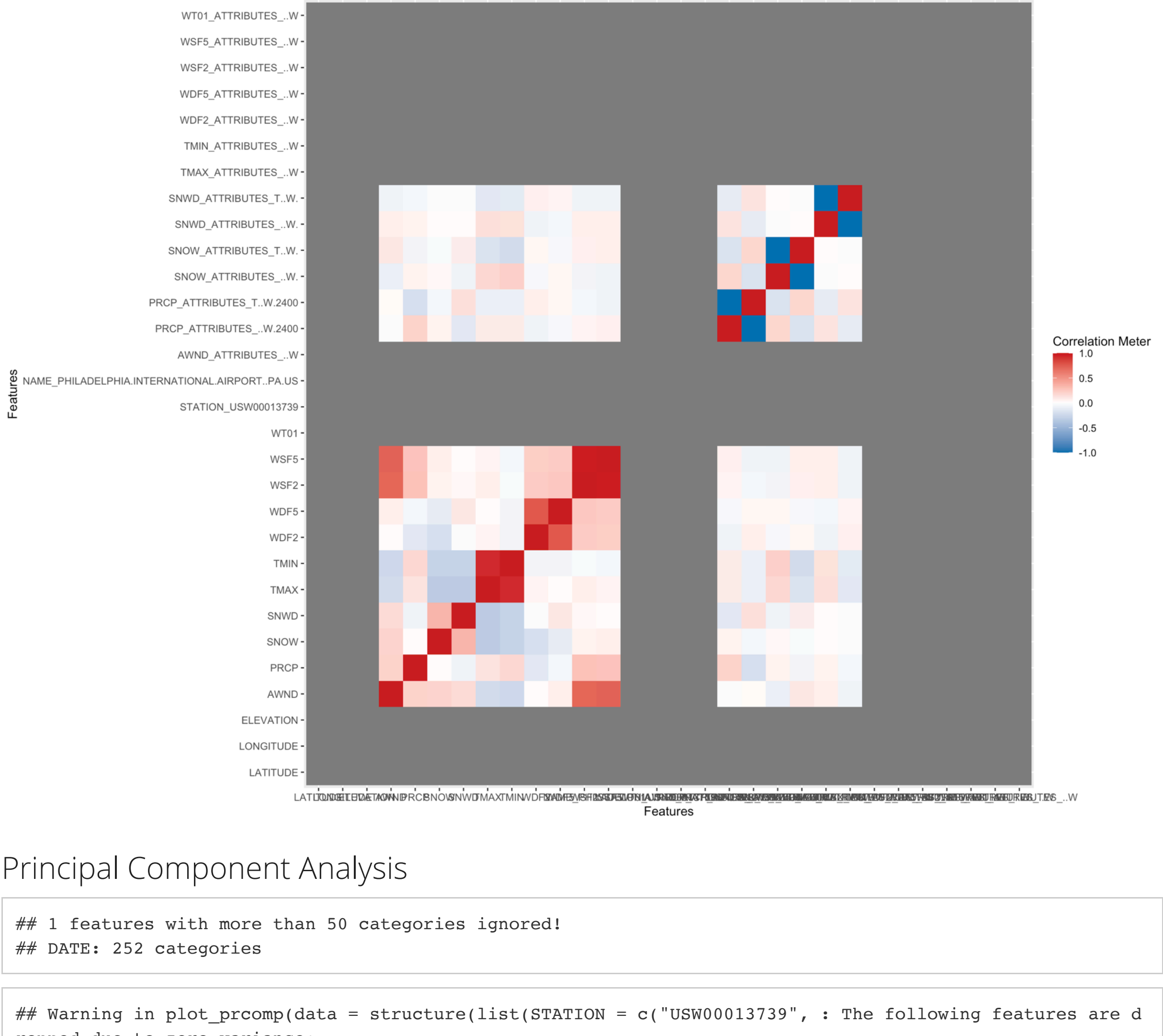
## Warning: Removed 482 rows containing non-finite values (stat\_qq).  
## Warning: Removed 482 rows containing non-finite values (stat\_qq\_line).



## Correlation Analysis

## 1 features with more than 20 categories ignored!  
## DATE: 252 categories

## Warning in cor(x = structure(list(LATITUDE = c(39.87326, 39.87326, 39.87326, :  
## the standard deviation is zero



## Principal Component Analysis

## 1 features with more than 50 categories ignored!  
## DATE: 252 categories

## Warning in plot\_prcomp(data = structure(list(STATION = c("USW00013739", : The following features are dropped due to zero variance:  
## \* LATITUDE  
## \* LONGITUDE  
## \* ELEVATION  
## \* WT01  
## \* STATION\_USW00013739  
## \* NAME\_PHLADELPHIA.INTERNATIONAL.AIRPORT..PA.US  
## \* AWND\_ATTRIBUTES\_..W  
## \* TMAX\_ATTRIBUTES\_..W  
## \* TMIN\_ATTRIBUTES\_..W  
## \* WDF2\_ATTRIBUTES\_..W  
## \* WDF5\_ATTRIBUTES\_..W  
## \* WSPF\_ATTRIBUTES\_..W  
## \* WT01\_ATTRIBUTES\_..W

