

Data Profiling Report

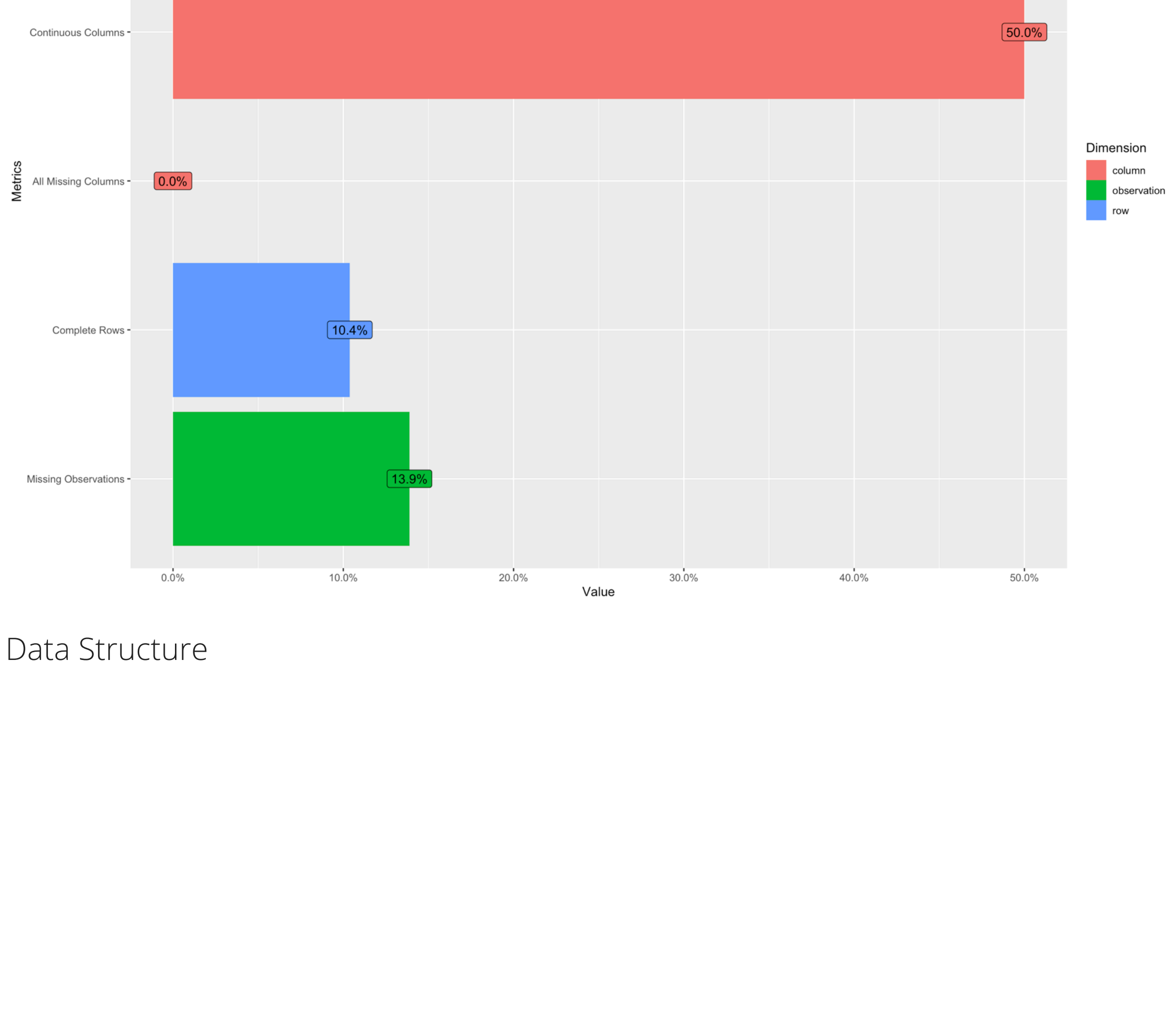
- Basic Statistics
 - Raw Counts
 - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
 - Histogram
 - Bar Chart (with frequency)
 - QQ Plot
- Correlation Analysis
- Principal Component Analysis

Basic Statistics

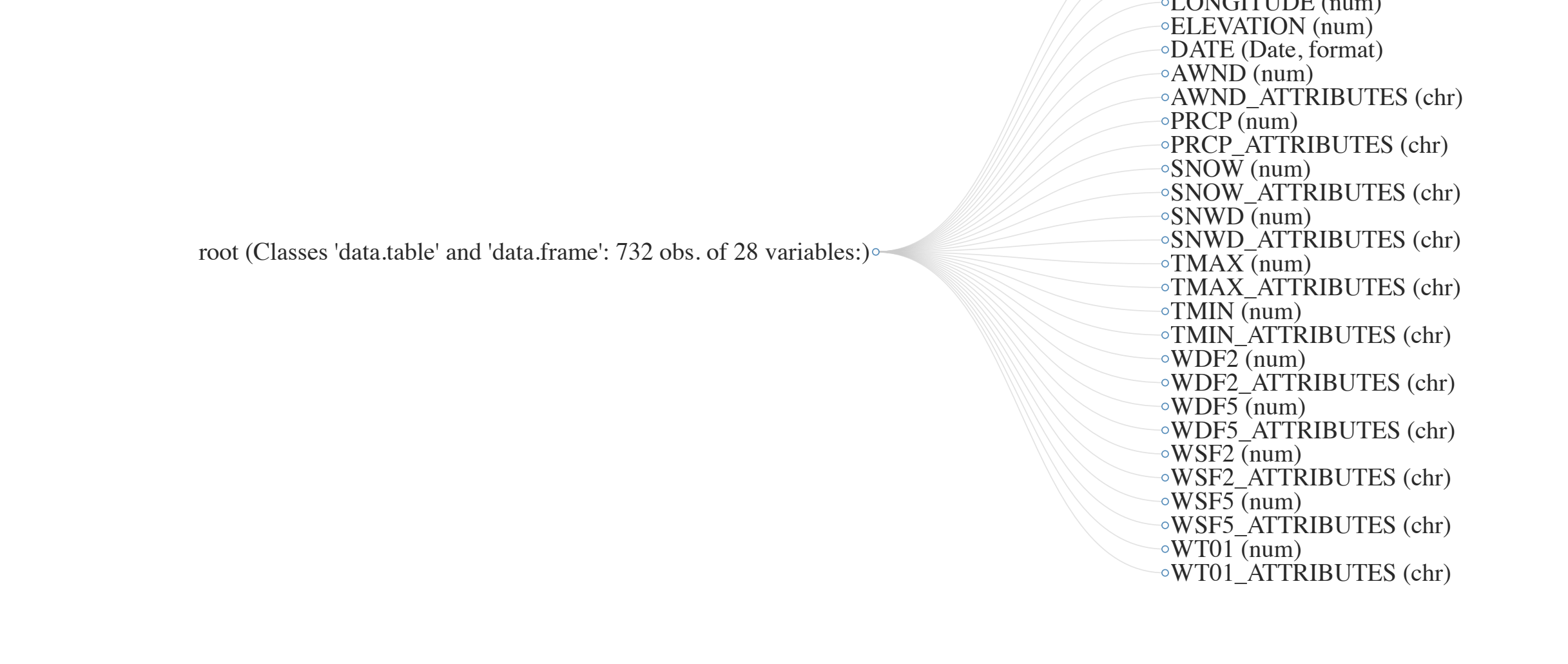
Raw Counts

Name	Value
Rows	732
Columns	28
Discrete columns	14
Continuous columns	14
All missing columns	0
Missing observations	2,846
Complete Rows	76
Total observations	20,496
Memory allocation	167.4 Kb

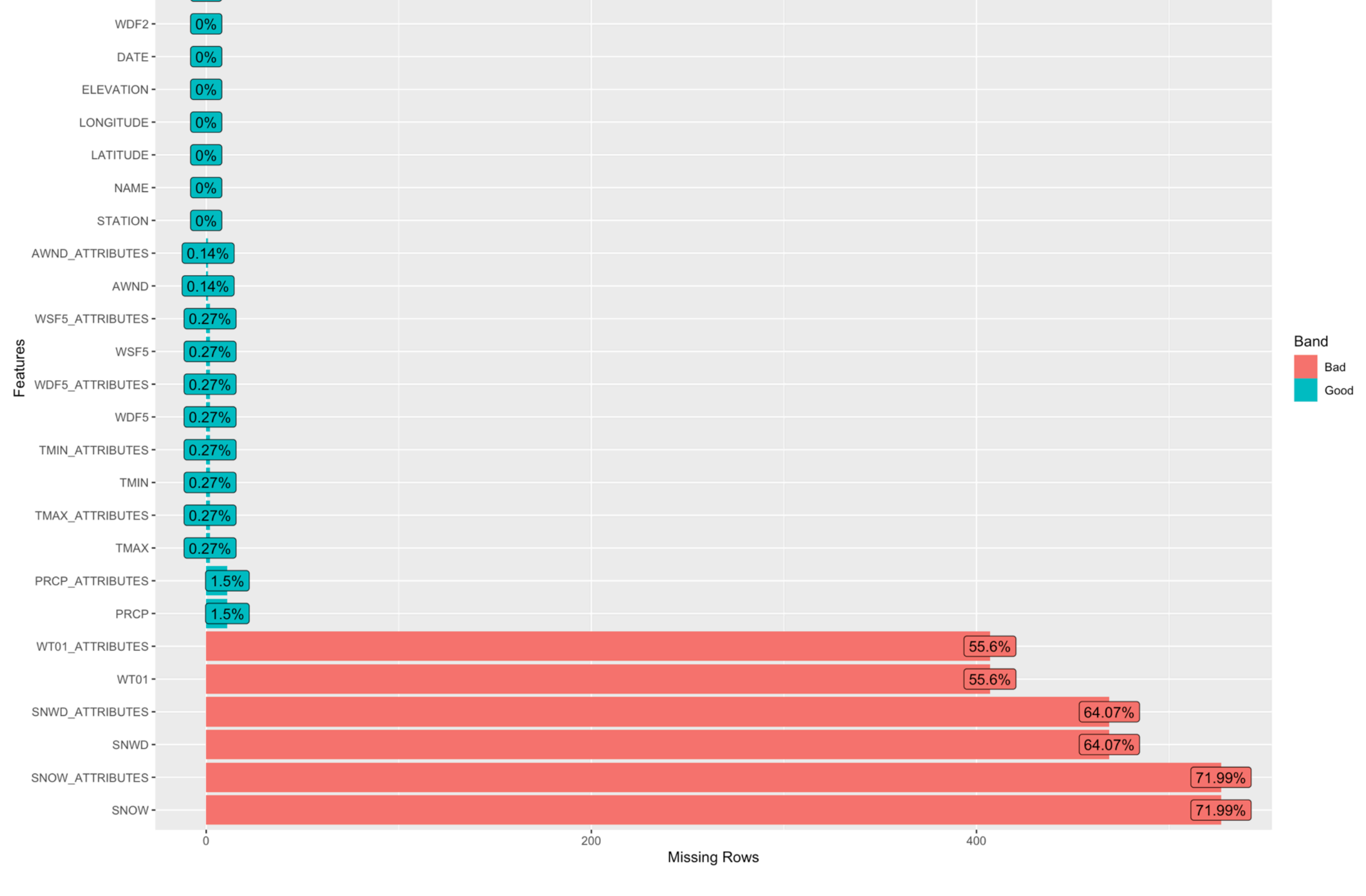
Percentages



Data Structure

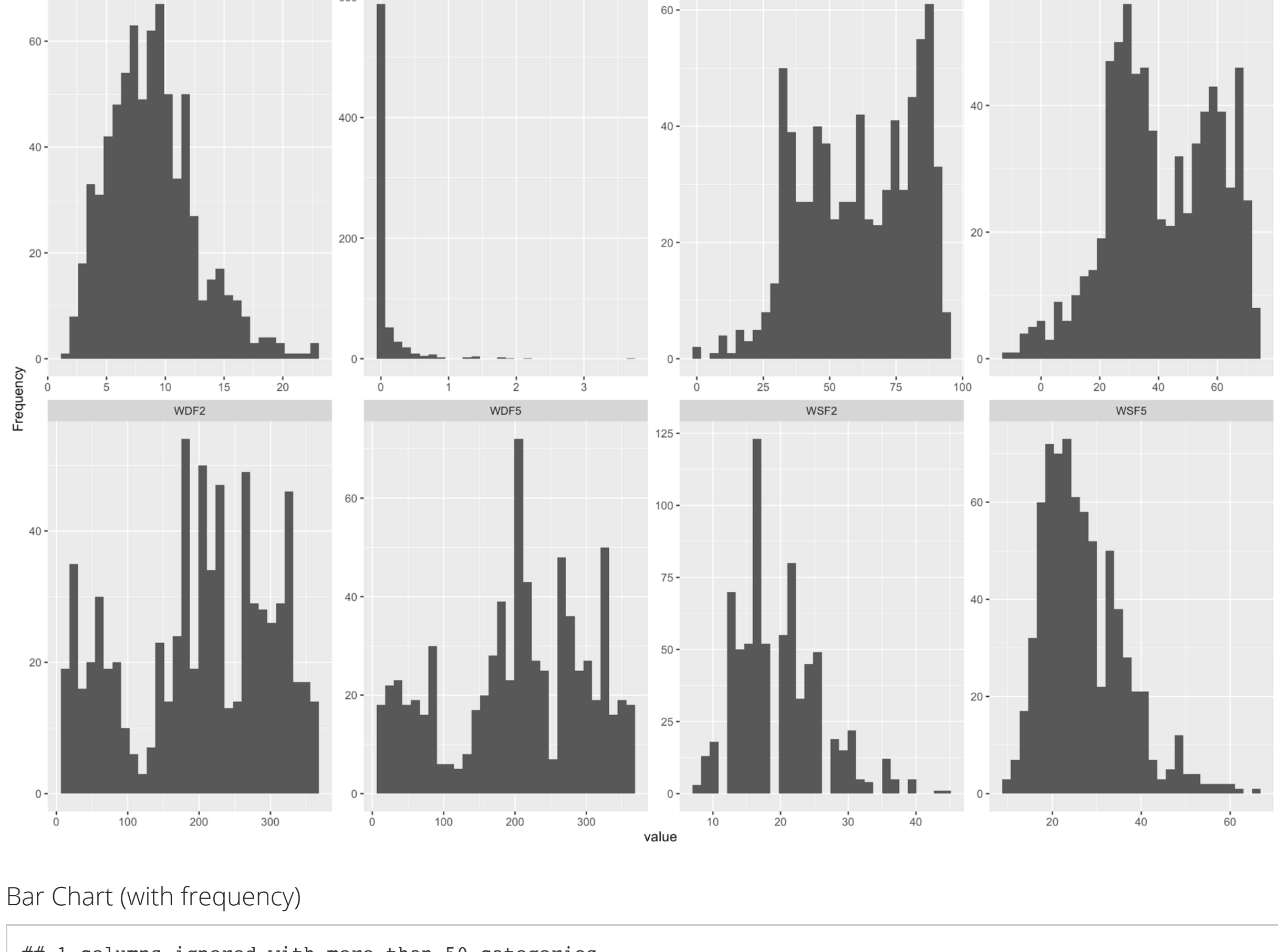


Missing Data Profile



Univariate Distribution

Histogram



Bar Chart (with frequency)

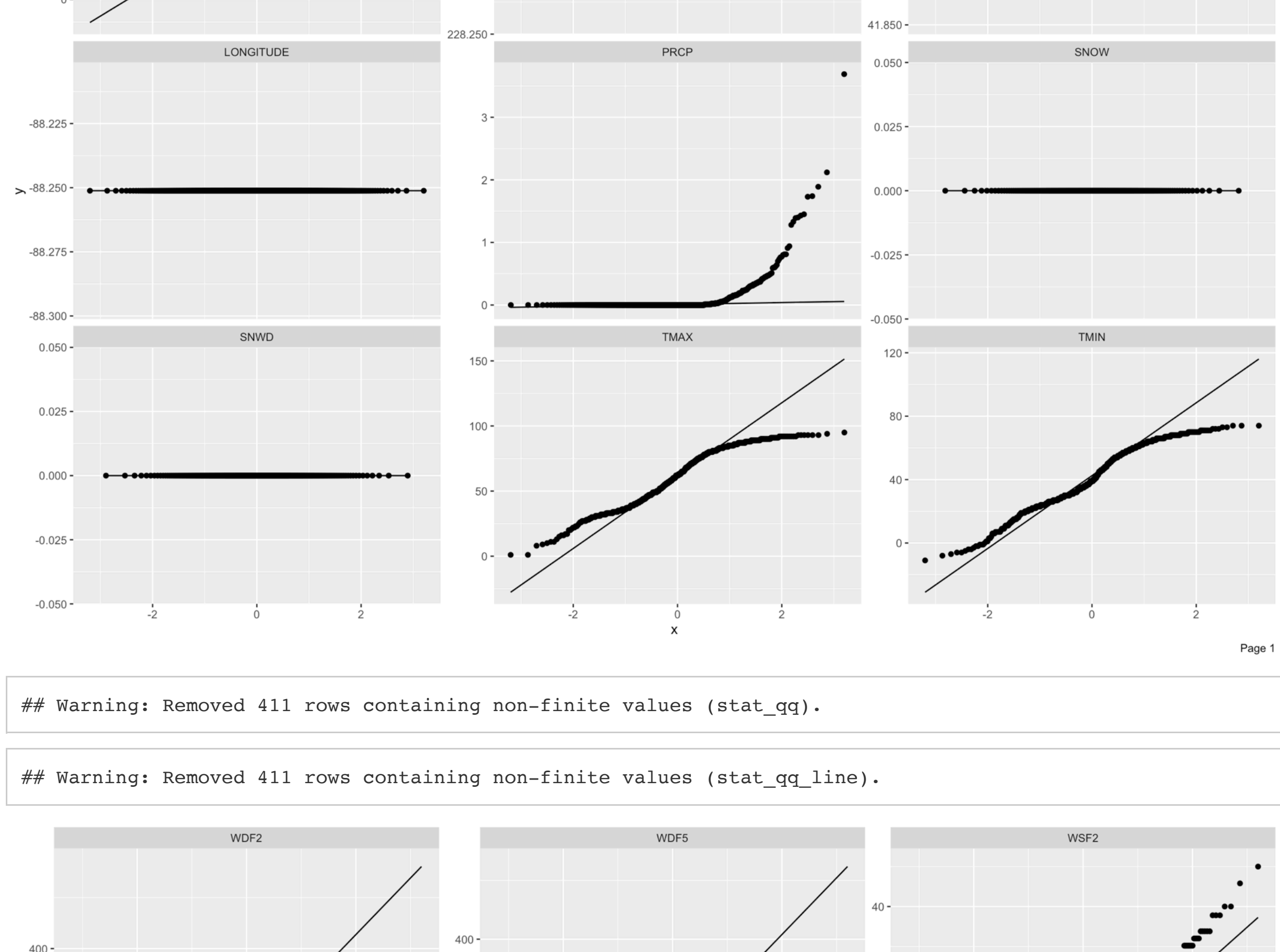
1 columns ignored with more than 50 categories.
DATE: 732 categories



QQ Plot

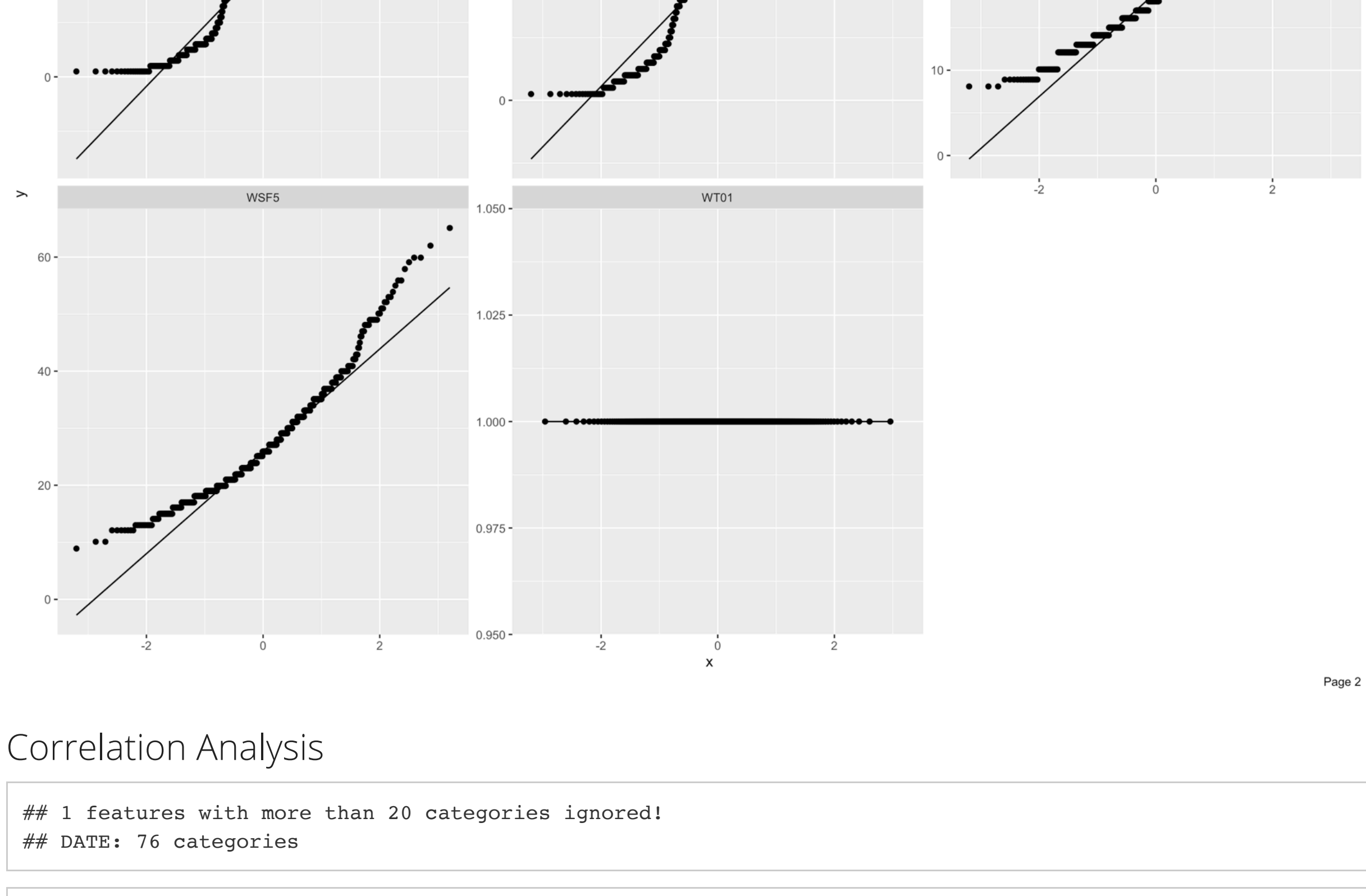
Warning: Removed 1012 rows containing non-finite values (stat_qq).

Warning: Removed 1012 rows containing non-finite values (stat_qq_line).



Warning: Removed 411 rows containing non-finite values (stat_qq).

Warning: Removed 411 rows containing non-finite values (stat_qq_line).

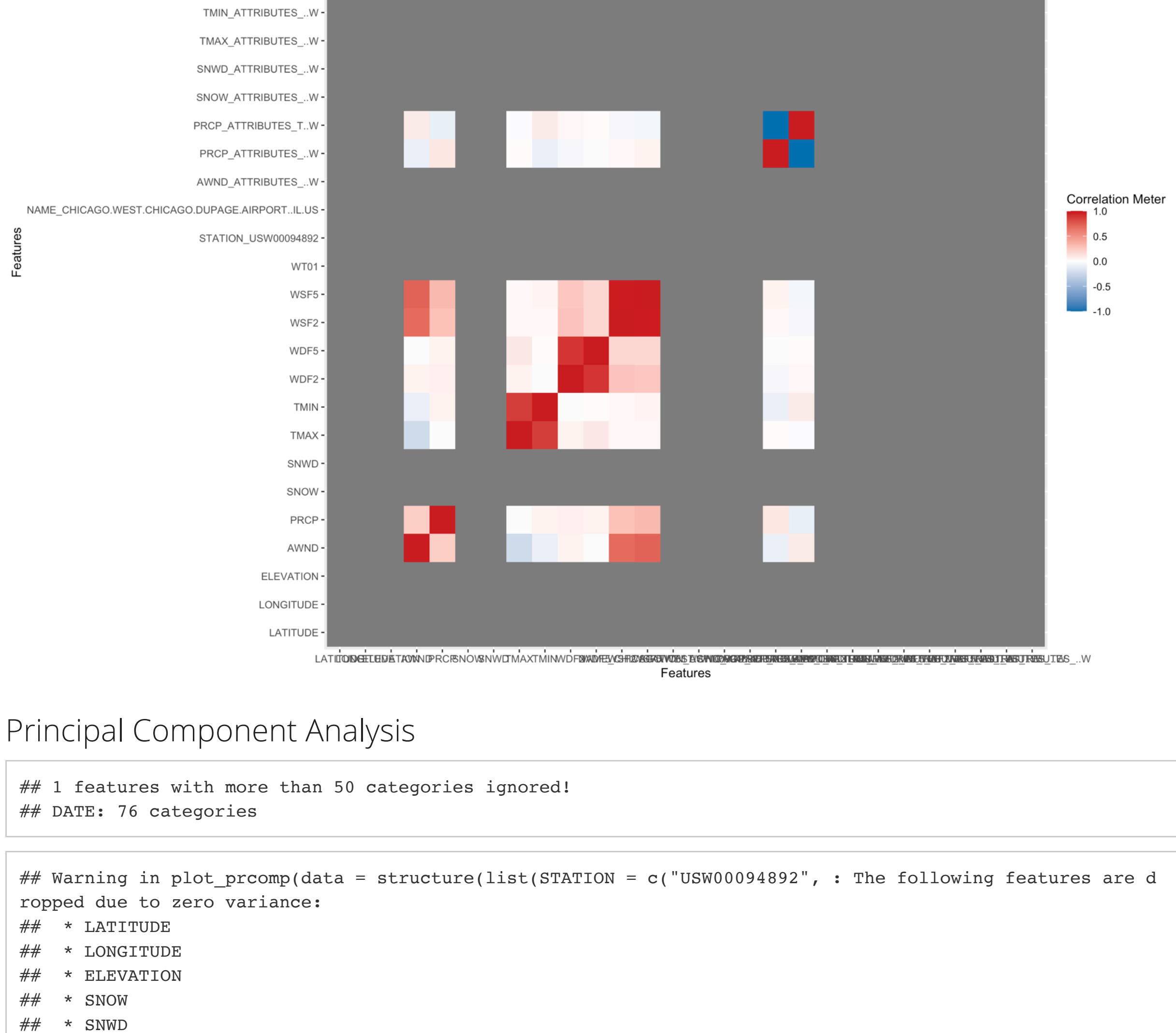


Correlation Analysis

1 features with more than 20 categories ignored!

DATE: 76 categories

Warning in cor(x = structure(list(LATITUDE = c(41.89641, 41.89641, 41.89641, :
the standard deviation is zero



Principal Component Analysis

1 features with more than 50 categories ignored!

DATE: 76 categories

Warning in plot_pcaomp(data = structure(list(STATION = c("USW00094892", ; the following features are dropped due to zero variance:
* LATITUDE
* LONGITUDE
* ELEVATION
* SNOW
* SNWD
* WT01
* STATION_USW00094892
* NAME_CHICAGO_WEST_CHICAGO_DUPAGE_AIRPORT_IL_US
* AWND_ATTRIBUTES_..W
* SNOW_ATTRIBUTES_..W
* SNWD_ATTRIBUTES_..W
* TMAX_ATTRIBUTES_..W
* TMIN_ATTRIBUTES_..W
* WDF2_ATTRIBUTES_..W
* WDF5_ATTRIBUTES_..W
* WSE2_ATTRIBUTES_..W
* WSE5_ATTRIBUTES_..W
* WT01_ATTRIBUTES_..W

% Variance Explained by Principal Components
(Note: Labels indicate cumulative % explained variance)

