Data Profiling Report

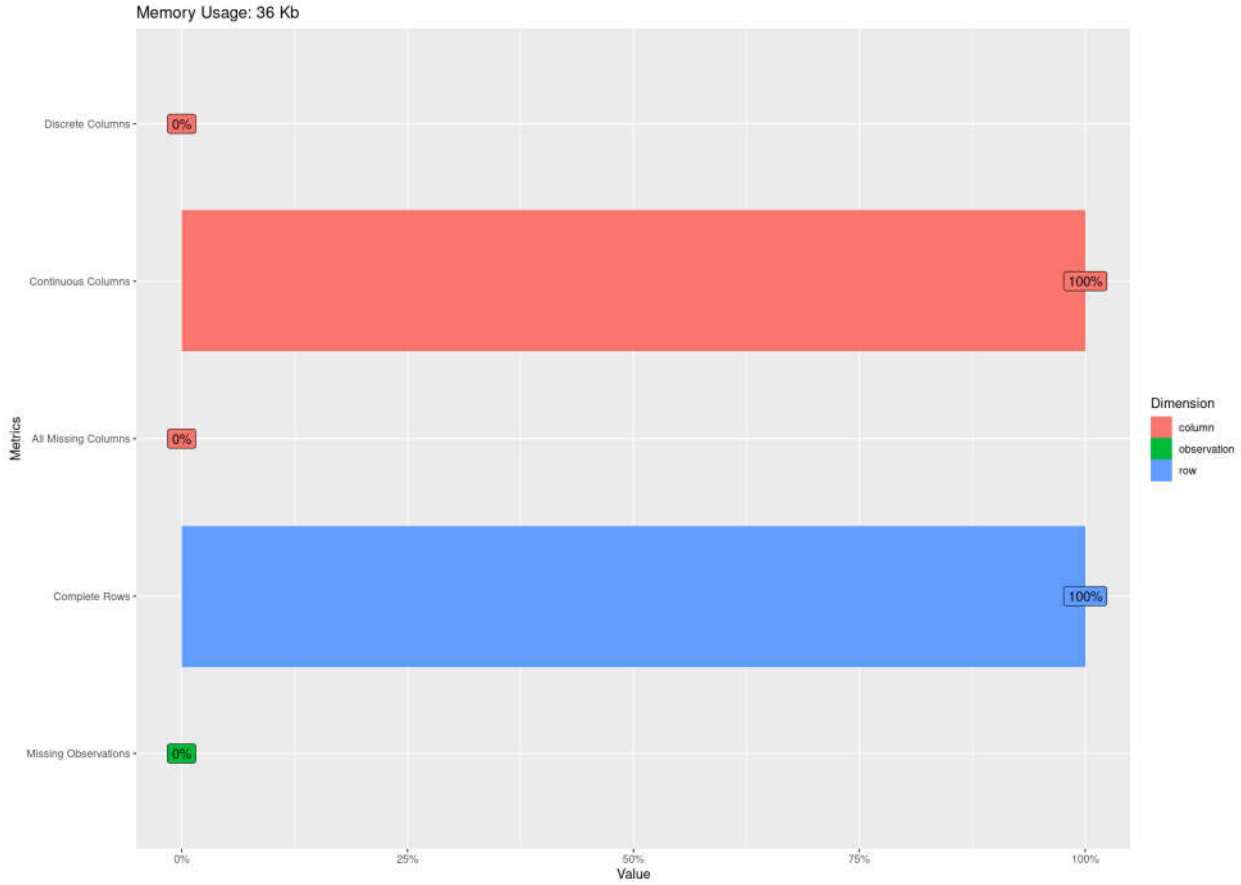
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

Basic Statistics

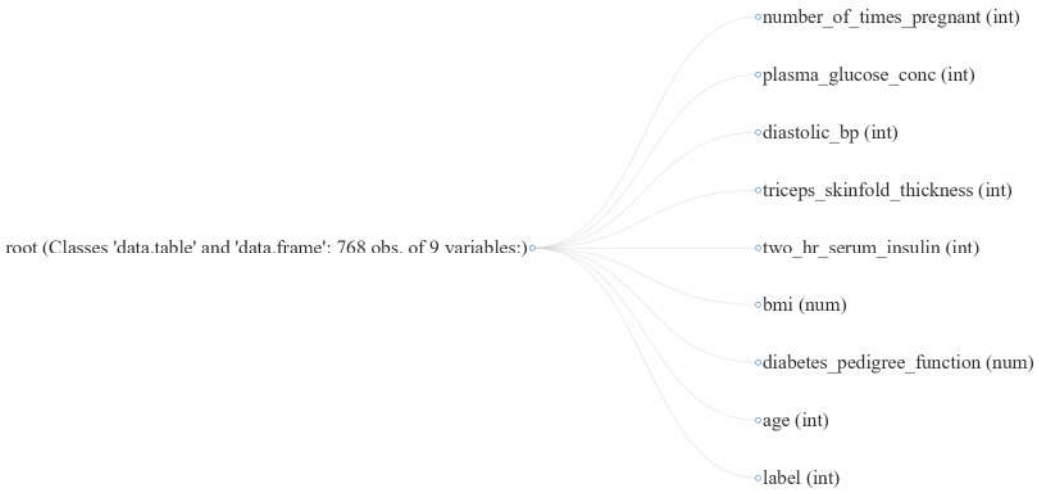
Raw Counts

Name	Value
Rows	768
Columns	9
Discrete columns	0
Continuous columns	9
All missing columns	0
Missing observations	0
Complete Rows	768
Total observations	6,912
Memory allocation	36 Kb

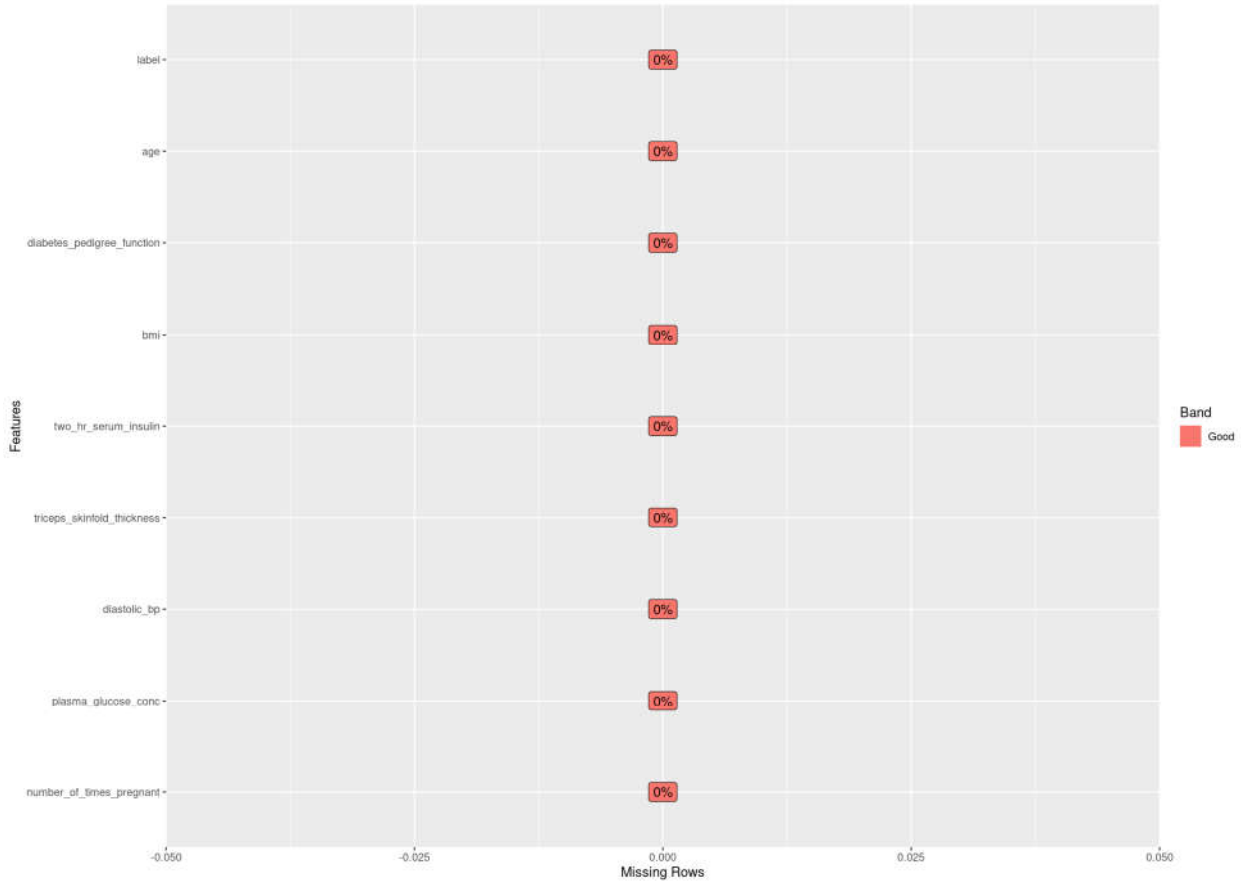
Percentages



Data Structure

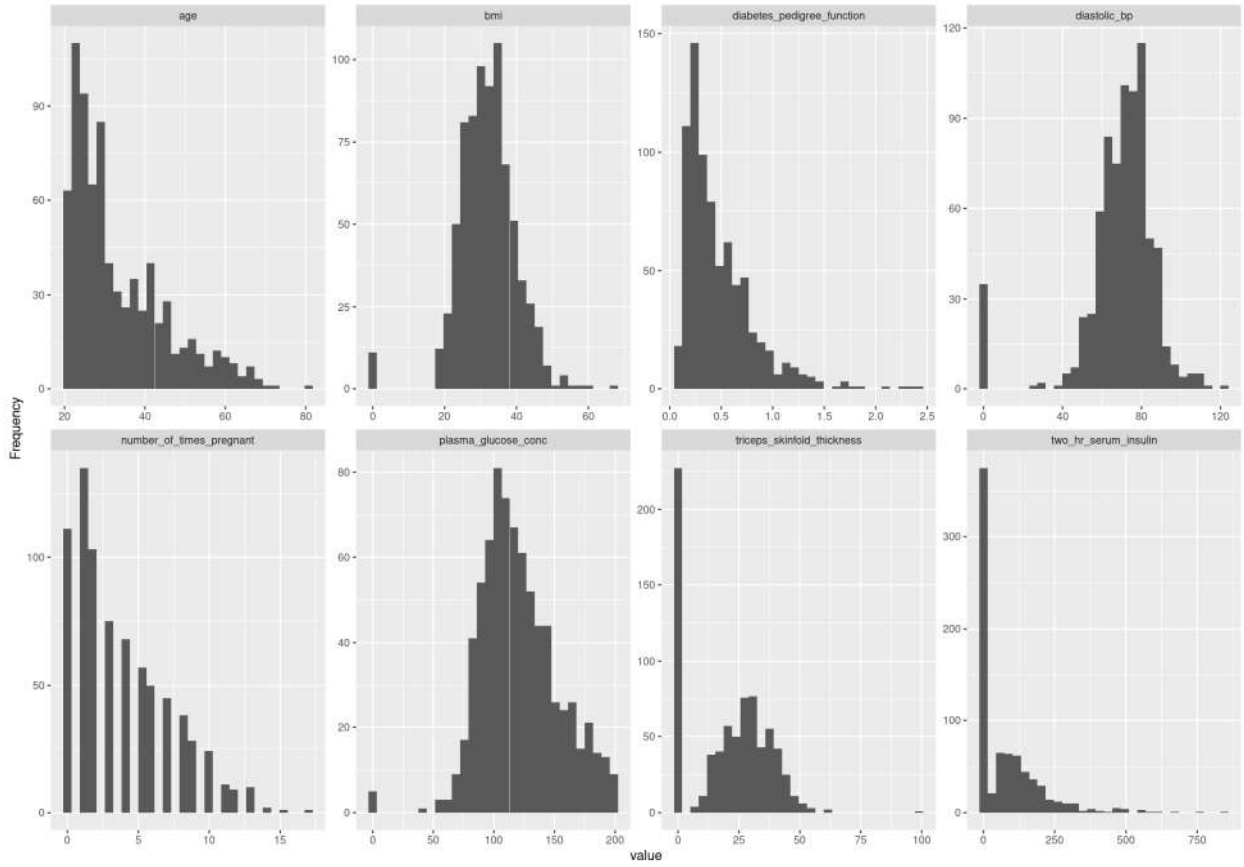


Missing Data Profile

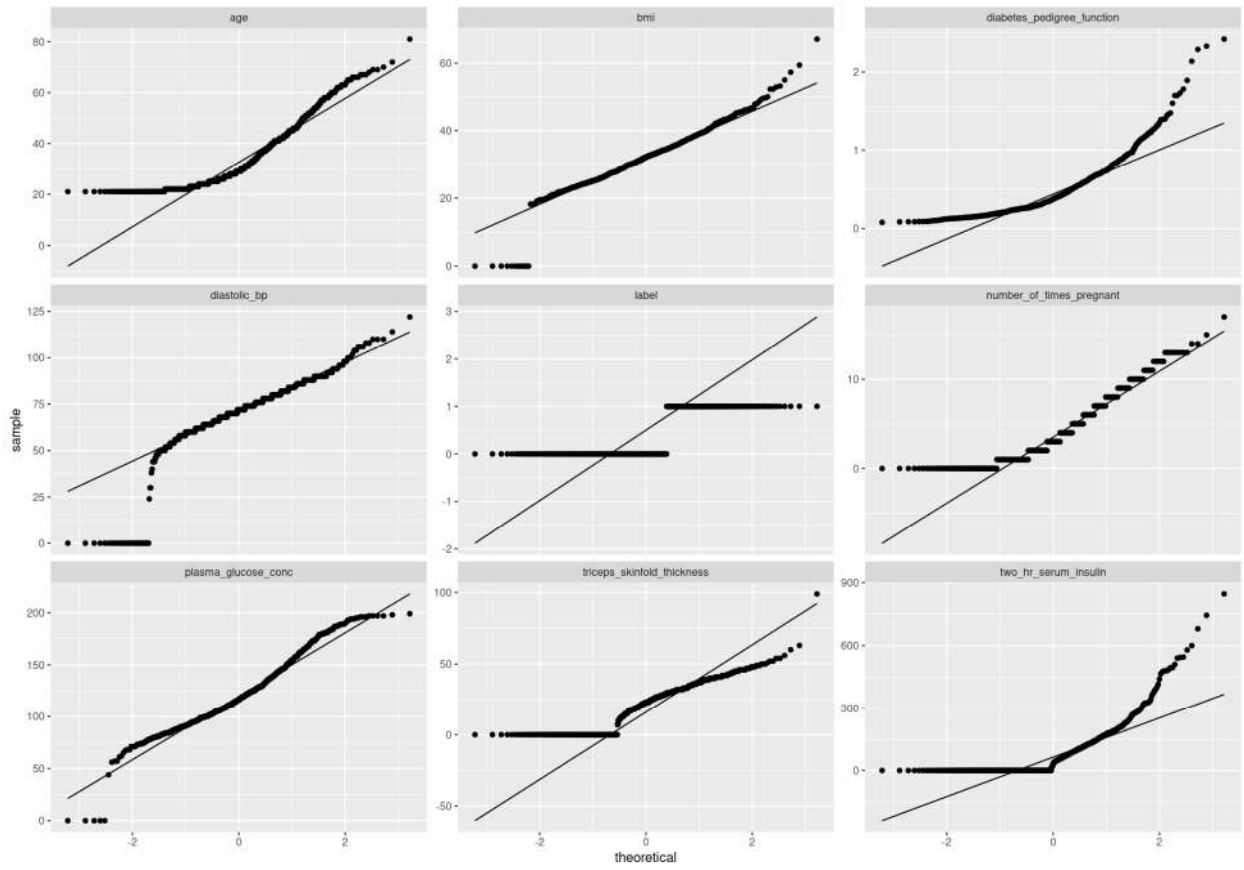


Univariate Distribution

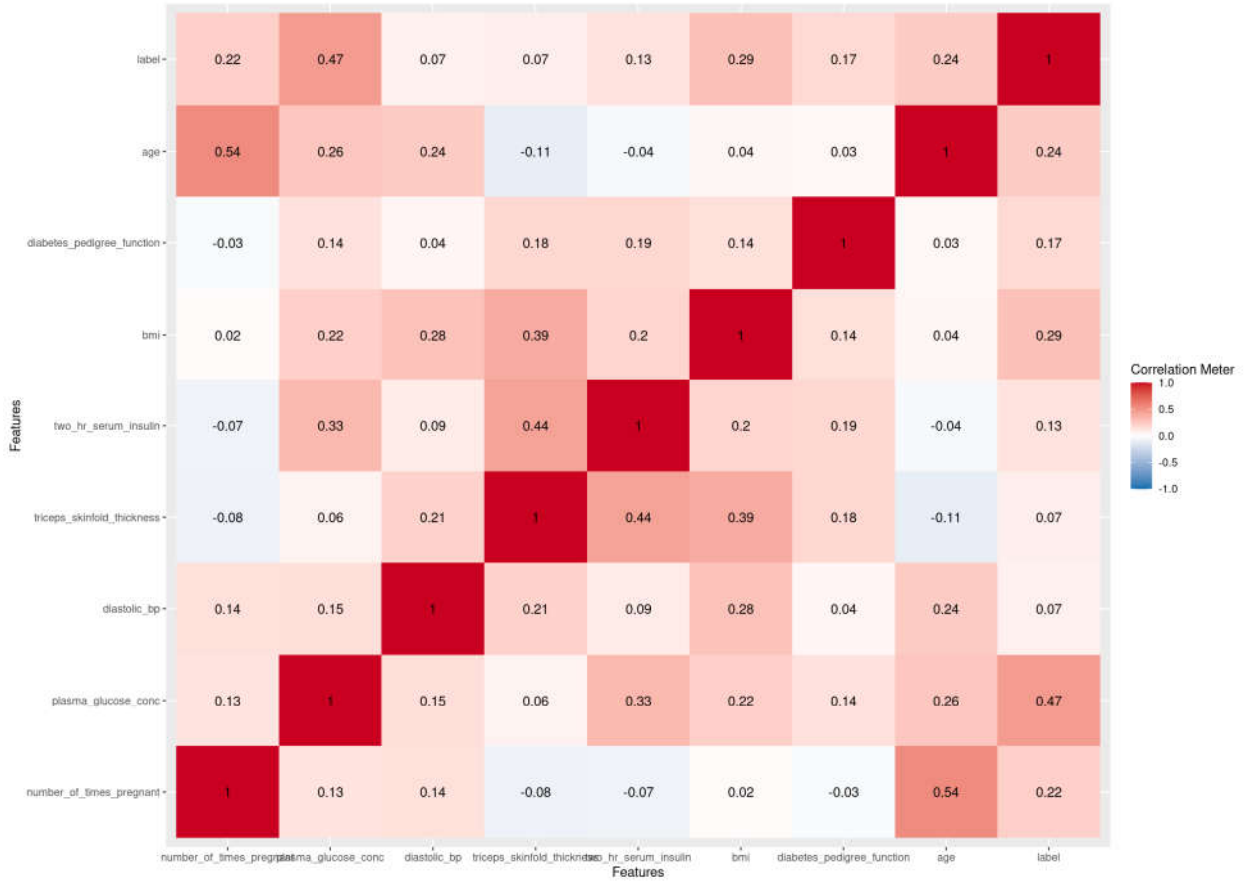
Histogram



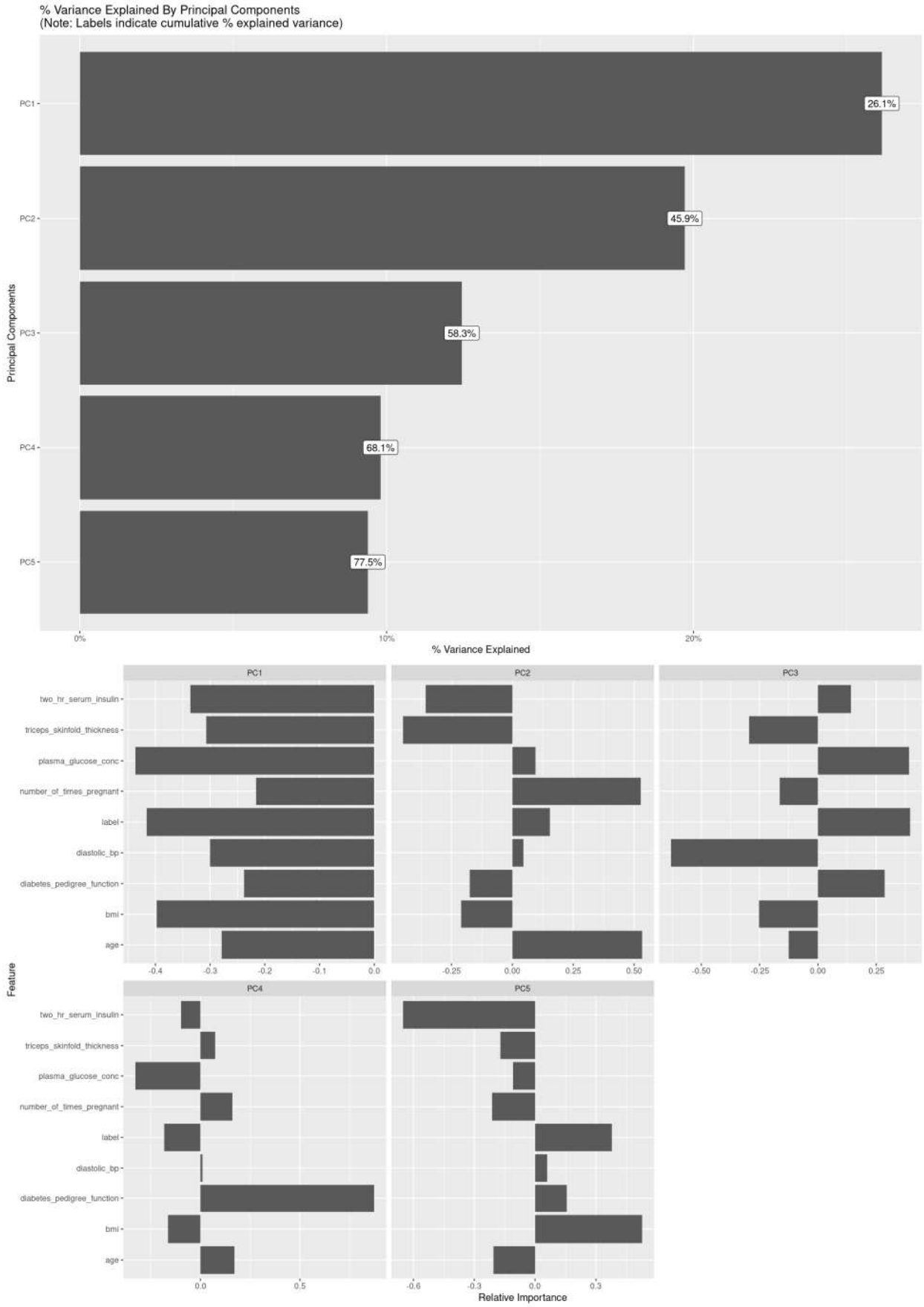
QQ Plot



Correlation Analysis



Principal Component Analysis


% Variance Explained By Principal Components
(Note: Labels indicate cumulative % explained variance)
Principal Components
PC1
PC2
PC3
PC4
PC5
26.1%
45.9%
58.3%
68.1%
77.5%
0%
10%
20%
% Variance Explained
PC1
PC2
PC3
PC4
PC5
two_hr_serum_insulin
triceps_skinfold_thickness
plasma_glucose_conc
number_of_times_pregnant
label
diastolic_bp
diabetes_pedigree_function
bmi
age
Feature
Relative Importance