Exploratory Data Analysis Report for the Boston Dataset with a Focus on the Per Capita Crime Rate

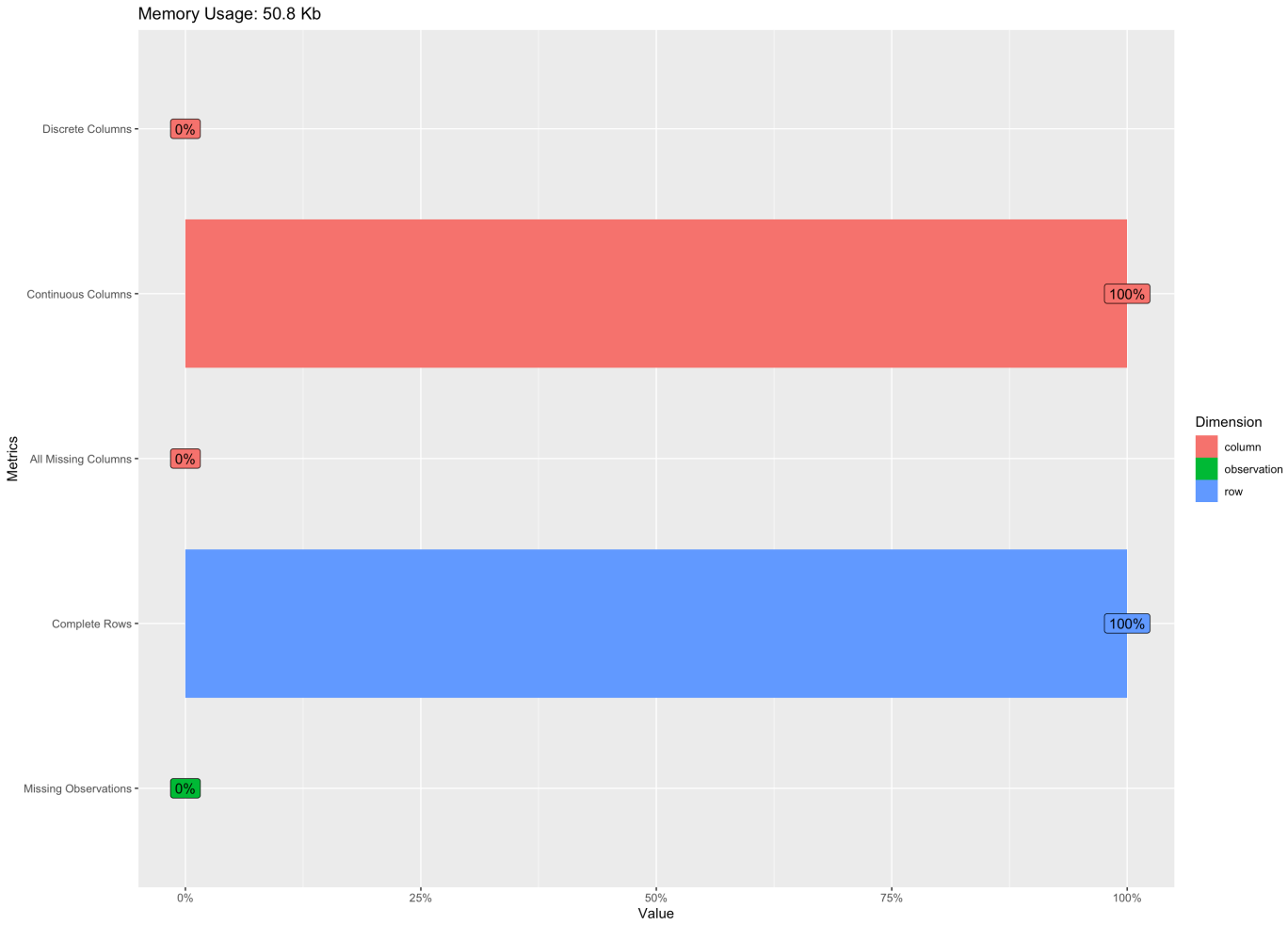
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - QQ Plot
  - QQ Plot (by crim)
- Correlation Analysis
- Principal Component Analysis
- Bivariate Distribution
  - Boxplot (by crim)
  - Scatterplot (by crim)

Basic Statistics

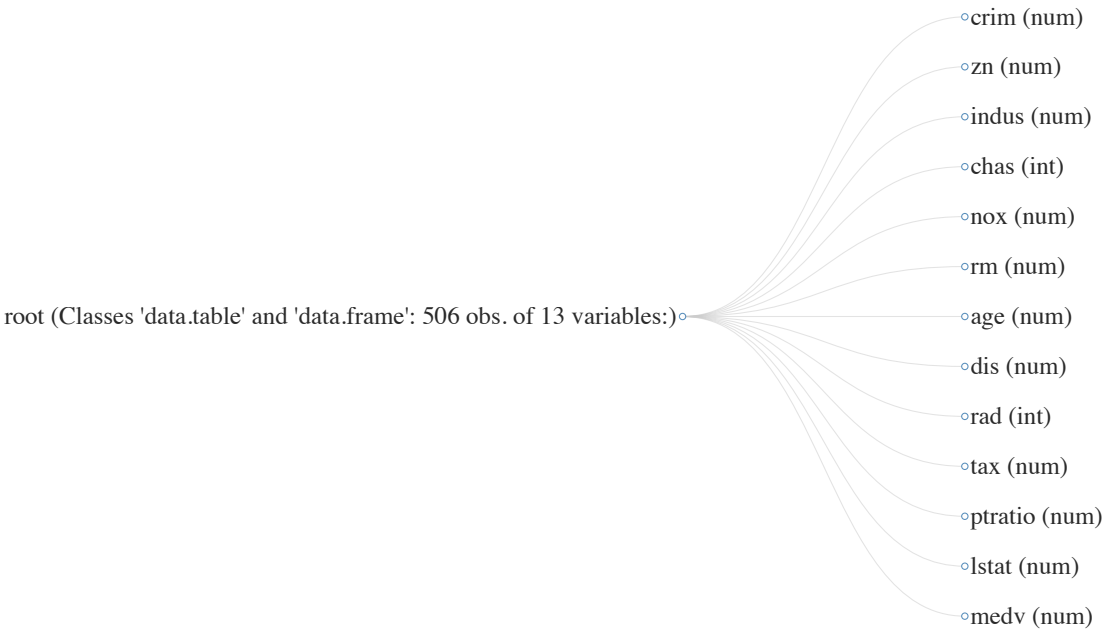
Raw Counts

Name	Value
Rows	506
Columns	13
Discrete columns	0
Continuous columns	13
All missing columns	0
Missing observations	0
Complete Rows	506
Total observations	6,578
Memory allocation	50.8 Kb

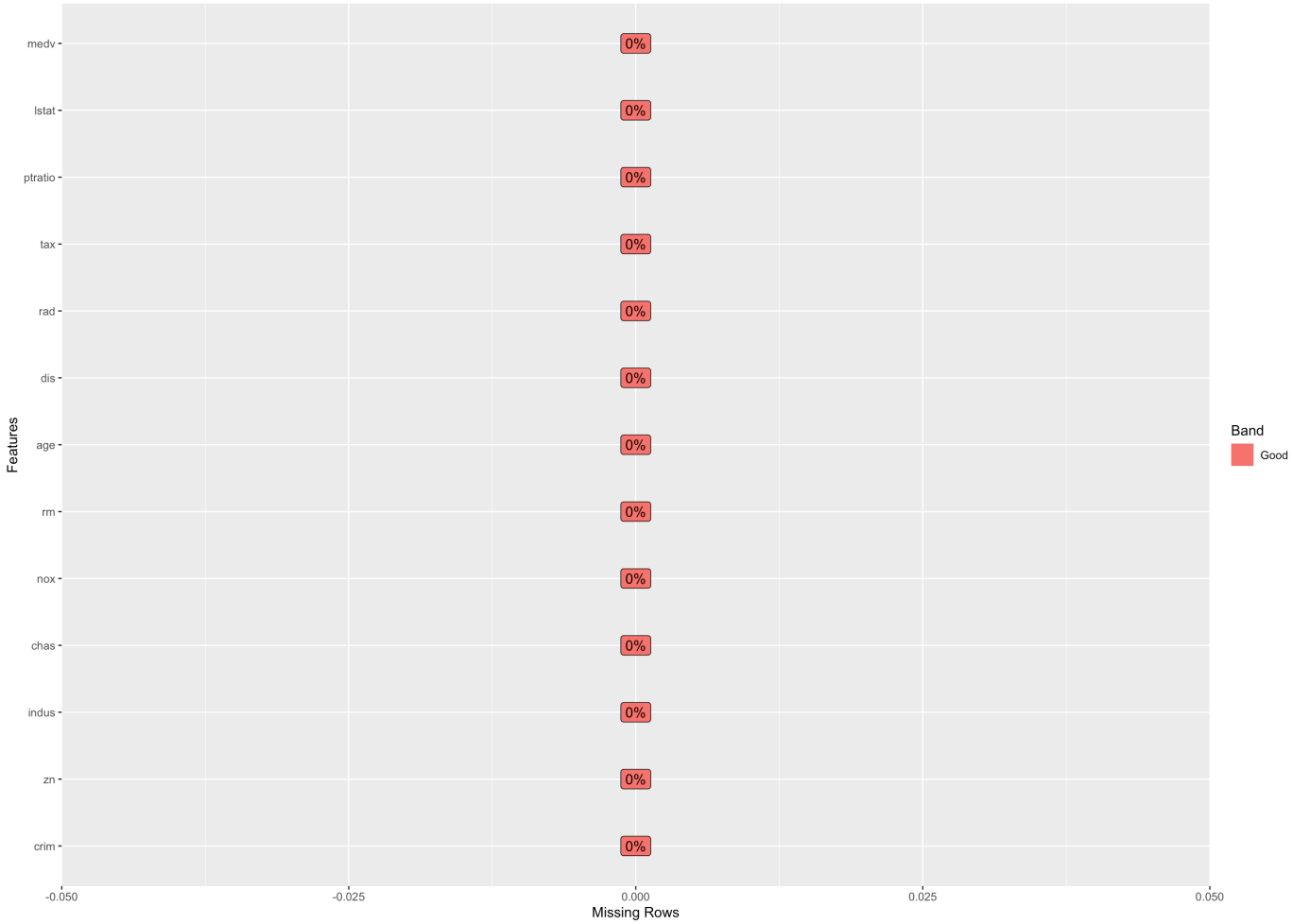
Percentages



Data Structure

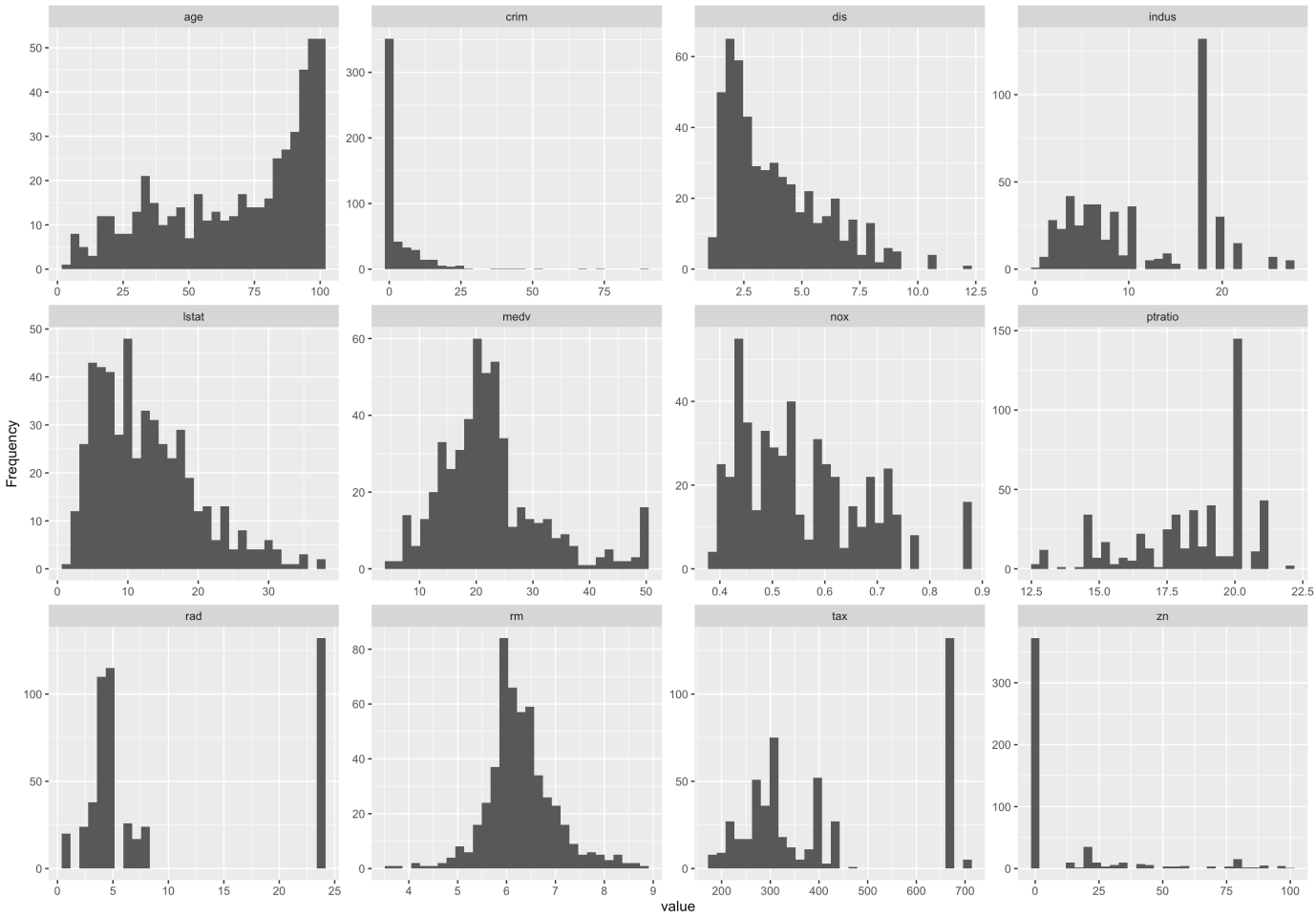


Missing Data Profile

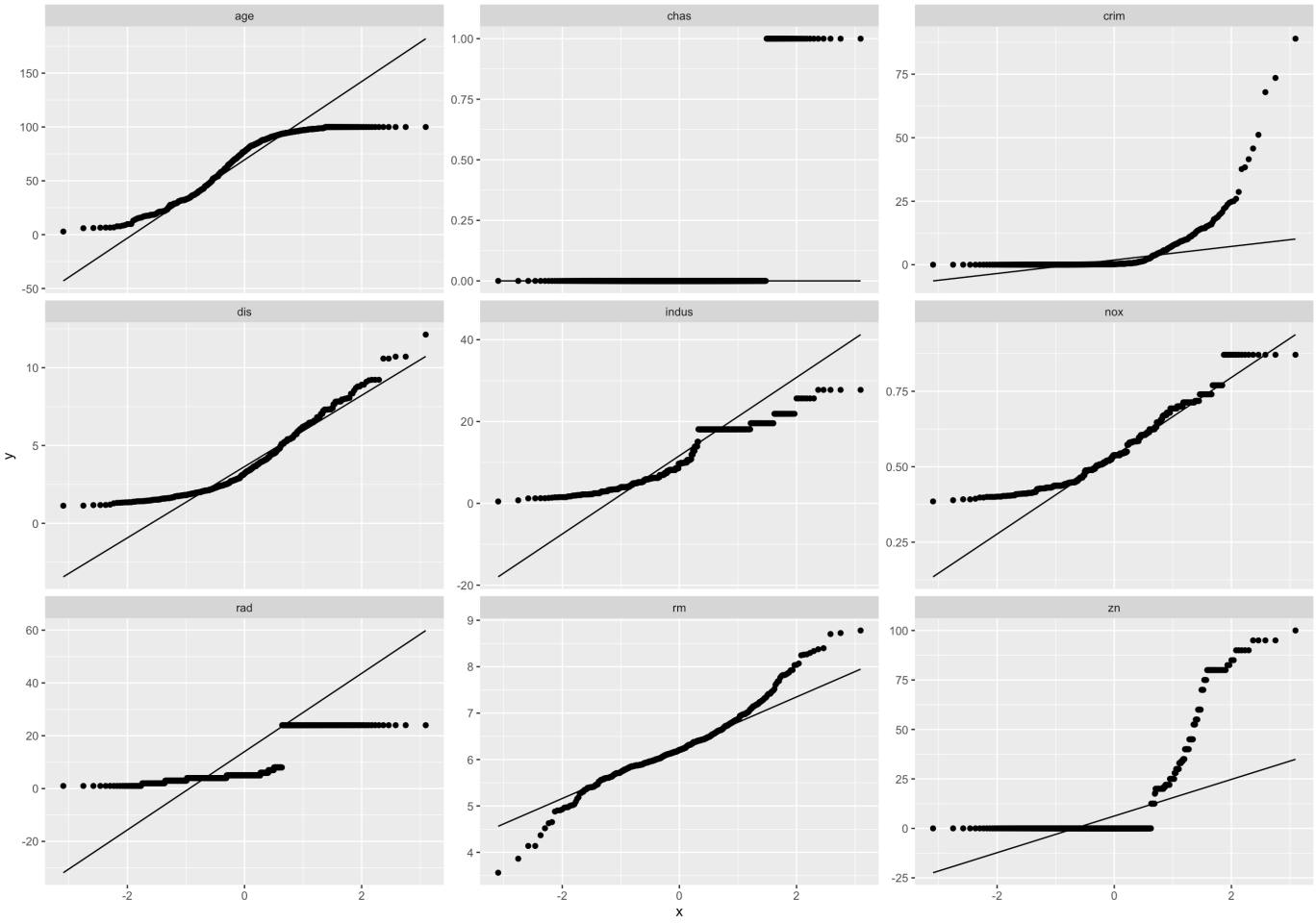


Univariate Distribution

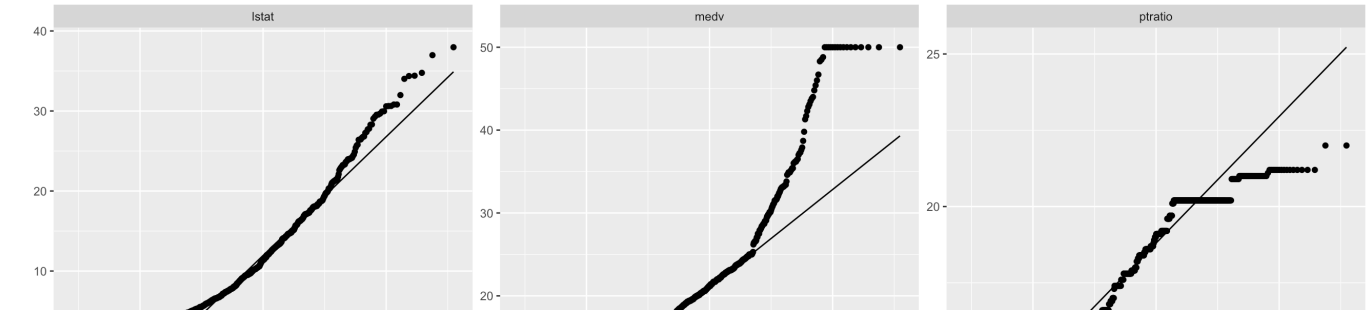
Histogram

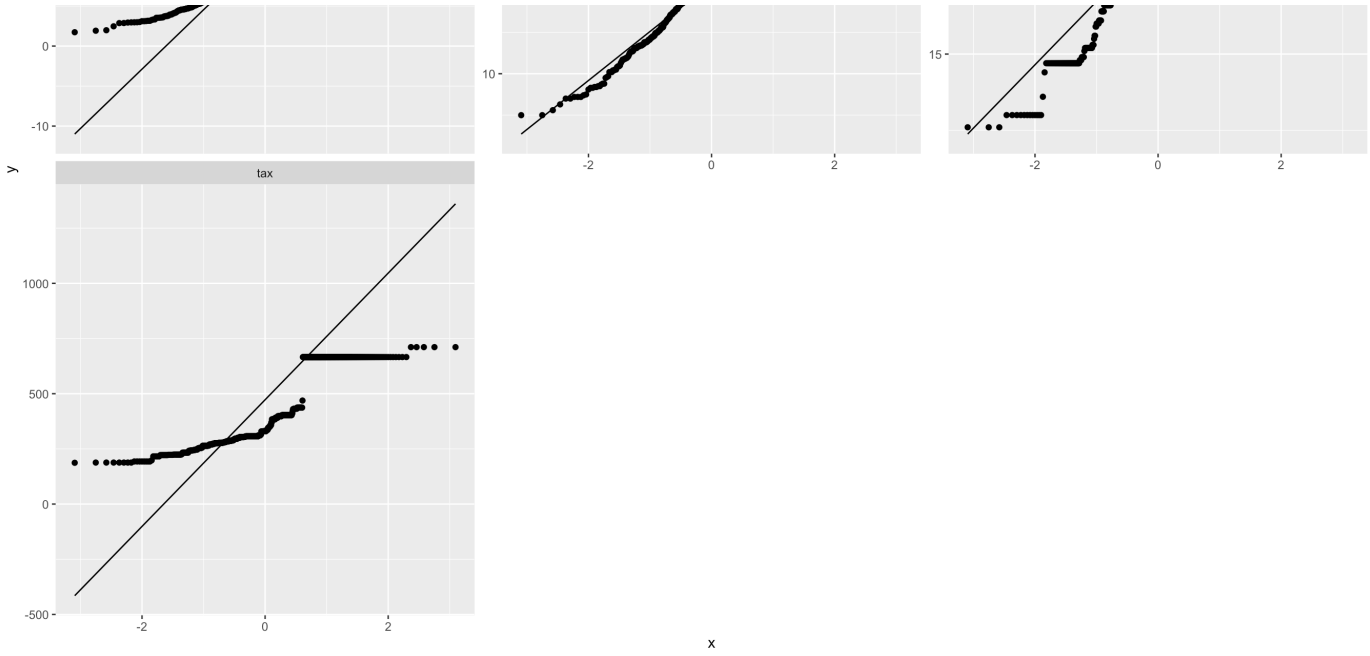

age
crim
dis
indus
lstat
medv
nox
ptratio
rad
rm
tax
zn
Frequency
value

QQ Plot



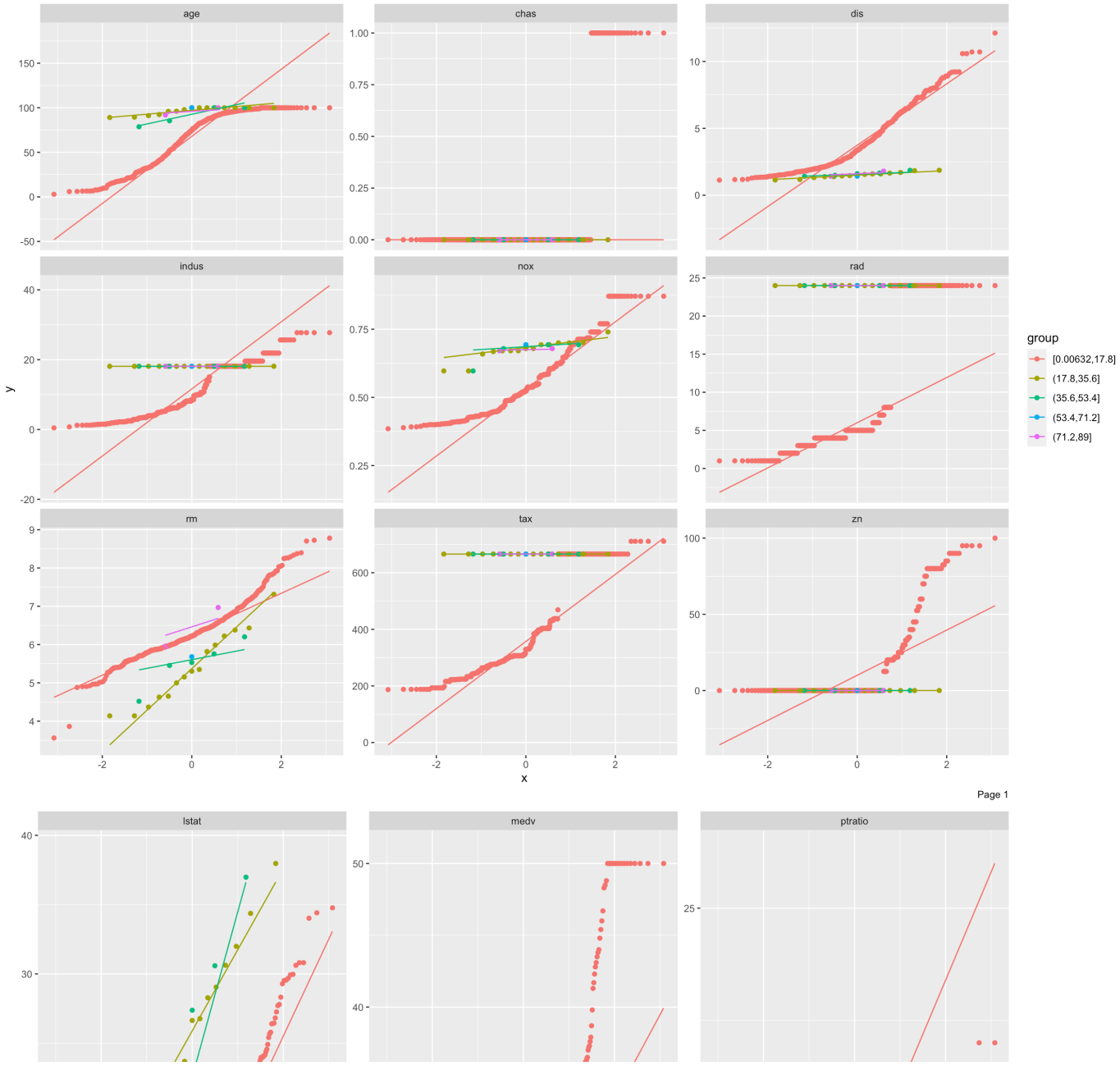
Page 1

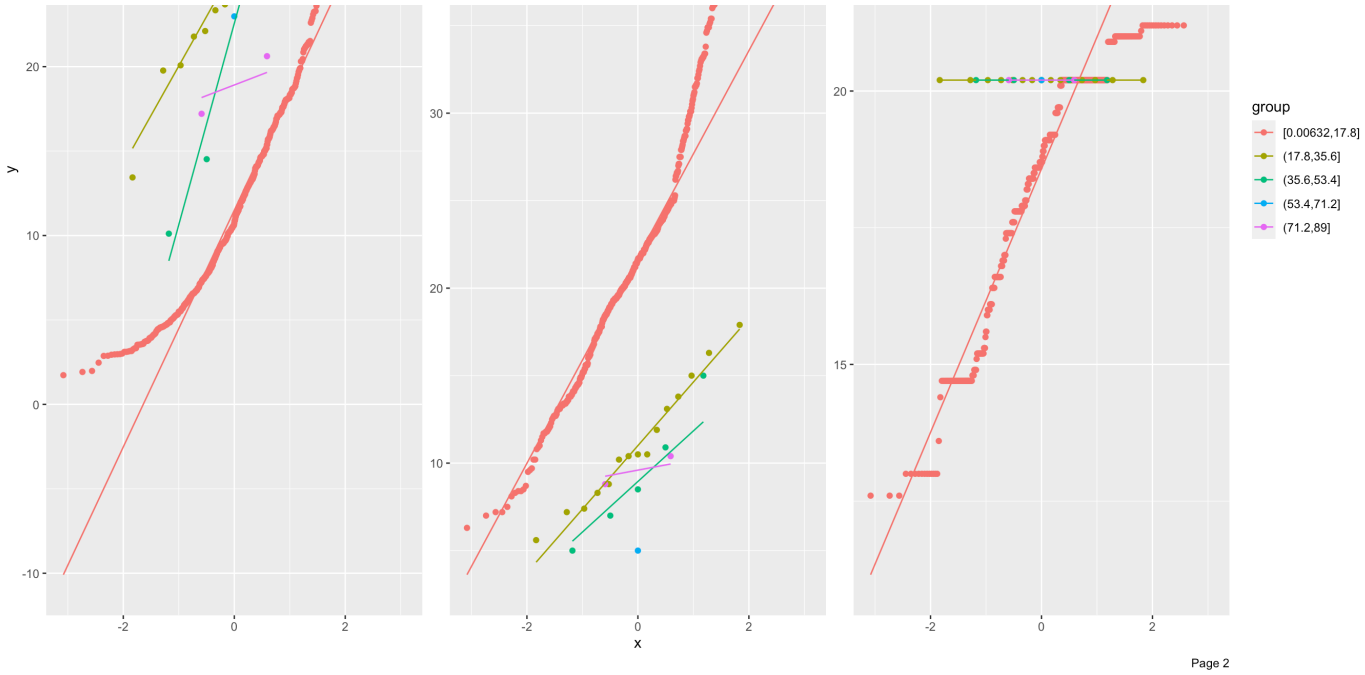




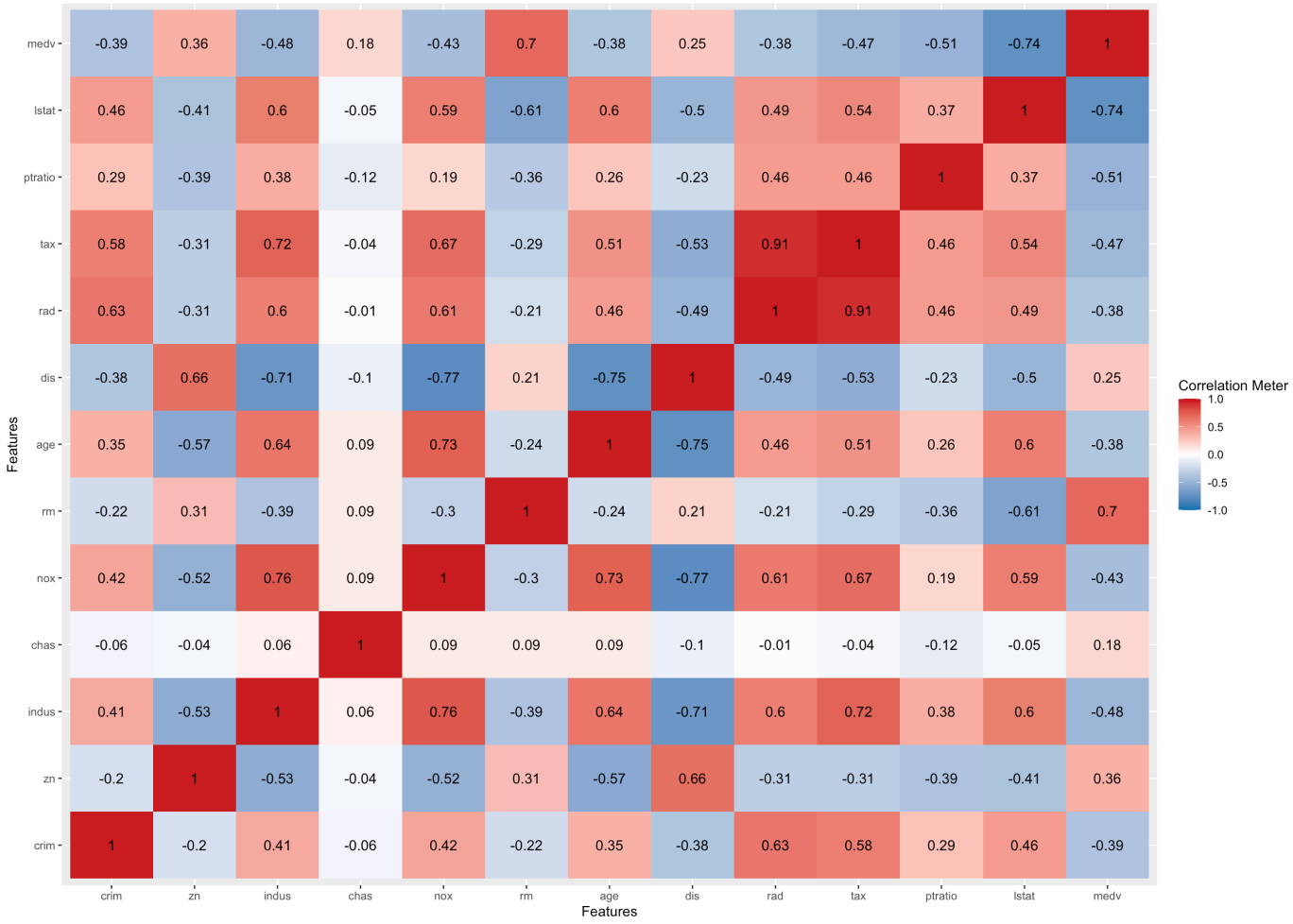
Page 2

QQ Plot (by crim)

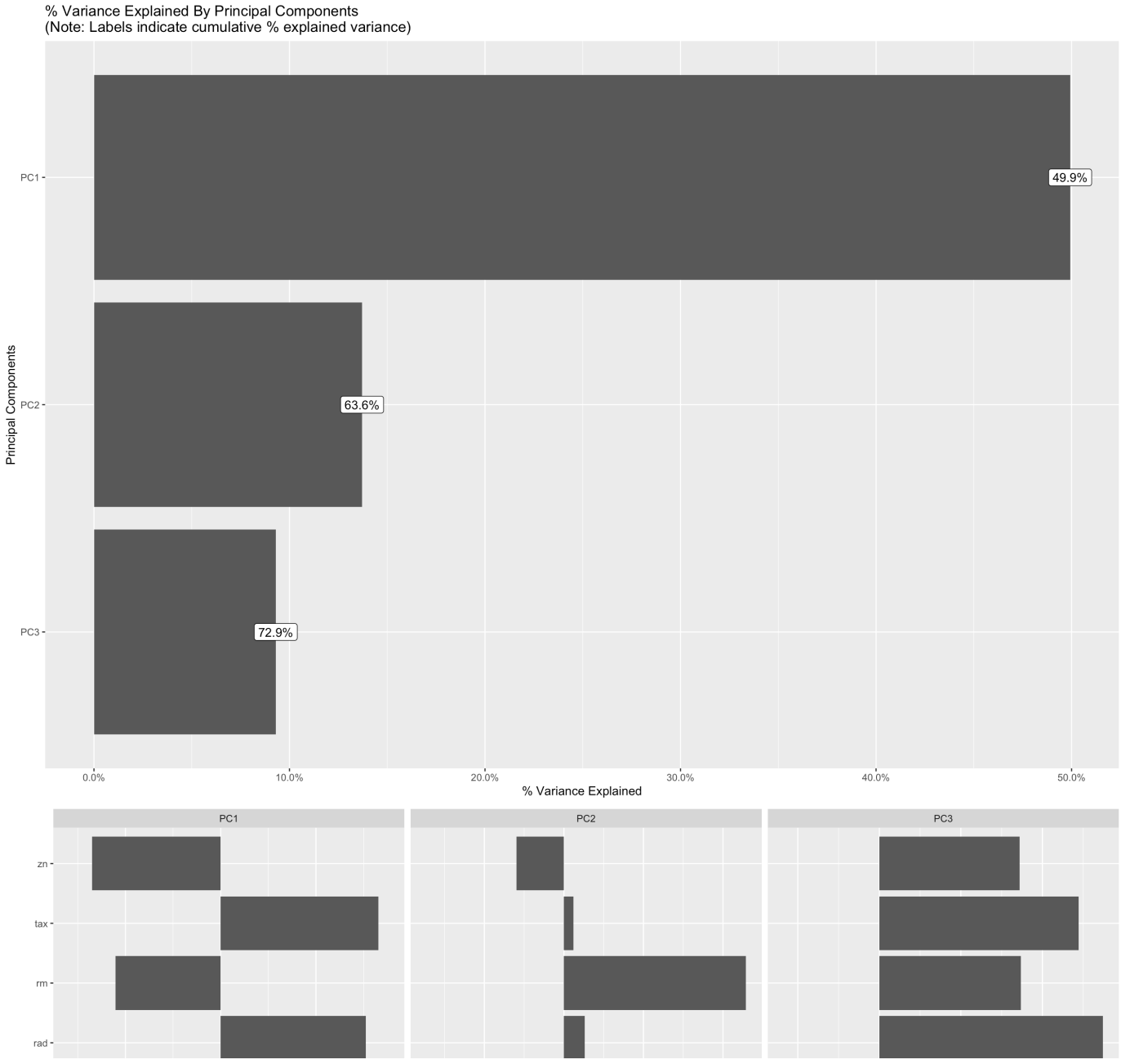

age
chas
dis
indus
nox
rad
rm
tax
zn
lstat
medv
ptratio
x
y
group
[0.00632,17.8]
(17.8,35.6]
(35.6,53.4]
(53.4,71.2]
(71.2,89]
Page 1

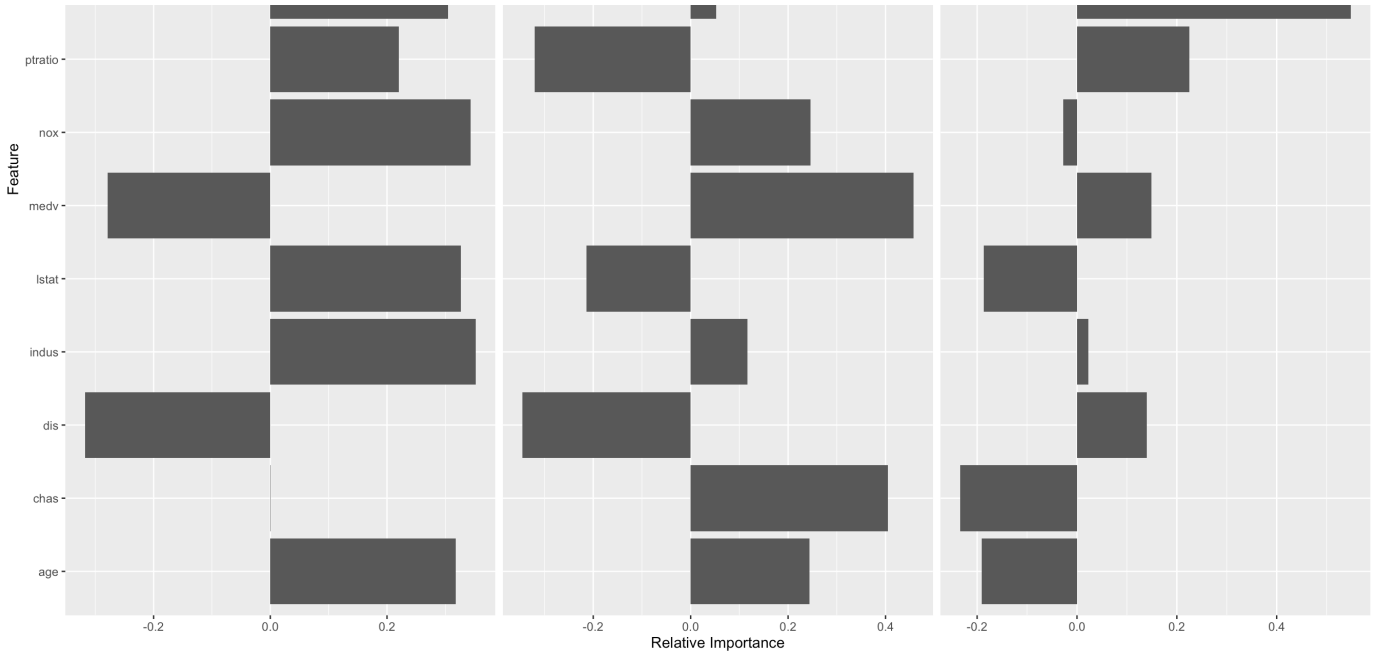


Correlation Analysis



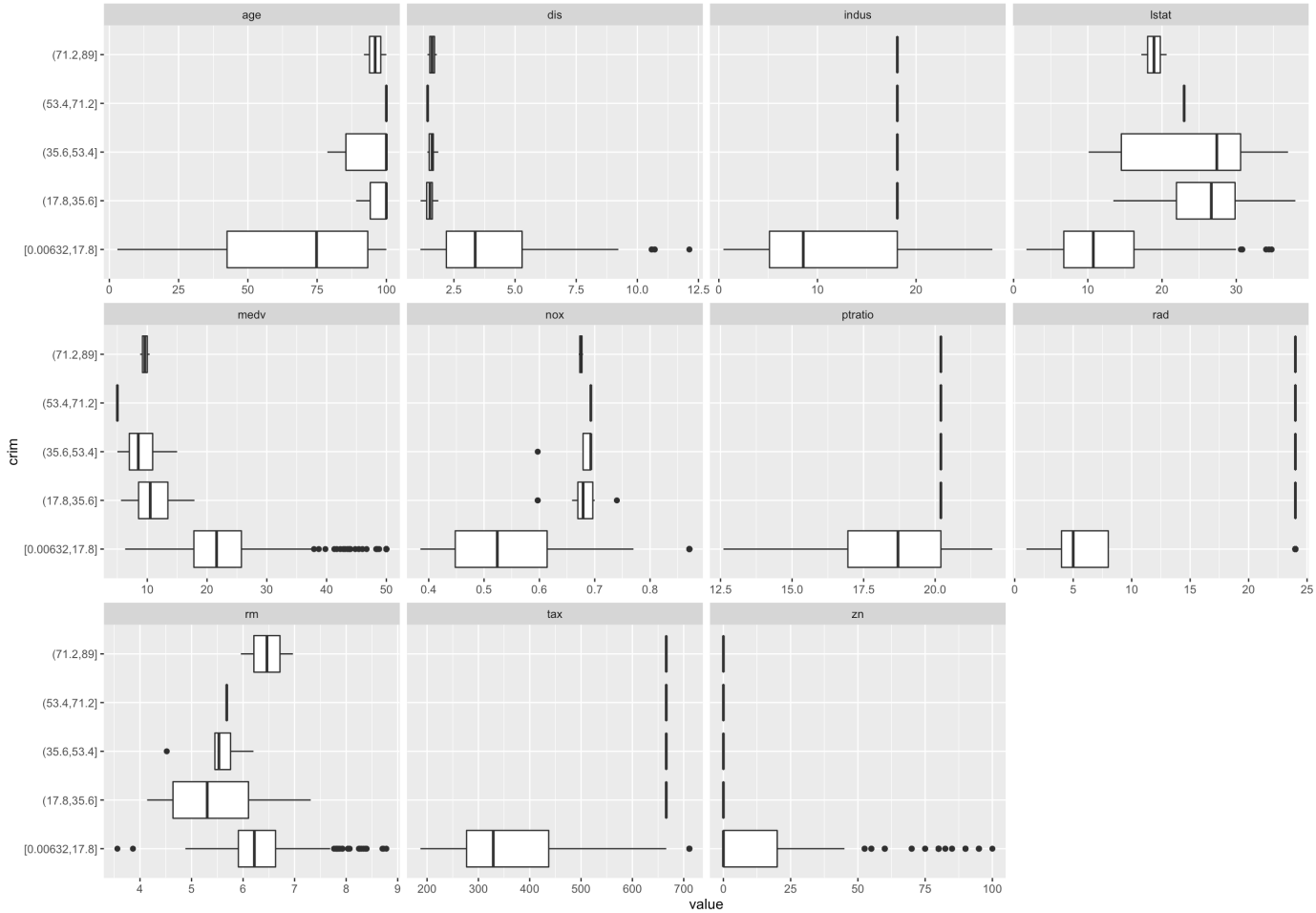
Principal Component Analysis


% Variance Explained By Principal Components
(Note: Labels indicate cumulative % explained variance)
Principal Components
PC1
PC2
PC3
49.9%
63.6%
72.9%
0.0%
10.0%
20.0%
30.0%
40.0%
50.0%
% Variance Explained
PC1
PC2
PC3
zn
tax
rm
rad

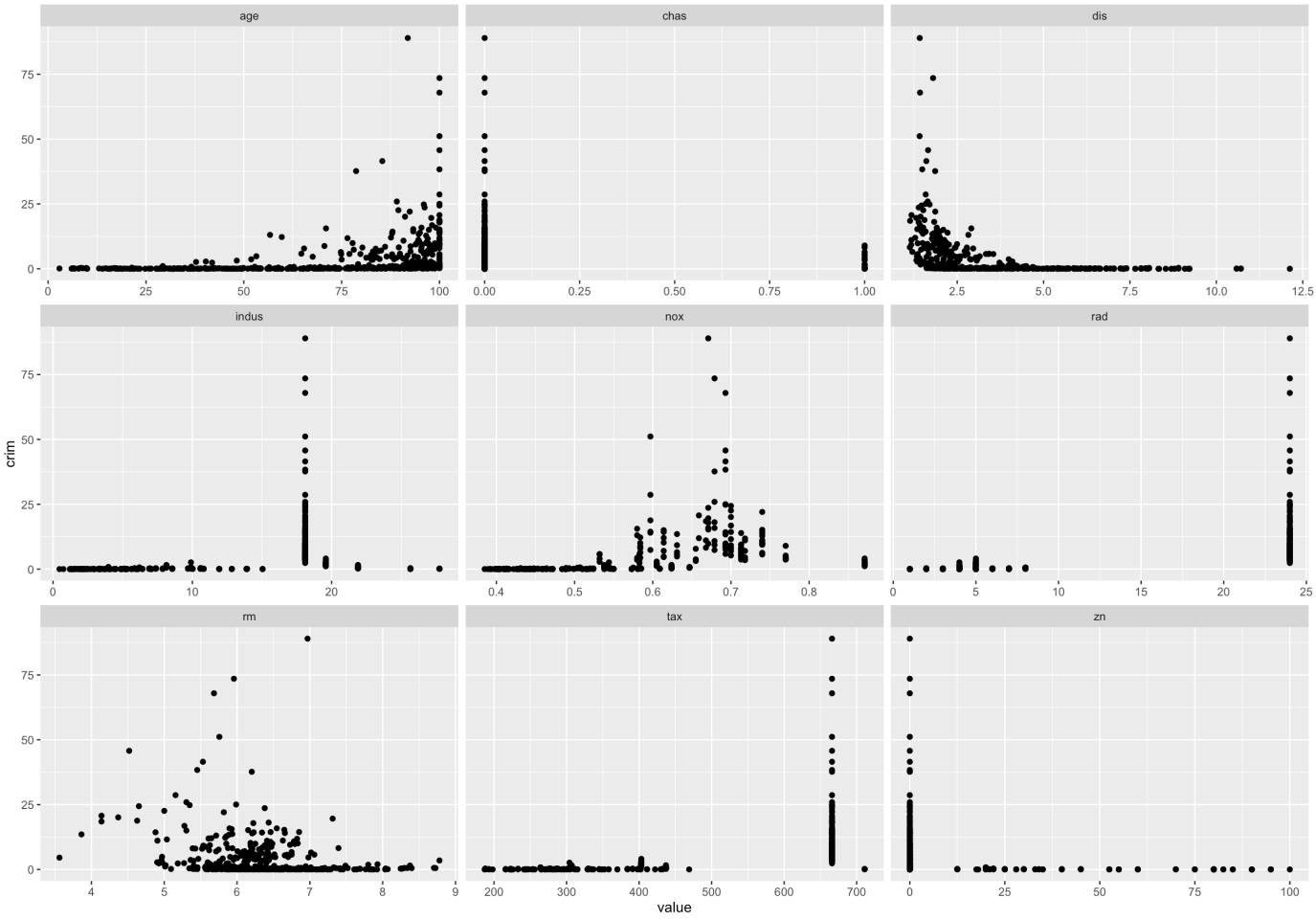


Bivariate Distribution

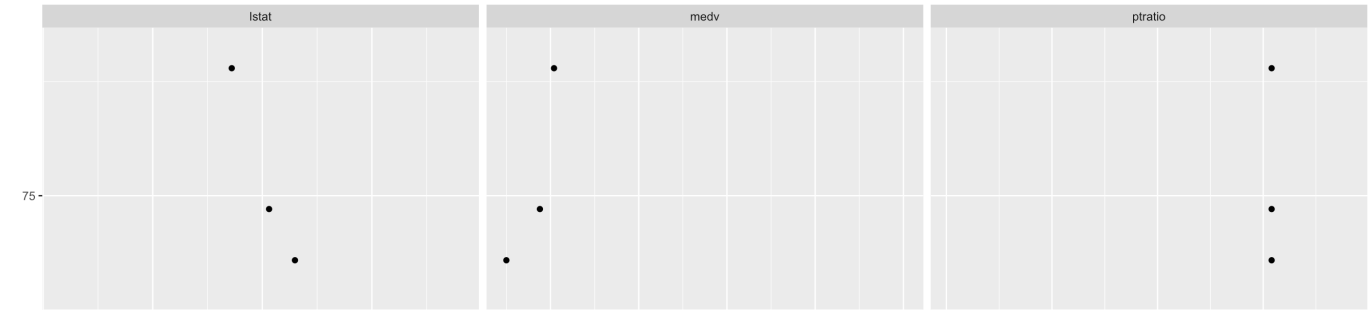
Boxplot (by crim)

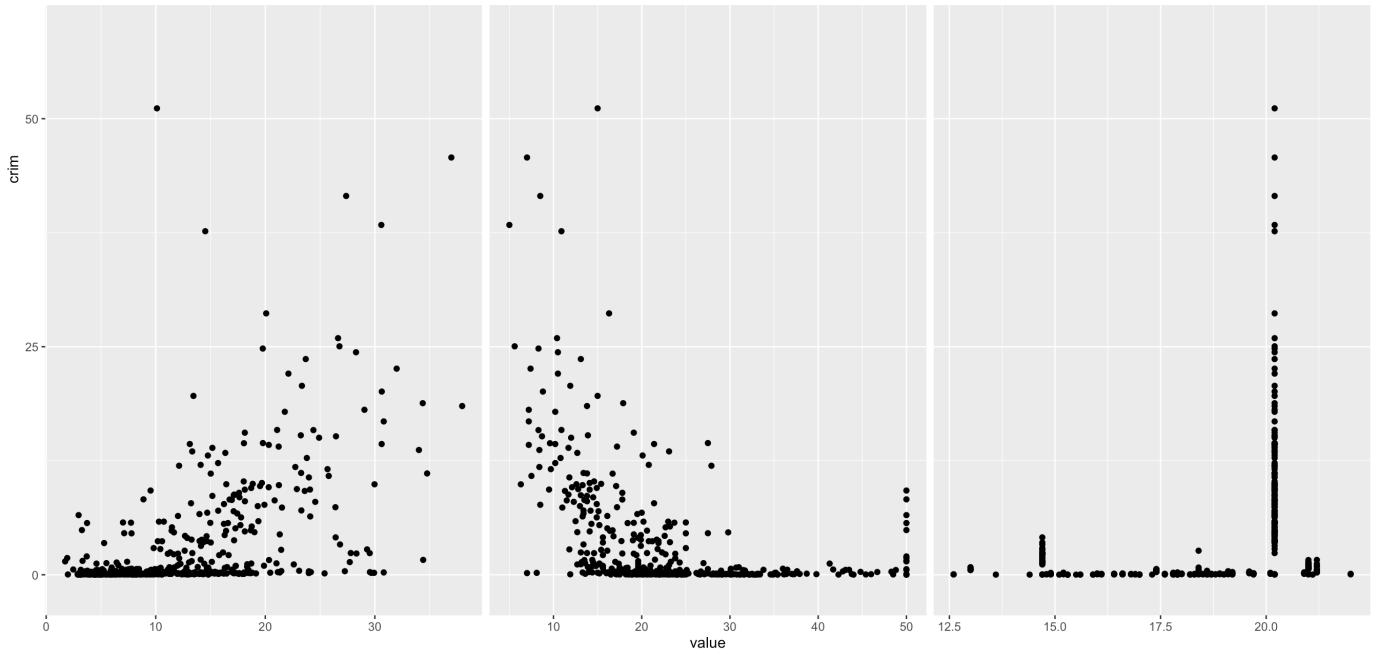


Scatterplot (by crim)



Page 1





Page 2