Data Profiling Report

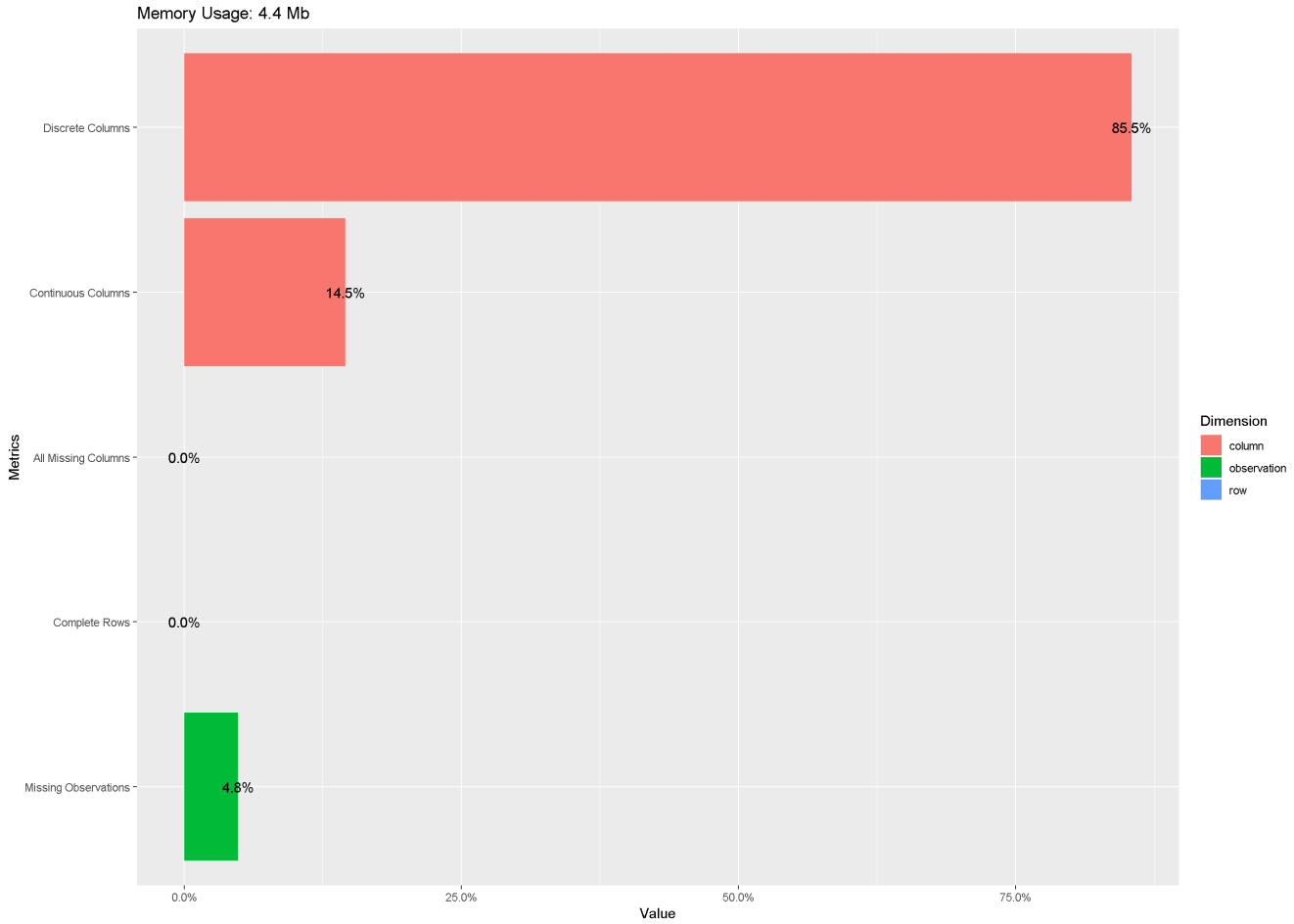
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (by frequency)
  - QQ Plot
- Correlation Analysis
- Principle Component Analysis

Basic Statistics

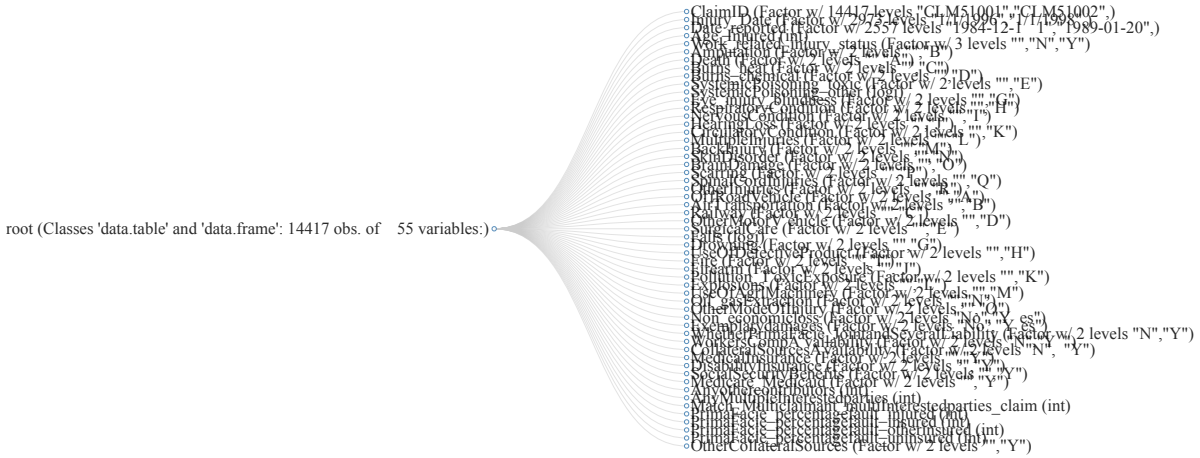
Raw Counts

Name	Value
Rows	14,417
Columns	55
Discrete columns	47
Continuous columns	8
All missing columns	0
Missing observations	38,454
Complete Rows	0
Total observations	792,935
Memory allocation	4.4 Mb

Percentages



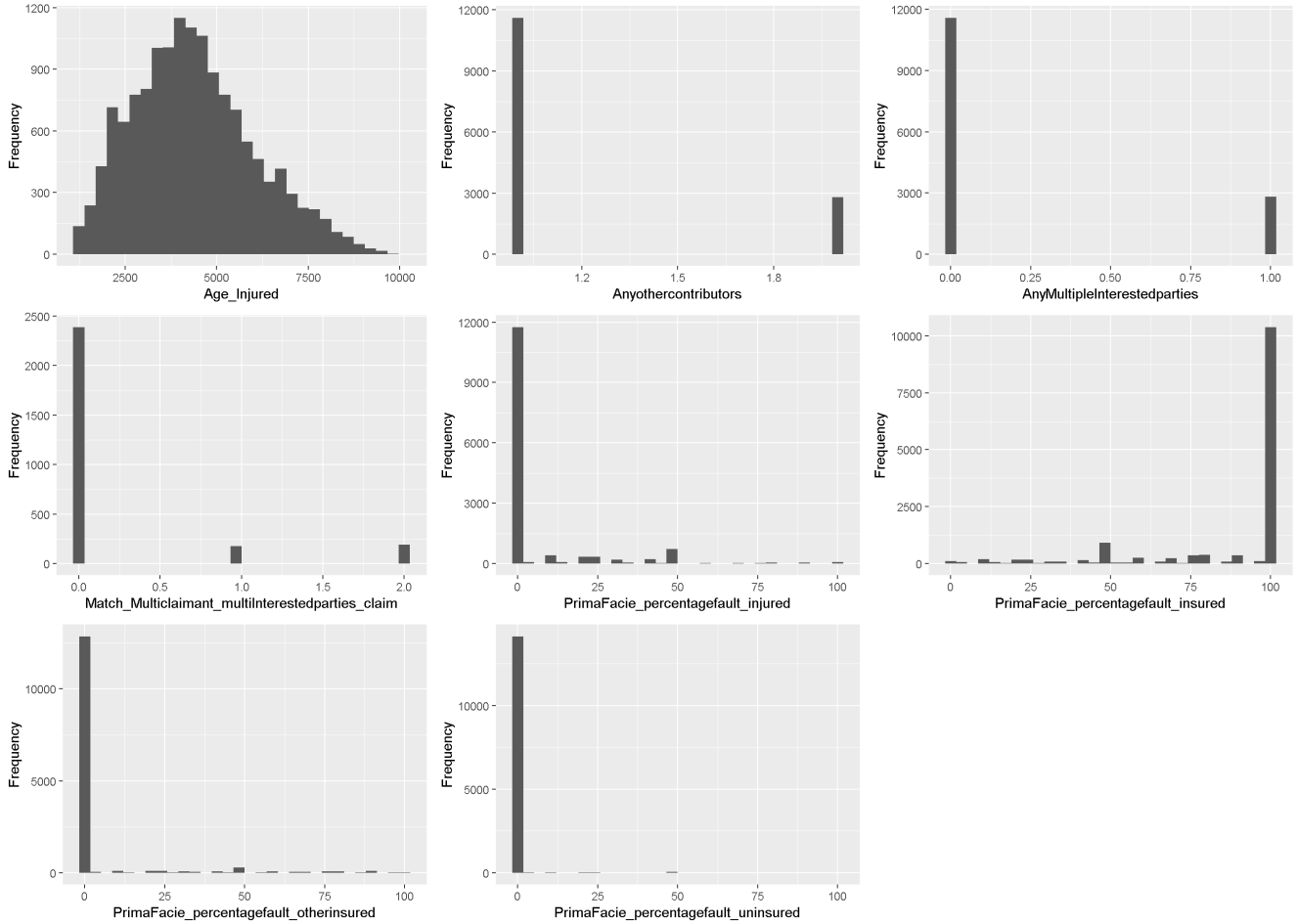
Data Structure



Missing Data Profile

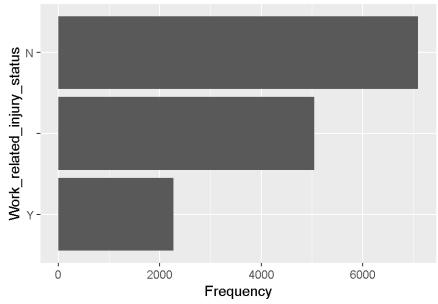
Univariate Distribution

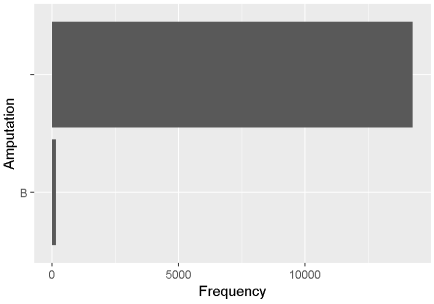
Histogram

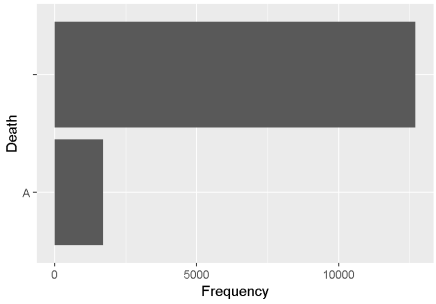


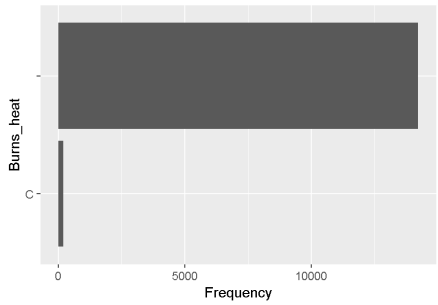
Bar Chart (by frequency)

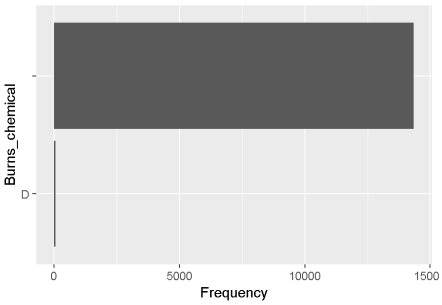
```
## 3 columns ignored with more than 50 categories.
## ClaimID: 14417 categories
## Injury_Date: 2973 categories
## Date_reported: 2557 categories
```

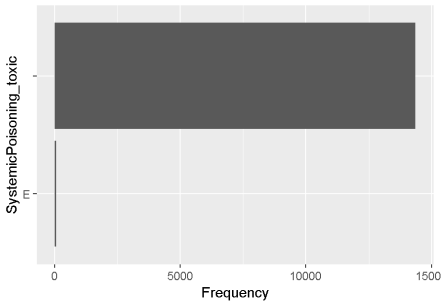

Work_related_injury_status
N
Y
0
2000
4000
6000
Frequency

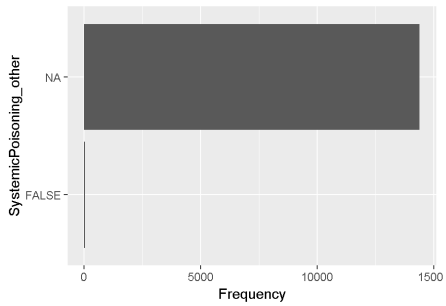

Amputation
B
0
5000
10000
Frequency

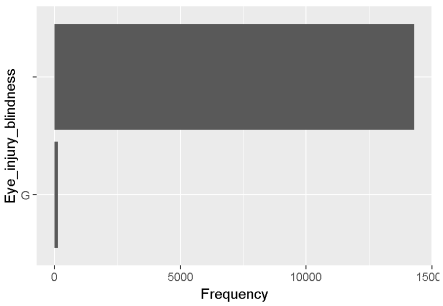

Death
A
0
5000
10000
Frequency

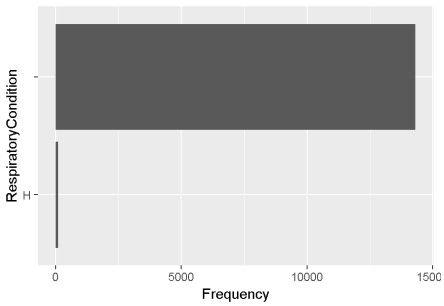

Burns_heat
C
0
5000
10000
Frequency

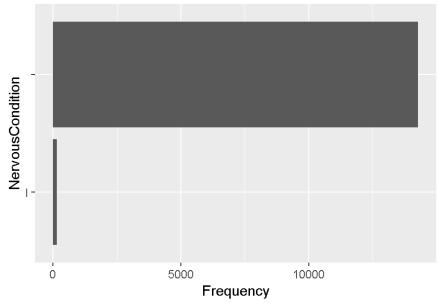

Burns_chemical
D
0
5000
10000
15000
Frequency

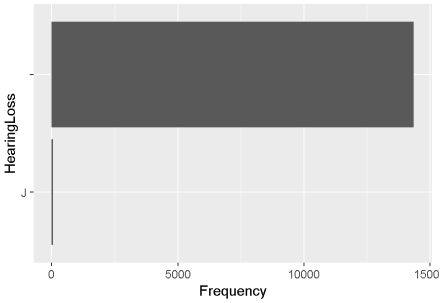

SystemicPoisoning_toxic
E
0
5000
10000
15000
Frequency

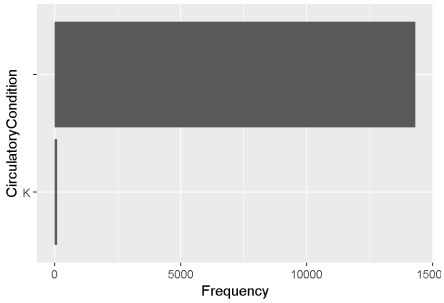

SystemicPoisoning_other
NA
FALSE
0
5000
10000
15000
Frequency

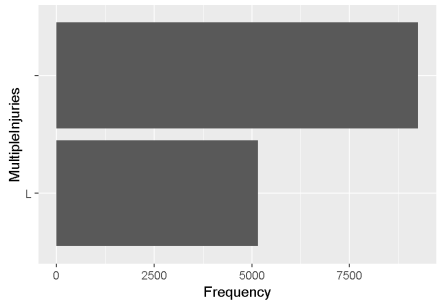

Eye_injury_blindness
G
0
5000
10000
15000
Frequency

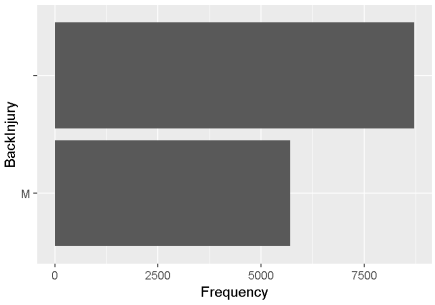

RespiratoryCondition
H
0
5000
10000
15000
Frequency

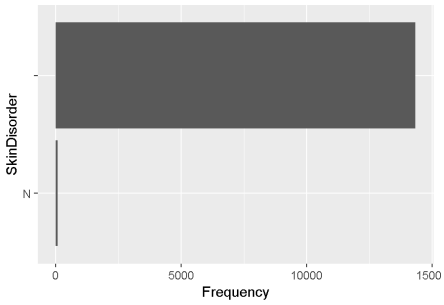

NervousCondition
I
0
5000
10000
Frequency

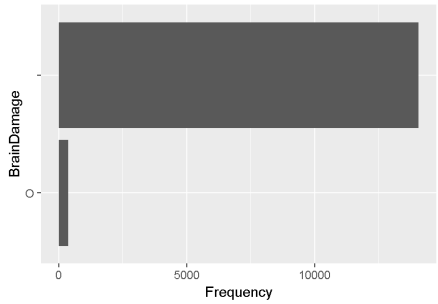

HearingLoss
J
0
5000
10000
15000
Frequency

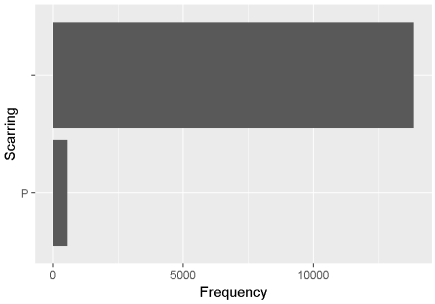

CirculatoryCondition
K
0
5000
10000
15000
Frequency

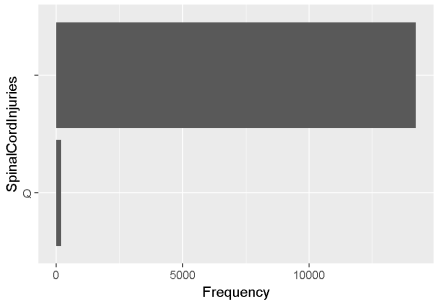

MultipleInjuries
L
0
2500
5000
7500
Frequency

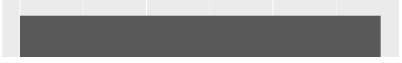

BackInjury
M
0
2500
5000
7500
Frequency

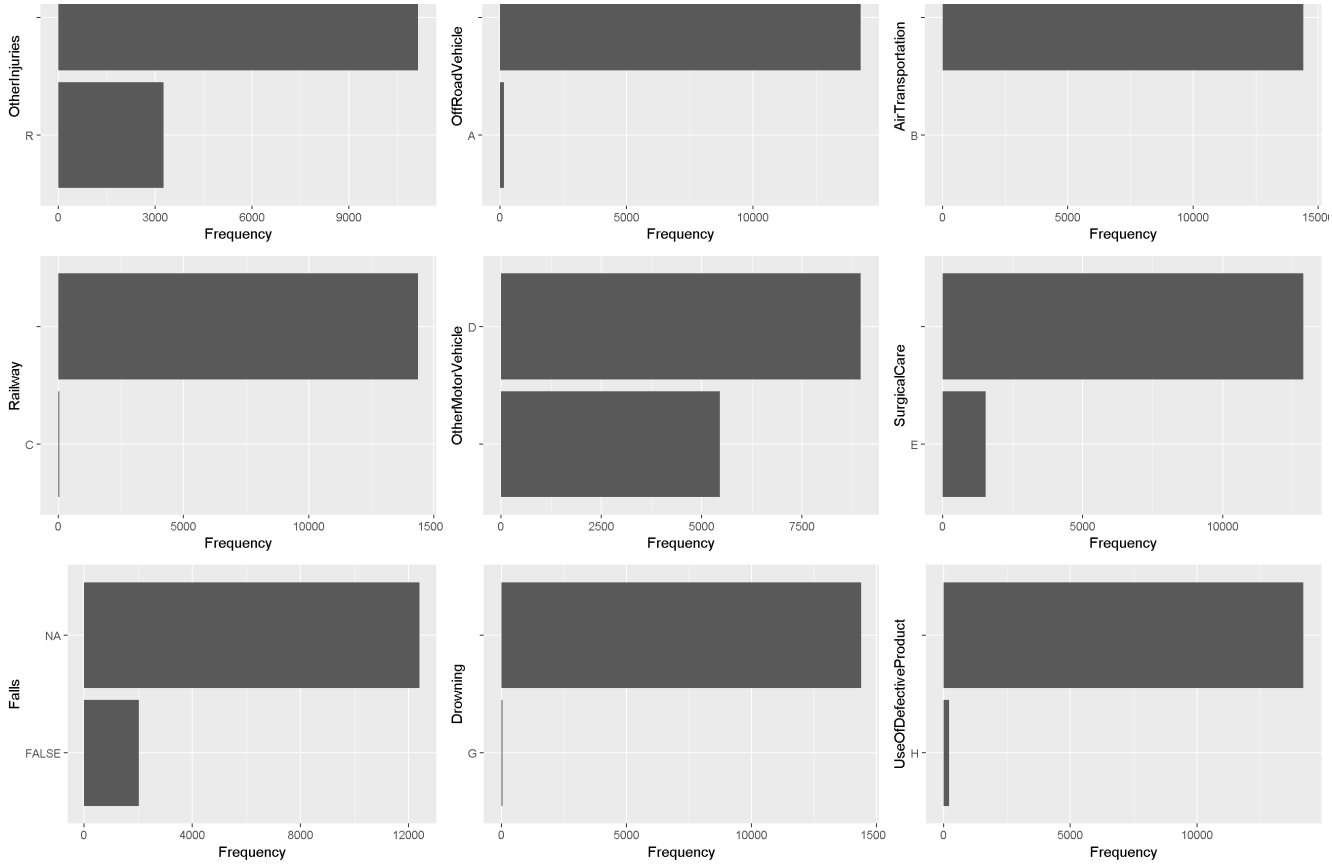

SkinDisorder
N
0
5000
10000
15000
Frequency


BrainDamage
O
0
5000
10000
Frequency

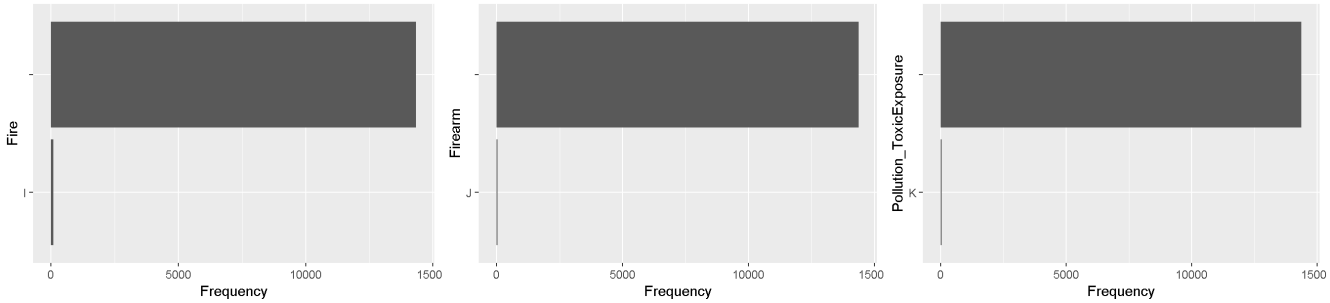

Scarring
P
0
5000
10000
Frequency

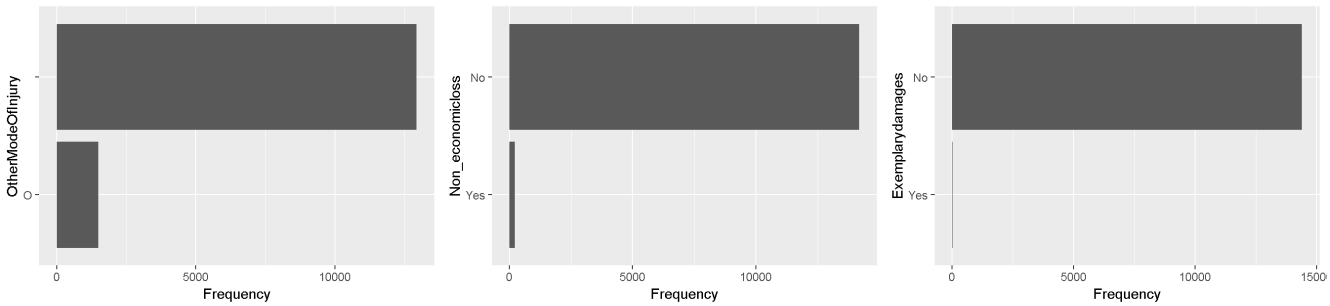

SpinalCordInjuries
Q
0
5000
10000
Frequency



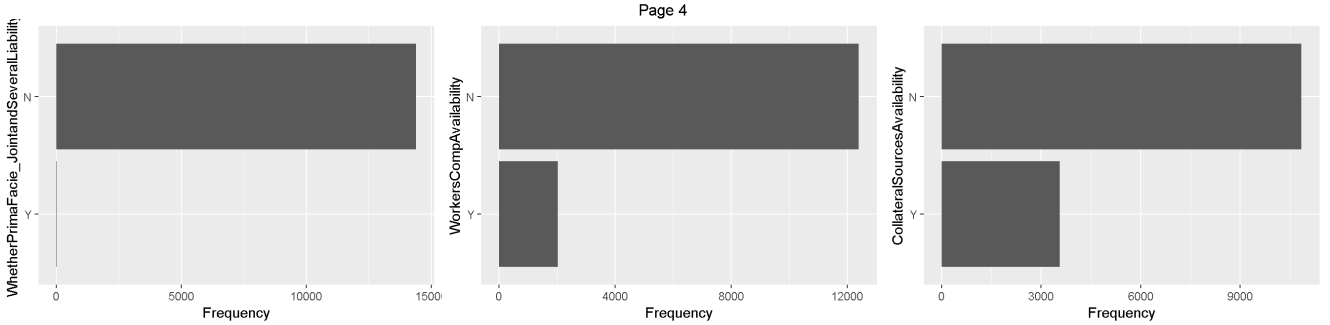

OtherInjuries
R
0
3000
6000
9000
Frequency
OffRoadVehicle
A
0
5000
10000
Frequency
AirTransportation
B
0
5000
10000
Frequency
Railway
C
0
5000
10000
Frequency
OtherMotorVehicle
D
0
2500
5000
7500
Frequency
SurgicalCare
E
0
5000
10000
Frequency
Falls
NA
FALSE
0
4000
8000
12000
Frequency
Drowning
G
0
5000
10000
Frequency
UseOfDefectiveProduct
H
0
5000
10000
Frequency

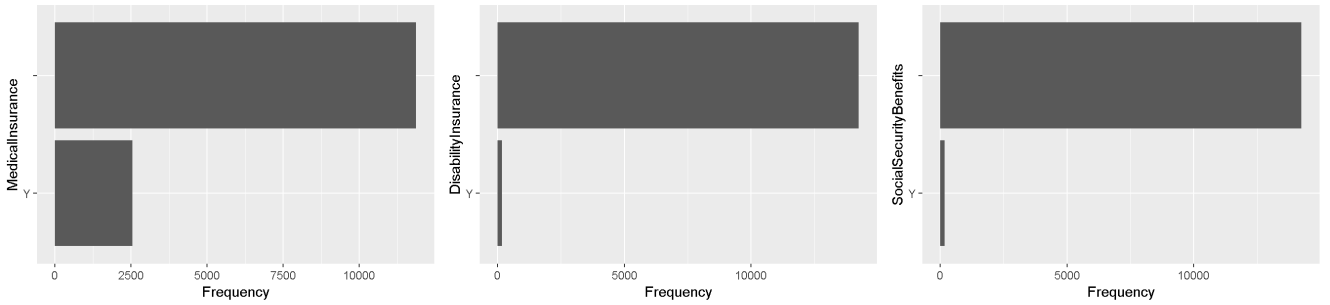
Page 3





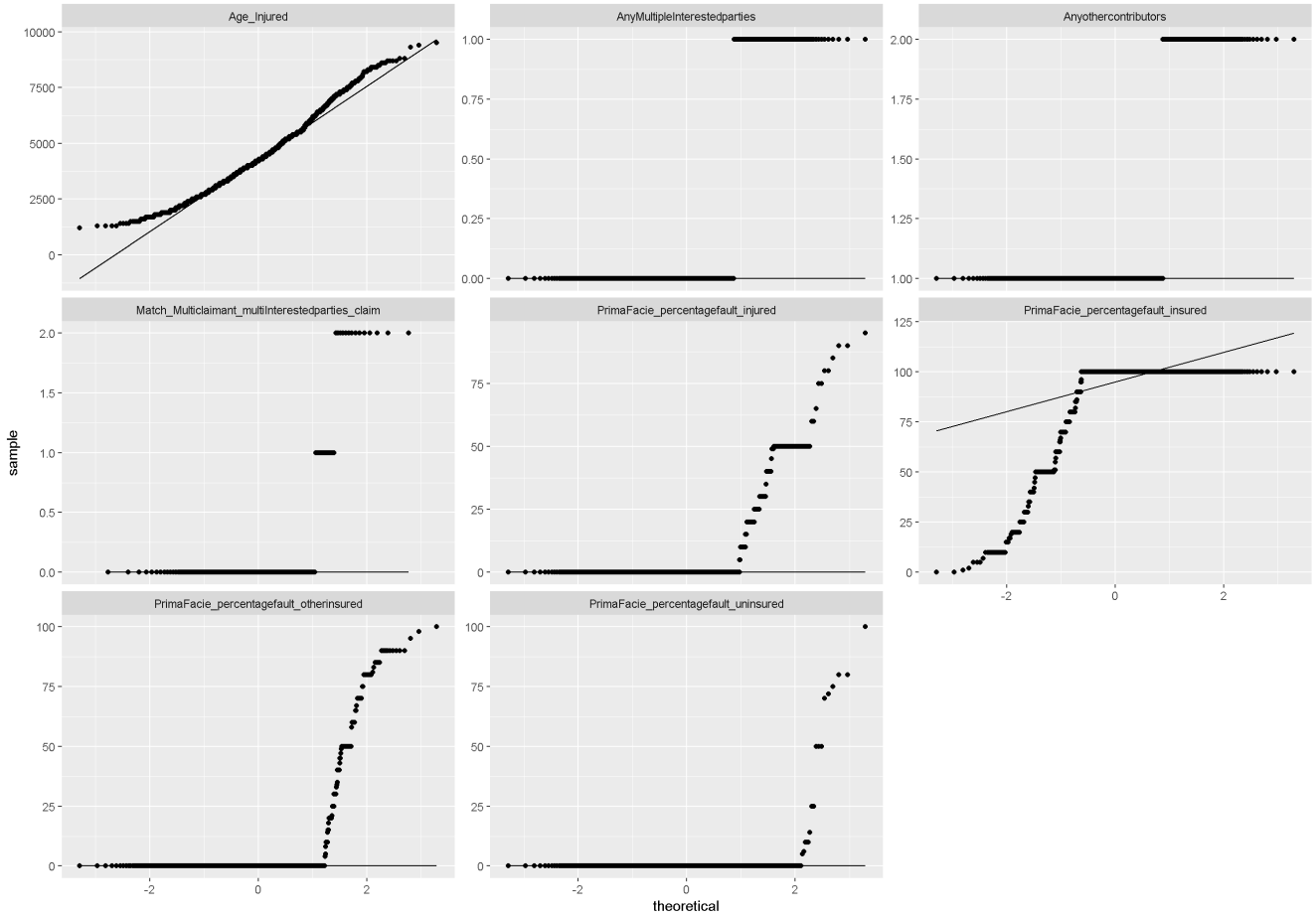
Page 4





Page 5

QQ Plot



Correlation Analysis

```
## Insufficient complete rows! Skipping correlation analysis.
```

Principle Component Analysis

```
## Insufficient complete rows! Skipping principle component analysis.
```