

# EDA Report - GSS Demographic Survey

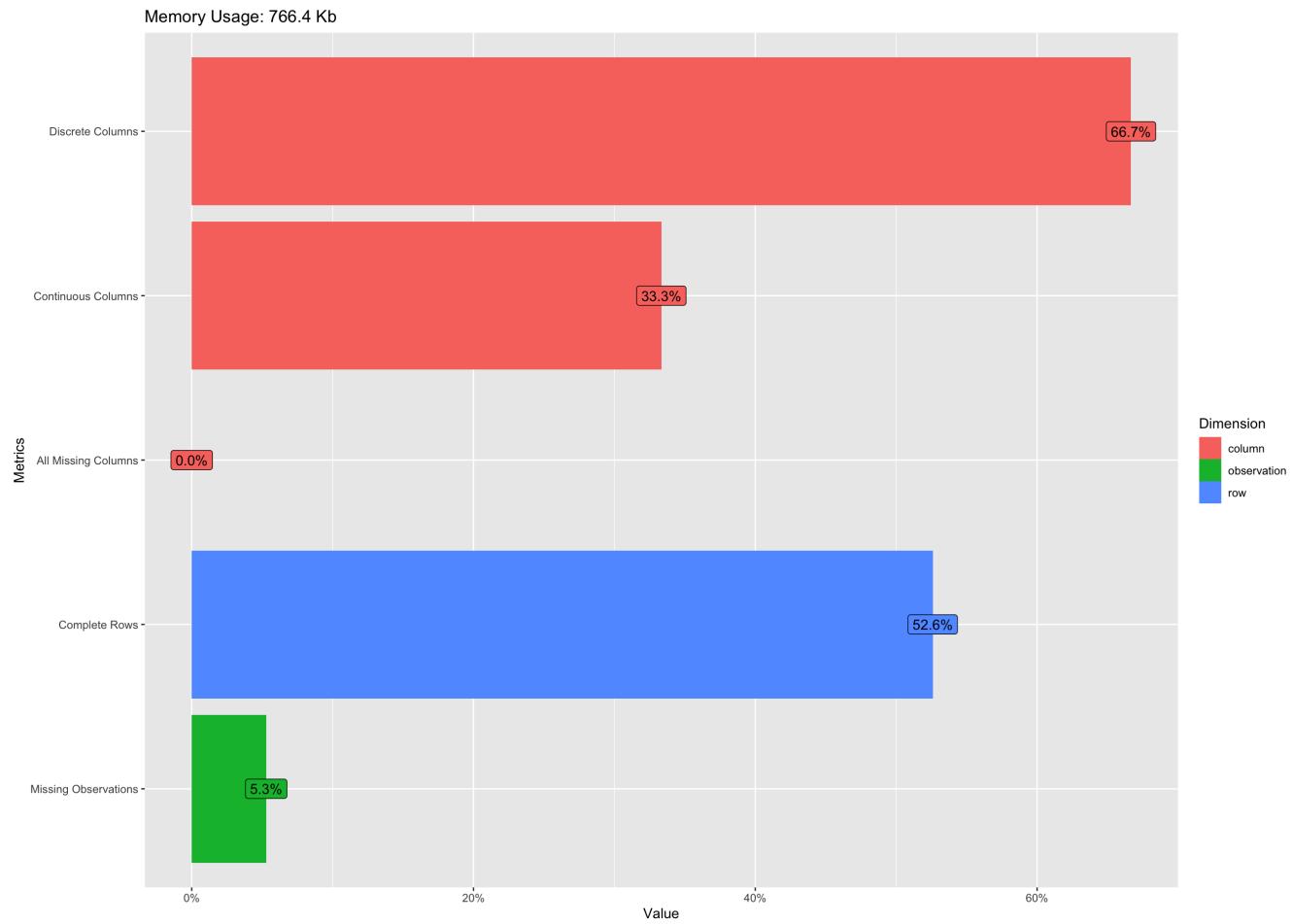
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (with frequency)
  - Bar Chart (by rincome)
  - QQ Plot
  - QQ Plot (by rincome)
- Correlation Analysis
- Principal Component Analysis
- Bivariate Distribution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

## Basic Statistics

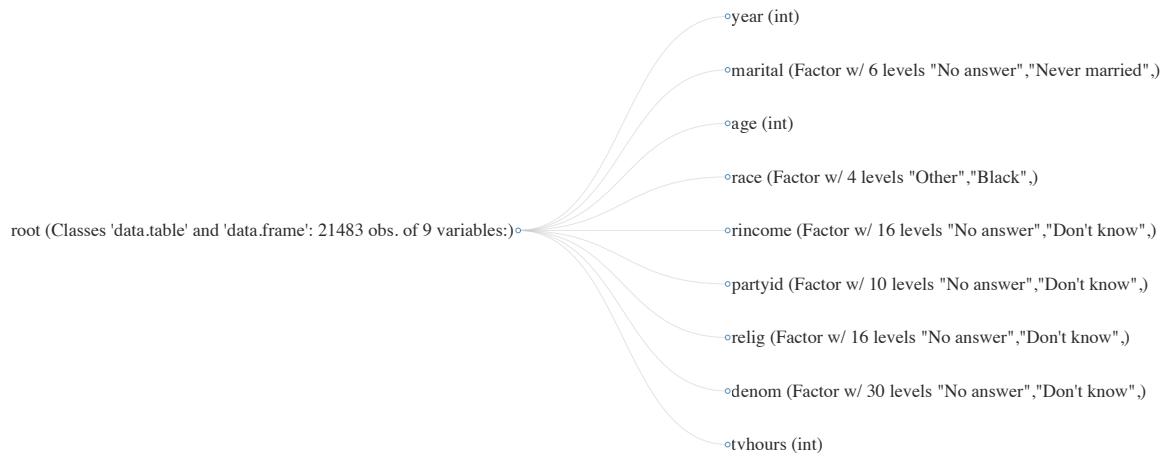
### Raw Counts

Name	Value
Rows	21,483
Columns	9
Discrete columns	6
Continuous columns	3
All missing columns	0
Missing observations	10,222
Complete Rows	11,299
Total observations	193,347
Memory allocation	766.4 Kb

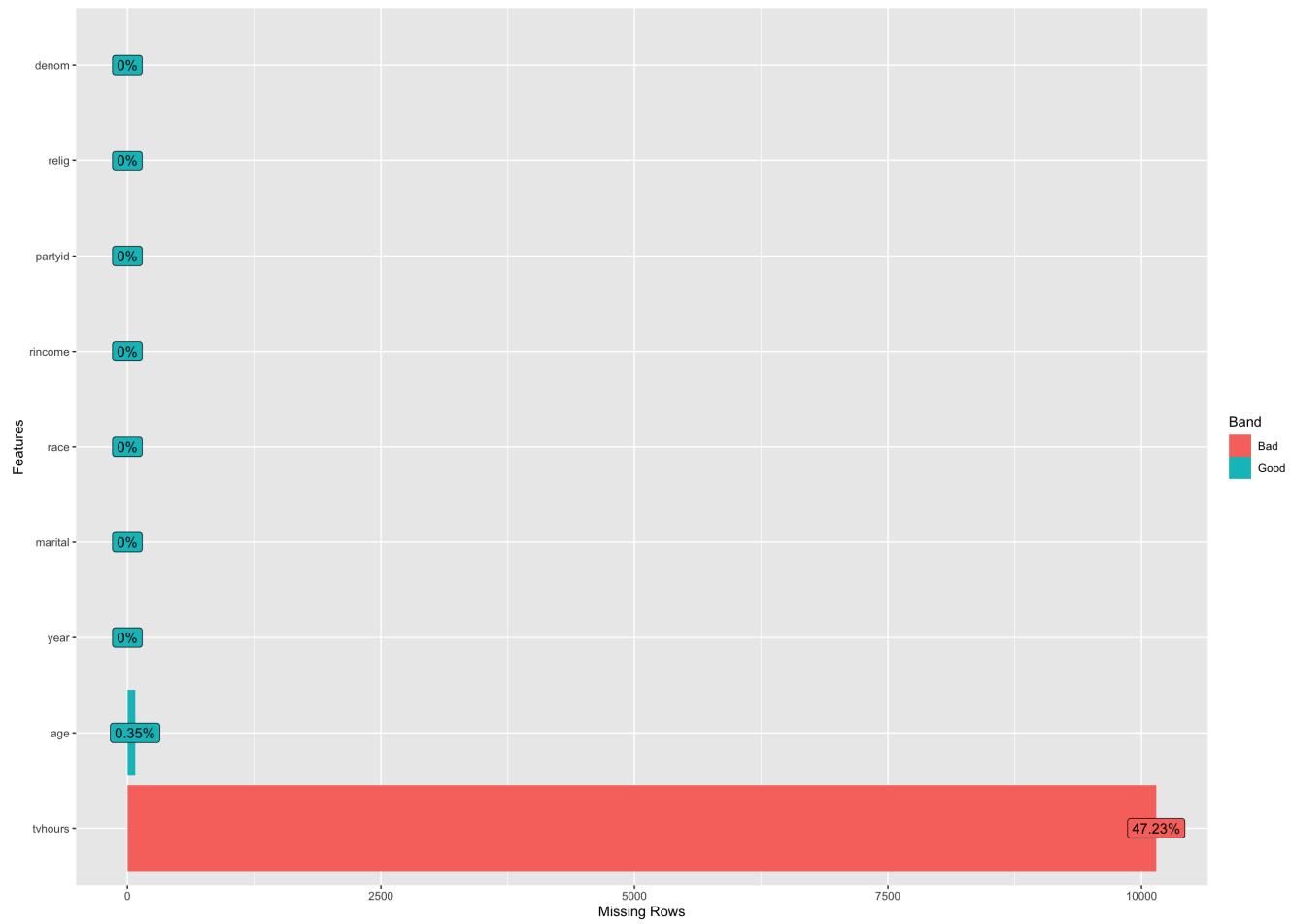
### Percentages



## Data Structure

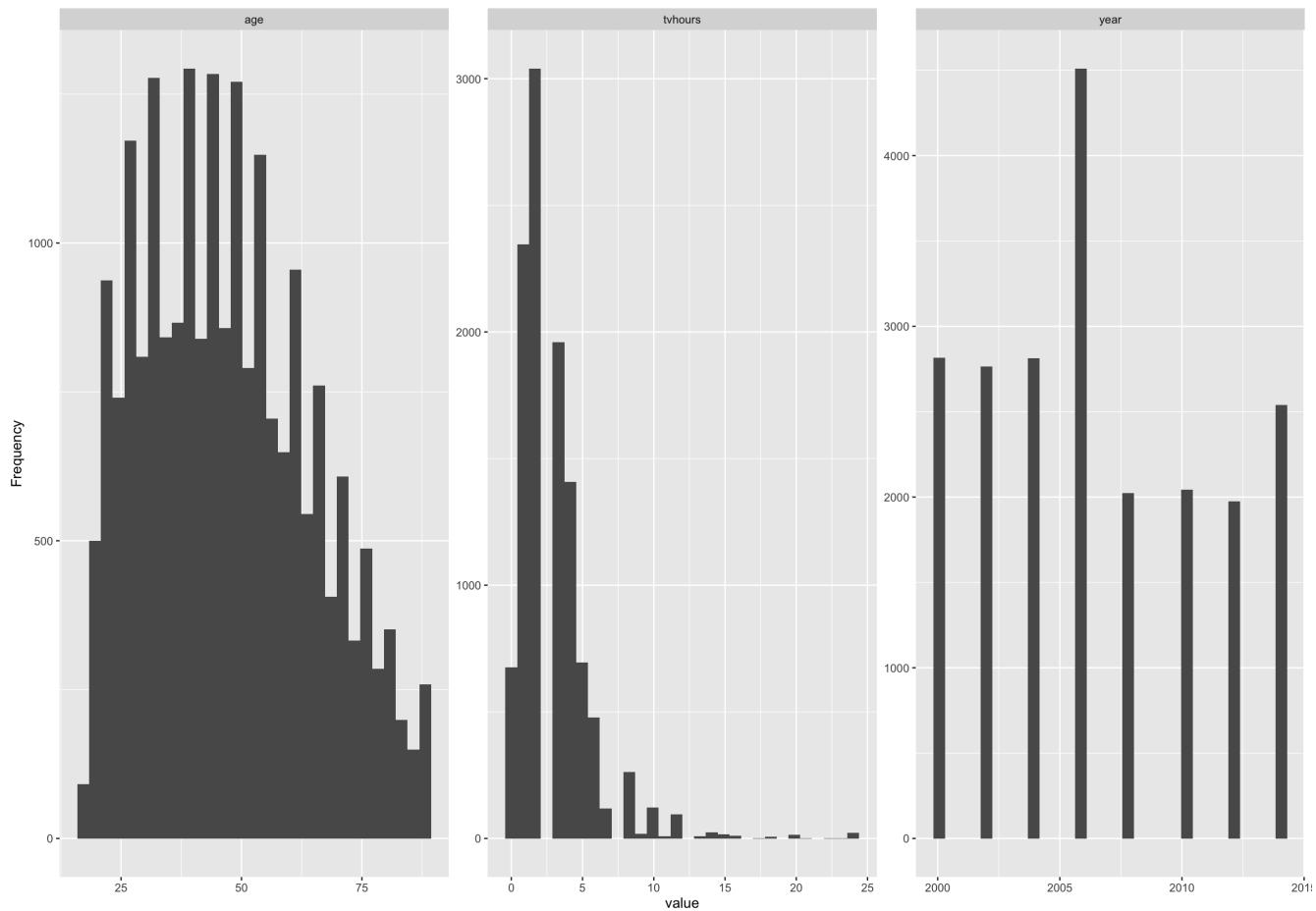


## Missing Data Profile

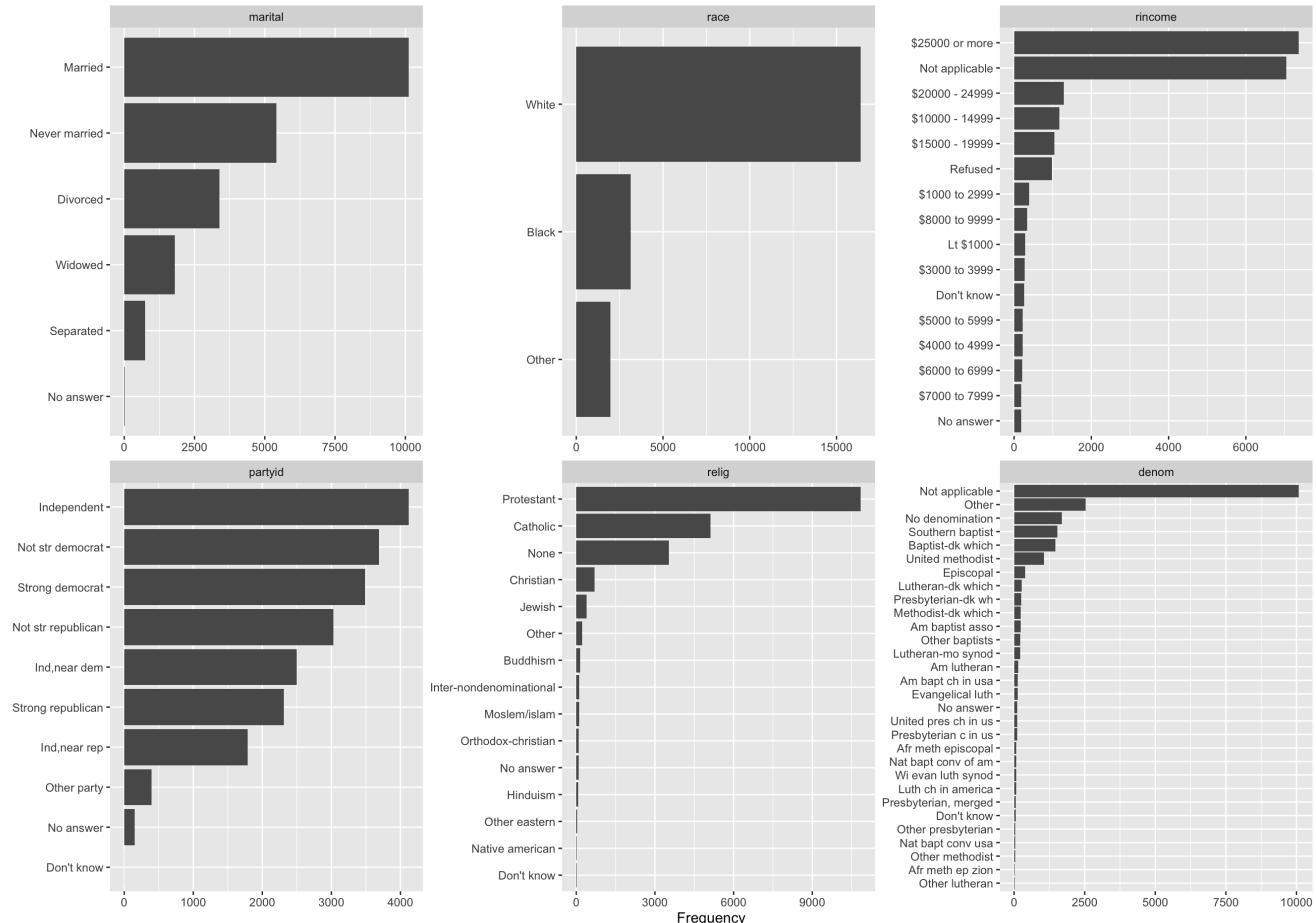


## Univariate Distribution

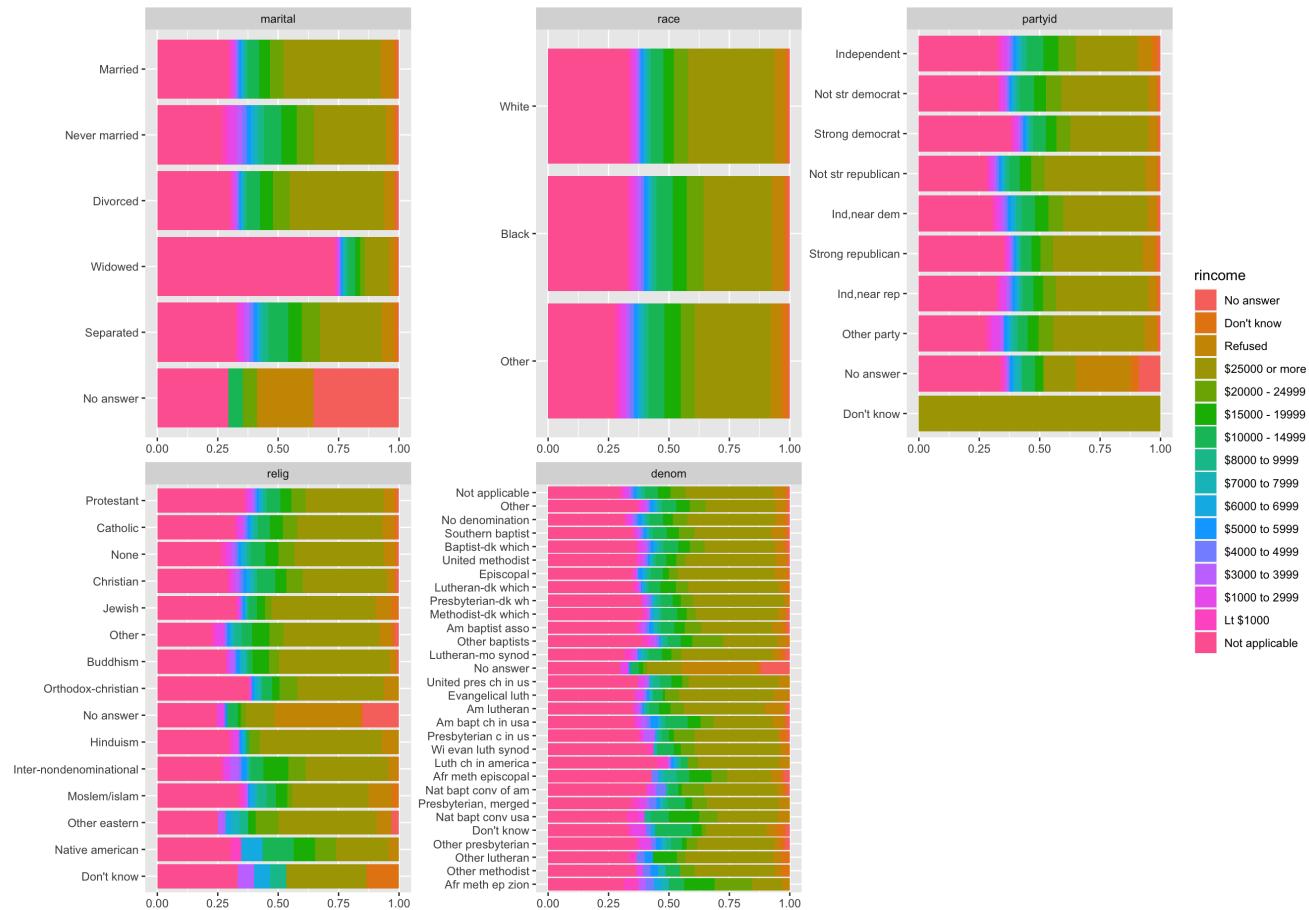
### Histogram



Bar Chart (with frequency)



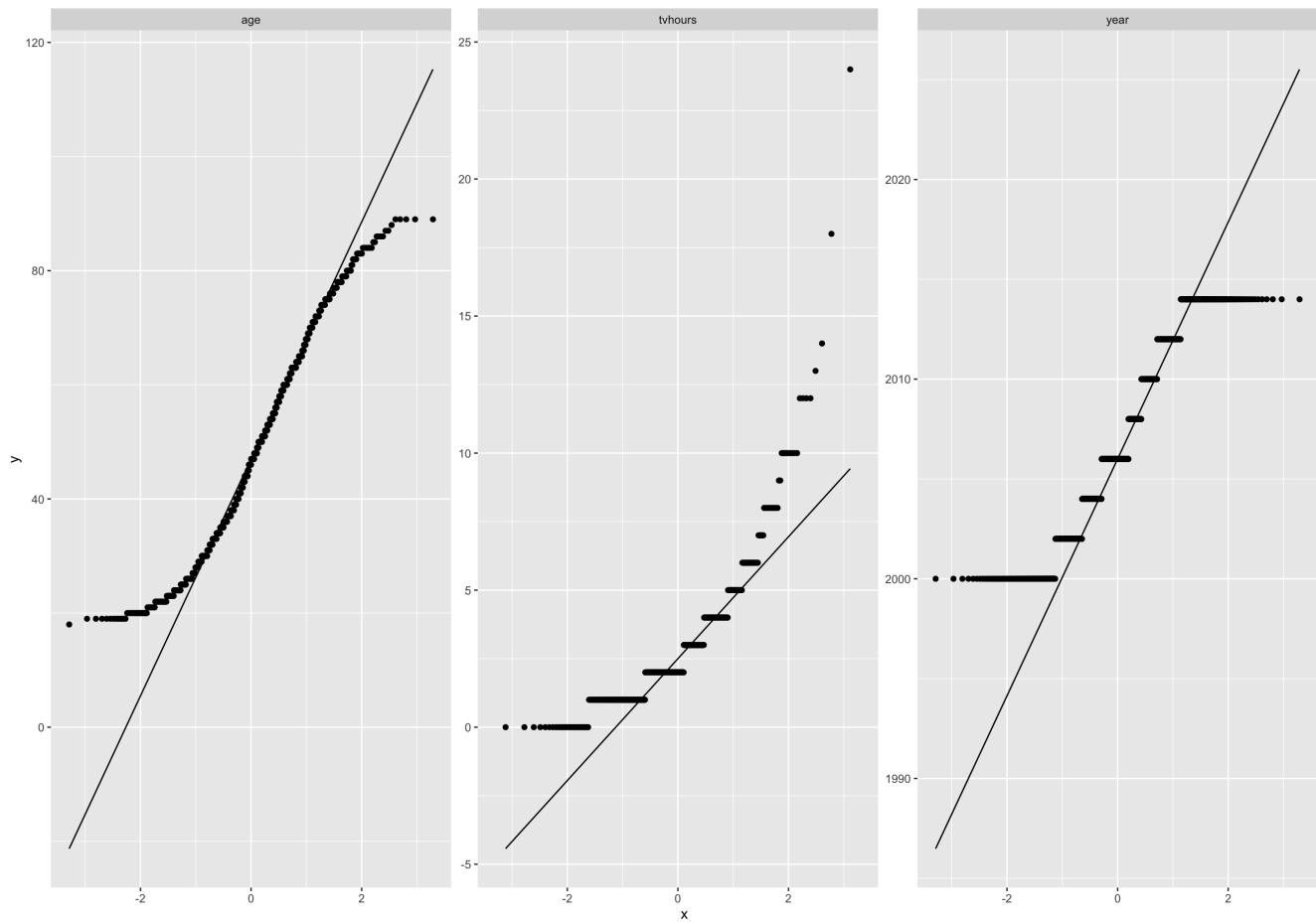
## Bar Chart (by rincome)



## QQ Plot

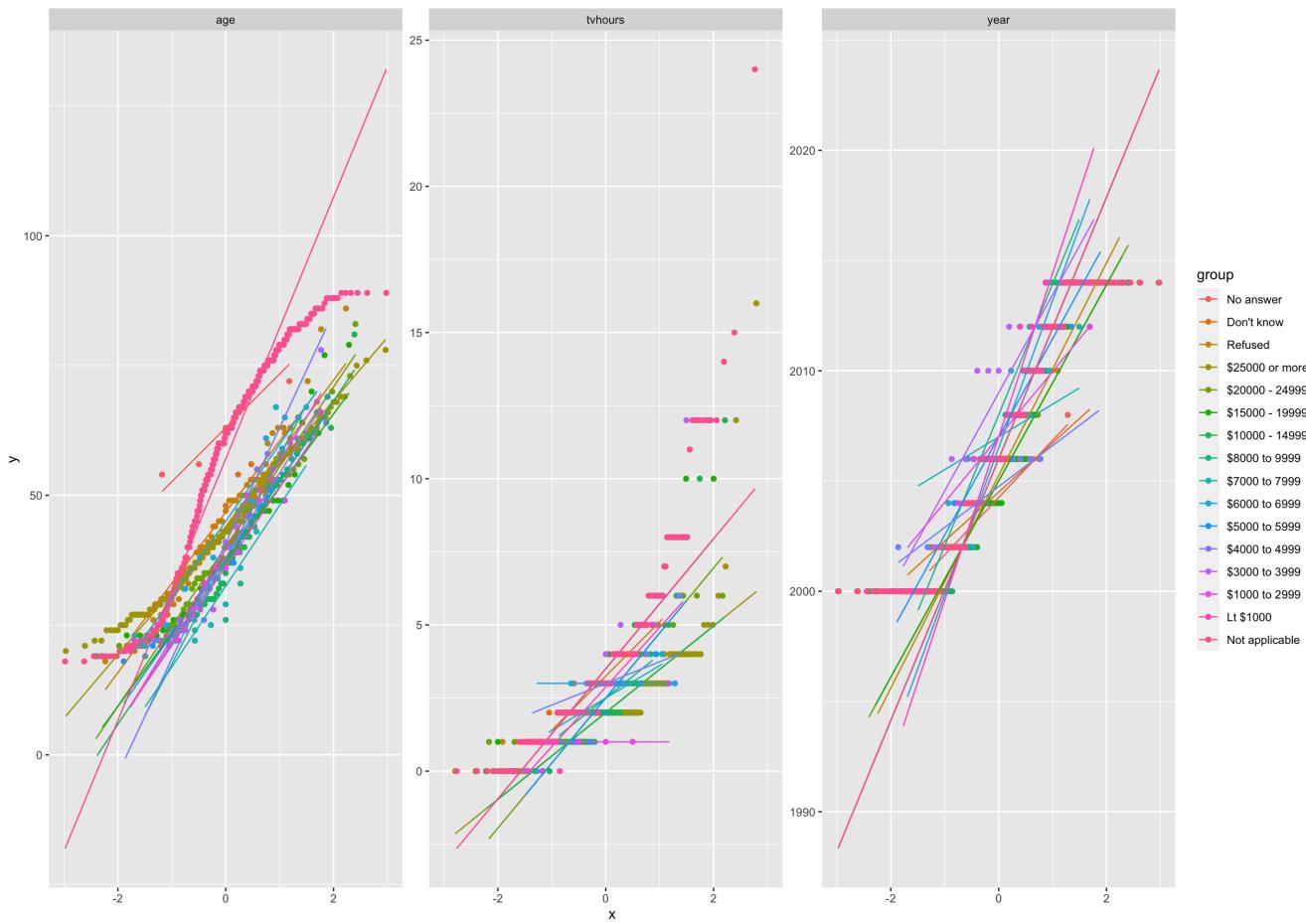
```
## Warning: Removed 458 rows containing non-finite values (stat_qq).
```

```
## Warning: Removed 458 rows containing non-finite values (stat_qq_line).
```



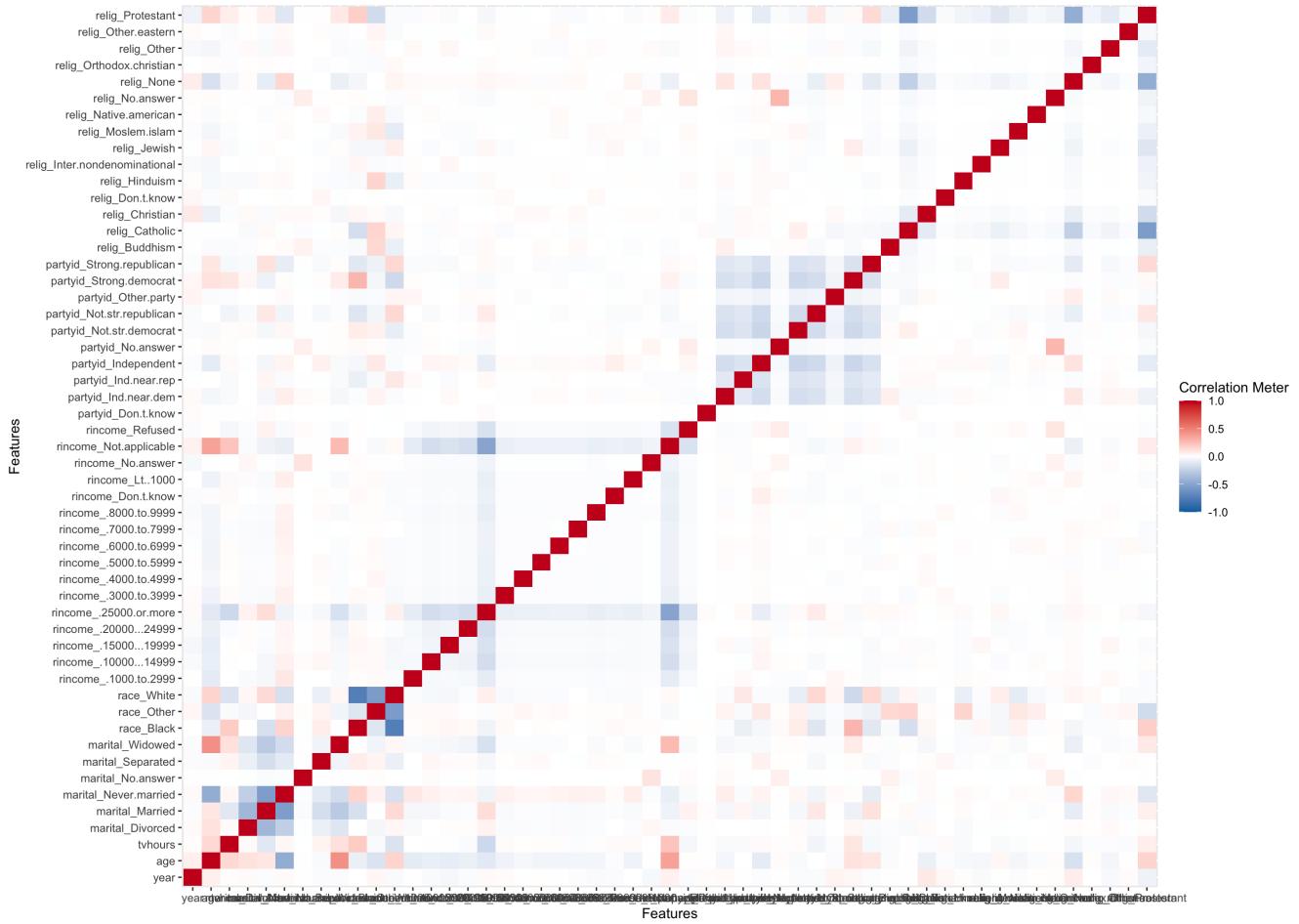
## QQ Plot (by rincome)

```
## Warning: Removed 470 rows containing non-finite values (stat_qq).
## Warning: Removed 470 rows containing non-finite values (stat_qq_line).
```



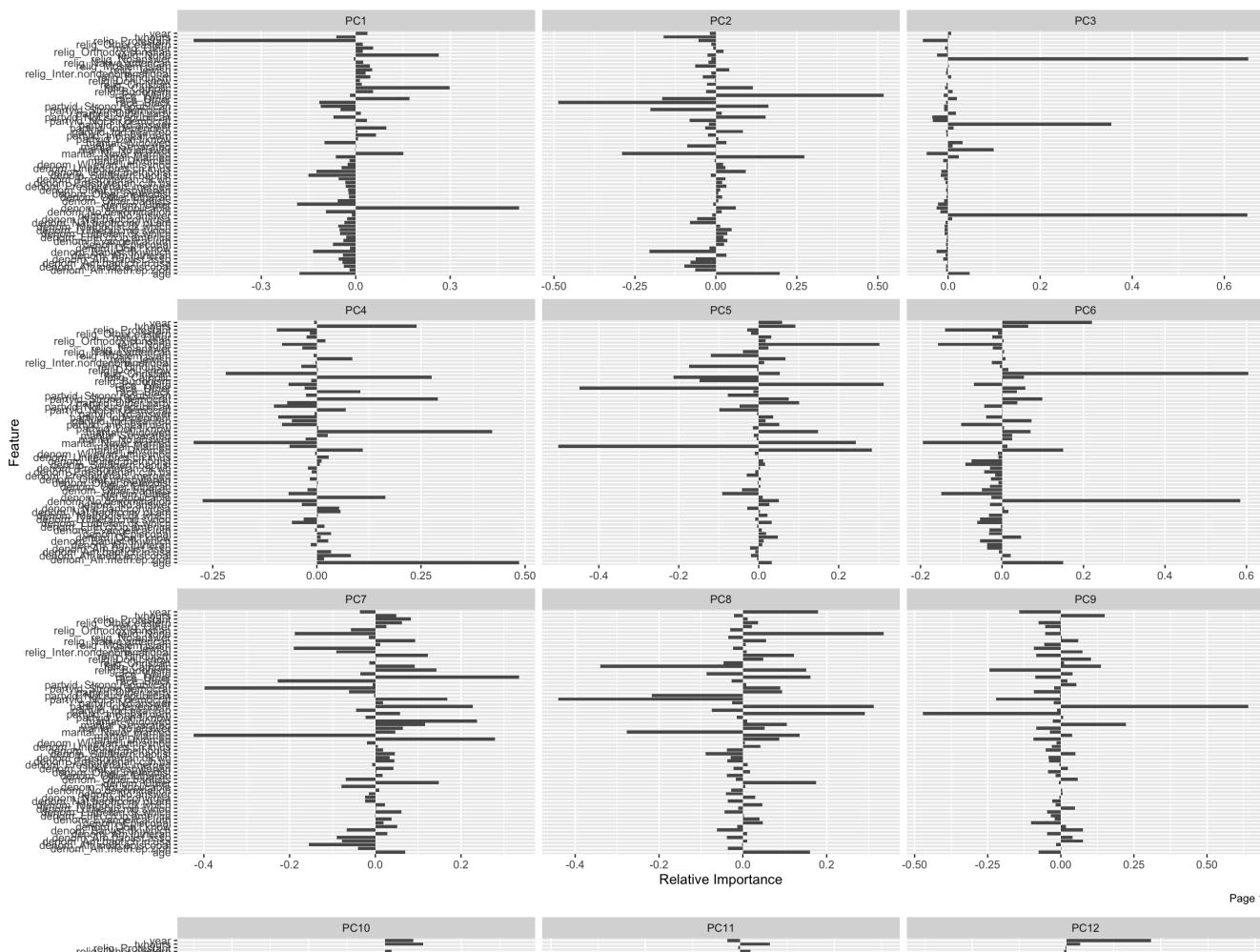
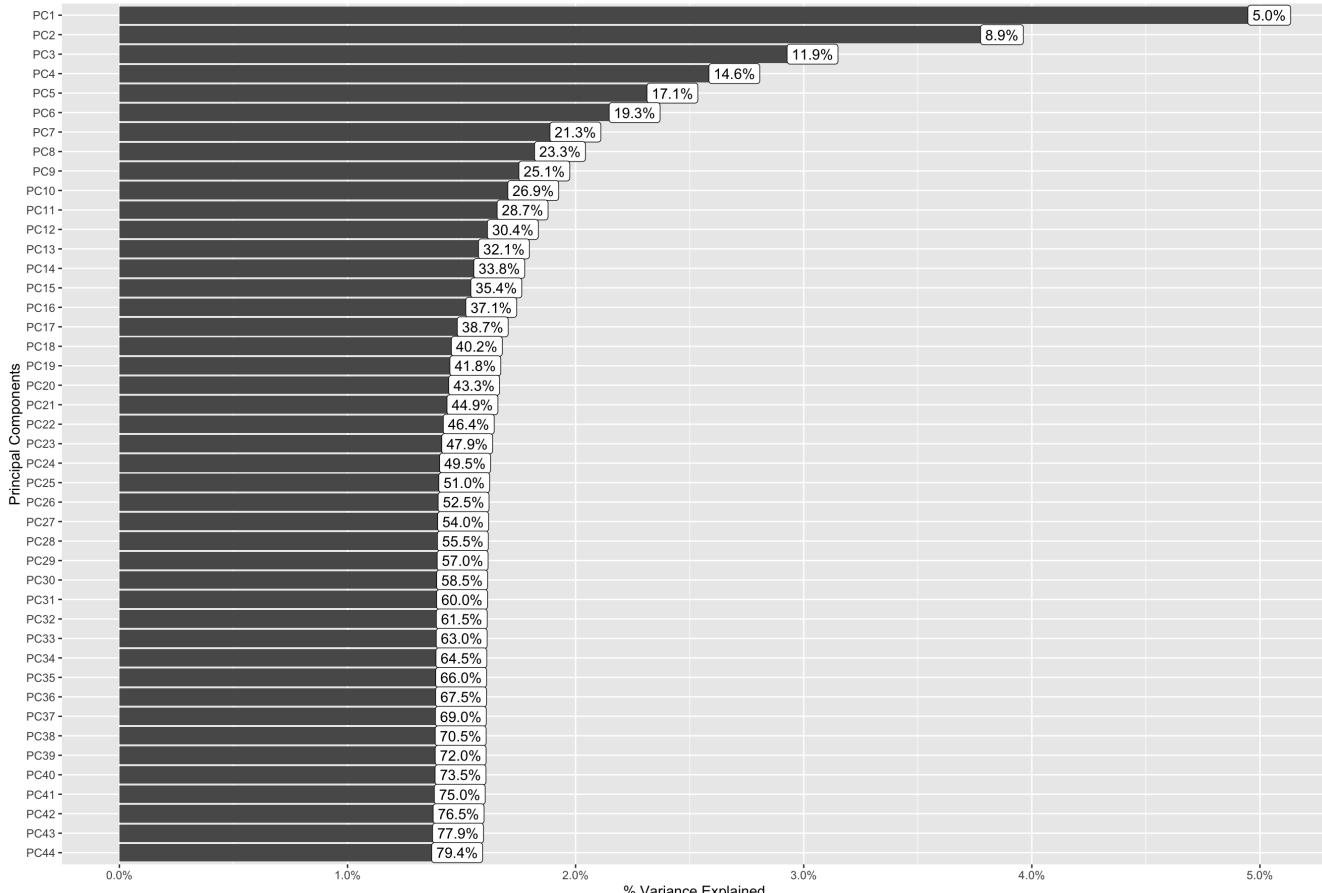
## Correlation Analysis

```
## 1 features with more than 20 categories ignored!
## denom: 30 categories
```

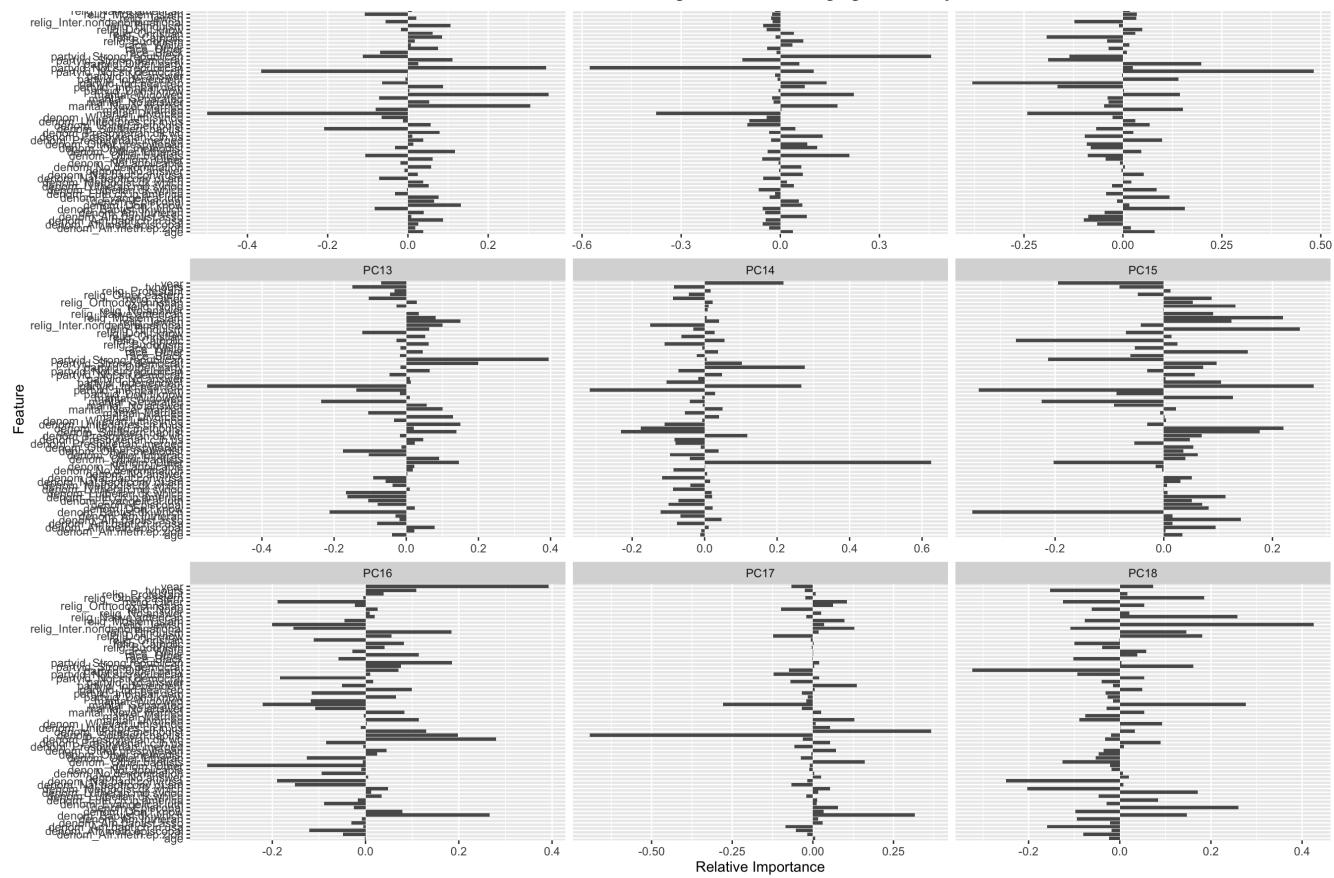


# Principal Component Analysis

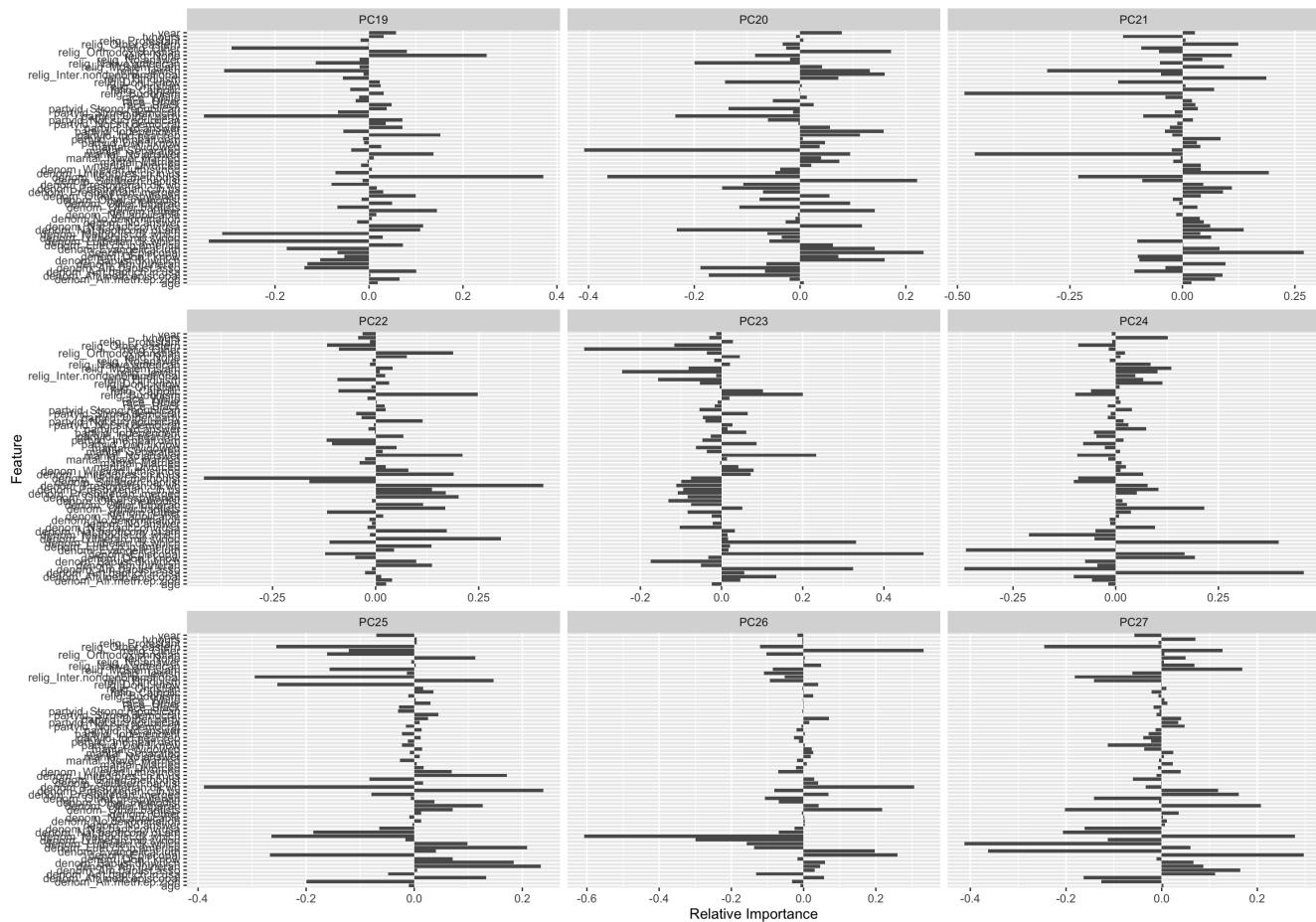
% Variance Explained By Principal Components  
 (Note: Labels indicate cumulative % explained variance)



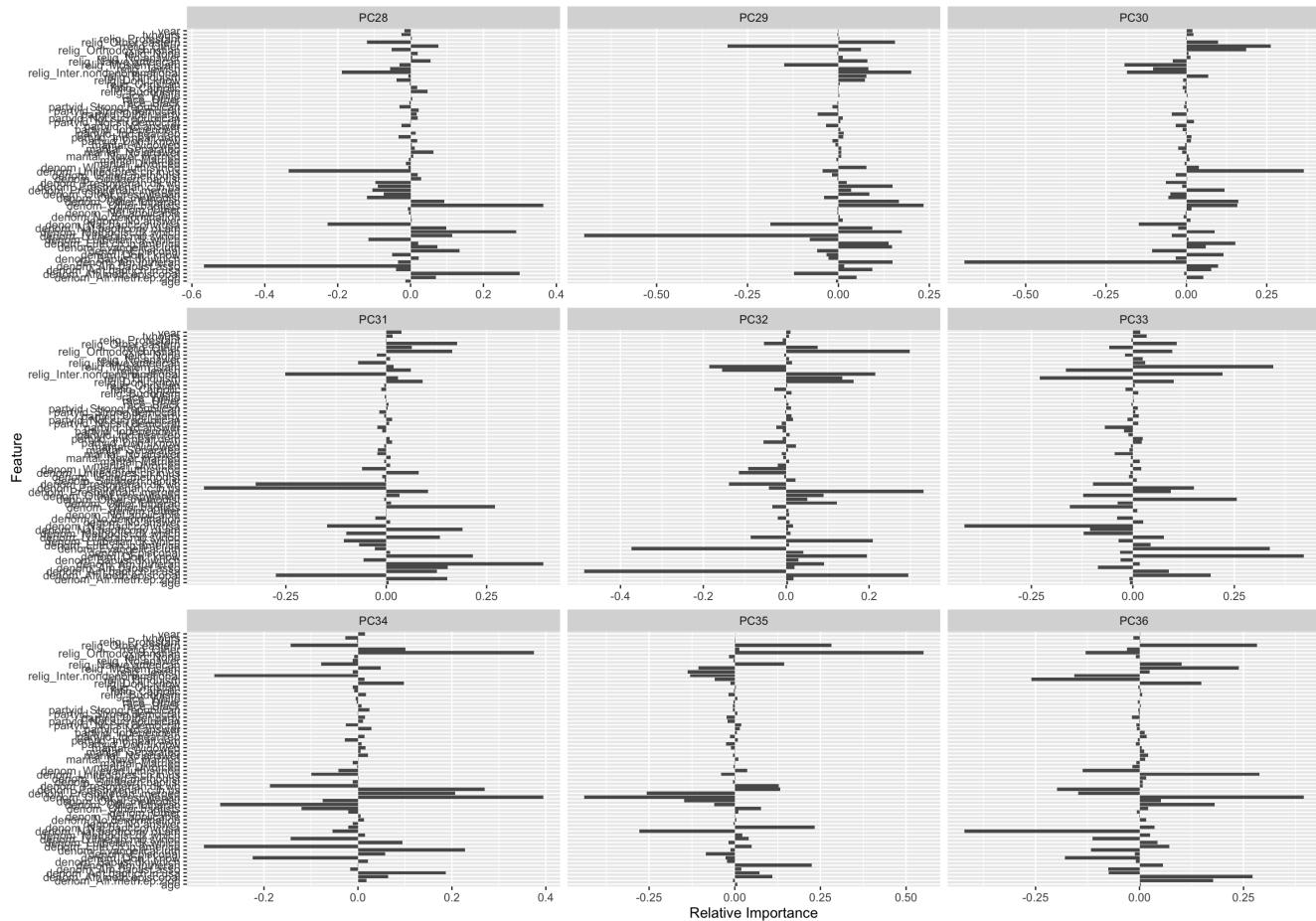
Page 1



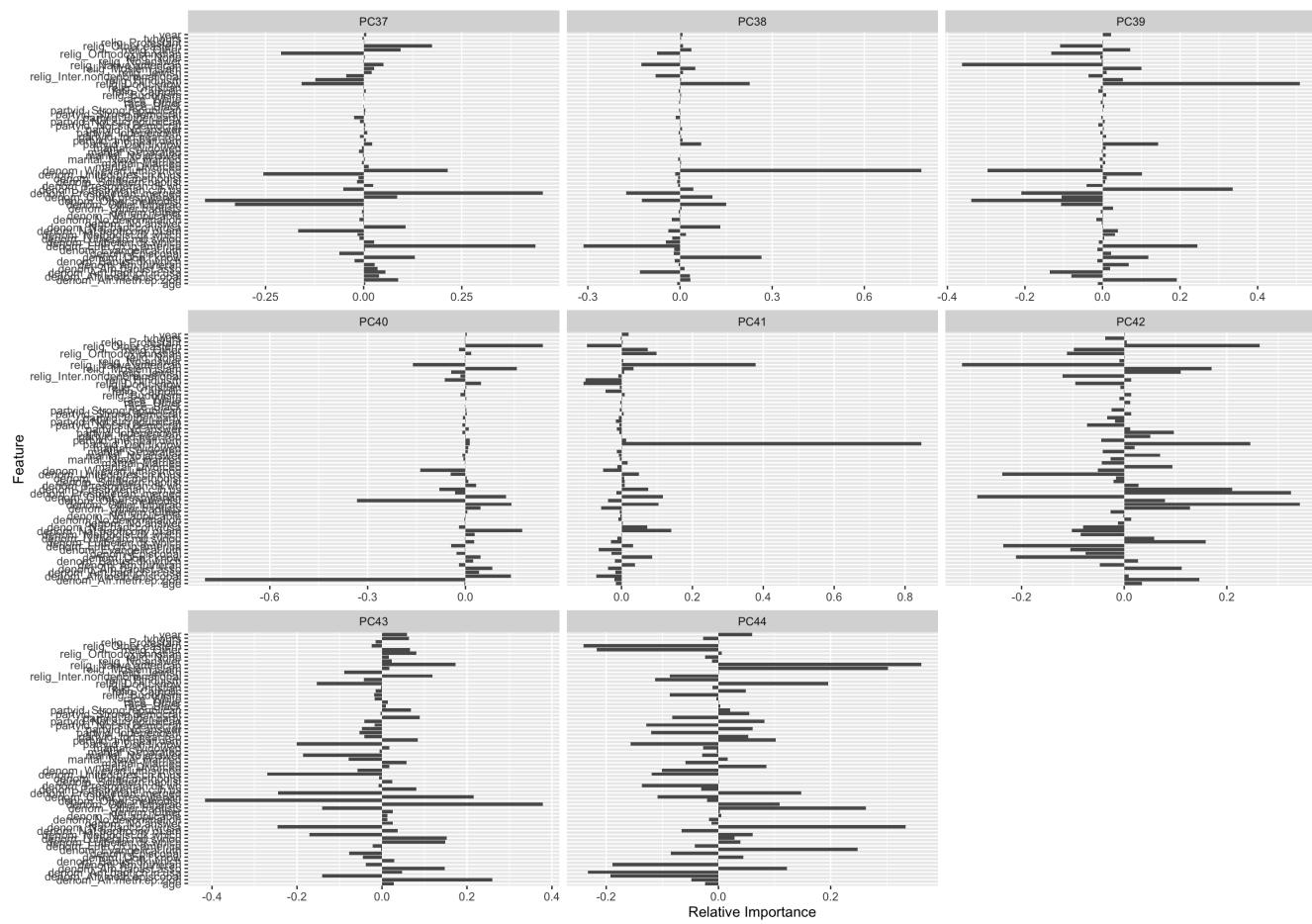
Page 2



Page 3



Page 4

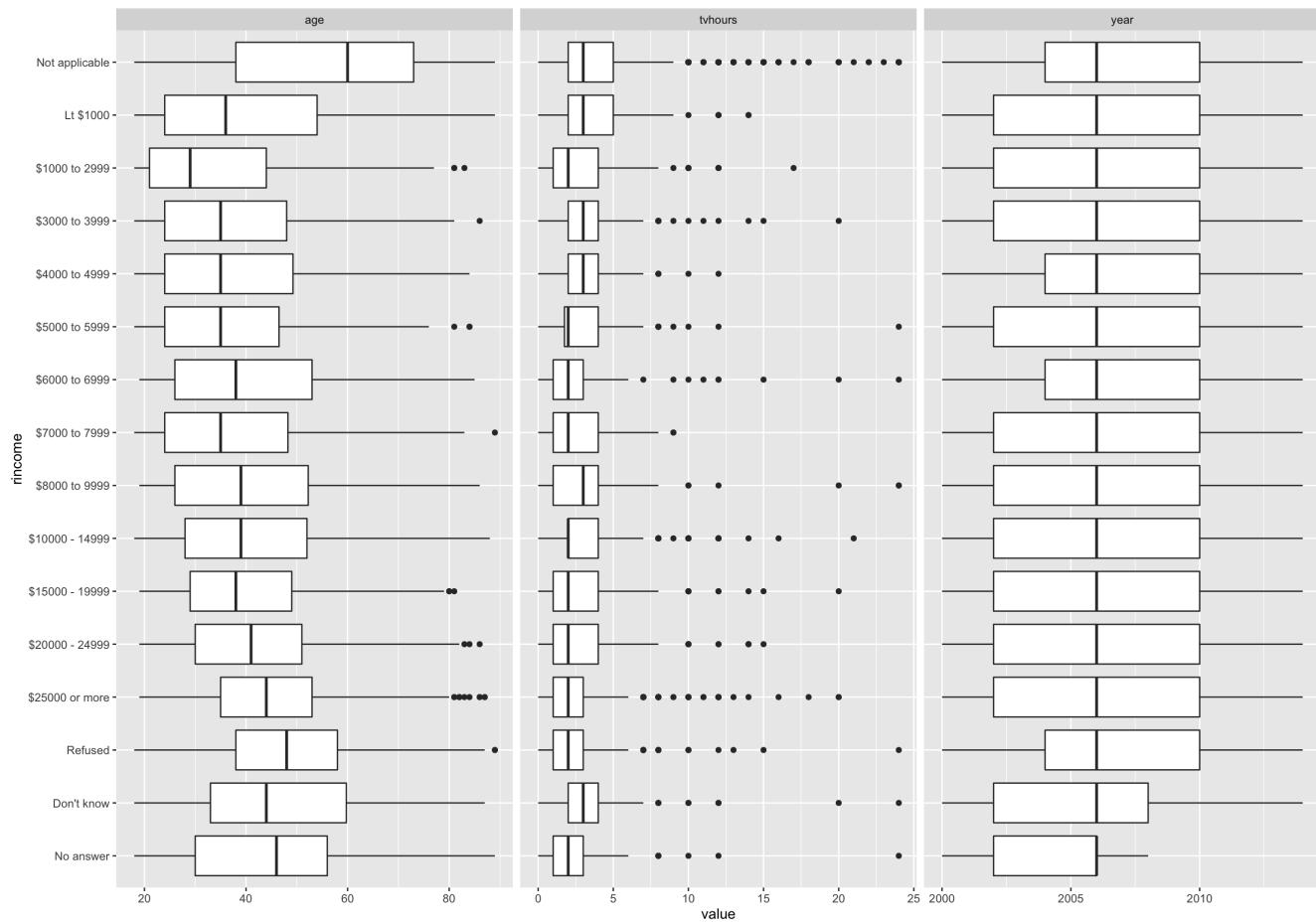


Page 5

## Bivariate Distribution

Boxplot (by rincome)

```
## Warning: Removed 10222 rows containing non-finite values (stat_boxplot).
```



Scatterplot (by rincome)

