

# Data Profiling Report

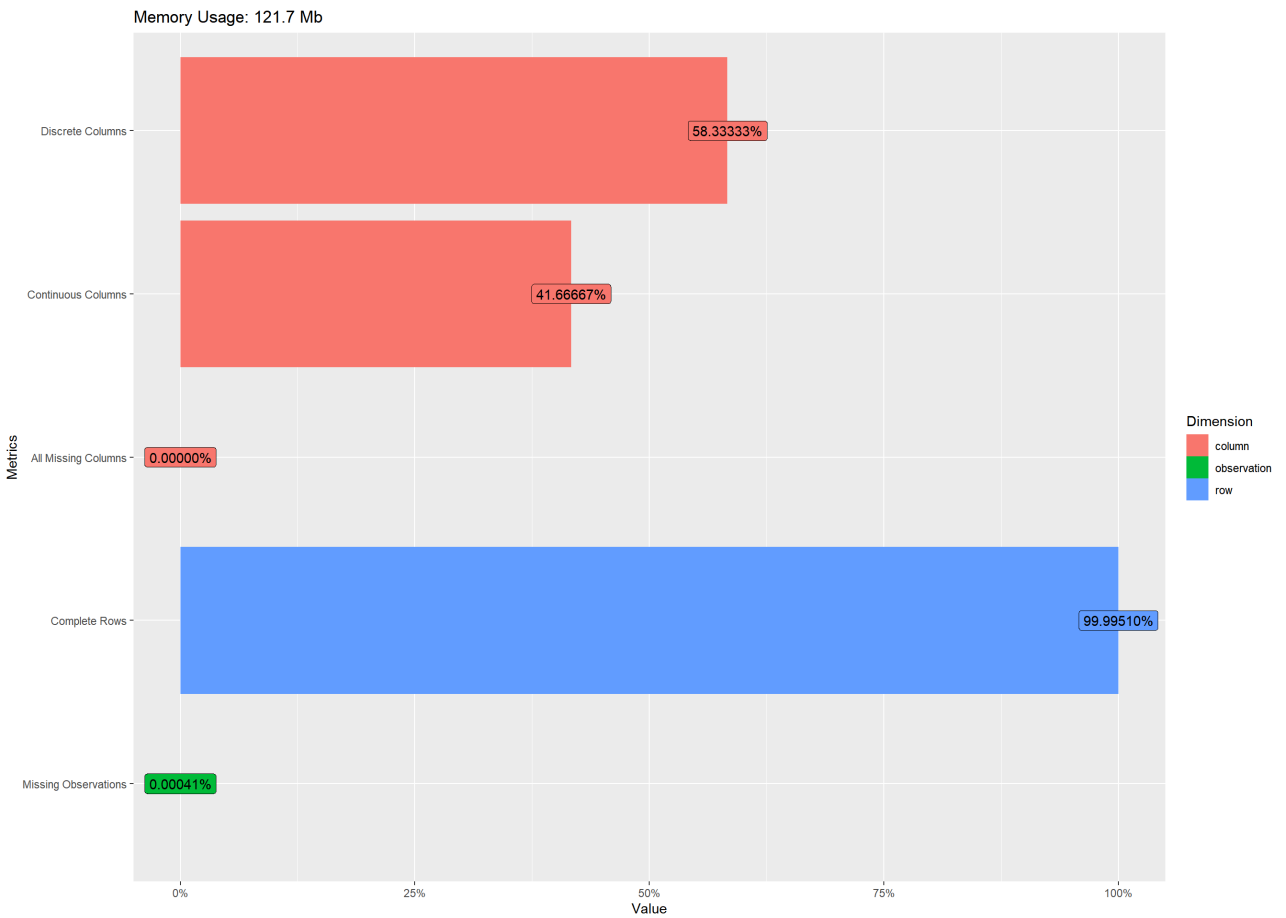
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (with frequency)
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

## Basic Statistics

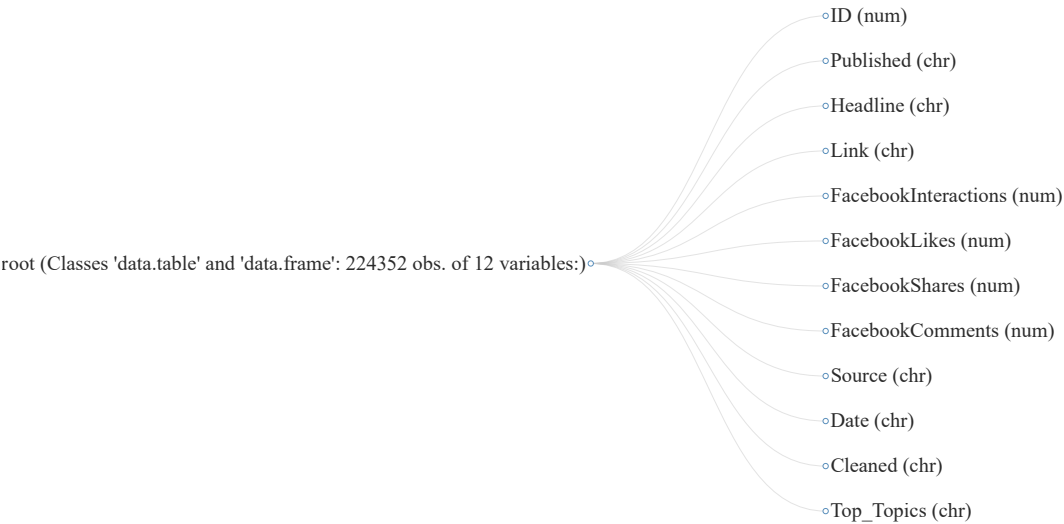
### Raw Counts

| Name                 | Value     |
|----------------------|-----------|
| Rows                 | 224,352   |
| Columns              | 12        |
| Discrete columns     | 7         |
| Continuous columns   | 5         |
| All missing columns  | 0         |
| Missing observations | 11        |
| Complete Rows        | 224,341   |
| Total observations   | 2,692,224 |
| Memory allocation    | 121.7 Mb  |

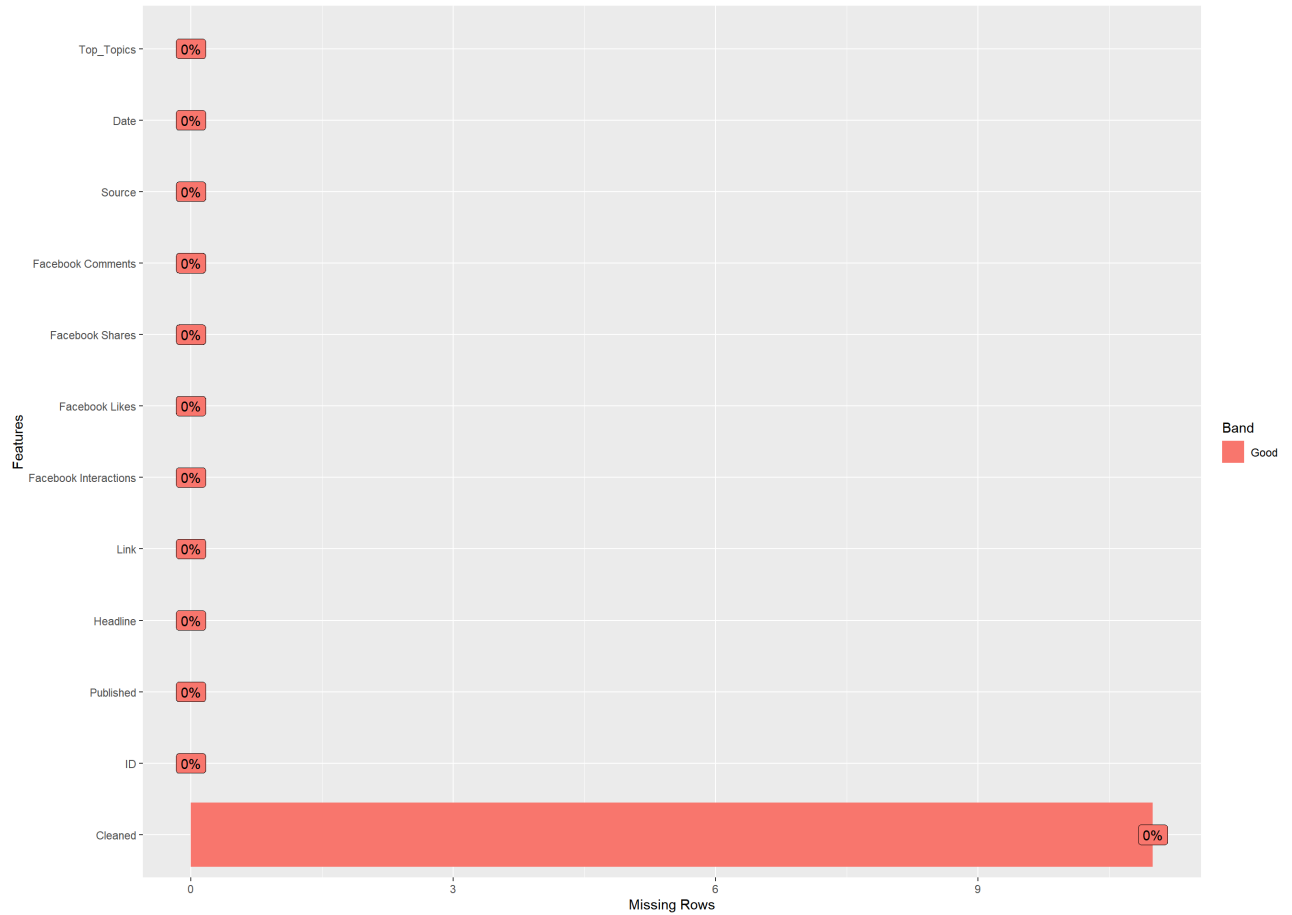
### Percentages



## Data Structure

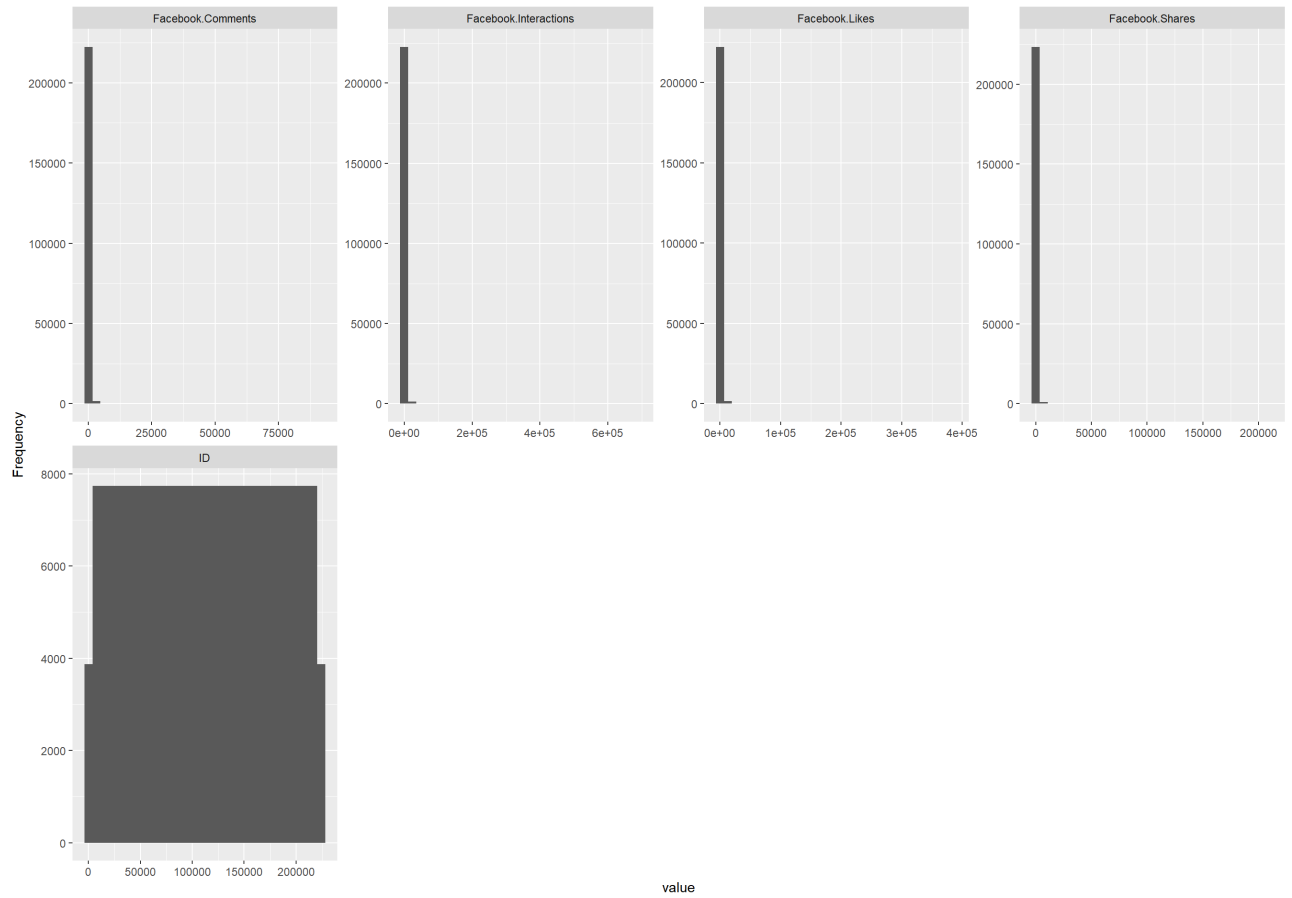


Missing Data Profile



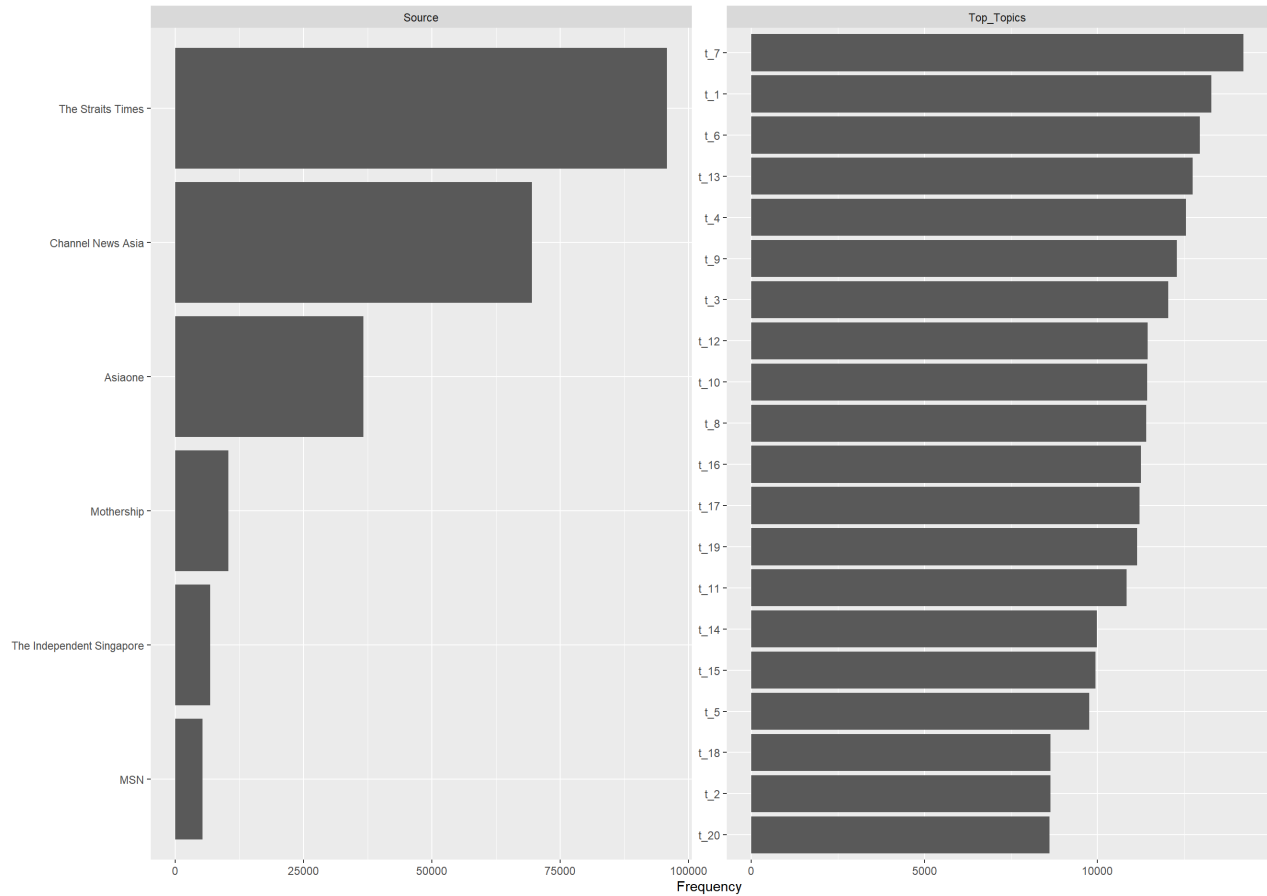
Univariate Distribution

Histogram

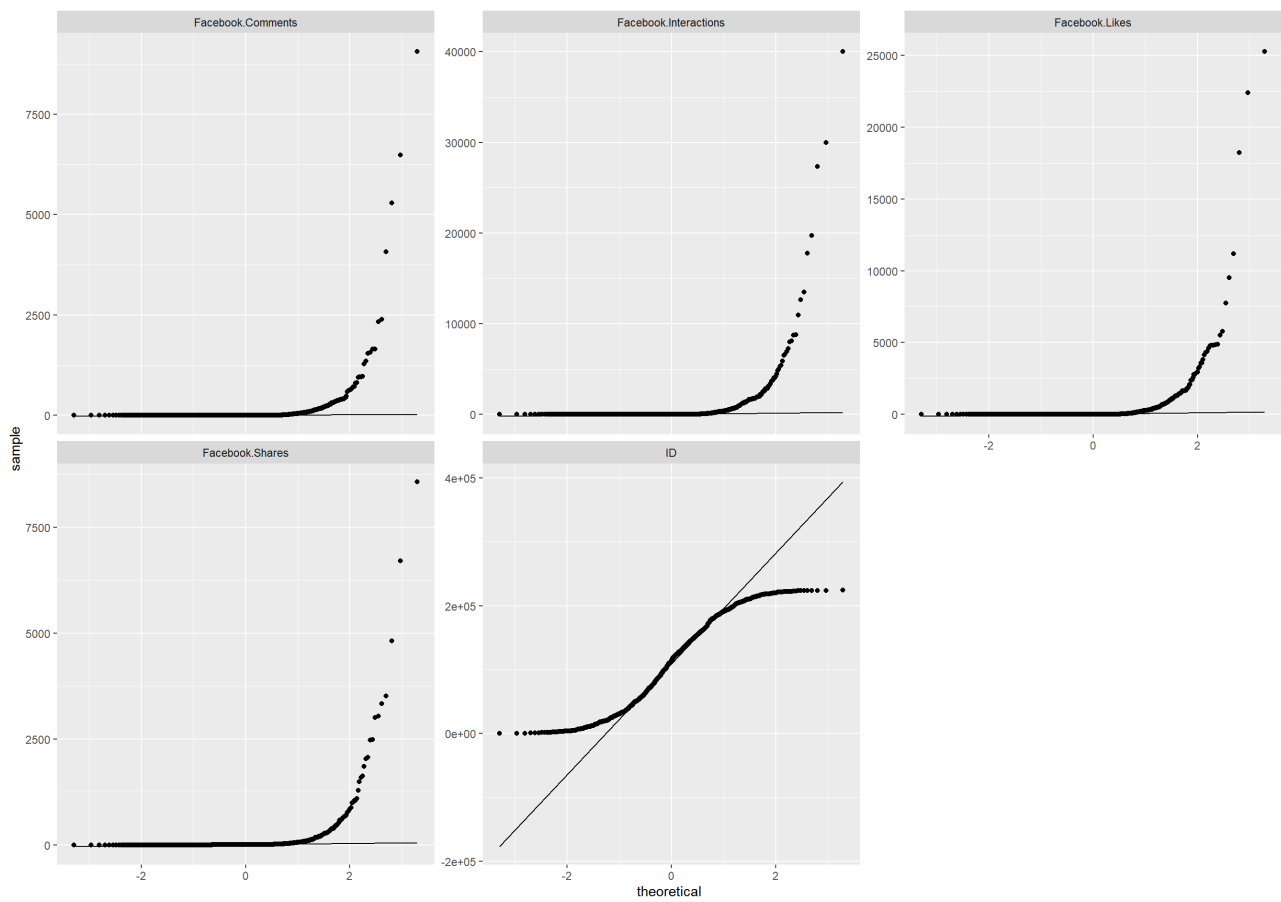


Bar Chart (with frequency)

```
## 5 columns ignored with more than 50 categories.  
## Published: 129618 categories  
## Headline: 213103 categories  
## Link: 224349 categories  
## Date: 367 categories  
## Cleaned: 208444 categories
```

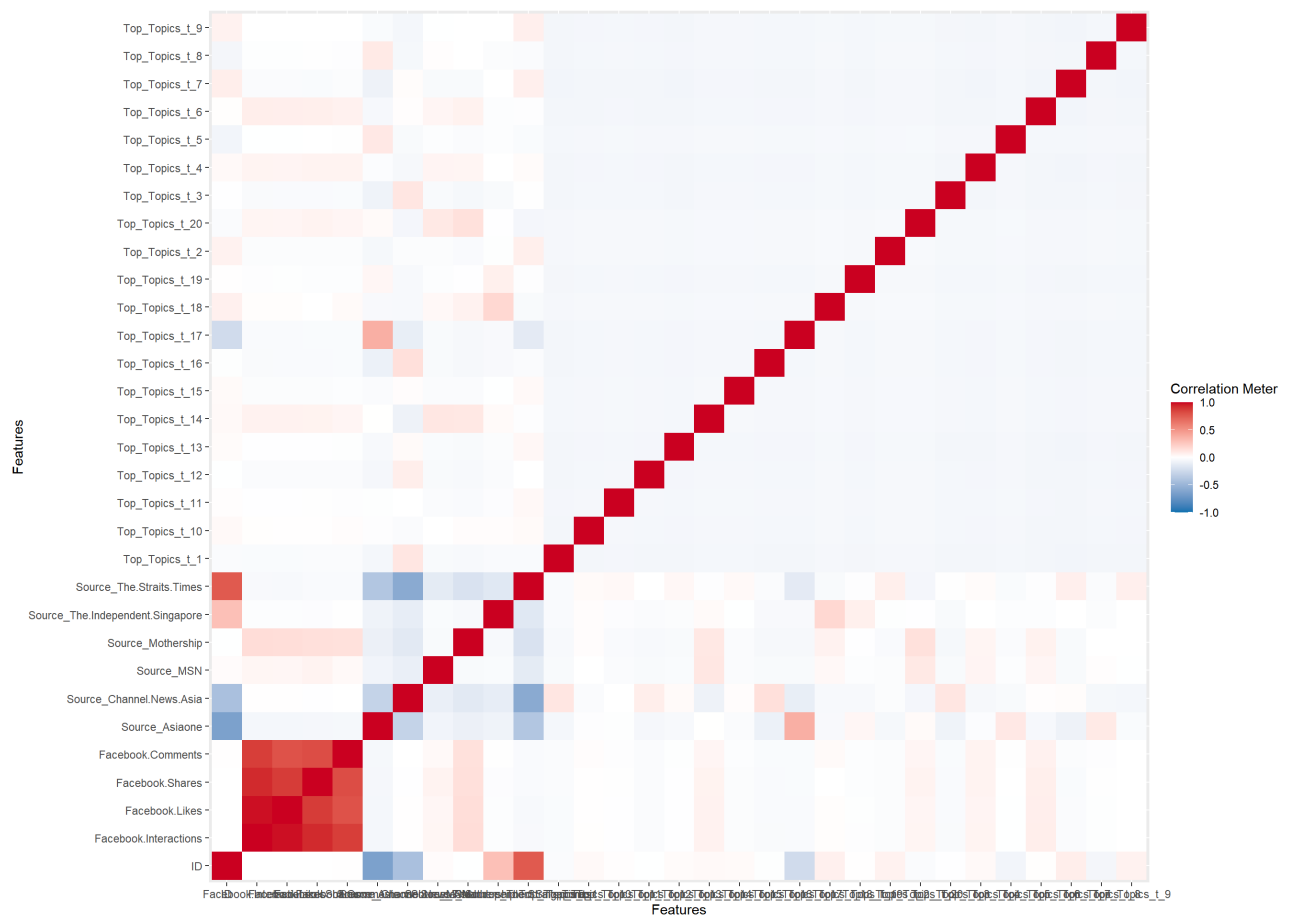


QQ Plot



## Correlation Analysis

```
## 5 features with more than 20 categories ignored!  
## Published: 129617 categories  
## Headline: 213100 categories  
## Link: 224338 categories  
## Date: 367 categories  
## Cleaned: 208443 categories
```



Principal Component Analysis

```
## 5 features with more than 50 categories ignored!  
## Published: 129617 categories  
## Headline: 213100 categories  
## Link: 224338 categories  
## Date: 367 categories  
## Cleaned: 208443 categories
```

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

