

# Data Profiling Report

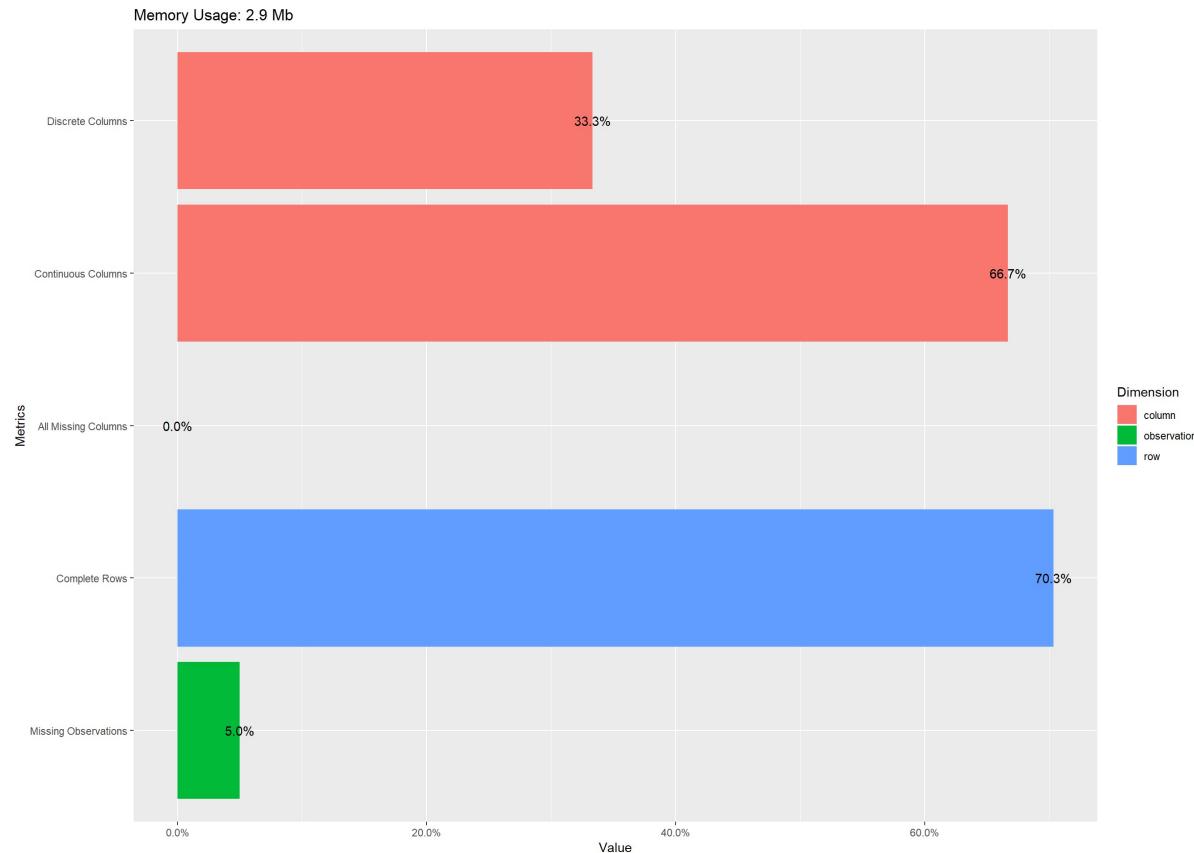
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (by frequency)
  - Bar Chart (by price)
  - QQ Plot
  - QQ Plot (by price)
- Correlation Analysis
- Bivariate Distribution
  - Boxplot (by price)
  - Scatterplot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

## Basic Statistics

### Raw Counts

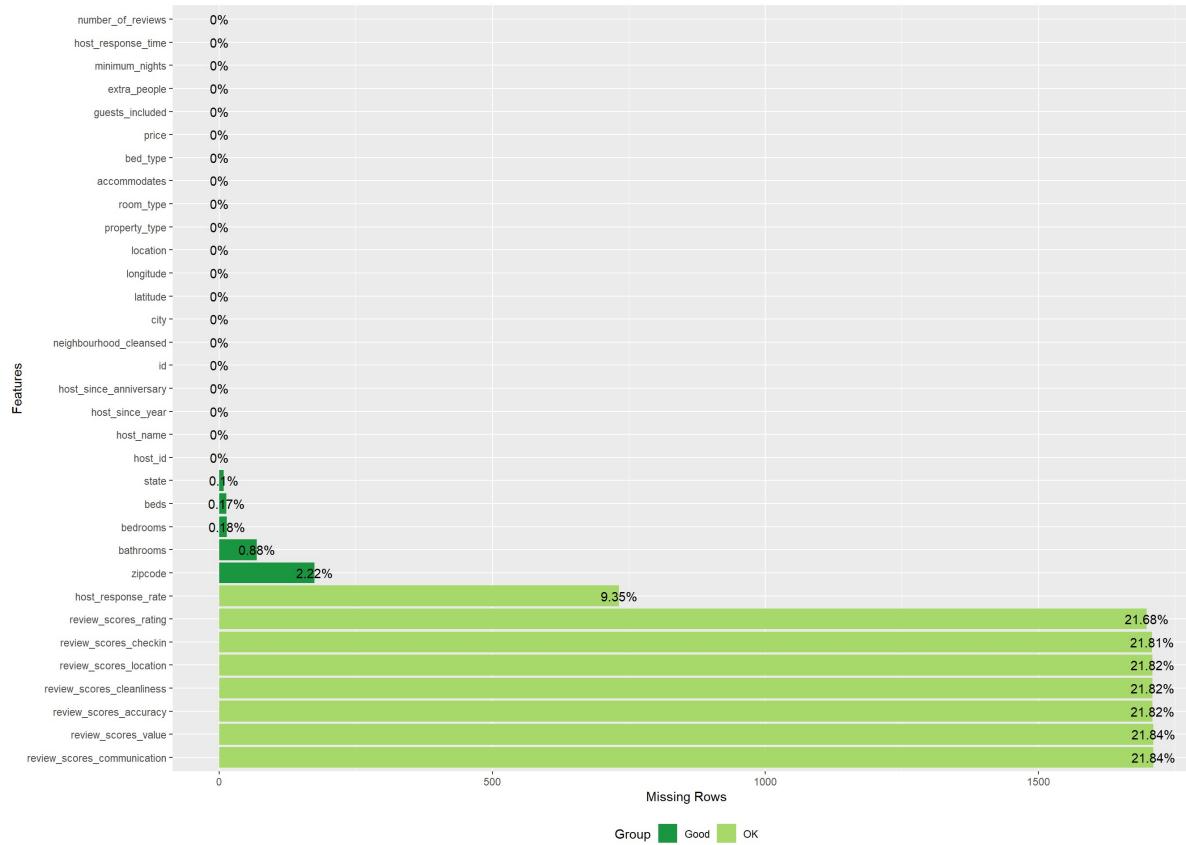
Name	Value
Rows	7,833
Columns	33
Discrete columns	11
Continuous columns	22
All missing columns	0
Missing observations	12,965
Complete Rows	5,508
Total observations	258,489
Memory allocation	2.9 Mb

### Percentages



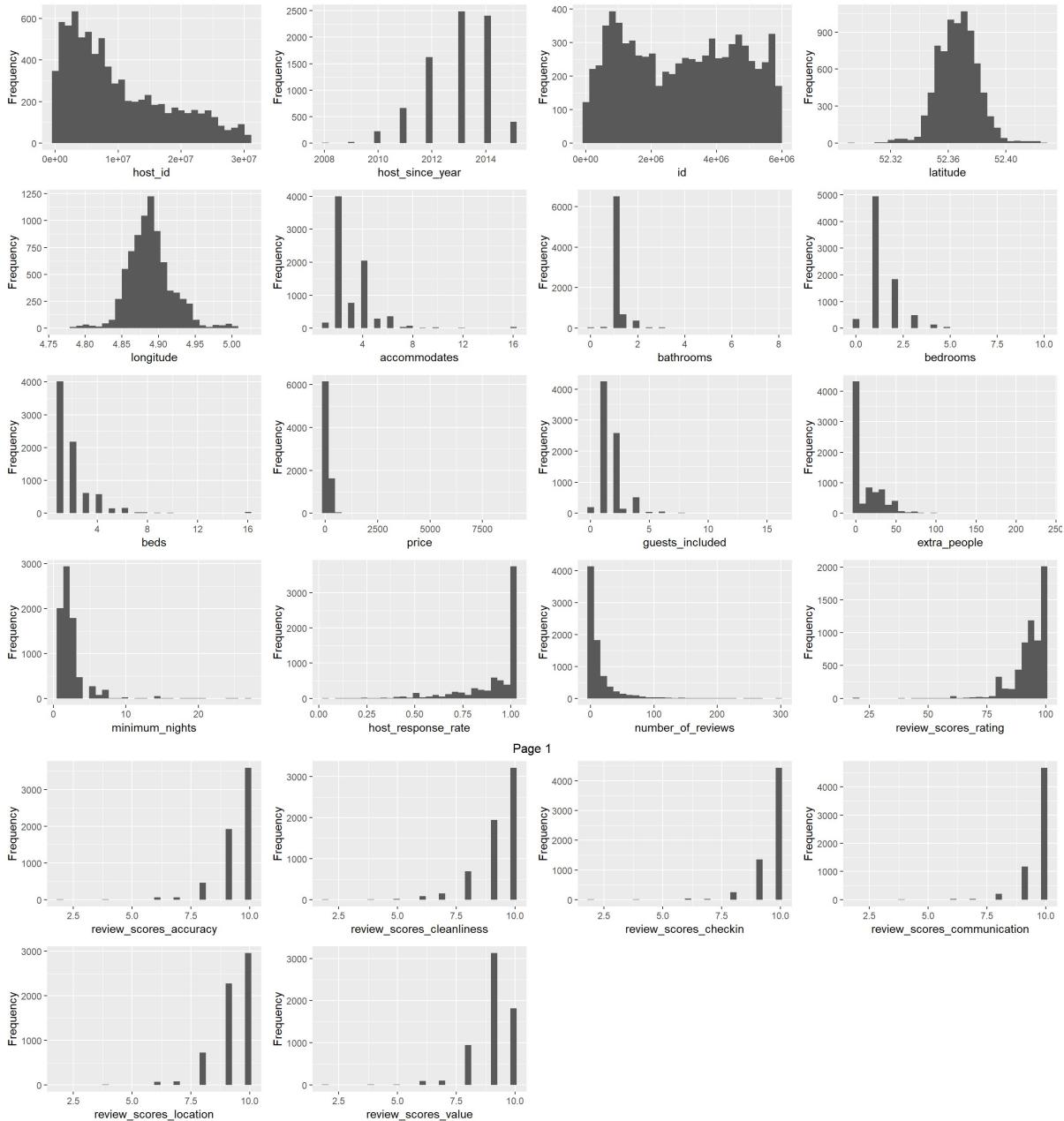
## Data Structure

Missing Data Profile



## Univariate Distribution

### Histogram

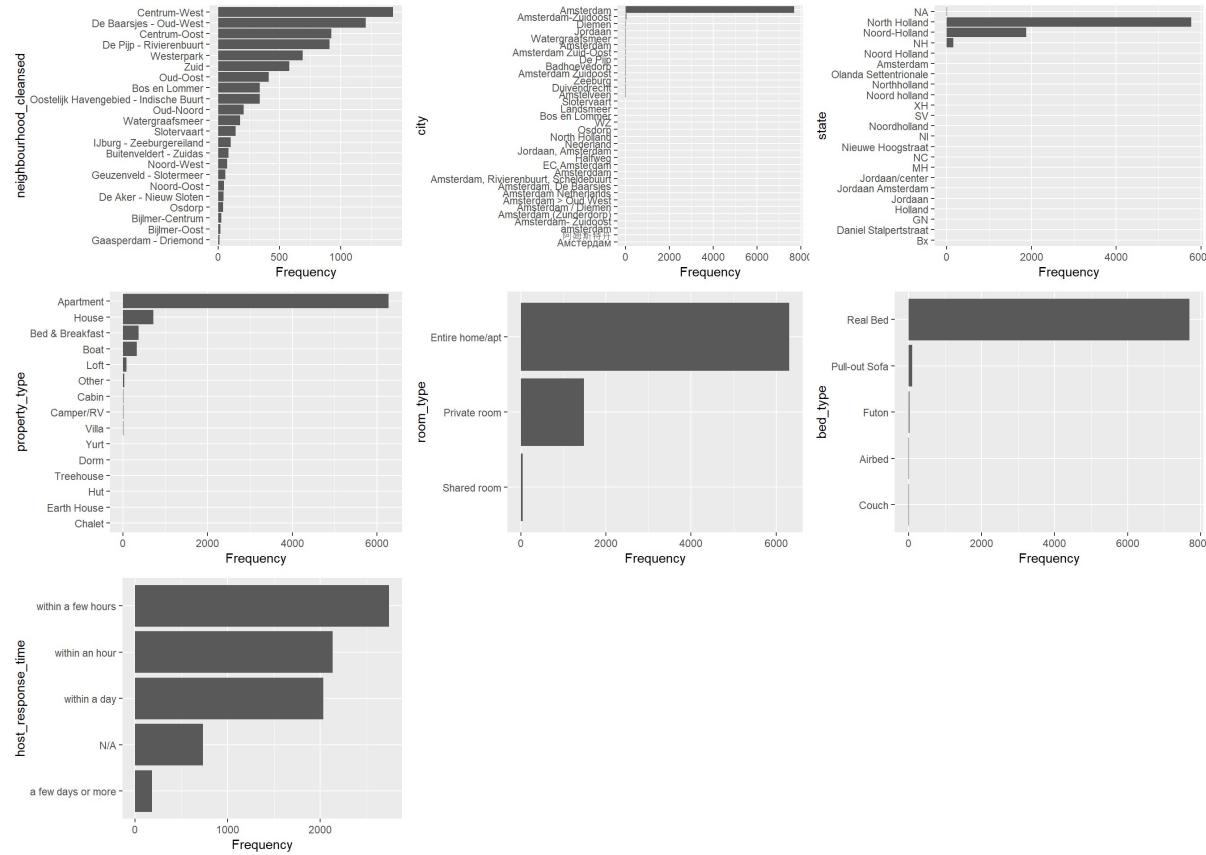


Page 1

Page 2

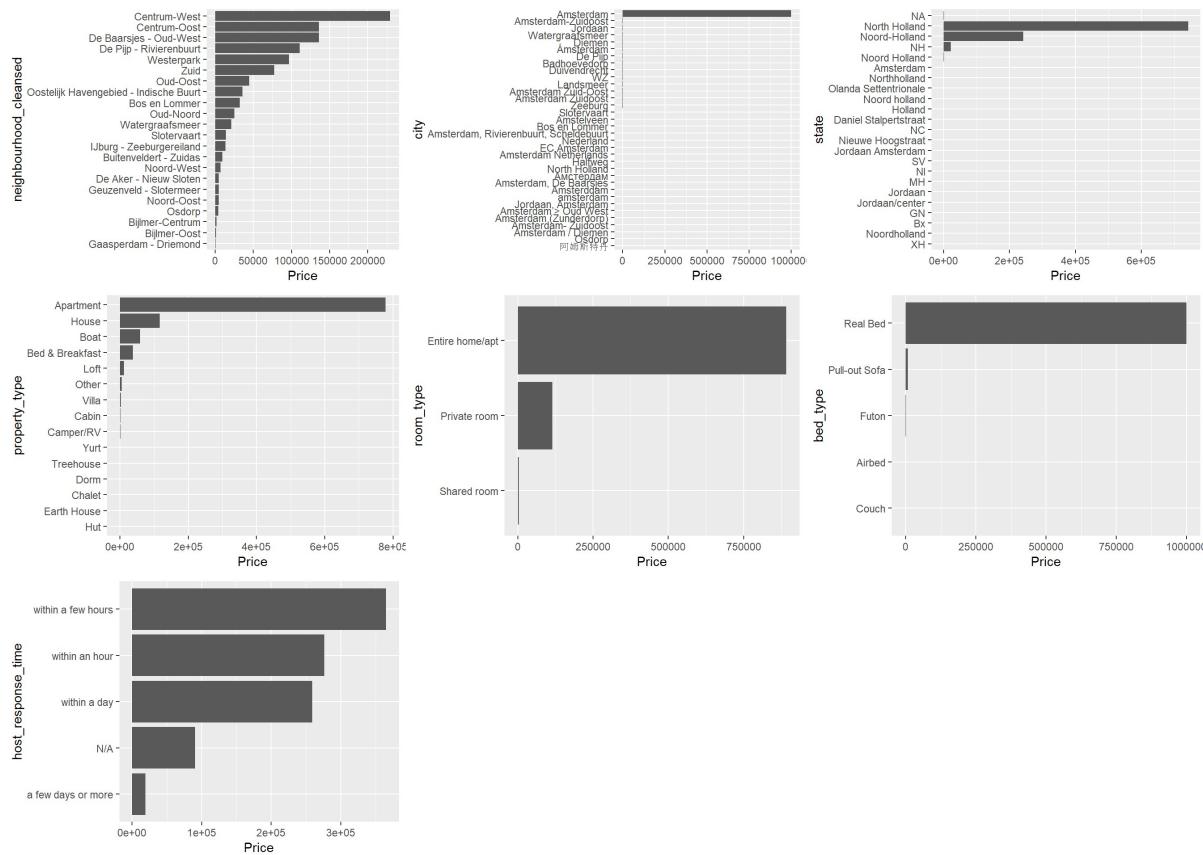
### Bar Chart (by frequency)

```
## 4 columns ignored with more than 50 categories.
## host_name: 2987 categories
## host_since_anniversary: 366 categories
## zipcode: 3275 categories
## location: 7813 categories
```

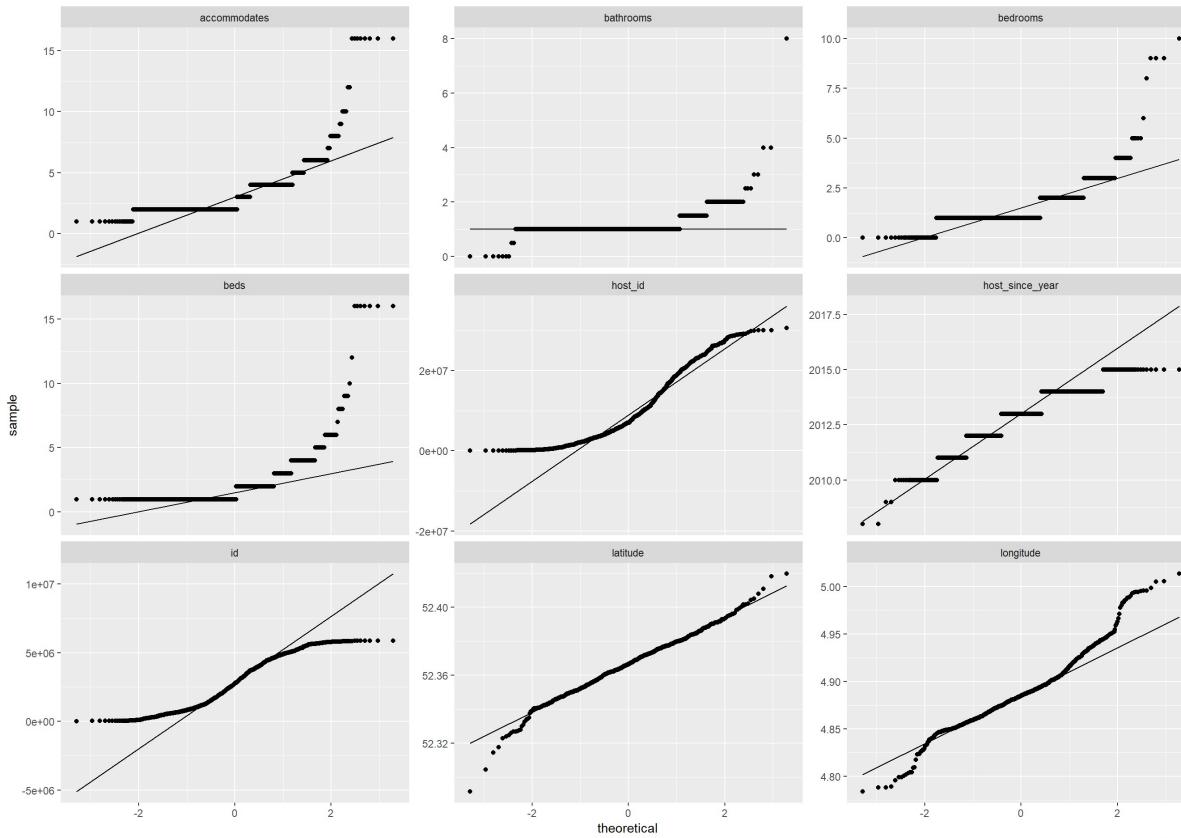


### Bar Chart (by price)

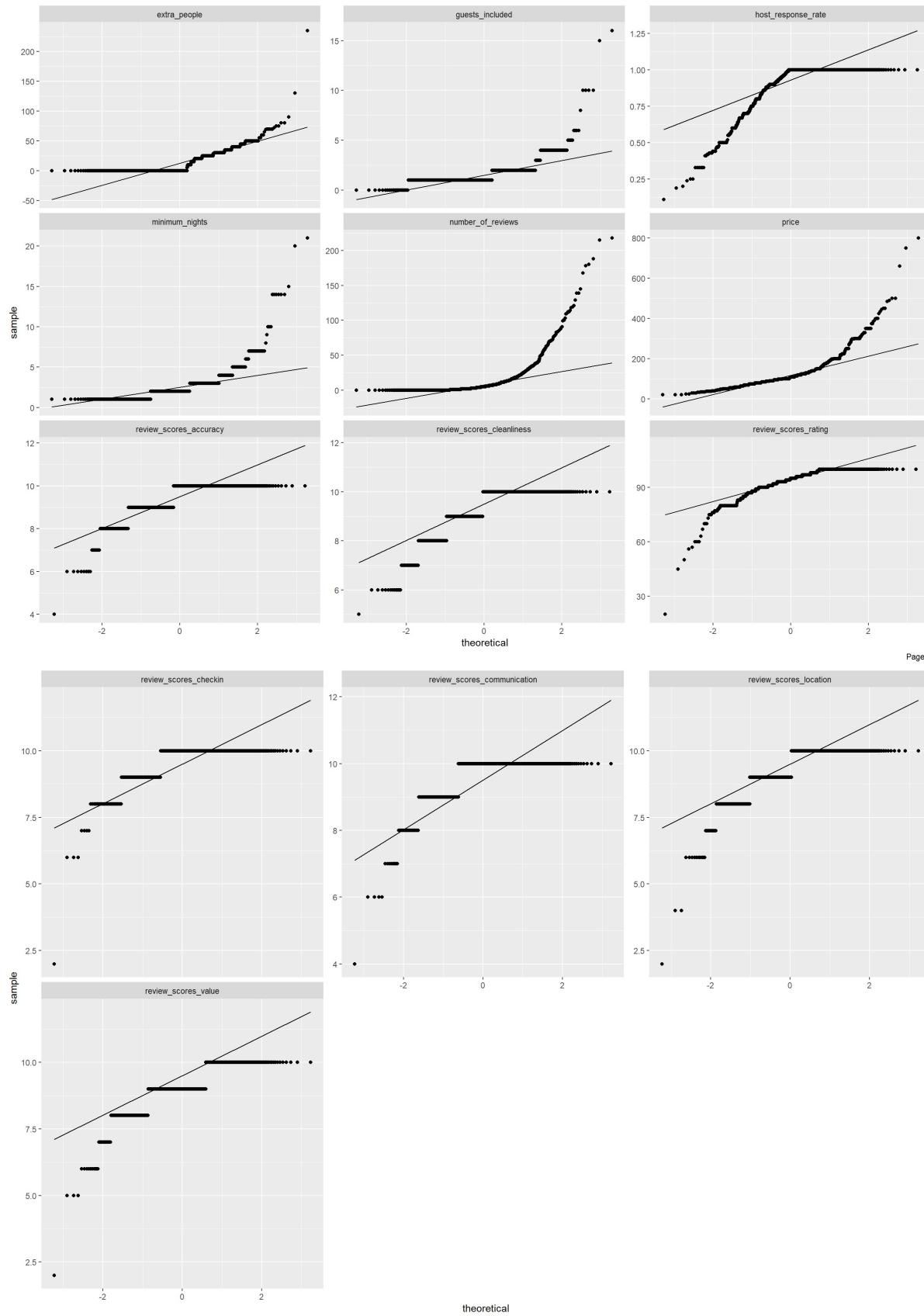
```
## 4 columns ignored with more than 50 categories.
## host_name: 2987 categories
## host_since_anniversary: 366 categories
## zipcode: 3275 categories
## location: 7813 categories
```



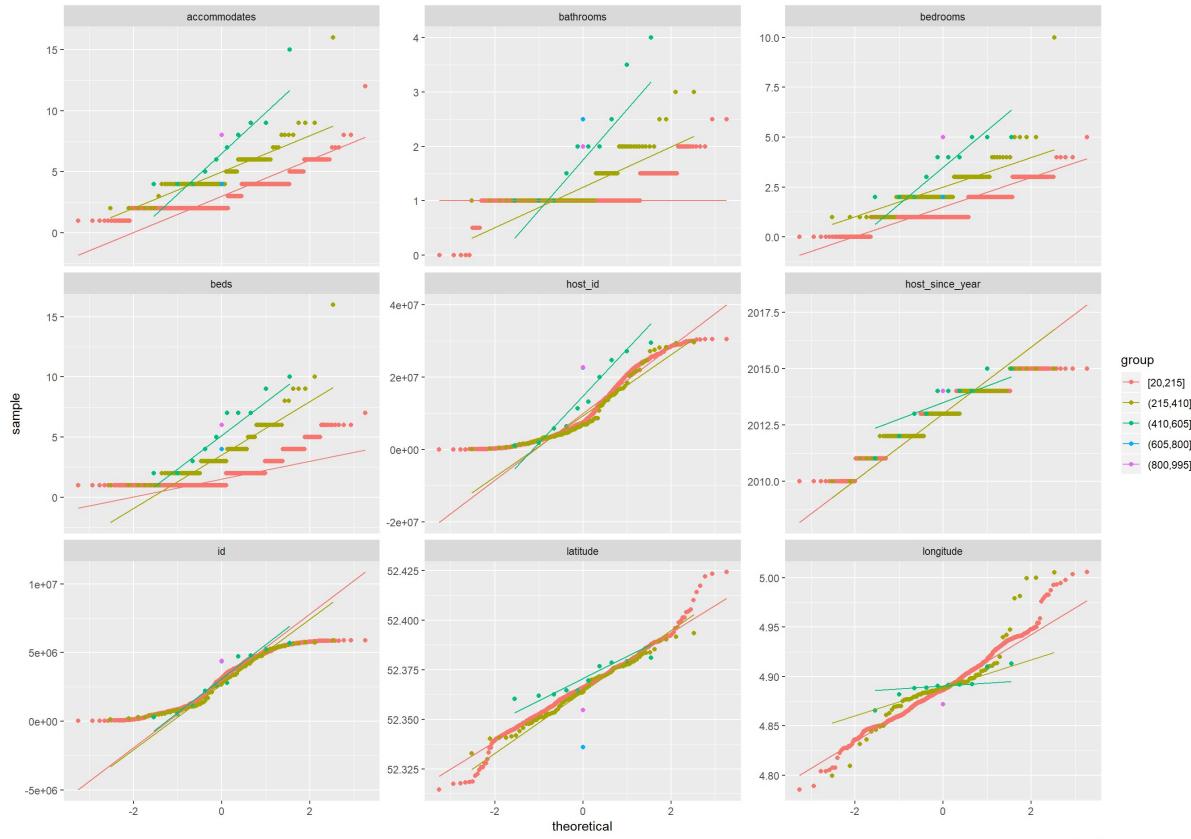
## QQ Plot



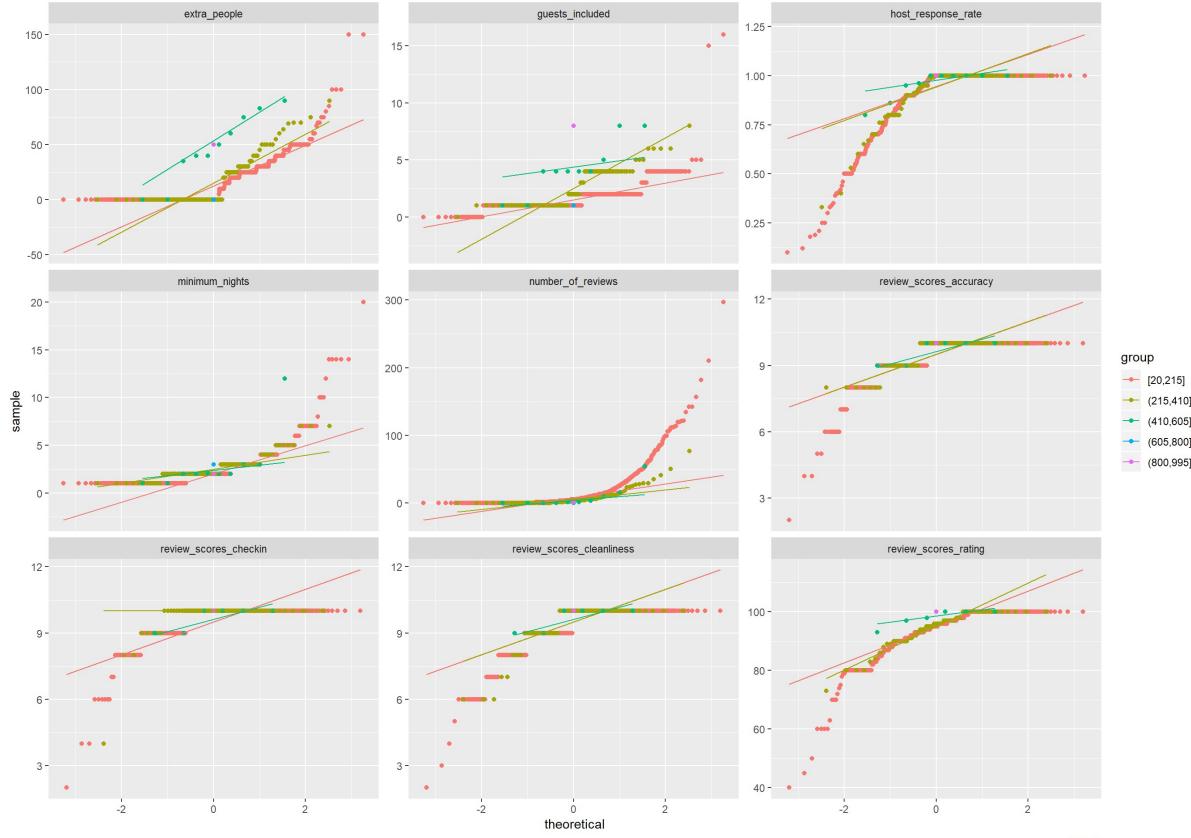
Page 1



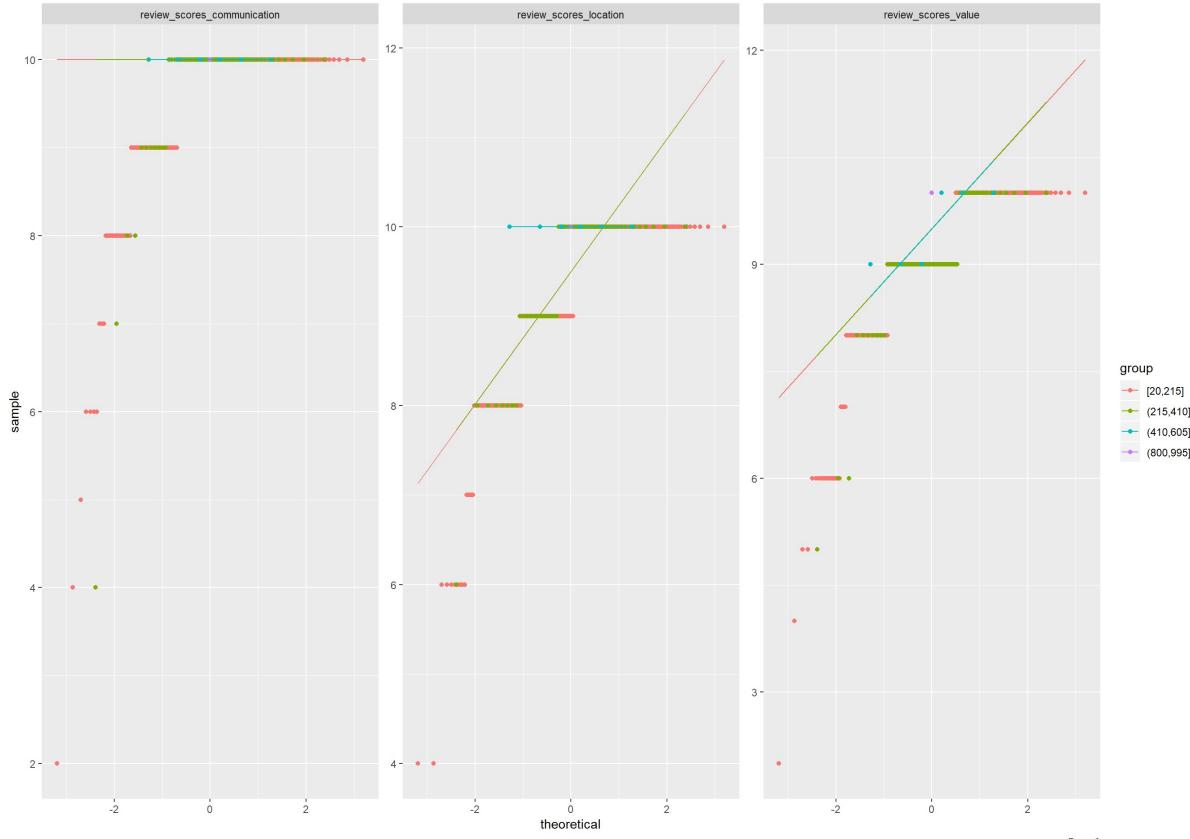
QQ Plot (by price)



Page 1



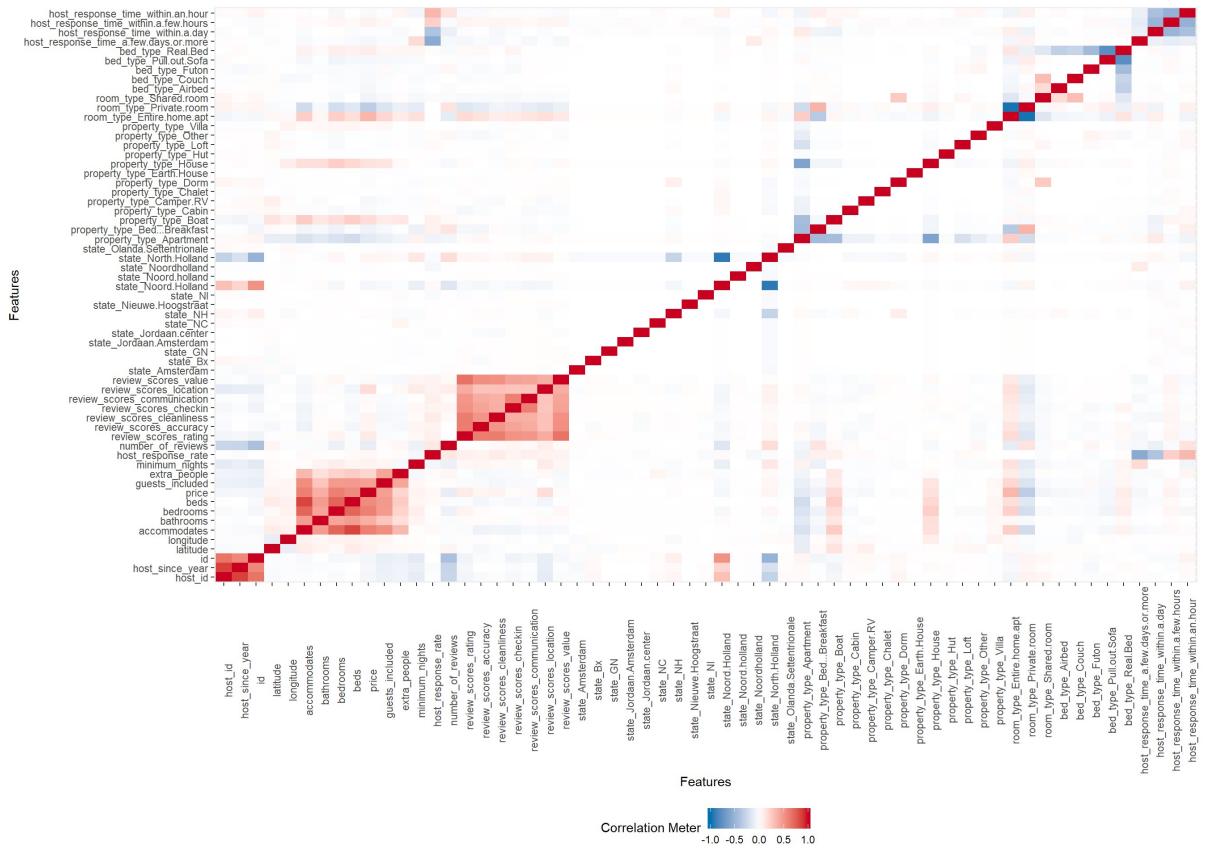
Page 2



Page 3

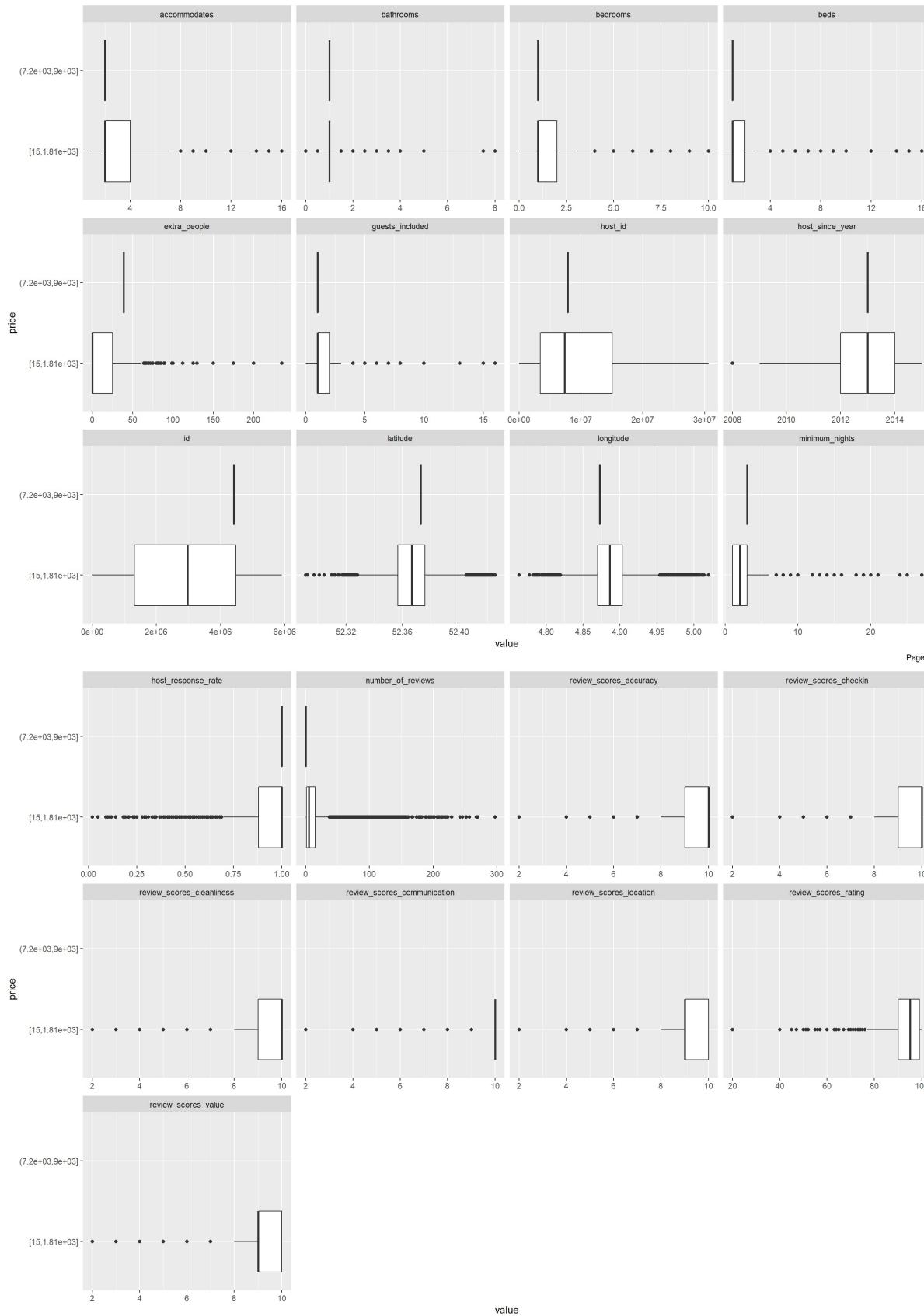
## Correlation Analysis

```
## 6 features with more than 20 categories ignored!
## host_name: 2389 categories
## host_since_anniversary: 366 categories
## neighbourhood_cleansed: 22 categories
## city: 28 categories
## zipcode: 2456 categories
## location: 5493 categories
```

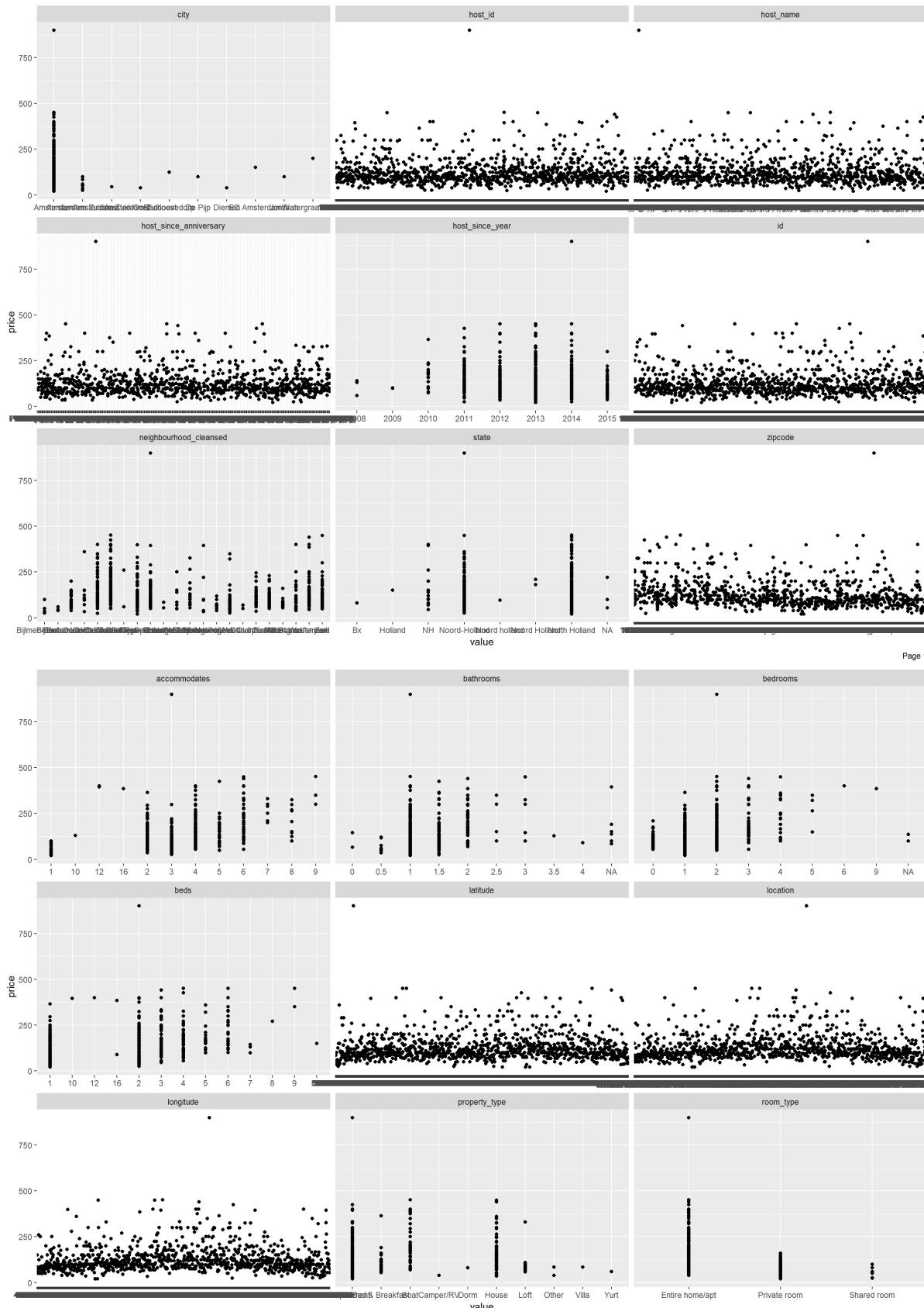


## Bivariate Distribution

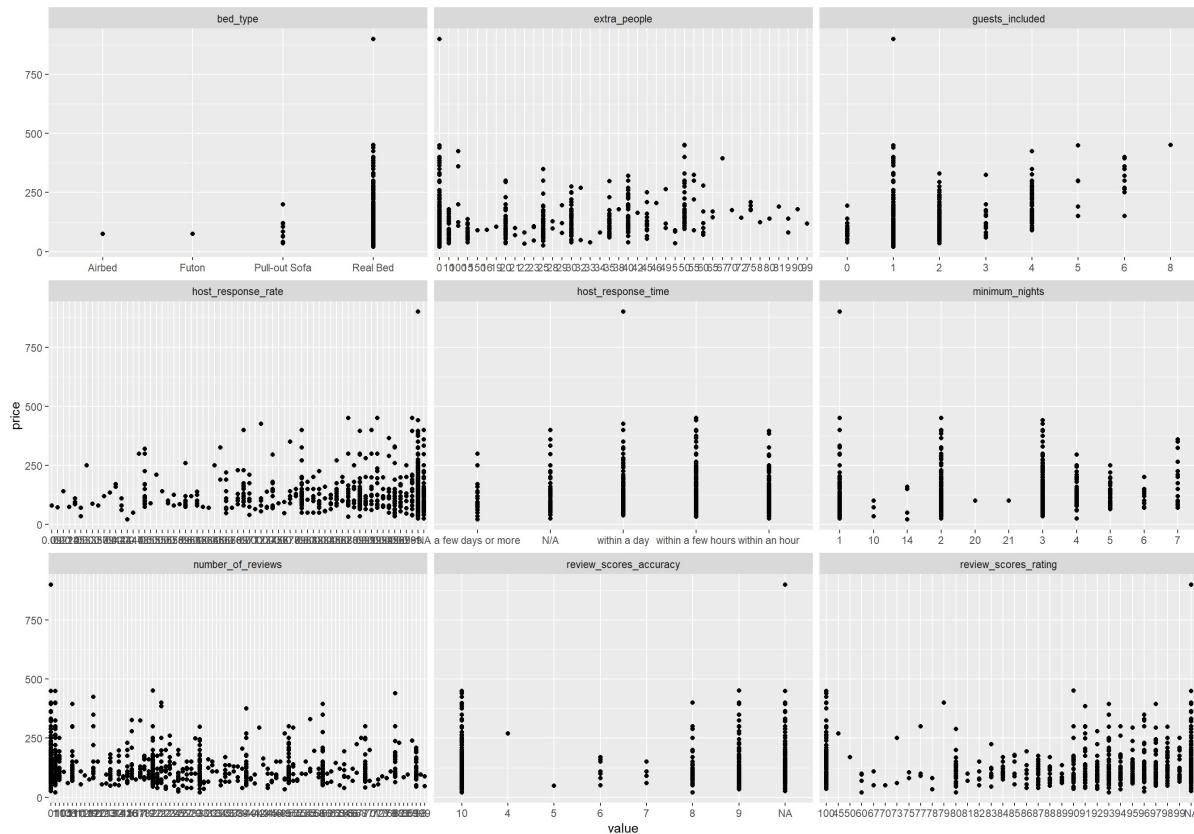
Boxplot (by price)



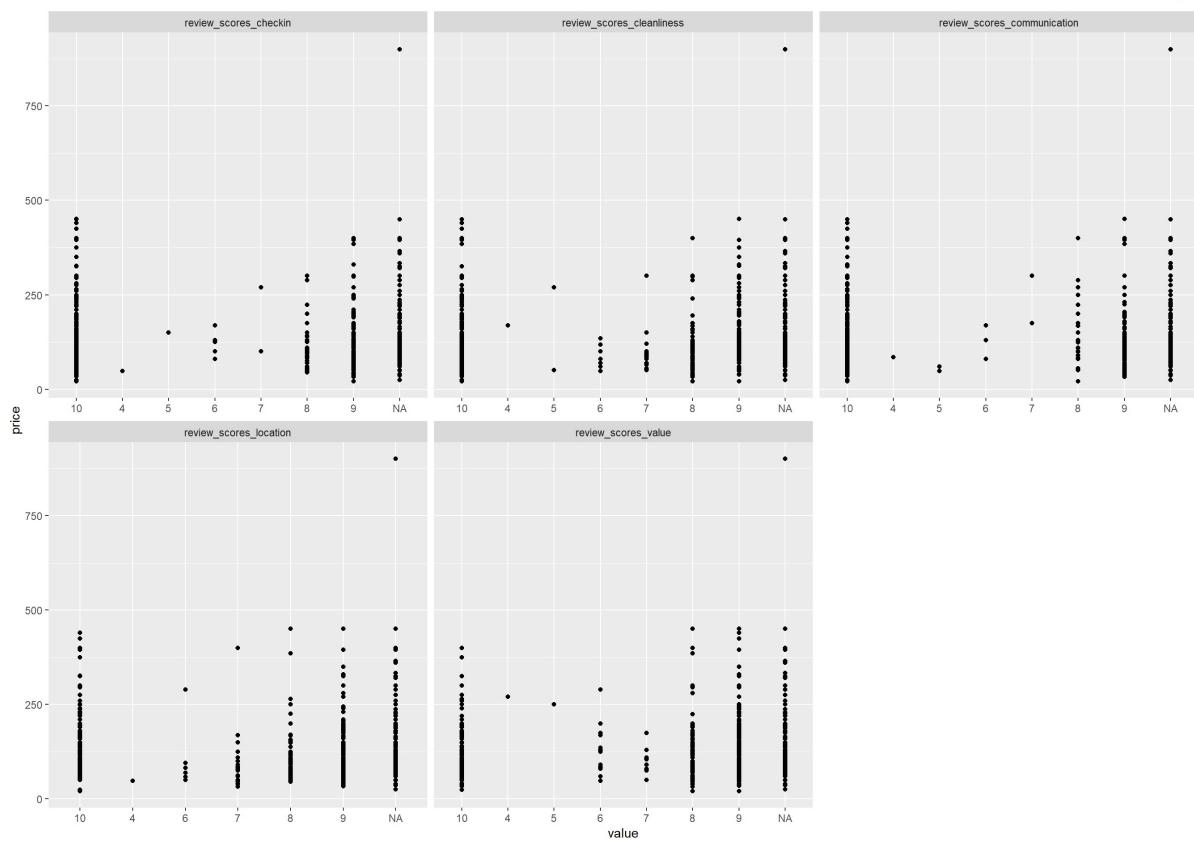
Scatterplot (by price)



Page 2



Page 3



Page 4