COURSE PROJECT

Taylor & Francis Rmarkdown template for authors (LaTeX-based **Interact** layout + Chicago author-date reference style)

Beining Wu[1]

[a]Department of Statistics and Finance, University of Science and Technology of China, Hefei, 230026

ARTICLE HISTORY
Compiled December 31, 2021

ABSTRACT
Remains

KEYWORDS
Remains

1. Preliminaries

```
data <- read.csv(file = "data.csv", head=TRUE, fileEncoding = "UTF8")
attach(data)
library(mapchina)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(sf)
```

```
## Linking to GEOS 3.9.1, GDAL 3.2.3, PROJ 7.2.1; sf_use_s2() is TRUE
```

CONTACT Beining Wu. Email: andrewwu@mail.ustc.edu.cn

```
library(fitdistrplus)
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
## Loading required package: survival
```

2. Data Overview

First we take a glance at the distrbution of the `rent` variable. d‘

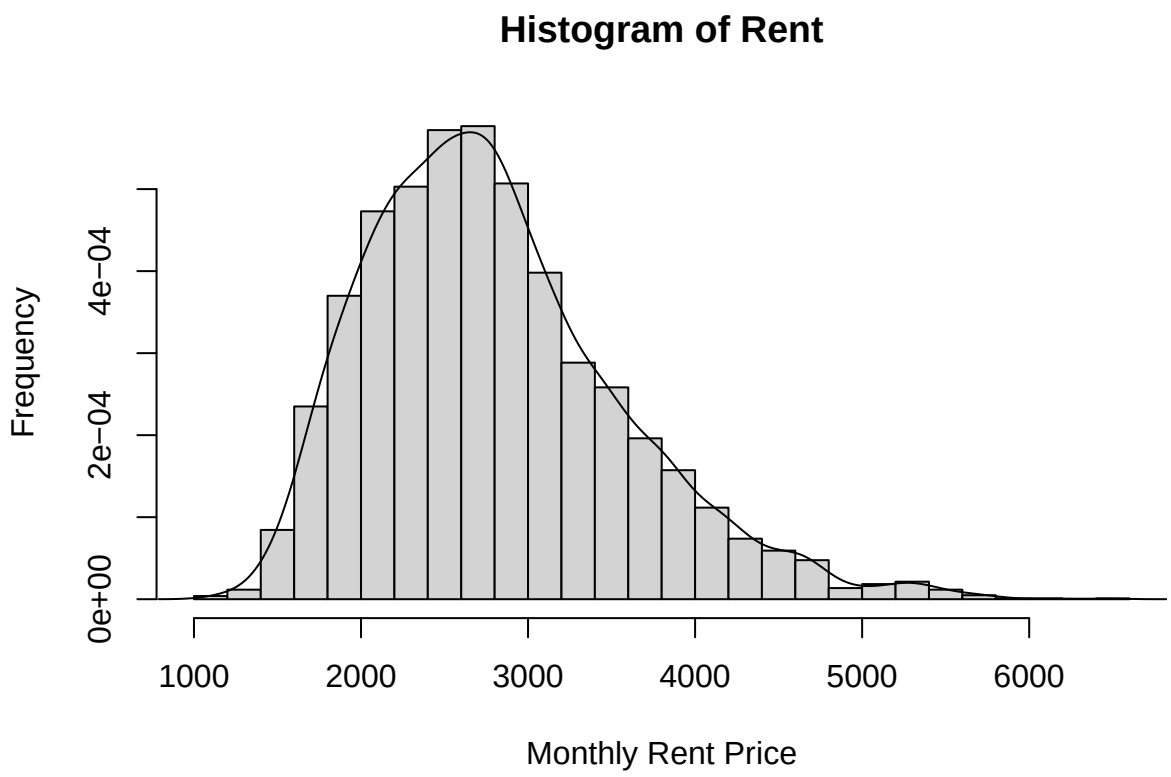
```
head(data)
```

```
##   rent bedroom livingroom bathroom area room floor_grp subway region  heating
## 1 2730       2          1        1   12 主卧     高楼层     是   通州 集中供暖
## 2 2740       3          1        1    9 次卧     低楼层     是   昌平 集中供暖
## 3 2810       3          1        1   14 主卧     低楼层     是   丰台 集中供暖
## 4 2650       4          1        1    8 次卧     低楼层     是   丰台 集中供暖
## 5 2670       4          1        1   13 主卧     高楼层     否   丰台 集中供暖
## 6 2530       3          1        1   12 次卧     高楼层     是   顺义 集中供暖
```

```
summary(data)
```

```
##       rent          bedroom        livingroom      bathroom          area
##  Min.   :1150   Min.   :2.000   Min.   :1.00   Min.   :1.000   Min.   : 5.00
##  1st Qu.:2240   1st Qu.:2.000   1st Qu.:1.00   1st Qu.:1.000   1st Qu.:10.00
##  Median :2690   Median :3.000   Median :1.00   Median :1.000   Median :12.00
##  Mean   :2798   Mean   :2.996   Mean   :1.01   Mean   :1.027   Mean   :12.85
##  3rd Qu.:3230   3rd Qu.:4.000   3rd Qu.:1.00   3rd Qu.:1.000   3rd Qu.:15.00
##  Max.   :6460   Max.   :5.000   Max.   :2.00   Max.   :2.000   Max.   :30.00
##      room            floor_grp           subway             region
##  Length:5149        Length:5149        Length:5149        Length:5149
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    heating
##  Length:5149
##  Class :character
##  Mode  :character
##
##
##
```

```
hist(rent, breaks = "scott", main = "Histogram of Rent", xlab = "Monthly Rent Price", yla
lines(density(rent))
```



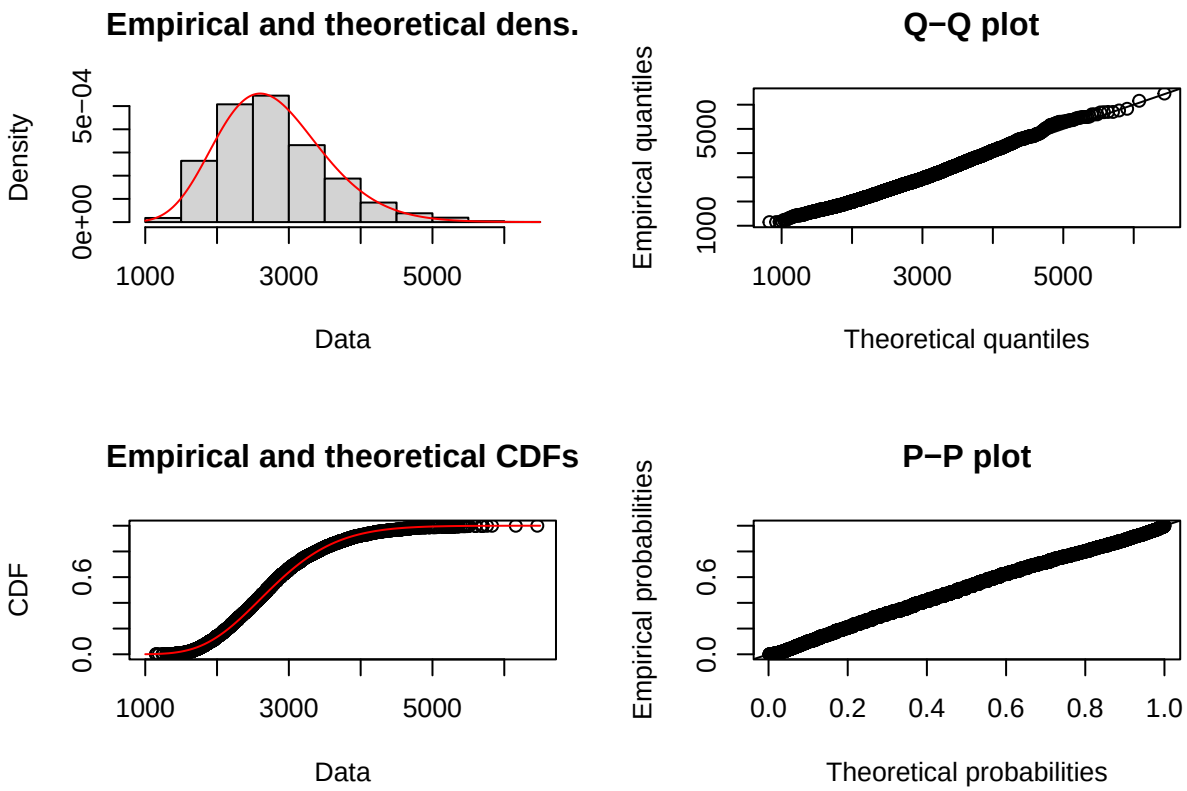
3. Fitting

From the histogram plot we may hypothesis that the distribution is gamma.

```
library(fitdistrplus)
library(MASS)
fit.gamma <- fitdist(rent, distr = "gamma", method = "mle")
summary(fit.gamma)
```

```
## Fitting of the distribution ' gamma ' by maximum likelihood
## Parameters :
##           estimate   Std. Error
## shape 14.212371899 0.1739956485
## rate   0.005080088 0.0000607289
## Loglikelihood:  -41215.48   AIC:  82434.97   BIC:  82448.06
## Correlation matrix:
##           shape      rate
## shape 1.0000000 0.9549845
## rate  0.9549845 1.0000000
```

```
plot(fit.gamma)
```



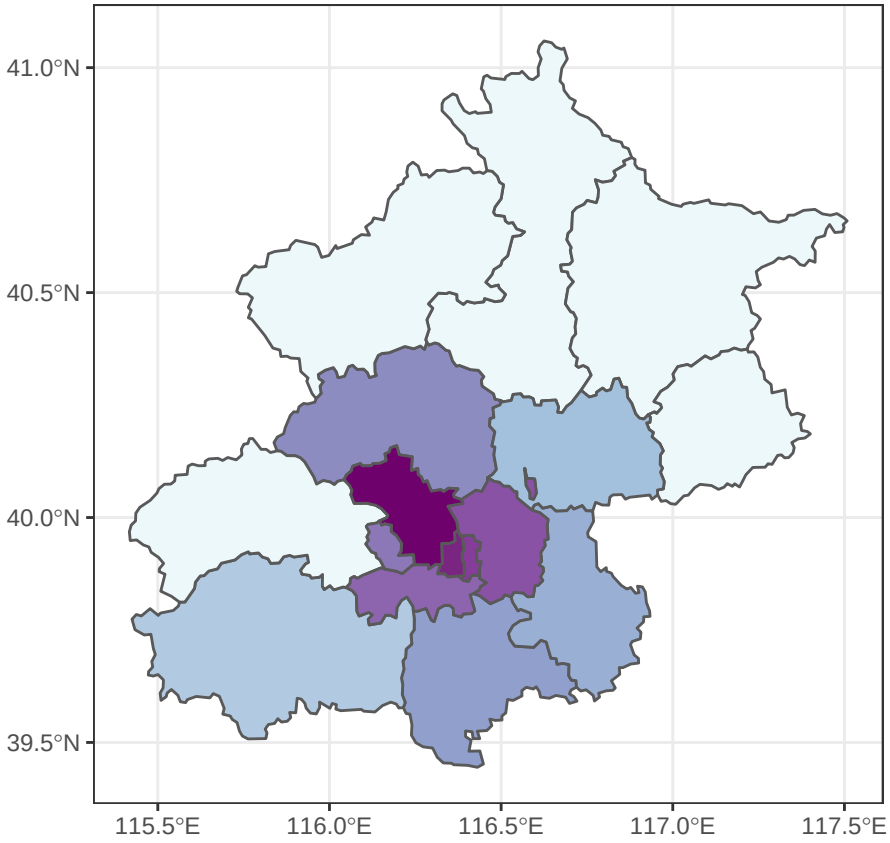
4. Mean Unit Value in Each District

Now we consider the mean rent in each district.

```
unit <- data$rent/data$area

mean_unit_reg <- c(mean(unit[region==" 东城"]),mean(unit[region==" 西城"]),mean(unit[regi


mean_unit_reg[is.na(mean_unit_reg)]=0
```

And we can visualize as follows.

```
df <- china %>%
        filter(Name_Province == " 北京市")

ggplot(data = df) +
        geom_sf(aes(fill = rank(mean_unit_reg))) +
        scale_fill_distiller(palette = "BuPu", direction = 1) +
        theme_bw() +
        theme(legend.position = "none")
```



From the map we know that the unit price is significantly **higher** when location is more central. Now the

5. Mean Unit Value and Subway

```
boxplot(rent~subway, main="Boxplot of Rent, Grouing by Subway")
```

```
## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
## substituted for <e5>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
## substituted for <90>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
## substituted for <a6>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
## substituted for <e5>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
```

```
## substituted for <90>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '否' in 'mbcsToSbcs': dot
## substituted for <a6>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <e6>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <98>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <af>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <e6>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <98>

## Warning in axis(side = base::quote(1), at = base::quote(1:2), labels =
## base::quote(c("否", : conversion failure on '是' in 'mbcsToSbcs': dot
## substituted for <af>
```