

# First Homework

Legolas

2023-10-19

1. 1
2. 2
3. 3

2023 9 12

- 100 30 3000
- 

```
lj1<- read_csv("2023-09-12_cleaned.csv")
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)
```

```
## Rows: 3000 Columns: 18
## -- Column specification -----
## Delimiter: ","
## chr (11): property_name, property_region, directions1, directions2, decorati...
## dbl (7): price_ttl, price_sqm, bedrooms, livingrooms, building_area, proper...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
theme_set(theme(text = element_text(family="",size = 10)))
summary(cars)
```

```
##      speed      dist
## Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean    : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.    :120.00
```

```
(lj1) property_name, property_region, price_ttl, price_sqm, bedrooms, livingrooms, building_area,
directions1, directions2, decoration, property_t_height, property_height, property_style, followers,
near_subway, if_2y, has_key, vr 18 , 3000 10
```

```
glimpse(lj1)
```

```
## Rows: 3,000
## Columns: 18
## $ property_name      <chr> " A ", " ", " ", " ", " ", " ", "~
## $ property_region    <chr> " ", " ", " ", " ", " ", " ", "~
## $ price_ttl          <dbl> 237.0, 127.0, 75.0, 188.0, 182.0, 122.0, 99.0, 193.8~
## $ price_sqm          <dbl> 18709, 14613, 15968, 15702, 17509, 10376, 12346, 163~
## $ bedrooms          <dbl> 3, 3, 1, 3, 3, 3, 2, 3, 4, 3, 5, 3, 4, 3, 3, 2, 3, 4~
## $ livingrooms        <dbl> 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2~
## $ building_area      <dbl> 126.68, 86.91, 46.97, 119.73, 103.95, 117.59, 80.19, ~
## $ directions1       <chr> " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", "~
## $ directions2       <chr> " ", NA, NA, " ", NA, " ", NA, " ", " ", " ", " ", ~
## $ decoration         <chr> " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", "~
## $ property_t_height  <dbl> 17, 28, 18, 32, 34, 34, 7, 34, 5, 7, 25, 32, 8, 31, ~
## $ property_height    <chr> " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", "~
## $ property_style     <chr> " ", " ", " ", " ", " ", " ", " ", " ", " ", " ", "~
## $ followers          <dbl> 3, 1, 3, 2, 3, 1, 0, 0, 2, 0, 0, 0, 10, 0, 0, 1, 0, ~
## $ near_subway        <chr> " ", NA, " ", " ", " ", NA, NA, " ", " ", ~
## $ if_2y              <chr> NA, " ", " ", NA, " ", " ", " ", " ", "~
## $ has_key            <chr> " ", " ", " ", " ", " ", " ", " ", " ", "~
## $ vr                <chr> NA, "VR ", NA, NA, "VR ", NA, "VR ", ~
```

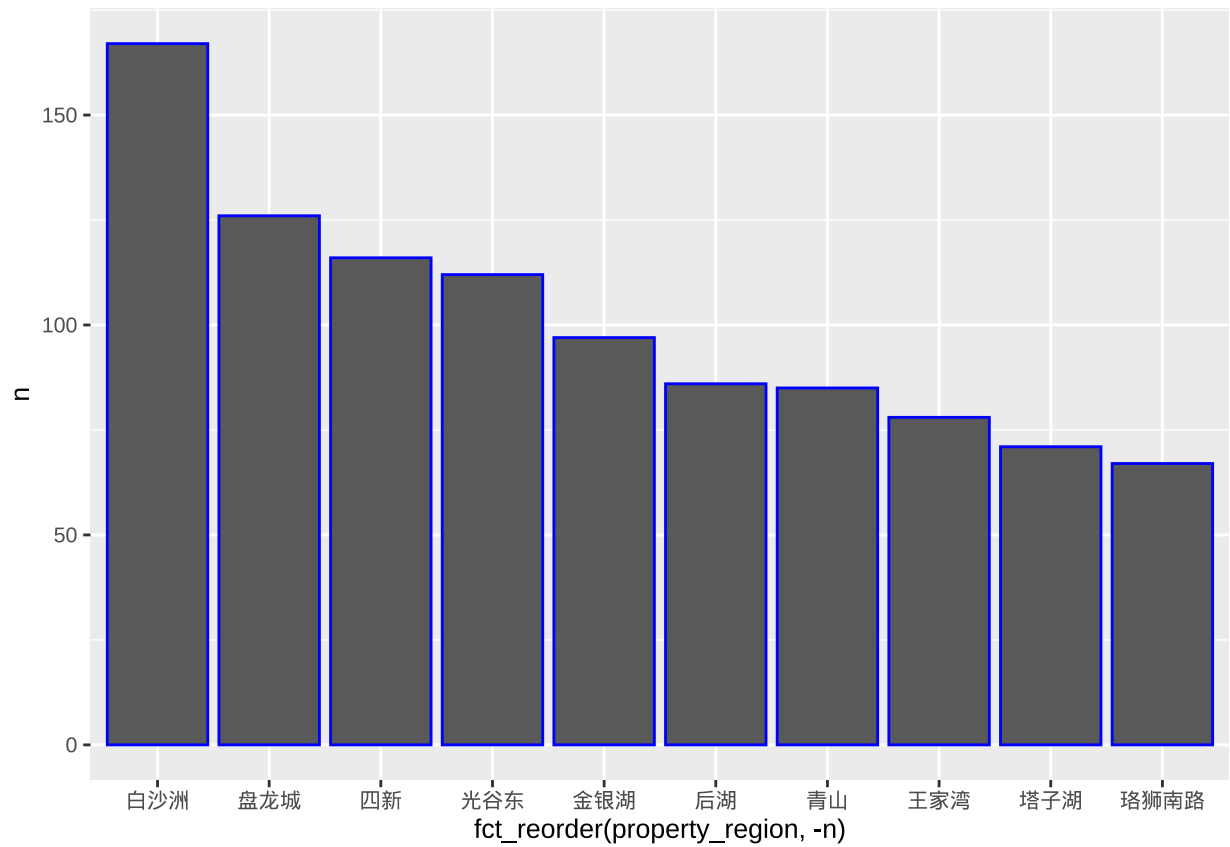
```
summary(lj1)
```

```
## property_name      property_region      price_ttl      price_sqm
## Length:3000      Length:3000      Min.   : 10.6   Min.   : 1771
## Class :character  Class :character  1st Qu.: 95.0   1st Qu.:10799
## Mode  :character  Mode  :character  Median : 137.0  Median :14404
##                                     Mean   : 155.9   Mean   :15148
##                                     3rd Qu.: 188.0  3rd Qu.:18211
##                                     Max.   :1380.0  Max.   :44656
## bedrooms          livingrooms      building_area  directions1
## Min.   :1.000     Min.   :0.000   Min.   : 22.77  Length:3000
## 1st Qu.:2.000     1st Qu.:1.000   1st Qu.: 84.92  Class :character
## Median :3.000     Median :2.000   Median : 95.55  Mode  :character
## Mean   :2.695     Mean   :1.709   Mean   :100.87
## 3rd Qu.:3.000     3rd Qu.:2.000   3rd Qu.:117.68
## Max.   :7.000     Max.   :4.000   Max.   :588.66
## directions2       decoration      property_t_height property_height
## Length:3000      Length:3000      Min.   : 2.00   Length:3000
## Class :character  Class :character  1st Qu.:11.00   Class :character
## Mode  :character  Mode  :character  Median :27.00   Mode  :character
##                                     Mean   :24.22
##                                     3rd Qu.:33.00
##                                     Max.   :62.00
## property_style     followers      near_subway      if_2y
```

```
## Length:3000      Min.   : 0.000      Length:3000      Length:3000
## Class :character 1st Qu.: 1.000      Class :character Class :character
## Mode  :character Median : 3.000      Mode  :character Mode  :character
##                  Mean   : 6.614
##                  3rd Qu.: 6.000
##                  Max.   :262.000
##   has_key          vr
## Length:3000      Length:3000
## Class :character Class :character
## Mode  :character Mode  :character
##
##
##
```

- 1 11 7
- 2 10.6 1380 155.9 137
- 3

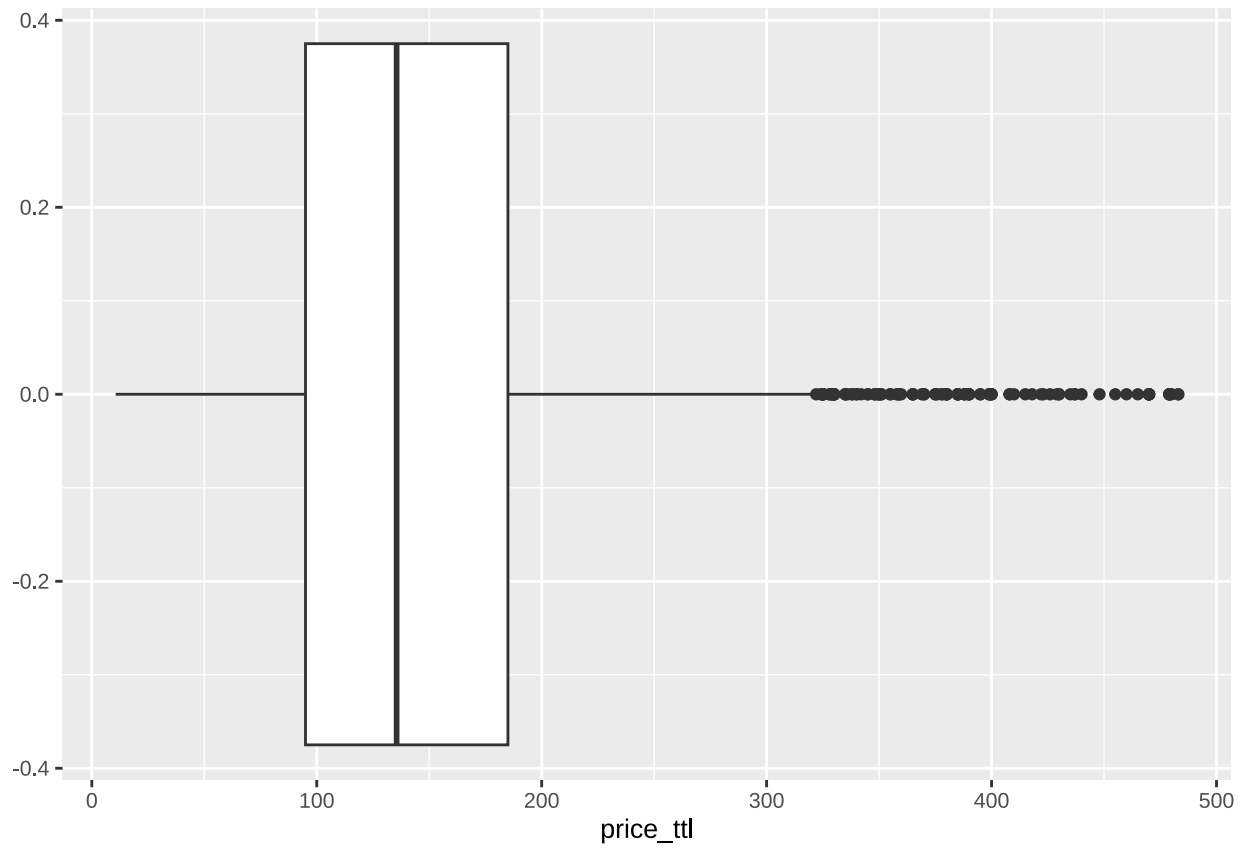
```
lj1 %>%
  group_by(property_region)%>%
  summarise(n=n()) %>%
  arrange(desc(n))%>%
  head(10)%>%
  ggplot(aes(fct_reorder(property_region,-n),n))+
  geom_col(color="blue")
```



1

- 1
- 2

```
lj_filtered<- lj1%>%
  filter(lj1[["price_ttl"]]<=500)
ggplot(lj_filtered,aes(x=price_ttl))+geom_boxplot()
```



2

- 1
- 2

...

- 1
- 2