

lj_investigate

wangsy

2025-09-15

One:Main Conclusions

This chart is extracted from the Lianjia database, showing the distribution of a scatter plot between the age of properties and their floor area. 80% of the houses are from the last 30 years, with a small portion exceeding 50 years in age. 95% of the properties have an area of less than 200 square meters, with the majority being under 100 square meters. For more details on using data analysis see https://github.com/tigerwangcqupt/r_projects-static_project_01.git.

Two:Analysis of ideas and method

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    4.0.0      v tibble    3.3.0
## v lubridate  1.9.4      v tidyr     1.3.1
## v purrr      1.1.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(dplyr)
library(conflicted)
library(showtext)
```

```
## 载入需要的程序包: sysfonts
## 载入需要的程序包: showtextdb
```

```
font_add("SimSun", "simsun.ttc") # 添加系统字体
showtext_auto() # 自动启用字体
conflicts_prefer(dplyr::filter)
```

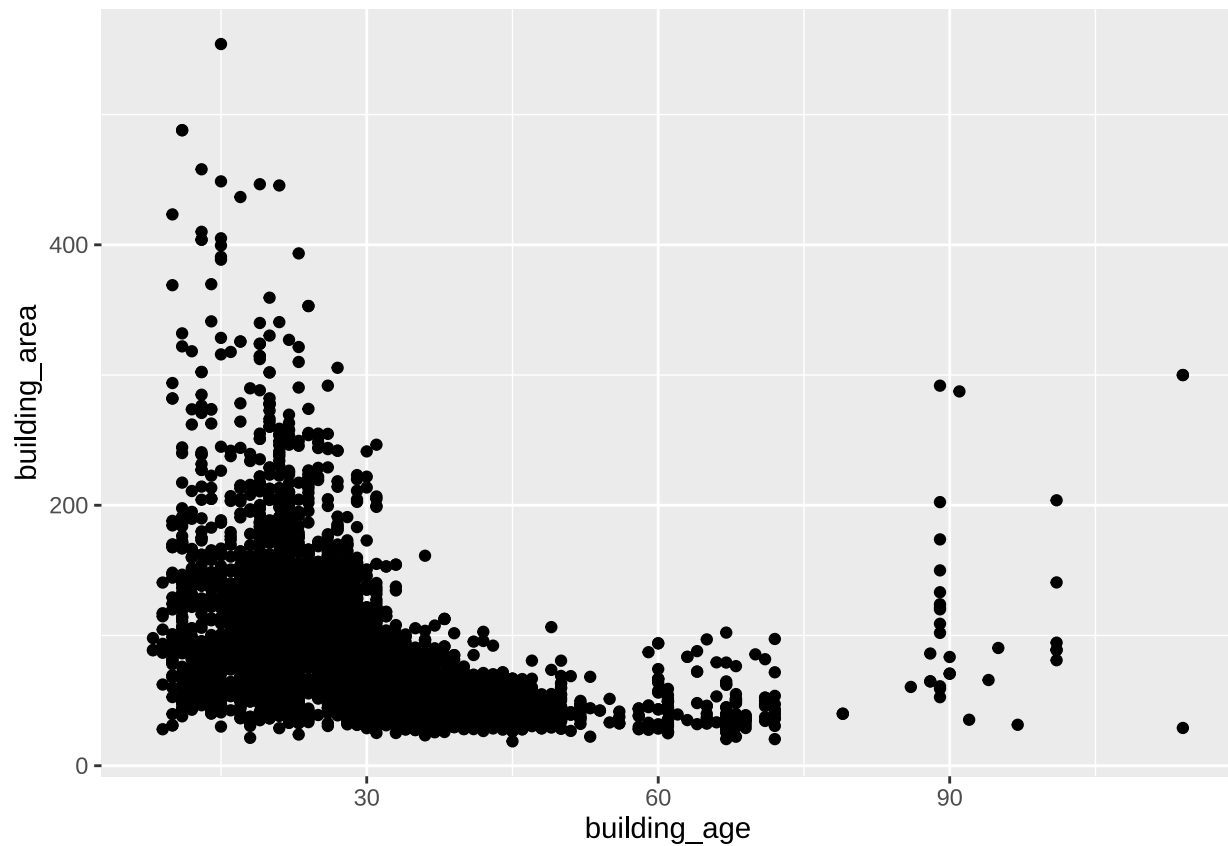
```
## [conflicted] Will prefer dplyr::filter over any other package.
```

```
conflicts_prefer(dplyr::lag)
```

```
## [conflicted] Will prefer dplyr::lag over any other package.
```

```
load("../data/lj_sh_2019.RData")
```

Specific chart display:



Three:Analysis Process

房子年代	建筑面积	占比
10	100	80%
20	80	70%
30	90	60%

Based on the provided data and R data modeling, the following report data is generated.