HW 01 - Airbnb listings in Edinburgh

Elijah Russell

09/19/2025

Load packages and data

```
library(tidyverse)
library(dsbox)
library(dplyr)
install.packages("ggridges")
library(ggridges)
library(ggplot2)
```

Exercises

Exercise 1

The dataset has 13245 observations (rows).

Exercise 2

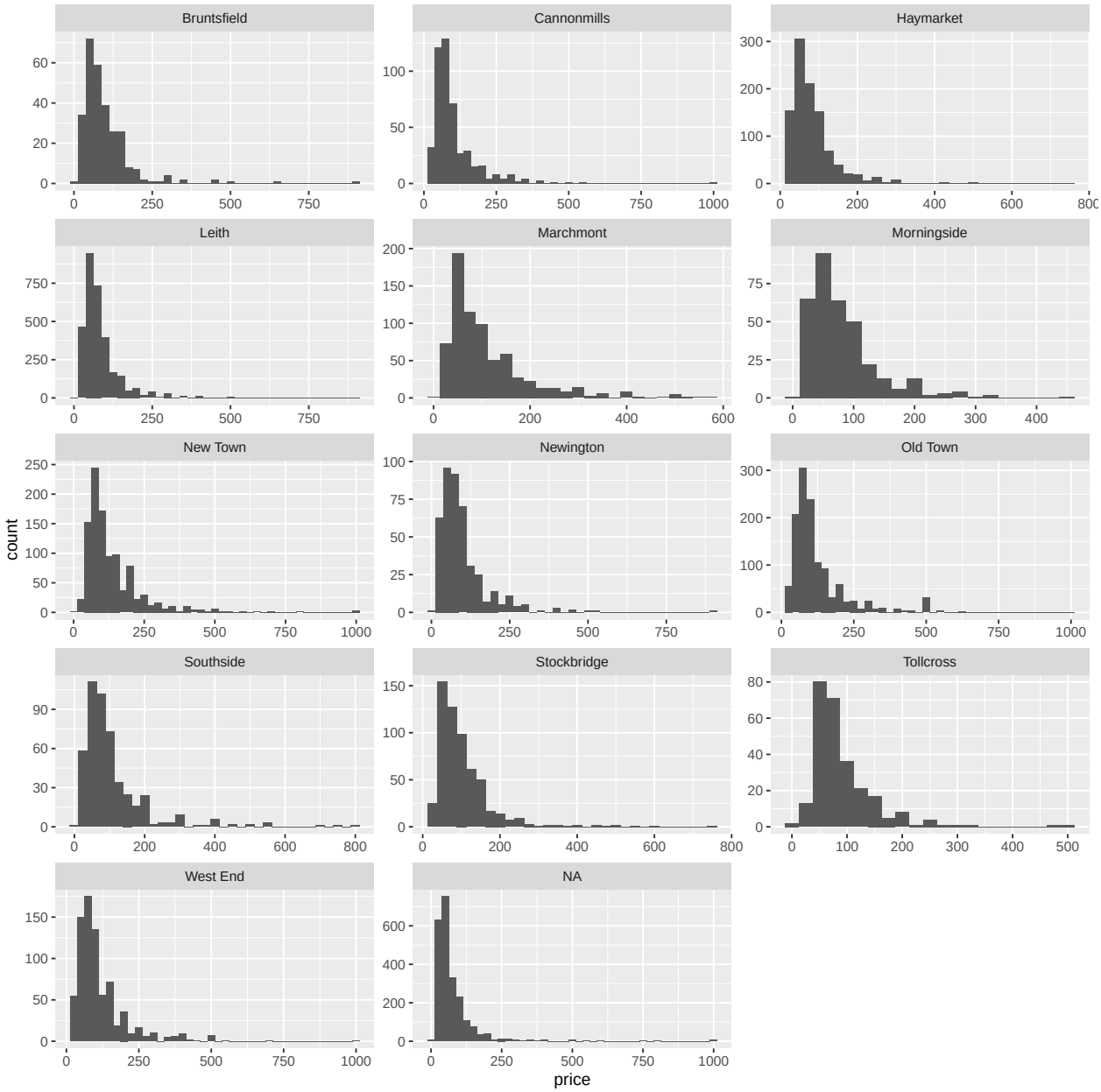
Each row represents and airbnb

Exercise 3

I chose a 3 column wide wrapped grid because it allowed me to show the most amount of data in the smallest space.

```
ggplot(data = edibnb, mapping = aes(x = price)) +
  geom_histogram(binwidth = 25) +
  facet_wrap(~neighbourhood,scales = "free",ncol=3)
```

```
## Warning: Removed 199 rows containing non-finite outside the scale range
## (`stat_bin()`).
```



Exercise 4

All of the top 5 neighbourhoods have relatively similar price distributions. New Town is the most expensive in terms of both mean and median, followed by Old Town. However, Stockbridge and West End have higher starting prices, while Stockbridge also has the lowest ending price.

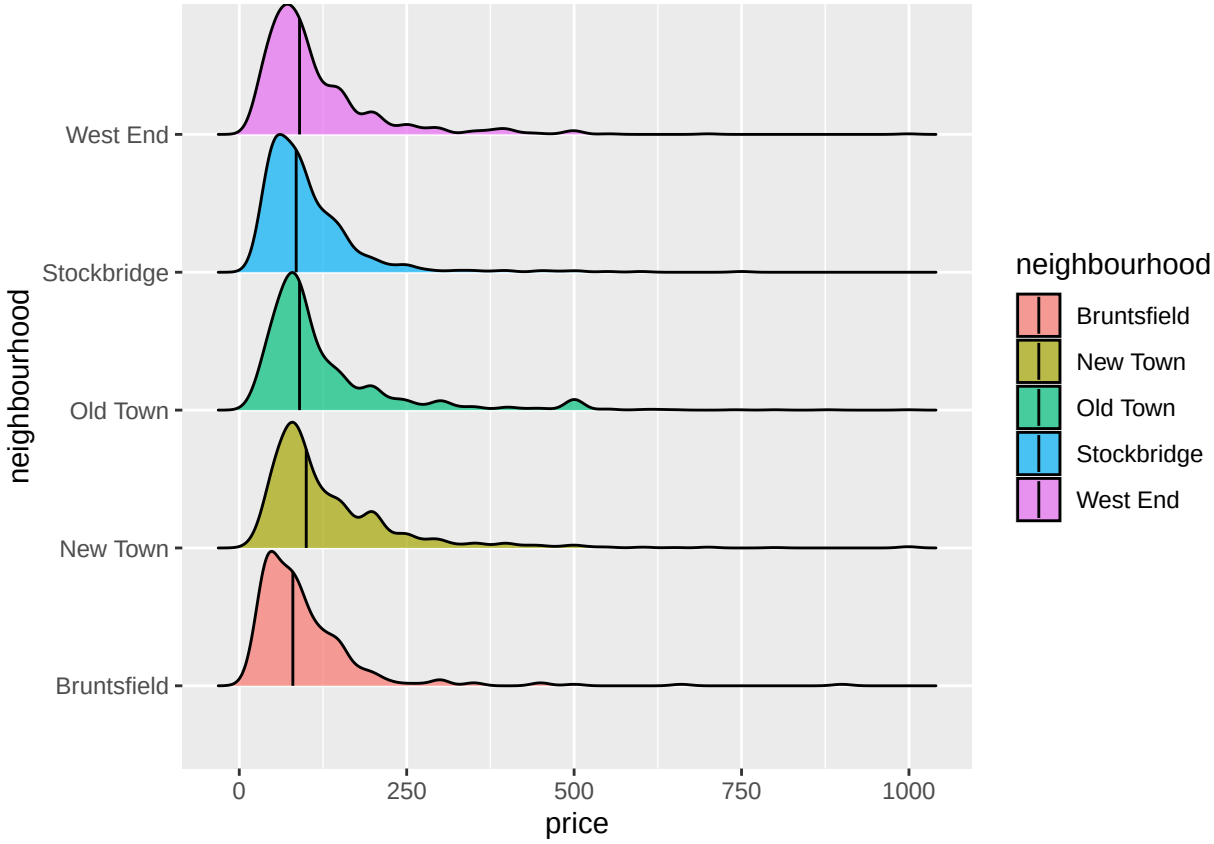
```
top5_neighbourhoods <- edibnb %>%
  group_by(neighbourhood) %>%
  summarise(median_price = median(price, na.rm = TRUE)) %>%
  arrange(desc(median_price)) %>%
  slice_head(n = 5)
```

```
edibnb %>%
  filter(neighbourhood %in% top5_neighbourhoods$neighbourhood) %>%
  ggplot(aes(x = price, y = neighbourhood, fill = neighbourhood)) +
```

```
  geom_density_ridges(scale = 1, alpha = 0.7, quantile_lines = TRUE, quantiles = 2)
```

```
## Picking joint bandwidth of 13.8
```

```
## Warning: Removed 104 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```




```
summary_stats <- edibnb %>%
  filter(neighbourhood %in% top5_neighbourhoods$neighbourhood) %>%
  group_by(neighbourhood) %>%
  summarise(
    min = min(price, na.rm = TRUE),
    mean = mean(price, na.rm = TRUE),
    median = median(price, na.rm = TRUE),
    sd = sd(price, na.rm = TRUE),
    IQR = IQR(price, na.rm = TRUE),
    max = max(price, na.rm = TRUE),
    .groups = "drop"
  )
summary_stats
```

```
## # A tibble: 5 x 7
##   neighbourhood   min  mean median    sd   IQR   max
##   <chr>         <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl>
## 1 Bruntsfield      10  99.4     80  90.2  72.5   900
## 2 New Town         12 136.     100 109.   86.5   999
## 3 Old Town         15 128.      90 110.   76     999
## 4 Stockbridge      21 104.      85  77.6  66     750
```

```
## 5 West End          19 116.      90  93.3  80     999
```

Exercise 5

For the most part the airbnbs in Edinburgh are rated highly consistently. Some neighbourhoods like Leith and Old Town appear to have more negative reviews than other neighbourhoods, but still have overwhelmingly positive reviews.

```
ggplot(edibnb, aes(x = neighbourhood, y = review_scores_rating, fill = neighbourhood)) +
  geom_boxplot(alpha = 0.7, outlier.alpha = 0.3)
```

```
## Warning: Removed 2177 rows containing non-finite outside the scale range
## (`stat_boxplot()`).
```