Tutorial 2

Iz Leitch

```
#this loads libraries
#|echo : FALSE
library(tidyverse)
```

```
-- Attaching packages --------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.4.0      v purrr   1.0.1
v tibble  3.1.8      v dplyr   1.0.10
v tidyr   1.2.1      v stringr 1.5.0
v readr   2.1.3      v forcats 0.5.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```
library(janitor)
```

```
Attaching package: 'janitor'

The following objects are masked from 'package:stats':

    chisq.test, fisher.test
```

```
library(opendatatoronto)
library(dplyr)
library(scales)
```

```
Attaching package: 'scales'

The following object is masked from 'package:purrr':

    discard

The following object is masked from 'package:readr':

    col_factor
```

Introduction

This is the paper about rentsafe toronto housing scores

Running Code

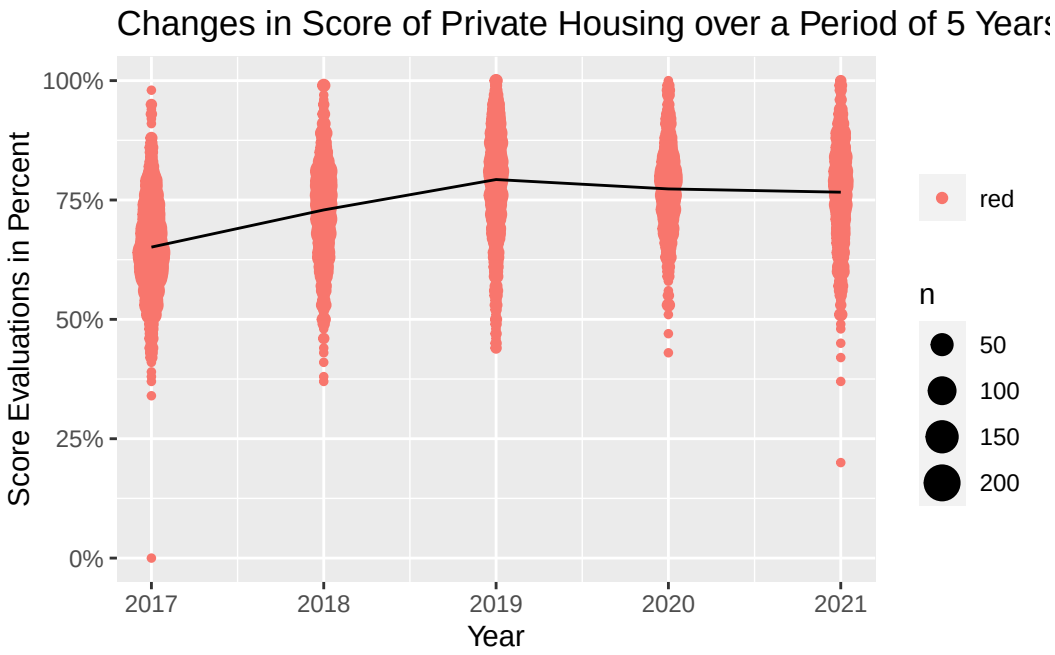
```
#this loads data from the raw data file
#| echo: FALSE
data <-
  read.csv(here::here("inputs/data/raw_data.csv"))

#This turns SCORE, YEAR_EVALUATED, and YEAR_BUILT into numeric and adds them to the datase
data <-
  data |>
  mutate(score = as.numeric(SCORE),
         year_eval = as.numeric(YEAR_EVALUATED),
         year_built = as.numeric(YEAR_BUILT))
```

```
# Graphing score by Year, Property type: Private
#| echo: false

data_average_property_type_private <-
  data |>
  filter(PROPERTY_TYPE == "PRIVATE") |>
  group_by(year_eval) |>
  summarise(mean = mean(score / 100))

data |>
  filter(PROPERTY_TYPE == "PRIVATE")|>
```

```
  ggplot(aes(x = year_eval,
             y = score / 100)) +
  scale_y_continuous(labels = percent_format(),
                     limits = c(0, 1)) +
  geom_count(aes(color = "red")) +
  geom_line(data = data_average_property_type_private, aes(x = year_eval, y = mean)) +
  labs(
    title = "Changes in Score of Private Housing over a Period of 5 Years",
    x = "Year",
    y = "Score Evaluations in Percent",
    color = ""
  )
```

```
Warning: Removed 1651 rows containing non-finite values (`stat_sum()`).
```

```
Warning: Removed 1 row containing missing values (`geom_line()`).
```

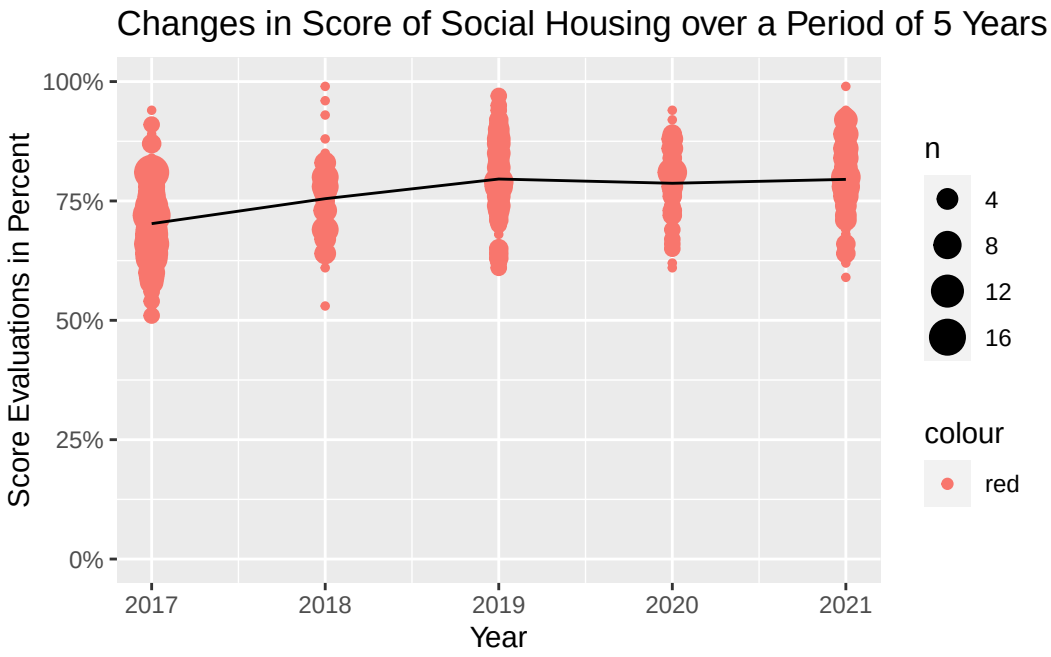



```
# Graphing score by Year, Property type: Social Housing
#| echo: false
```

```r
data_average_property_type_social <-
  data |>
  filter(PROPERTY_TYPE == "SOCIAL HOUSING") |>
  group_by(year_eval) |>
  summarise(mean = mean(score / 100))

data |>
  filter(PROPERTY_TYPE == "SOCIAL HOUSING")|>
  ggplot(aes(x = year_eval,
             y = score / 100)) +
  scale_y_continuous(labels = percent_format(),
                     limits = c(0, 1)) +
  geom_count(aes(color="red")) +
  geom_line(data = data_average_property_type_social, aes(x = year_eval, y = mean)) +
  labs(
    title = "Changes in Score of Social Housing over a Period of 5 Years",
    x = "Year",
    y = "Score Evaluations in Percent",
  )
```

```
Warning: Removed 110 rows containing non-finite values (`stat_sum()`).
```

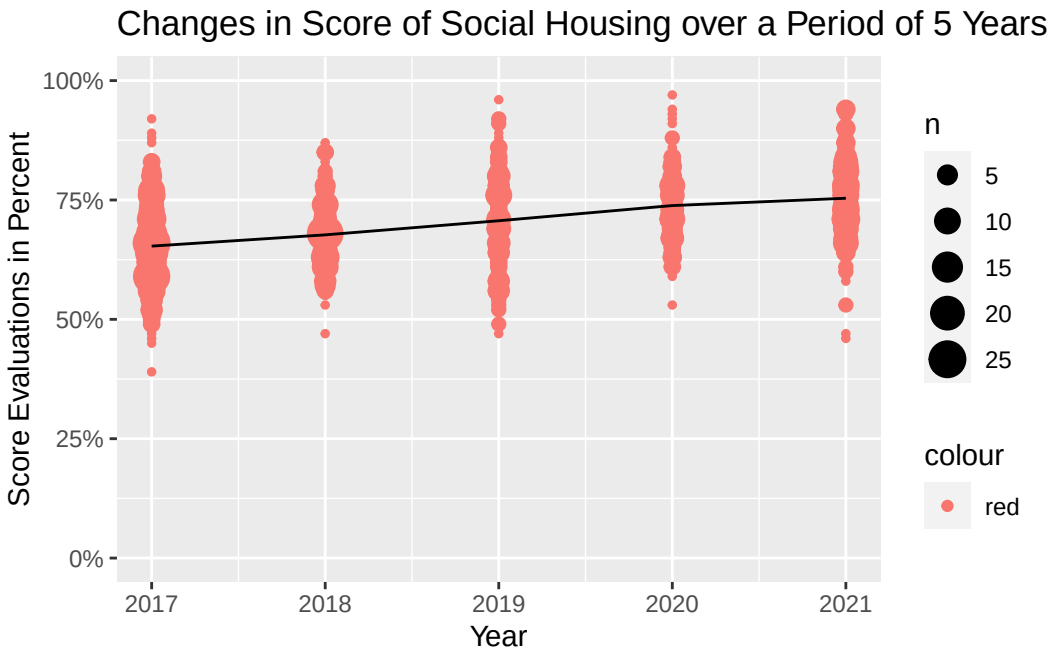
```
Warning: Removed 1 row containing missing values (`geom_line()`).
```




```
#Graphing score by year of evaluation, Property type:TCHC
#| echo: false

data_average_property_type_tchc <-
  data |>
  filter(PROPERTY_TYPE == "TCHC") |>
  group_by(year_eval) |>
  summarise(mean = mean(score / 100))

data |>
  filter(PROPERTY_TYPE == "TCHC")|>
  ggplot(aes(x = year_eval,
             y = score / 100)) +
  scale_y_continuous(labels = percent_format(),
                     limits = c(0, 1)) +
  geom_count(aes(color="red")) +
  geom_line(data = data_average_property_type_tchc, aes(x = year_eval, y = mean)) +
  labs(
    title = "Changes in Score of Social Housing over a Period of 5 Years",
    x = "Year",
    y = "Score Evaluations in Percent",
  )
```

```
Warning: Removed 140 rows containing non-finite values (`stat_sum()`).
```

```
Warning: Removed 1 row containing missing values (`geom_line()`).
```

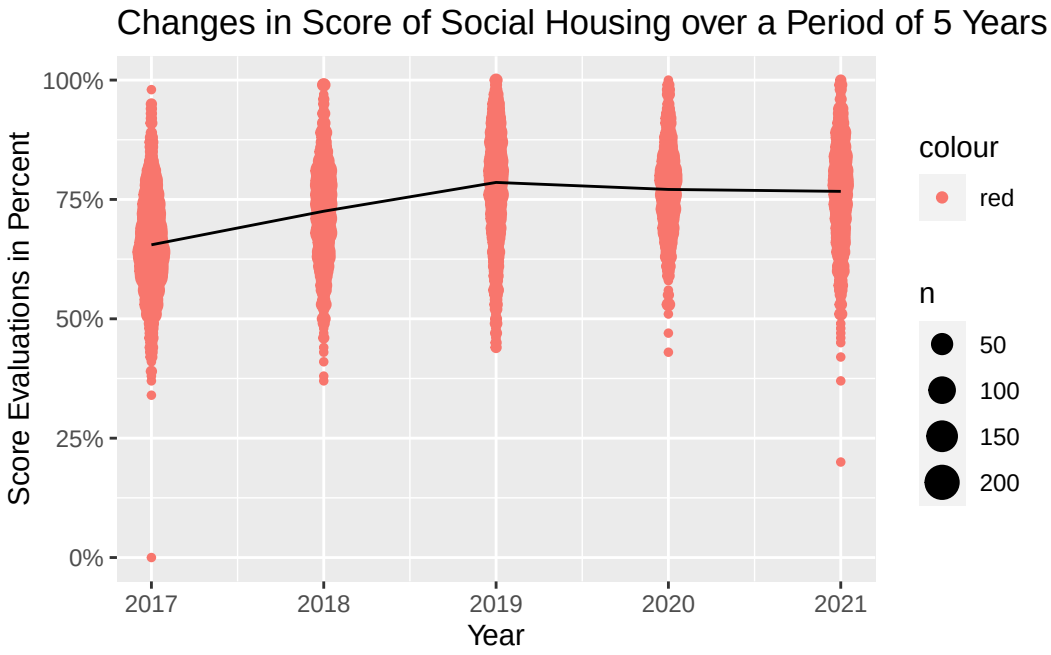



```
# Graphing score by year of evaluation: All Property types
#| echo: false

data_average_property_type <-
  data |>
  group_by(year_eval) |>
  summarise(mean = mean(score / 100))

data |>
  ggplot(aes(x = year_eval,
             y = score / 100)) +
  scale_y_continuous(labels = percent_format(),
                     limits = c(0, 1)) +
  geom_count(aes(color="red")) +
  geom_line(data = data_average_property_type, aes(x = year_eval, y = mean)) +
  labs(
    title = "Changes in Score of Social Housing over a Period of 5 Years",
```

```
    x = "Year",
    y = "Score Evaluations in Percent",
  )
```

```
Warning: Removed 1901 rows containing non-finite values (`stat_sum()`).
```

```
Warning: Removed 1 row containing missing values (`geom_line()`).
```



References