# First presentation of my data set

Callum Coode

## Effect of social distancing on cases of flu

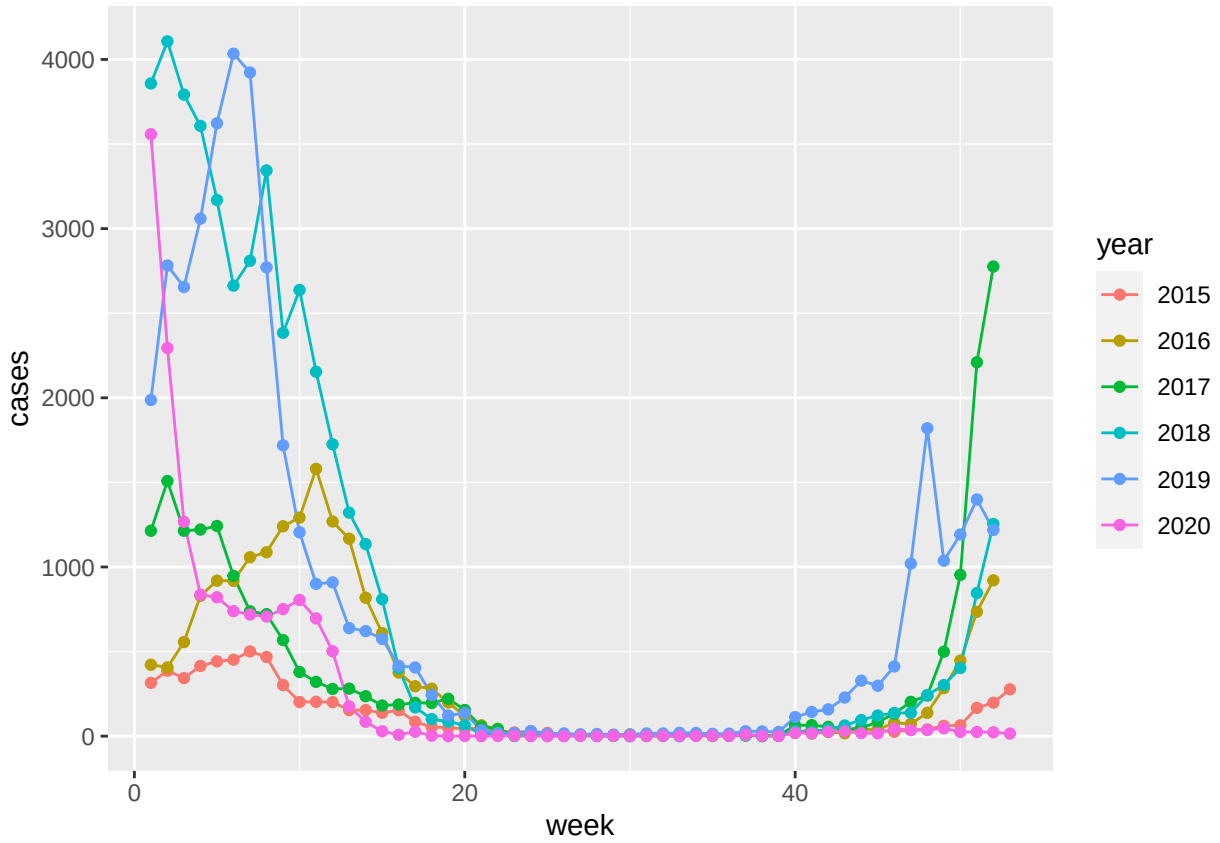
I wanted to find something related to Covid-19 but looking at something positive. I found data from *WHO on cases of influenza* as a lot of people are saying that they haven't got flu or colds this year due to social distancing. I have processed the data to just look at the total cases of all types of flu on a weekly basis.

```
df <- read.csv(here("data", "rawfludata.csv"))

df <- df %>%
  rename(year = Year, week = Week, start = Start.date, end = End.date,
         cases = Total.number.of.influenza.positive.viruses)

df <- df %>%
  select(year, week, start, end, cases)

head(df, 10)
```

```
##    year week      start        end cases
## 1  2015    1 2014-12-29 2015-01-04   315
## 2  2015    2 2015-01-05 2015-01-11   387
## 3  2015    3 2015-01-12 2015-01-18   343
## 4  2015    4 2015-01-19 2015-01-25   415
## 5  2015    5 2015-01-26 2015-02-01   442
## 6  2015    6 2015-02-02 2015-02-08   452
## 7  2015    7 2015-02-09 2015-02-15   501
## 8  2015    8 2015-02-16 2015-02-22   468
## 9  2015    9 2015-02-23 2015-03-01   302
## 10 2015   10 2015-03-02 2015-03-08   202
```

From initially looking at the data, it does look like the cases in 2020/2021 are a lot lower. For example, the mean number of cases in 2020 is 272.17, whereas in 2019 the mean is 816.29.

```
df$year <- as.character(df$year)

p1 <- ggplot(df, aes(x = week, y = cases, group = year, colour = year))
p1 + geom_line() + geom_point()
```




Since cases are usually a lot higher in the winter and lockdown started on 23rd March (week 13), I may display this data from week 13 to week 13 to better show the impact of social distancing on cases. I may also include bar graphs showing highest or total cases for each of these periods.

**Please note this does not include data from 2021.** There is currently no data for the most recent 2 weeks of 2021, but this will be added as soon as this becomes available, hopefully in the next few weeks.