Exercise 12.6.1

Xiru Lyu

2/20/2018

```
# load in required package
library(tidyverse)
```

3. I claimed that `iso2` and `iso3` were redundant with country. Confirm this claim.

```
# import the dataset

who <- who %>% gather(code, value, new_sp_m014:newrel_f65, na.rm = TRUE) %>%
  mutate(code = stringr::str_replace(code, "newrel", "new_rel")) %>%
  separate(code, c("new", "var", "sexage"))

# check to see if `iso2` and `iso3` are redundant
who.2 <- who %>% select(country, iso2,iso3) %>% distinct()
length(unique(who.2$country))
```

```
## [1] 219
```

```
length(unique(who.2$iso2))
```

```
## [1] 219
```

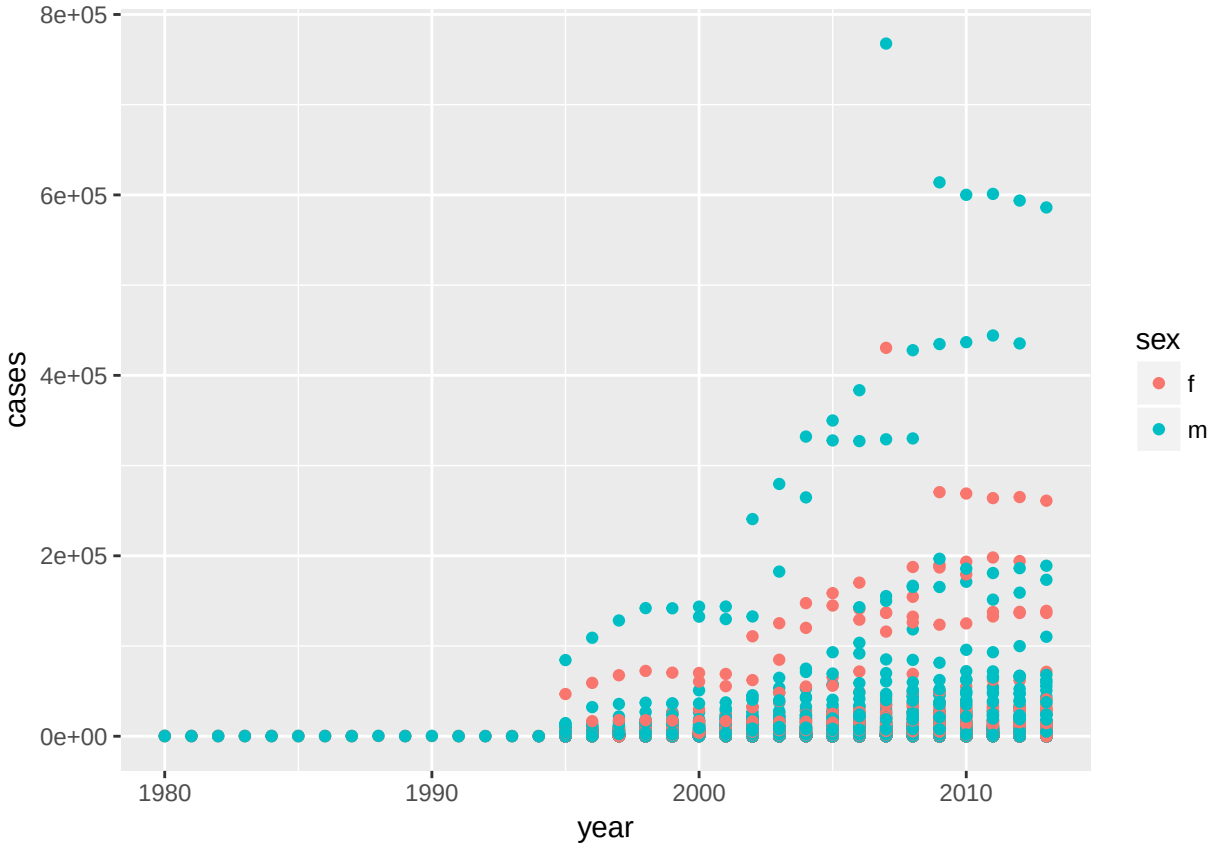
```
length(unique(who.2$iso3))
```

```
## [1] 219
```

4. For each country, year, and sex compute the total number of cases of TB. Make an informative visualisation of the data.

```
who <- who %>% select(-iso2,-iso3,-new) %>% separate(sexage,c('sex','age'),sep=1)

who %>% group_by(country,year,sex) %>% summarize(cases=sum(value)) %>% ggplot(aes(x=year,y=cases,group=
```


8e+05
6e+05
4e+05
2e+05
0e+00
cases
1980
1990
2000
2010
year
sex
f
m