```
  mutate(neighbourhood_name=str_replace(neighbourhood_name,
                                        "weston pellam park",
                                        "weston pelham park"))
# back and make data corrections


nbhoods_temp2 <- nbhoods_temp %>%
  full_join(nbhood_raw_temp, by="neighbourhood_name")
nbhoods_all <- nbhoods_temp2 %>%
  full_join(income, by="neighbourhood_name")

# checking duplicated rows
nrow(nbhoods_all)
```

```
## [1] 140
```

```
nrow(nbhood_raw_temp)
```

```
## [1] 140
```

```
nrow(income)
```

```
## [1] 140
```

```
nrow(nbhoods_temp)
```

```
## [1] 140
```

```
nrow(nbhoods_temp2)
```

```
## [1] 140
```

```
problems <- nbhoods_all %>%
  filter(is.na(neighbourhood_id) | is.na(AREA_ID) | is.na(id))
na <- nbhoods_all %>%
  filter_all(any_vars(is.na(.)))
```

Data wrangling: part 3

```
med_inc <- median(nbhoods_all$percentage);med_inc
```

```
## [1] 16.55
```

```
med_rate <- median(nbhoods_all$rate_per_100000);med_rate
```

```
## [1] 2486.039
```

```
nbhoods_final <- nbhoods_all %>%
  select(neighbourhood_name,percentage,rate_per_100000) %>%
  mutate(nbhood_type = case_when(
    ((percentage >= med_inc)&(rate_per_100000 >= med_rate))
    ~ "Higher low income rate, higher case rate",
    ((percentage >= med_inc)&(rate_per_100000 < med_rate))
    ~ "Higher low income rate, lower case rate",
    ((percentage < med_inc)&(rate_per_100000 >= med_rate))
    ~ "Lower low income rate, higher case rate",
    ((percentage < med_inc)&(rate_per_100000 < med_rate))
```