Code last run 2021-02-17.
Daily: Data as of January 29, 2021.
Neighbourhood: Data as of January 31, 2021.

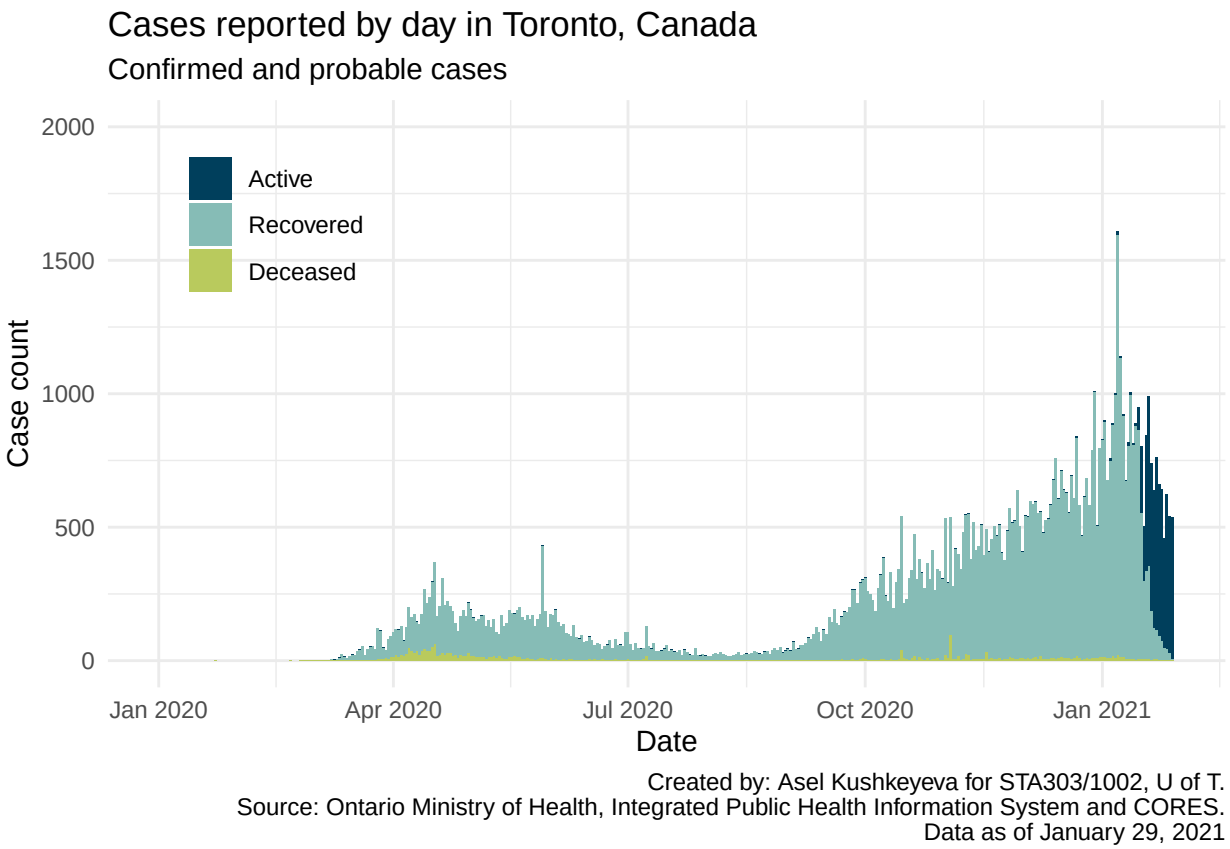
Task 1: Daily cases

Data wrangling

```
reported <- reported_raw %>%
  mutate_if(is.numeric, replace_na, replace = 0) %>%
  mutate(reported_date = date(reported_date)) %>%
  rename(Active = active, Recovered = recovered, Deceased = deceased) %>%
  pivot_longer(col = c(Active, Recovered, Deceased), names_to = "cases") %>%
  mutate(cases = fct_relevel(cases, "Recovered", after = 1))
```

Data visualization

```
reported %>%
  ggplot(aes(reported_date, value, fill = cases)) +
  geom_bar(stat = "identity") +
  lims(x= c(date("2020-01-01"), date("2021-01-29")), y = c(0, 2000)) +
  labs(title = "Cases reported by day in Toronto, Canada",
       subtitle = "Confirmed and probable cases",
       caption = str_c(
         "Created by: Asel Kushkeyeva for STA303/1002, U of T.\n",
         "Source: Ontario Ministry of Health, Integrated Public Health Information System and CORES.\n"
         date_daily[1,1]),
       x = "Date",
       y = "Case count") +
  theme_minimal() +
  theme(legend.title = element_blank(),
        legend.position = c(.15, .8)) +
  scale_fill_manual(values = c("#003F5C", "#86BCB6", "#B9CA5D"))
```



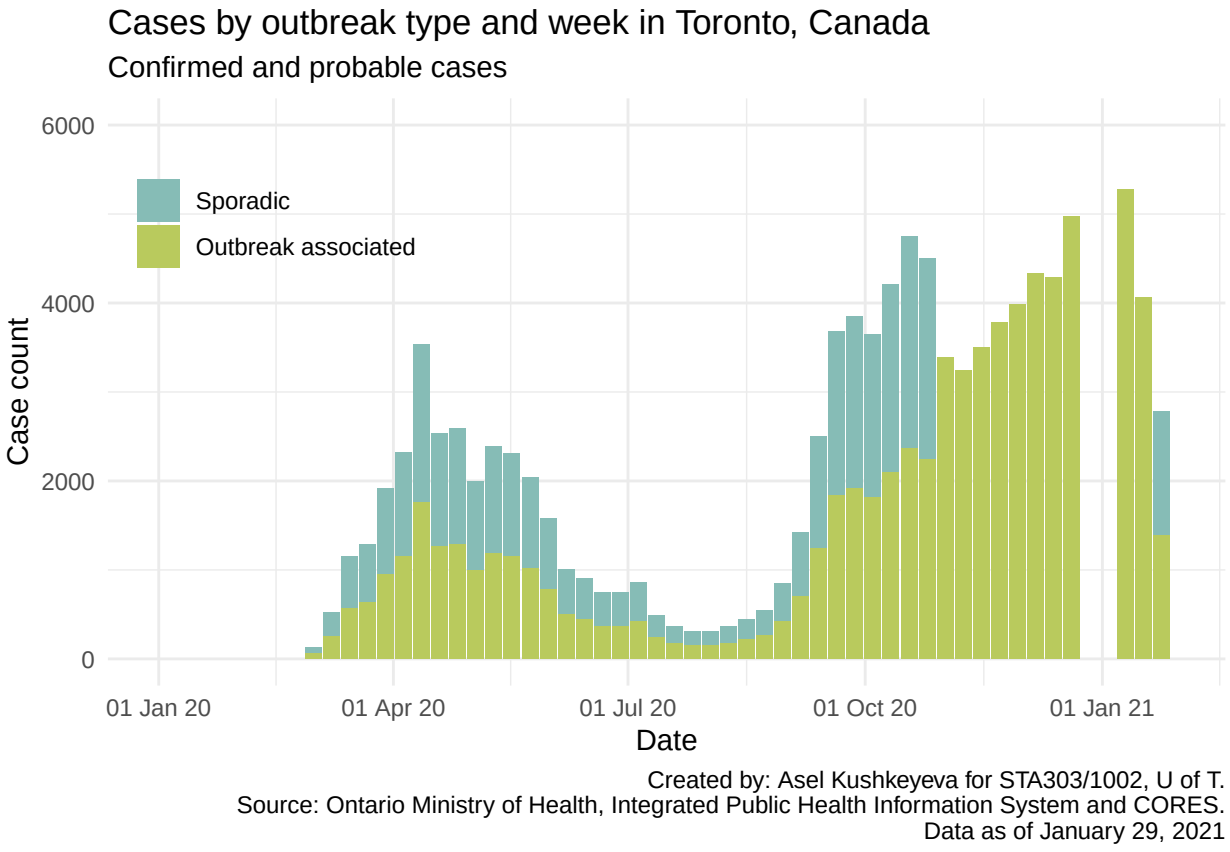
Task 2: Outbreak type

Data wrangling

```
outbreak <- outbreak_raw %>%
  mutate(episode_week = date(episode_week)) %>%
  pivot_wider(names_from = outbreak_or_sporadic, values_from = cases) %>%
  mutate(total_cases = `OB Associated` + `Sporadic`) %>%
  rename(`Outbreak associated` = `OB Associated`) %>%
  pivot_longer(col = c(`Outbreak associated`, `Sporadic`), names_to = "cases") %>%
  mutate(cases = fct_rev(cases))
```

Data visualization

```
outbreak %>%
  ggplot(aes(episode_week, total_cases, fill = cases)) +
  geom_bar(stat = "identity") +
  lims(x= c(date("2020-01-01"), date("2021-01-29")), y = c(0, 6000)) +
  labs(title = "Cases by outbreak type and week in Toronto, Canada",
       subtitle = "Confirmed and probable cases",
       caption = str_c(
         "Created by: Asel Kushkeyeva for STA303/1002, U of T.\n",
         "Source: Ontario Ministry of Health, Integrated Public Health Information System and CORES.\n"
         date_daily[1,1]),
       x = "Date",
       y = "Case count") +
  theme_minimal() +
  theme(legend.title = element_blank(),
        legend.position = c(.15, .8)) +
  scale_fill_manual(values = c("#86BCB6", "#B9CA5D")) +
  scale_x_date(labels = scales::date_format("%d %b %y"),
               limits = c(date("2020-01-01"), date("2021-01-29")))
```



Task 3: Neighbourhoods

Data wrangling: part 1

```
income <- nbhood_profile %>%
  mutate(Characteristic = gsub(" ", "", Characteristic)) %>%
  filter(Topic == "Low income in 2015", Characteristic == "18to64years(%)",
         `_id` == 1143) %>%
  select(!`City of Toronto`) %>%
  pivot_longer(col = c(6:145), names_to = "neighbourhood_name") %>%
  mutate(value = parse_number(value))
```

Data wrangling: part 2

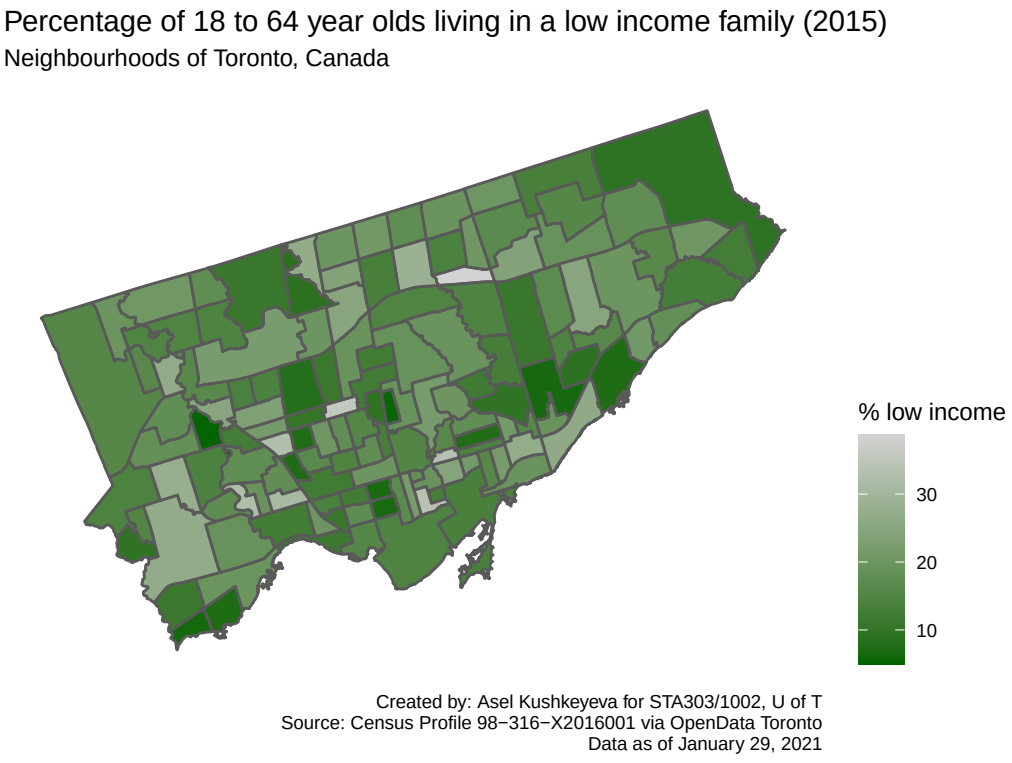
```
nbhoods_all1 <-  nbhoods_shape_raw %>%
  mutate(neighbourhood_name = str_remove(AREA_NAME, "[\\d]+")) %>%
  mutate(neighbourhood_name = gsub("\\(", "", neighbourhood_name)) %>%
  mutate(neighbourhood_name = gsub("\\)", "", neighbourhood_name)) %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\']")) %>%
  mutate(neighbourhood_name = gsub("Pellam", "Pelham", neighbourhood_name)) %>%
  mutate(neighbourhood_name = gsub("St.James", "St. James", neighbourhood_name)) %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\s]$")) %>%
  mutate(neighbourhood_name = sort(neighbourhood_name))


nbhood_raw1 <- nbhood_raw %>%
  filter(neighbourhood_name != "Missing Address/Postal Code") %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\']")) %>%
  mutate(neighbourhood_name = sort(neighbourhood_name)) %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\(]")) %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\)]"))%>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\s]$"))


income1 <- income %>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\']")) %>%
  mutate(neighbourhood_name = gsub("\\(", "", neighbourhood_name)) %>%
  mutate(neighbourhood_name = gsub("\\)", "", neighbourhood_name))%>%
  mutate(neighbourhood_name = str_remove(neighbourhood_name, "[\\s]$"))

nbhoods_all <- merge(nbhood_raw1,income1, by = "neighbourhood_name")
nbhoods_all <- merge(nbhoods_all1, nbhoods_all, by = "neighbourhood_name")

nbhoods_all <- nbhoods_all %>%
  rename(rate_per_100000 = rate_per_100_000_people)

rm(nbhoods_all1, nbhood_raw1, income1)
```

Data wrangling: part 3

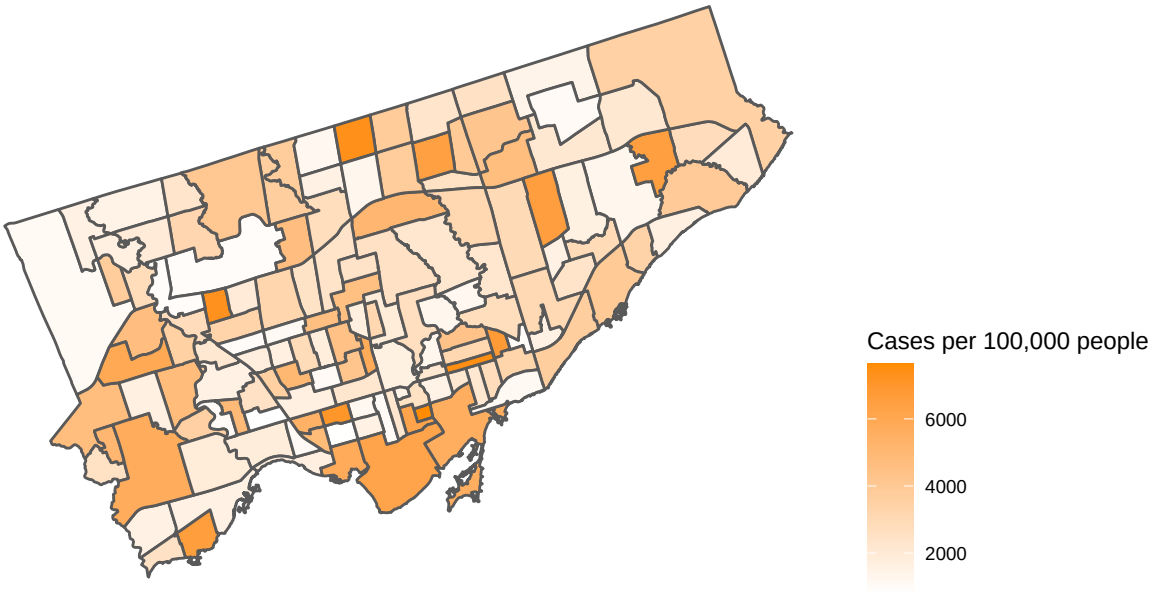
```
nbhoods_final <- nbhoods_all %>%
  mutate(med_inc = median(value), med_rate = median(rate_per_100000),
         nbhood_type = case_when(
           value >= med_inc & rate_per_100000 >= med_rate ~
             "Higher low income rate, higher case rate",
           value >= med_inc & rate_per_100000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

Data visualization

```
ggplot(data = nbhoods_final, aes(geometry = geometry)) +
  geom_sf(aes(fill = value)) +
  theme_map() +
  labs(title = "Percentage of 18 to 64 year olds living in a low income family (2015)",
       subtitle = "Neighbourhoods of Toronto, Canada",
       caption = str_c(
         "Created by: Asel Kushkeyeva for STA303/1002, U of T\n",
         "Source: Census Profile 98-316-X2016001 via OpenData Toronto\n",
         date_daily[1,1])) +
  theme(legend.position = "right") +
  scale_fill_gradient(name= "% low income", low = "darkgreen", high = "lightgrey")
```



```
ggplot(data = nbhoods_final, aes(geometry = geometry)) +
  geom_sf(aes(fill = rate_per_100000)) +
  theme_map() +
  labs(title = "COVID-19 cases per 100,000, by neighbourhood in Toronto, Canada",
       caption = str_c(
         "Created by: Asel Kushkeyeva for STA303/1002, U of T\n",
         "Source: Ontario Ministry of Health, Integrated Public Health Information System and CORES\n",
         date_daily[1,1])) +
  theme(legend.position = "right") +
  scale_fill_gradient(name= "Cases per 100,000 people", low = "white",
                      high = "darkorange")
```



```
ggplot(data = nbhoods_final, aes(geometry = geometry)) +
  geom_sf(aes(fill = nbhood_type)) +
  theme_map() +
  labs(title = "COVID-19 cases per 100,000, by neighbourhood in Toronto, Canada",
       caption = str_c(
         "Created by: Asel Kushkeyeva for STA303/1002, U of T\n",
         "Income data source: Census Profile 98-316-X2016001 via OpenData Toronto\n",
         "COVID data source: Ontario Ministry of Health, Integrated Public\n",
         "Health Information System and CORES\n",
         date_daily[1,1])) +
  theme(legend.position = "right") +
  scale_fill_brewer(
    palette = "Set1",
    name = str_c(
      "% of 18 to 64 year-olds in\n",
      "low income families and\n", "COVID-19 case rates"))
```