Code last run 2021-02-10.
Daily: Data as of January 29, 2021.
Neighbourhood: Data as of January 28, 2021.

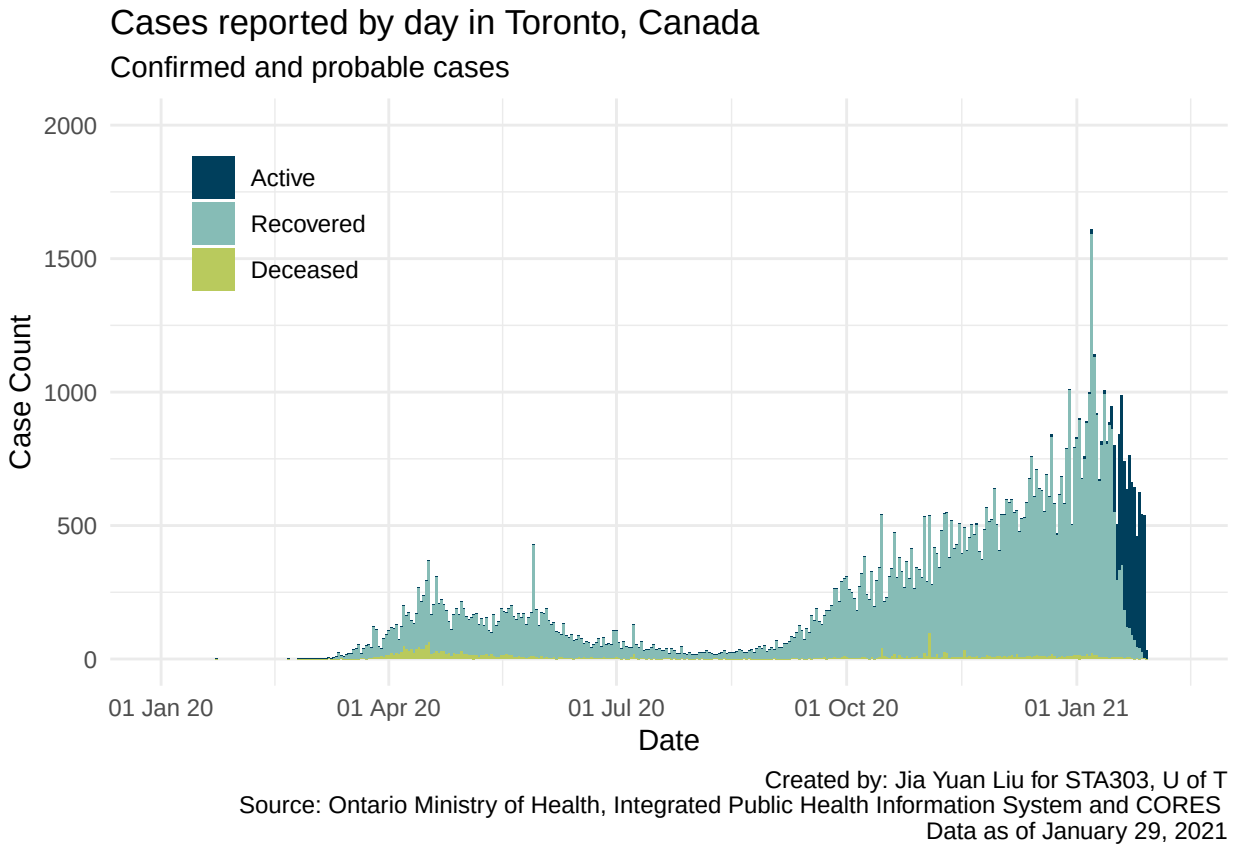
Task 1: Daily cases

Data wrangling

```
reported <- reported_raw %>%
  mutate_if(is.numeric, replace_na, replace=0) %>%
  mutate(reported_date = date(reported_date)) %>%
  pivot_longer(-c(reported_date), names_to = "CaseType", values_to = "Cases") %>%
  mutate(CaseType = str_to_sentence(CaseType)) %>%
  mutate(CaseType = fct_relevel(CaseType,"Deceased", after = 2))
```

Data visualization

```
reported %>%
  ggplot(aes(x=reported_date, y = Cases, fill = CaseType)) + geom_bar(stat = "identity") +
  theme_minimal() +
  labs(title = "Cases reported by day in Toronto, Canada",
  subtitle = "Confirmed and probable cases",
  x = "Date",
  y = "Case Count",
  caption = str_c("Created by: Jia Yuan Liu for STA303, U of T
  Source: Ontario Ministry of Health, Integrated Public Health Information System and CORES \n",
                  date_daily[1,1]))  +
  scale_x_date(labels = scales::date_format("%d %b %y"),
               limits = c(date("2020-01-01"),Sys.Date( ))) +
  scale_y_continuous(limits = c(0,2000)) +
  theme(legend.title = element_blank(), legend.position = c(0.15, 0.8)) +
  scale_fill_manual(values = c("#003F5C","#86BCB6","#B9CA5D"))
```



Task 2: Outbreak type

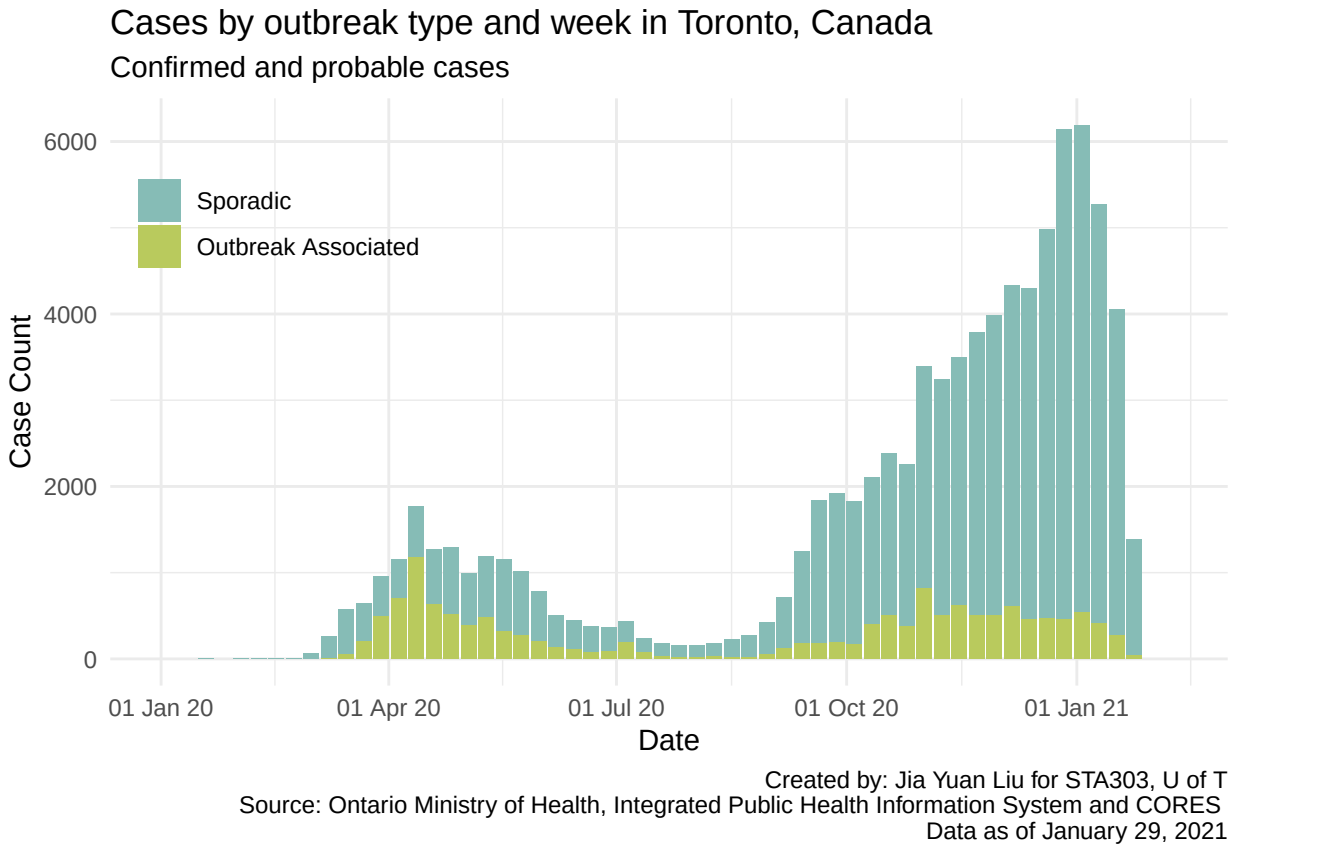
Data wrangling

```
outbreak <- outbreak_raw %>%
  mutate_if(is.numeric, replace_na, replace=0) %>%
  mutate(episode_week = date(episode_week)) %>%
  mutate(outbreak_or_sporadic = str_replace(outbreak_or_sporadic,
                                            "OB Associated", "Outbreak Associated")) %>%
  mutate(outbreak_or_sporadic = fct_relevel(outbreak_or_sporadic,
                                            "Outbreak Associated", after = 2))

total_cases <- outbreak %>%
  group_by(episode_week) %>%
  summarise(total = sum(total_cases = cases),.groups = "drop")
```

Data visualization

```
outbreak%>%
  ggplot(aes(x=episode_week, y = cases, fill= outbreak_or_sporadic)) +
  geom_bar(stat = "identity") +
  theme_minimal() +
  labs(title = "Cases by outbreak type and week in Toronto, Canada",
  subtitle = "Confirmed and probable cases",
  x = "Date",
  y = "Case Count",
  caption = str_c("Created by: Jia Yuan Liu for STA303, U of T
  Source: Ontario Ministry of Health, Integrated Public Health Information System and CORES \n",
                  date_daily[1,1]))  +
  scale_x_date(labels = scales::date_format("%d %b %y"),
               limits = c(date("2020-01-01"),Sys.Date( ))) +
  scale_y_continuous(limits = c(0,max(total_cases$total))) +
  theme(legend.title = element_blank(), legend.position = c(0.15, 0.8)) +
  scale_fill_manual(values = c("#86BCB6","#B9CA5D"))
```



Task 3: Neighbourhoods

Data wrangling: part 1

```
income <- nbhood_profile%>%
  filter(`_id` == 1143) %>%
  pivot_longer(-c(`_id`, Category, Topic, `Data Source`,Characteristic),
               names_to = "neighbourhood", values_to = "percentage") %>%
  mutate(percentage = parse_number(percentage))
```

Data wrangling: part 2

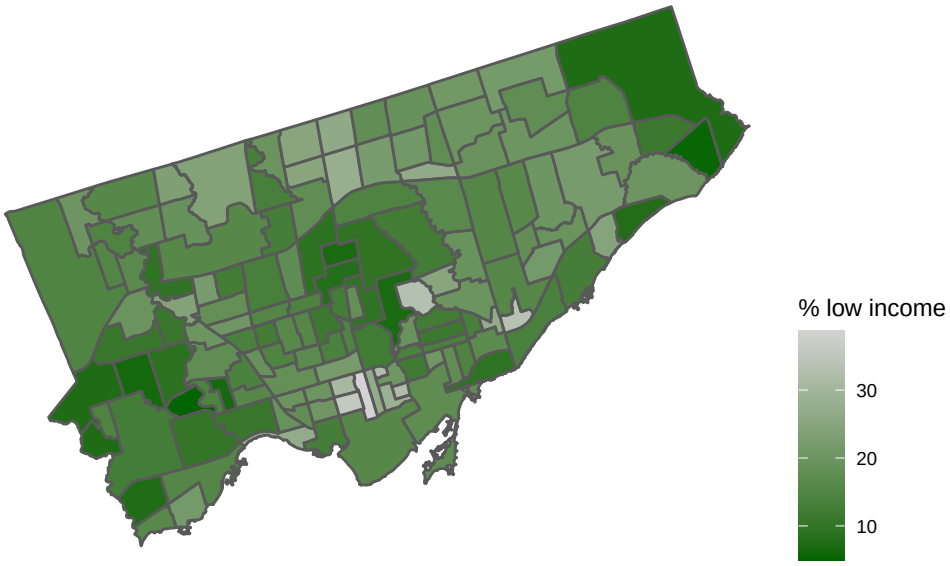
```
nbhoods_all <- nbhoods_shape_raw %>%
  mutate(neighbourhood_name = str_remove(AREA_NAME, "\\s\\(\\d+\\)$")) %>%
  mutate(neighbourhood_name = str_replace(neighbourhood_name, "St.James", "St. James")) %>%
  mutate(neighbourhood_name = str_replace(neighbourhood_name, "Weston-Pellam", "Weston-Pelham")) %>%
  left_join(income, by = c("neighbourhood_name"="neighbourhood")) %>%
  left_join(nbhood_raw, by = "neighbourhood_name") %>%
  rename("rate_per_100000" = "rate_per_100_000_people")
```

Data wrangling: part 3

```
nbhoods_final <- nbhoods_all %>%
  mutate(
    med_inc = median(nbhoods_all$percentage, na.rm = TRUE),
         med_rate = median(nbhoods_all$rate_per_100000, na.rm = TRUE),
         nbhood_type = case_when(
           percentage >= med_inc & rate_per_100000 >= med_rate ~ "Higher low income rate, higher case ra
           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 ~ "Lower low income rate, lower case rate"
```

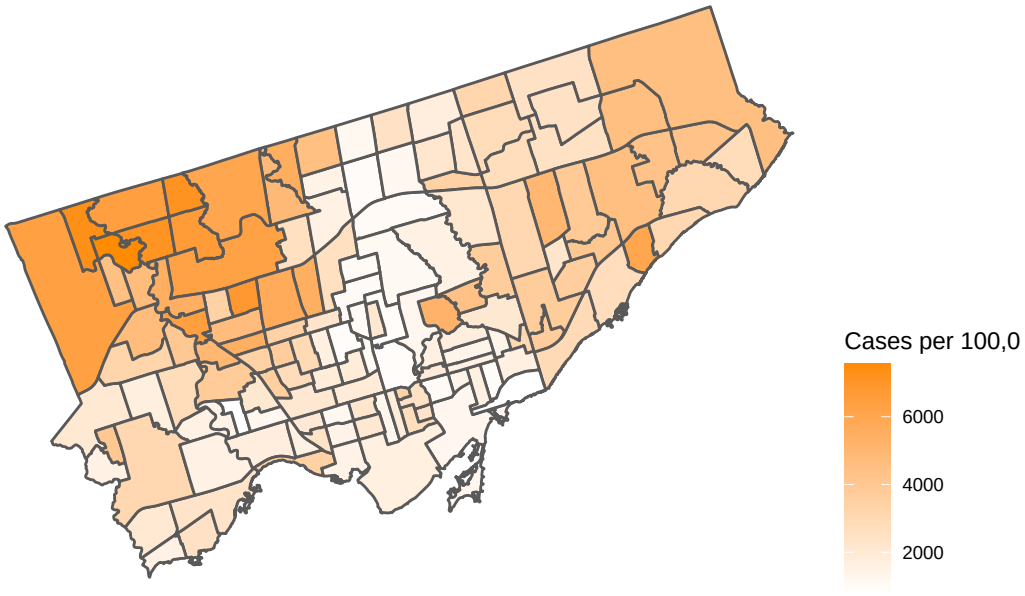
Data visualization

```
nbhoods_final %>%
ggplot() +
geom_sf(aes(fill = percentage)) +
scale_fill_gradient(name = "% low income", low = "darkgreen", high = "lightgrey")+
               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: Jia Yuan Liu for STA303/1002, U of T\nSource: Census Profile 98-316-X20160
theme_map() + theme(legend.position = c(1, 0))
```



```
nbhoods_final %>%
ggplot() +
geom_sf(aes(fill = rate_per_100000)) +
scale_fill_gradient(name = "Cases per 100,000 people", low = "white", high = "darkorange") +
  labs(
title = "COVID-19 cases per 100,000, by neighbourhood in Toronto, Canada",
caption = str_c("Created by:Jia Yuan Liu for STA303/1002, U of T\nSource: Ontario Ministry of Health,In
theme_map() + theme(legend.position = c(1, 0))
```

COVID−19 cases per 100,000, by neighbourhood in Toronto, Canada



Created by:Jia Yuan Liu for STA303/1002, U of T
Source: Ontario Ministry of Health,Integrated System and CORES
Data as of January 29, 2021

```
nbhoods_final %>%
  ggplot() +
  geom_sf(aes(fill = nbhood_type)) +
  scale_fill_brewer(name = "% of 18 to 64 year-olds in\nlow income families and\nCOVID-19 case rates", p
labs(
  title = "COVID-19 cases per 100,000, by neighbourhood in Toronto, Canada",
  caption = str_c("Created by: Jia Yuan Liu for STA303/1002, U of T\nIncome data source: Census   Profil
COVID data sourceL Ontario Ministry of Health, Integrated Public Health Information System and CORES\n"
  theme_map() + theme(legend.position = c(1, 0))
```