

# Assingment 2 Main

Smitesh Patil

2023-03-06

```
#loading colorblind colors
palette <- colorblindr::palette_OkabeIto

#normalising dailyccases and confirmdccases
IRL_Covid19_2021_12_21<- IRL_counties_Covid19%>%
  filter(TimeStamp == ymd("2021-12-21"))%>%
  mutate(ConfirmedC_per_100k = round(100000 * ConfirmedC/Population,1))%>%
  mutate(DailyCCase_per_100k = round(100000 * DailyCCase/Population, 1))

#getting the mean of confirmed cases per 100k for plotting
mean_daily_cases <- IRL_Covid19_2021_12_21 %>%
  select(ConfirmedC_per_100k) %>%
  st_drop_geometry() %>%
  unlist() %>%
  mean()

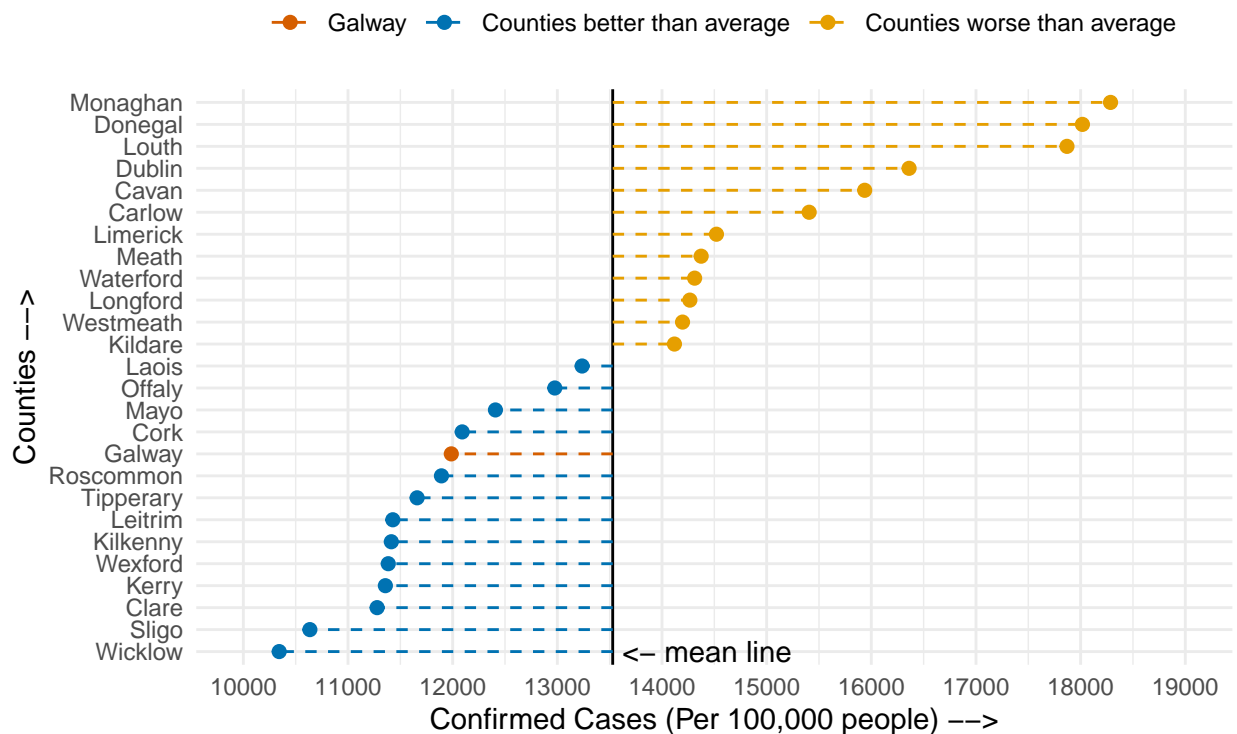
#for plotting the graph
IRL_Covid19_2021_12_21 %>%
  # creating color column for plotting color different for galay, less than mean and
  #more than mean
  mutate(color = ifelse(CountyName == "Galway", "1",
                        ifelse(ConfirmedC_per_100k < mean_daily_cases, "2", "3"))) %>%
  #loading aesthetics for the graphy reordering for sorting
  ggplot(aes(x = ConfirmedC_per_100k, y = reorder(CountyName,ConfirmedC_per_100k)))+
  #geom_point for dot plot
  geom_point(size = 2, aes(color = color))+
  #mean line
  geom_vline(aes(xintercept = mean_daily_cases))+
  #support line
  geom_linerange(aes(xmin = mean_daily_cases, xmax = ConfirmedC_per_100k, color = color),
                linetype = "dashed")+
  #setting colors and labels
  scale_color_manual(values = c(palette[6], palette[5], palette[1]),
                    labels = c("Galway", "Counties better than average",
                              "Counties worse than average"))+
  #changeing x ticks for graph
  scale_x_continuous(limits = c(10000, 19000),
                    breaks = seq(10000, 19000, by = 1000),
                    name = "Confirmed Cases (Per 100,000 people -->")+
  scale_y_discrete(name = "Counties -->")+
  #adding the mean line mark
  annotate(x=mean_daily_cases+900, y=1, label="<- mean line", color="black",
```

```

    geom = "text", lineheight = .6)+
#theme set
theme_minimal()+
#title
ggtitle("Confirmed Cases for counties compared to average on 21st Dec 21")+
#caption
labs(caption = "Source: Covid Data Ireland")+
#for legend position
theme(legend.position = "top",
      legend.title = element_blank())

```

## Confirmed Cases for counties compared to average on 21st Dec 21



Source: Covid Data Ireland

```

#normalising for plotting
IRL_Covid19_plot2 <- IRL_counties_Covid19%>%
  mutate(ConfirmedC_per_100k = round(100000 * ConfirmedC/Population,1))%>%
  mutate(DailyCCase_per_100k = round(100000 * DailyCCase/Population, 1))

# getting a list of means by counties and taking the first and last form highest
#and lowest values of cumulative data
mean <- IRL_Covid19_plot2 %>%
  st_drop_geometry() %>%
  filter(TimeStamp == "2021-12-21") %>%
  arrange(ConfirmedC_per_100k) %>%
  select(CountyName, ConfirmedC_per_100k)

# get the data of galway and county with lowest and highest cumulative scores

```

```

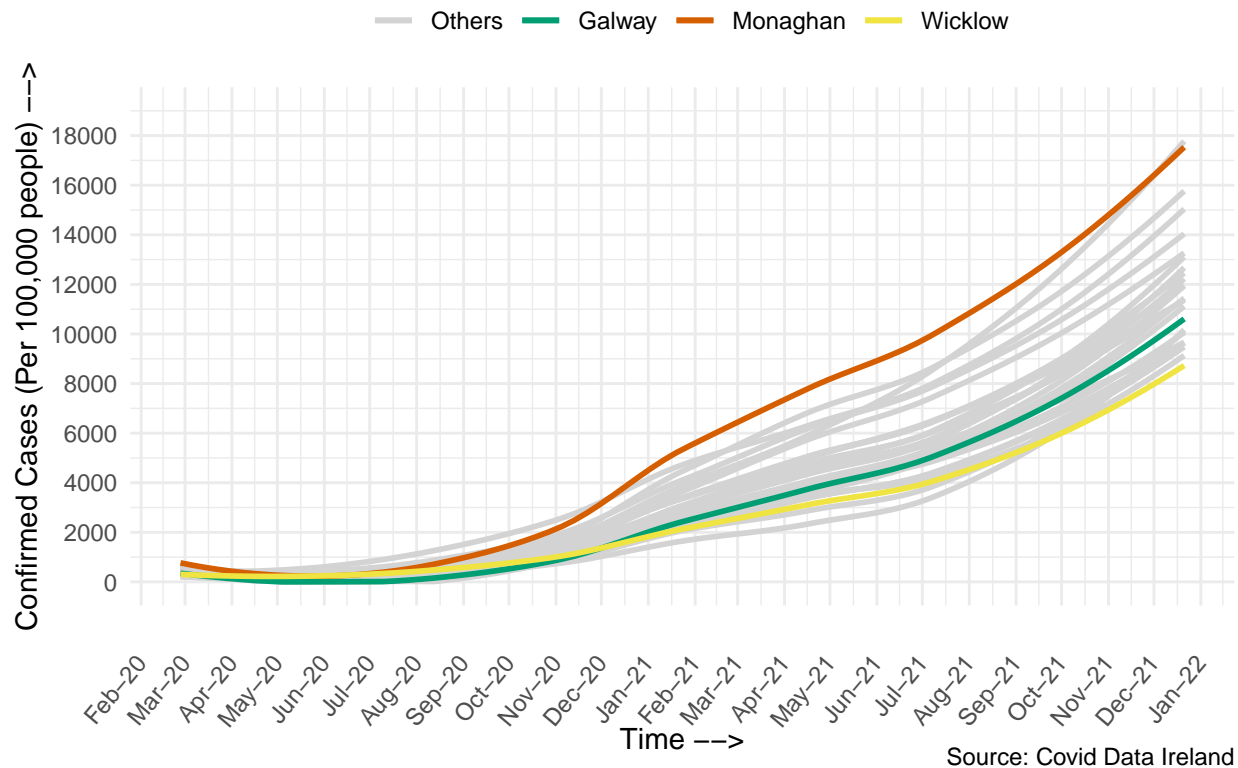
select_county_data<- IRL_Covid19_plot2%>%
  filter(CountyName %in% c("Galway", head(mean$CountyName, 1), tail(mean$CountyName, 1)))

# for rest of the counties
other_counties<- IRL_Covid19_plot2%>%
  filter(!CountyName %in% c("Galway", head(mean$CountyName, 1), tail(mean$CountyName, 1)))

#for plotting plot2
IRL_Covid19_plot2 %>%
  #loading aesthetics for graphy
  ggplot(aes(x = TimeStamp, y=ConfirmedC_per_100k, color = color))+
  # for other counties a faded grey shade
  geom_smooth(data = other_counties, aes(group = CountyName, colour = "#d3d3d3" ),
              size = 1, alpha = 0.9, na.rm = TRUE, method = "loess", se = FALSE)+
  # individual colors for galway and highest, lowest counties
  geom_smooth(data = select_county_data, aes(group = CountyName, color = CountyName),
              size = 1, alpha = 0.8, na.rm = TRUE, method = "loess", se = FALSE)+
  # settings colors
  scale_color_manual(values = c("#d3d3d3", palette[3], palette[6] , palette[4]),
                    labels = c("Others", "Galway", "Monaghan", "Wicklow"))+
  # setting labels on a sequence of values for y axis
  scale_y_continuous(limits = c(0, 19000),
                    breaks = seq(0, 19000, by = 2000),
                    name = "Confirmed Cases (Per 100,000 people -->))+
  # setting dates on the x axis
  scale_x_date(date_breaks = "months", date_labels = "%b-%y",
              name = "Time -->))+
  #title
  ggtitle("Cumulative Covid cases in Galway compared to Other counties")+
  #labels
  labs(caption = "Source: Covid Data Ireland")+
  theme_minimal()+
  theme(axis.text.x = element_text(angle = 50, vjust = 0.5, hjust=1),
        axis.title.x = element_text(vjust = -2.5),
        legend.position = "top",
        legend.title = element_blank())

```

## Cumulative Covid cases in Galway compared to Other counties



```
library(tidyr)
library(RColorBrewer)

#normalise on per 100k population
plot_3 <- IRL_counties_Covid19 %>%
  mutate(ConfirmedC_per_100k = round(100000 * ConfirmedC/Population,1))%>%
  mutate(DailyCCase_per_100k = round(100000 * DailyCCase/Population, 1))

#get the cases for all counties on a 4 week period
plot_3 <- plot_3[plot_3$TimeStamp == ymd("2021-01-01") | plot_3$TimeStamp ==ymd("2021-01-29"),]%>%
  select(CountyName, ConfirmedC_per_100k, TimeStamp) %>%
  #subtract the previous value from current as in the dataframe we have two rows for one county
  # with the data between the range of 4 weeks
  mutate(new_cases_4weeks = ConfirmedC_per_100k - lag(ConfirmedC_per_100k,
                                                    default = first(ConfirmedC_per_100k)))

# as we have two entries for each county bases on difference we only keep the difference
# value between the same county and discard the the other
plot_3 <- plot_3[seq(2, nrow(plot_3), 2), ]
#get the mean value of new cases found during the 4 week period
mean = mean(plot_3$new_cases_4weeks)

# selecting columns necessary for plotting
plot_3<- plot_3 %>%
  select(CountyName, geometry, new_cases_4weeks)
```

```

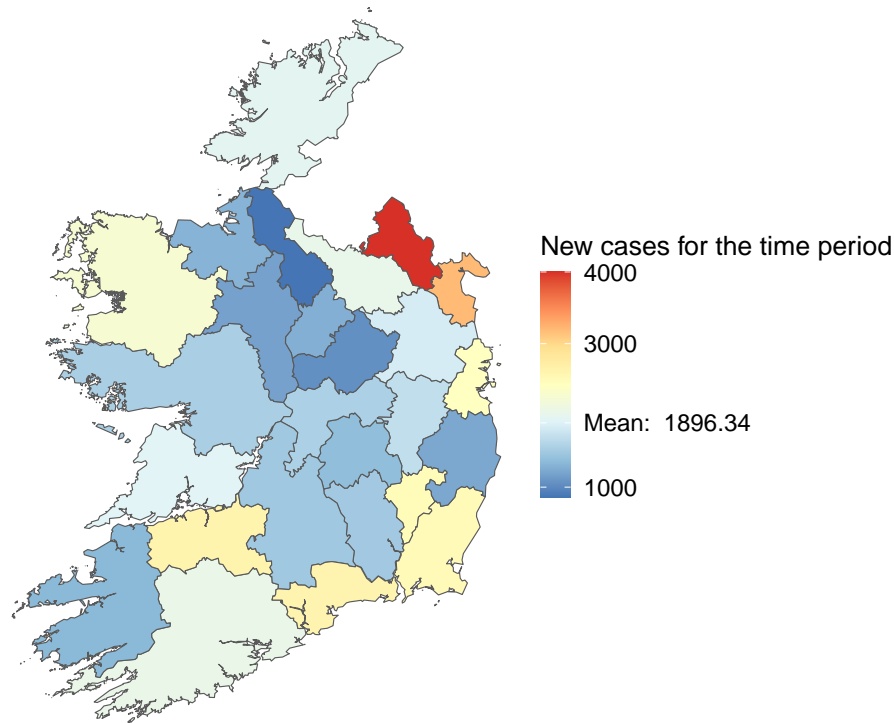
ggplot(plot_3) +
  #sf geom for map based on geometry filling the case data
  geom_sf(aes(fill = new_cases_4weeks))+
  #using the diverging color palette
  scale_fill_distiller(palette = "RdYlBu",
    #manually defining breaks

    breaks = c(4000,3000 ,mean, 1000),

    #labels for the break
    labels = c("4000", "3000", paste("Mean: ", round(mean,2)), 1000),
    # for positioning the colorbar
    guide = guide_colorbar(
      label.position = "right",
      title = "New cases for the time period",
      barwidth = grid::unit(0.4, "cm"),
      barheight = grid::unit(3, "cm")))+
  ggtitle("Comparing new covid cases in counties of Ireland from 1st Jan 21 to 29th Jan around the mean")
  labs(caption = "Source: Covid Data Ireland")+
  theme_void()+
  #formatting the graph
  theme(plot.title = element_text(size = 10),
    legend.title = element_text(size = 10))

```

Comparing new covid cases in counties of Ireland from 1st Jan 21 to 29th Jan around the m



Source: Covid Data Ireland