Covid-19: A Look at Texas Cities

Samuel Higgins

```
library(tidyverse)
```

```
## -- Attaching packages ------------------------------------ tidyverse 1.3.0 --

## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.1     v dplyr   1.0.0
## v tidyr   1.1.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## -- Conflicts --------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
library(ggmap)
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.

## Please cite ggmap if you use it! See citation("ggmap") for details.
```

```
library(maps)
```

```
##
## Attaching package: 'maps'

## The following object is masked from 'package:purrr':
##
##     map
```

```r
library(mapdata)

covid <- read.csv("C:/Users/samue/Documents/College Notes/PyRe/Data sets/COVID-19/us_counties_Jan17_21.
glimpse(covid)
```

```
## Rows: 933,498
## Columns: 6
## $ date   <chr> "2020-01-21", "2020-01-22", "2020-01-23", "2020-01-24", "202...
## $ county <chr> "Snohomish", "Snohomish", "Snohomish", "Cook", "Snohomish", ...
## $ state  <chr> "Washington", "Washington", "Washington", "Illinois", "Washi...
## $ fips   <int> 53061, 53061, 53061, 17031, 53061, 6059, 17031, 53061, 4013,...
## $ cases  <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ deaths <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
```

```r
txcovid <- covid %>%
  filter(state == "Texas") %>%
  pivot_wider(names_from = state, values_from = county) %>%
  rename(county_tx = Texas)

glimpse(txcovid)
```

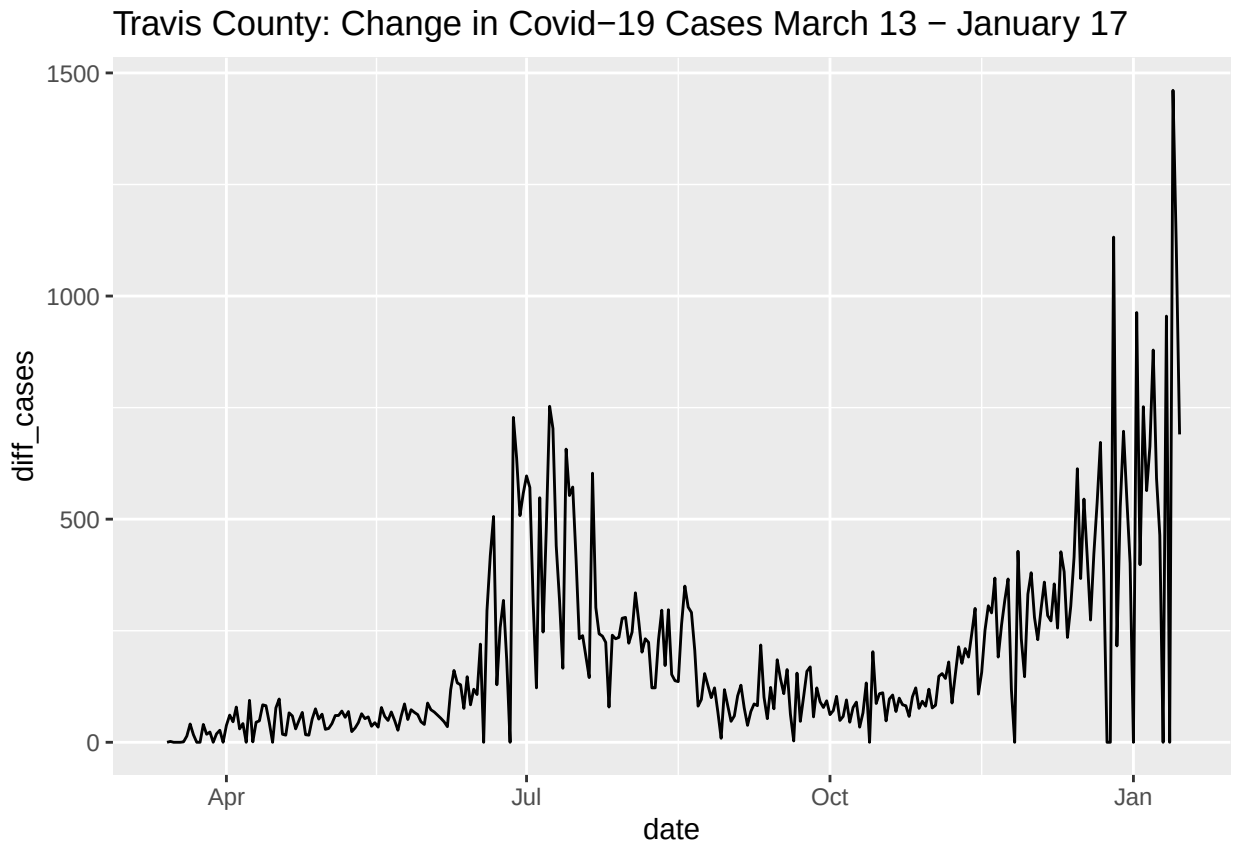
```
## Rows: 70,812
## Columns: 5
## $ date      <chr> "2020-02-12", "2020-02-13", "2020-02-14", "2020-02-15", "...
## $ fips      <int> 48029, 48029, 48029, 48029, 48029, 48029, 48029, 48029, 4...
## $ cases     <int> 1, 2, 2, 2, 2, 2, 2, 2, 4, 4, 4, 10, 10, 10, 10, 11, 1...
## $ deaths    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
## $ county_tx <chr> "Bexar", "Bexar", "Bexar", "Bexar", "Bexar", "Bexar", "Be...
```

```r
txcovid$date <- as.Date(txcovid$date)
```

Travis County

```r
txcovid_travis <- txcovid %>%
  filter(county_tx == "Travis") %>%
  mutate(diff_cases = cases - lag(cases),
         diff_deaths = deaths - lag(deaths))


# Difference of Covid-19 cases over time for Travis County
txcovid_travis %>%
  ggplot(aes(x = date, y = diff_cases)) +
  geom_line() +
  ggtitle("Travis County: Change in Covid-19 Cases March 13 - January 17")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

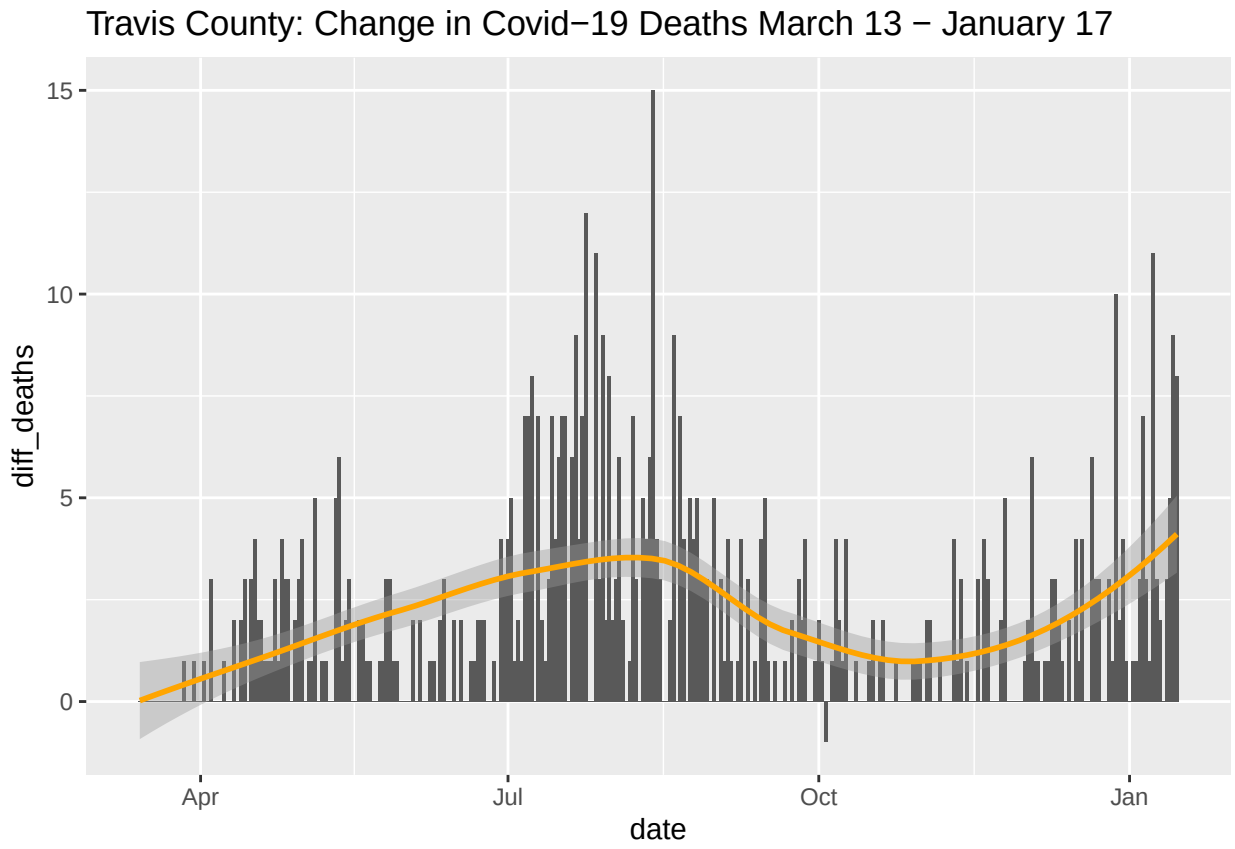



```
# Difference of Covid-19 deaths over time for Travis County
txcovid_travis %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  stat_smooth(color = "orange") +
  ggtitle("Travis County: Change in Covid-19 Deaths March 13 - January 17")
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

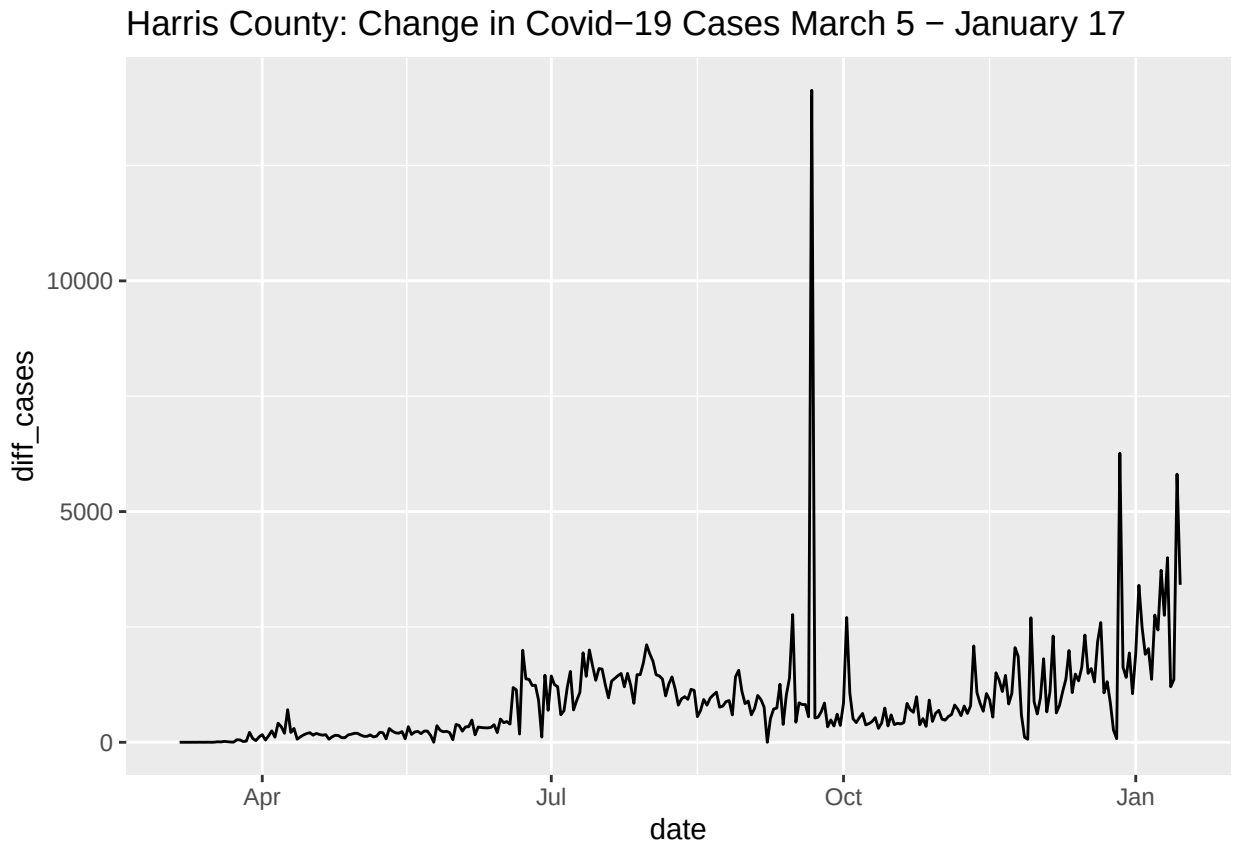
```
## Warning: Removed 1 rows containing missing values (position_stack).
```



Harris County

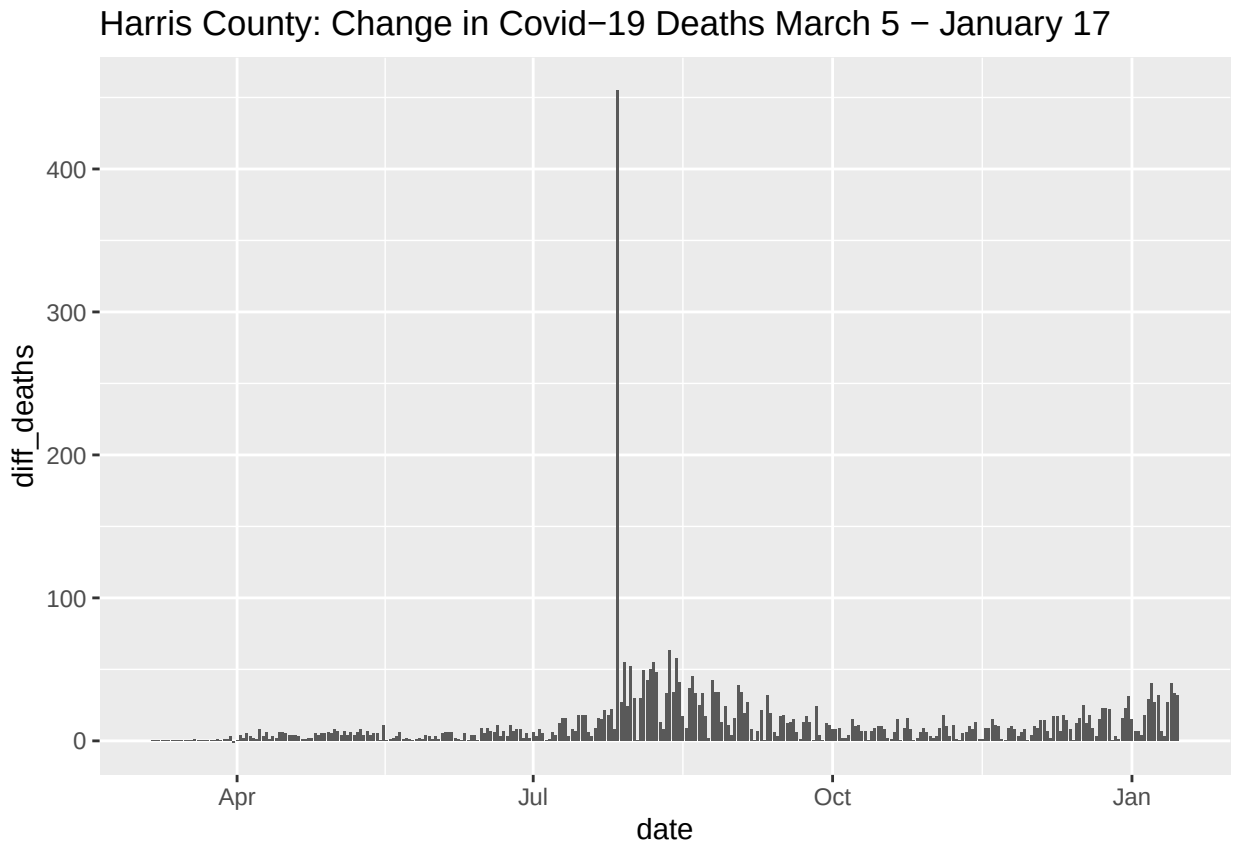
```
txcovid_harris <- txcovid %>%
  filter(county_tx == "Harris") %>%
  mutate(diff_cases = cases - lag(cases),
         diff_deaths = deaths - lag(deaths))

# Difference of Covid-19 cases over time for Harris County
txcovid_harris %>%
  ggplot(aes(x = date, y = diff_cases)) +
  geom_line() +
  ggtitle("Harris County: Change in Covid-19 Cases March 5 - January 17")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```




```
# Difference of Covid-19 deaths over time for Harris County
txcovid_harris %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  ggtitle("Harris County: Change in Covid-19 Deaths March 5 - January 17")
```

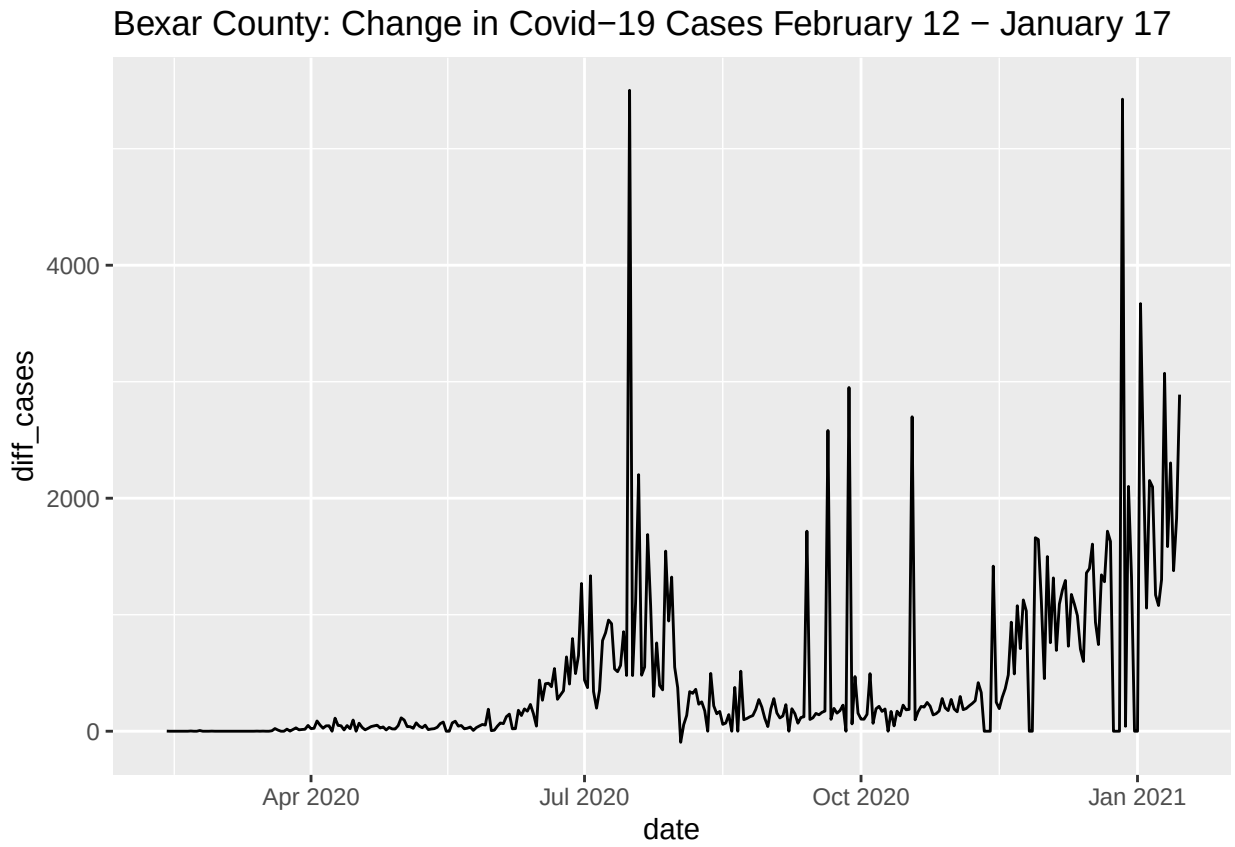
```
## Warning: Removed 1 rows containing missing values (position_stack).
```



Bexar County

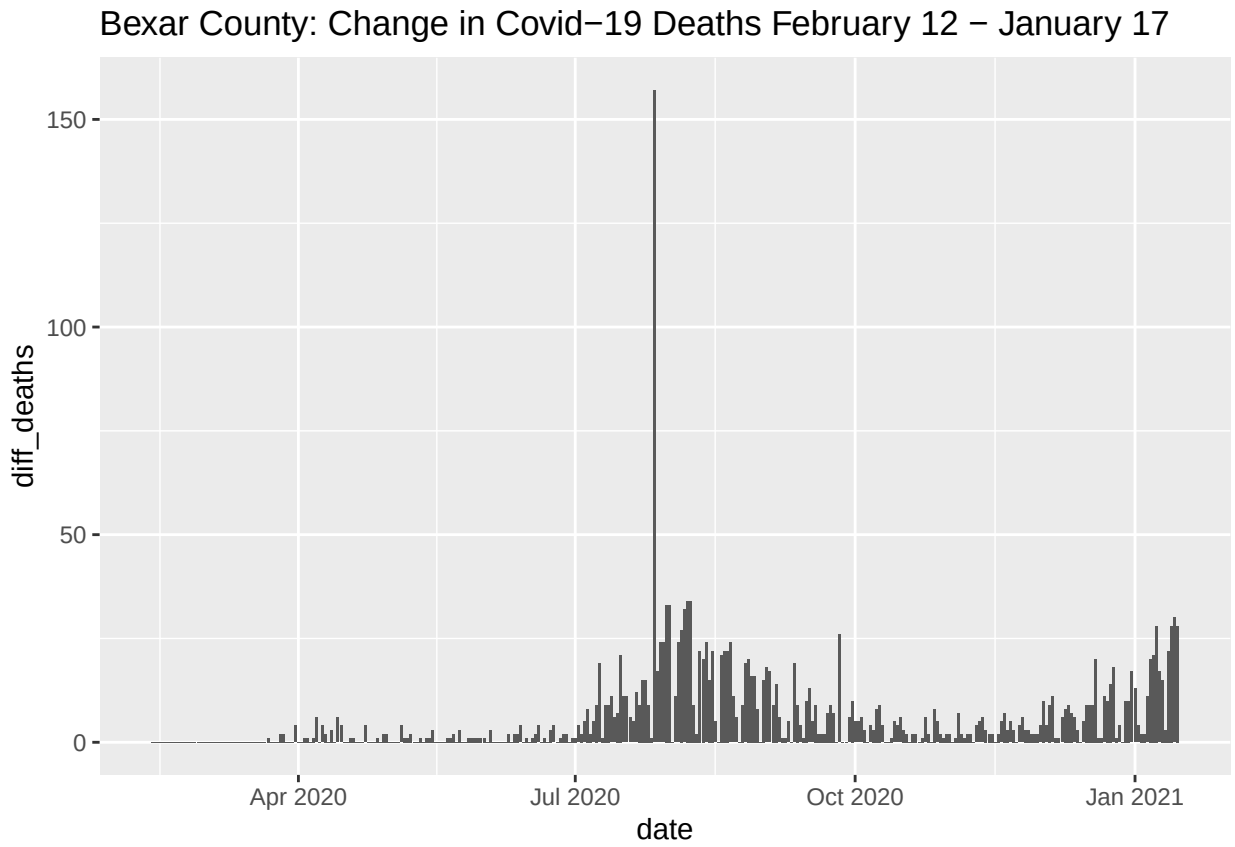
```
txcovid_bexar <- txcovid %>%
  filter(county_tx == "Bexar") %>%
  mutate(diff_cases = cases - lag(cases),
         diff_deaths =  deaths - lag(deaths))

# Difference of Covid-19 cases over time for Bexar County
txcovid_bexar %>%
  ggplot(aes(x = date, y = diff_cases)) +
  geom_line() +
  ggtitle("Bexar County: Change in Covid-19 Cases February 12 - January 17")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```




```
# Difference of Covid-19 deaths over time for Bexar County
txcovid_bexar %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  ggtitle("Bexar County: Change in Covid-19 Deaths February 12 - January 17")
```

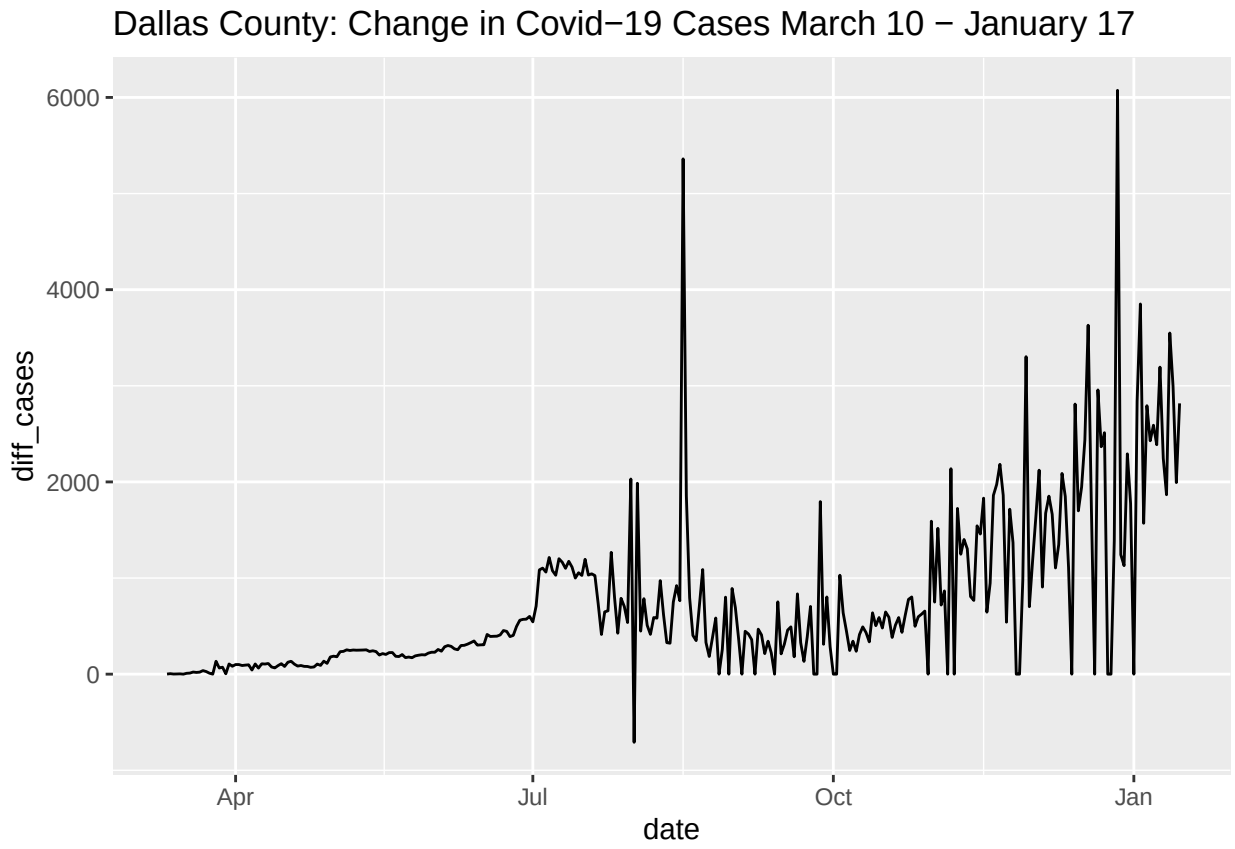
```
## Warning: Removed 1 rows containing missing values (position_stack).
```



Dallas County

```
txcovid_dallas <- txcovid %>%
  filter(county_tx == "Dallas") %>%
  mutate(diff_cases = cases - lag(cases),
         diff_deaths = deaths - lag(deaths))

# Difference of Covid-19 cases over time for Dallas County
txcovid_dallas %>%
  ggplot(aes(x = date, y = diff_cases)) +
  geom_line() +
  ggtitle("Dallas County: Change in Covid-19 Cases March 10 - January 17")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

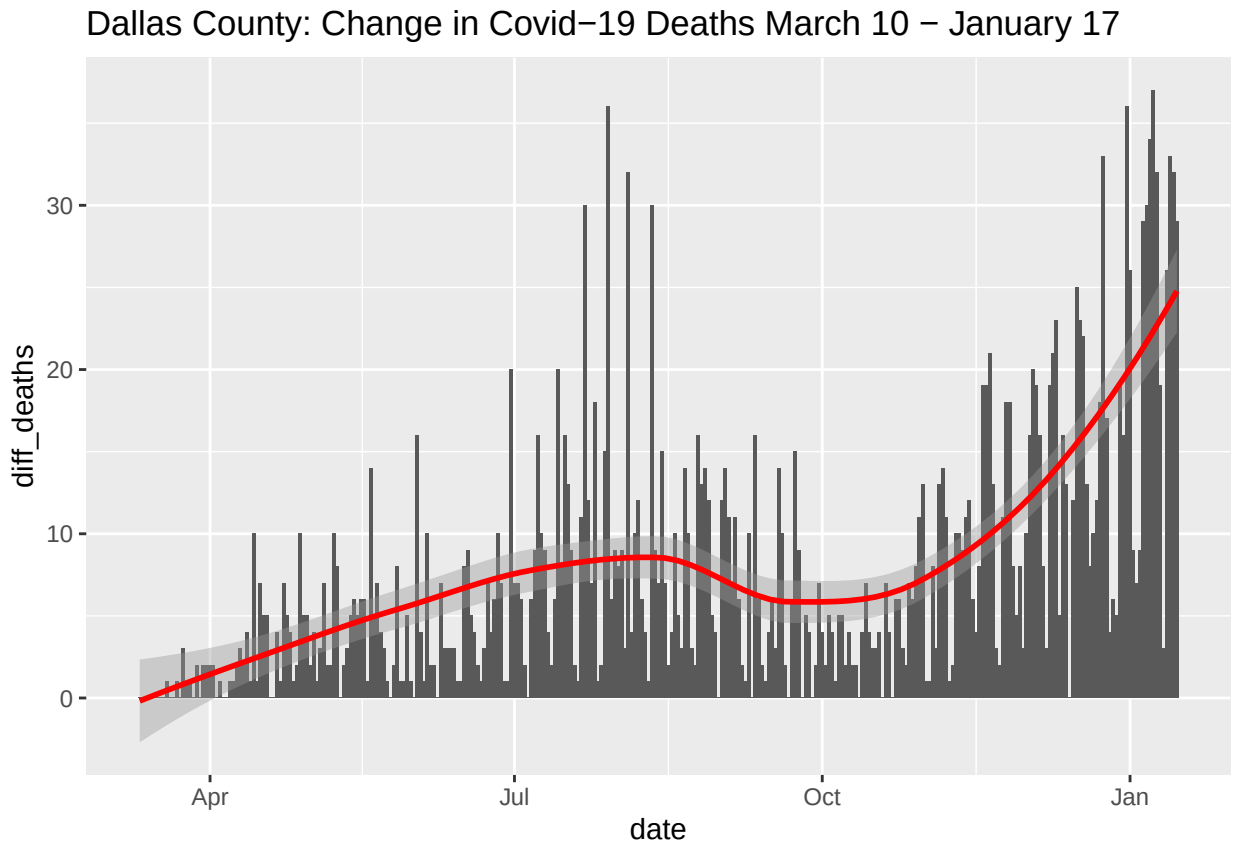



```
# Difference of Covid-19 deaths over time for Dallas County
txcovid_dallas %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  stat_smooth(color = "red") +
  ggtitle("Dallas County: Change in Covid-19 Deaths March 10 - January 17")
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

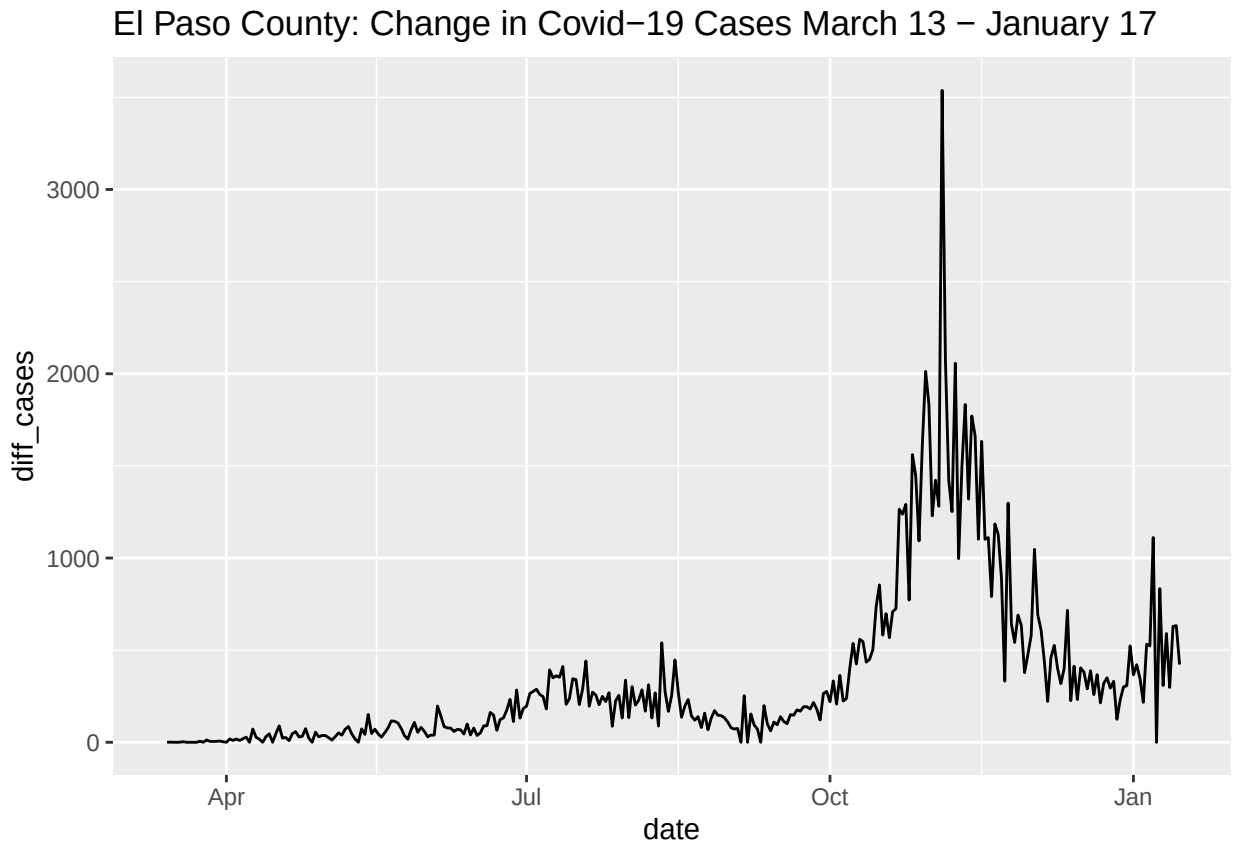
```
## Warning: Removed 1 rows containing missing values (position_stack).
```



El Paso County

```
txcovid_elpaso <- txcovid %>%
  filter(county_tx == "El Paso") %>%
  mutate(diff_cases = cases - lag(cases),
         diff_deaths = deaths - lag(deaths))

# Difference of Covid-19 cases over time for El Paso County
txcovid_elpaso %>%
  ggplot(aes(x = date, y = diff_cases)) +
  geom_line() +
  ggtitle("El Paso County: Change in Covid-19 Cases March 13 - January 17")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

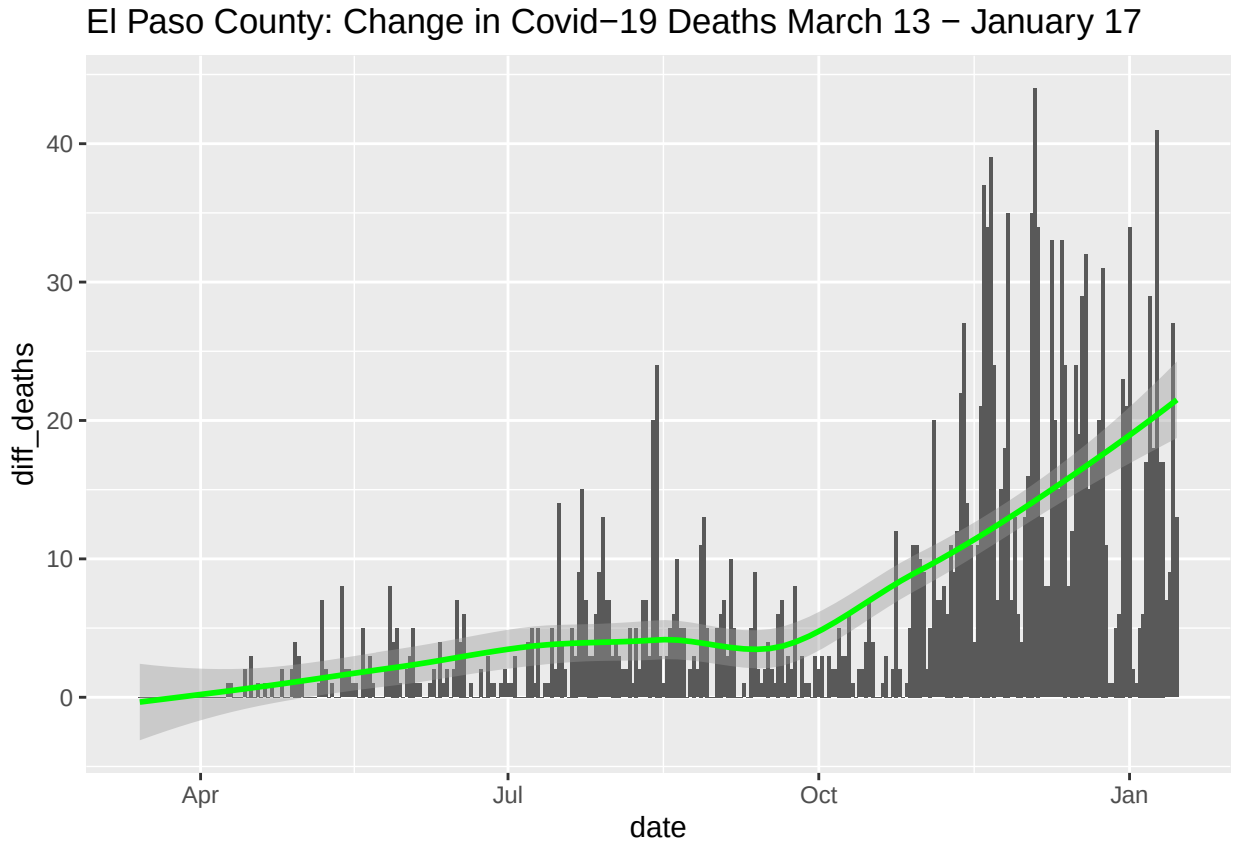



```
# Difference of Covid-19 deaths over time for El Paso County
txcovid_elpaso %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  stat_smooth(color = "green") +
  ggtitle("El Paso County: Change in Covid-19 Deaths March 13 - January 17")
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (position_stack).
```



Covid-19 Cases Choropleth Map

```
txcovid_current <- txcovid %>%
  group_by(county_tx, fips) %>%
  summarise(current_cases = max(cases),
            total_deaths = max(deaths)) %>%
  na.omit
```

```
## `summarise()` regrouping output by 'county_tx' (override with `.groups` argument)
```

```
glimpse(txcovid_current)
```

```
## Rows: 254
## Columns: 4
## Groups: county_tx [254]
## $ county_tx     <chr> "Anderson", "Andrews", "Angelina", "Aransas", "Archer...
## $ fips          <int> 48001, 48003, 48005, 48007, 48009, 48011, 48013, 4801...
## $ current_cases <int> 5542, 1558, 6552, 867, 688, 125, 3711, 1331, 740, 807...
## $ total_deaths  <int> 73, 36, 189, 26, 8, 6, 85, 17, 14, 19, 54, 9, 53, 196...
```

```r
states <- map_data("state")
tx_state <- subset(states, region == "texas")
counties <- map_data("county")
tx_county <- subset(counties, region == "texas")

tx_base <- ggplot(data = tx_state, mapping = aes(x = long, y = lat, group = group)) +
  coord_fixed(1.3) +
  geom_polygon(color = "black", fill = "gray")

remove_axis <- theme(
  axis.text = element_blank(),
  axis.line = element_blank(),
  axis.ticks = element_blank(),
  panel.border = element_blank(),
  panel.grid = element_blank(),
  axis.title = element_blank()
)

capfirst <- function(s){
  paste(toupper(substring(s, 1, 1)), substring(s, 2), sep = "")
}

tx_county$subregion <- capfirst(tx_county$subregion)

covid_map <- left_join(txcovid_current, tx_county, by = c("county_tx" = "subregion"))

cmap <- tx_base +
  geom_polygon(data = covid_map, aes(fill = current_cases), color = "white") +
  geom_polygon(color = "black", fill = NA) +
  theme_bw() +
  remove_axis

cmap2 <- cmap +
  scale_fill_gradient(low = "#FEDFDF", high = "#FD100F",
                      breaks = c(250000, 200000, 150000, 100000, 50000, 10000),
                      name = "Current Covid-19 Cases")
cmap2
```

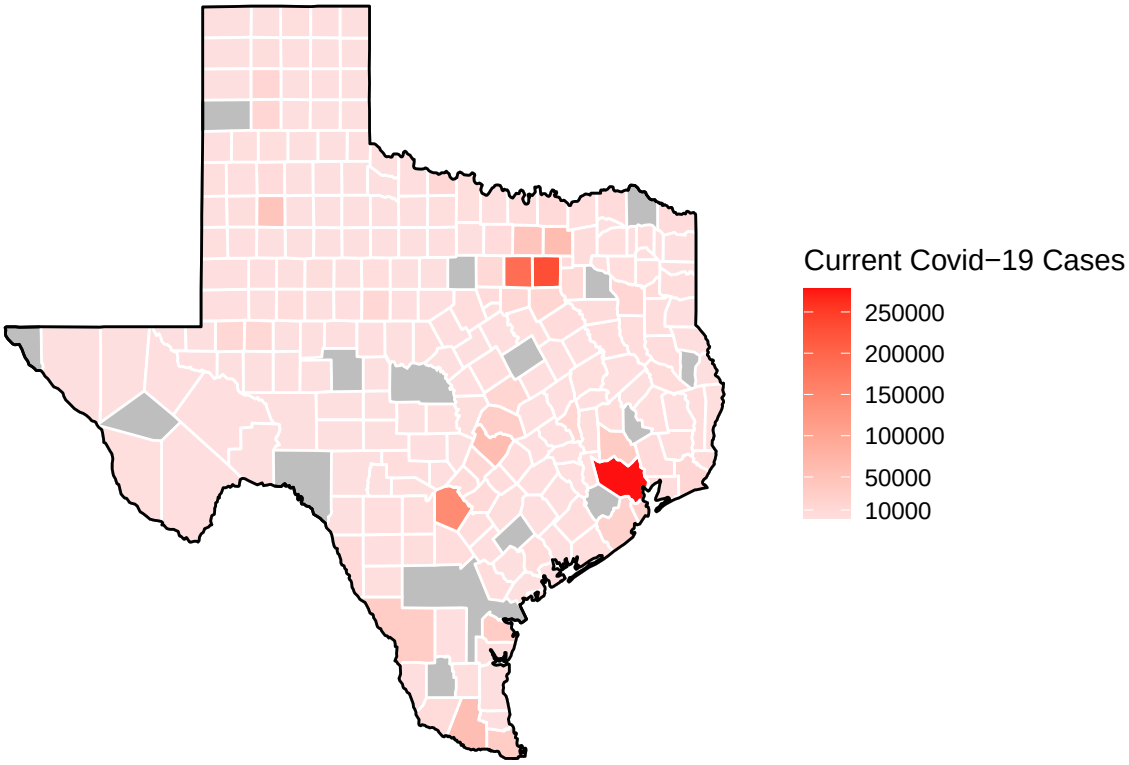


Table of Current Covid-19 Cases and Total Deaths

```
kable(head(txcovid_current, n = 50), caption = "Covid-19 Current Cases and Deaths by County") %>%
  kable_styling(latex_options = c("striped", "condensed"), stripe_color = "#B2BEB5")
```

```
kable(tail(txcovid_current, n = 50)) %>% kable_styling(latex_options = c("striped", "condensed"),
                                                       stripe_color = "#B2BEB5")
```

Table 1: Covid-19 Current Cases and Deaths by County

county_tx	fips	current_cases	total_deaths
Anderson	48001	5542	73
Andrews	48003	1558	36
Angelina	48005	6552	189
Aransas	48007	867	26
Archer	48009	688	8
Armstrong	48011	125	6
Atascosa	48013	3711	85
Austin	48015	1331	17
Bailey	48017	740	14
Bandera	48019	807	19
Bastrop	48021	4298	54
Baylor	48023	272	9
Bee	48025	3017	53
Bell	48027	16406	196
Bexar	48029	143515	1970
Blanco	48031	317	11
Borden	48033	18	0
Bosque	48035	979	20
Bowie	48037	4720	137
Brazoria	48039	25117	269
Brazos	48041	16634	157
Brewster	48043	923	8
Briscoe	48045	111	3
Brooks	48047	676	29
Brown	48049	3382	93
Burleson	48051	1089	21
Burnet	48053	2165	33
Caldwell	48055	2705	54
Calhoun	48057	1343	11
Callahan	48059	926	27
Cameron	48061	32453	1222
Camp	48063	928	32
Carson	48065	329	11
Cass	48067	1446	57
Castro	48069	829	21
Chambers	48071	3431	31
Cherokee	48073	3469	86
Childress	48075	1276	13
Clay	48077	870	13
Cochran	48079	214	11
Coke	48081	425	10
Coleman	48083	572	19
Collin	48085	61803	461
Collingsworth	48087	234	8
Colorado	48089	1204	17
Comal	48091	7234	201
Comanche	48093	1116	35
Concho	48095	240	4
Cooke	48097	2794	43
Coryell	48099	4239	39

county_tx	fips	current_cases	total_deaths
San Patricio	48409	3248	109
San Saba	48411	479	15
Schleicher	48413	220	6
Scurry	48415	2266	43
Shackelford	48417	179	1
Shelby	48419	1233	45
Sherman	48421	117	11
Smith	48423	15407	298
Somervell	48425	877	8
Starr	48427	7167	212
Stephens	48429	624	18
Sterling	48431	100	4
Stonewall	48433	128	3
Sutton	48435	404	6
Swisher	48437	674	14
Tarrant	48439	187863	1825
Taylor	48441	13084	247
Terrell	48443	72	1
Terry	48445	1475	46
Throckmorton	48447	56	3
Titus	48449	3049	53
Tom Green	48451	14610	201
Travis	48453	60084	606
Trinity	48455	474	12
Tyler	48457	1418	18
Upshur	48459	2356	39
Upton	48461	282	7
Uvalde	48463	2538	47
Val Verde	48465	6281	157
Van Zandt	48467	3154	74
Victoria	48469	6600	132
Walker	48471	6808	93
Waller	48473	2529	25
Ward	48475	920	12
Washington	48477	1641	69
Webb	48479	32658	543
Wharton	48481	2893	81
Wheeler	48483	426	9
Wichita	48485	13179	255
Wilbarger	48487	1686	41
Willacy	48489	1909	69
Williamson	48491	29796	262
Wilson	48493	2752	40
Winkler	48495	627	15
Wise	48497	5336	81
Wood	48499	2517	85
Yoakum	48501	801	23
Young	48503	1662	31
Zapata	48505	1291	18
Zavala	48507	1233	28