Covid-19: A Look at Texas Cities

Samuel Higgins

```r
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()
```

```r
library(kableExtra)
```

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

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

```r
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.
```

```r
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/us-counties.csv")
glimpse(covid)
```

```
## Rows: 644,678
## 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: 48,235
## 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, 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, 0, ...
## $ county_tx <chr> "Bexar", "Bexar", "Bexar", "Bexar", "Bexar", "Bexar", "Be...
```

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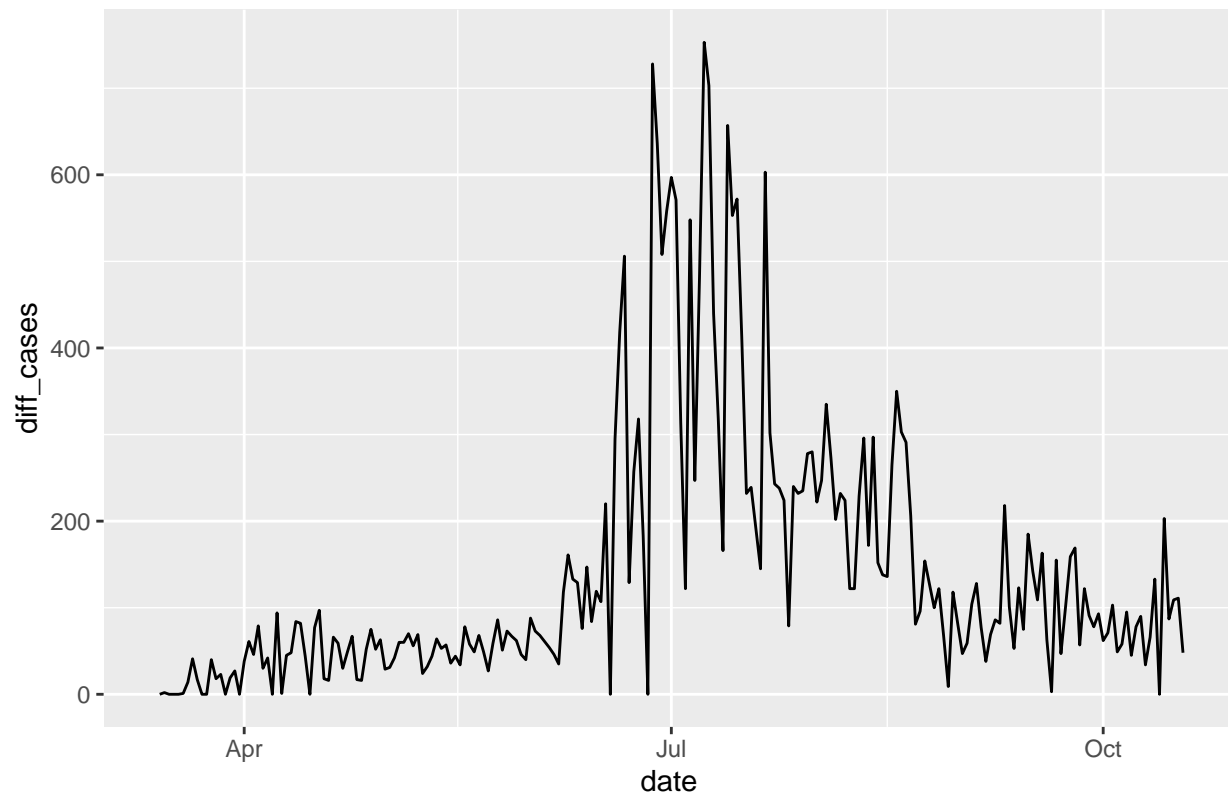
Travis County

```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 - October 18")
```

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

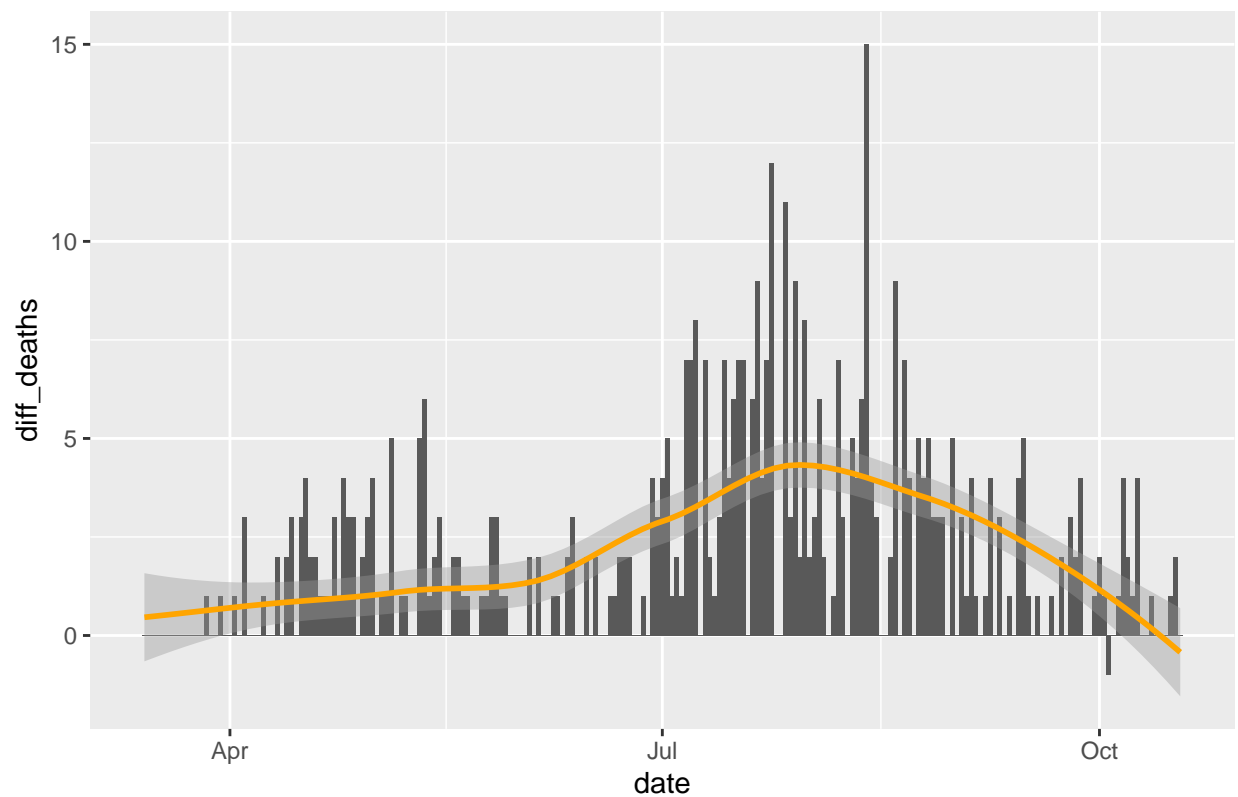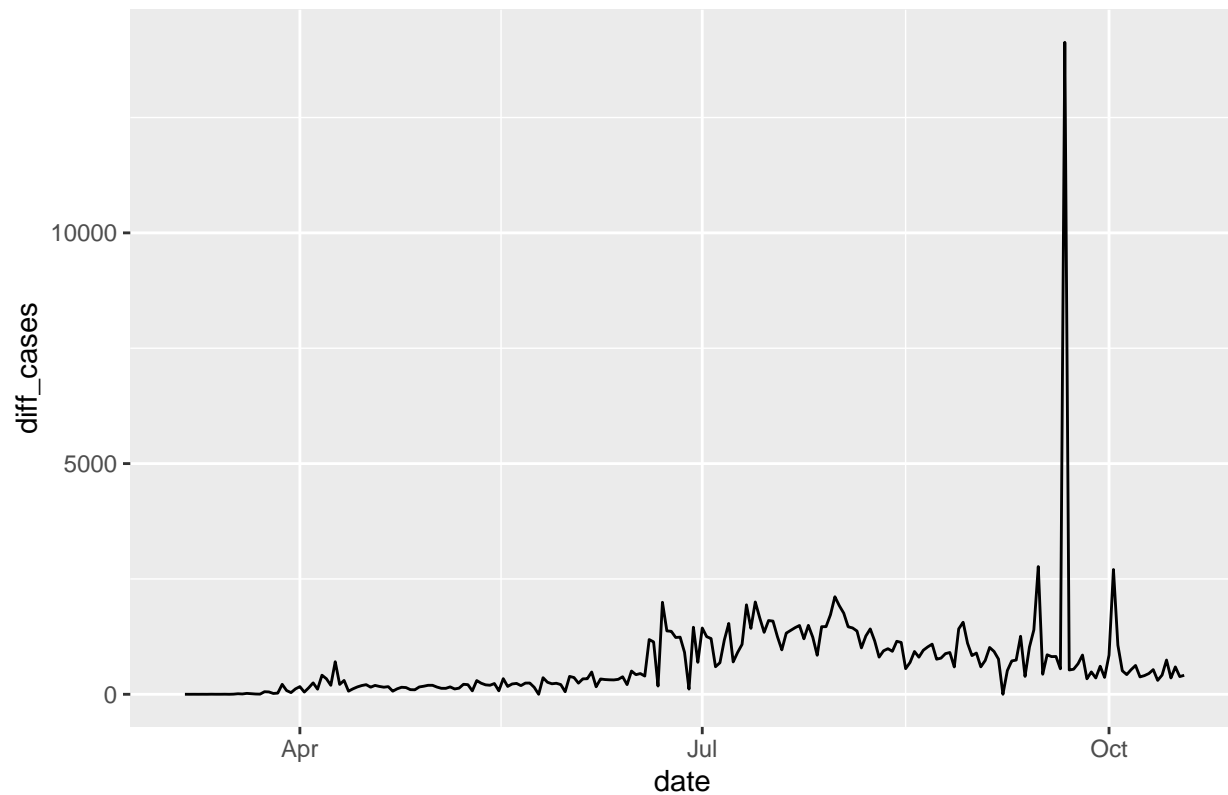## Travis County: Change in Covid−19 Cases March 13 − October 18


```r
# 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 - October 18")
```

```
## `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).
```

Travis County: Change in Covid−19 Deaths March 13 − October 18

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 - October 18")
```

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

Harris County: Change in Covid−19 Cases March 5 – October 18


```r
# 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 - October 18")
```

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

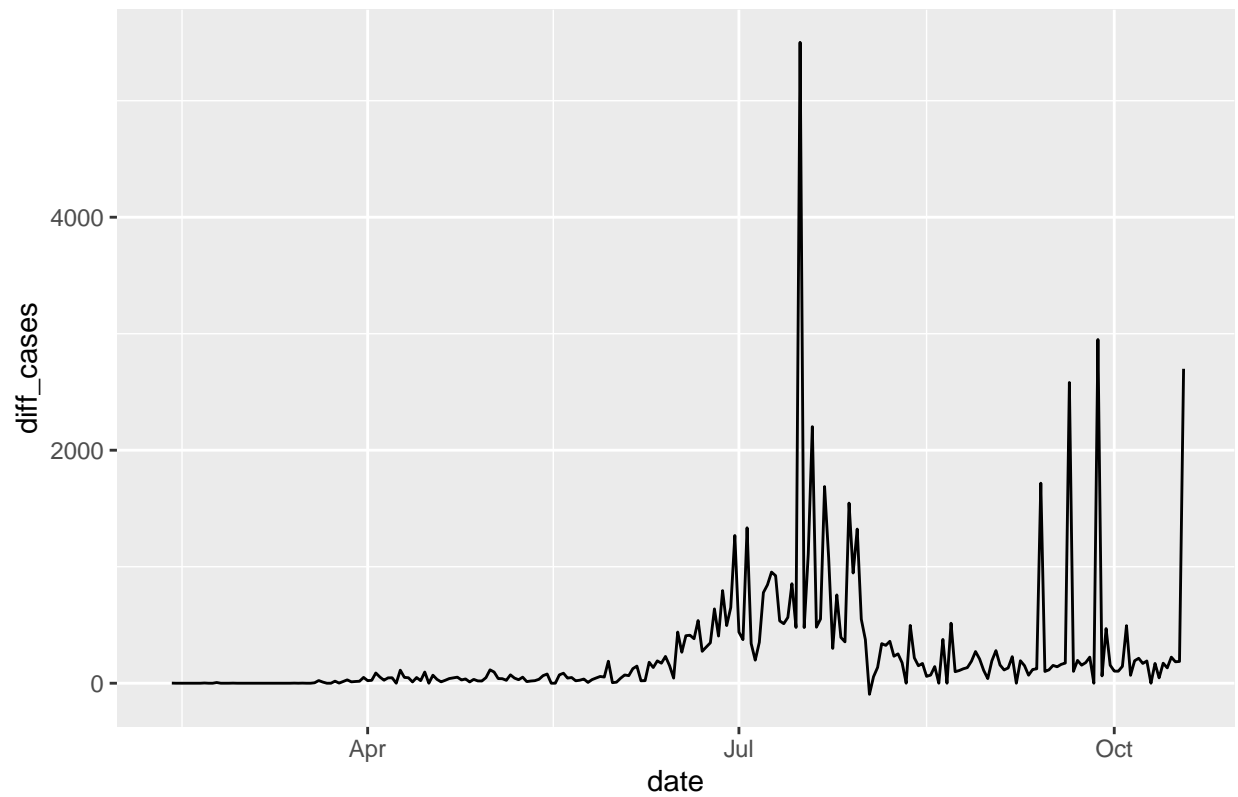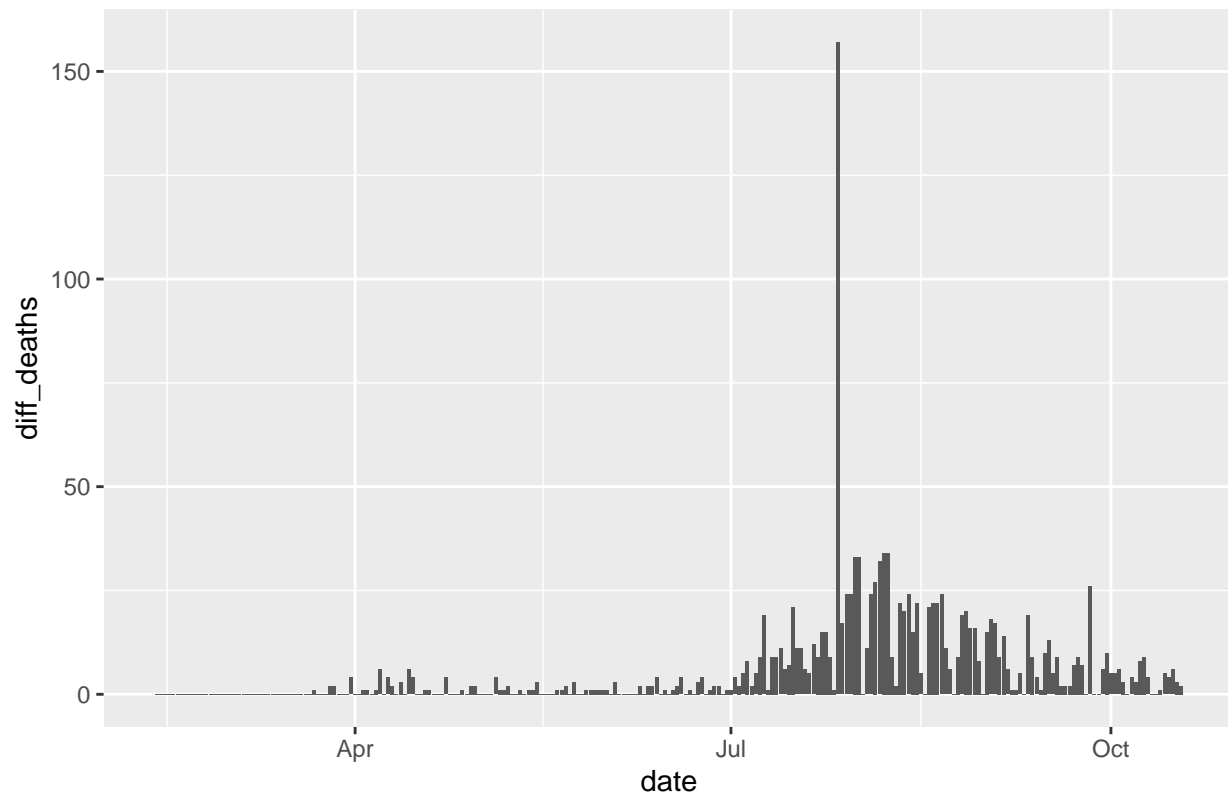Harris County: Change in Covid−19 Deaths March 5 − October 18
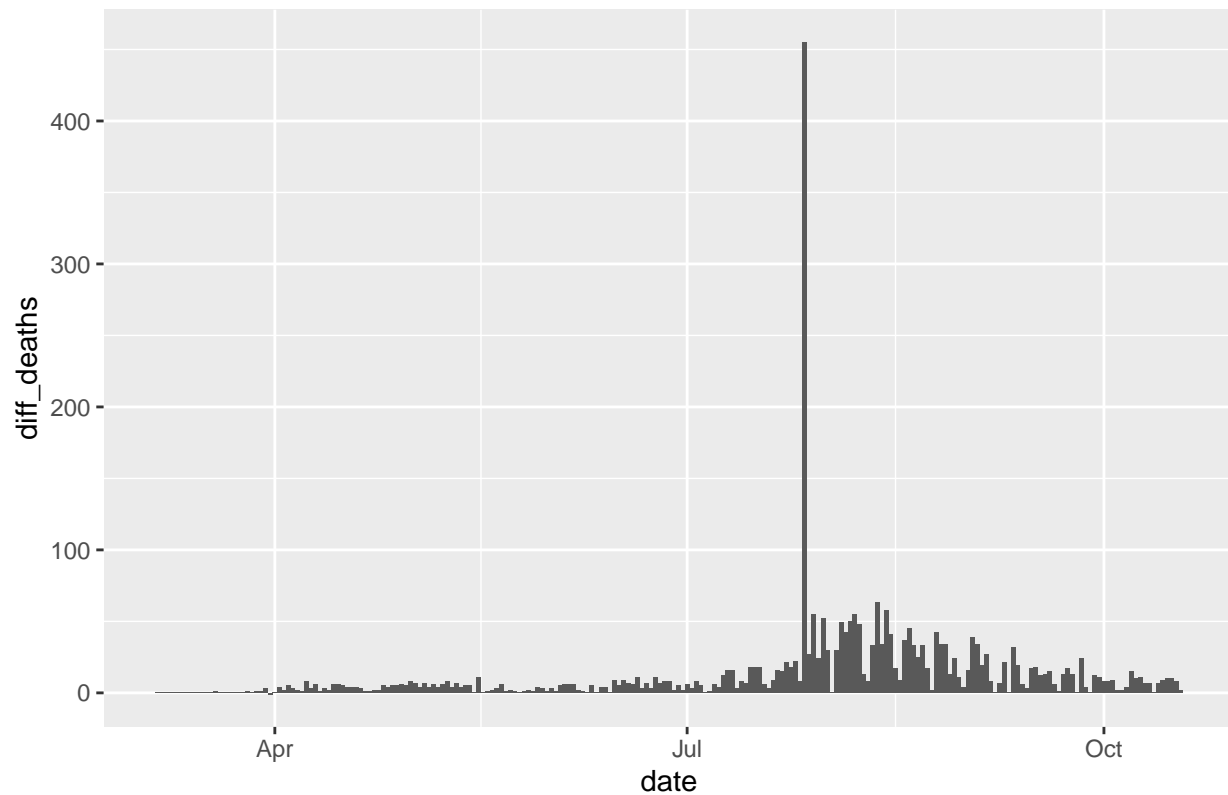
Bexar County

```
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 - October 18")
```

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

Bexar County: Change in Covid−19 Cases February 12 − October 18


```
# 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 - October 18")
```

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

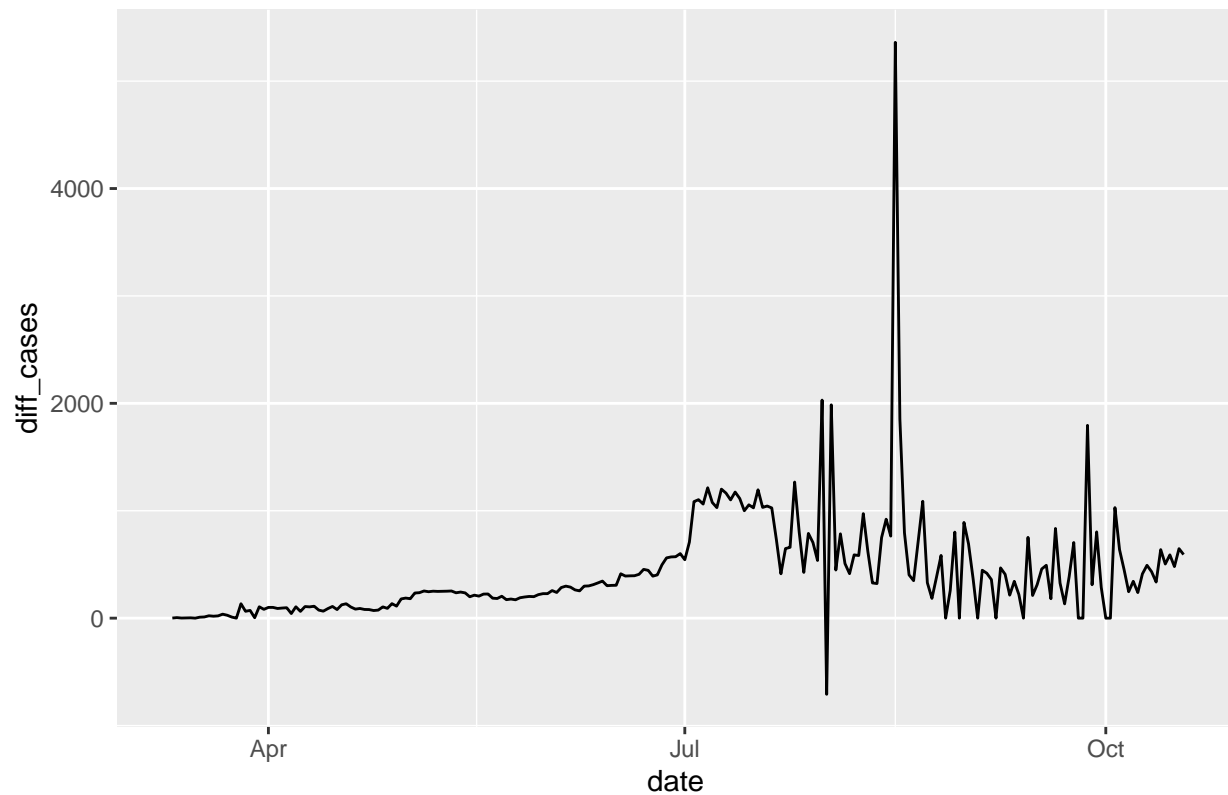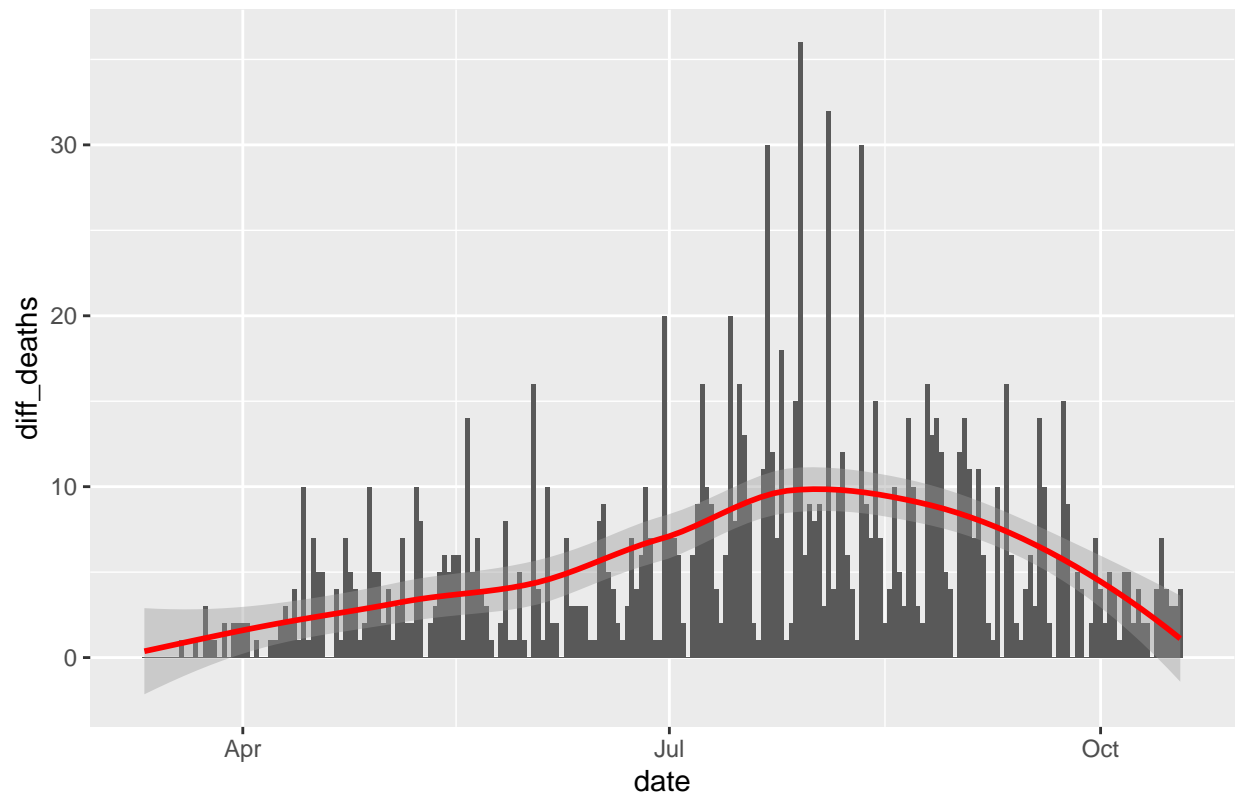Bexar County: Change in Covid−19 Deaths February 12 – October 18


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 - October 18")
```

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

## Dallas County: Change in Covid−19 Cases March 10 − October 18


```r
# 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 - October 18")
```

```
## '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).
```

Dallas County: Change in Covid−19 Deaths March 10 − October 18

El Paso County

```r
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 - October 18")
```

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

El Paso County: Change in Covid−19 Cases March 13 − October 18
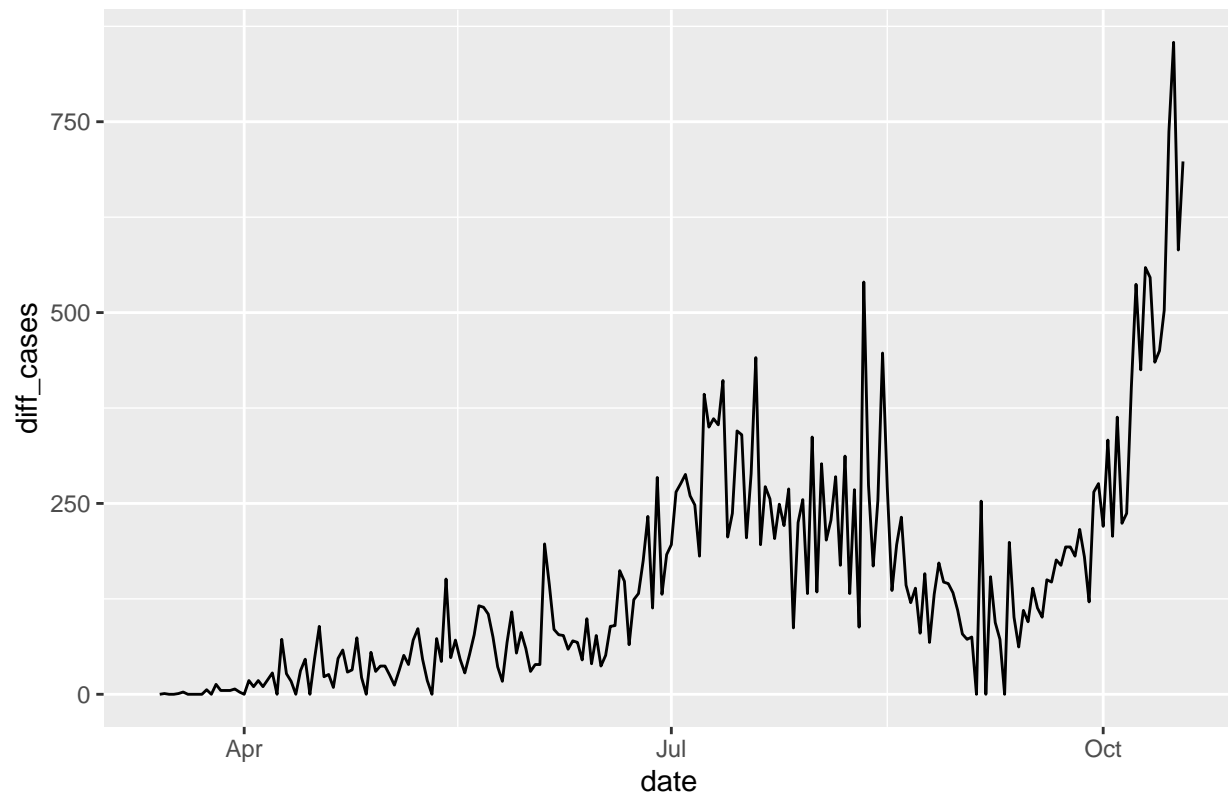

```r
# Difference of Covid-19 deaths over time for El Paso County
txcovid_elpaso %>%
  ggplot(aes(x = date, y = diff_deaths)) +
  geom_bar(stat = "identity") +
  stat_smooth(color = "green") +
  ggtitle("El Paso County: Change in Covid-19 Deaths March 13 - October 18")
```

```
## `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: Change in Covid–19 Deaths March 13 – October 18
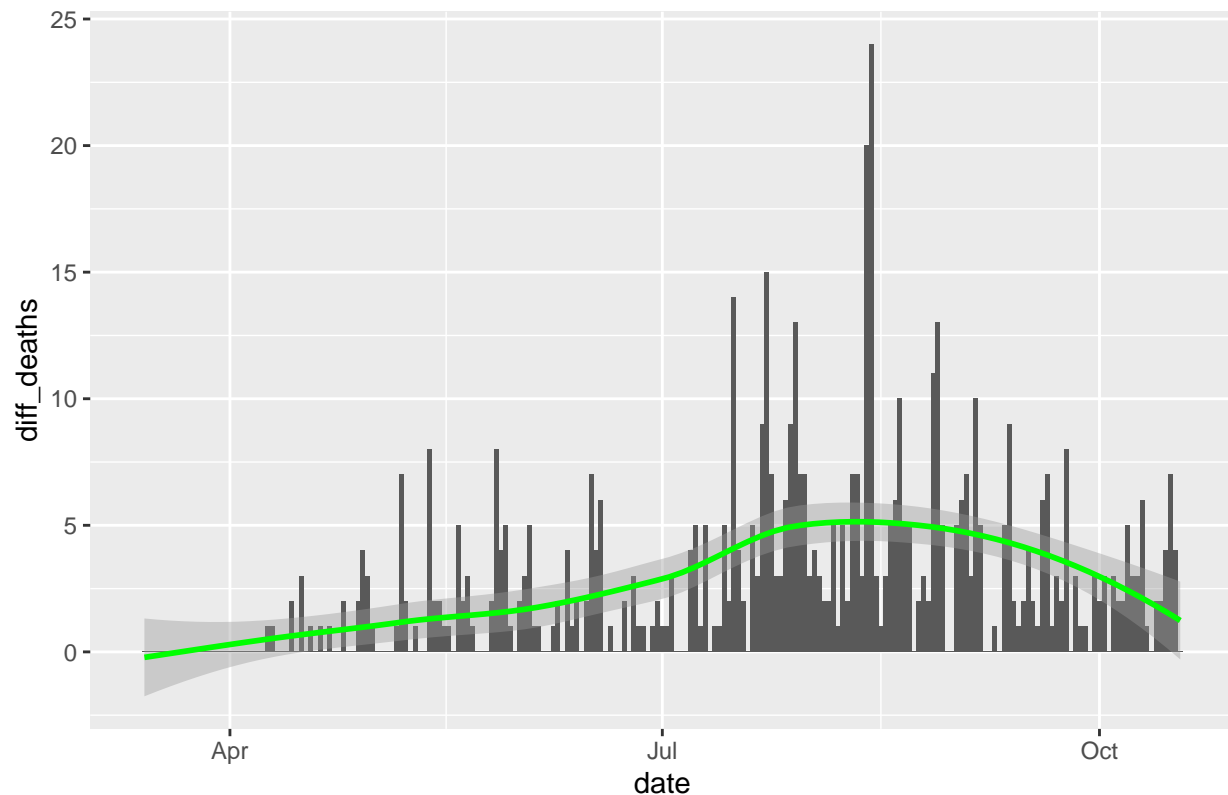
Covid-19 Cases Choropleth Map

```
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: 253
## Columns: 4
## Groups: county_tx [253]
## $ county_tx     <chr> "Anderson", "Andrews", "Angelina", "Aransas", "Archer...
## $ fips          <int> 48001, 48003, 48005, 48007, 48009, 48011, 48013, 4801...
## $ current_cases <int> 2909, 547, 2753, 341, 114, 18, 1254, 507, 264, 179, 1...
## $ total_deaths  <int> 38, 10, 88, 18, 1, 1, 35, 9, 5, 8, 31, 2, 37, 96, 138...
```

```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(150000, 125000, 100000, 75000, 50000, 25000),
                      name = "Current Covid-19 Cases")
cmap2
```
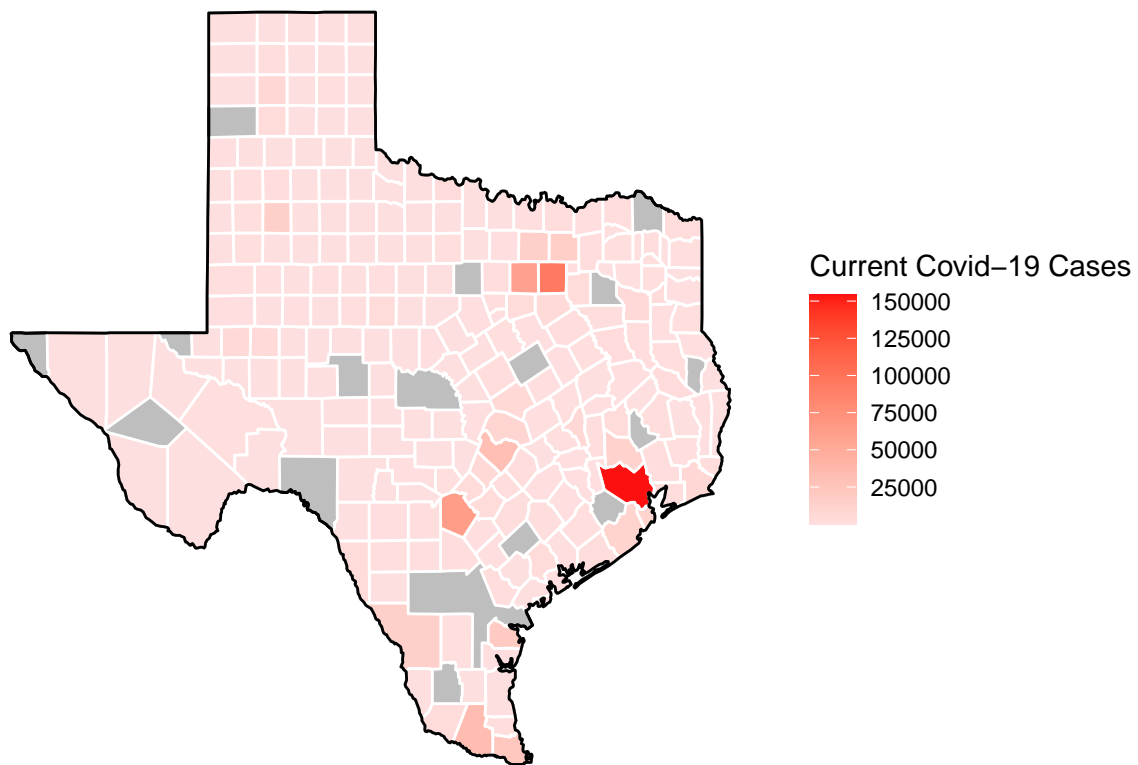
Table of Current Covid-19 Cases and Total Deaths

```
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	2909	38
Andrews	48003	547	10
Angelina	48005	2753	88
Aransas	48007	341	18
Archer	48009	114	1
Armstrong	48011	18	1
Atascosa	48013	1254	35
Austin	48015	507	9
Bailey	48017	264	5
Bandera	48019	179	8
Bastrop	48021	1945	31
Baylor	48023	42	2
Bee	48025	1840	37
Bell	48027	6064	96
Bexar	48029	63340	1381
Blanco	48031	153	6
Borden	48033	2	0
Bosque	48035	385	10
Bowie	48037	1668	74
Brazoria	48039	11954	179
Brazos	48041	8167	65
Brewster	48043	246	5
Briscoe	48045	18	1
Brooks	48047	350	23
Brown	48049	932	33
Burleson	48051	420	6
Burnet	48053	854	12
Caldwell	48055	1681	37
Calhoun	48057	879	9
Callahan	48059	110	4
Cameron	48061	23709	1066
Camp	48063	383	13
Carson	48065	49	0
Cass	48067	461	21
Castro	48069	287	6
Chambers	48071	1497	19
Cherokee	48073	1611	31
Childress	48075	91	0
Clay	48077	97	2
Cochran	48079	81	1
Coke	48081	84	4
Coleman	48083	84	4
Collin	48085	16937	171
Collingsworth	48087	19	0
Colorado	48089	498	4
Comal	48091	3642	120
Comanche	48093	343	5
Concho	48095	140	2
Cooke	48097	505	7
Coryell	48099	1941	13

county_tx	fips	current_cases	total_deaths
San Patricio	48409	1549	79
San Saba	48411	110	0
Schleicher	48413	55	2
Scurry	48415	957	11
Shackelford	48417	34	0
Shelby	48419	538	30
Sherman	48421	72	1
Smith	48423	5765	124
Somervell	48425	154	5
Starr	48427	3870	179
Stephens	48429	151	4
Sterling	48431	11	0
Stonewall	48433	18	1
Sutton	48435	95	3
Swisher	48437	131	3
Tarrant	48439	59274	783
Taylor	48441	3814	62
Terrell	48443	4	0
Terry	48445	440	6
Throckmorton	48447	16	2
Titus	48449	1527	35
Tom Green	48451	4290	74
Travis	48453	30956	445
Trinity	48455	200	7
Tyler	48457	252	4
Upshur	48459	456	12
Upton	48461	40	3
Uvalde	48463	819	36
Val Verde	48465	2384	124
Van Zandt	48467	829	28
Victoria	48469	4193	97
Walker	48471	3855	61
Waller	48473	948	15
Ward	48475	156	2
Washington	48477	692	48
Webb	48479	14669	337
Wharton	48481	1415	58
Wheeler	48483	59	0
Wichita	48485	2861	26
Wilbarger	48487	281	3
Willacy	48489	1224	55
Williamson	48491	9327	150
Wilson	48493	977	28
Winkler	48495	130	4
Wise	48497	848	15
Wood	48499	590	34
Yoakum	48501	284	5
Young	48503	470	6
Zapata	48505	342	6
Zavala	48507	488	18