Time series analysis with updated window definition

Cheyenne Ehman, Ziyan Zhu

3/29/2021

Select varible of interests

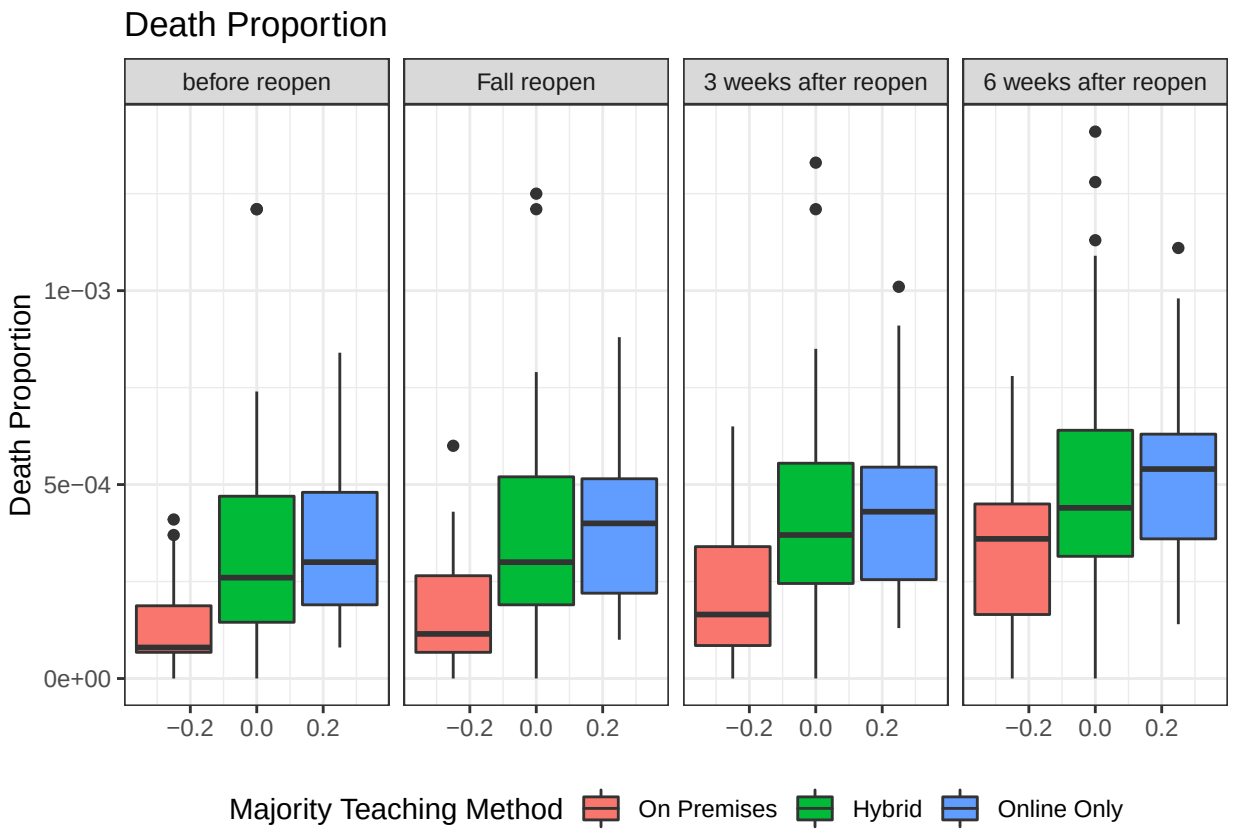
```
## If you don't have the covidcast package, run following line

#devtools::install_github("cmu-delphi/covidcast", ref = "main", subdir = "R-packages/covidcast",depende
```

```
source("step2_data_wrangle.R")
################### school reopen dates ###################
district_policies <- OH_K12 %>%
  distinct(county,county_enroll,leaid,district_enroll,schooltemporaryshutdown,opendategrouped,teachingm
# Calculate the proportion and generate date brackets
prop_opendate <- district_policies%>%
  filter(!schooltemporaryshutdown %in% c('Closed indefinitely','Pending','Unknown'))%>%
  group_by(county,county_enroll,opendategrouped)%>%
  summarise(n_opendate = sum(district_enroll))%>% # number of students under certain date for each coun
  mutate(prop_opendate = round(n_opendate/county_enroll,2))%>% # proportion
  group_by(county)%>%
  #filter(prop_opendate>0.6)%>%
  slice(which.max(prop_opendate))%>% # filter large proportions of students with same reopen dates #can
  mutate(threeweeks_lag_open = opendategrouped+21,sixweeks_lag_open = opendategrouped+42,twomonths_lag_
  select(-n_opendate)
```

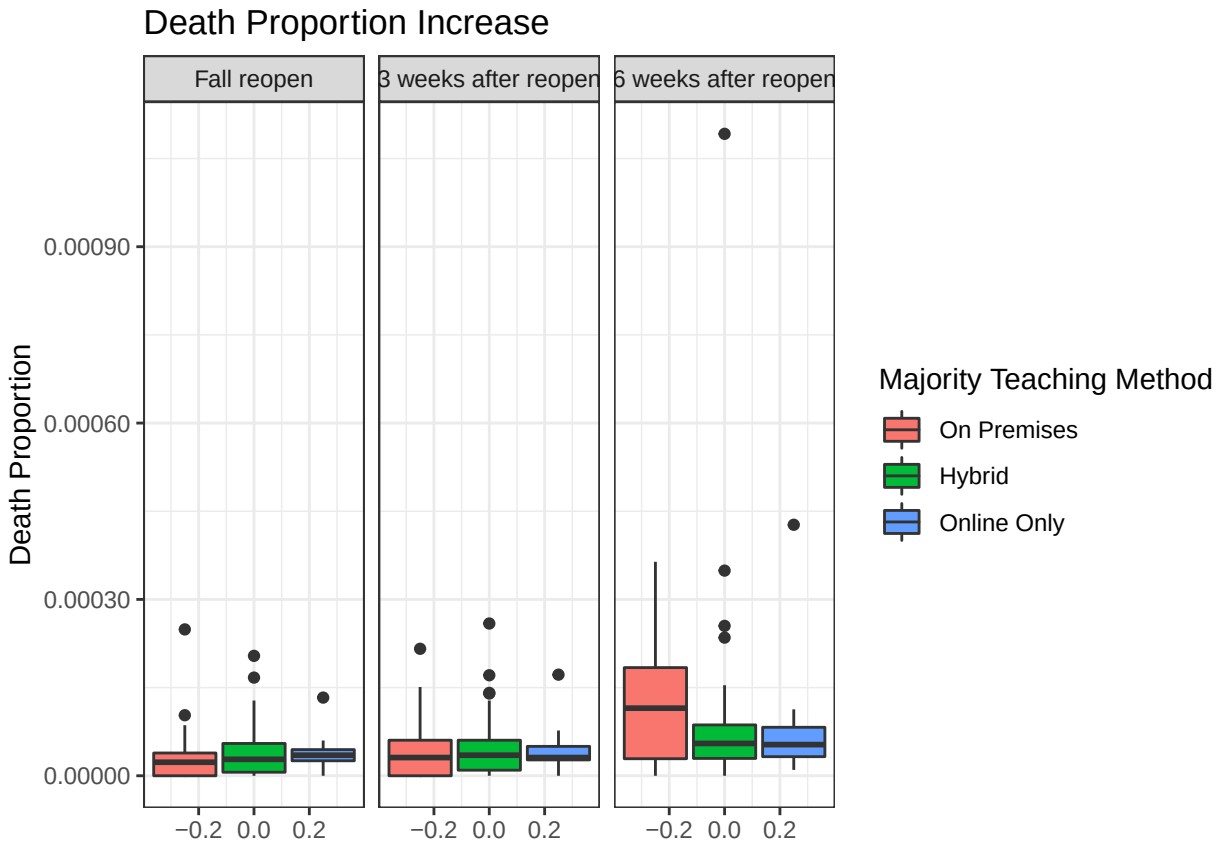
```
opendate_cases <- case_mobility%>%
  inner_join(prop_opendate,by=c('COUNTY'='county'))%>%
  group_by(COUNTY)%>%
  filter(DATE>=opendategrouped & DATE<=beforechristmas)%>%
  mutate(window_id = case_when(DATE<threeweeks_lag_open~"fall_reopento21d",
    DATE>=threeweeks_lag_open & DATE<sixweeks_lag_open~'reopen_21dto42d',
    DATE>= sixweeks_lag_open & DATE<twomonths_lag_open ~ 'reopen_42dto63d',
    TRUE ~ 'before_christmas'
  ))%>%
  mutate(death_prop = round(CUMDEATHS/POPULATION,5),
         window_id = as.factor(window_id))%>%
  left_join(wide_teaching_enroll,by=c('COUNTY'='county','county_enroll'))
# select the start date and end date data for each window of time

start_of_window <- opendate_cases%>%
  group_by(COUNTY,window_id)%>%
  arrange(DATE)%>%
  filter(row_number()==1)%>%
  ungroup()%>%
  mutate(y_label = case_when(window_id == "fall_reopento21d"~'before reopen',
```

```
    window_id == "reopen_21dto42d" ~ 'Fall reopen',
    window_id == "reopen_42dto63d" ~ '3 weeks after reopen',
    window_id == "before_christmas" ~ '6 weeks after reopen'
  ))%>%
  select(-opendategrouped,-threeweeks_lag_open,-sixweeks_lag_open,-twomonths_lag_open,-beforechristmas,
```

```
start_of_window$y_label <- factor(start_of_window$y_label,
         levels = c("before reopen",
         "Fall reopen",
         "3 weeks after reopen",
         "6 weeks after reopen"))

start_of_window$major_teaching <- factor(start_of_window$major_teaching,
         levels = c("On Premises",
         "Hybrid",
         "Online Only"))

start_of_window %>%
  ggplot(aes(y = death_prop,
             fill = major_teaching))+
  geom_boxplot(na.rm = T) +
  facet_grid(~y_label)+
  theme_bw()+
  labs(y = "Death Proportion",
       fill = "Majority Teaching Method",
       title = "Death Proportion")+theme(legend.position = "bottom")
```



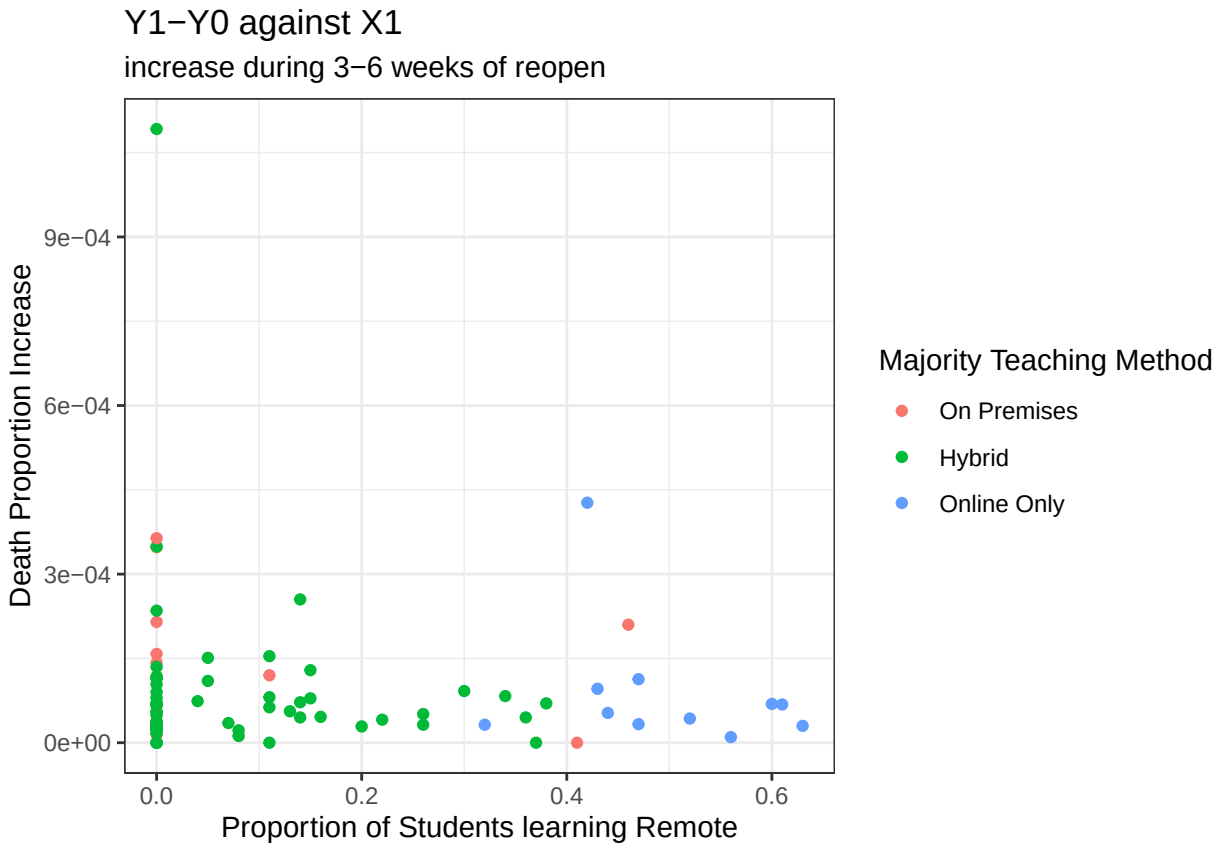
```
#ggsave("Deathprop_box.jpg", width = 8.5, height = 5)

require(scales)
start_of_window %>%
  group_by(COUNTY)%>%
  mutate(death_inc = CUMDEATHS-lag(CUMDEATHS))%>%
  drop_na(death_inc)%>%
  mutate(death_prop_inc = round(death_inc/POPULATION,6))%>%
  ggplot(aes(y = death_prop_inc,
             fill = major_teaching))+
  geom_boxplot(na.rm = T) +
  theme_bw()+
  labs(y = "Death Proportion",
       fill = "Majority Teaching Method",
       title = "Death Proportion Increase") +
  facet_grid(~y_label)+
  scale_y_continuous(labels = comma)
```




```
#ggsave("Deathpropinc_box.jpg", width = 8.5, height = 5)
```

```
start_of_window %>%
  group_by(COUNTY)%>%
  mutate(death_inc = CUMDEATHS-lag(CUMDEATHS))%>%
  drop_na(death_inc)%>%
  mutate(death_prop_inc = round(death_inc/POPULATION,6))%>%
  filter(y_label == "6 weeks after reopen") %>%
  ggplot(aes(x = Online_Only, y = death_prop_inc, color = major_teaching)) +
```

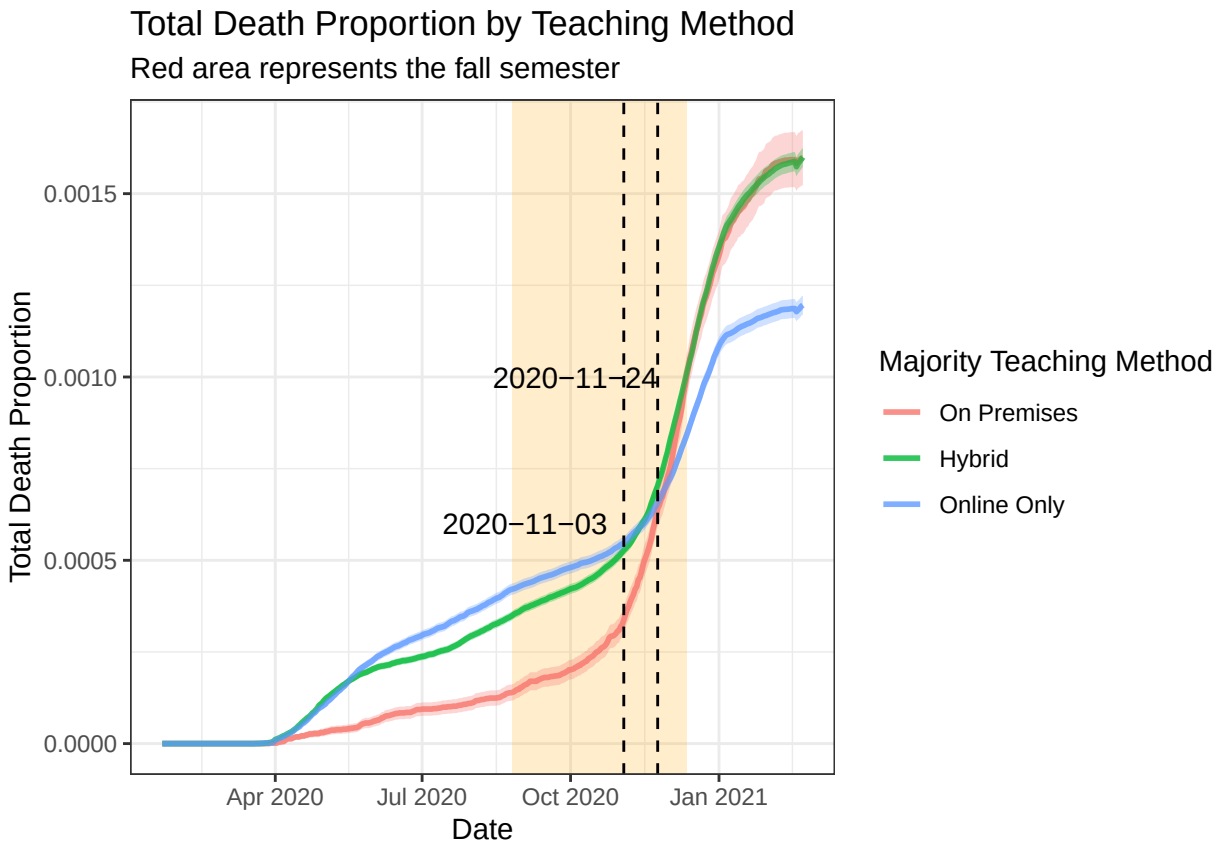
```
  geom_point()+
  theme_bw() +
  labs(x = "Proportion of Students learning Remote",
       y = "Death Proportion Increase",
       title = "Y1-Y0 against X1",
       subtitle = "increase during 3-6 weeks of reopen",
       color = "Majority Teaching Method")
```




```
ggsave("y1x1.jpg", width = 7, height = 5)
```

```
county_policy_wide$major_teaching <- factor(county_policy_wide$major_teaching,
         levels = c("On Premises",
         "Hybrid",
         "Online Only"))
# see when the intesection happens
date.intercept <- as.Date("2020-11-24")

# add 95% confidence bans
confidence_level <- .95
z_cl <- qnorm(confidence_level)

# case_policy_wide
case_policy_wide <- case_mobility %>%
  left_join(county_policy_wide[,c("county","major_teaching","Online_Only","Hybrid","On_Premises")],by =
  mutate(death_prop = CUMDEATHS/POPULATION)

# plot
```

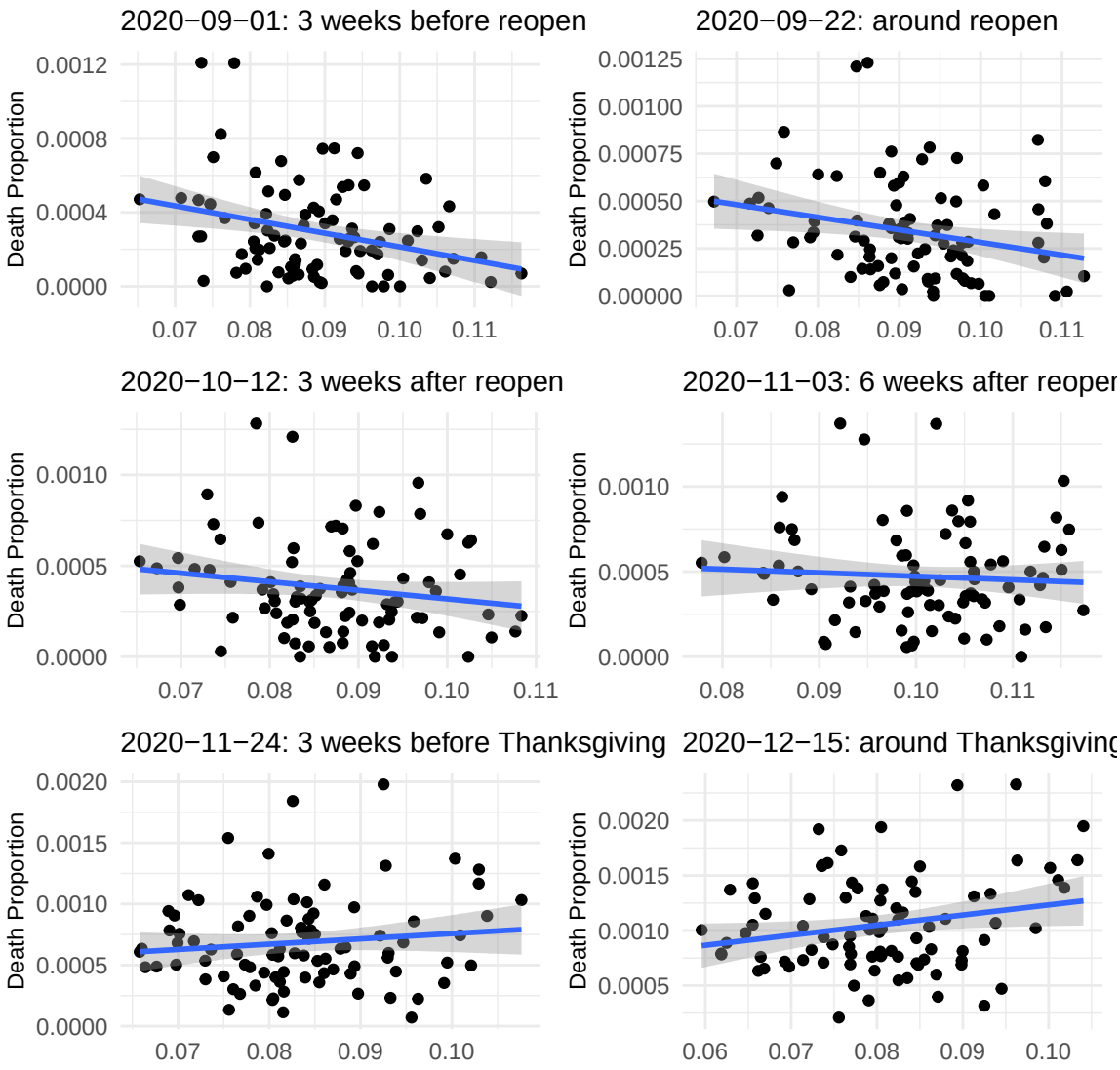
```r
case_policy_wide%>%
  group_by(DATE, major_teaching) %>%
  drop_na(major_teaching)%>%
  summarise(total_deaths = sum(CUMDEATHS),
            total_pop = sum(POPULATION),
            death_prop = total_deaths/total_pop,
            death_prop_upper = death_prop + z_cl*sqrt(death_prop*(1 - death_prop)/total_pop),
            death_prop_lower = death_prop - z_cl*sqrt(death_prop*(1 - death_prop)/total_pop),
            .groups = "drop") %>%
  ggplot(aes(x = DATE, y = death_prop, group = major_teaching))+
    geom_rect(data=opendate_cases[1,],
            aes(xmin=as.Date("2020/08/26"), xmax=as.Date("2020/12/12"),
                ymin=-Inf,ymax=Inf),
            color = NA,alpha=0.2, show.legend = F, fill = "orange") +
    geom_line(aes(color = major_teaching),size = 1, alpha = .8) +
    geom_ribbon(aes(ymin = death_prop_lower, ymax = death_prop_upper,
                    fill= major_teaching),
                alpha = .3, show.legend = F)+
    geom_vline(xintercept = date.intercept, linetype = "dashed") +
    annotate("text",x = date.intercept,y = .001,
             label = date.intercept,
             hjust = 1.0) +
    geom_vline(xintercept = as.Date('2020/11/03'), linetype = "dashed") +
    annotate("text",x = as.Date('2020/11/03'),y = .0006,
              label =as.Date('2020/11/03'),
              hjust = 1.1) +
    theme_bw() +
    labs(x = "Date", y = "Total Death Proportion",
         title = "Total Death Proportion by Teaching Method",
         subtitle = "Red area represents the fall semester",
         color = "Majority Teaching Method")
```



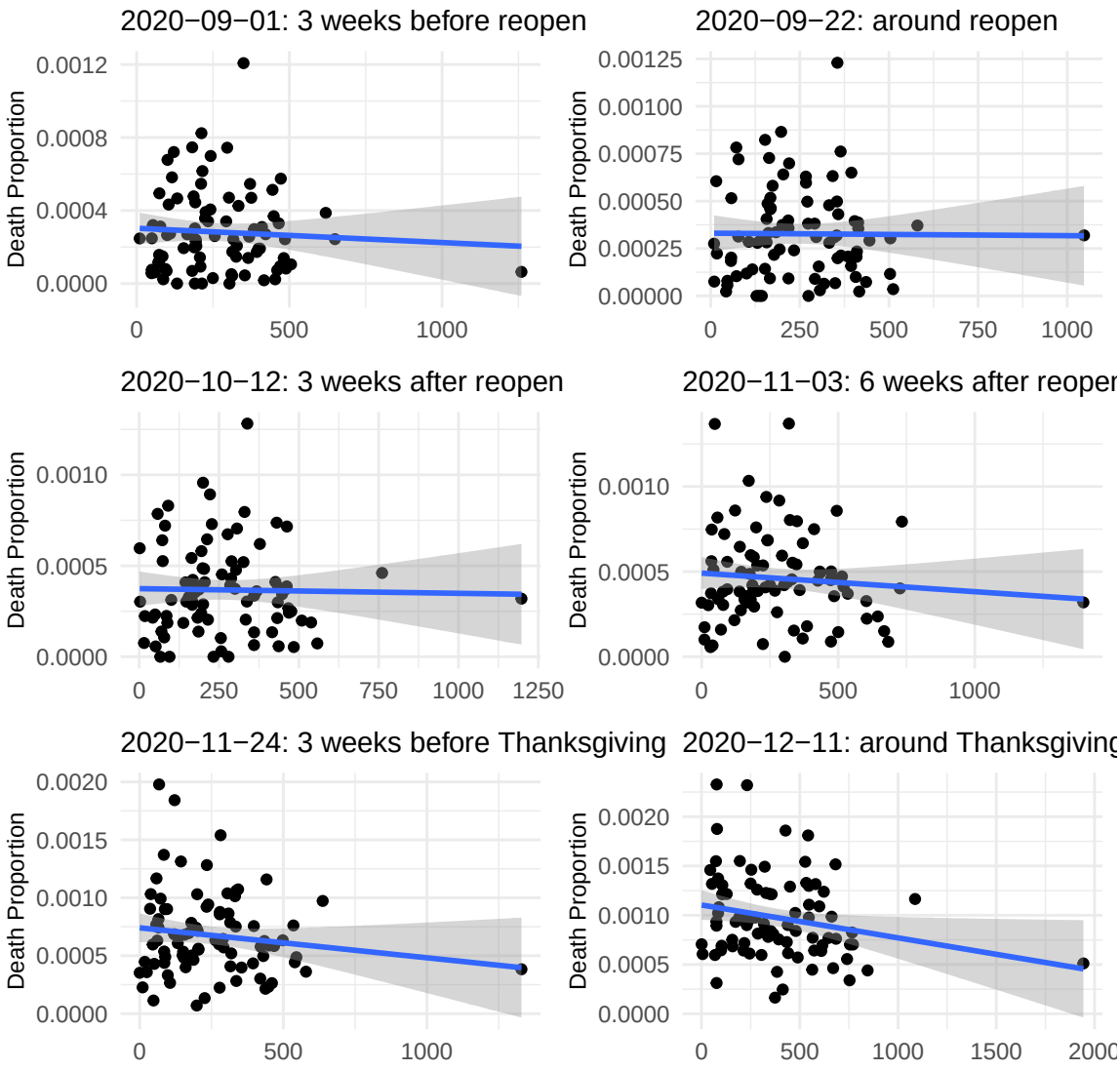

```
# +
#   scale_x_date(date_breaks = "2 month", date_labels = "%b-%y")
```

```
library(gridExtra)

date_plot <- function(date,var,title_text){
  p <- case_mobility%>%
  filter(DATE == date)%>%
  mutate(death_prop = CUMDEATHS/POPULATION)%>%
  ggplot(aes(x= get(var),y = death_prop))+theme_minimal()+
  labs(x=NULL,y="Death Proportion",title=paste0(date,": ",title_text))+geom_point()+geom_smooth(method=
return(p)
}

p11 <- date_plot(date = '2020-09-01',var = "work_prop_7d",title_text = "3 weeks before reopen")

p12 <- date_plot(date = '2020-09-01',var = "res_visit_prop",title_text = "3 weeks before reopen")

p13 <- date_plot(date = '2020-09-01',var = "bar_visit_prop",title_text = "3 weeks before reopen")

#############################################################################
p21 <- date_plot(date = '2020-09-22',var = "work_prop_7d",title_text = "around reopen")

p22 <- date_plot(date = '2020-09-22',var = "res_visit_prop",title_text = "around reopen")

p23 <- date_plot(date = '2020-09-22',var = "bar_visit_prop",title_text = "around reopen")
```

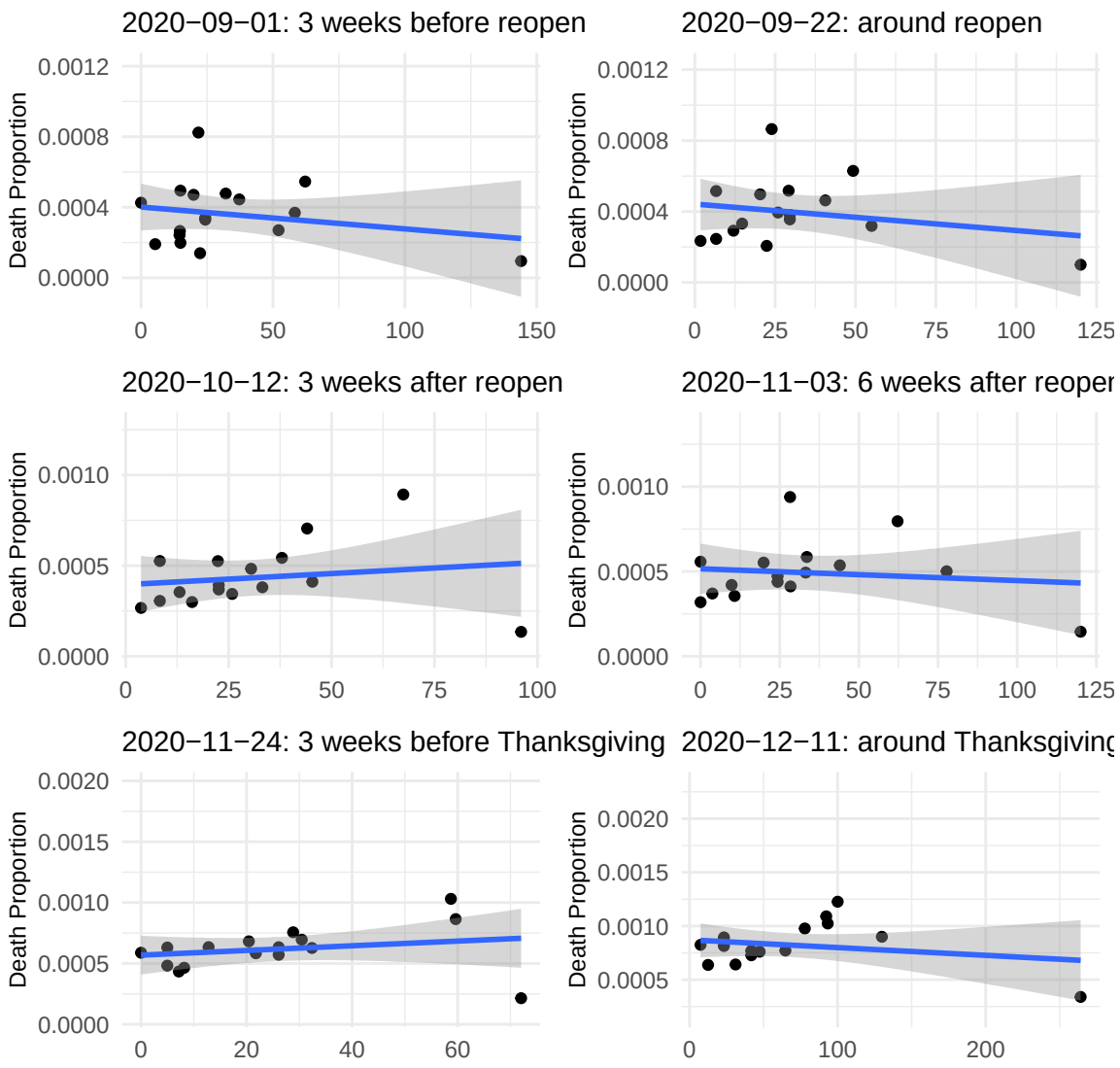
```
##############################################################################
p31 <- date_plot(date = '2020-10-12',var = "work_prop_7d",title_text = "3 weeks after reopen")

p32 <- date_plot(date = '2020-10-12',var = "res_visit_prop",title_text = "3 weeks after reopen")

p33 <- date_plot(date = '2020-10-12',var = "bar_visit_prop",title_text = "3 weeks after reopen")

##############################################################################
p41 <- date_plot(date = '2020-11-03',var = "work_prop_7d",title_text = "6 weeks after reopen")

p42 <- date_plot(date = '2020-11-03',var = "res_visit_prop",title_text = "6 weeks after reopen")

p43 <- date_plot(date = '2020-11-03',var = "bar_visit_prop",title_text = "6 weeks after reopen")

##############################################################################
p51 <- date_plot(date = '2020-11-24',var = "work_prop_7d",title_text = "3 weeks before Thanksgiving")

p52 <- date_plot(date = '2020-11-24',var = "res_visit_prop",title_text = "3 weeks before Thanksgiving")

p53 <- date_plot(date = '2020-11-24',var = "bar_visit_prop",title_text = "3 weeks before Thanksgiving")
##############################################################################
p61 <- date_plot(date = '2020-12-15',var = "work_prop_7d",title_text = "around Thanksgiving")

p62 <- date_plot(date = '2020-12-11',var = "res_visit_prop",title_text = "around Thanksgiving")

p63 <- date_plot(date = '2020-12-11',var = "bar_visit_prop",title_text = "around Thanksgiving")

##############################################################################
p71 <- date_plot(date = '2021-02-22',var = "work_prop_7d",title_text = "")

p72 <- date_plot(date = '2021-02-22',var = "res_visit_prop",title_text = "")

p73 <- date_plot(date = '2021-02-22',var = "bar_visit_prop",title_text = "")

##############################################################################
options(scipen=10000)
grid.arrange(p11, p21, p31,p41,p51,p61,nrow = 3,ncol=2,widths = c(5,5),heights = c(3,3,3),bottom= "Frac
```




```
grid.arrange(p12, p22, p32,p42,p52,p62,nrow = 3,ncol=2,widths = c(5,5),heights = c(3,3,3),bottom="Resta
```




```
grid.arrange(p13,p23,p33,p43,p53,p63,nrow = 3,ncol=2,widths = c(5,5),heights = c(3,3,3),bottom="Bar visi
```



Bar visits by population

```
avg_work <- case_policy_wide %>%
  drop_na(major_teaching)%>%
  group_by(DATE, major_teaching) %>%
  summarise(avg_ptwork_prop_7d = mean(work_prop_7d))
```

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

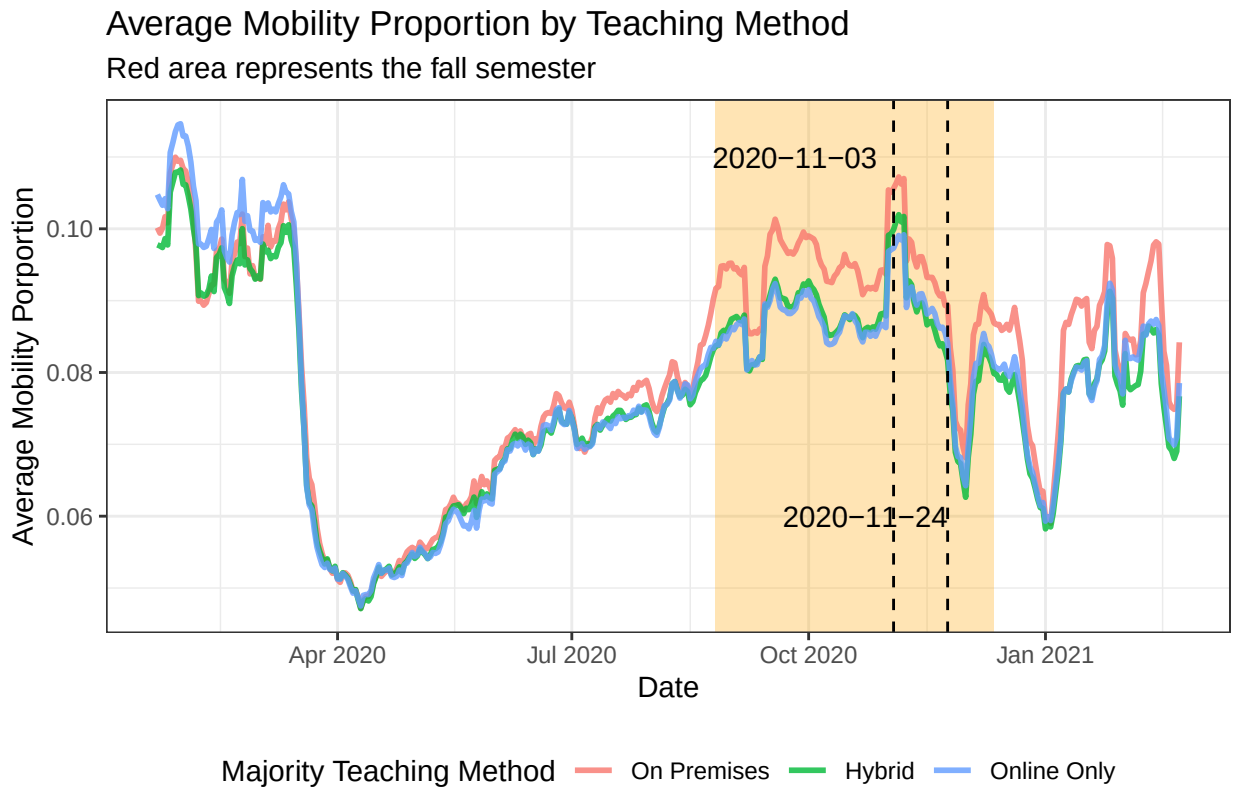
```
avg_work%>%
  ggplot(aes(x = DATE, y = avg_ptwork_prop_7d, color = major_teaching))+
    geom_rect(data = avg_work[1,],
              aes(xmin=as.Date("2020/08/26"), xmax=as.Date("2020/12/12")),
                  ymin=-Inf,ymax=Inf,color = NA,alpha=0.3,show.legend = F, fill = "orange") +
    geom_line(aes(color = major_teaching),size = 1, alpha = .8) +
        geom_vline(xintercept = date.intercept, linetype = "dashed") +
    annotate("text",x = date.intercept,y = .06,
             label = date.intercept,
             hjust = 1.0) +
    geom_vline(xintercept = as.Date('2020/11/03'), linetype = "dashed") +
```

```r
  annotate("text",x = as.Date('2020/11/03'),y = .11,
              label =as.Date('2020/11/03'),
              hjust = 1.1) +
  theme_bw() +
  labs(x = "Date", y = "Average Mobility Porportion",
       title = "Average Mobility Proportion by Teaching Method",
       subtitle = "Red area represents the fall semester",
       color = "Majority Teaching Method",
       caption = "The fraction of devices that spent between 3 and 6 hours at a location other than th
theme(legend.position = "bottom")
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph’s)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
```

```
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <99>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
```

```
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'The fraction of devices that spent between 3 and 6
## hours at a location other than their home during the daytime (SafeGraph's)' in
## 'mbcsToSbcs': dot substituted for <99>
```



n of devices that spent between 3 and 6 hours at a location other than their home during the daytime (SafeGraph...s)

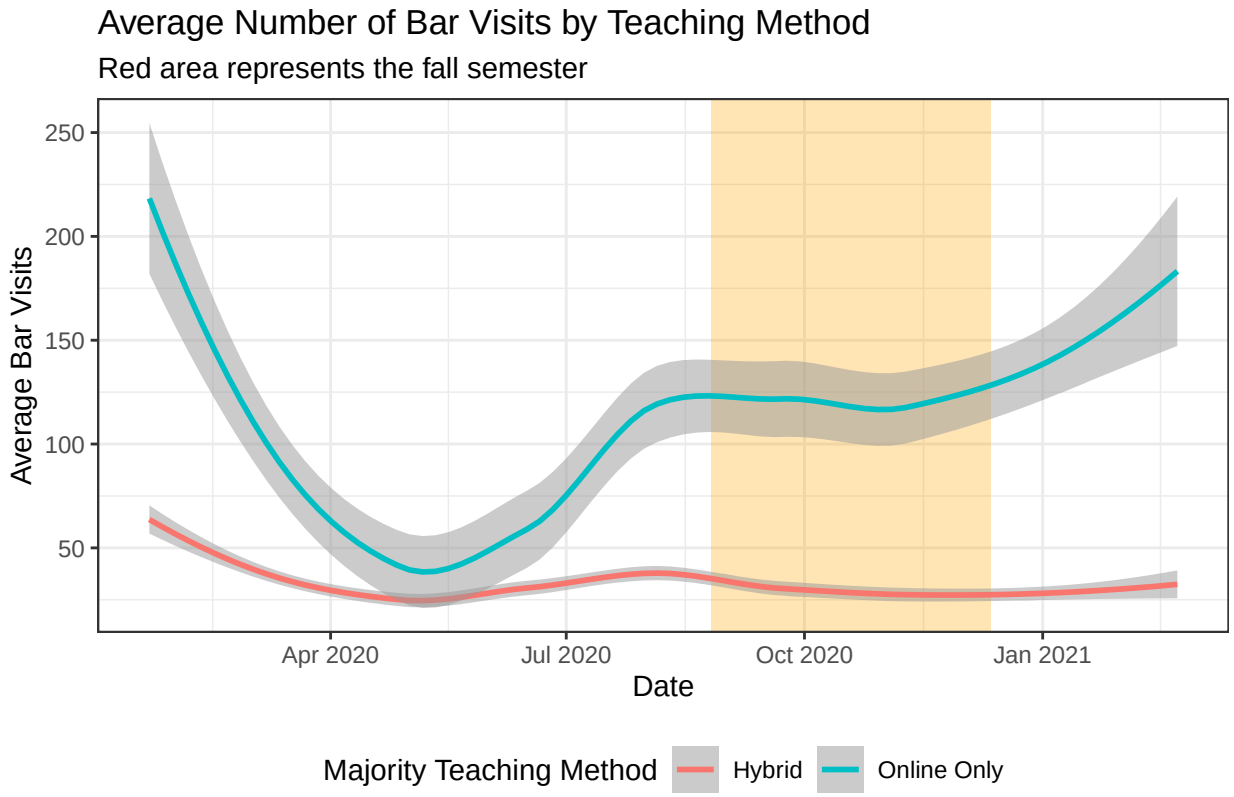
```
avg_bar_visit <- case_policy_wide %>%
  drop_na(bar_visit_prop)%>%
  group_by(DATE, major_teaching) %>%
  summarise(avg_bar_visit = mean(bar_visit_prop))
```

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

```
avg_bar_visit%>%
  ggplot(aes(x = DATE, y = avg_bar_visit, color = major_teaching))+
    geom_rect(data = avg_bar_visit[1,],
              aes(xmin=as.Date("2020/08/26"), xmax=as.Date("2020/12/12")),
                  ymin=-Inf,ymax=Inf,color = NA,alpha=0.3, show.legend = F, fill = "orange") +
    #geom_point(size = 2, alpha = .8)+
    geom_smooth(size = 1, alpha = .5) +
    theme_bw() +
    labs(x = "Date", y = "Average Bar Visits",
```

```
        title = "Average Number of Bar Visits by Teaching Method",
        subtitle = "Red area represents the fall semester",
        color = "Majority Teaching Method",
        caption = "Weekly counts of visits and normalized by population size (SafeGraph)")+
  theme(legend.position = "bottom")
```

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



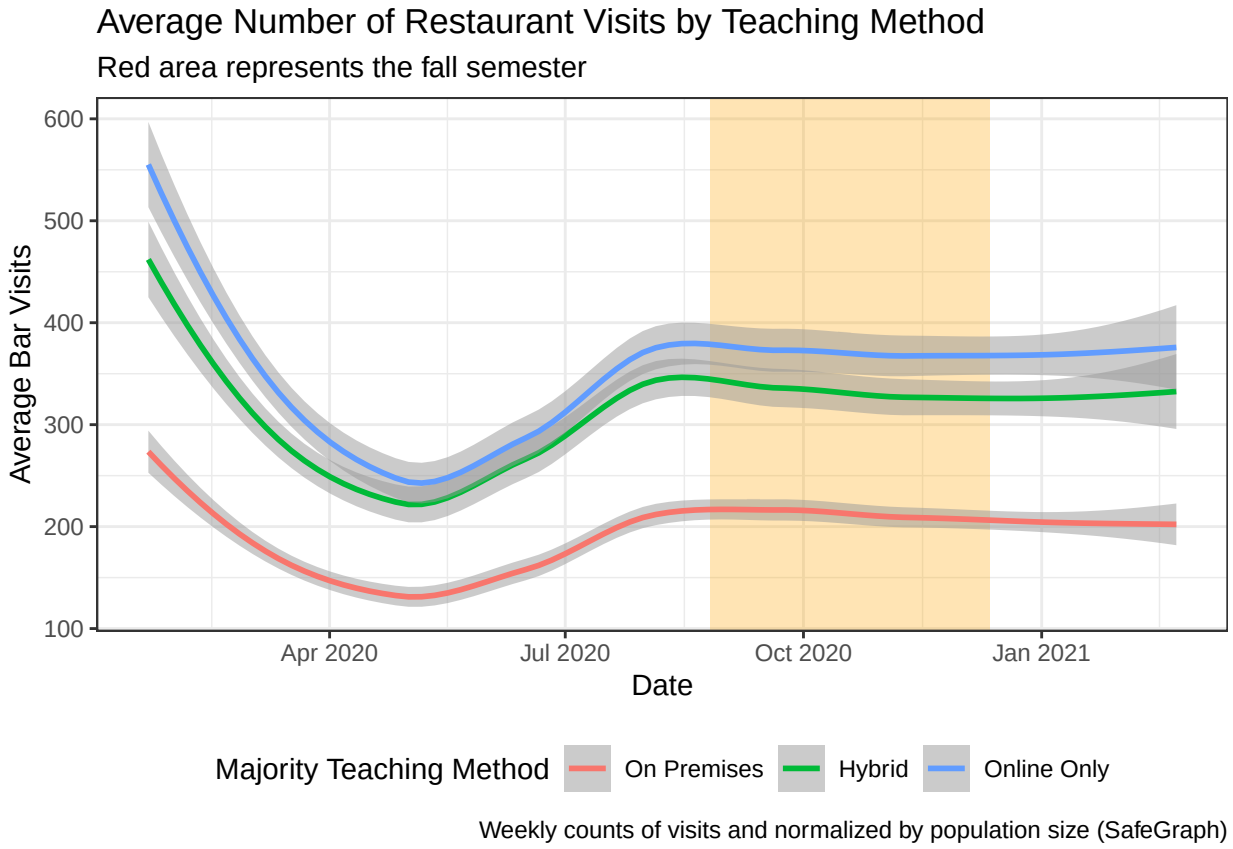

```
avg_res_visit <- case_policy_wide %>%
  drop_na(res_visit_prop)%>%
  group_by(DATE, major_teaching) %>%
  summarise(avg_res_visit = mean(res_visit_prop))
```

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

```
avg_res_visit%>%
  drop_na(major_teaching)%>%
  ggplot(aes(x = DATE, y = avg_res_visit, color = major_teaching))+
    geom_rect(data = avg_res_visit[1,],
            aes(xmin=as.Date("2020/08/26"), xmax=as.Date("2020/12/12")),
                ymin=-Inf,ymax=Inf,color = NA,alpha=0.3, show.legend = F, fill = "orange") +
    #geom_point(size = 1, alpha = .8)+
    geom_smooth(size = 1, alpha = .5) +
    theme_bw() +
    labs(x = "Date", y = "Average Bar Visits",
         title = "Average Number of Restaurant Visits by Teaching Method",
         subtitle = "Red area represents the fall semester",
         color = "Majority Teaching Method",
```

```
        caption = "Weekly counts of visits and normalized by population size (SafeGraph)")+
  theme(legend.position = "bottom")
```

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



Notice

After removing missing values in each of the mobility measures, we end up with different sample size for each linear regression

```
summary(lm(death_prop~work_prop_7d,data = case_policy_wide,na.action='na.omit'))
```

```
##
## Call:
## lm(formula = death_prop ~ work_prop_7d, data = case_policy_wide,
##     na.action = "na.omit")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0005834 -0.0003986 -0.0002491  0.0002207  0.0024472
##
## Coefficients:
##                Estimate Std. Error t value            Pr(>|t|)
## (Intercept)  0.00027520 0.00001481   18.58 <0.0000000000000002 ***
## work_prop_7d 0.00211781 0.00018531   11.43 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0005767 on 35021 degrees of freedom
## Multiple R-squared:  0.003716,   Adjusted R-squared:  0.003687
## F-statistic: 130.6 on 1 and 35021 DF,  p-value: < 0.00000000000000022
```

```
summary(lm(death_prop~res_visit_prop,data = case_policy_wide,na.action='na.omit'))
```

```
##
## Call:
## lm(formula = death_prop ~ res_visit_prop, data = case_policy_wide,
##     na.action = "na.omit")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0005306 -0.0004127 -0.0002491  0.0001850  0.0024543
##
## Coefficients:
##                      Estimate    Std. Error t value             Pr(>|t|)
## (Intercept)    0.00041176670 0.00000464860  88.579 < 0.0000000000000002 ***
## res_visit_prop 0.00000003558 0.00000001156   3.077              0.00209 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005665 on 33324 degrees of freedom
##   (1697 observations deleted due to missingness)
## Multiple R-squared:  0.000284,   Adjusted R-squared:  0.000254
## F-statistic: 9.468 on 1 and 33324 DF,  p-value: 0.002093
```

```
summary(lm(death_prop~bar_visit_prop,data = case_policy_wide,na.action='na.omit'))
```

```
##
## Call:
## lm(formula = death_prop ~ bar_visit_prop, data = case_policy_wide,
##     na.action = "na.omit")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0003484 -0.0003422 -0.0001275  0.0001626  0.0013851
##
## Coefficients:
##                       Estimate     Std. Error t value            Pr(>|t|)
## (Intercept)     0.00034837898  0.00000520124  66.980 <0.0000000000000002 ***
## bar_visit_prop -0.00000001162  0.00000003757  -0.309               0.757
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000403 on 7221 degrees of freedom
##   (27800 observations deleted due to missingness)
## Multiple R-squared:  1.324e-05,  Adjusted R-squared:  -0.0001252
## F-statistic: 0.09559 on 1 and 7221 DF,  p-value: 0.7572
```

NOTES:

Though we are getting all significance at the end, I am suspective about the results. Please double check the code and input data, make sure it is not random.

```r
summary(lm(death_prop~work_prop_7d+res_visit_prop,data = case_policy_wide,na.action='na.omit'))
```

```
##
## Call:
## lm(formula = death_prop ~ work_prop_7d + res_visit_prop, data = case_policy_wide,
##     na.action = "na.omit")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0005708 -0.0003810 -0.0002459  0.0001793  0.0024678
##
## Coefficients:
##                     Estimate    Std. Error t value            Pr(>|t|)
## (Intercept)    0.00025842346 0.00001481203  17.447 <0.0000000000000002 ***
## work_prop_7d   0.00204538960 0.00018762626  10.901 <0.0000000000000002 ***
## res_visit_prop 0.00000001370 0.00000001172   1.169               0.242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005655 on 33323 degrees of freedom
##   (1697 observations deleted due to missingness)
## Multiple R-squared:  0.003837,   Adjusted R-squared:  0.003777
## F-statistic: 64.17 on 2 and 33323 DF,  p-value: < 0.00000000000000022
```

```r
summary(lm(death_prop~Hybrid+On_Premises+Online_Only,data = case_policy_wide))
```

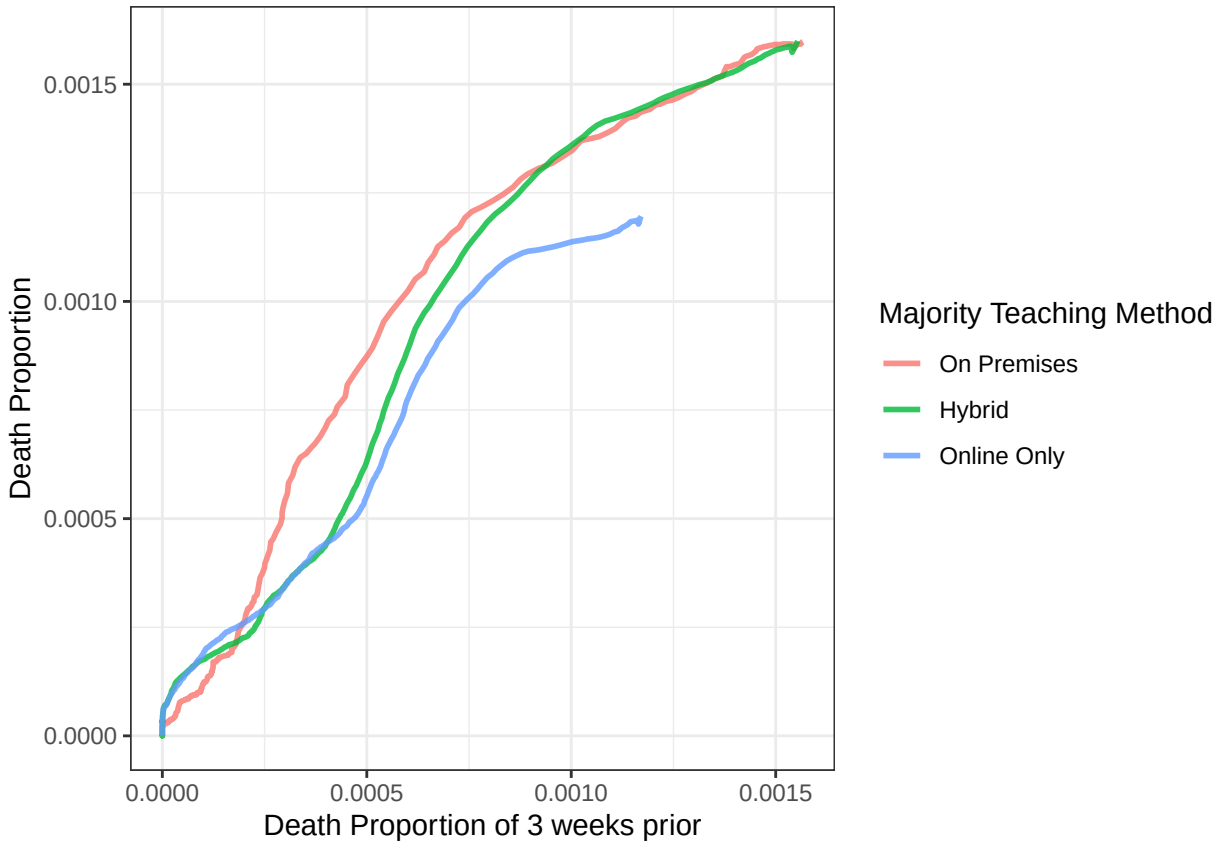
```
##
## Call:
## lm(formula = death_prop ~ Hybrid + On_Premises + Online_Only,
##     data = case_policy_wide)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0006384 -0.0004175 -0.0002487  0.0002245  0.0024486
##
## Coefficients:
##                Estimate  Std. Error t value            Pr(>|t|)
## (Intercept)  0.00063839  0.00001584  40.300 <0.0000000000000002 ***
## Hybrid      -0.00022088  0.00001812 -12.192 <0.0000000000000002 ***
## On_Premises -0.00018479  0.00001916  -9.643 <0.0000000000000002 ***
## Online_Only -0.00027022  0.00002387 -11.320 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005783 on 34223 degrees of freedom
##   (796 observations deleted due to missingness)
## Multiple R-squared:  0.004825,   Adjusted R-squared:  0.004738
## F-statistic: 55.31 on 3 and 34223 DF,  p-value: < 0.00000000000000022
```

```r
summary(lm(death_prop~work_prop_7d+Hybrid+On_Premises+Online_Only,data = case_policy_wide))
```

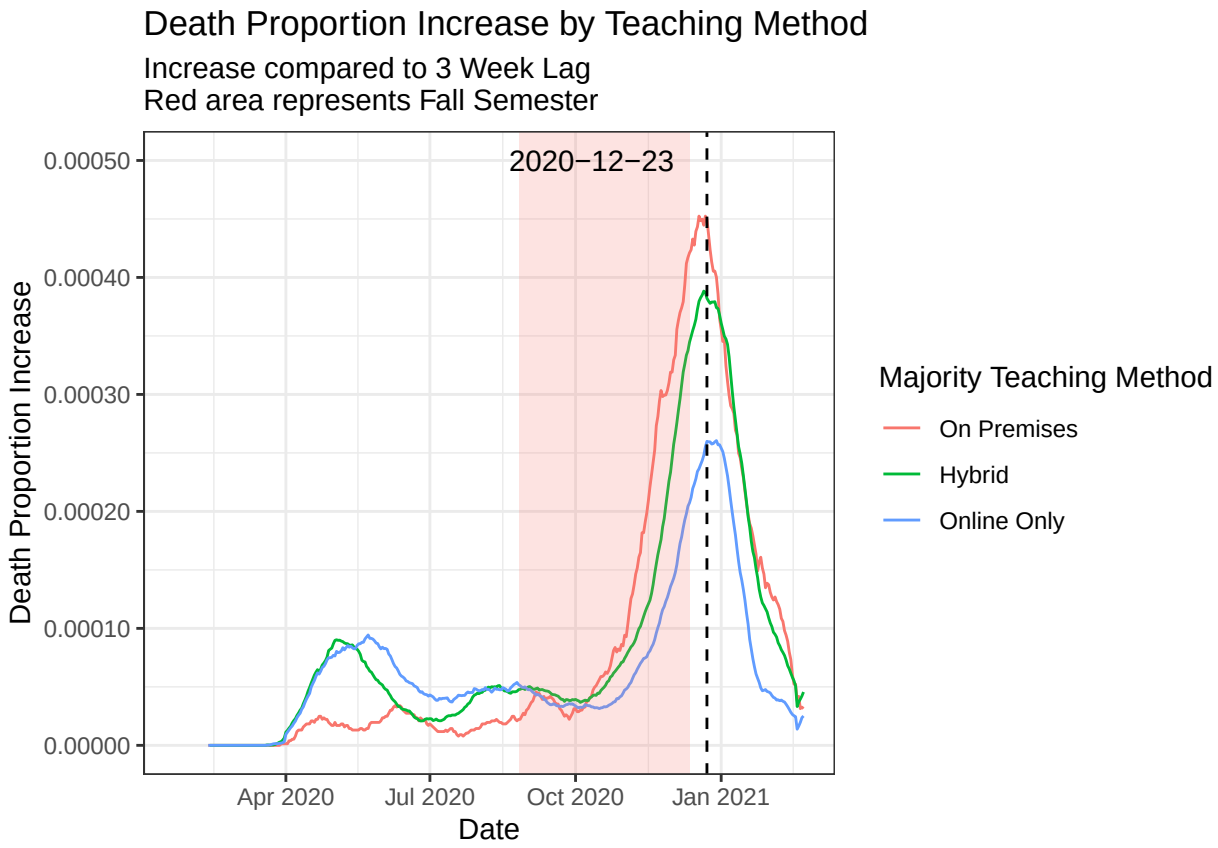
```
##
## Call:
## lm(formula = death_prop ~ work_prop_7d + Hybrid + On_Premises +
##     Online_Only, data = case_policy_wide)
```

```
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0006691 -0.0003983 -0.0002533  0.0002164  0.0024690
##
## Coefficients:
##                 Estimate  Std. Error t value            Pr(>|t|)
## (Intercept)   0.00047488  0.00002109   22.52 <0.0000000000000002 ***
## work_prop_7d  0.00222574  0.00018994   11.72 <0.0000000000000002 ***
## Hybrid       -0.00022885  0.00001809  -12.65 <0.0000000000000002 ***
## On_Premises  -0.00021334  0.00001928  -11.06 <0.0000000000000002 ***
## Online_Only  -0.00027112  0.00002382  -11.38 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005771 on 34222 degrees of freedom
##   (796 observations deleted due to missingness)
## Multiple R-squared:  0.008802,   Adjusted R-squared:  0.008686
## F-statistic: 75.97 on 4 and 34222 DF,  p-value: < 0.00000000000000022
```

```r
lag_cases <- case_mobility %>%
  left_join(county_policy_wide[,c("county","major_teaching")],
            by = c("COUNTY" = "county")) %>%
  drop_na(major_teaching)%>%
  select(COUNTY,DATE,CUMDEATHS,POPULATION,major_teaching)%>%
  group_by(COUNTY) %>%
  mutate(lag_total_deaths = lag(CUMDEATHS,21)) %>%
  ungroup()%>%
  group_by(DATE,major_teaching) %>%
  summarise(total_deaths = sum(CUMDEATHS),
            total_deaths_lag = sum(lag_total_deaths),
            total_pop = sum(POPULATION),
            death_prop = total_deaths/total_pop,
            lag_death_prop = total_deaths_lag/total_pop,
            death_prop_inc = (total_deaths-total_deaths_lag)/total_pop,
            .groups = "drop")

ggplot(lag_cases, aes(x = lag_death_prop, y = death_prop, color = major_teaching)) +
    geom_line(size = 1,alpha = .8, na.rm=T)+
  theme_bw() +
  labs(x = "Death Proportion of 3 weeks prior", y  = "Death Proportion",
       color = "Majority Teaching Method")
```




```
peak.date <- as.Date("2020-12-23")
ggplot(lag_cases,aes(x = DATE, y = death_prop_inc,
                     color = major_teaching,
                     fill = "red")) +
    geom_line(na.rm = T) +
    geom_rect(data = lag_cases[1,],
              aes(xmin=as.Date("2020/08/26"), xmax=as.Date("2020/12/12"),
                  ymin=-Inf,ymax=Inf),
              color = NA,alpha=0.2, show.legend = F) +
    geom_vline(xintercept = peak.date, linetype = "dashed")+
    annotate("text",x = peak.date,y = .0005,
             label = peak.date,
             hjust = 1.2)  +
  theme_bw() +
  labs(x = "Date",
       y = "Death Proportion Increase",
       title = "Death Proportion Increase by Teaching Method",
       subtitle = "Increase compared to 3 Week Lag \nRed area represents Fall Semester",
       color = "Majority Teaching Method") +
  scale_y_continuous(labels = comma)
```


Death Proportion Increase by Teaching Method
Increase compared to 3 Week Lag
Red area represents Fall Semester
2020−12−23
0.00050
0.00040
0.00030
0.00020
0.00010
0.00000
Death Proportion Increase
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date
Majority Teaching Method
On Premises
Hybrid
Online Only