

COVID-19-Economics EDA

Benjamin Chang, Nicolas Min, Emmanuel Rodriguez

February 19, 2021

For Mobility Dataset

Data not included, can be found here

```
mobility=read.csv('Data/2020_US_Region_Mobility_Report.csv', header = T, na.strings = "")

#convert to date format
mobility$date=as.Date(mobility$date)

##Visualization

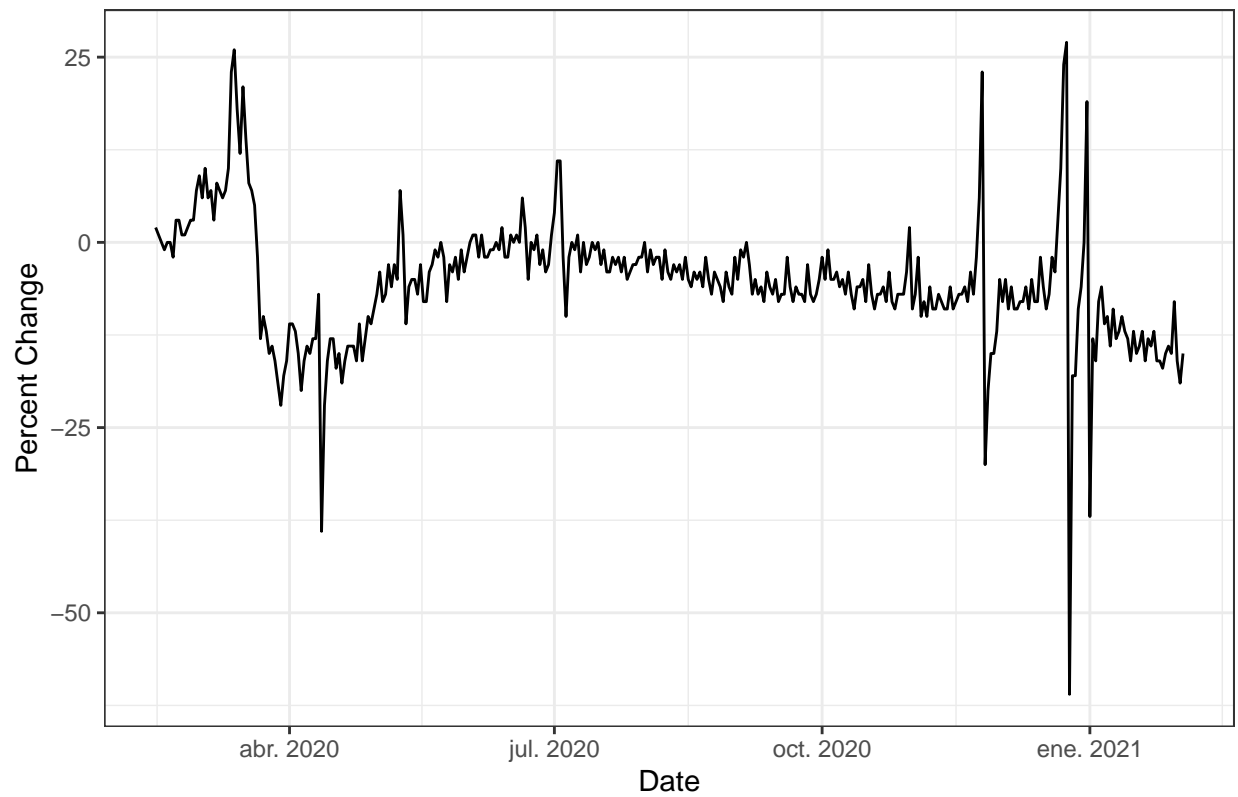
#subset data
usMob=mobility[is.na(mobility$sub_region_1),]
caliMob=mobility[which(mobility$iso_3166_2_code=='US-CA'),]

#US

#Grocery and Pharmacy

ggplot(usMob, aes(x = date, y = grocery_and_pharmacy_percent_change_from_baseline, group=1)) +
  geom_path()+ theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x = 'Date', y = 'Percent Change',
       title = 'US Grocery and Pharmacy Percent Change')+
  theme_bw()
```

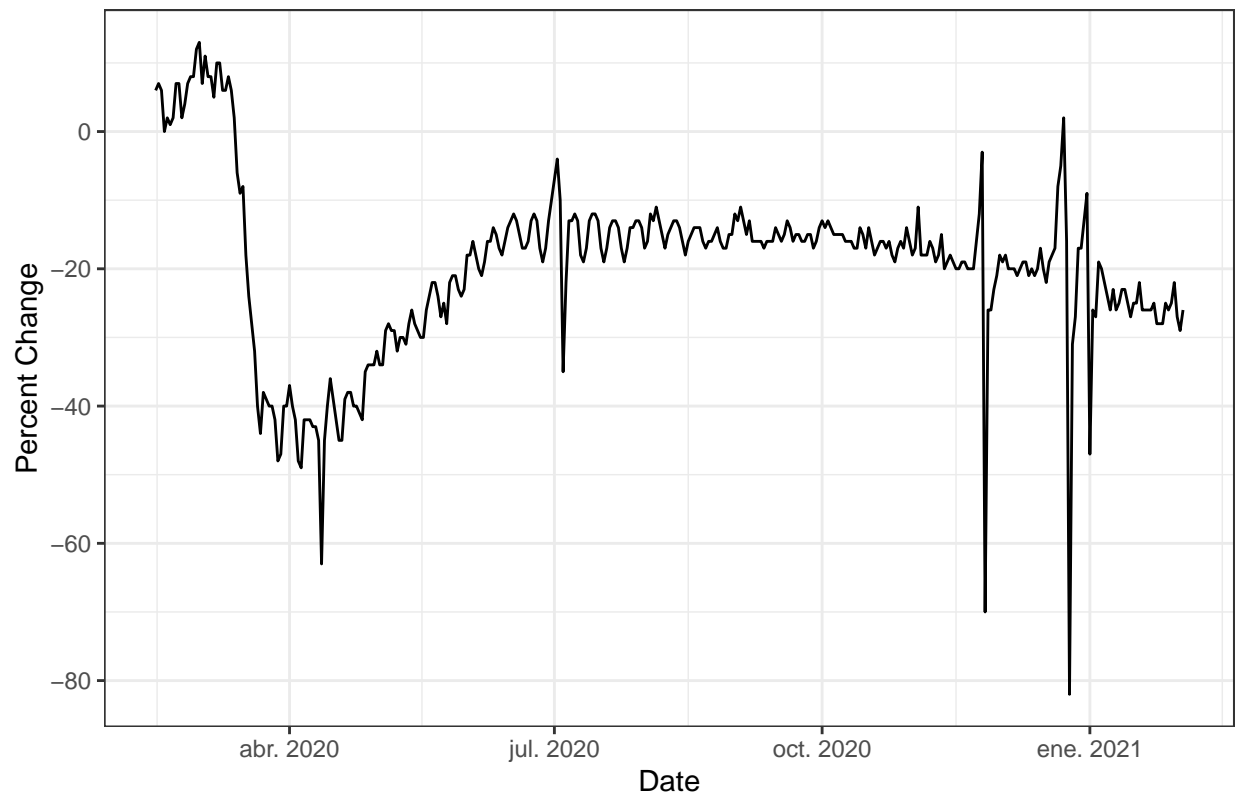
US Grocery and Pharmacy Percent Change



#Recreation Percent Change

```
ggplot(usMob, aes(x = date, y = retail_and_recreation_percent_change_from_baseline, group=1)) +
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x = 'Date', y = 'Percent Change',
  title = 'US Recreation Percent Change') +
  theme_bw()
```

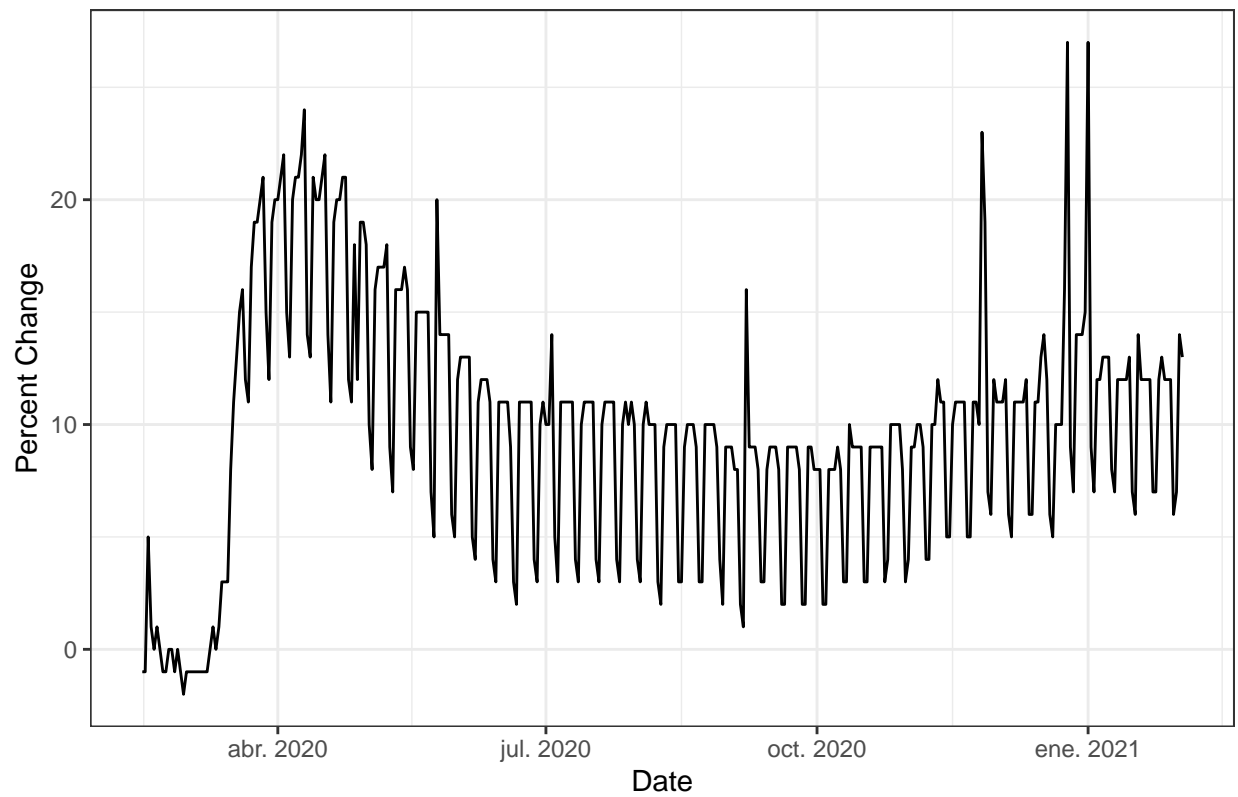
US Recreation Percent Change



#Residential Percent Change

```
ggplot(usMob, aes(x = date, y = residential_percent_change_from_baseline, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'Percent Change',  
       title = 'US Residential Percent Change') +  
  theme_bw()
```

US Residential Percent Change

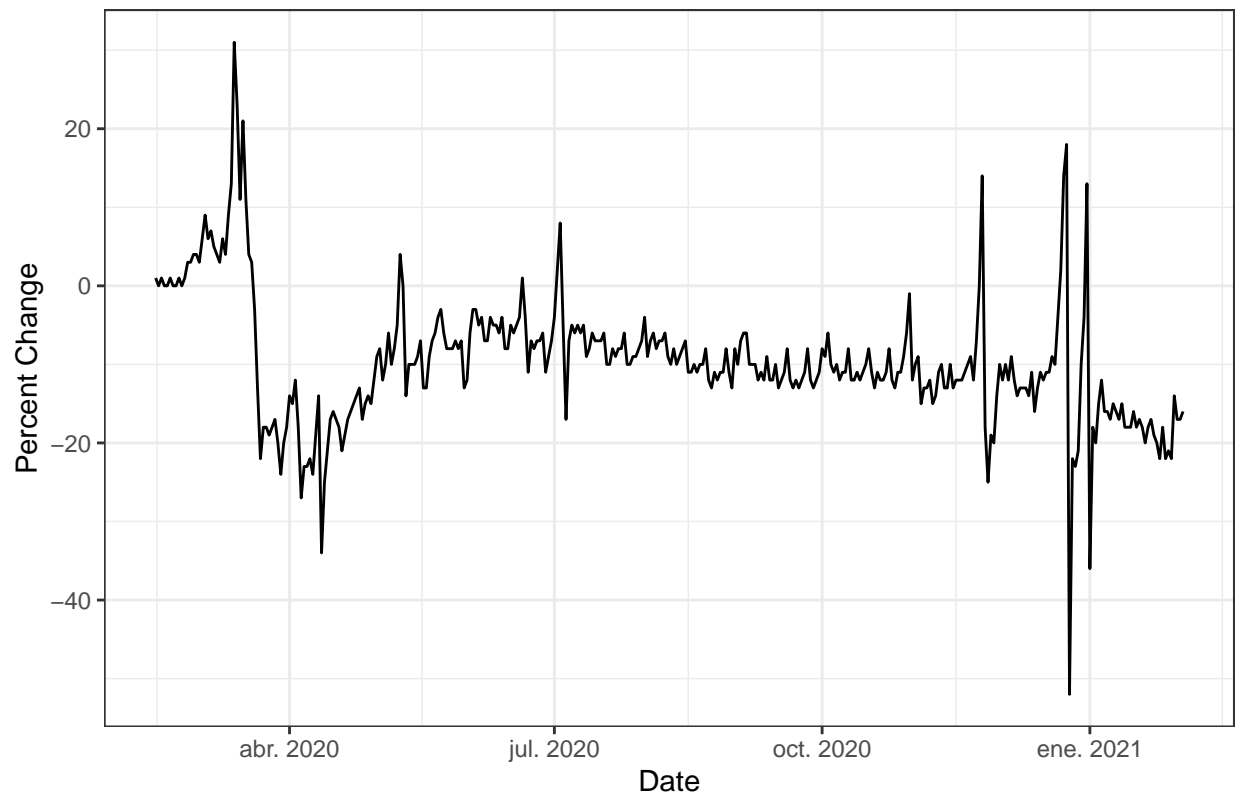


```
#California
```

```
#Recreation Percent Change
```

```
ggplot(caliMob, aes(x = date, y = grocery_and_pharmacy_percent_change_from_baseline, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'Percent Change',  
       title = 'California Grocery and Pharmacy Percent Change') +  
  theme_bw()
```

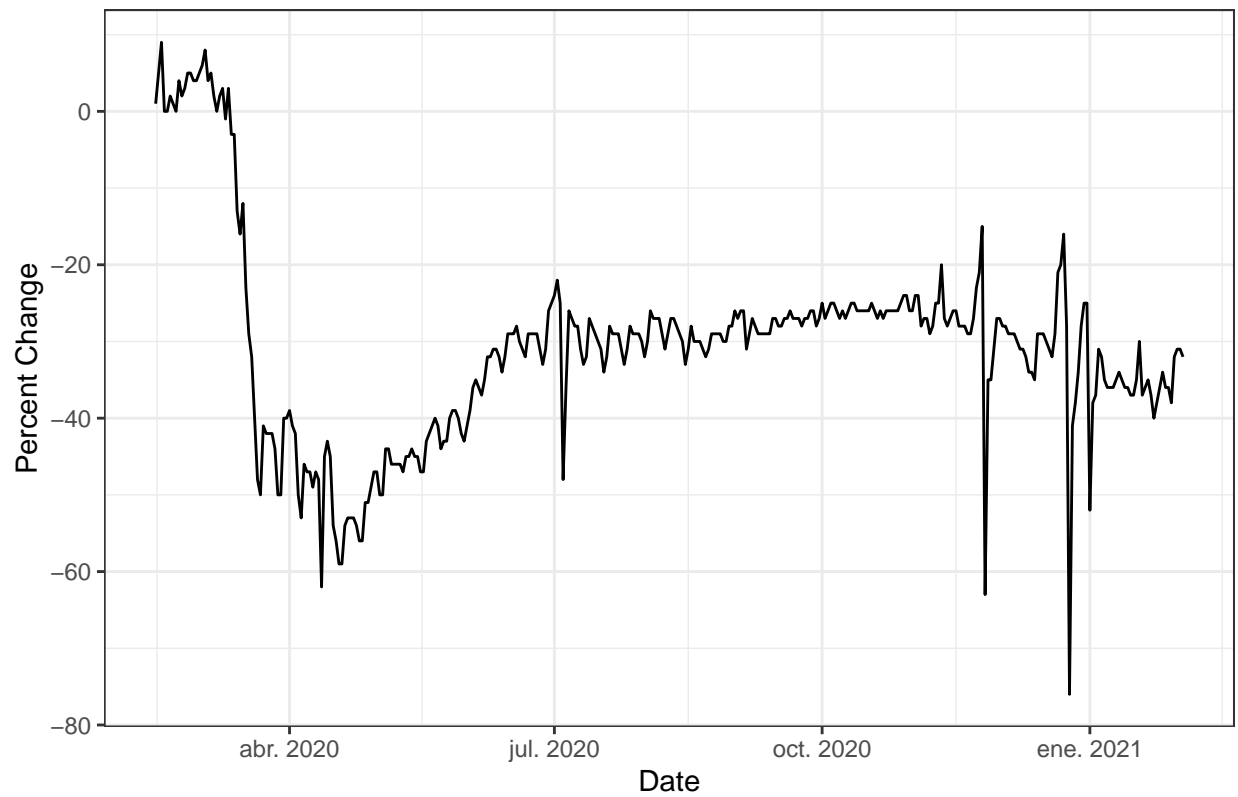
California Grocery and Pharmacy Percent Change



#Recreation Percent Change

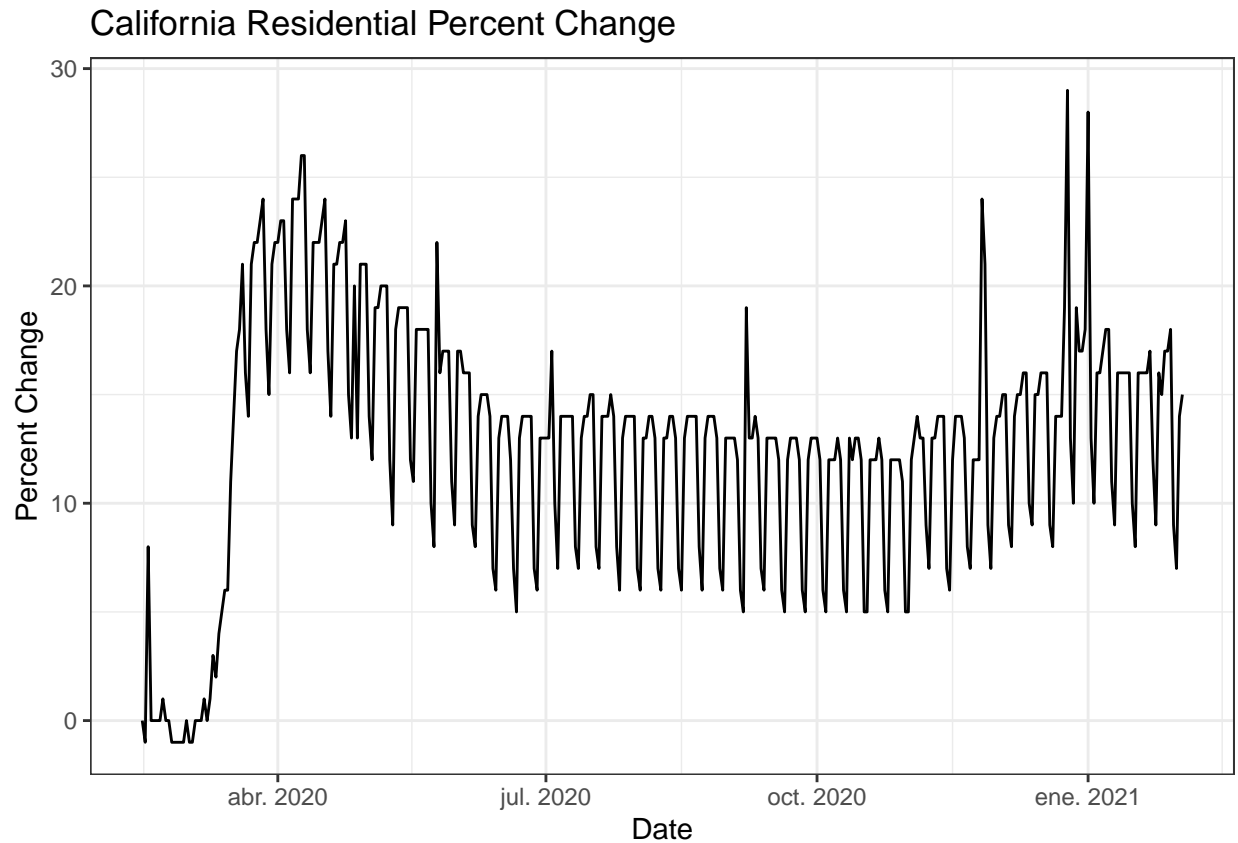
```
ggplot(caliMob, aes(x = date, y = retail_and_recreation_percent_change_from_baseline, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'Percent Change',  
    title = 'California Recreation Percent Change') +  
  theme_bw()
```

California Recreation Percent Change



#Residential Percent Change

```
ggplot(caliMob, aes(x = date, y = residential_percent_change_from_baseline, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'Percent Change',  
       title = 'California Residential Percent Change') +  
  theme_bw()
```



Some Findings:

- The same pattern occurred in the US and California for all different mobility analysis
- For mobility data regarding Grocery and Pharmacy, there was an initial increase due to people rushing to stock up for lockdown, then a general decrease
 - Possibly due to restrictions for indoor capacity, so lines were formed and it causes an incentive to only have 1 family member go grocery shopping
- For Recreation, there is a sharp decrease followed by a level off
 - This can be due to people complying to health restrictions in the beginning, but people were tired of staying indoors all day
- For Residential, there is an opposite pattern as recreation, which might occur due to the same reasoning as mentioned above

For Unemployment Dataset

Link to Data

```
unemploy=read.csv('Data/UnemploySA.csv', header = T, na.strings = "")

## Data Cleaning

#title of columns
nameTitle=c('FIPS Code','State and area','Year','Month', 'Civilian non-institutional population',
            'Civilian labor force','Percent of population','Employment','Percent of population',
            'Unemployment Total','Unemployment Rate' )

names(unemploy) <- nameTitle

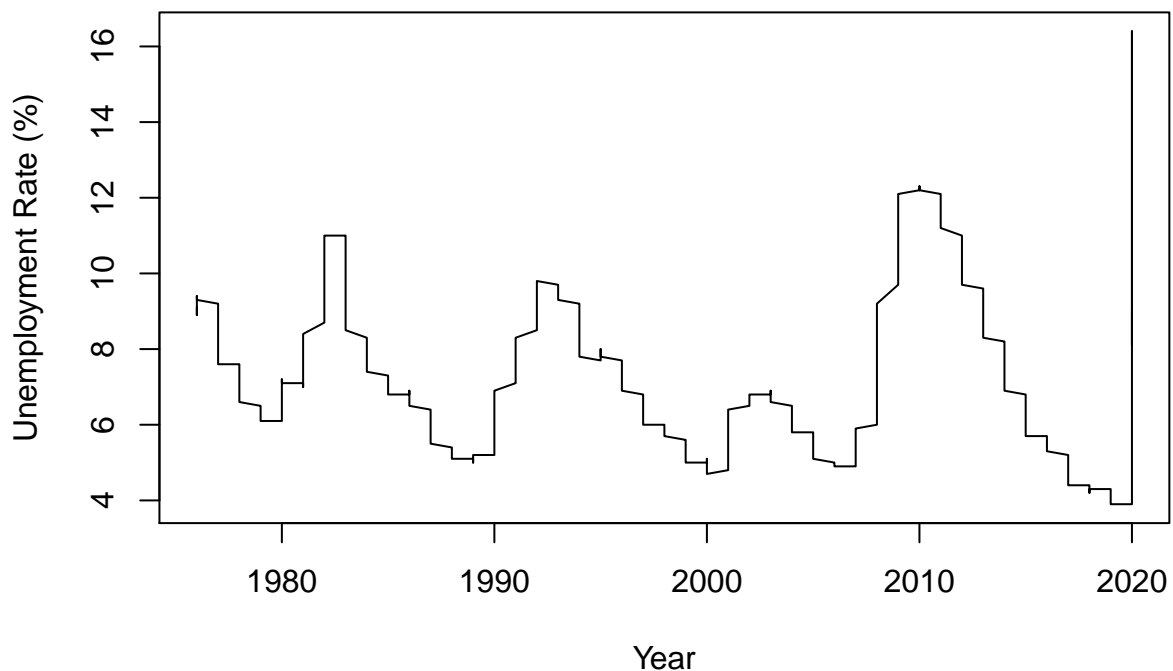
unemploy=unemploy[-c(seq(1,7)),]

## Visualization

#subset data
caliUnem=unemploy[which(unemploy$'State and area'=='California'),]

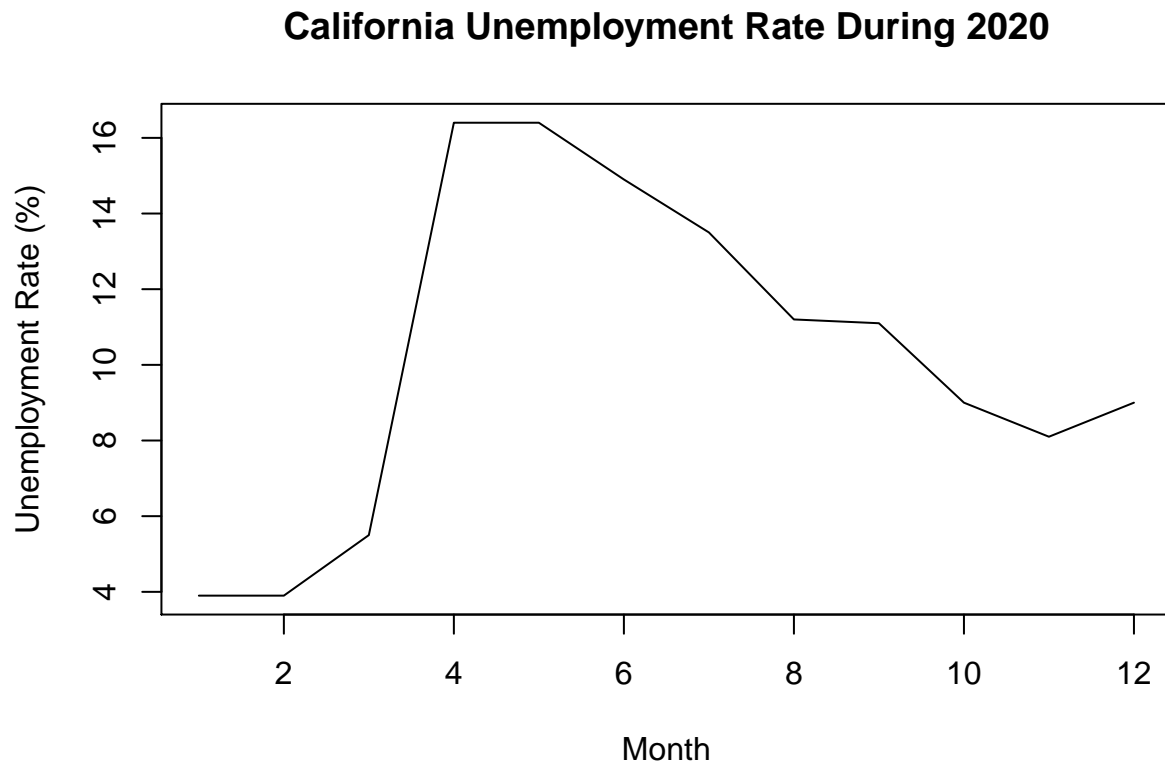
#for cali
plot(caliUnem$Year, caliUnem$'Unemployment Rate', col = 'black', type = 'l',
     main = 'California Unemployment Rate by Year', xlab = 'Year',
     ylab = 'Unemployment Rate (%)')
```

California Unemployment Rate by Year




```
#zoomed cali
```

```
plot(caliUnem$Month[caliUnem$Year>2019],  
     caliUnem$'Unemployment Rate'[caliUnem$Year>2019], col = 'black', type = 'l',  
     main = 'California Unemployment Rate During 2020', xlab = 'Month',  
     ylab = 'Unemployment Rate (%)')
```



Some Findings:

- The unemployment rate in this dataset was seasonally adjusted in California
- For the graph that plotted years since 1976:
 - There seems to be an increase in unemployment rate then there is a recession
- For the graph that plotted the months in 2020:
 - There is a huge spike in unemployment when health restrictions were created, followed by a decrease until around Thanksgiving
 - This might be due to a fiscal impact on multiple companies as people are staying home and not going outside

For Unemployment (Sex, Race) Dataset

Link to Data

```
unemploym=read.csv('Data/unemploymicro.csv', header = T, na.strings = "")

##Cleaning

#corresponding Id to sex,race
nameID=c('LNS14000000','LNS14000003','LNS14000006','LNS14000009',
         'LNS14000025','LNS14000026','LNS14032183')
nameTitle=c('Overall','White','Black','Latino', 'Men','Women','Asian')

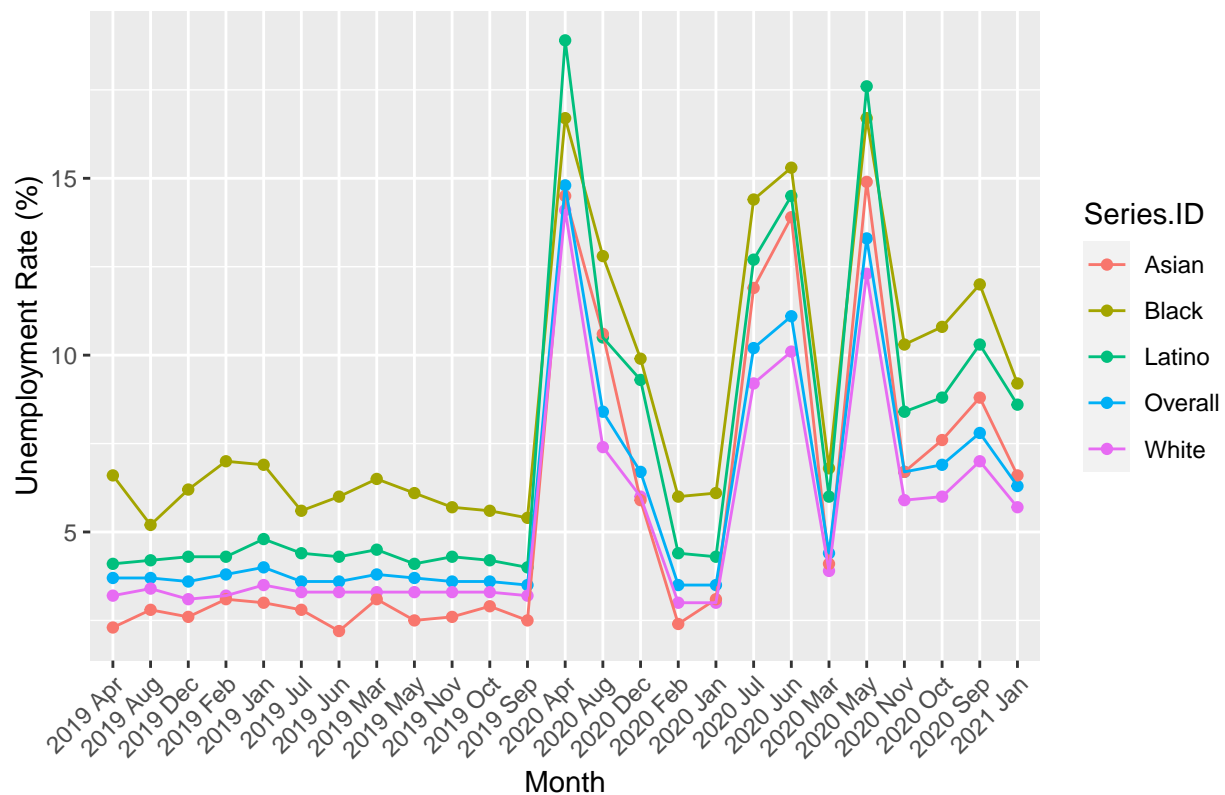
for (i in 1:length(nameID)){
  unemploym$Series.ID[unemploym$Series.ID==nameID[i]]=nameTitle[i]
}

##Visualizations

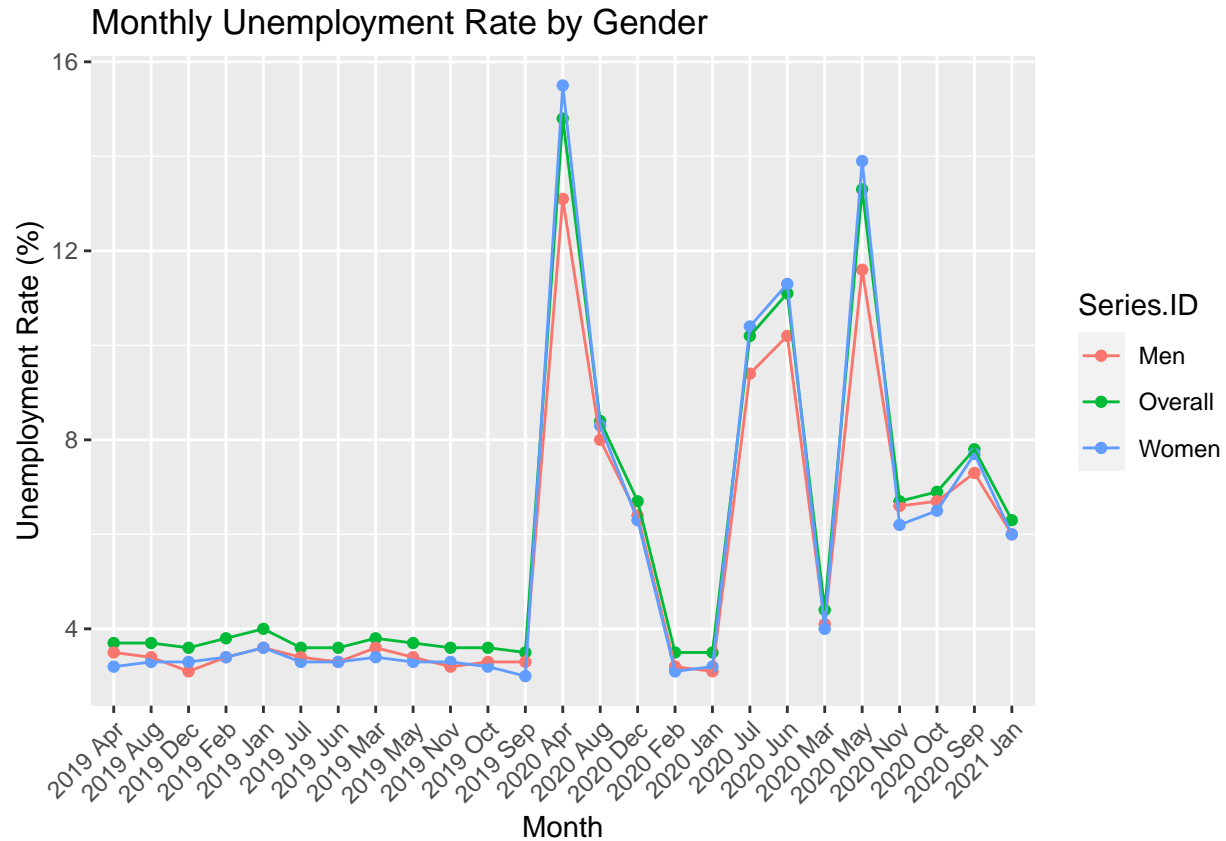
dfrace=unemploym[which(unemploym$Series.ID %in% c('Overall','White','Black',
         'Latino', 'Asian'))],]
dfgender=unemploym[which(unemploym$Series.ID %in% c('Overall','Men','Women'))],]

ggplot(dfrace, aes(x = Label, y = Value, group=Series.ID, color=Series.ID)) +
  geom_point() + geom_line()+theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x = 'Month', y = 'Unemployment Rate (%)',
       title = 'Monthly Unemployment Rate by Race')
```

Monthly Unemployment Rate by Race



```
ggplot(dfgender, aes(x = Label, y = Value, group=Series.ID, color=Series.ID)) +
  geom_point() + geom_line()+theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x = 'Month', y = 'Unemployment Rate (%)',
       title = 'Monthly Unemployment Rate by Gender')
```



```
##Stat Tests
```

```
#subset data
```

```
dfwhite=unemploy[which(unemploy$Series.ID=='White'),]
dflatino=unemploy[which(unemploy$Series.ID=='Latino'),]
dfasian=unemploy[which(unemploy$Series.ID=='Asian'),]
dfblack=unemploy[which(unemploy$Series.ID=='Black'),]
```

```
dfwomen=unemploy[which(unemploy$Series.ID=='Women'),]
dfmen=unemploy[which(unemploy$Series.ID=='Men'),]
```

```
#t tests
```

```
##white and latino
```

```
t.test(dfwhite$Value,dflatino$Value)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: dfwhite$Value and dflatino$Value
```

```
## t = -1.9378, df = 43.007, p-value = 0.05923
```

```
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
## -4.30999071 0.08599071
## sample estimates:
## mean of x mean of y
## 5.320 7.432

## black and asian
t.test(dfblack$Value,dfasian$Value)

##
## Welch Two Sample t-test
##
## data: dfblack$Value and dfasian$Value
## t = 2.6788, df = 47.356, p-value = 0.01013
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.7614006 5.3505994
## sample estimates:
## mean of x mean of y
## 8.792 5.736

## asian and latino
t.test(dfasian$Value,dflatino$Value)

##
## Welch Two Sample t-test
##
## data: dfasian$Value and dflatino$Value
## t = -1.3743, df = 47.9, p-value = 0.1757
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -4.1774453 0.7854453
## sample estimates:
## mean of x mean of y
## 5.736 7.432

## women and men
t.test(dfwomen$Value,dfmen$Value)

##
## Welch Two Sample t-test
##
## data: dfwomen$Value and dfmen$Value
## t = 0.24757, df = 46.292, p-value = 0.8056
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.653979 2.117979
## sample estimates:
## mean of x mean of y
## 5.680 5.448
```

Some Findings:

- On average, the unemployment rate between women were a bit larger than men
- Latinos were affected the most in terms of unemployment rate due to covid, while white was affected the least

- a t-test was ran against the unemployment rate and we obtained a p-value of 0.06, which means that the difference is not statistically significant

For Covid Cases over Time

Link to Data

```
covid=read.csv('Data/uscases.csv', header = T, na.strings = "")
covidstate=read.csv('Data/usstatecases.csv', header = T, na.strings = "")

##Initial Cleaning

#convert to date format
covidstate$date=as.Date(covidstate$date)
covid$date=as.Date(covid$date)

calicovid=covidstate[which(covidstate$state=='California'),]

#new cases
calicovid=calicovid %>% mutate(newcases = c(0,diff(cases)))

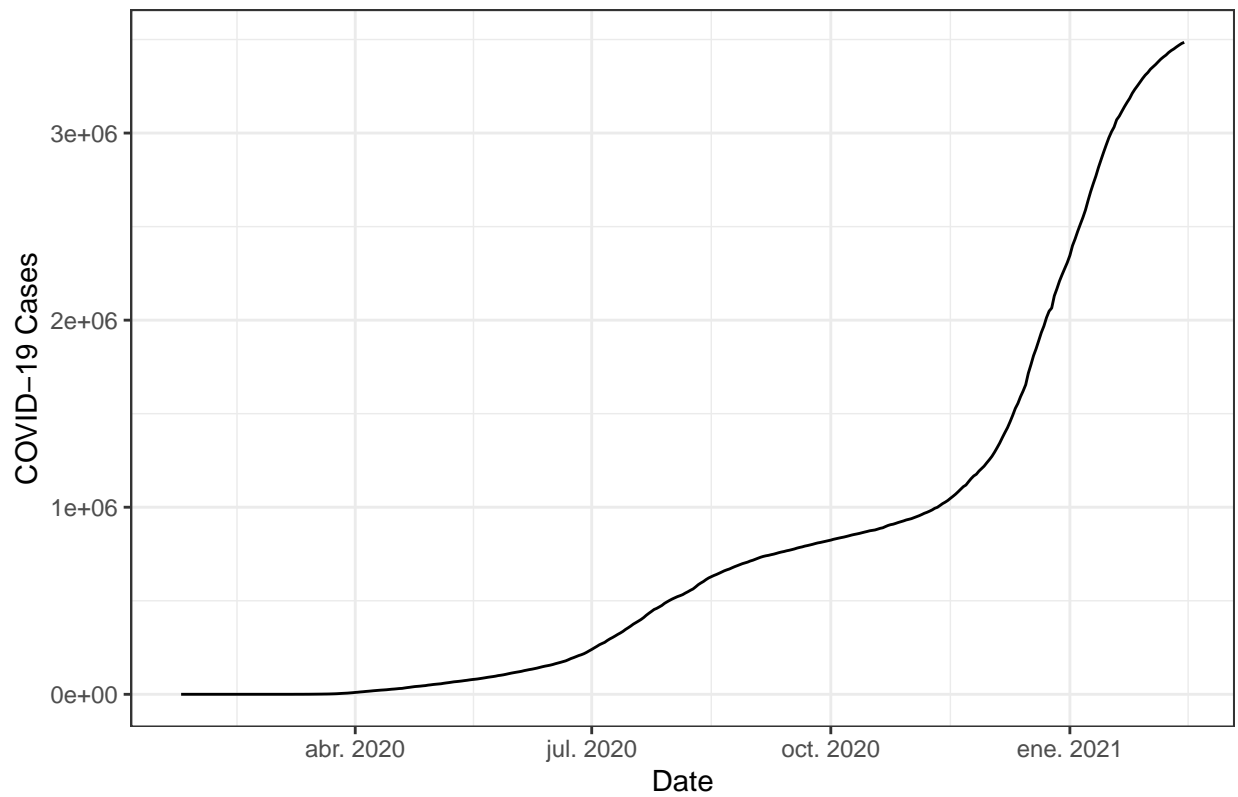
covid=covid %>% mutate(newcases = c(0,diff(cases)))

## Visualizations

#visualizations for cali

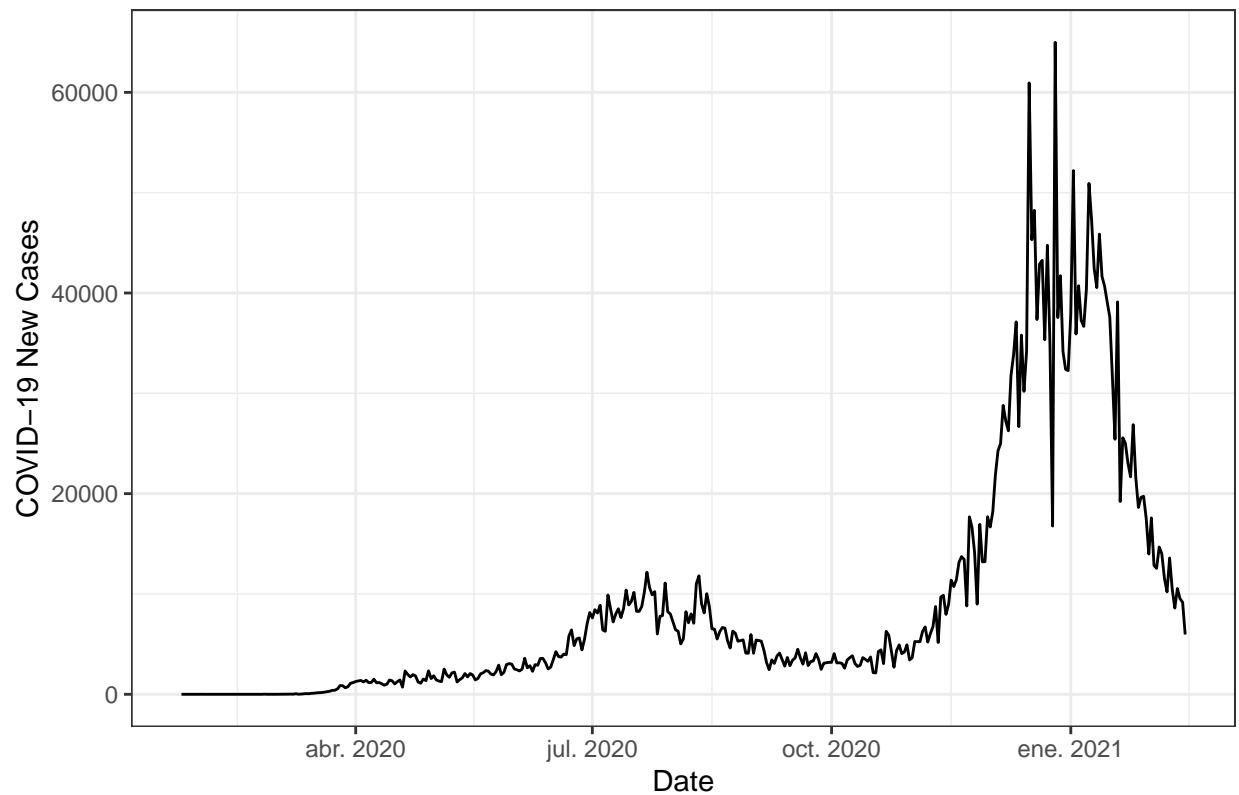
ggplot(calicovid, aes(x = date, y = cases, group=1)) +
  geom_path()+ theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x = 'Date', y = 'COVID-19 Cases',
       title = 'California COVID-19 Cases over Time')+
  theme_bw()
```

California COVID-19 Cases over Time



```
ggplot(calicovid, aes(x = date, y = newcases, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'COVID-19 New Cases',  
       title = 'California COVID-19 New Cases over Time') +  
  theme_bw()
```

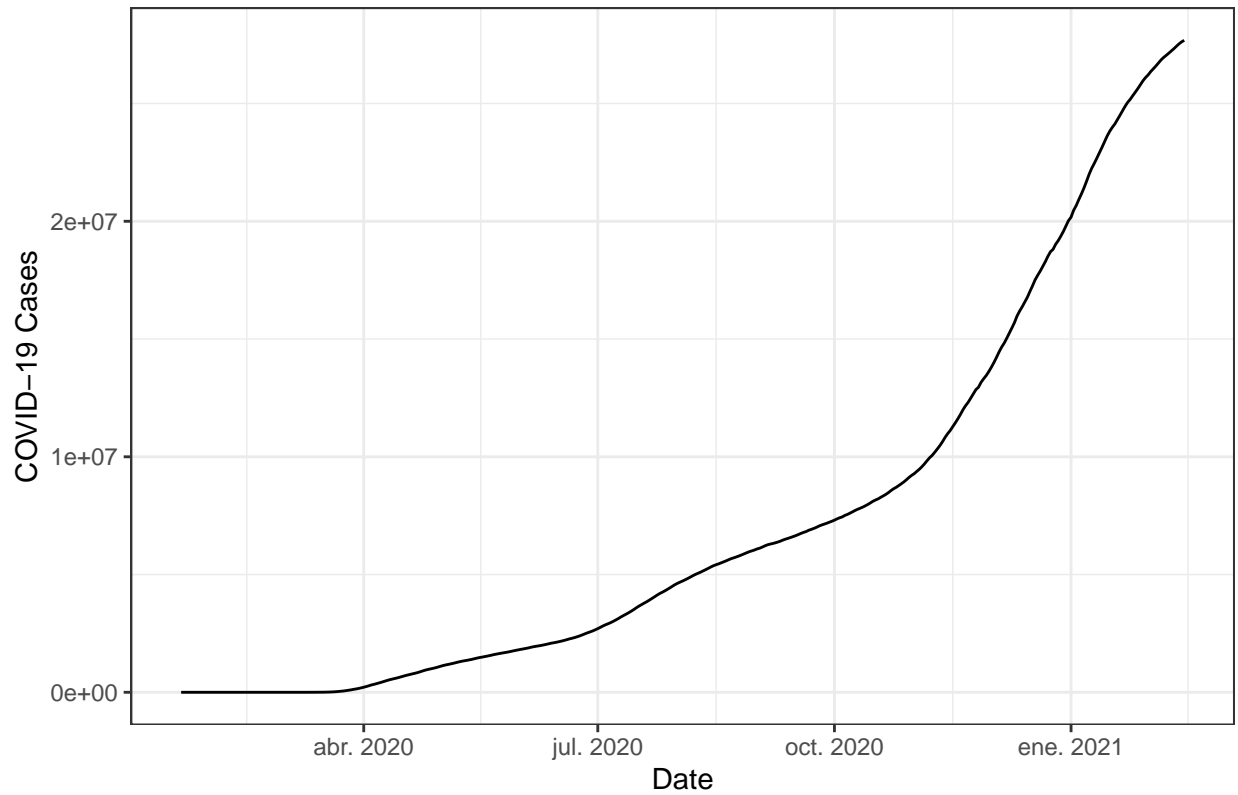
California COVID-19 New Cases over Time



#For US

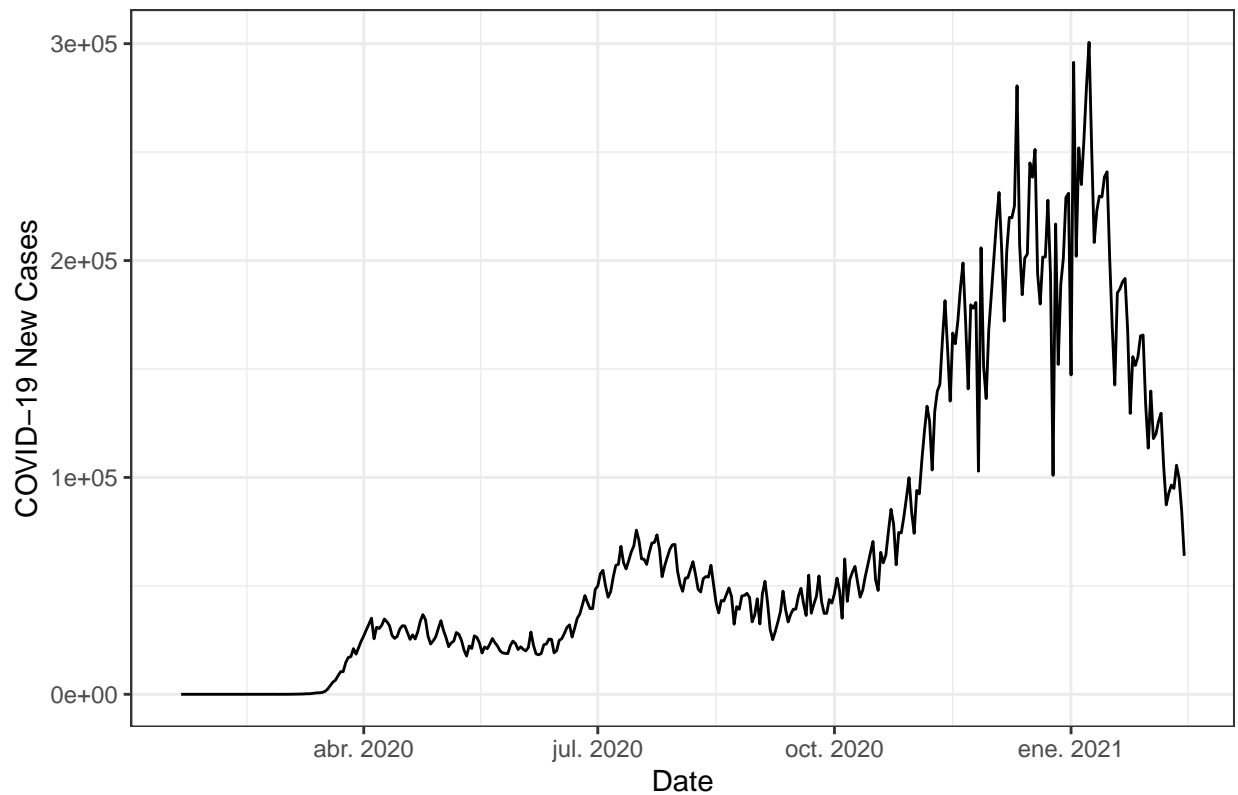
```
ggplot(covid, aes(x = date, y = cases, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'COVID-19 Cases',  
       title = 'US COVID-19 Cases over Time') +  
  theme_bw()
```


US COVID-19 Cases over Time



```
ggplot(covid, aes(x = date, y = newcases, group=1)) +  
  geom_path() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  labs(x = 'Date', y = 'COVID-19 New Cases',  
       title = 'US COVID-19 New Cases over Time') +  
  theme_bw()
```

US COVID-19 New Cases over Time



Some Findings:

- For the COVID cases in US:
 - There seems to be a small spike around March and one around June
 - This might be due to non-strict health restrictions in some states, causing the virus to spread
 - There is a huge spike around the holidays of 2020
 - * People might have traveled home to family and ignored recommendations
- For the COVID cases in California:
 - There is only one small spike around June due to effective health restrictions
 - The same spike occurred over the holidays of 2020 due to people traveling home

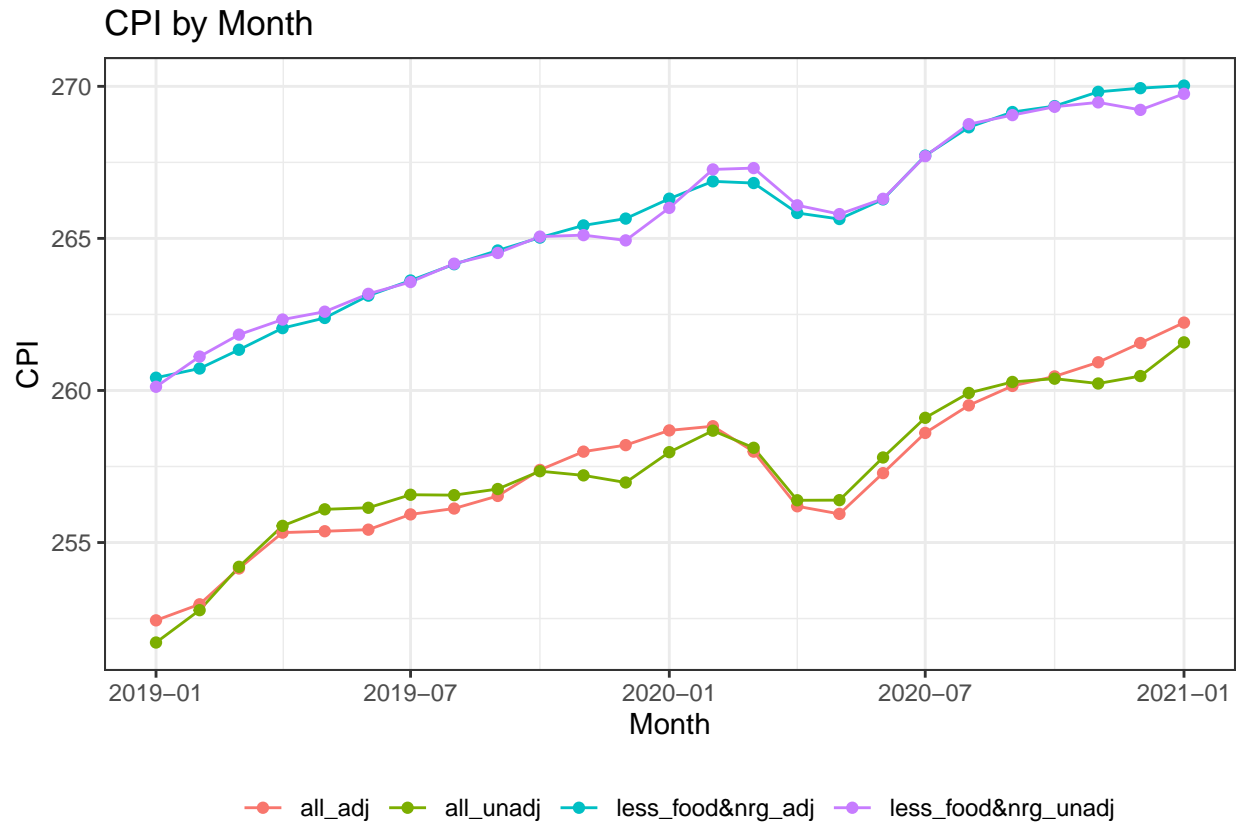
For CPI Dataset

Link to Data

```
#import data
cpi_all_adj <- read.csv('Data/cpi_all_adj.csv', header = T, na.strings = "")
cpi_all_adj$Series.ID <- 'all_adj'
cpi_all_notadj <- read.csv('Data/cpi_all_notadj.csv', header = T, na.strings = "")
cpi_all_notadj$Series.ID <- 'all_unadj'
cpi_lessfood_adj <- read.csv('Data/cpi_lessfood_adj.csv', header = T, na.strings = "")
cpi_lessfood_adj$Series.ID <- 'less_food&nrg_adj'
cpi_lessfood_notadj <- read.csv('Data/cpi_lessfood_notadj.csv', header = T, na.strings = "")
cpi_lessfood_notadj$Series.ID <- 'less_food&nrg_unadj'
cpi <- rbind(cpi_all_adj, cpi_all_notadj, cpi_lessfood_adj, cpi_lessfood_notadj)

#clean data - get the dates
cpi$Period <- str_replace(cpi$Period, "M", "")
cpi$month <- paste(cpi$Year, cpi$Period, '01', sep = '-')
cpi$month <- as.Date(cpi$month)

#plot
ggplot(cpi, aes(x=month, y=Value, group=Series.ID, color=Series.ID)) +
  geom_line() +
  geom_point() +
  ylab("CPI") +
  xlab("Month") +
  ggtitle("CPI by Month") +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank())
```



Some Findings:

- We see a dip in the price index. Perhaps this is due to people getting unemployed.
- the aggregate demand cannot keep up with the supply of the nation.

```
#import 2008 data
cpi_lessfood_adj_2008 <- read.csv('Data/cpi_lessfood_adj_2008.csv', header = T,
                                na.strings = "")
cpi_lessfood_adj_2008$Series.ID <- 'less_food&nrg_adj'
cpi_lessfood_notadj_2008 <- read.csv('Data/cpi_lessfood_notadj_2008.csv', header = T,
                                    na.strings = "")
cpi_lessfood_notadj_2008$Series.ID <- 'less_food&nrg_unadj'
cpi_less_food_2008 <- rbind(cpi_lessfood_adj_2008, cpi_lessfood_notadj_2008)
cpi_less_food_2020 <- rbind(cpi_lessfood_adj, cpi_lessfood_notadj)

#clean data - get the dates
cpi_less_food_2008$Period <- str_replace(cpi_less_food_2008$Period, "M", "")
cpi_less_food_2008$month <- paste(cpi_less_food_2008$Year, cpi_less_food_2008$Period,
                                '01', sep = '-')
cpi_less_food_2008$month <- as.Date(cpi_less_food_2008$month)
cpi_less_food_2020$Period <- str_replace(cpi_less_food_2020$Period, "M", "")
cpi_less_food_2020$month <- paste(cpi_less_food_2020$Year, cpi_less_food_2020$Period,
                                '01', sep = '-')
cpi_less_food_2020$month <- as.Date(cpi_less_food_2020$month)

p1 <- ggplot(cpi_less_food_2008, aes(x=month, y=X12.Month...Change, group=Series.ID,
```

```

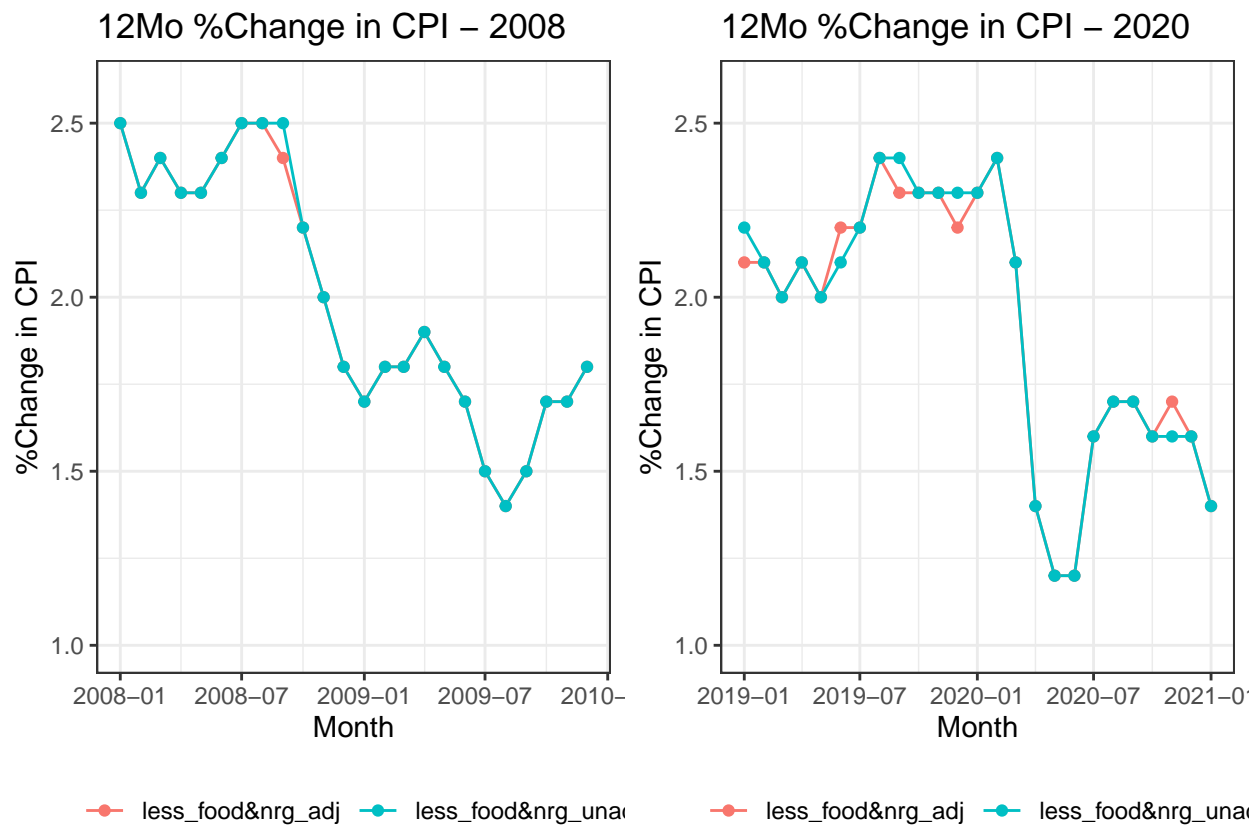
                                color=Series.ID)) +
  geom_line()+
  geom_point()+
  ylab("%Change in CPI") +
  xlab("Month") +
  ylim(1,2.6) +
  ggtitle("12Mo %Change in CPI - 2008") +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank())

p2 <- ggplot(cpi_less_food_2020, aes(x=month, y=X12.Month...Change, group=Series.ID,
                                color=Series.ID)) +

  geom_line()+
  geom_point()+
  ylab("%Change in CPI") +
  xlab("Month") +
  ylim(1,2.6) +
  ggtitle("12Mo %Change in CPI - 2020") +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank())

grid.arrange(p1, p2, nrow = 1)

```



Some Findings:

- Using the less food&energy data, we compared the % change (12 month) in CPI of 2020 to that of 2008.
- The 2020's dip happened over much shorter period of time (in 2-3 months) than 2008's.
- This emphasizes the magnitude of the impact of COVID-19 - how quickly it dropped the price level of goods as demands dropped (again, due to unemployment, stay at home order, and etc.). As expected, the pandemic outbreak had a greater impact on consumer goods than the financial crisis.

For Prime rate and FFR Dataset

[Link to Data](#)

[Link to Data](#)

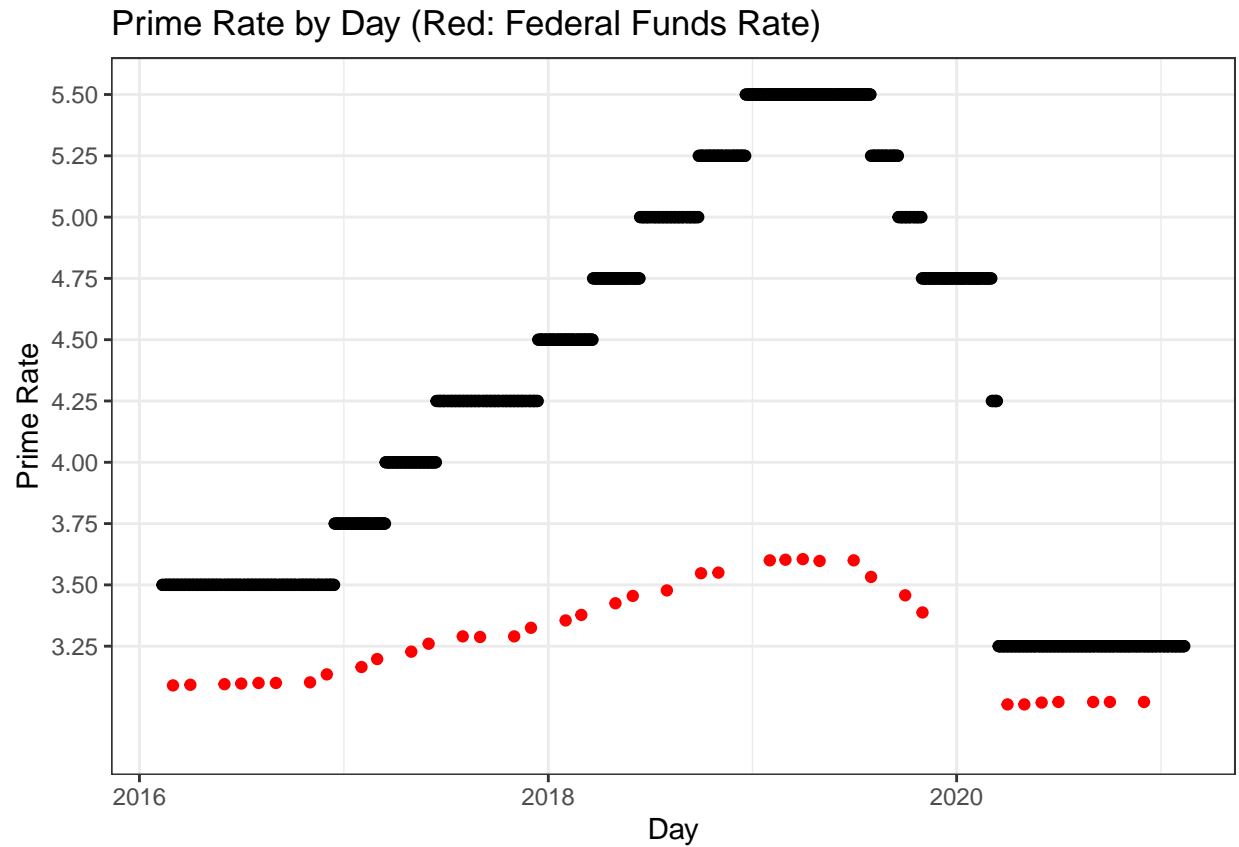
```
#read in data
prime_rate <- read.csv('Data/DPRIME.csv', header = T, na.strings = "")
fed_funds_rate <- read.csv('Data/FEDFUNDS.csv', header = T, na.strings = "")
rates <- left_join(prime_rate, fed_funds_rate)
```

```
## Joining, by = "DATE"
```

```
#clean
rates$DATE <- as.Date(prime_rate$DATE)
rates <- rates %>% filter(DPRIME != '.')
```

```
#plot
rates %>%
  ggplot(aes(x = DATE)) +
  geom_point(aes(y = DPRIME), color = 'black') +
  geom_point(aes(y = FEDFUNDS), color = 'red') +
  coord_cartesian(ylim=c(-0.5,10))+
  ylab("Prime Rate") +
  xlab("Day") +
  ggtitle("Prime Rate by Day (Red: Federal Funds Rate)") +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank())
```

```
## Warning: Removed 1223 rows containing missing values (geom_point).
```



Some Findings:

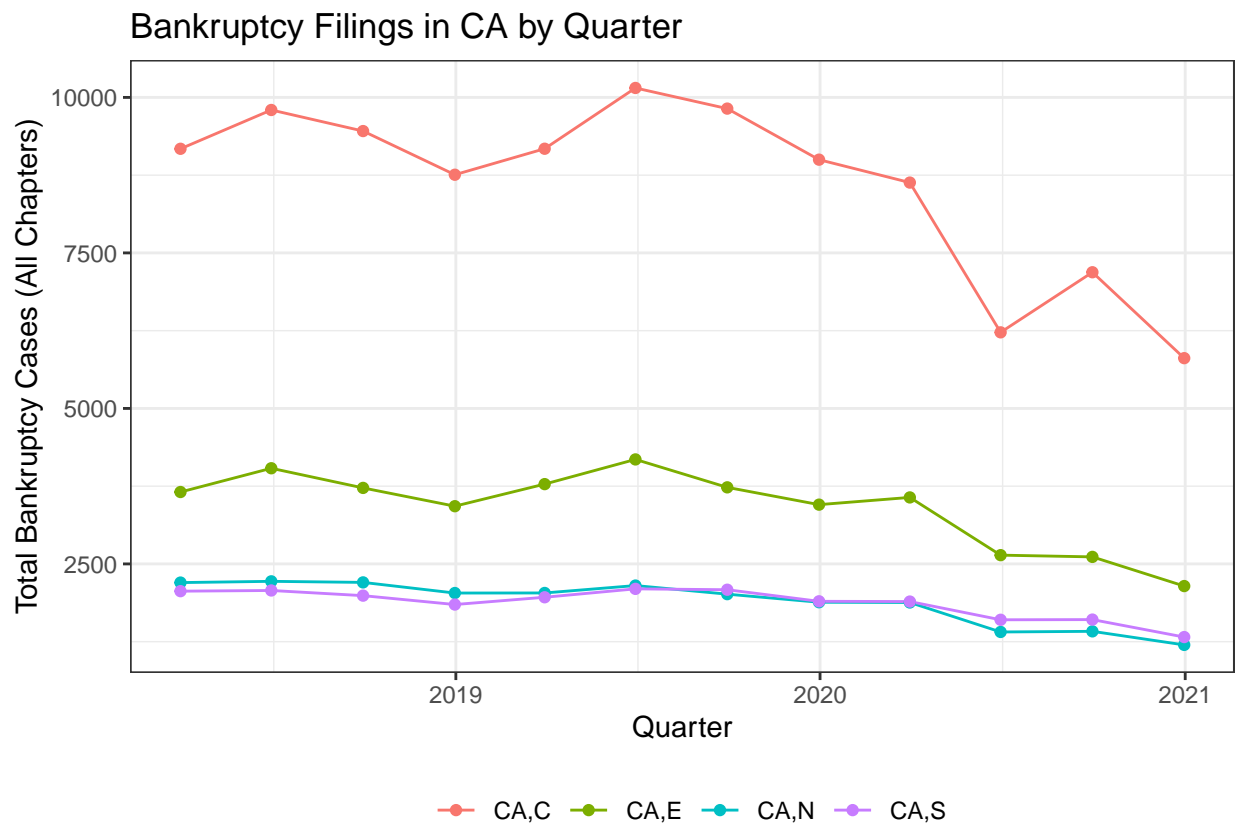
-As the Fed decided to adjust its federal funds rates (the rate that banks charge each other for short-term loans) back down to near zero level just after the pandemic outbreak, we see banks setting their prime rates low at 3.25% as well.

For Bankruptcy Dataset

Link to Data

```
#Read data
bankruptcy_california <- read_excel("Data/bankruptcy_california.xlsx")
bankruptcy_california$time=as.Date(bankruptcy_california$time)

#Plot
ggplot(bankruptcy_california, aes(x=time, y=TotalAllChapters,
                                group=CircuitandDistrict, color=CircuitandDistrict)) +
  geom_line()+
  geom_point()+
  ylab("Total Bankruptcy Cases (All Chapters)") +
  xlab("Quarter") +
  ggtitle("Bankruptcy Filings in CA by Quarter") +
  theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank())
```



Some Findings:

- Data only is by quarter, so we have manually combined the California data from 2018 to 2020.
- All in all, the bankruptcy dropped after the COVID-19! This is counter intuitive.
- We looked at the bankruptcy data to see the impact of COVID-19 on the most economically vulnerable people. However, looking at the decrease of bankruptcy after the pandemic, we noticed perhaps bankruptcy data is not a very representative economic indicator. There are too many variables that are

in play: significant slowdown in the government/administrative process, PPP loan, and etc. In fact, the slowdown in the government is easily observed in many places. One good example is from the SSA (Social Security Administration)'s disability processing data (not included in this report)

- We observed that SSA has been seeing significant delays in processing claims and transferring disability applications to the state agencies.

For Stringency dataset

Link to Data

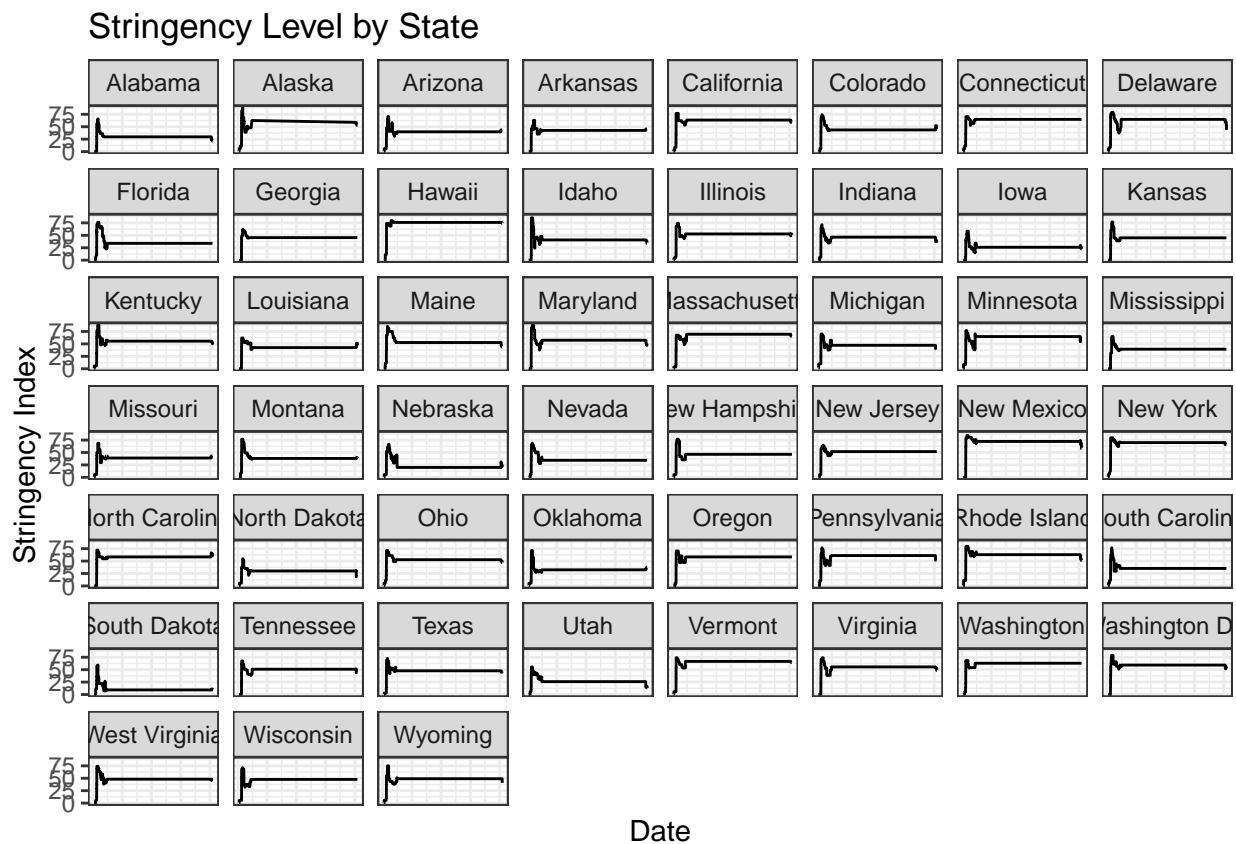
Link to Data

```
#load data
stringency_us <- read.csv("https://raw.githubusercontent.com/OxCGRT/USA-covid-policy/master/data/OxCGRT")

stringency_states <- stringency_us[stringency_us$RegionName!="",]

#plot (stringency)
ggplot(stringency_states) +
  geom_line(aes(x=Date, y=StringencyIndex), color = "black")+
  ylab("Stringency Index") +
  xlab("Date") +
  ggtitle("Stringency Level by State") +
  theme_bw() +
  theme(legend.title = element_blank(), axis.text.x = element_blank(), axis.ticks.x = element_blank()) +
  facet_wrap("RegionName")
```

Warning: Removed 4 row(s) containing missing values (geom_path).



Some findings:

- We realized that the Baek et al paper looks at Stay-at-Home (SAH) order almost as a binary variable. During the time when the paper was written, this approach made sense. However, the SAH order has taken many forms and shapes since. And we noticed different states in the U.S. impose different

restrictions. Oxford U. provides an interesting data set that looks at the stringency level at a state level (calculated based on various indicators such as: school closure, workplace closure, restrictions on gathering, and etc.).

- Using the data set, we were able to identify that some states have more stringent orders than others.
- We plan to dig deeper into whether or not the indices that Oxford U. has come up with is valid and reliable. If it is, we think it could be a good complement to the Google's Mobility Dataset.

For Personal Income and its disposition

Link to data (Only tain table T20600)

```
##Only taking worksheet
#### Income data, US ###

#Read data and select columns of interest

### Index (Line, Data)
## A065RC. Personal Income (current dollars)
## A033RC. Compensation of employees
## A041RC. Proprietors' income with inventory valuation and capital consumption adjustments
## A048RC. Rental income of persons with capital consumption adjustment
## W210RC. Personal income receipts on assets
## A577RC. Personal current transfer receipts
## W825RC. Unemployment insurance benefit
## DPCERC. Personal Consumption expenditure (current)
## A072RC. Savings as a percentage of disposable personal income
## A067RX. Personal Income (chained)
## A229RX. Personal Income per capita (chained)

income_table <- read_excel("Data/T20600 - Personal Income and its Disposition.xlsx", skip=6)[,-c(1:2)]

## New names:
## * `` -> ...2
## * `` -> ...3

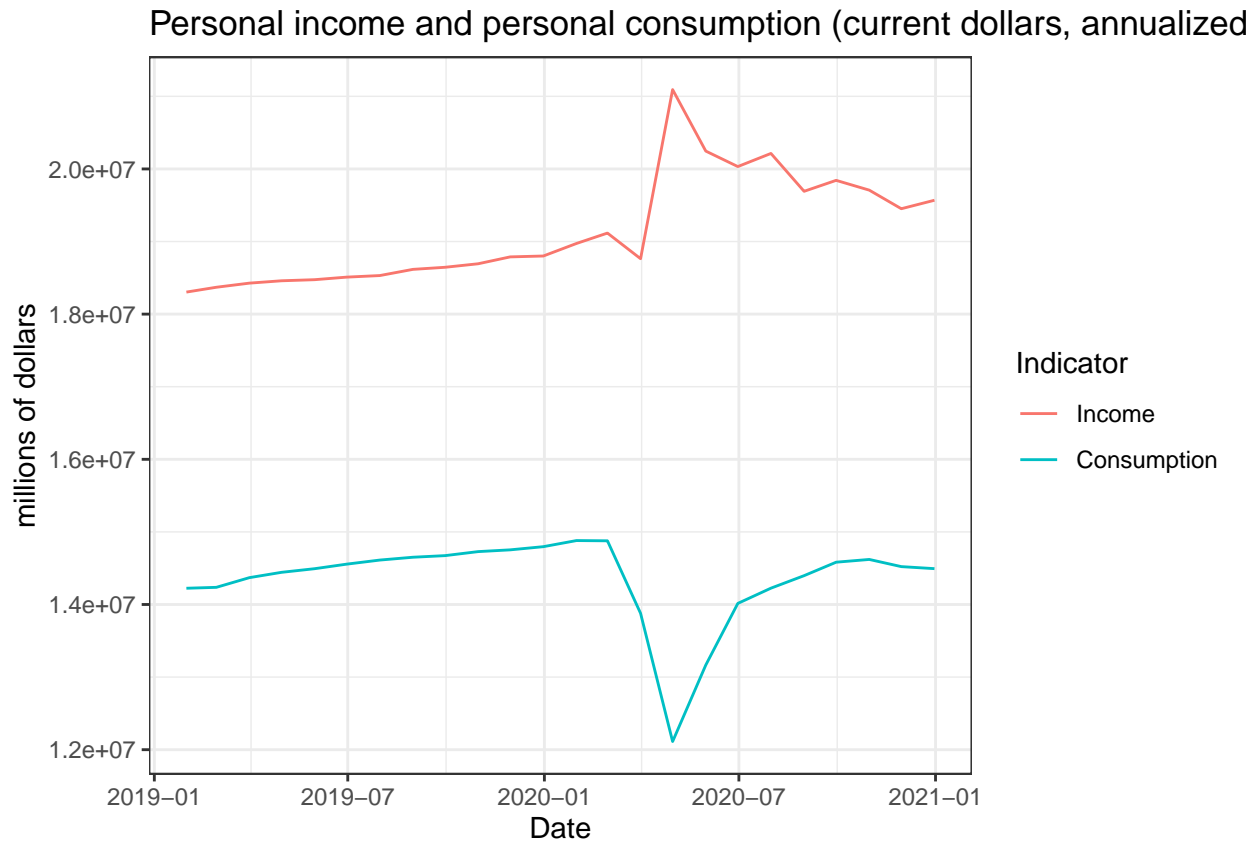
colnames(income_table)[1] <- 'id'
income_table <- as.data.frame(income_table) %>% filter(id %in%
  c('A065RC','A033RC','A041RC',
    'A048RC','W210RC','A577RC','W825RC',
    'DPCERC','A072RC',
    'A067RX','A229RX'))

#Create data table for income
income <- data.frame(date = seq(as.Date("1959-02-01"),length=12*(2020-1959+1),by="months")-1,
  income_total = as.numeric(filter(income_table,id=='A065RC')[-1]),
  income_wages = as.numeric(filter(income_table,id=='A033RC')[-1]),
  income_interest = as.numeric(filter(income_table,id=='W210RC')[-1]),
  income_govbenefits = as.numeric(filter(income_table,id=='A577RC')[-1]),
  income_others = as.numeric(filter(income_table,id=='A041RC')[-1])+
    as.numeric(filter(income_table,id=='A048RC')[-1]),
  income_govbenefits_unemployment = as.numeric(filter(income_table,id=='W825RC')[-1]),
  consumption_total = as.numeric(filter(income_table,id=='DPCERC')[-1]),
  savings_rate = as.numeric(filter(income_table,id=='A072RC')[-1]),
  income_chained = as.numeric(filter(income_table,id=='A067RX')[-1]),
  income_per_capita = as.numeric(filter(income_table,id=='A229RX')[-1]))

#Plots for Income

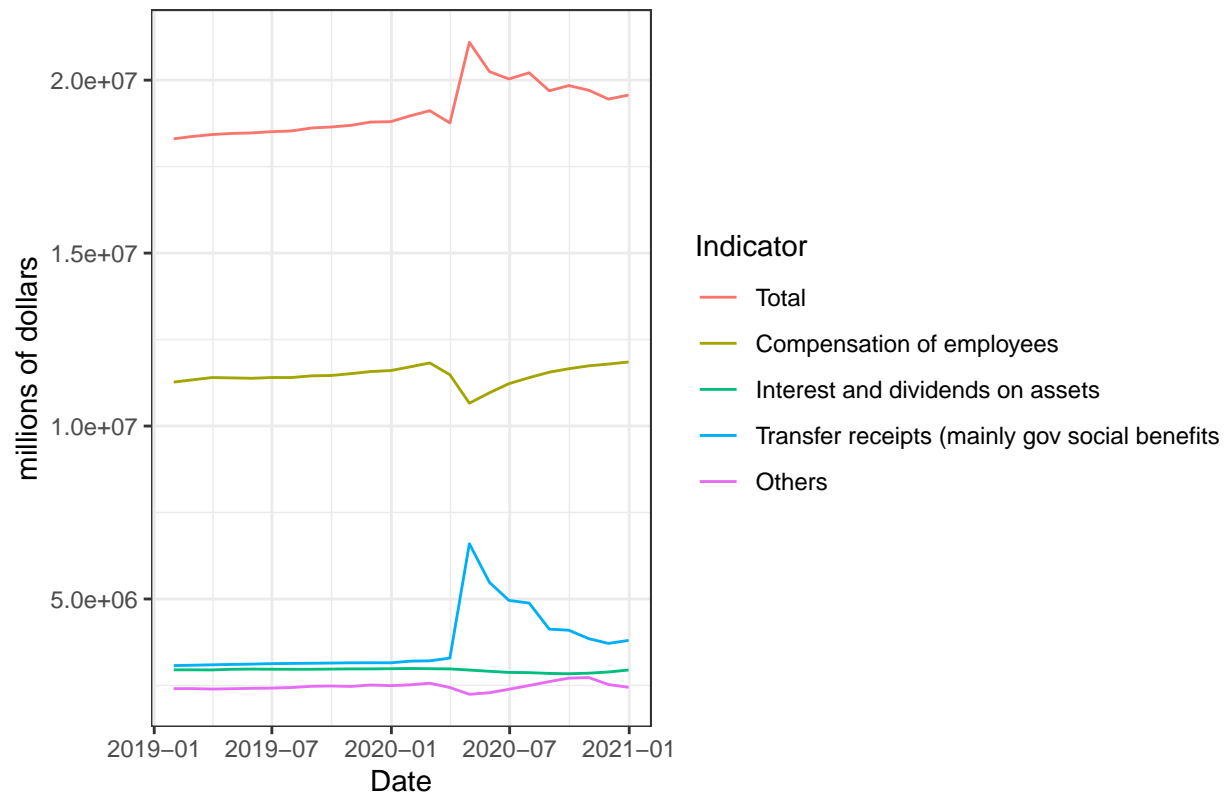
#Income and Consumption (current dollars)
income %>% select(date, income_total, consumption_total) %>% filter(date >= '2019-01-01') %>%
  melt(id='date') %>%
```

```
ggplot(aes(x=date, y=value, colour = variable))+
  geom_line()+
  ggtitle("Personal income and personal consumption (current dollars, annualized)")+
  labs(y = "millions of dollars", x = "Date")+
  scale_colour_discrete(name = 'Indicator', labels = c('Income', 'Consumption'))+
  theme_bw()
```



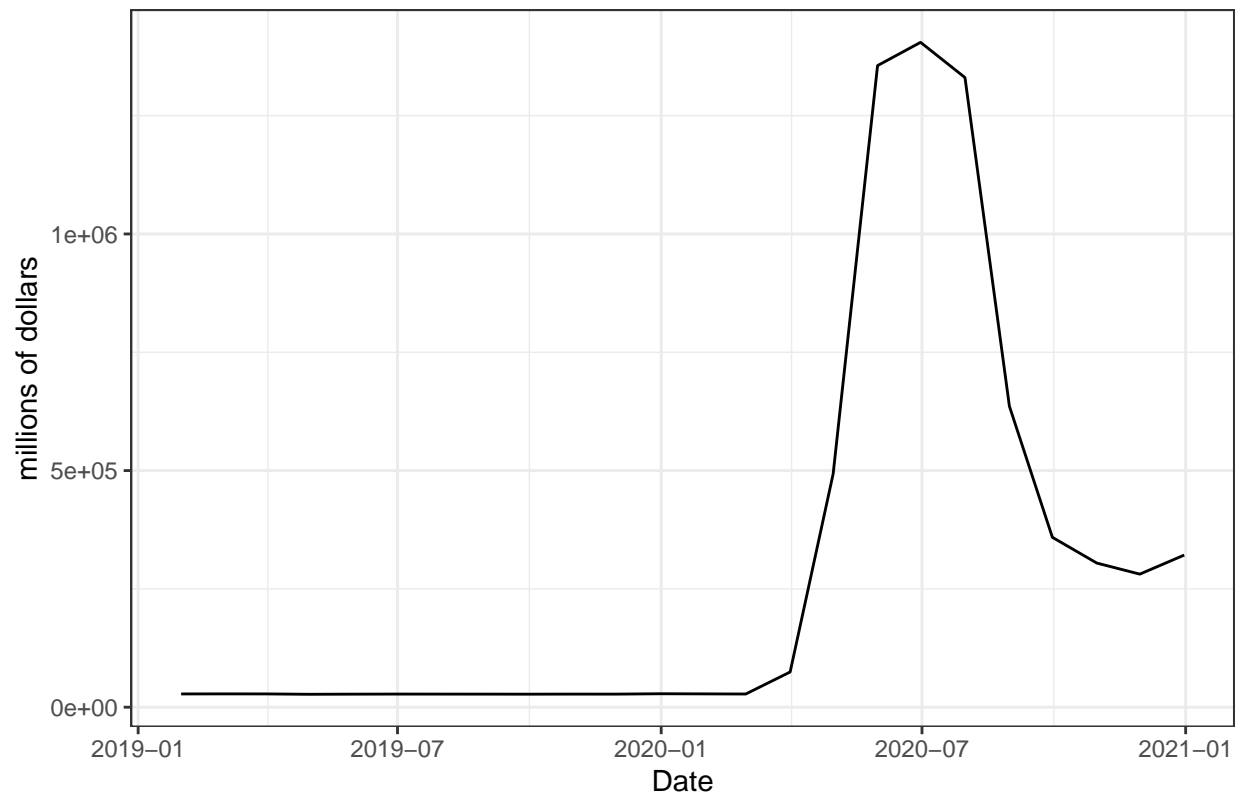
```
#Income decomposition (current dollars)
income %>% select(date, income_total, income_wages, income_interest,
                  income_govbenefits, income_others) %>%
  filter(date >= '2019-01-01') %>%
  melt(id='date') %>%
  ggplot(aes(x=date, y=value, colour = variable))+
  geom_line()+
  ggtitle("Personal income decomposition (current dollars, annualized)")+
  labs(y = "millions of dollars", x = "Date")+
  scale_colour_discrete(name = 'Indicator', labels = c('Total', 'Compensation of employees',
                                                       'Interest and dividends on assets',
                                                       'Transfer receipts (mainly gov social benefits',
                                                       'Others')))+
  theme_bw()
```

Personal income decomposition (current dollars, annualized)



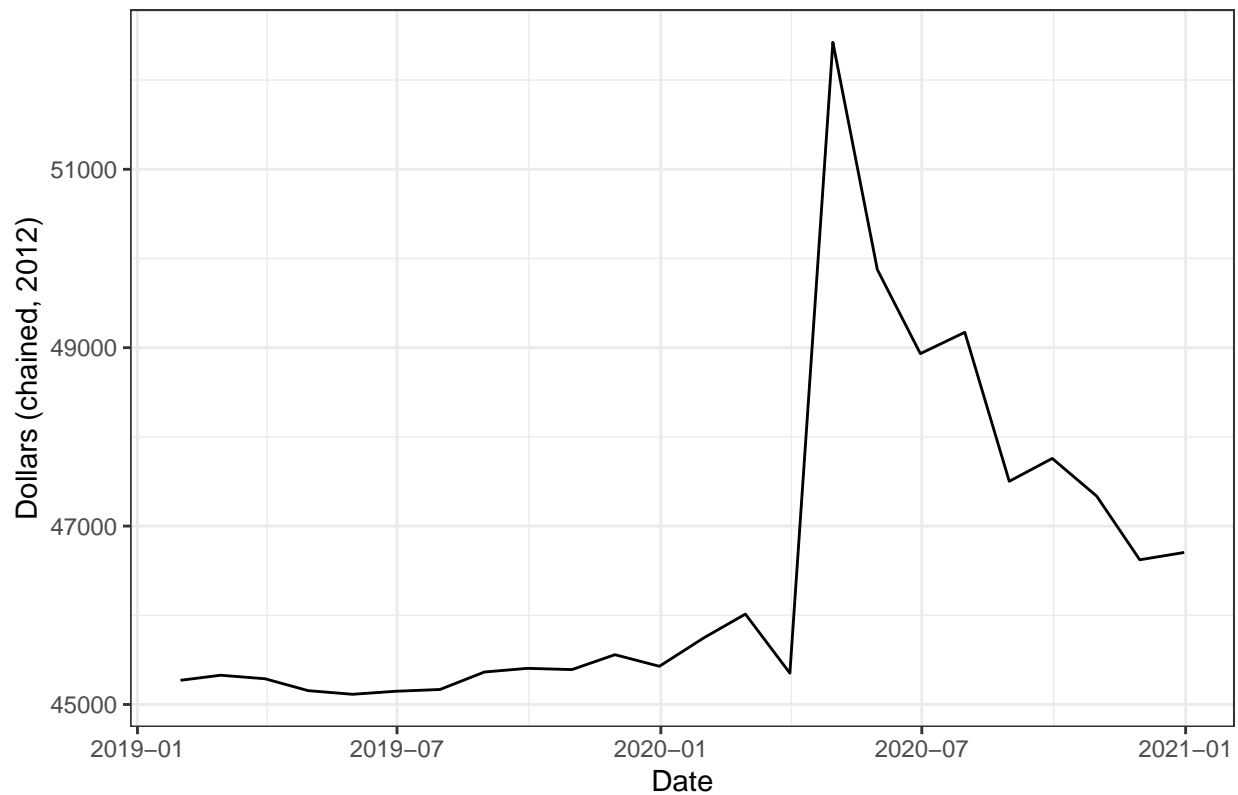
```
#Income from unemployment insurance benefits
income %>% filter(date >= '2019-01-01') %>%
  ggplot(aes(x = date, y= income_govbenefits_unemployment)) +
  geom_line() +
  ggtitle("Personal income from Unemployment Insurance benefits (current dollars, annualized)") +
  labs(y = "millions of dollars", x = "Date") +
  theme_bw()
```

Personal income from Unemployment Insurance benefits (current dollars,



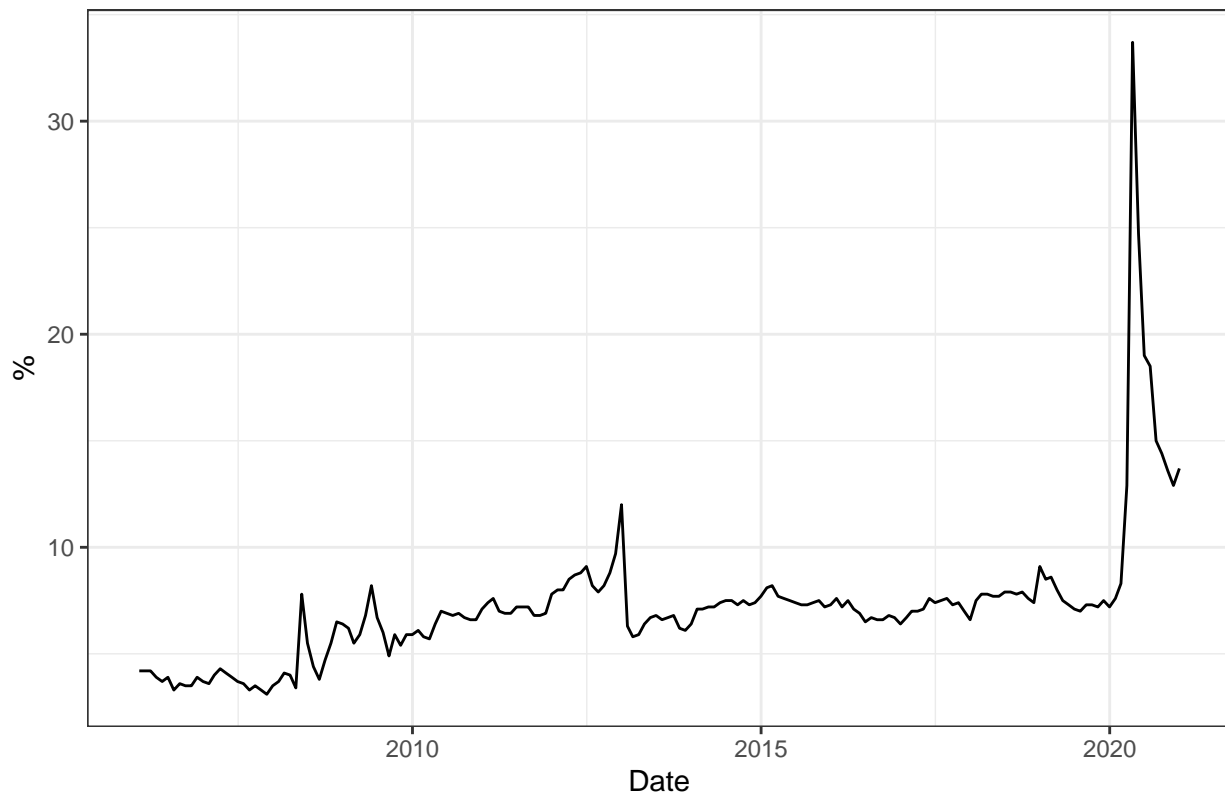
```
#Income per capita (chain dollars)
income %>% filter(date >= '2019-01-01') %>%
  ggplot(aes(x = date, y= income_per_capita)) +
  geom_line() +
  ggtitle("Personal income per capita (2012 chained dollars, annualized)") +
  labs(y = "Dollars (chained, 2012)", x = "Date") +
  theme_bw()
```


Personal income per capita (2012 chained dollars, annualized)



```
#Savings rate
income %>% filter(date >= '2006-01-01') %>%
  ggplot(aes(x = date, y= savings_rate)) +
  geom_line() +
  ggtitle("Savigs rate regarding the disposable income (Percentage)") +
  labs(y = "%", x = "Date") +
  theme_bw()
```

Savings rate regarding the disposable income (Percentage)



Some Findings:

- Counterintuitively, income increased significantly during March and April 2020, although looking at the decomposition of the income we observe that this increase is mainly justified by the social benefits by the government.
- As expected, we see a decrease in consumption (we will see more details in the next section).
- As a consequence of the prior 2 points, the savings rate had an important increase during the same period.
- In line with the other indicators that we have analyzed, the important effects in these were temporary, having the greatest impact during March and April 2020. However, it can be seen that the indicators have not been able to return to their preCOVID trend yet.

For Personal Consumption Expenditures

Link to data (using table U20406M)

```
#### PCE decomposition (chained 2012 dollars, annual rates), US ####

#Read data and select columns of interest

### Index (Line, Data)
## DPCERX. Personal Consumption expenditures
## DGDSRX. Goods
##   DDURRX. Durable Goods
##     DMOTRX. Motor vehicles
##     DFDHRX. Furnishing
##     DREQRX. Recreation Goods and Services
##   DNDGRX. Non-durable Goods
##     DFXARX. Food and beverages
##     DCLORX. Clothing and footwear
##     DONGRX. Other non durable
## DSERRX. Services
##   DHUTRX. Housing and utilities
##   DHLCRX. Health care
##   DTRSRX. Transportation
##   DRCARX. Recreation
##   DFSARX. Food services and accommodations
##   DIFSRX. Financial services and insurance
##   DOTSRX. Other services

pce_table <- read_excel("Data/U20406M Real PCE by Type of Product.xlsx", skip=7)[1:365,-c(1:2)]

## New names:
## * `` -> ...2
## * `` -> ...3

colnames(pce_table)[1] <- 'id'
pce_table <- as.data.frame(pce_table) %>% filter(id %in%
                                                    c('DPCERX','DGDSRX','DDURRX',
                                                      'DMOTRX','DFDHRX','DREQRX',
                                                      'DNDGRX','DFXARX','DCLORX','DONGRX',
                                                      'DSERRX','DHUTRX','DHLCRX',
                                                      'DTRSRX','DRCARX','DFSARX',
                                                      'DIFSRX','DOTSRX'))

#Create data table for pce
pce <- data.frame(date = seq(as.Date("2002-02-01"),length=12*(2020-2002+1),by="months")-1,
                  pce_total = as.numeric(filter(pce_table,id=='DPCERX')[-1]),
                  goods = as.numeric(filter(pce_table,id=='DGDSRX')[-1]),
                  goods_dur = as.numeric(filter(pce_table,id=='DDURRX')[-1]),
                  goods_dur_motor = as.numeric(filter(pce_table,id=='DMOTRX')[-1]),
                  goods_dur_fort = as.numeric(filter(pce_table,id=='DFDHRX')[-1]),
                  goods_dur_rec = as.numeric(filter(pce_table,id=='DREQRX')[-1]),
                  goods_ndur = as.numeric(filter(pce_table,id=='DNDGRX')[-1]),
                  goods_ndur_food = as.numeric(filter(pce_table,id=='DFXARX')[-1]),
                  goods_ndur_clothes = as.numeric(filter(pce_table,id=='DCLORX')[-1]),
                  goods_ndur_other = as.numeric(filter(pce_table,id=='DONGRX')[-1]),
```

```

serv = as.numeric(filter(pce_table,id=='DSERRX')[-1]),
serv_housing = as.numeric(filter(pce_table,id=='DHUTRX')[-1]),
serv_health = as.numeric(filter(pce_table,id=='DHLCRX')[-1]),
serv_transp = as.numeric(filter(pce_table,id=='DTRSXR')[-1]),
serv_recre = as.numeric(filter(pce_table,id=='DRCARX')[-1]),
serv_foodsnacco = as.numeric(filter(pce_table,id=='DFSARX')[-1]),
serv_financial = as.numeric(filter(pce_table,id=='DIFSRX')[-1]),
serv_others = as.numeric(filter(pce_table,id=='DOTSRX')[-1])
)

```

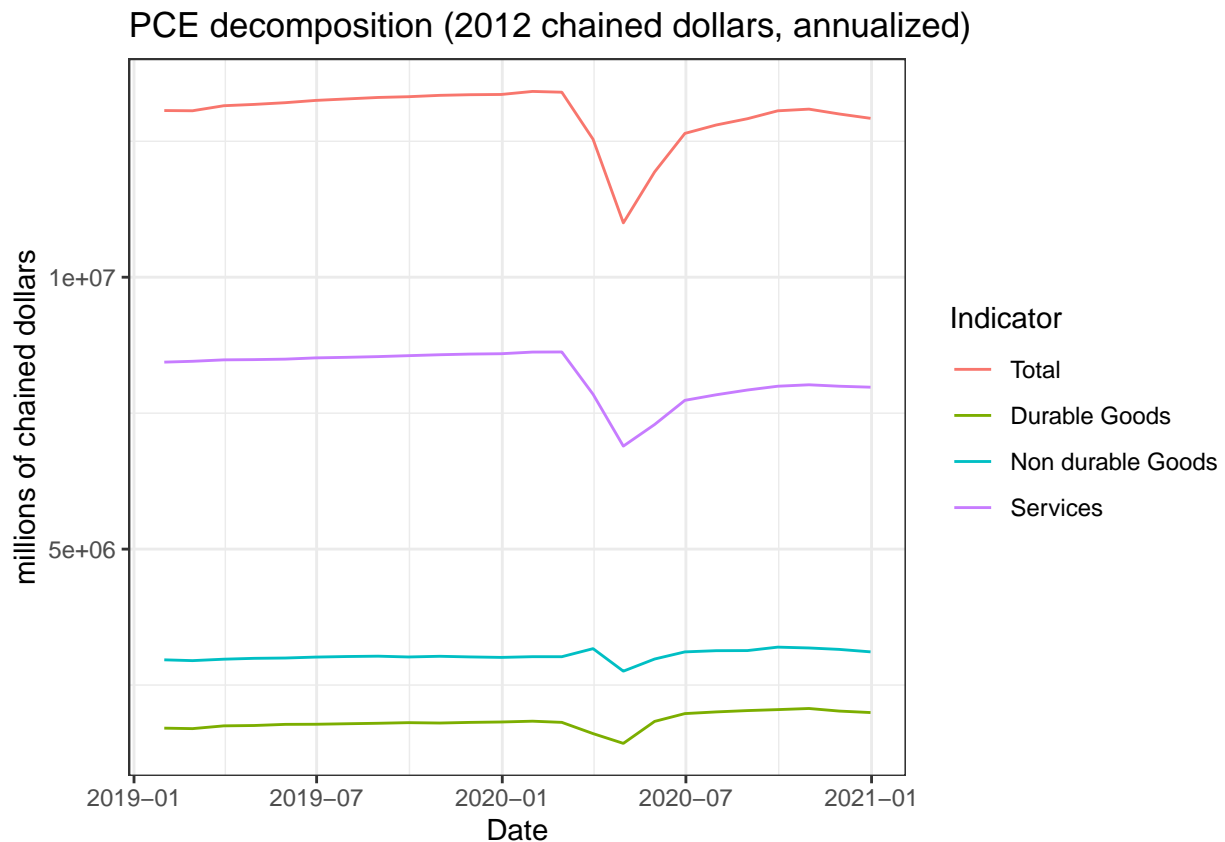
#Plots for PCE

#PCE (Total, Goods and Services)

```

pce %>% select(date, pce_total, goods_dur,goods_ndur, serv) %>%
  filter(date >= '2019-01-01') %>%
  melt(id='date') %>%
  ggplot(aes(x=date, y=value, colour = variable))+
  geom_line()+
  ggtitle("PCE decomposition (2012 chained dollars, annualized)")+
  labs(y = "millions of chained dollars", x = "Date")+
  scale_colour_discrete(name = 'Indicator', labels = c('Total', 'Durable Goods',
                                                    'Non durable Goods','Services'))+
  theme_bw()

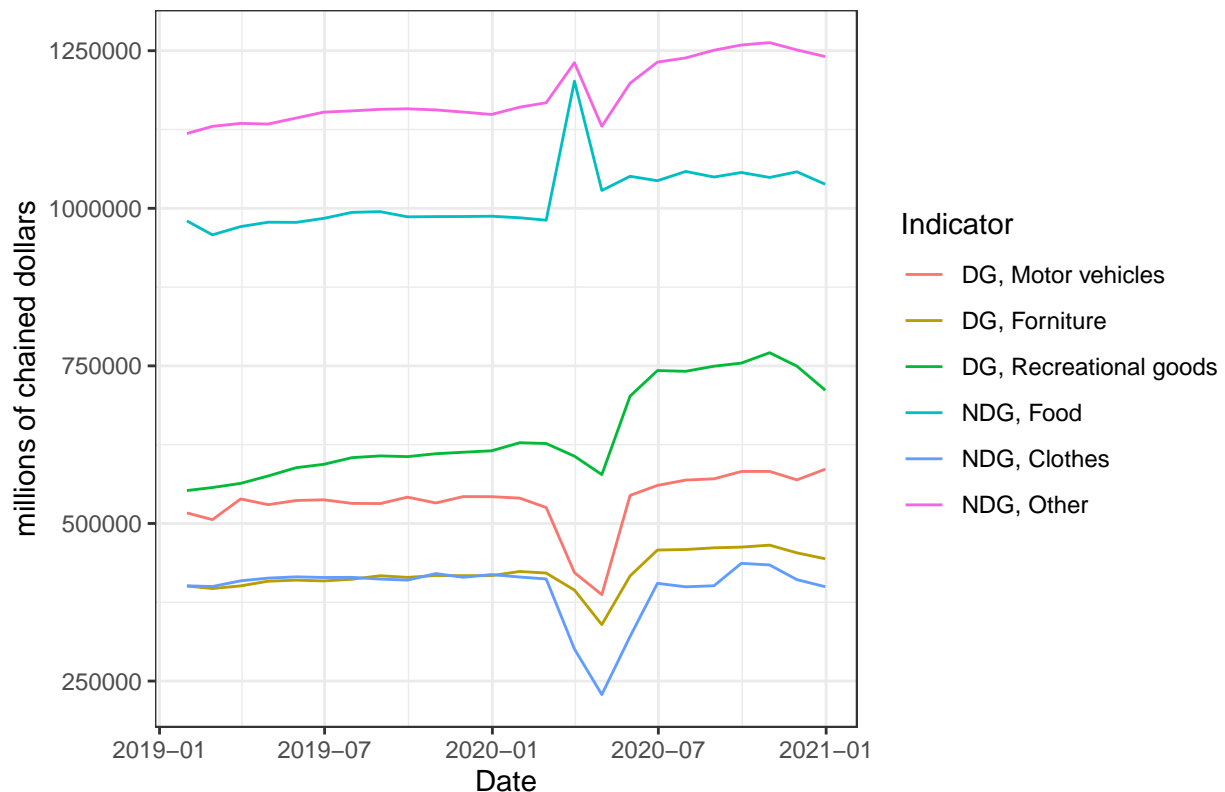
```



```
#PCE. Goods decomposition (More particular decomposition)
pce %>% select(date, goods_dur_motor, goods_dur_fort, goods_dur_rec, goods_ndur_food,
              goods_ndur_clothes, goods_ndur_other) %>%
  filter(date >= '2019-01-01') %>%
  melt(id='date') %>%
  ggplot(aes(x=date, y=value, colour = variable))+
  geom_line()+
  ggtitle("PCE Goods decomposition (2012 chained dollars, annualized)")+
  labs(y = "millions of chained dollars", x = "Date")+
  scale_colour_discrete(name = 'Indicator', labels = c('DG, Motor vehicles', 'DG, Furniture',
                                                    'DG, Recreational goods',
                                                    'NDG, Food', 'NDG, Clothes',
                                                    'NDG, Other'))+

theme_bw()
```

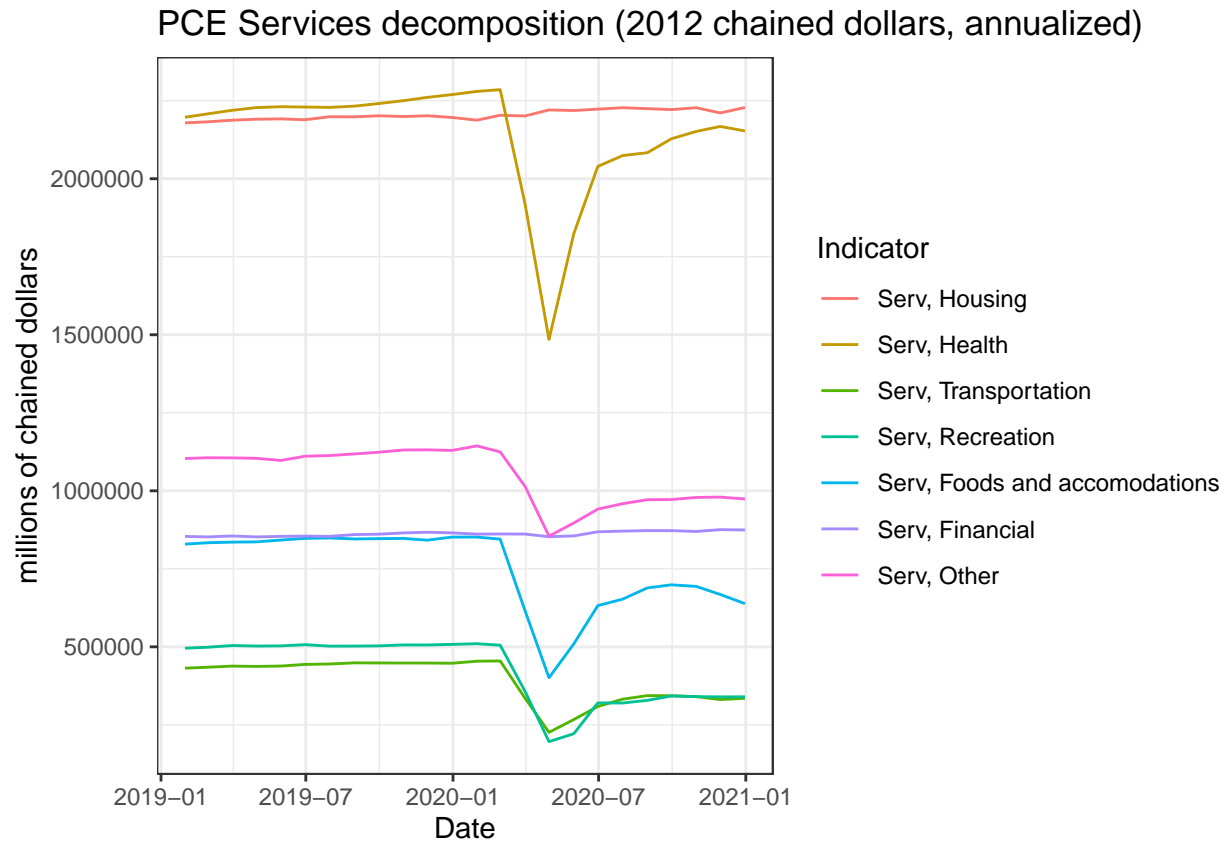
PCE Goods decomposition (2012 chained dollars, annualized)



```
#PCE. Services decomposition (More particular decomposition)
pce %>% select(date, serv_housing, serv_health, serv_transp, serv_recre,
              serv_foodsnacco, serv_financial, serv_others) %>%
  filter(date >= '2019-01-01') %>%
  melt(id='date') %>%
  ggplot(aes(x=date, y=value, colour = variable))+
  geom_line()+
  ggtitle("PCE Services decomposition (2012 chained dollars, annualized)")+
  labs(y = "millions of chained dollars", x = "Date")+
  scale_colour_discrete(name = 'Indicator', labels = c('Serv, Housing', 'Serv, Health',
                                                    'Serv, Transportation',
```

```
theme_bw()
```

```
'Serv, Recreation',  
'Serv, Foods and accomodations',  
'Serv, Financial',  
'Serv, Other'))+
```



Some Findings:

- Looking further in the consumption data, we can observe that the main contributor to its decrease is the Services category
- For Goods category, we see a decrease en almost all categories, the exception being food and beverages.

For Jobless claims

Link to data

```
#read data
jobless_us <- read.csv('Data/Jobless_claims_nat.csv', header = T, na.strings = "",
                      skip=2, stringsAsFactors = F) %>%
  'colnames<-'(c('Date', 'I_NSA', 'I_SF', 'I_SA', 'I_SA4W',
                 'C_NSA', 'C_SF', 'C_SA', 'C_SA4W',
                 'IUR_NSA', 'IUR_SA', 'COVEMP')) %>% select(Date, I_SA, C_SA)

jobless_cal <- read.csv('Data/Jobless_claims_cal.csv', header = T, na.strings = "",
                       skip=4, stringsAsFactors = F) %>%
  'colnames<-'(c('State', 'Date', 'Init', 'RWE', 'Cont', 'CovEmp',
                 'IUR')) %>% select(Date, Init, Cont)

#clean
jobless_us$Date <- as.Date(jobless_us$Date, format='%m/%d/%Y')
jobless_us$I_SA <- as.numeric(gsub(",", "", jobless_us$I_SA))
jobless_us$C_SA<- as.numeric(gsub(",", "", jobless_us$C_SA))

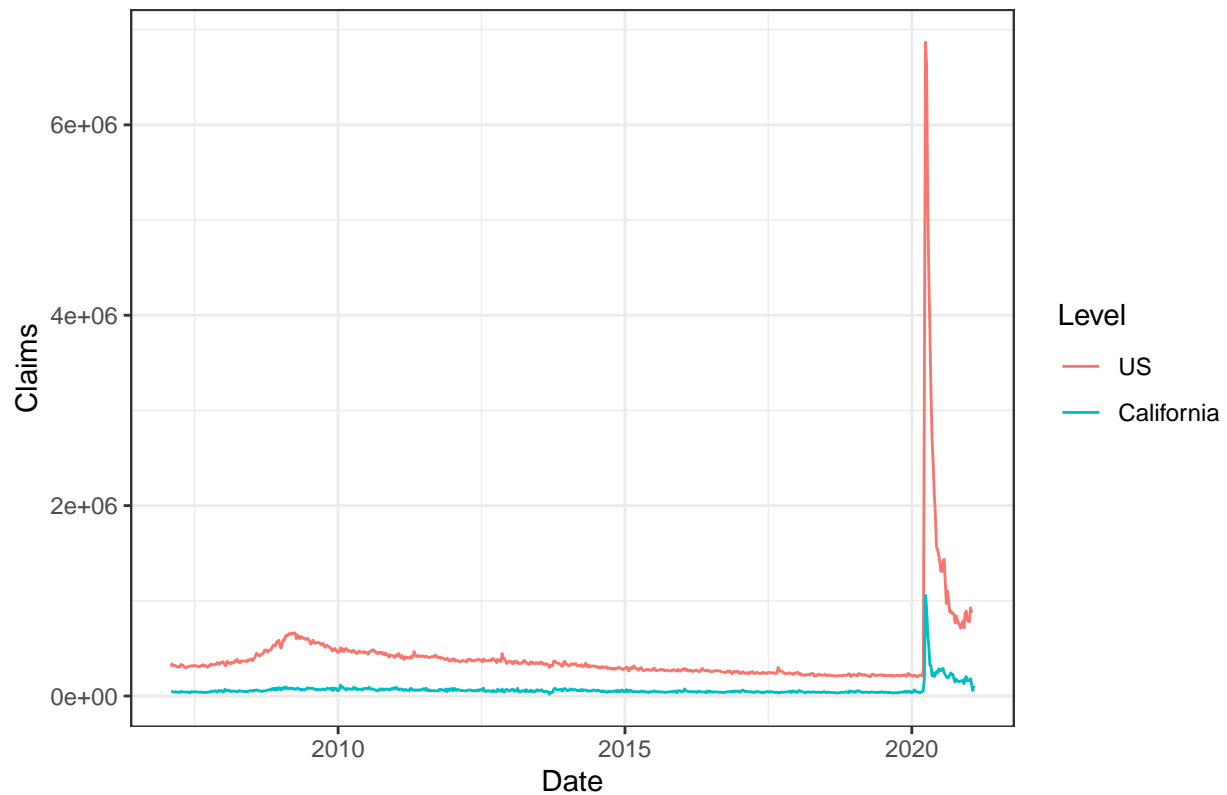
jobless_cal$Date <- as.Date(jobless_cal$Date, format='%m/%d/%Y')
jobless_cal$Init <- as.numeric(gsub(",", "", jobless_cal$Init))
jobless_cal$Cont<- as.numeric(gsub(",", "", jobless_cal$Cont))

#Merge
jobless <- merge(jobless_us, jobless_cal, by='Date') %>%
  'colnames<-'(c('Date', 'init_US', 'cont_US', 'init_CA', 'cont_CA'))

#plot Initial Jobless Claims (US and California)
jobless %>% select(Date, init_US, init_CA) %>% filter(Date>='2007-01-31') %>%
  melt(id='Date') %>%
  ggplot(aes(x = Date, y=value, colour=variable)) +
  geom_line()+
  ggtitle("Initial Jobless Claims (US and California)") +
  labs(y='Claims', x = "Date")+
  scale_colour_discrete(name = 'Level', labels = c('US', 'California'))+
  theme_bw()
```

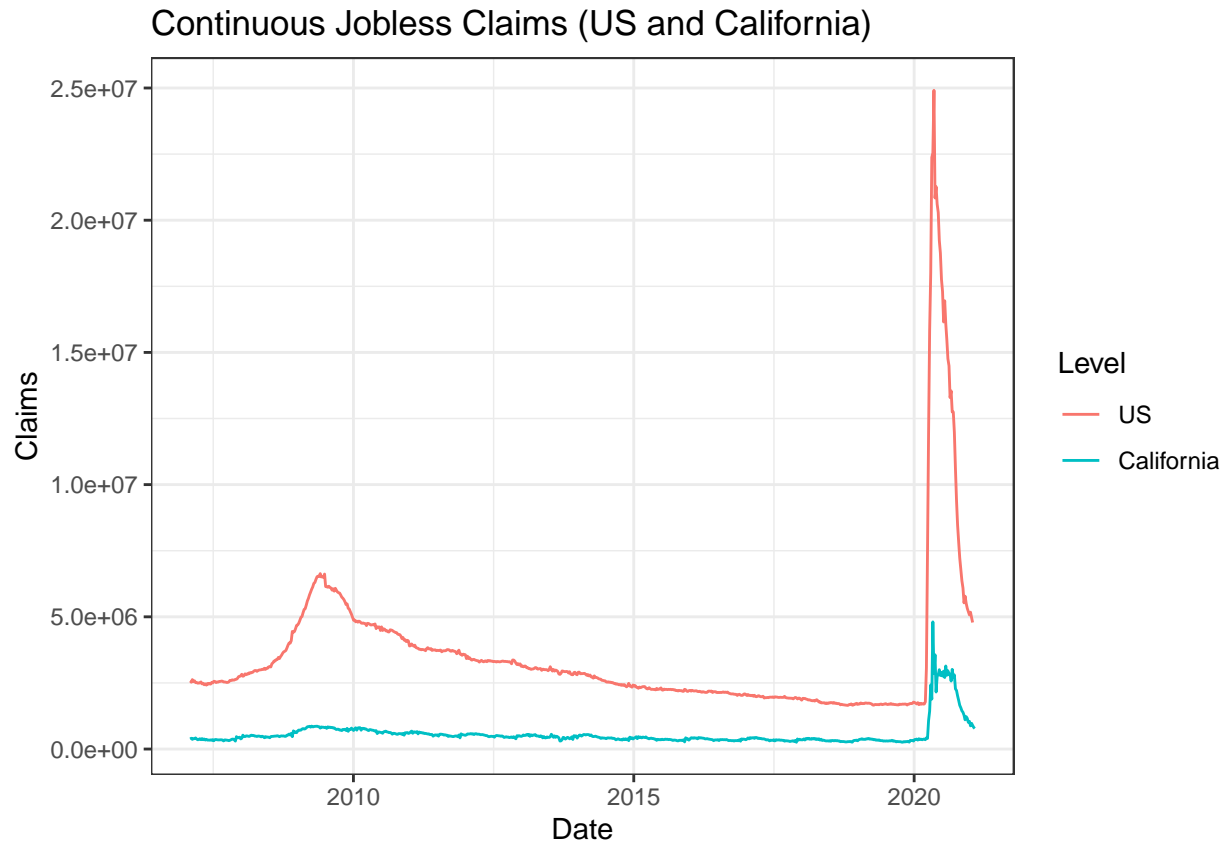
```
## Warning: Removed 2 row(s) containing missing values (geom_path).
```

Initial Jobless Claims (US and California)



```
#plot Continuous Jobless Claims (US and California)
jobless %>% select(Date, cont_US, cont_CA) %>% filter(Date>='2007-01-31') %>%
  melt(id='Date') %>%
  ggplot(aes(x = Date, y=value, colour=variable)) +
  geom_line()+
  ggtitle("Continuous Jobless Claims (US and California)") +
  labs(y='Claims', x = "Date") +
  scale_colour_discrete(name = 'Level', labels = c('US', 'California')) +
  theme_bw()
```

```
## Warning: Removed 2 row(s) containing missing values (geom_path).
```

Some Findings:

- Claims for unemployment insurance benefits had an increase not seen in this century, not even in the crisis of 2008
- However, in line with the other indicators, this change was quickly reversed, although it remains at high levels compared to before the quarantine measures.
- The comparison with what happened in the 2008 crisis, whose effects were more sustained for a longer time, is interesting.
- Some bias might be present in the continuous unemployment claims graph due to extensions in people receiving benefits