State Unemployment vs COVID

Robi Rahman

2021-04-19

Unemployment Data by State, January 2020 - March 2021

Importing and cleaning data on state unemployment rates during the coronavirus pandemic:

```r
bls_unemployment_data <- read.csv("Unemployment by state 2020-2021.csv")[,c(1,2,3,4,11)]

# List of state names:
state_names <- bls_unemployment_data$State[1:53]

# Initialize empty matrix
unemployment_data <- data.frame(matrix(ncol=55,nrow=15))

# Label matrix with state names
colnames(unemployment_data) <- c("Year", "Month", state_names)

# Initialize dates
unemployment_data$Year <- c(rep(2020,12),rep(2021,3))
unemployment_data$Month <- c(1:12, 1:3)


# Fill in matrix with data from the BLS file
for (m in 1:15) {
  for (s in 1:53) {
    unemployment_data[m,s+2]<-bls_unemployment_data[s+53*(m-1),5]
  }
}

# Save the cleaned data
write.csv(unemployment_data,"unemployment data.csv")
```

COVID Cases by State

Importing and cleaning data on state COVID case and death counts:

```r
# Define helper functions to extract the year, month, and day of a COVID report:
year_num <- function(x) {
  as.numeric(substr(x,1,4))
}
month_num <- function(x) {
  as.numeric(substr(x,6,7))
}
day_num <- function(x) {
  as.numeric(substr(x,9,10))
}
```

```r
# Read dataset of daily COVID cases and deaths for each state
covid_by_state <- read.csv("COVID cases by state.csv")[,c(1,2,3,6,8,11)]

# Parse date strings from the file
covid_by_state$submission_date <- as.Date(covid_by_state$submission_date, format = "%m/%d/%Y")

# Get the year, month, and day using the helper functions:
covid_by_state$year <- year_num(covid_by_state$submission_date)
covid_by_state$month <- month_num(covid_by_state$submission_date)
covid_by_state$day <- day_num(covid_by_state$submission_date)

# Sort dataframe by location and date
attach(covid_by_state)
covid_by_state <- covid_by_state[order(year, month, day, state),]
detach(covid_by_state)
```

Now we need to group and sum the daily cases in each state by month:

```r
covid_data <- covid_by_state %>%
  group_by(state, year, month) %>%
  summarize(cases=sum(new_case),deaths=sum(new_death))
```

`summarise()` has grouped output by 'state', 'year'. You can override using the `.groups` argument.

We now have COVID data by state and month, but need to organize it.

```r
# Copy the monthly unemployment data table:
cases_data <- unemployment_data[,1:2]
cases_data[16,] <- c(2021,4)
deaths_data <- unemployment_data[,1:2]
deaths_data[16,] <- c(2021,4)

# Fill in the monthly COVID case and deaths dataframes:
for (s in 1:60) {
  cases_data[s+2] <- rep(0,16)
  colnames(cases_data)[s+2] <- covid_data$state[1+16*(s-1)]
  deaths_data[s+2] <- rep(0,16)
  colnames(deaths_data)[s+2] <- covid_data$state[1+16*(s-1)]
  for (m in 1:16) {
    cases_data[m,s+2] <- covid_data[1+16*(s-1)+m,4]
    deaths_data[m,s+2] <- covid_data[1+16*(s-1)+m,5]
  }
}
```

Remove American Samoa, Federated States of Micronesia, Guam, MP, PW, RMI, Virgin Islands, and Los Angeles County. (I'm retaining New York City because it is present in both the unemployment and COVID cases/deaths datasets.)

```r
cases_data <- cases_data[c(-16),c(-6,-14,-16,-31,-47,-48,-50,-57)]
deaths_data <- deaths_data[c(-16),c(-6,-14,-16,-31,-47,-48,-50,-57)]
unemployment_data <- unemployment_data[,c(-8)]
```

Sorting the unemployment data by abbreviation to match the case data:

```r
colnames(unemployment_data) <- c(
  "Year","Month","AL","AK","AZ","AR","CA","CO","CT","DE","DC","FL","GA","HI",
  "ID","IL","IN","IA","KS","KY","LA","ME","MD","MA","MI","MN","MS","MO","MT",
```

```
  "NE","NV","NH","NJ","NM","NY","NYC","NC","ND","OH","OK","OR","PA","RI","SC",
  "SD","TN","TX","UT","VT","VA","WA","WV","WI","WY"
)
unemployment_data <- unemployment_data[,
                                       c(1,2,4,3,6,5,7:9,11,10,12:14,18,15:17,
                                         19:21,24,23,22,25:26,28,27,29,37:38,
                                         30,32,33:34,31,35,36,39:48,
                                         50,49,51,53,52,54)
                                      ]
```

We can subtract the previous month's unemployment rate from the next month's unemployment rate to find the monthly change in unemployment:

```
unemployment_changes <- unemployment_data
unemployment_changes[1:14,3:54] <- unemployment_changes[2:15,3:54] - unemployment_changes[1:14,3:54]
unemployment_changes <- unemployment_changes[c(-15),]
```

Now let's find the correlation between COVID cases and deaths, and state unemployment rates.

```
for (i in 1:52) {
  cases <- cases_data[1:14,i+2]
  deaths <- deaths_data[1:14,i+2]
  unemp_chg <- unemployment_changes[,i+2]
  cases_vs_unemp <- lm(unemp_chg ~ cases)
  summary(cases_vs_unemp)
  deaths_vs_unemp <- lm(unemp_chg ~ deaths)
  summary(deaths_vs_unemp)
  plot(cases,unemp_chg, main=paste(colnames(unemployment_changes)[i+2],"unemployment change vs cases"))
  abline(cases_vs_unemp$coefficients[[1]],cases_vs_unemp$coefficients[[2]])
  plot(deaths, unemp_chg, main=paste(colnames(unemployment_changes)[i+2],"unemployment change vs deaths
  abline(deaths_vs_unemp$coefficients[[1]],deaths_vs_unemp$coefficients[[2]])
}
```
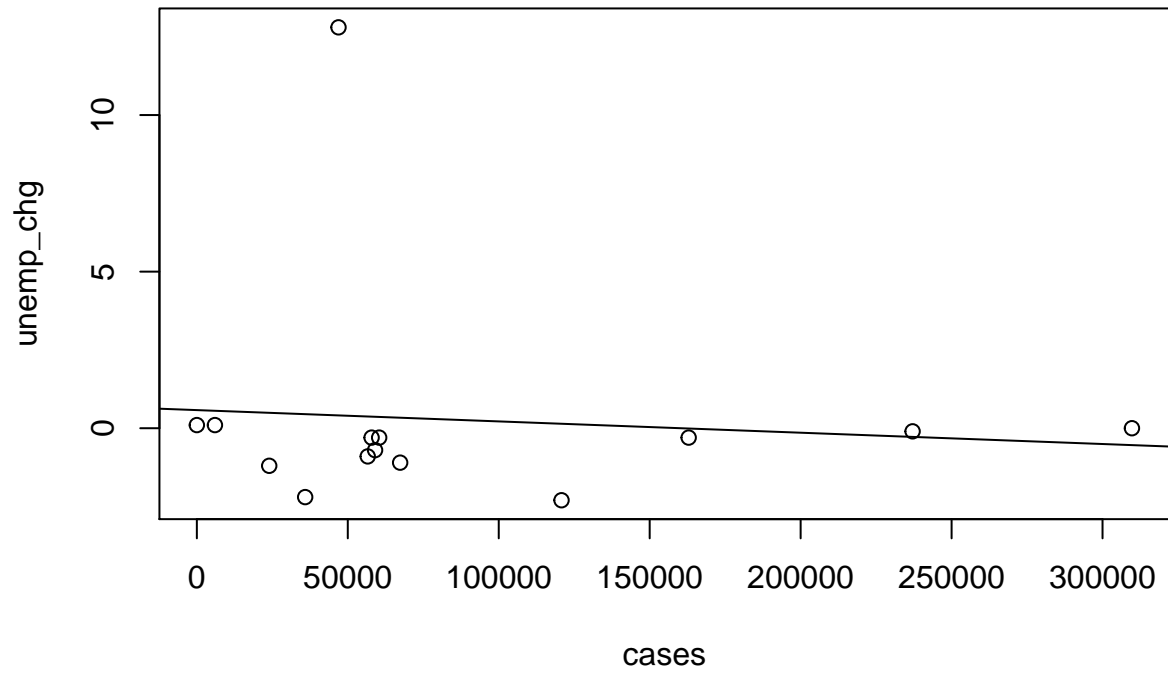
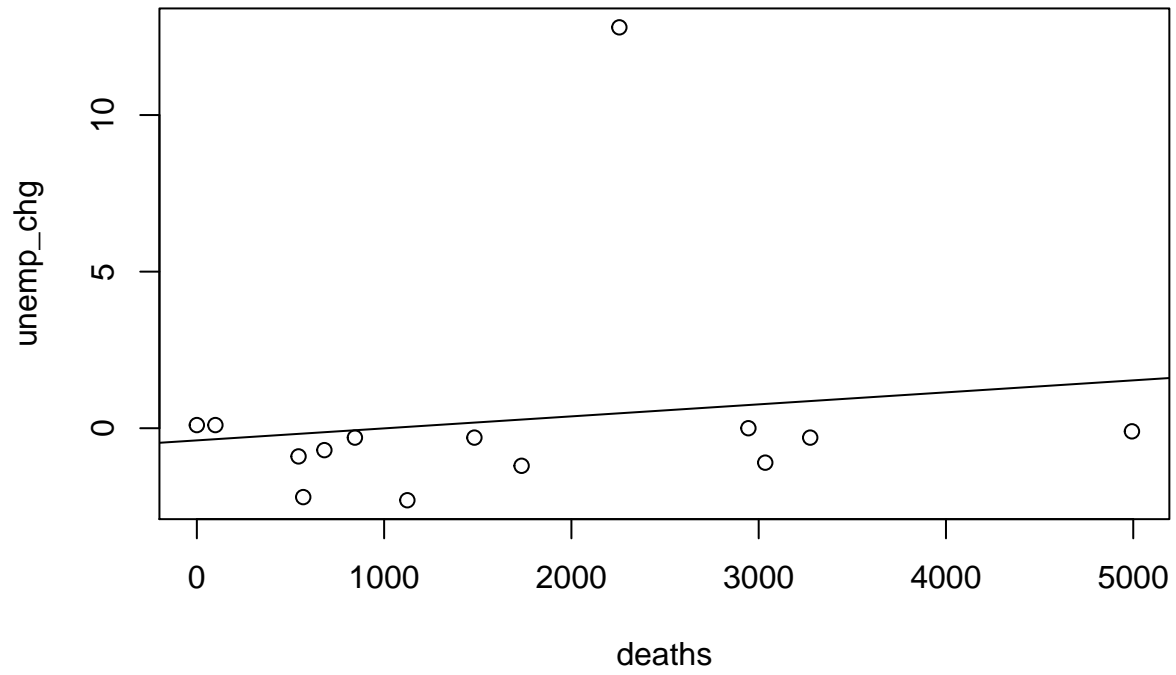**AK unemployment change vs cases**
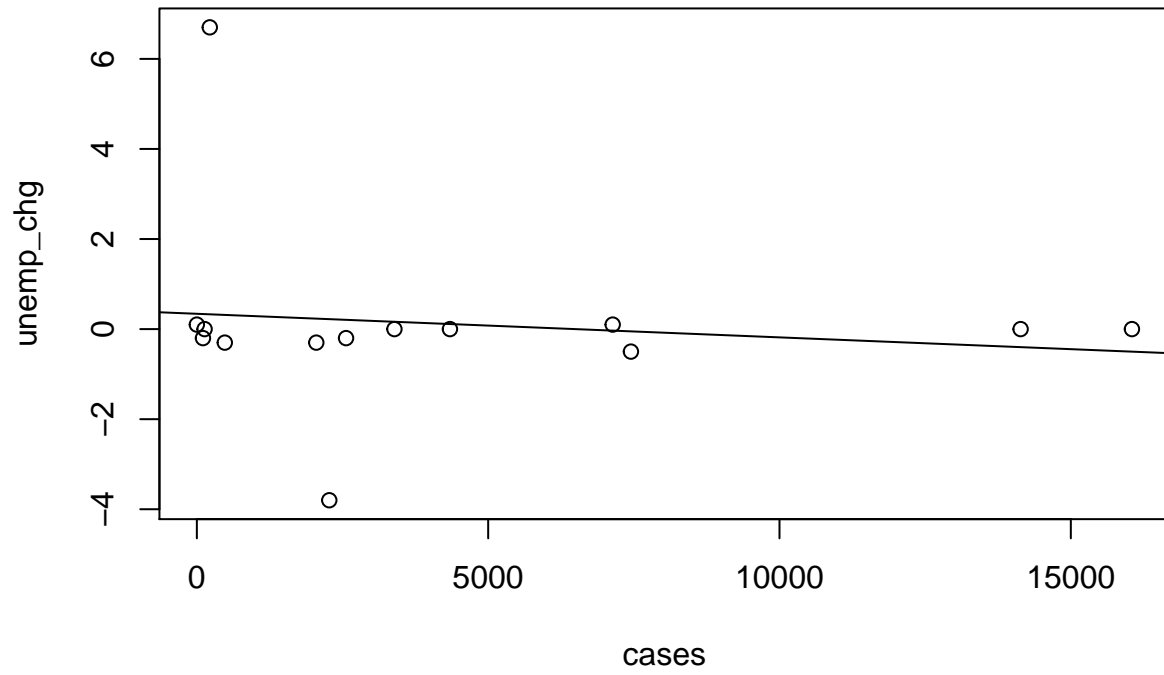
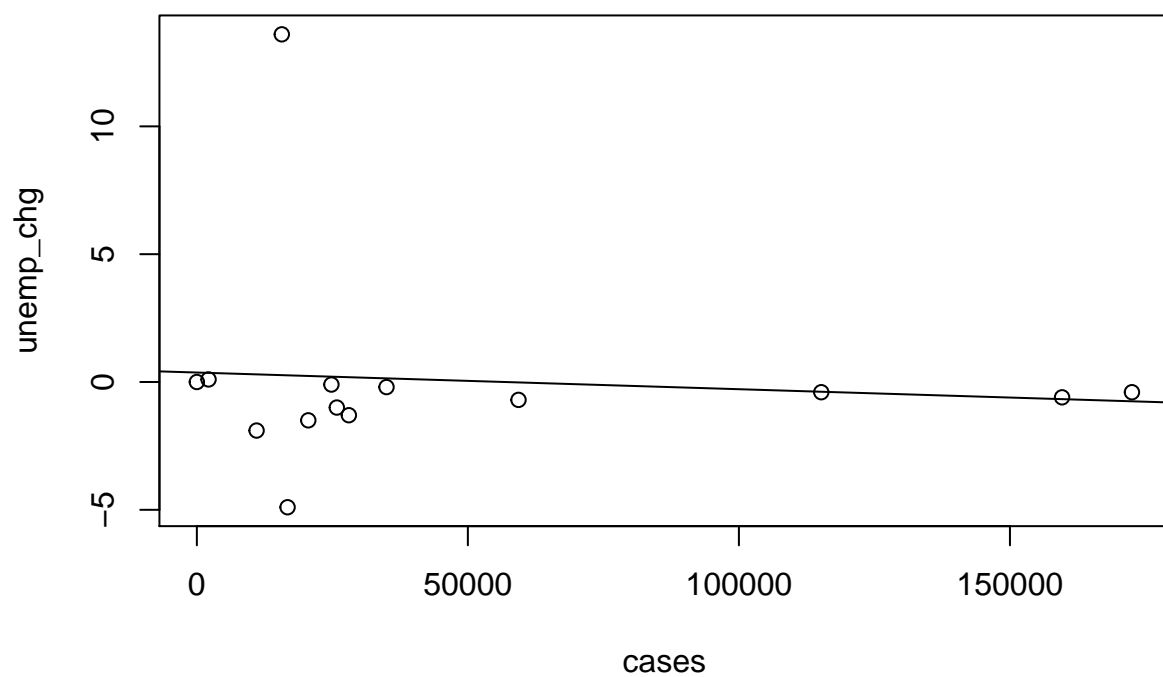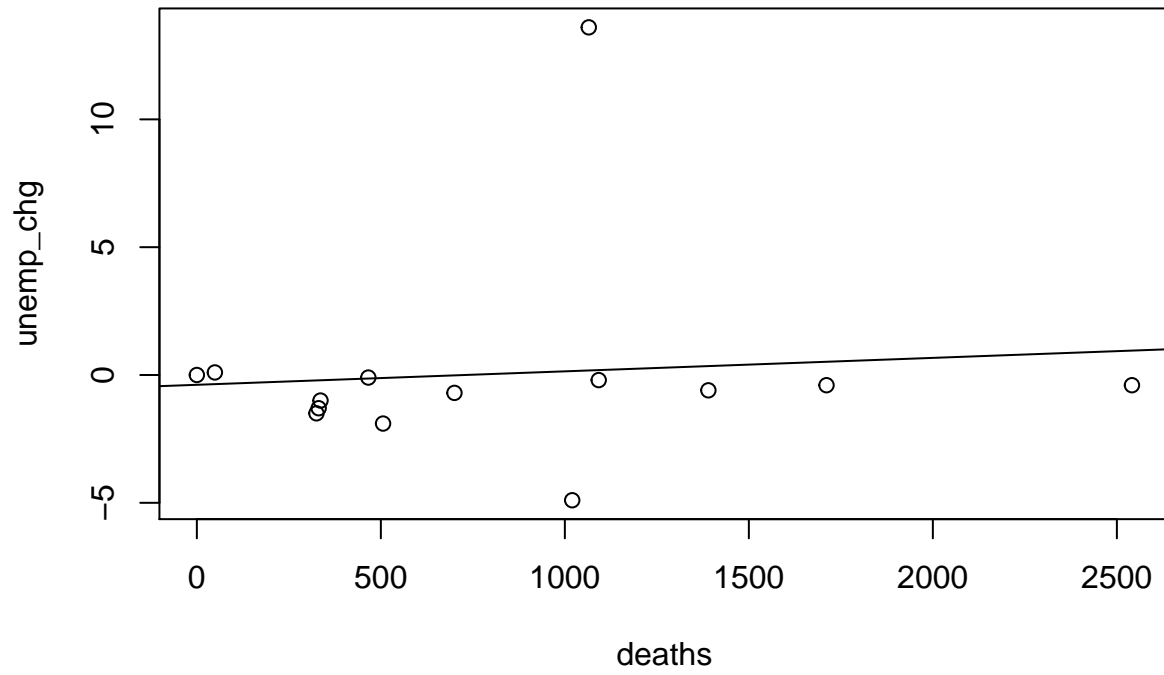# AK unemployment change vs deaths
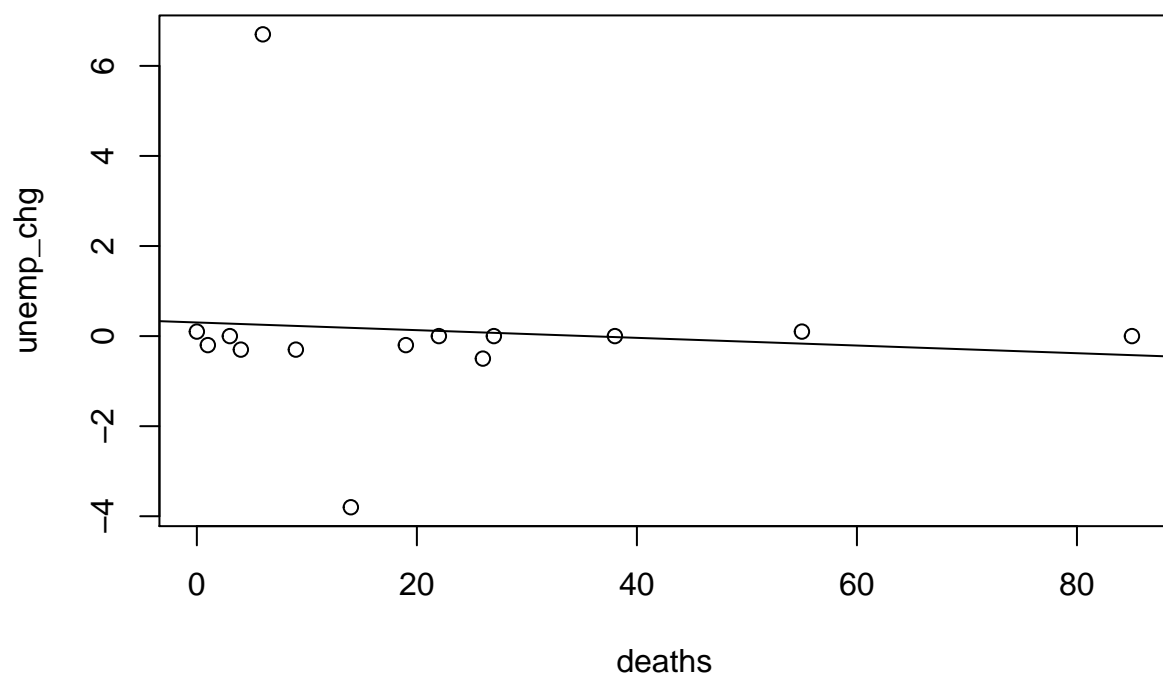
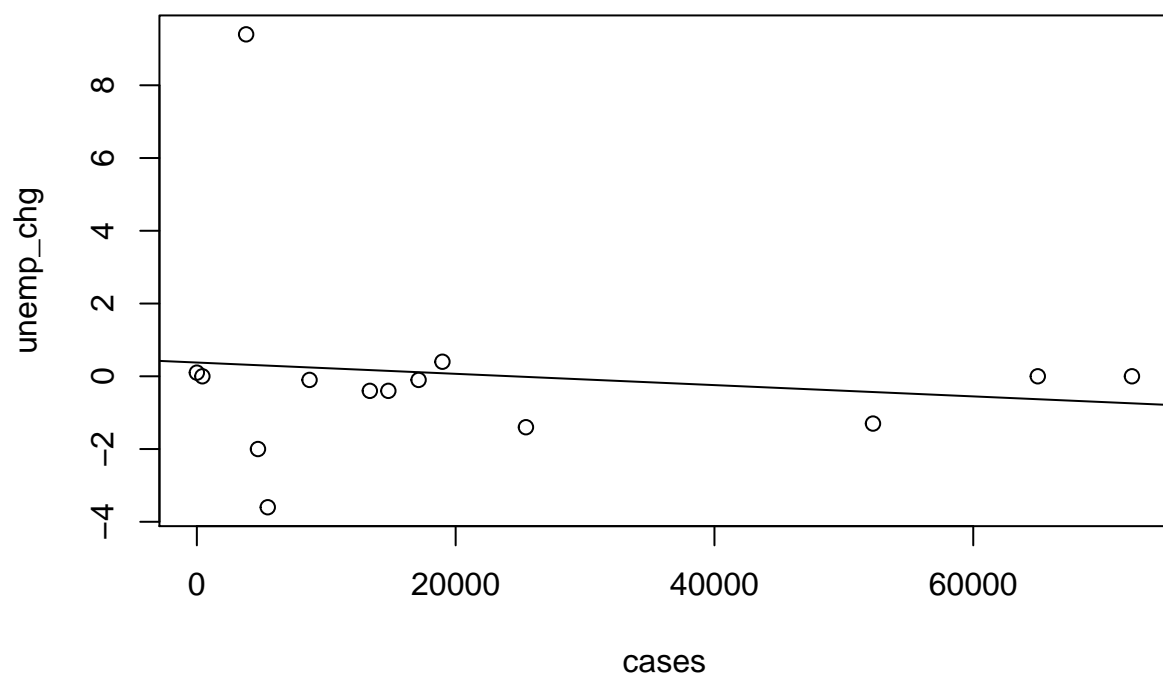# AL unemployment change vs cases
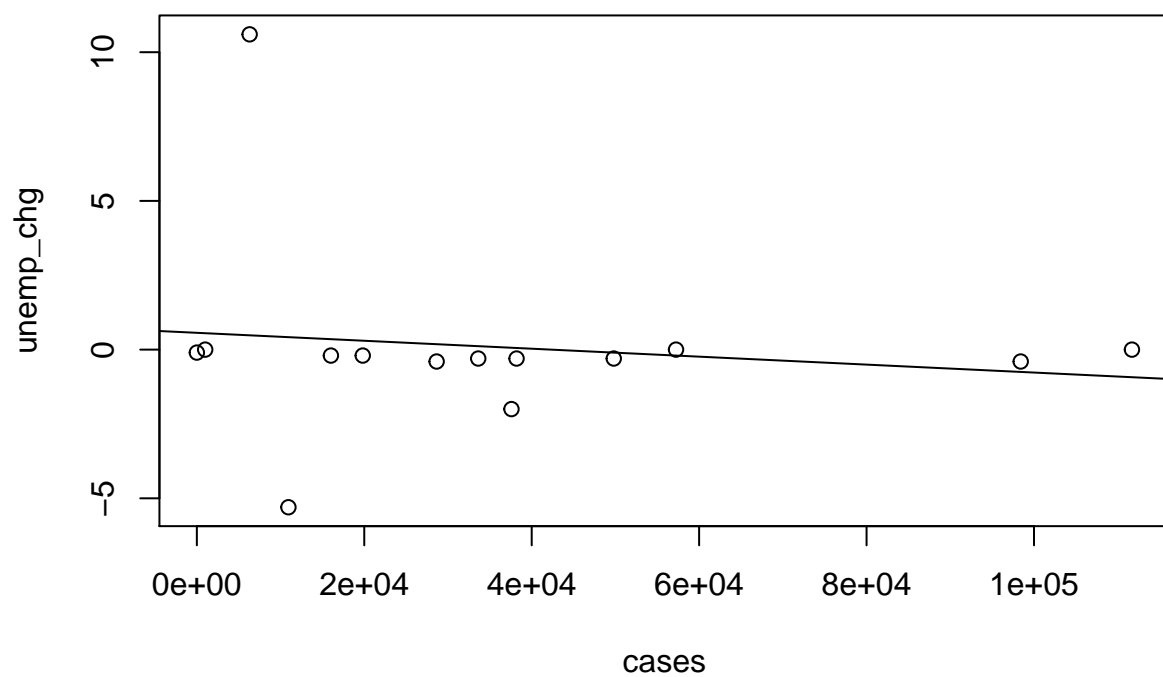
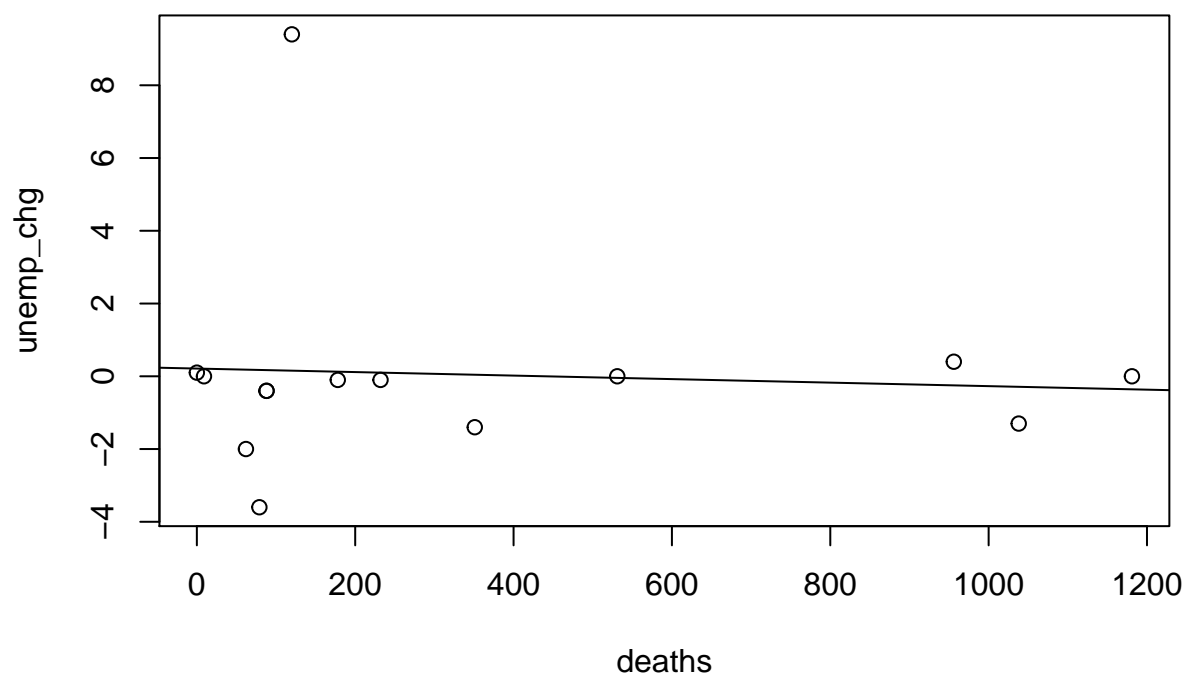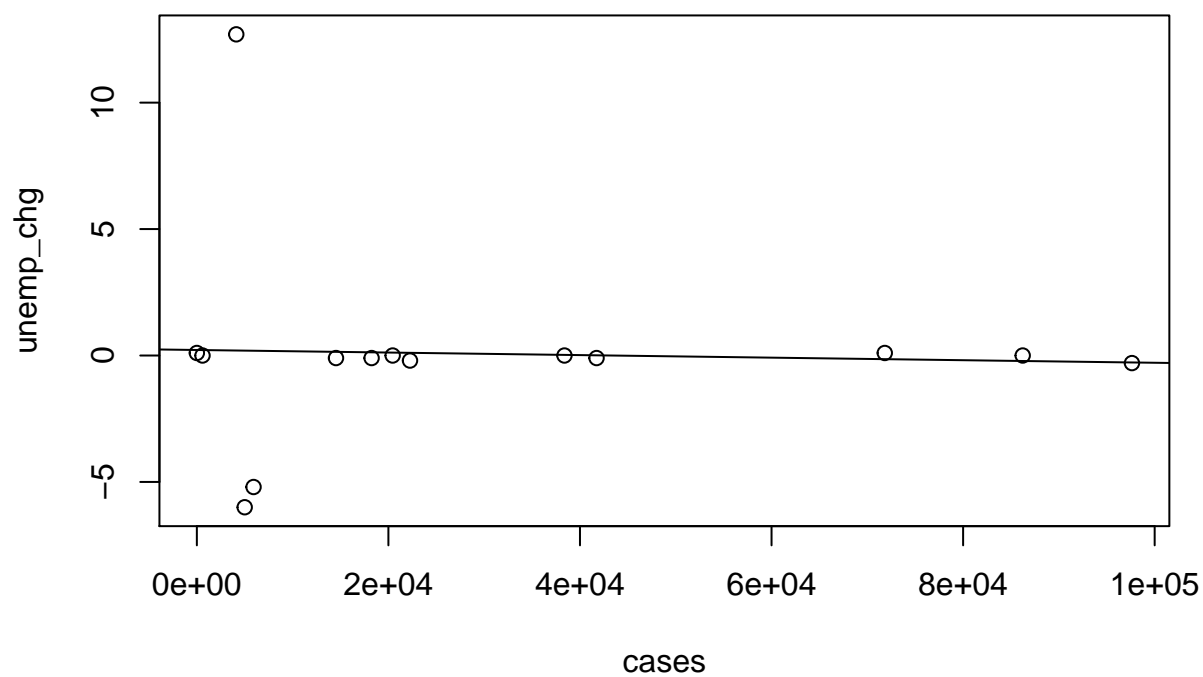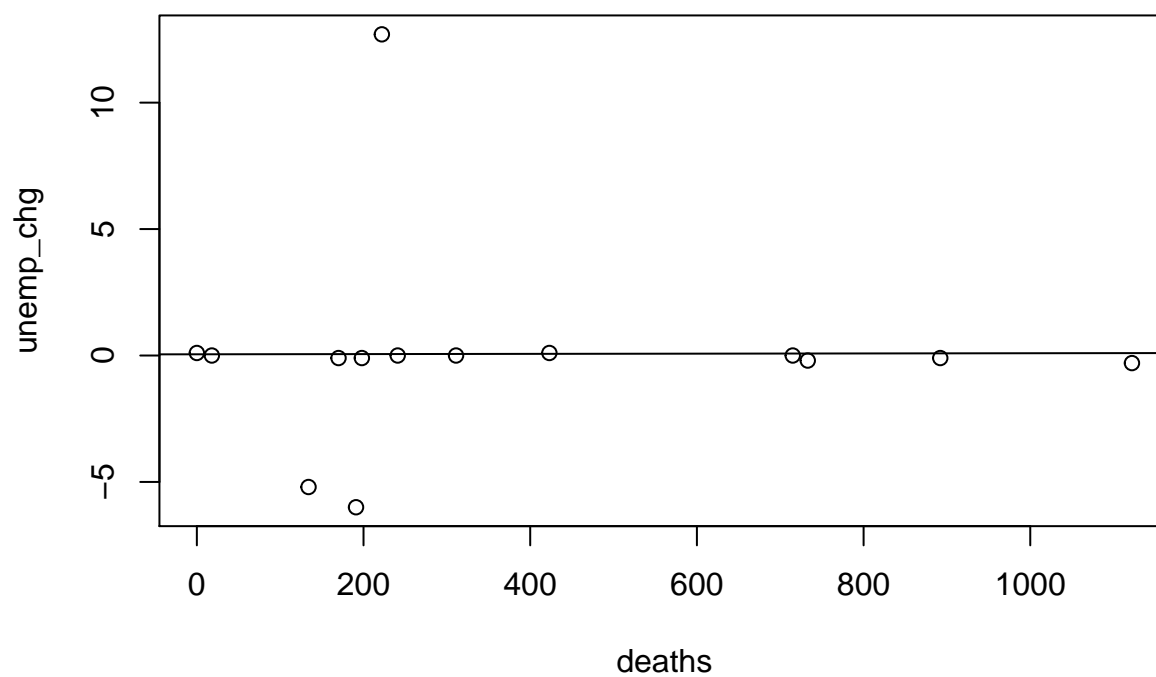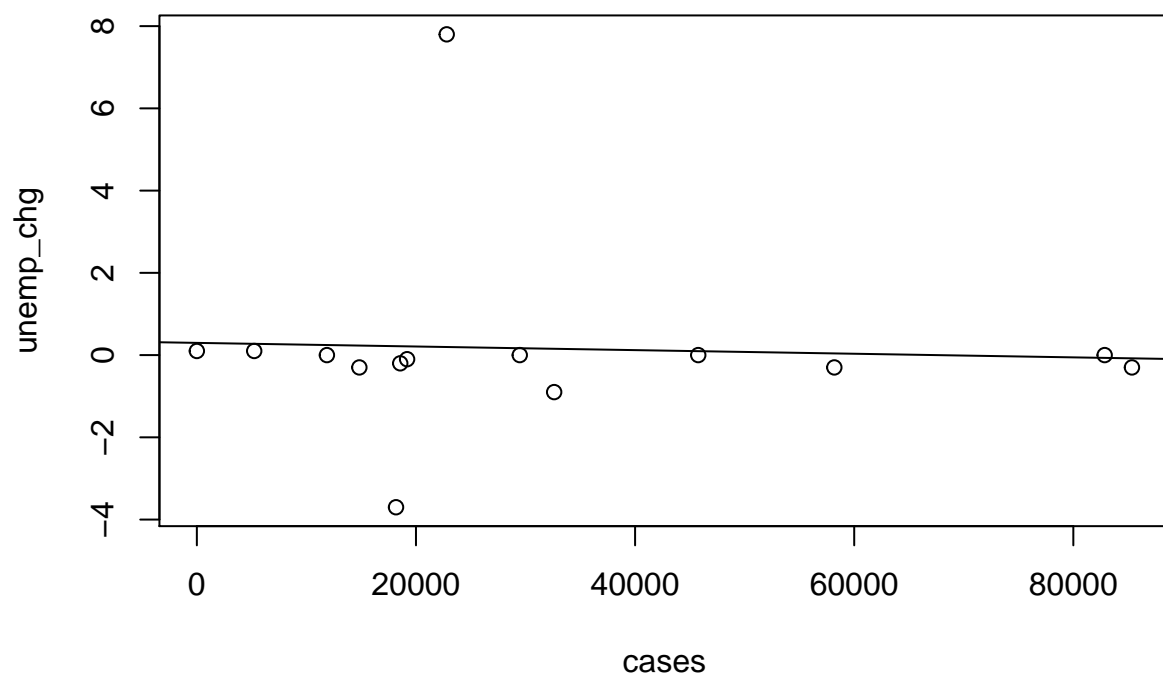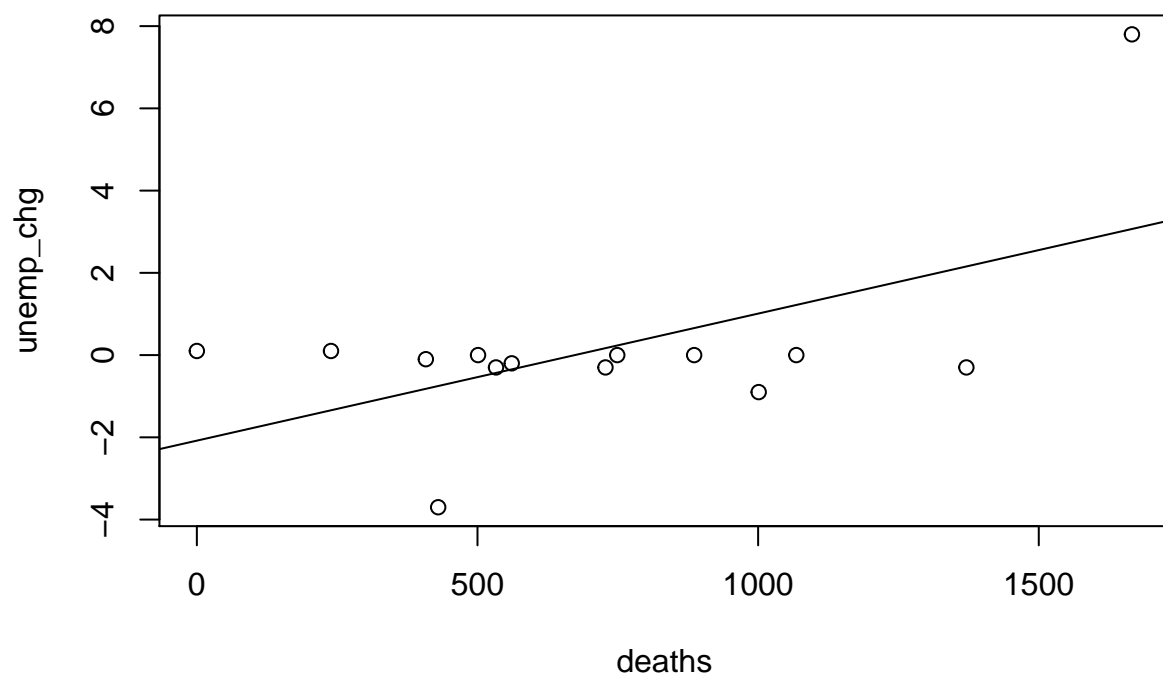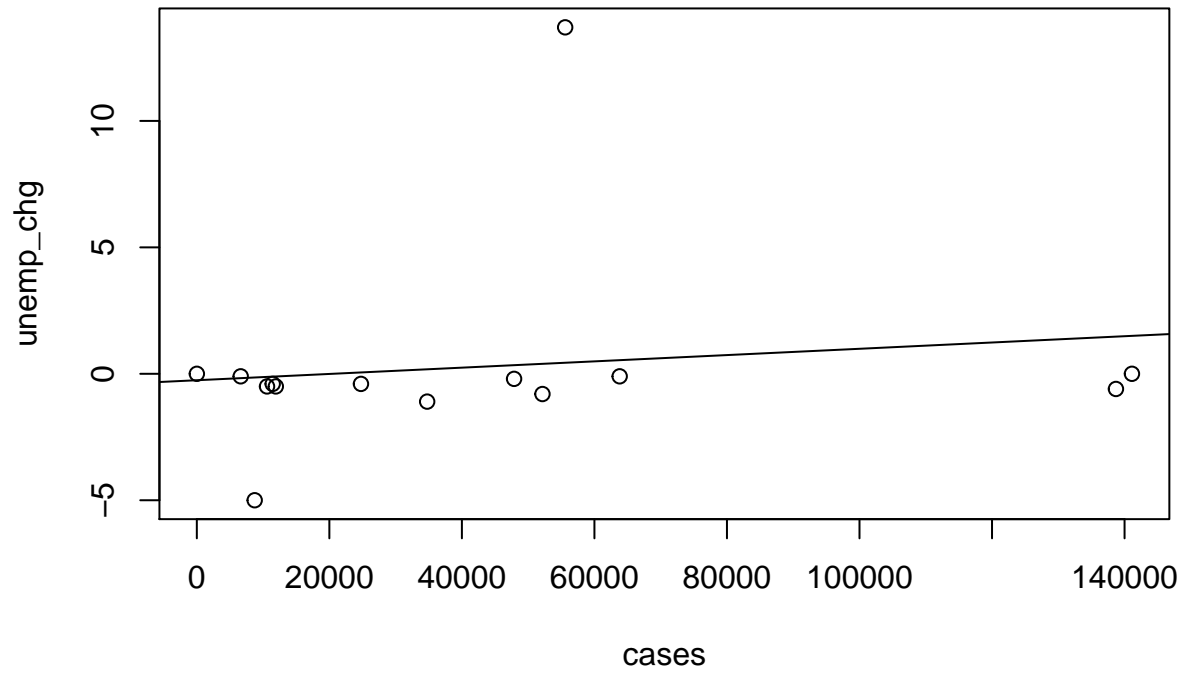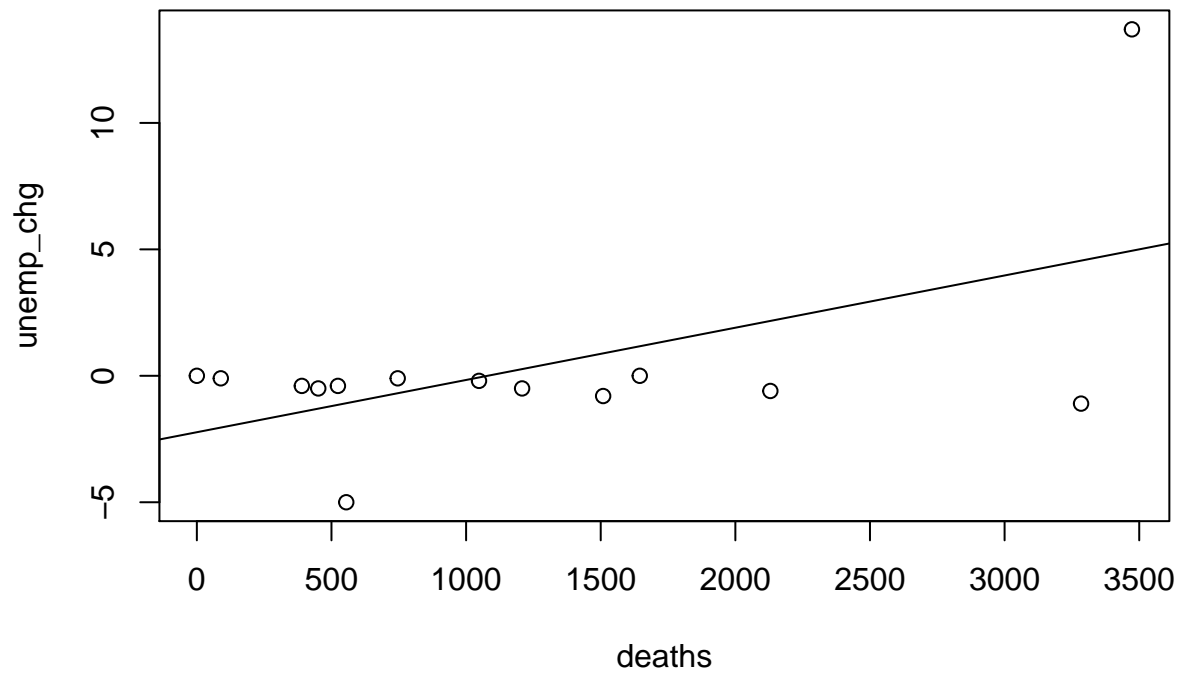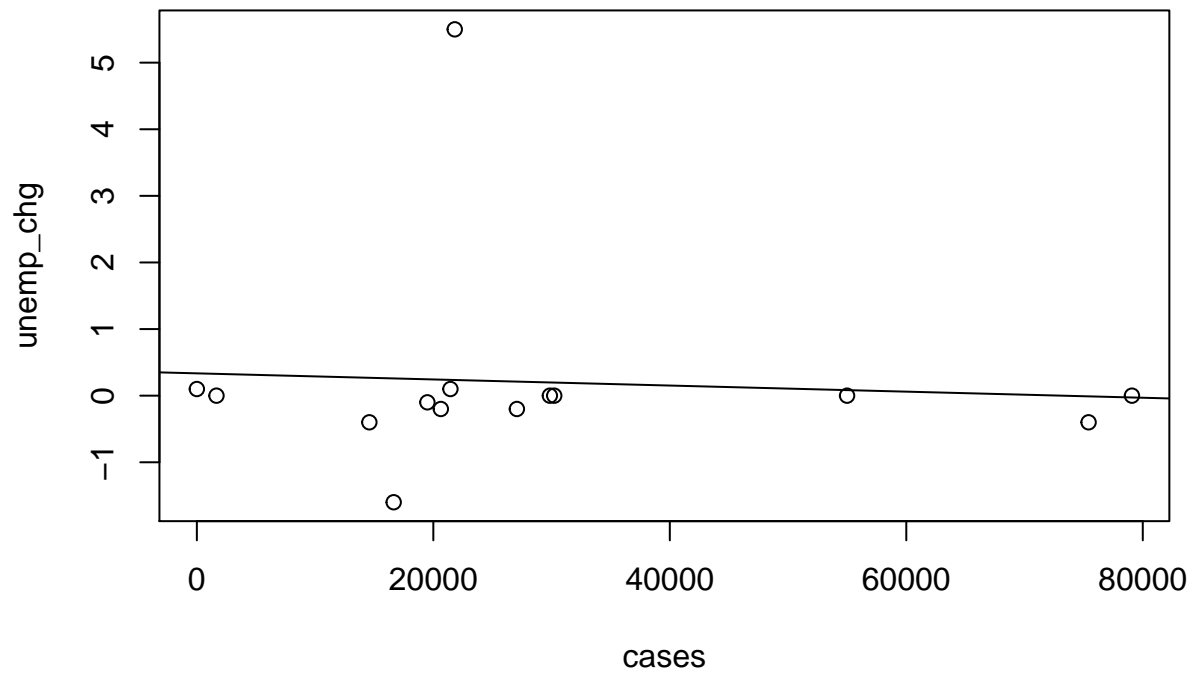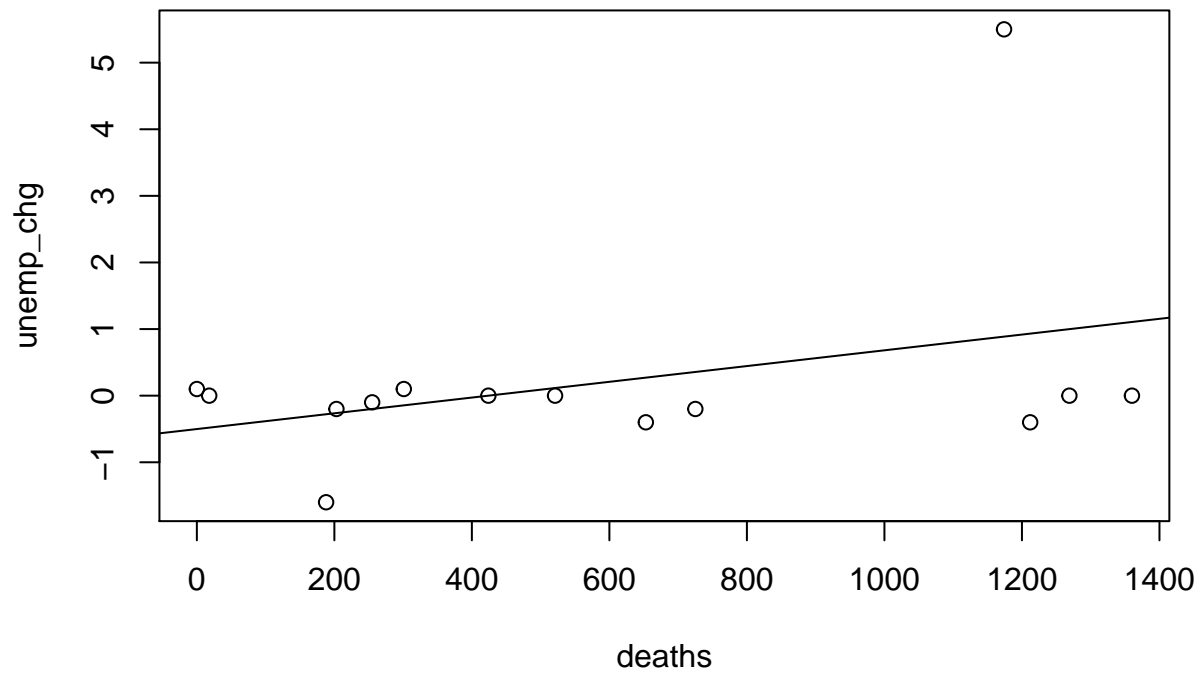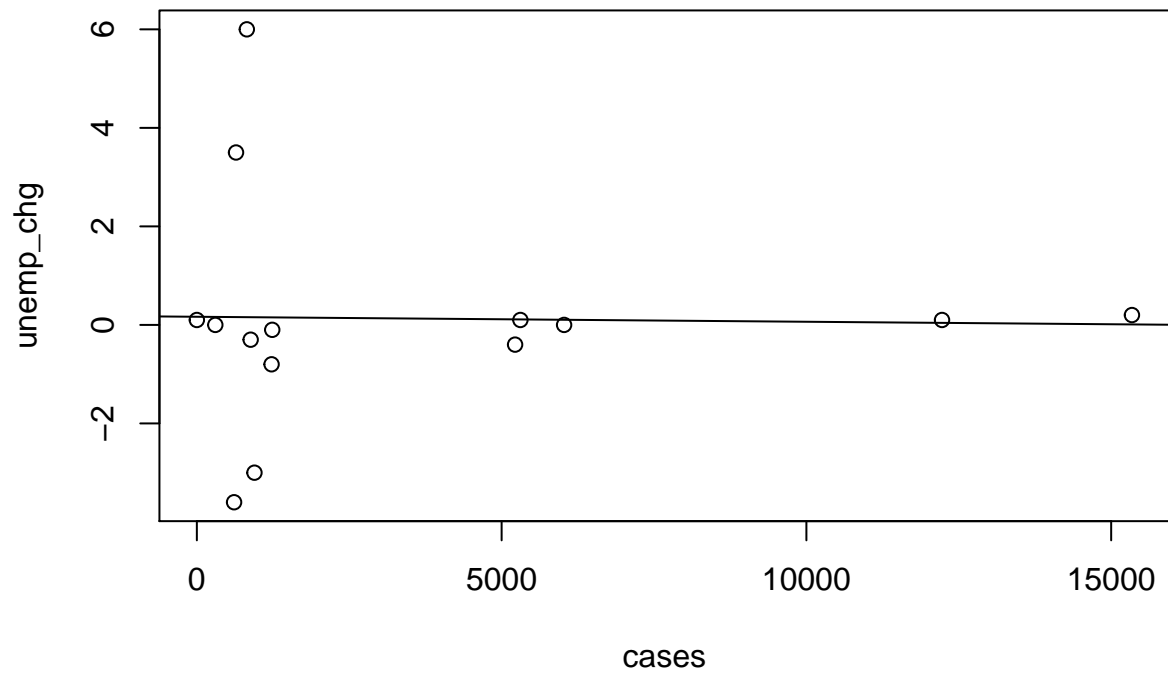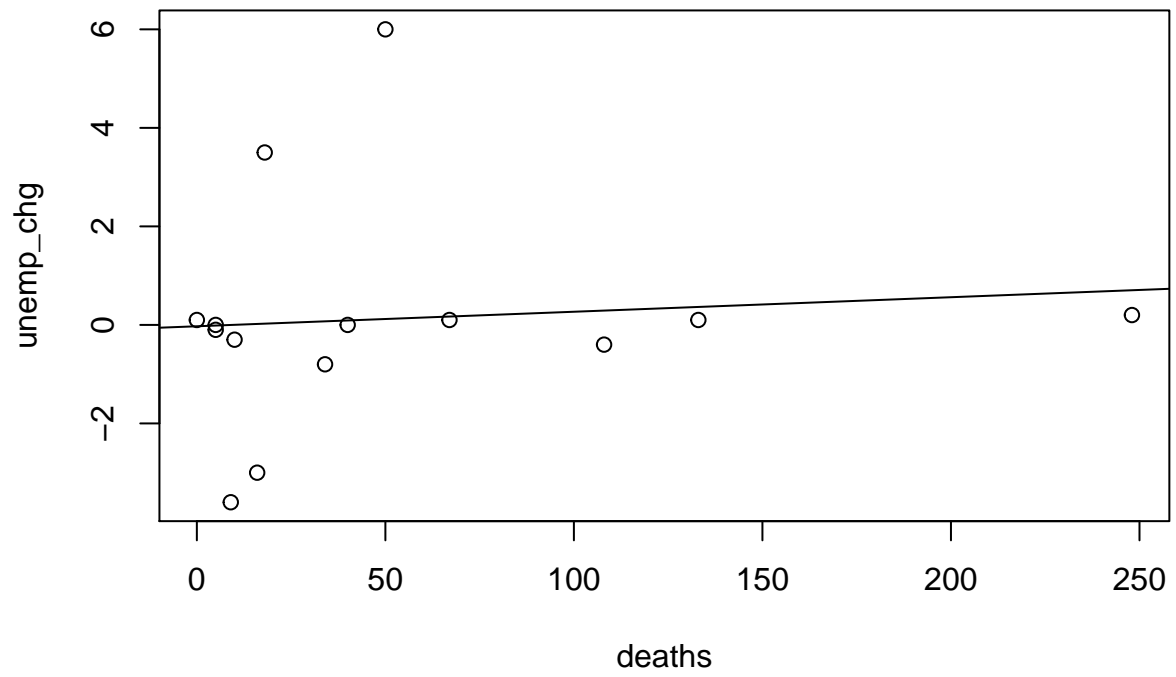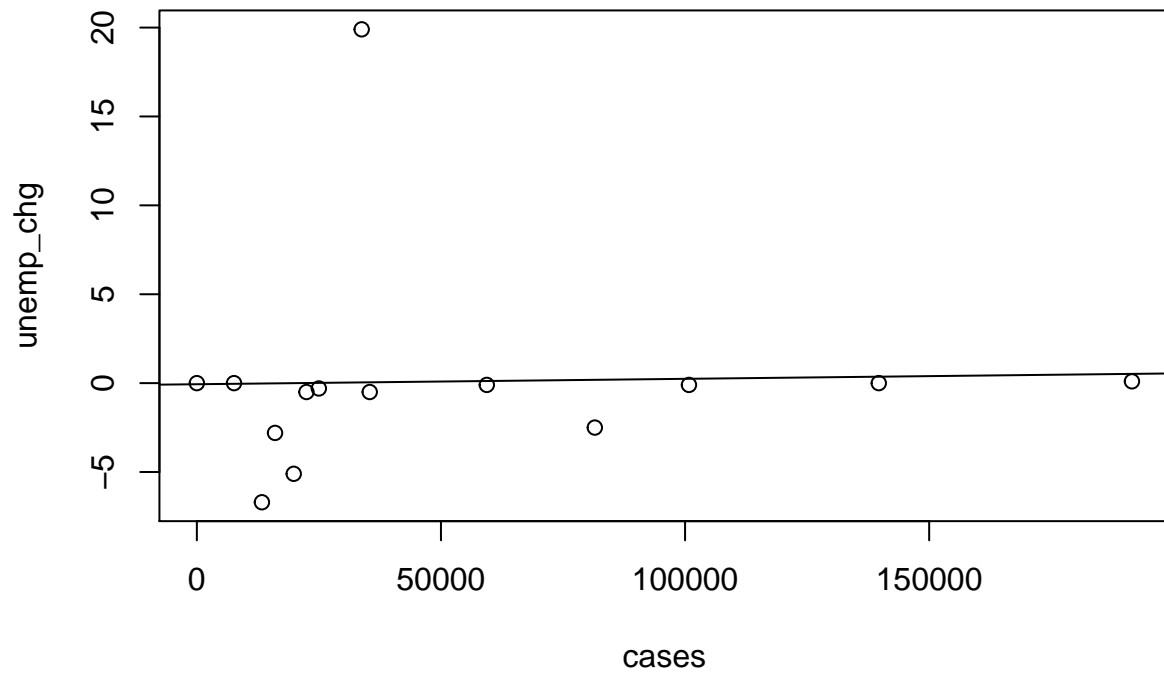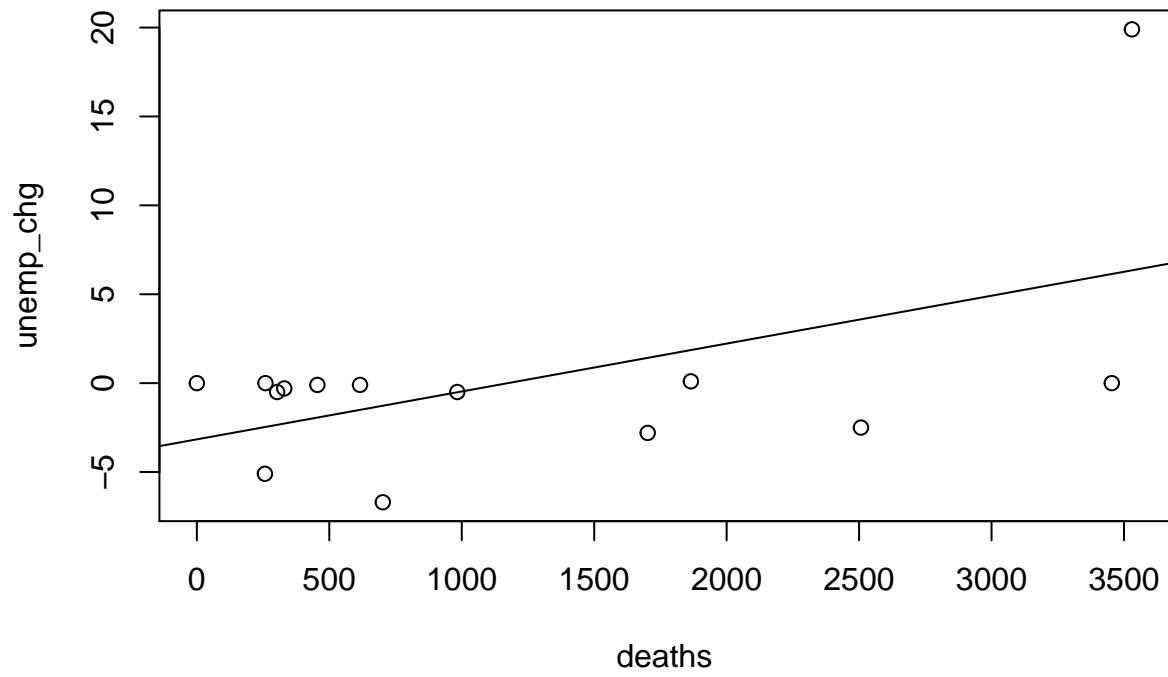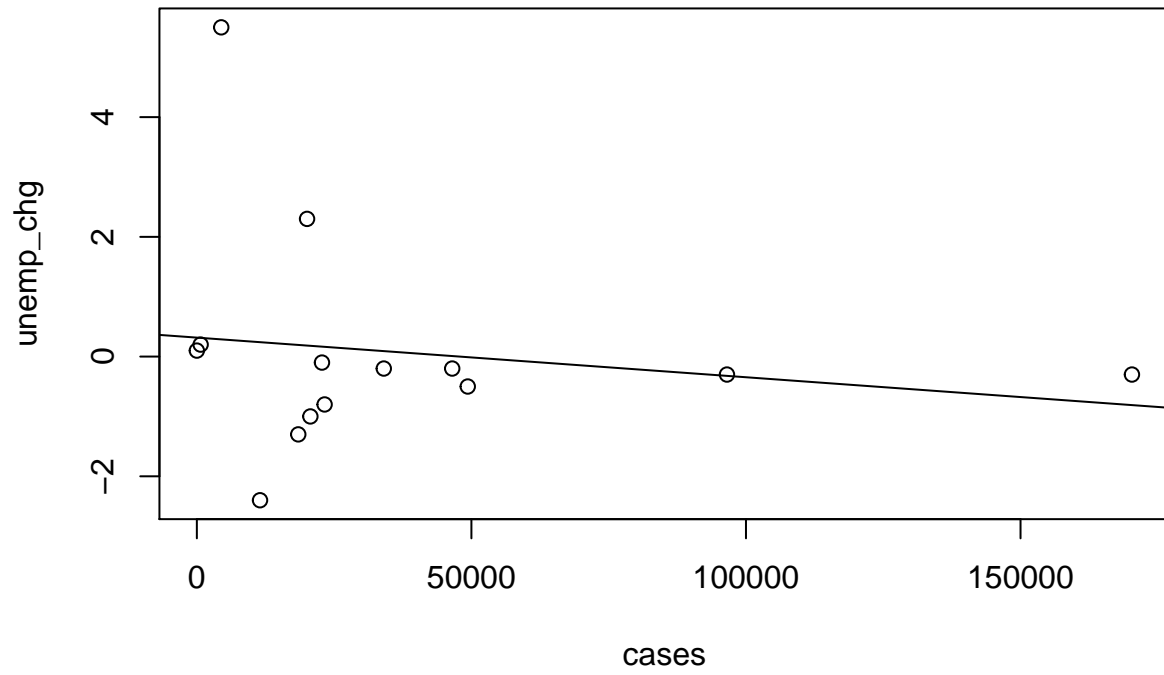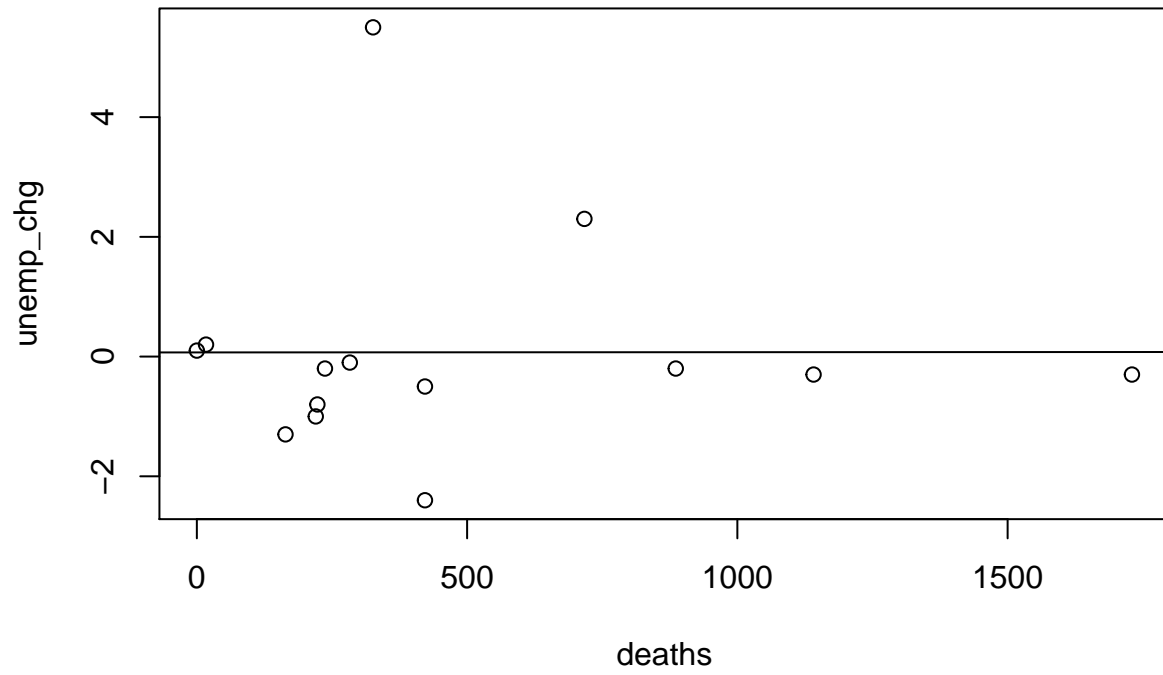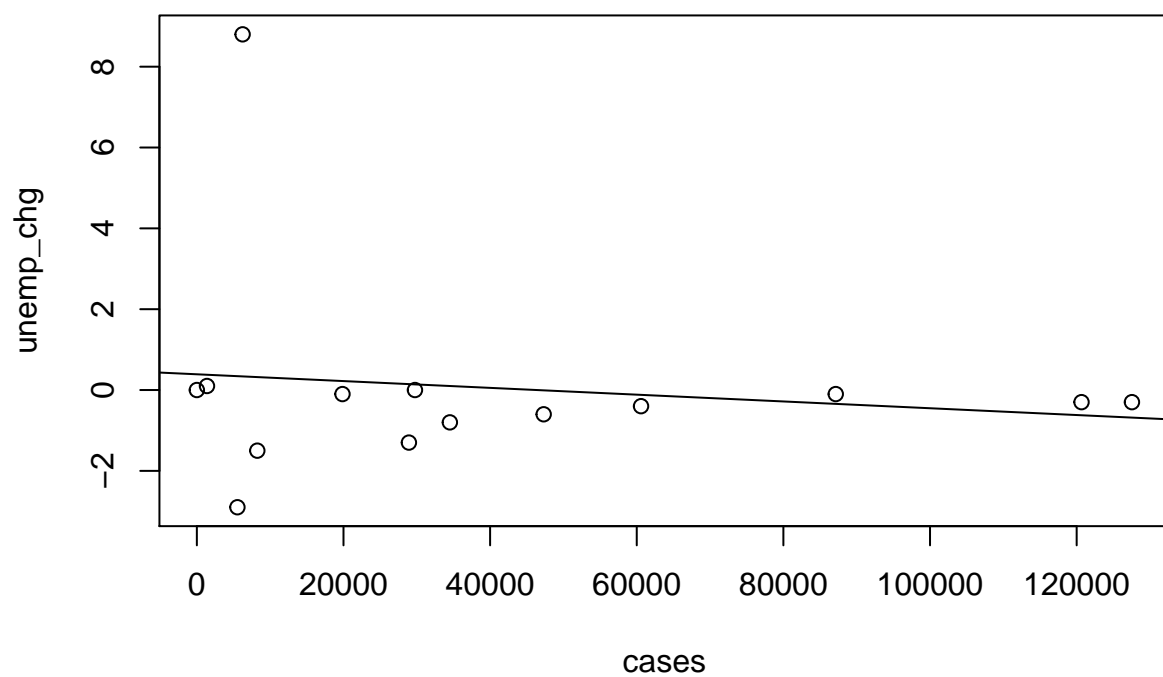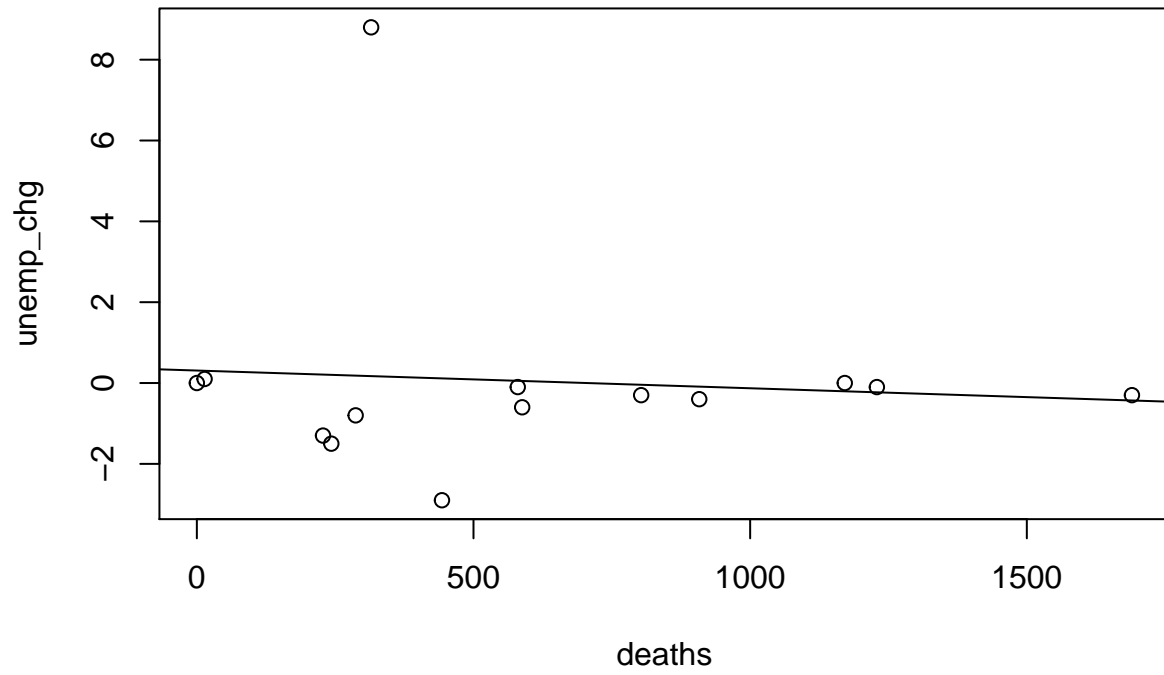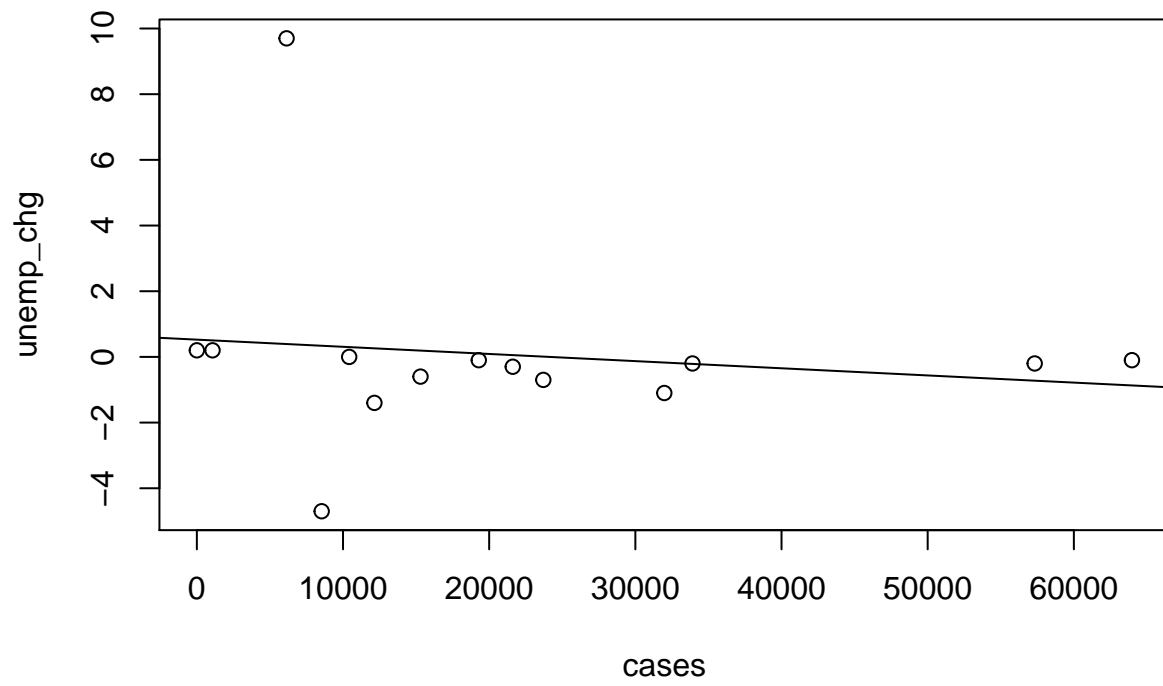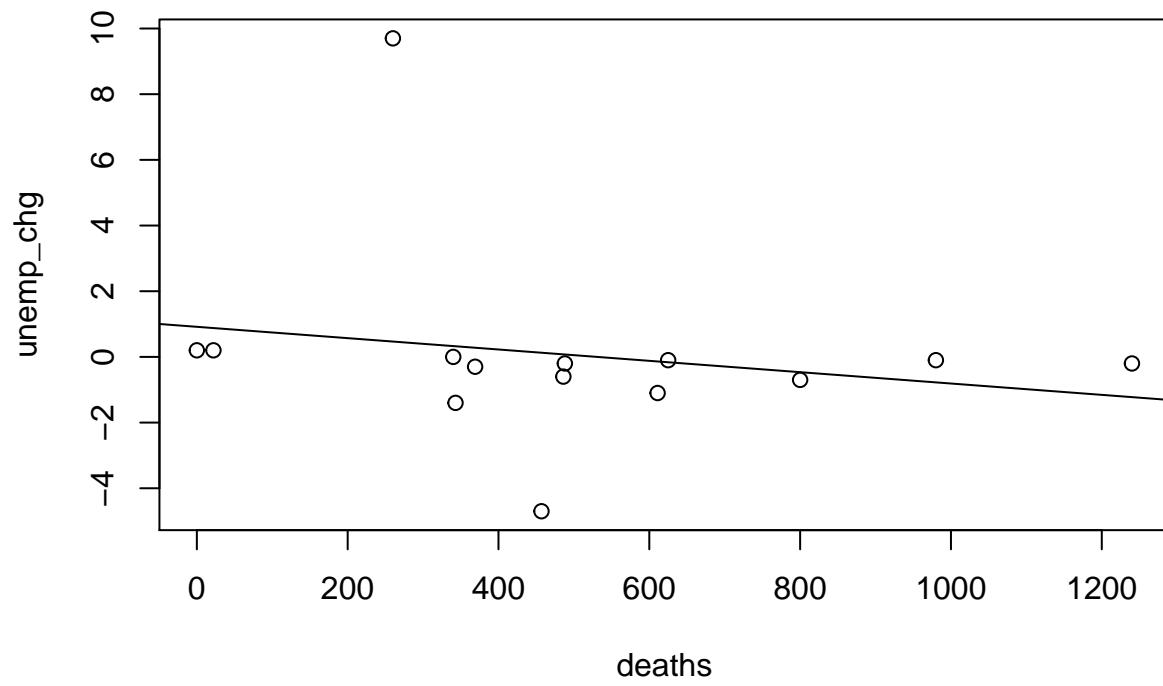**AL unemployment change vs deaths**
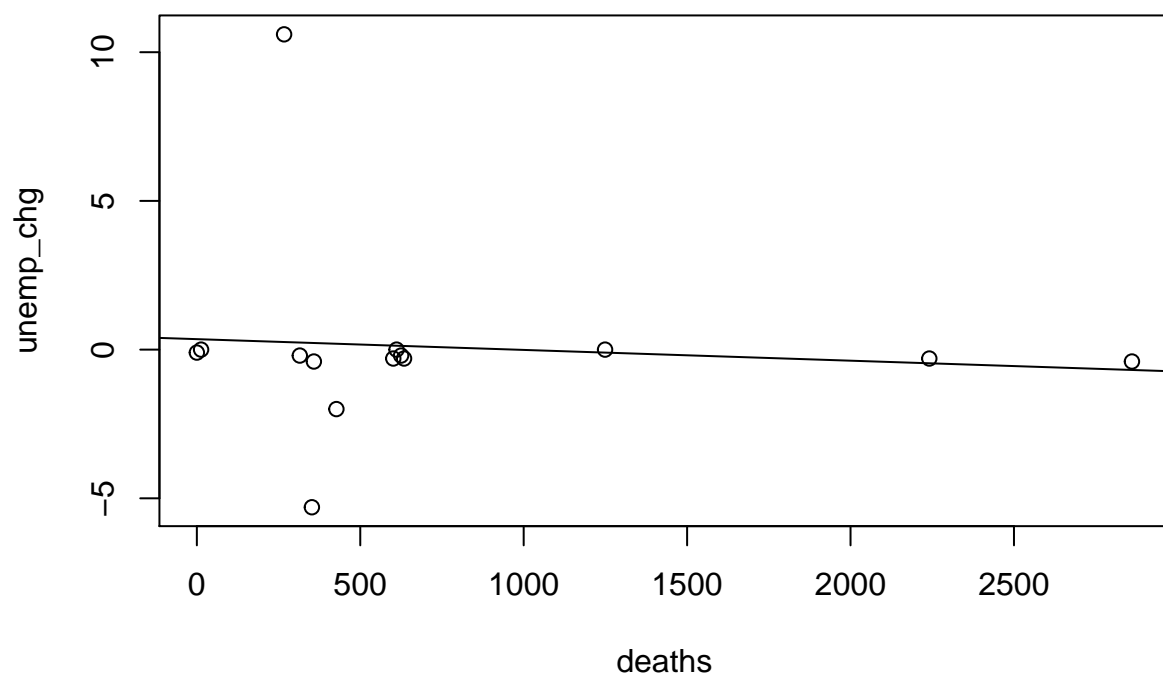
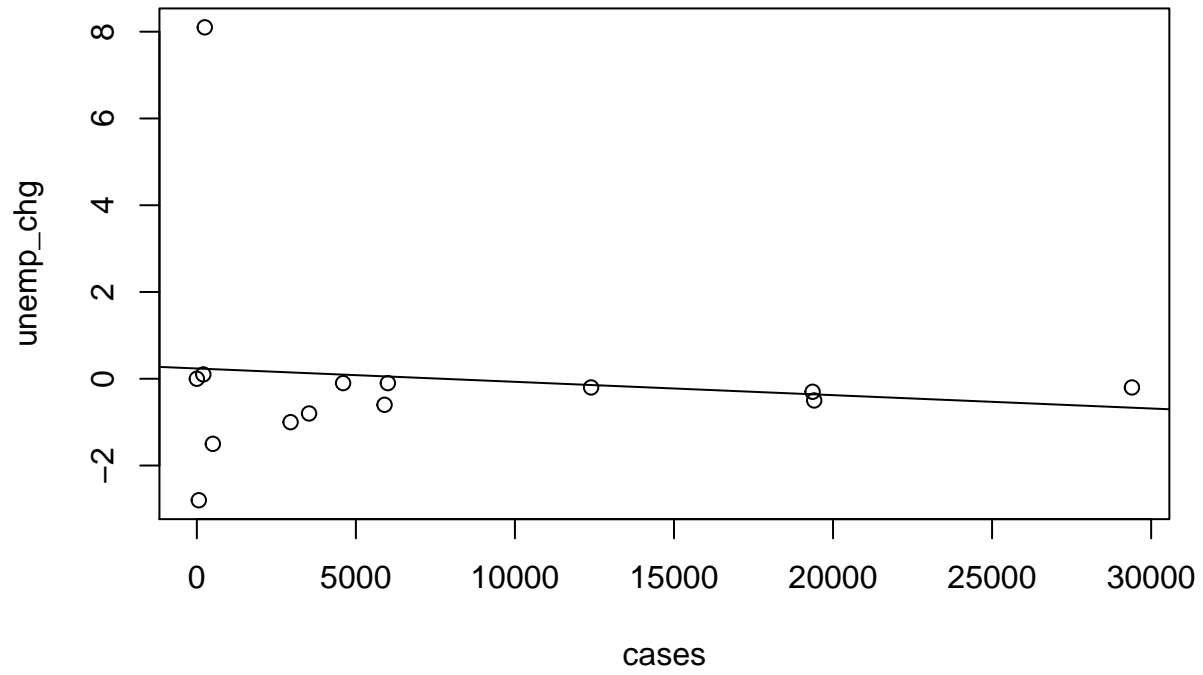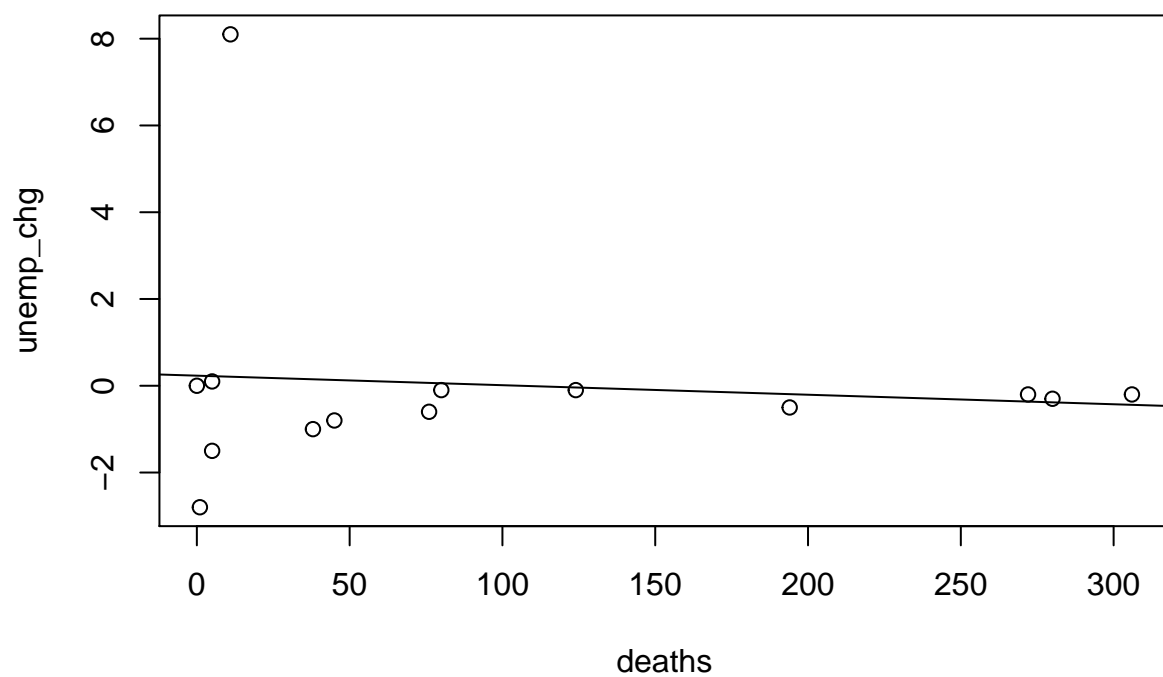## AR unemployment change vs cases
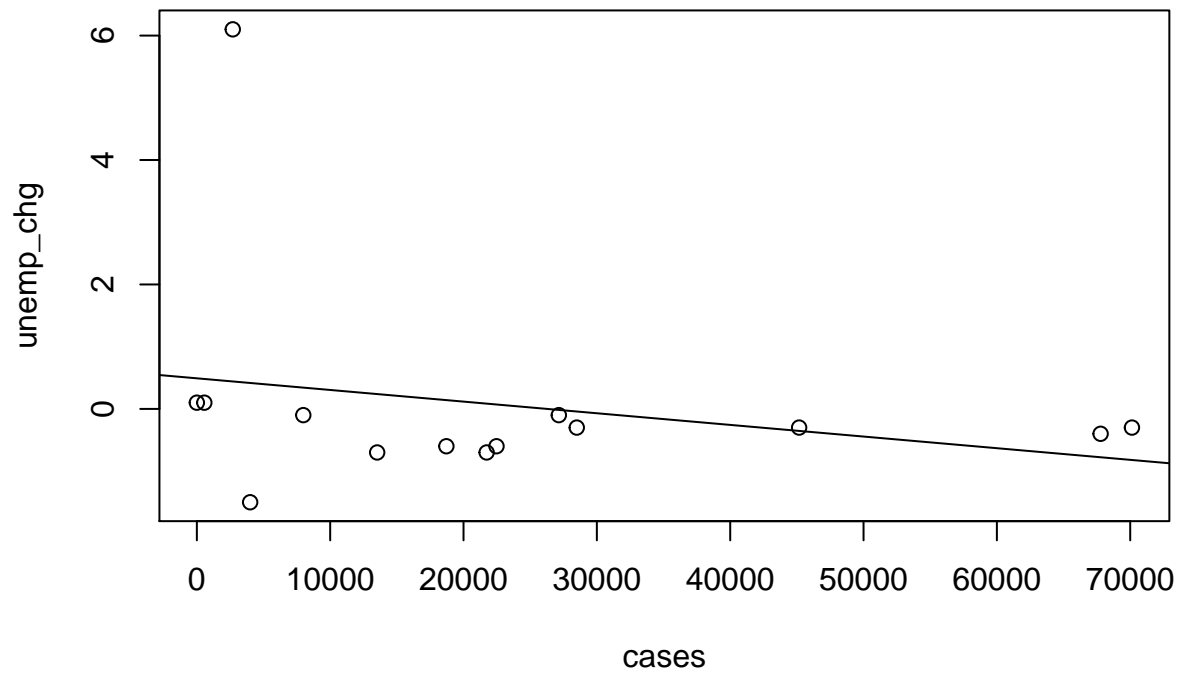
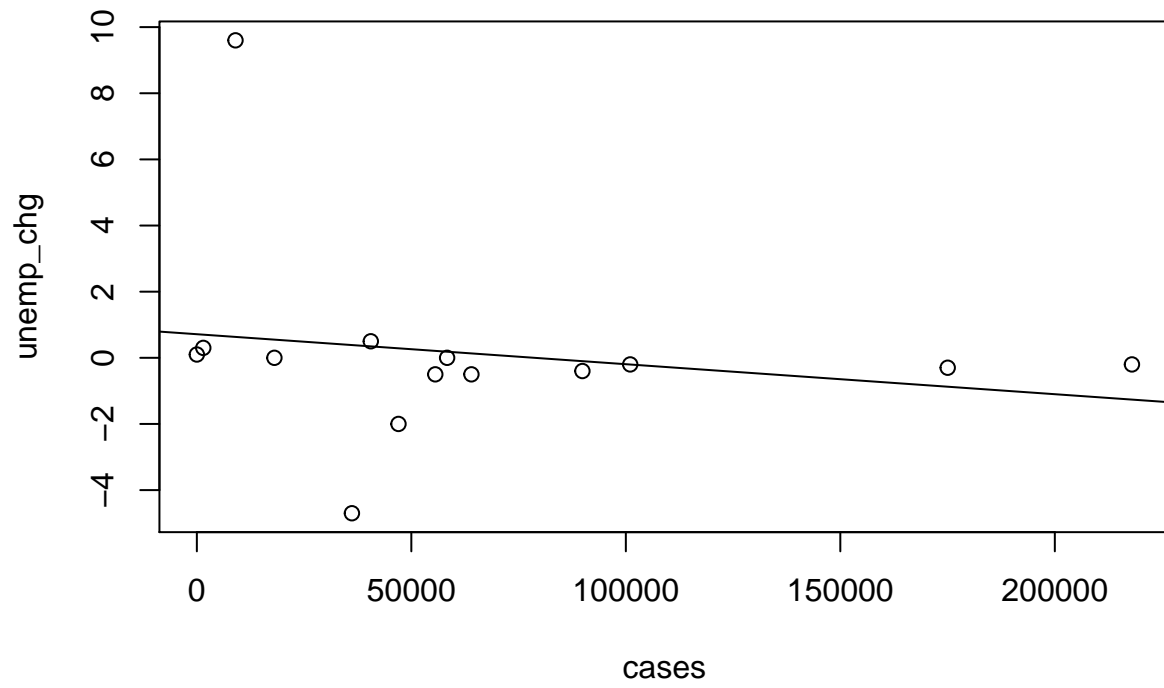# AR unemployment change vs deaths
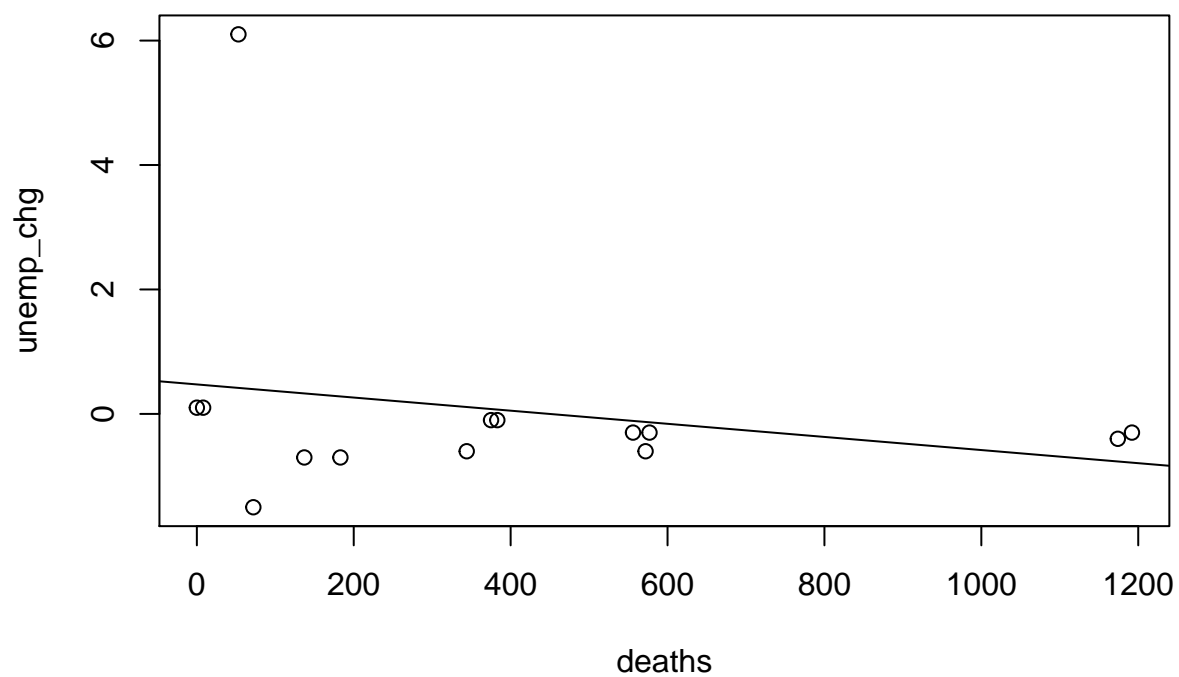
AZ unemployment change vs cases
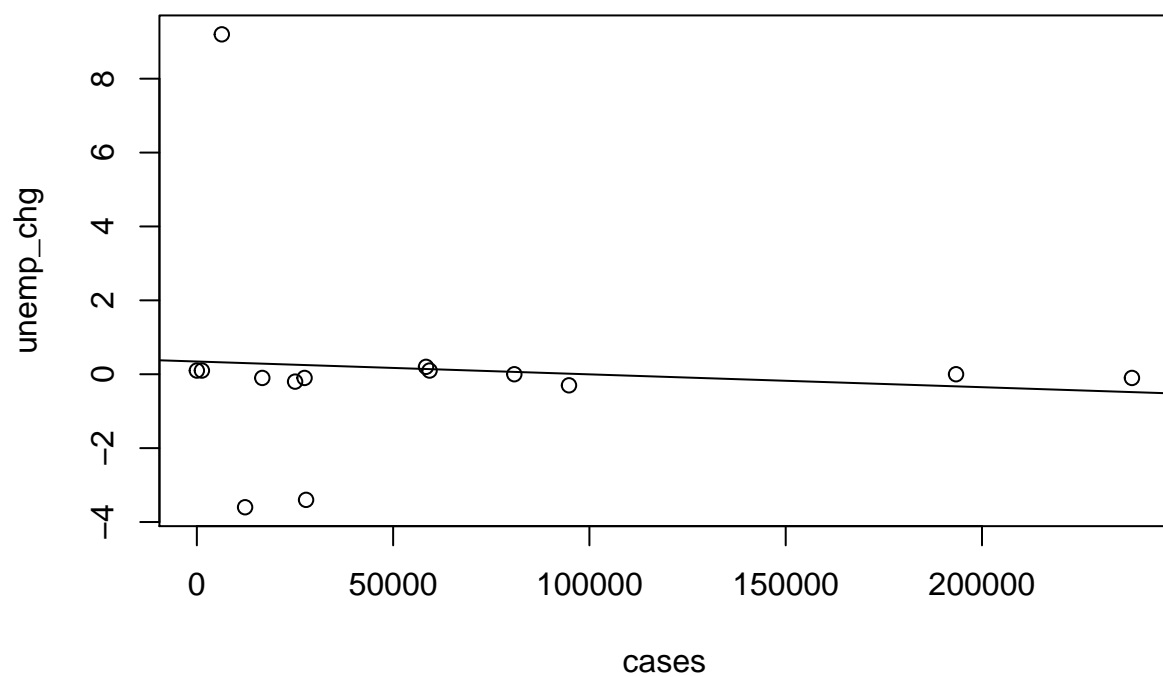
AZ unemployment change vs deaths
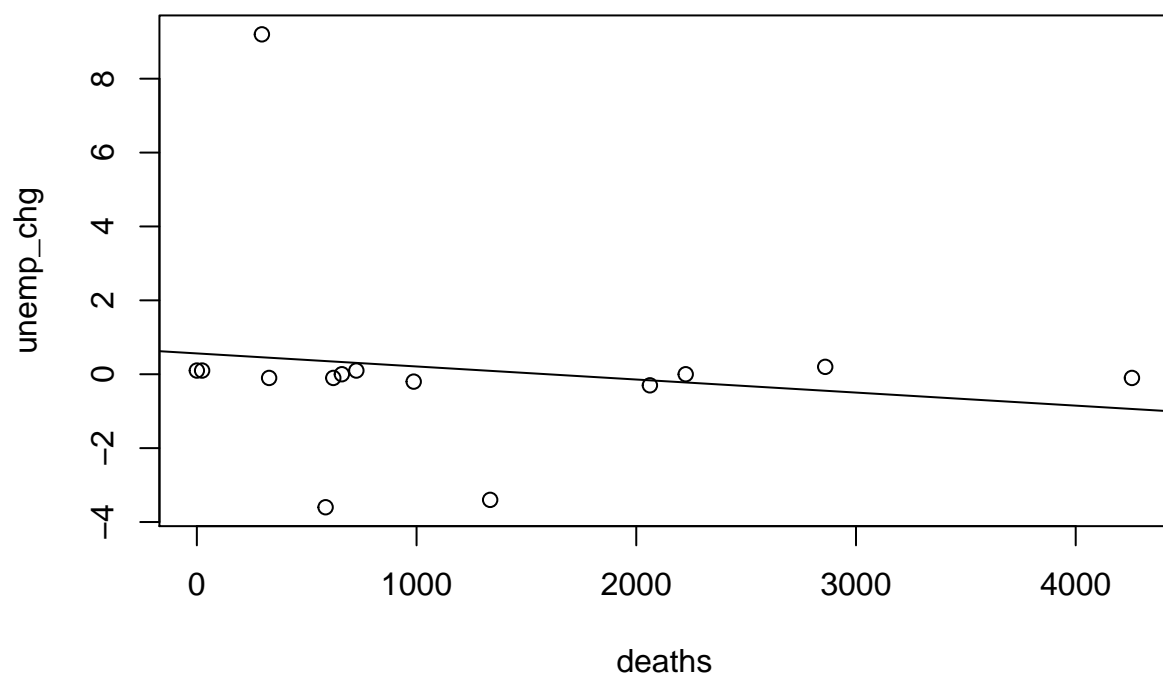
CA unemployment change vs cases
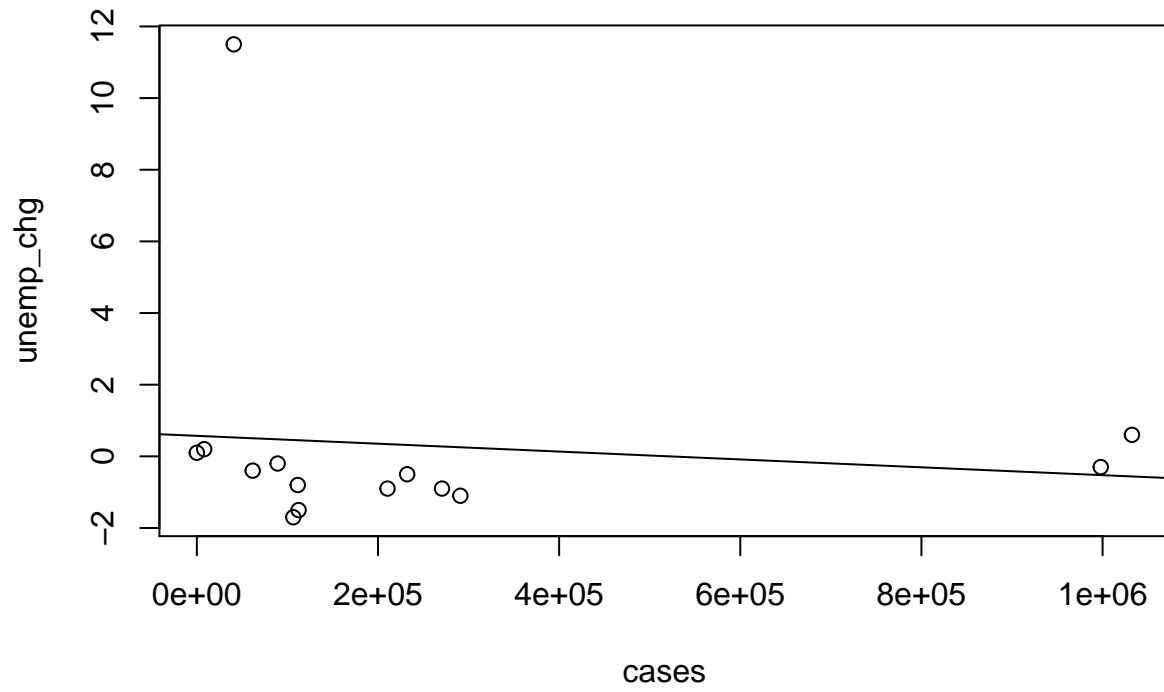
CA unemployment change vs deaths
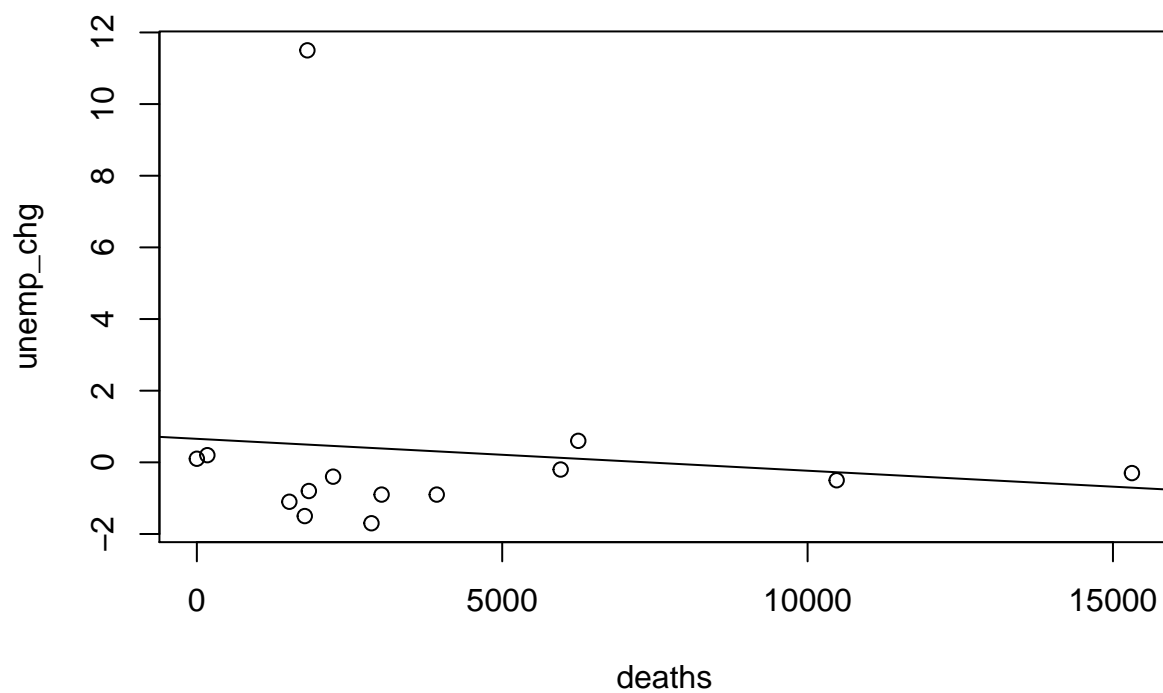
CO unemployment change vs cases
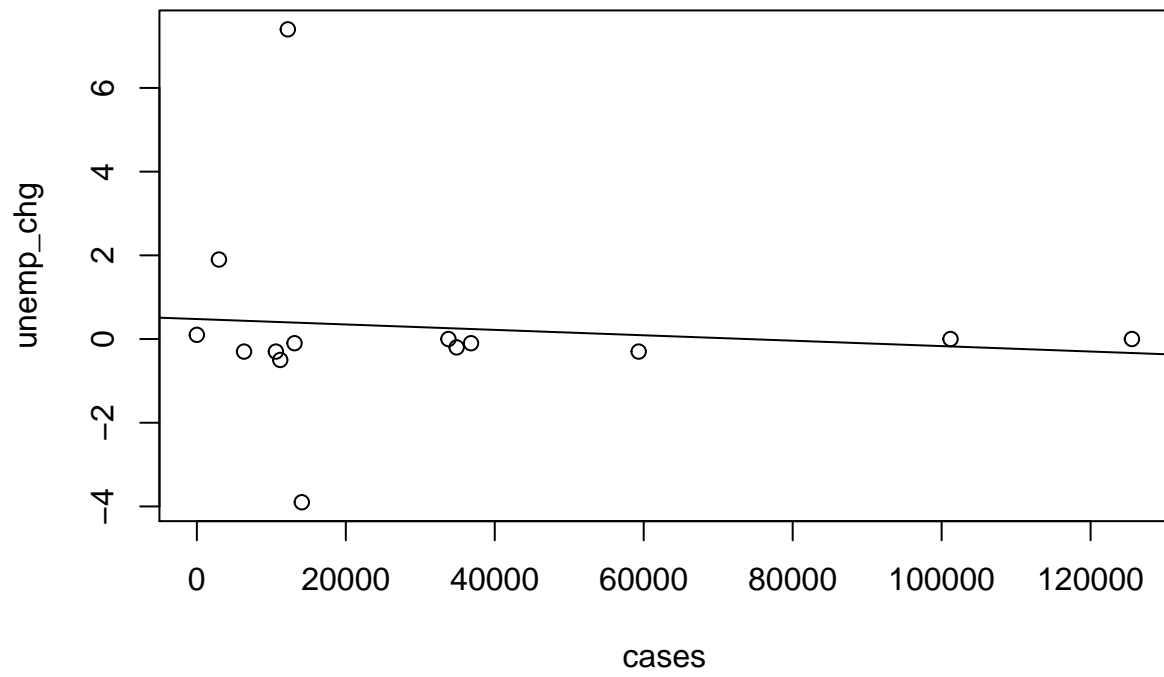
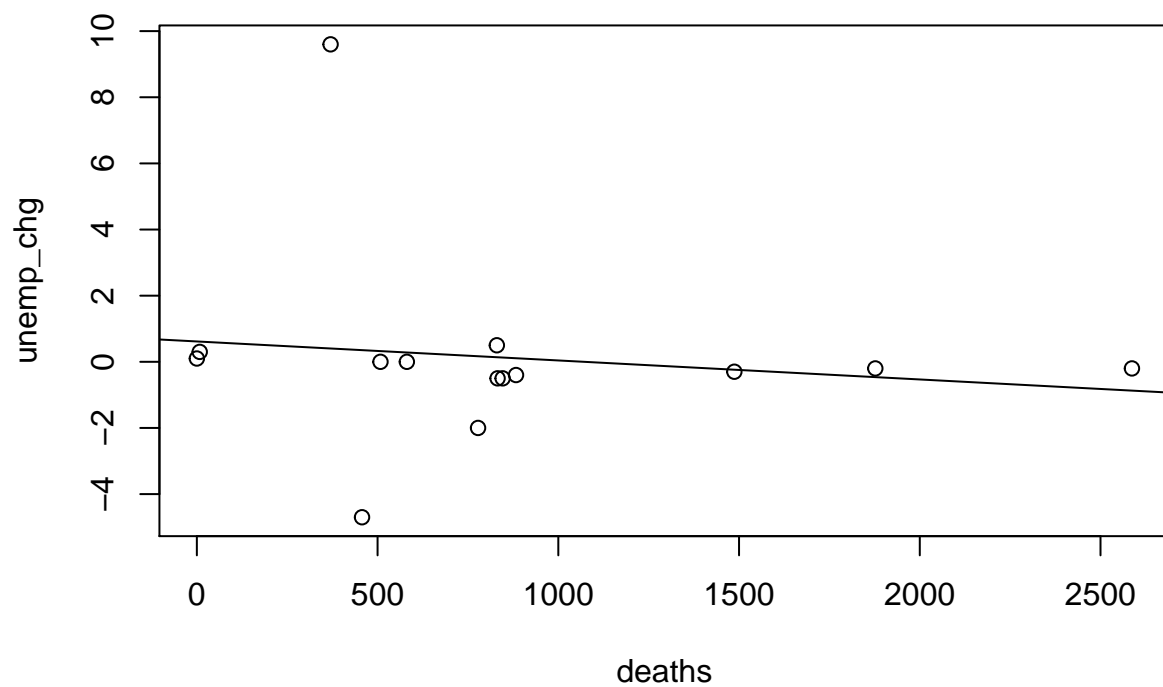# CO unemployment change vs deaths
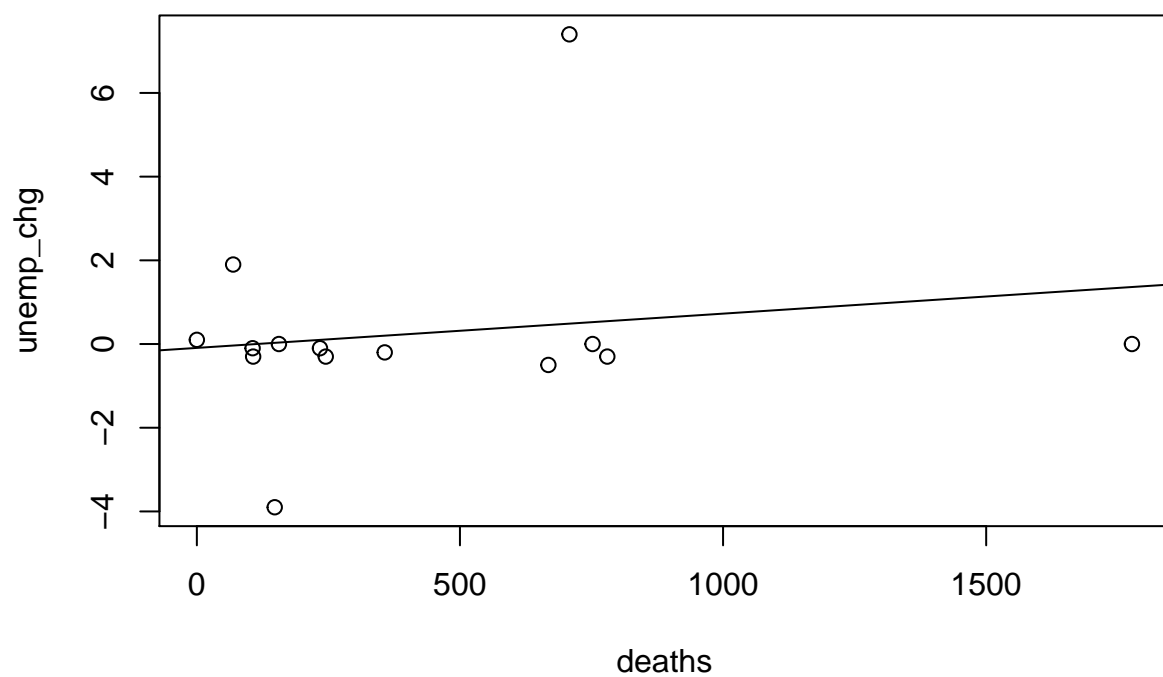
CT unemployment change vs cases
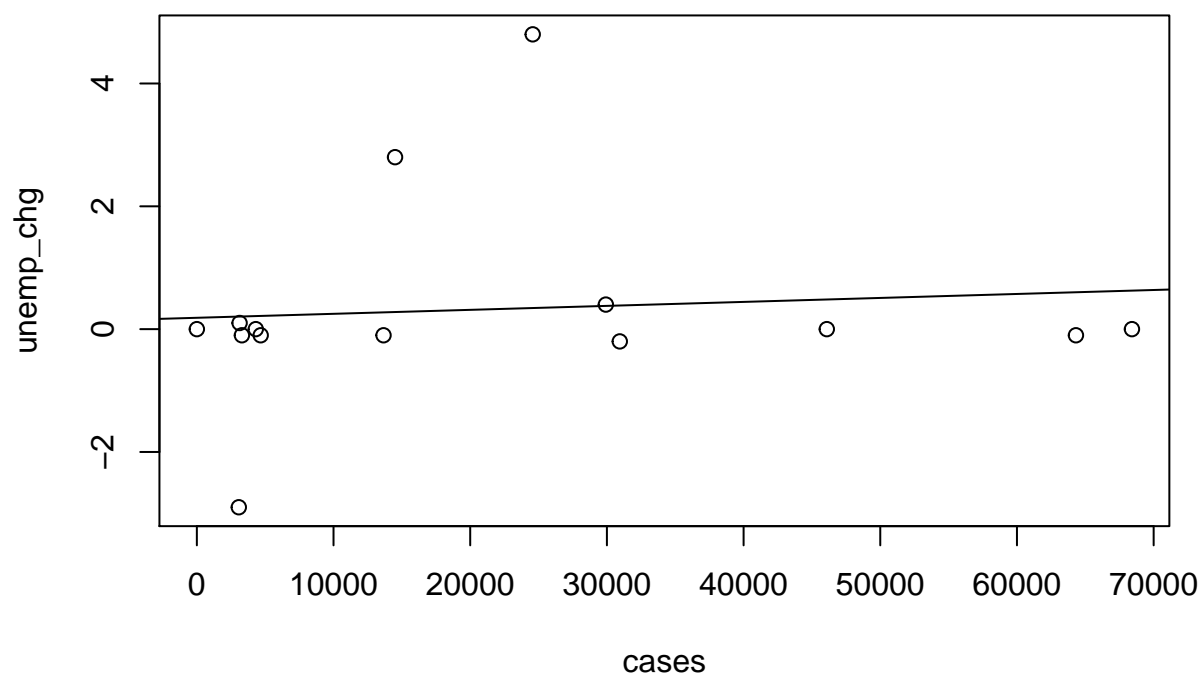
CT unemployment change vs deaths
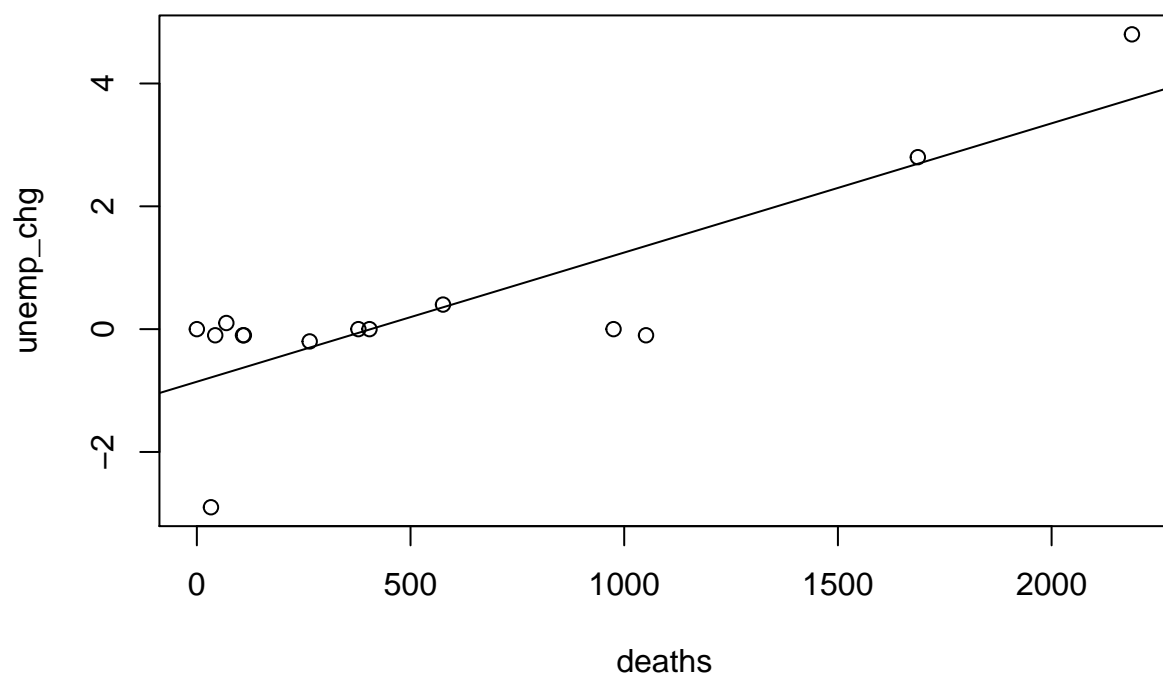
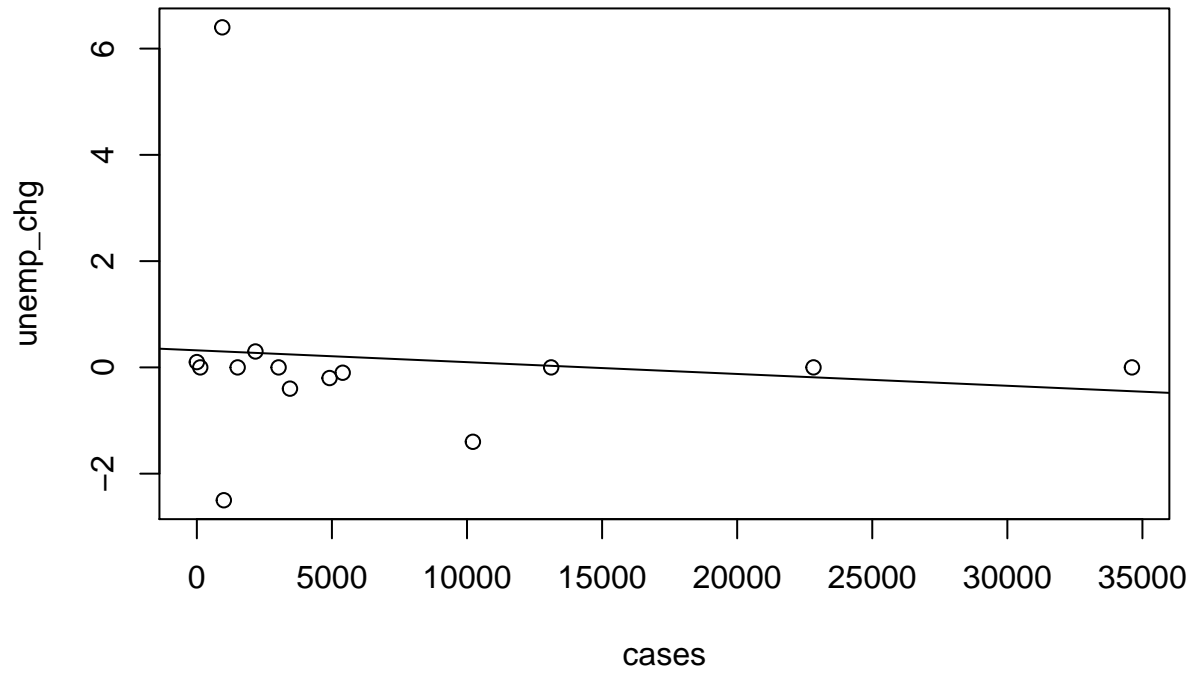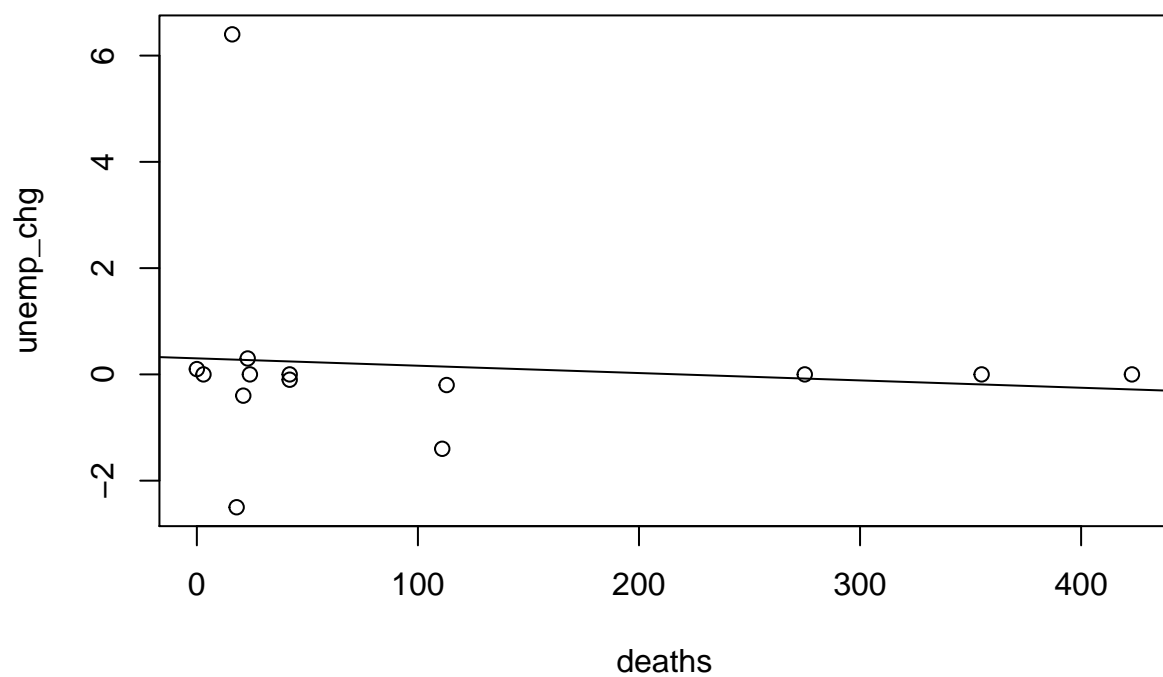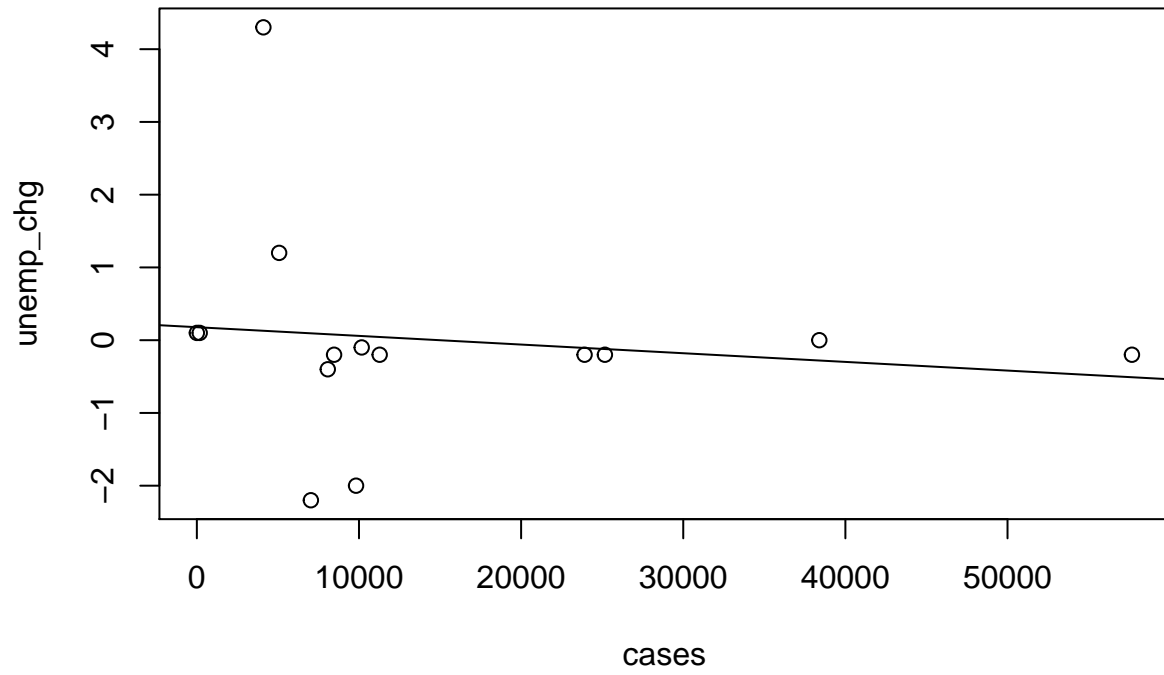# DC unemployment change vs cases
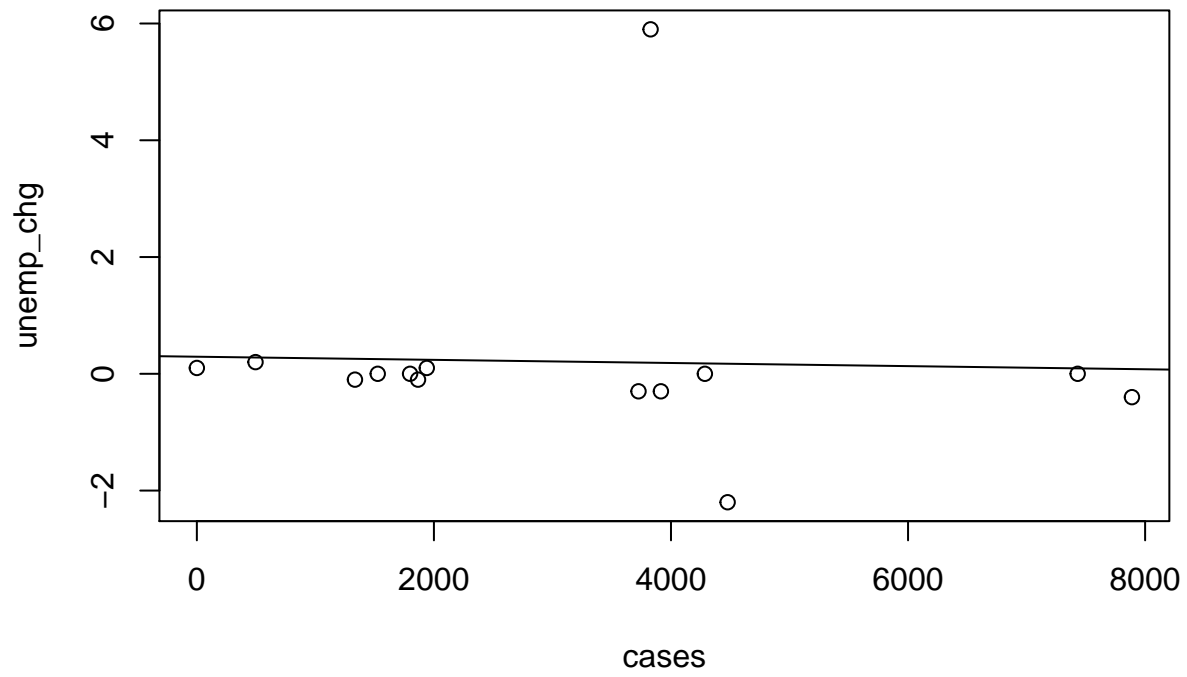
DC unemployment change vs deaths
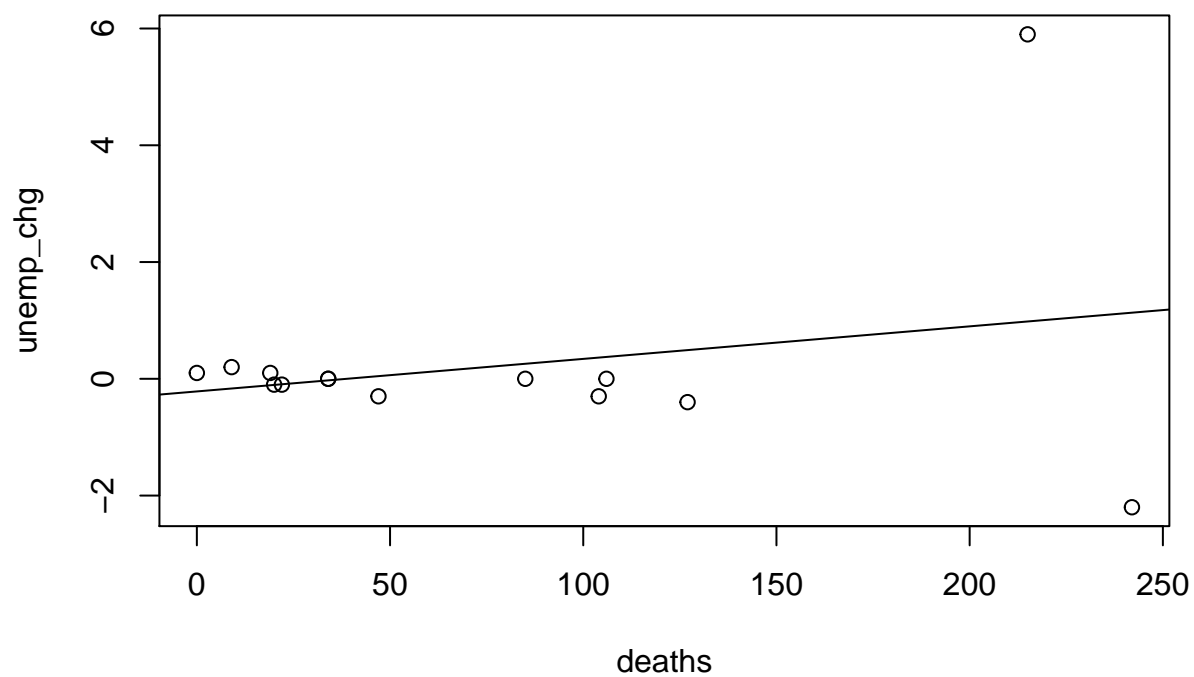
DE unemployment change vs cases
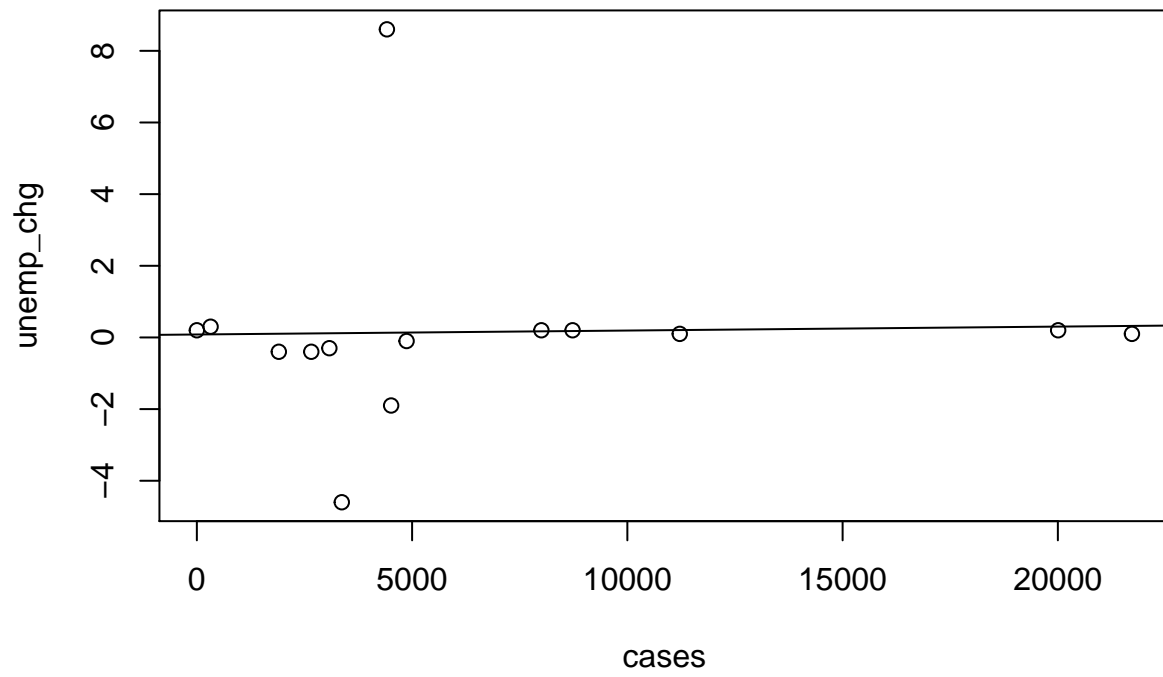
DE unemployment change vs deaths
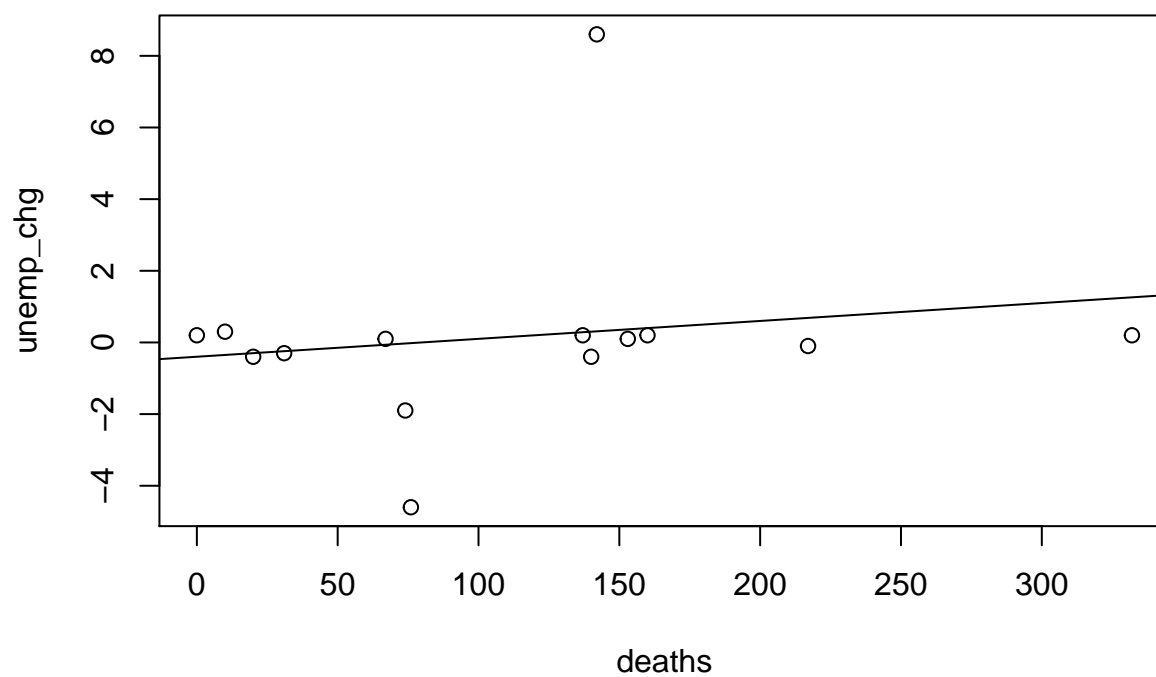
FL unemployment change vs cases
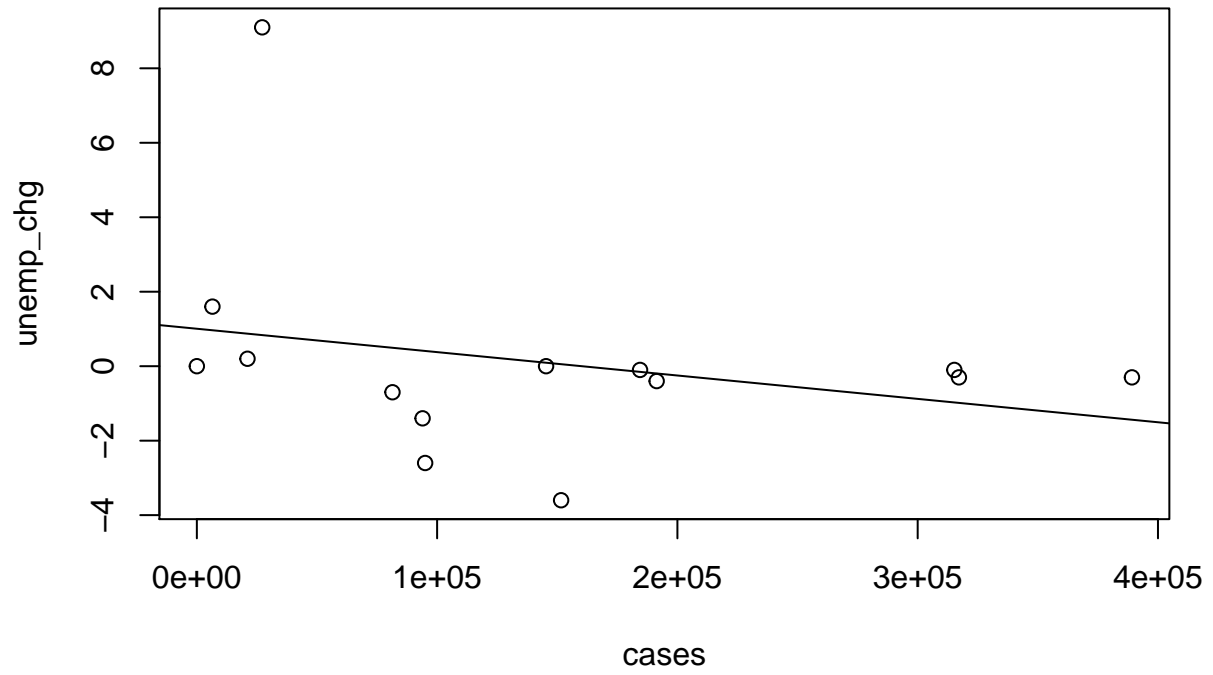
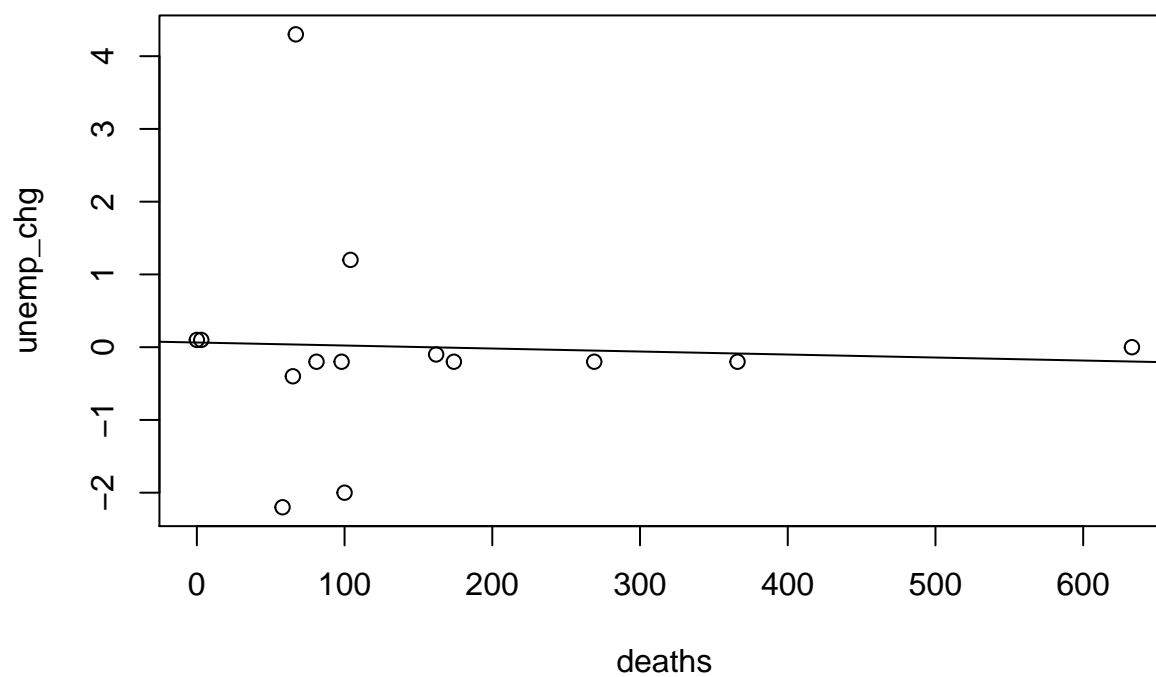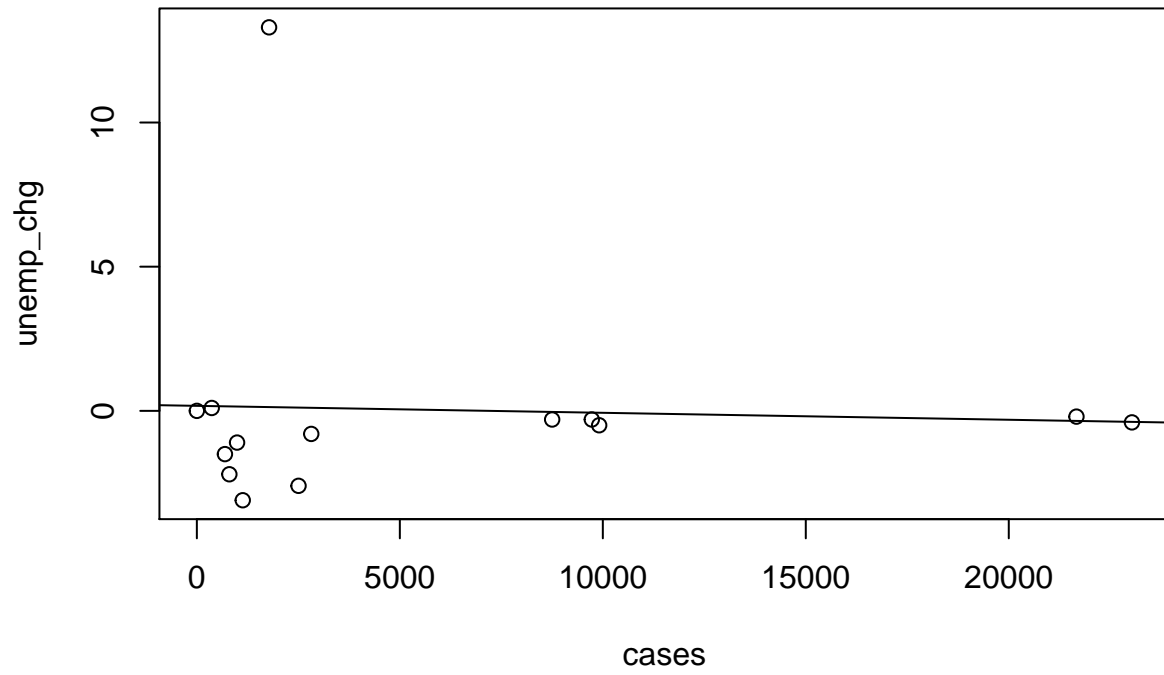# FL unemployment change vs deaths
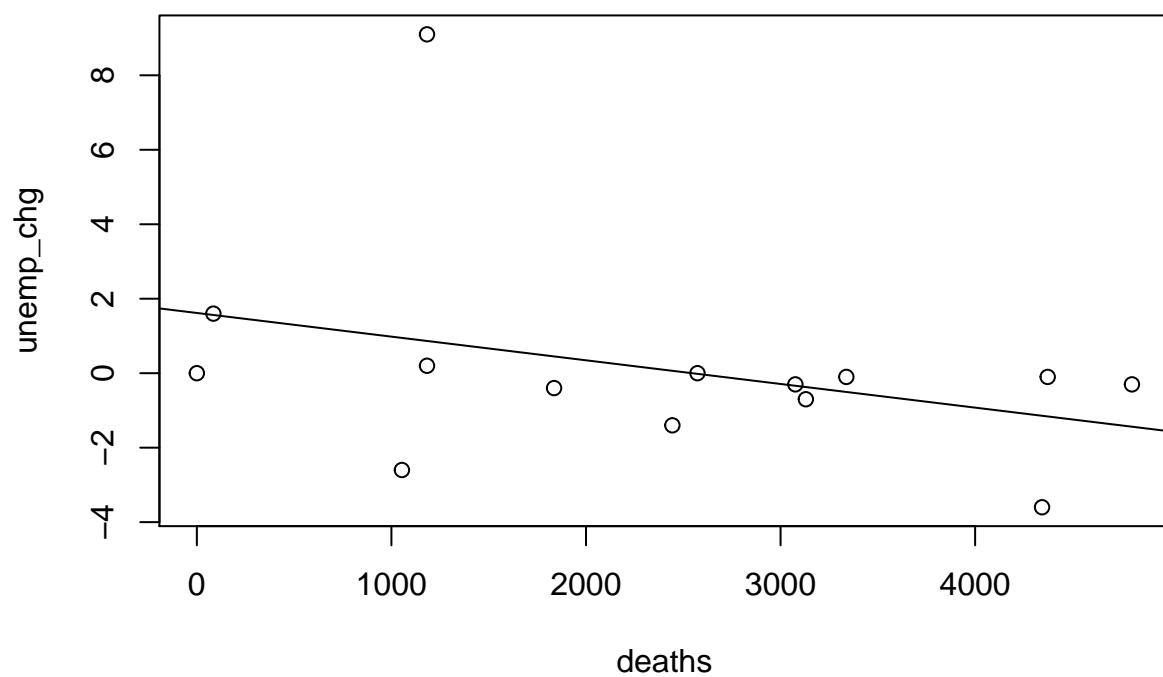
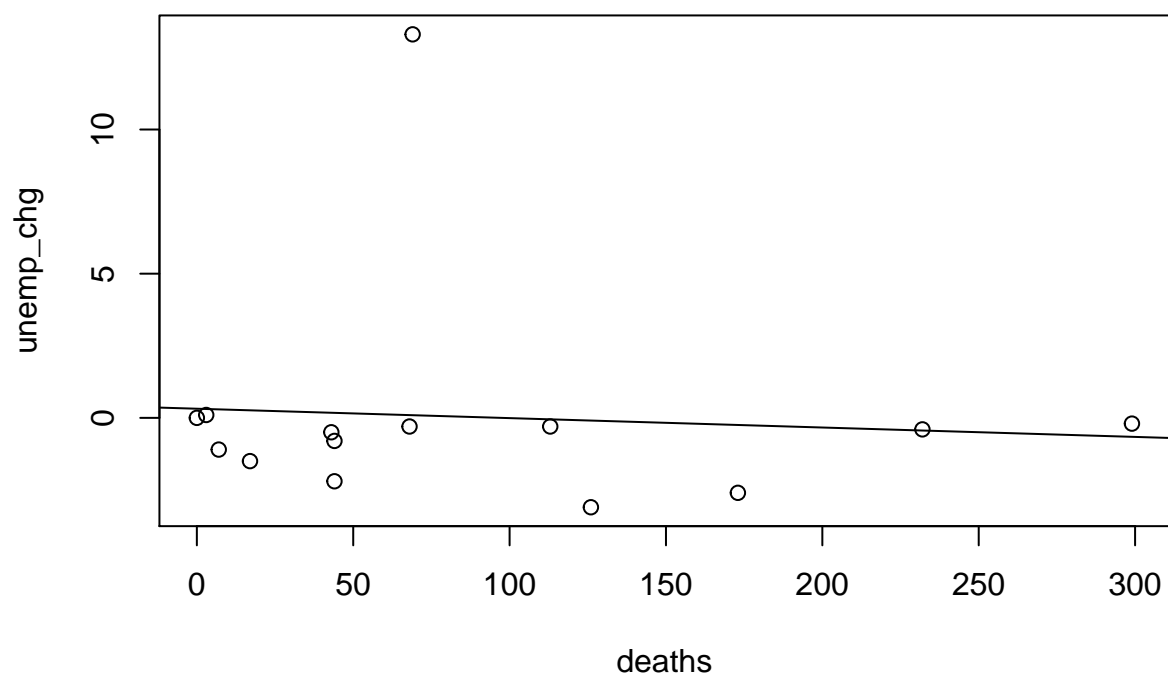# GA unemployment change vs cases
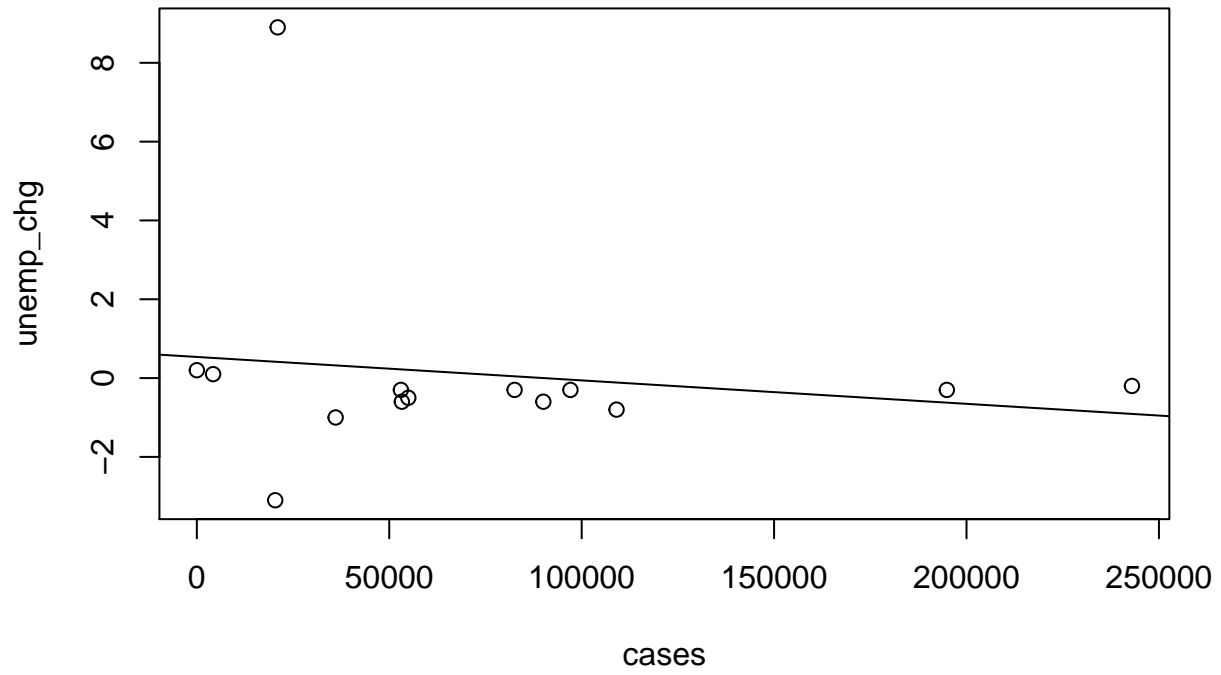
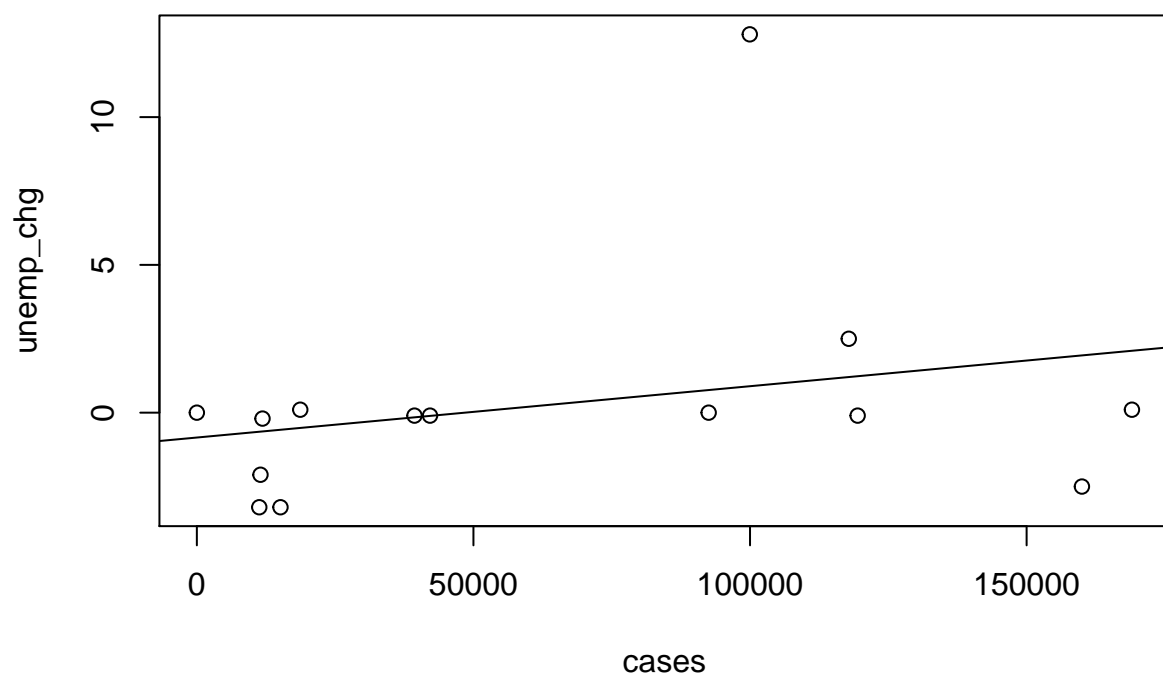# GA unemployment change vs deaths
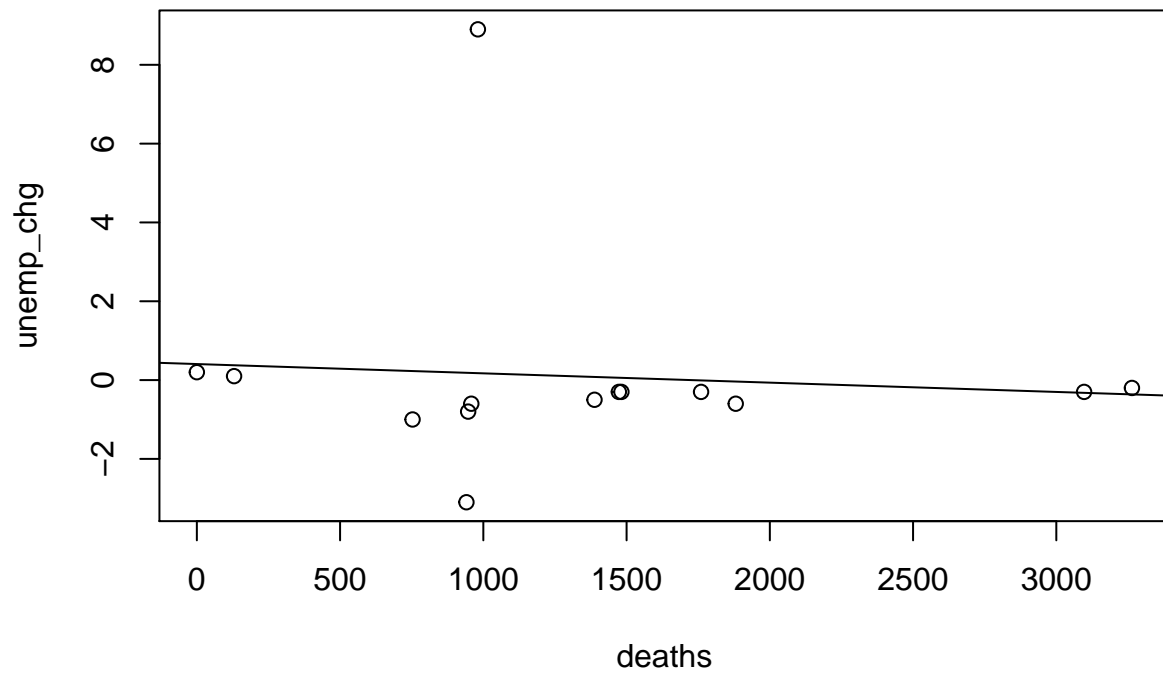
HI unemployment change vs cases
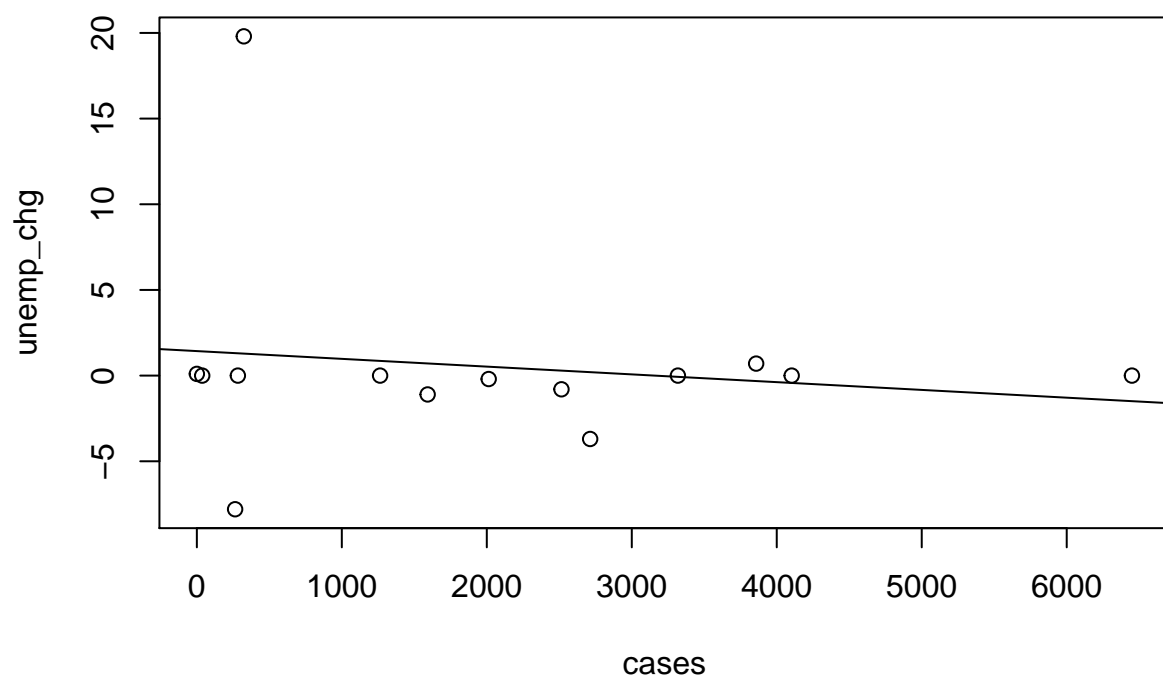
HI unemployment change vs deaths
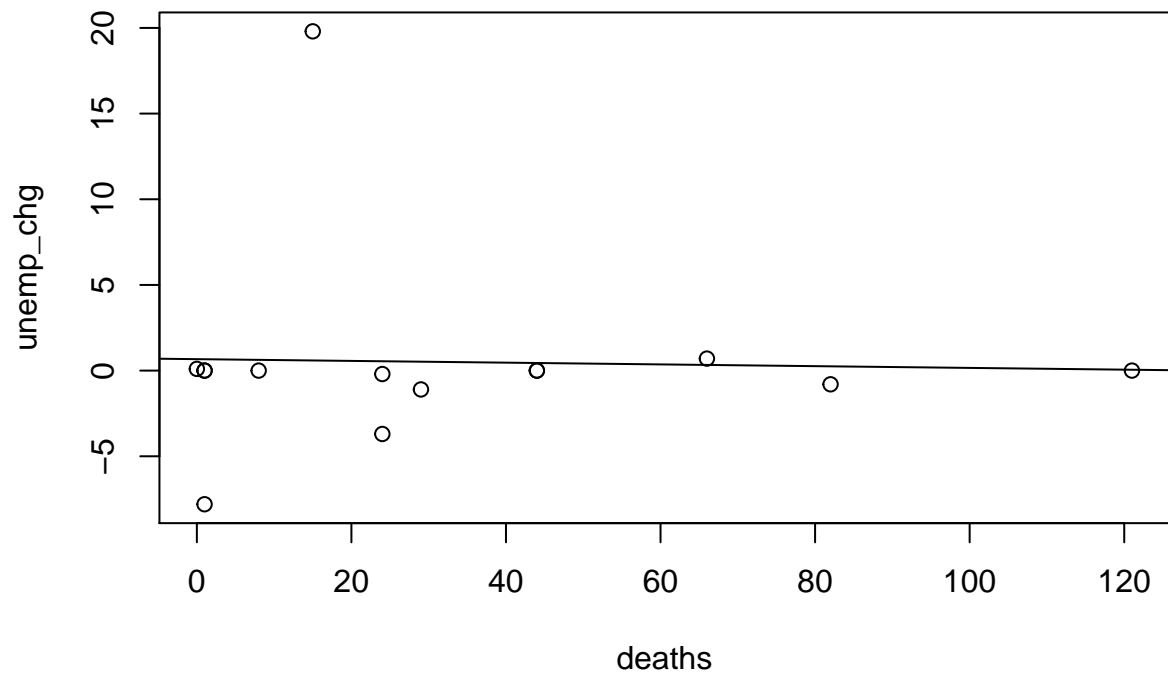
IA unemployment change vs cases
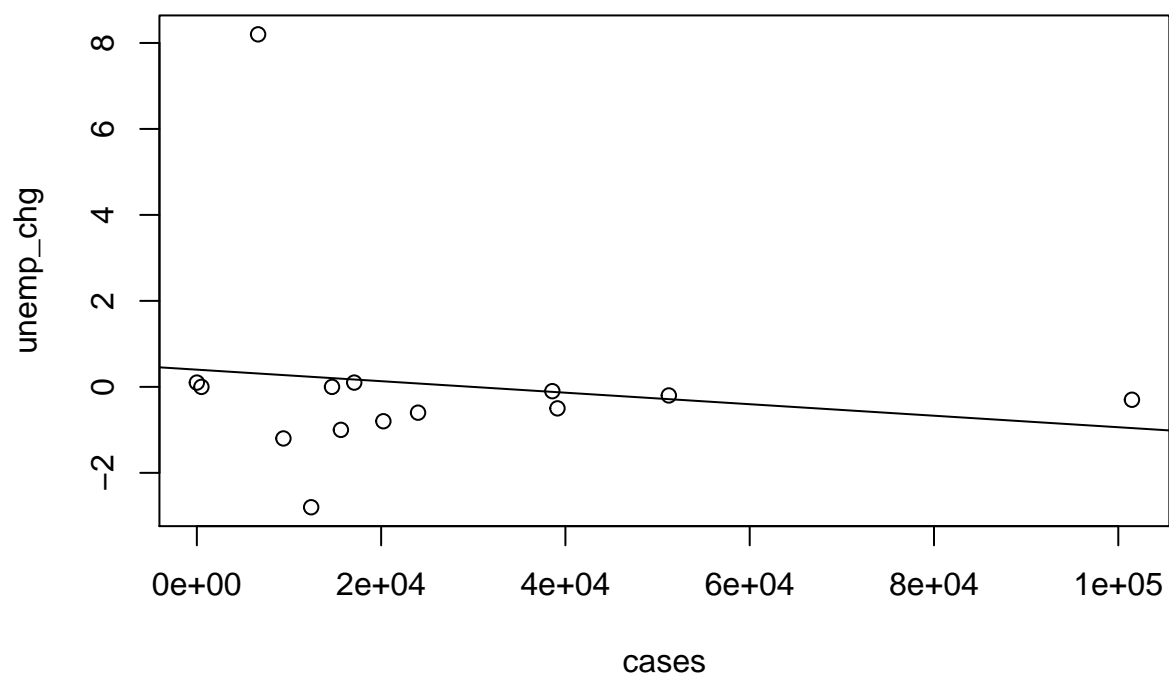
IA unemployment change vs deaths
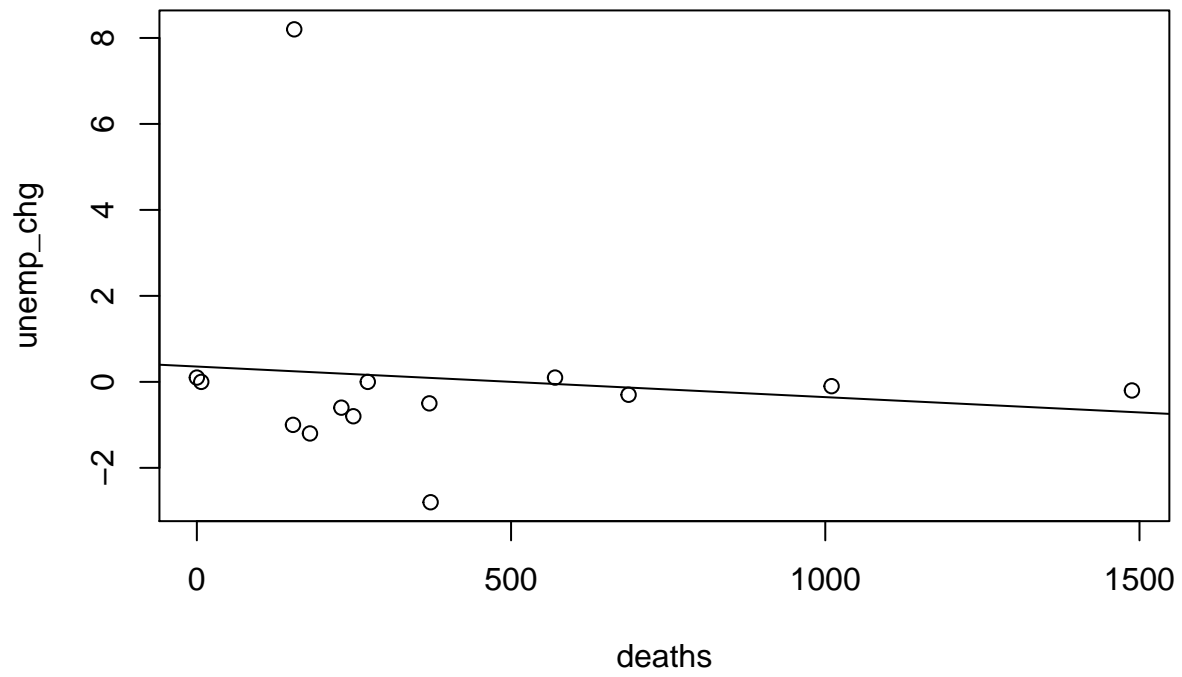
ID unemployment change vs cases
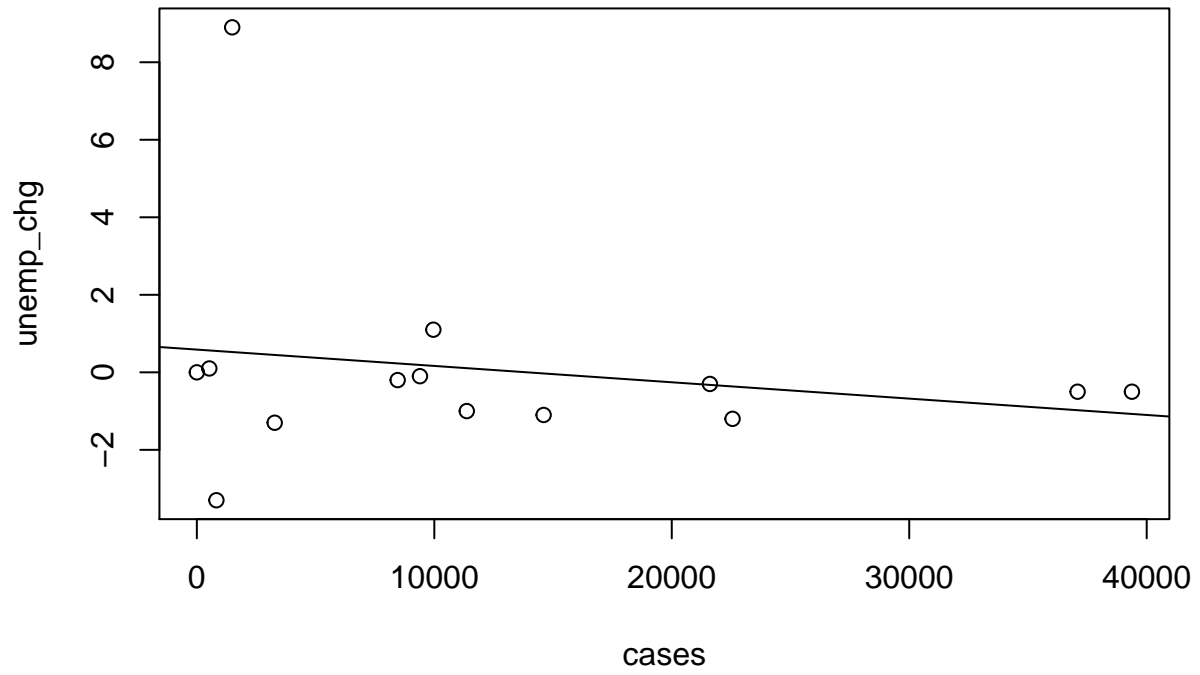
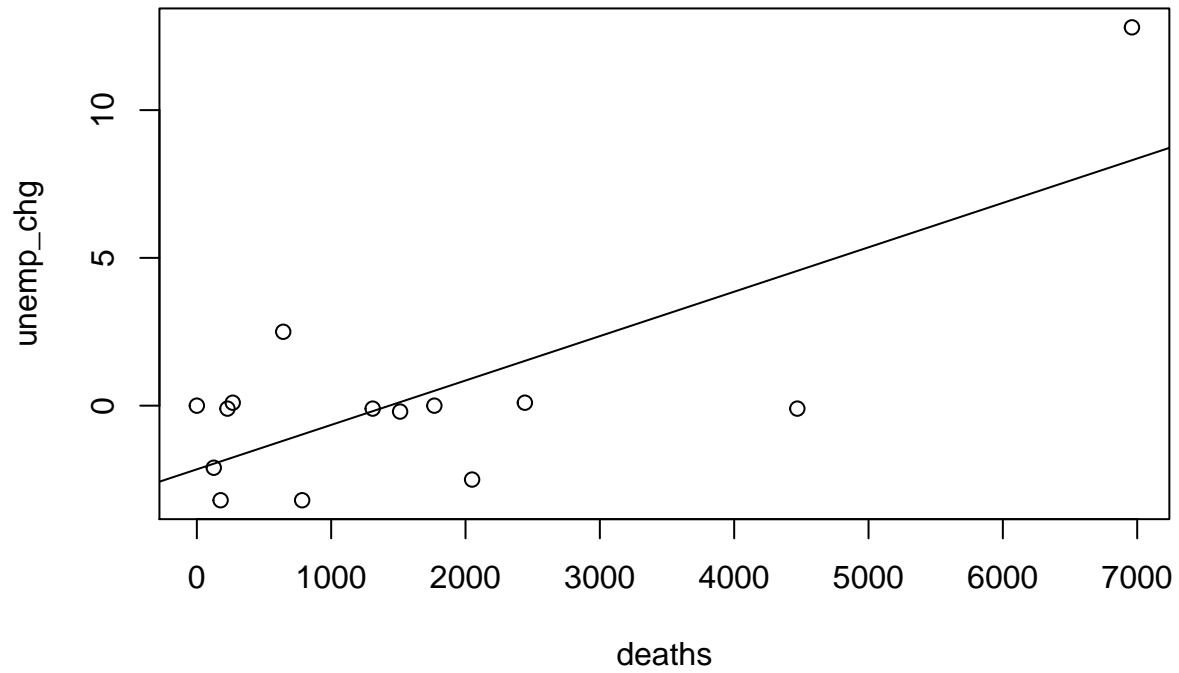# ID unemployment change vs deaths
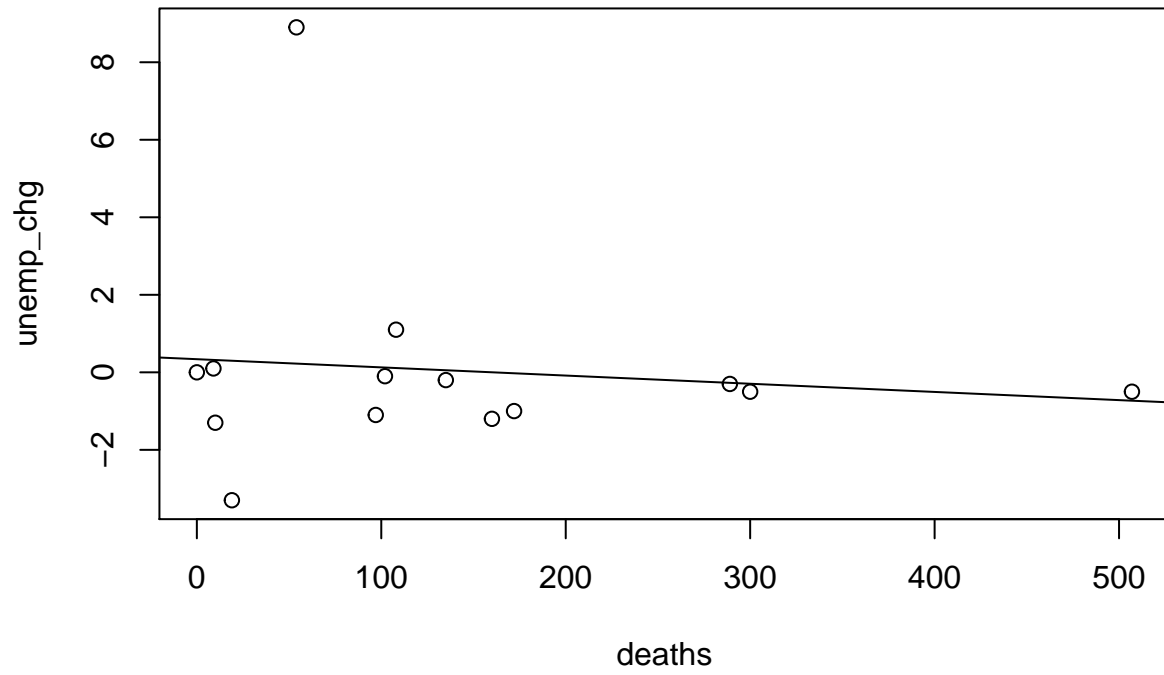
IL unemployment change vs cases

IL unemployment change vs deaths

IN unemployment change vs cases

IN unemployment change vs deaths

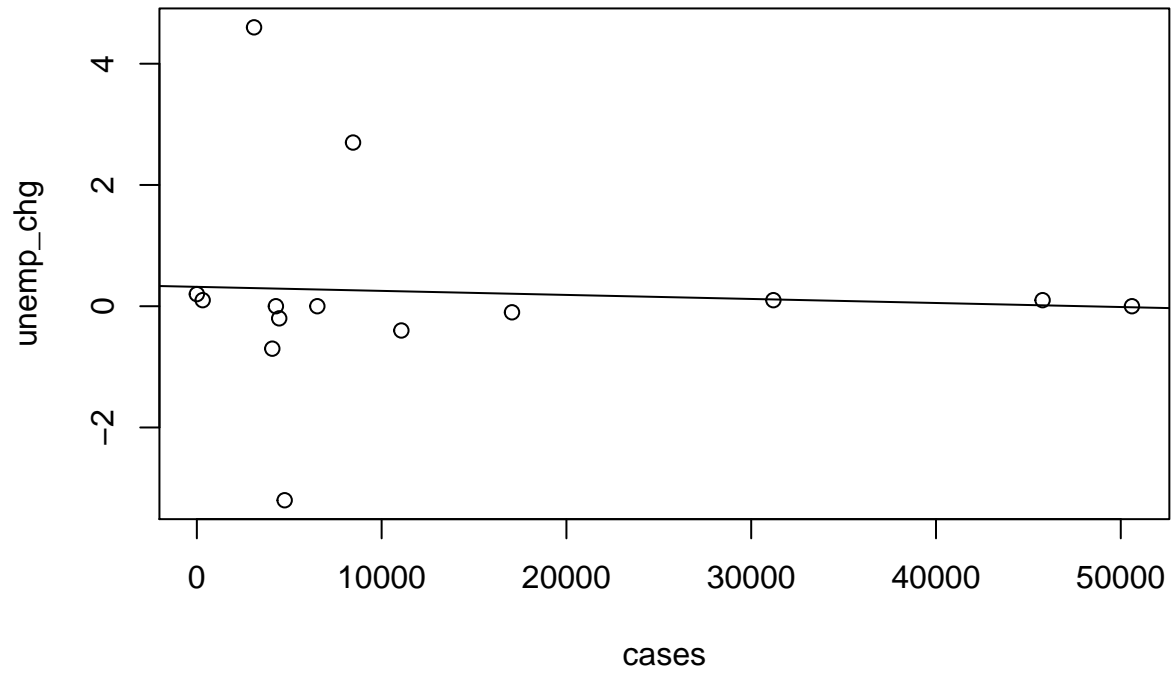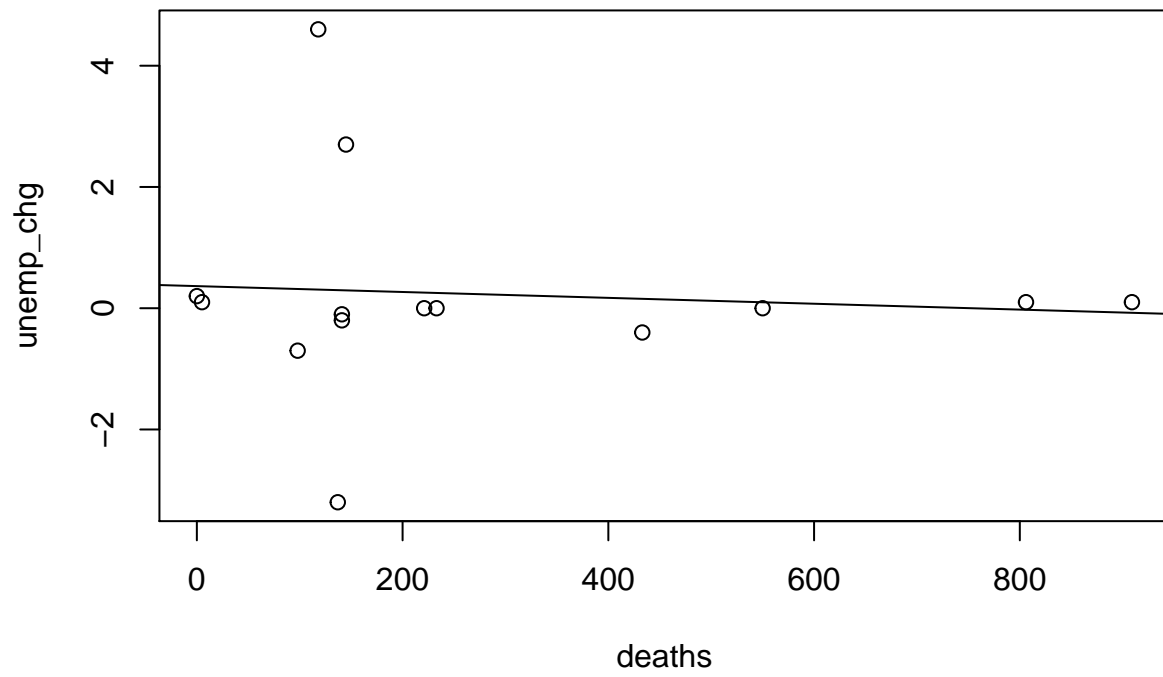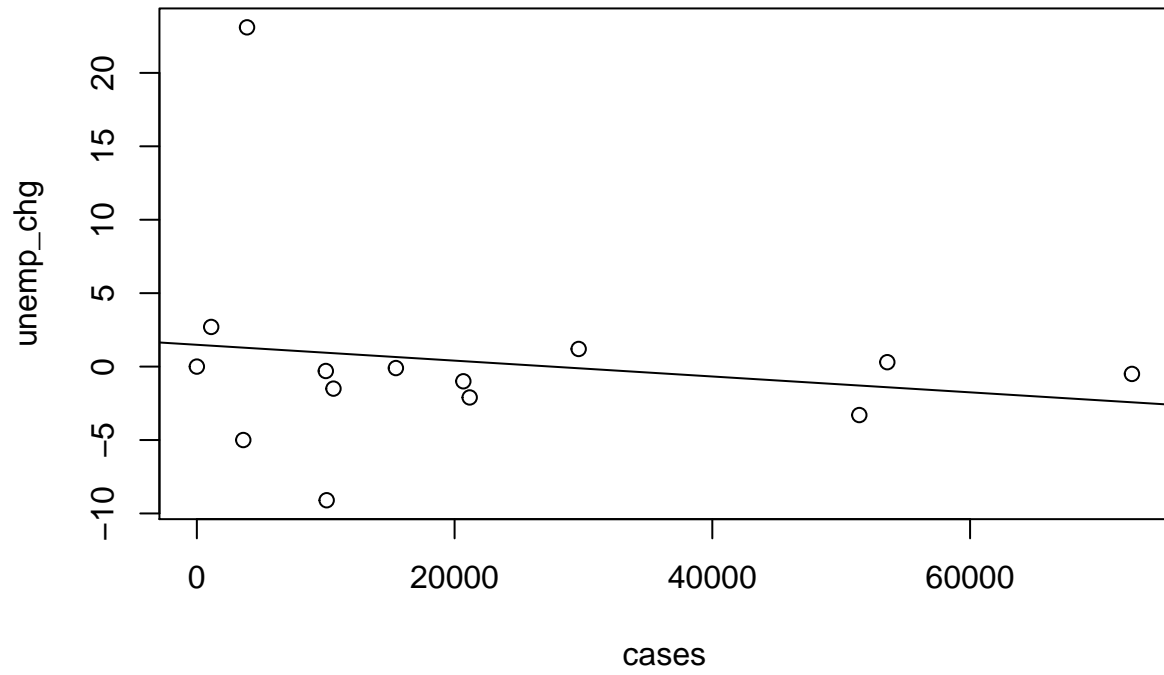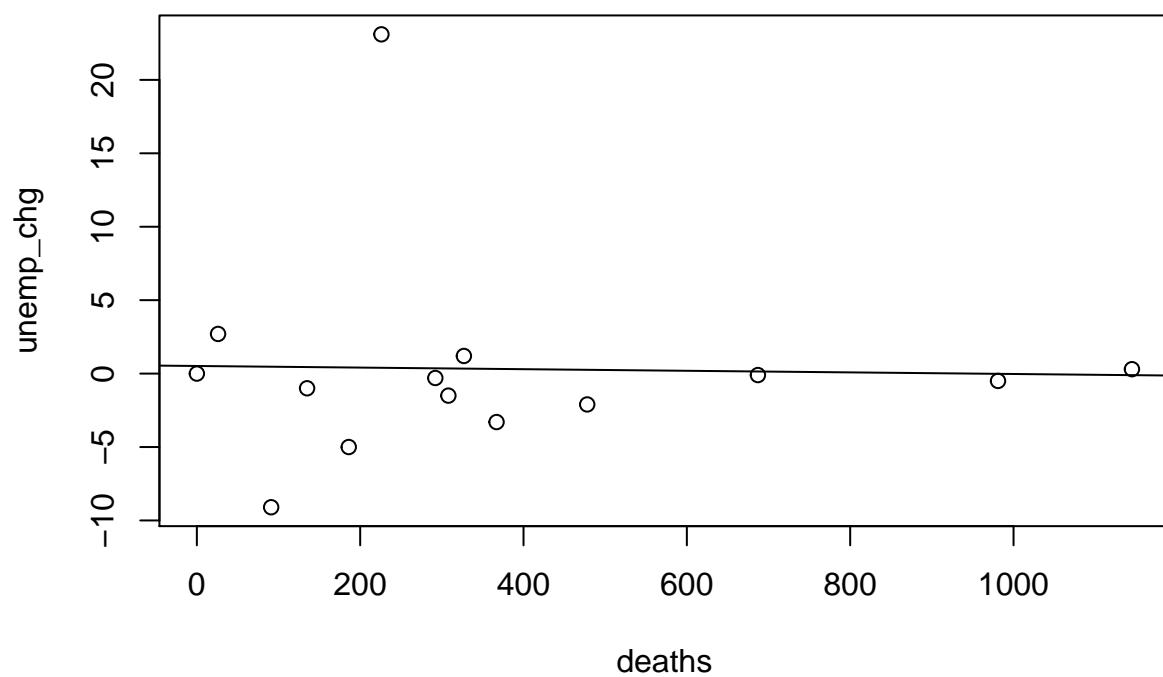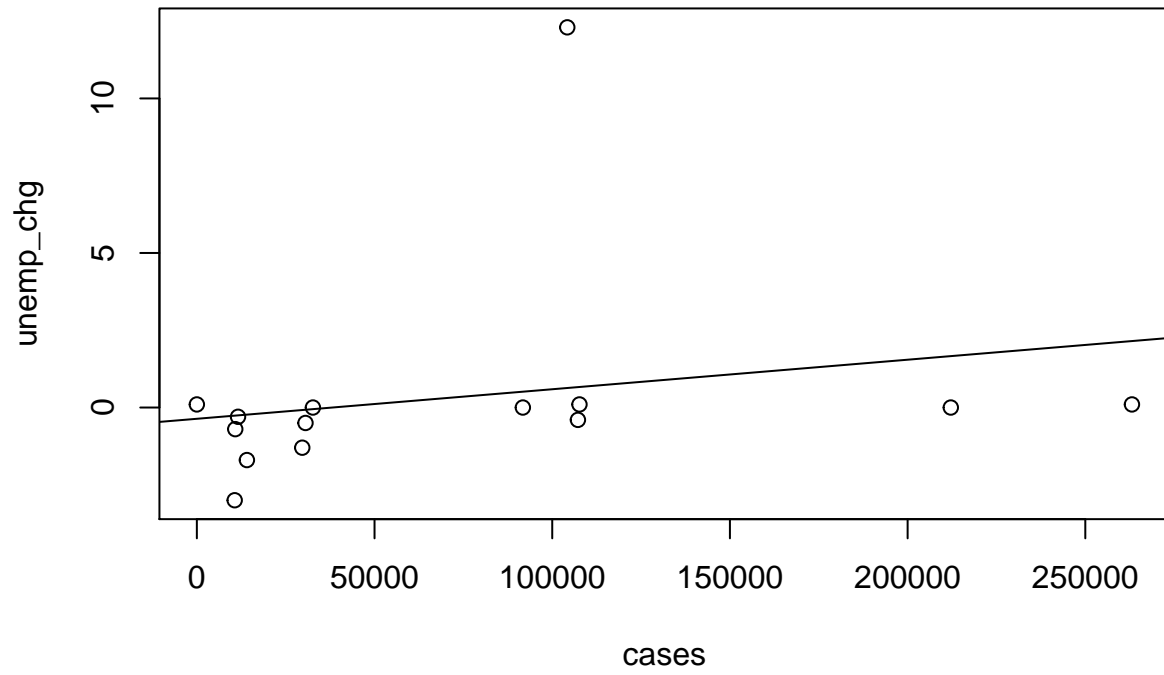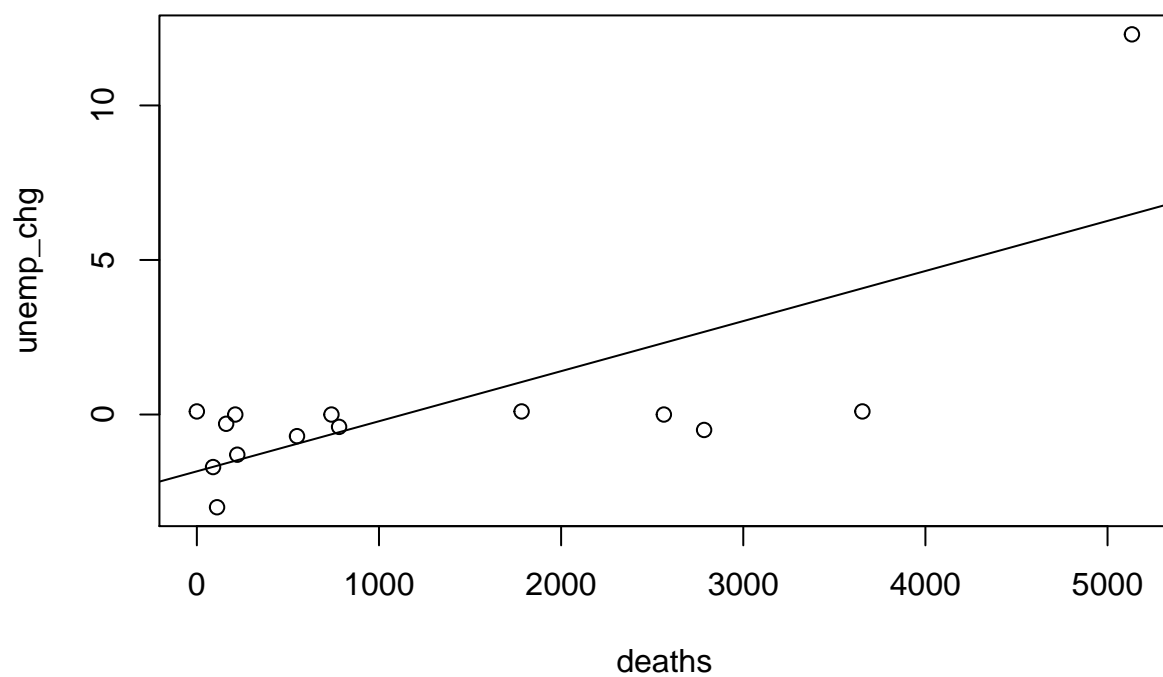# KS unemployment change vs cases

KS unemployment change vs deaths

KY unemployment change vs cases

KY unemployment change vs deaths

LA unemployment change vs cases

LA unemployment change vs deaths

MA unemployment change vs cases

MA unemployment change vs deaths

MD unemployment change vs cases

MD unemployment change vs deaths

ME unemployment change vs cases

ME unemployment change vs deaths

MI unemployment change vs cases

MI unemployment change vs deaths

MN unemployment change vs cases

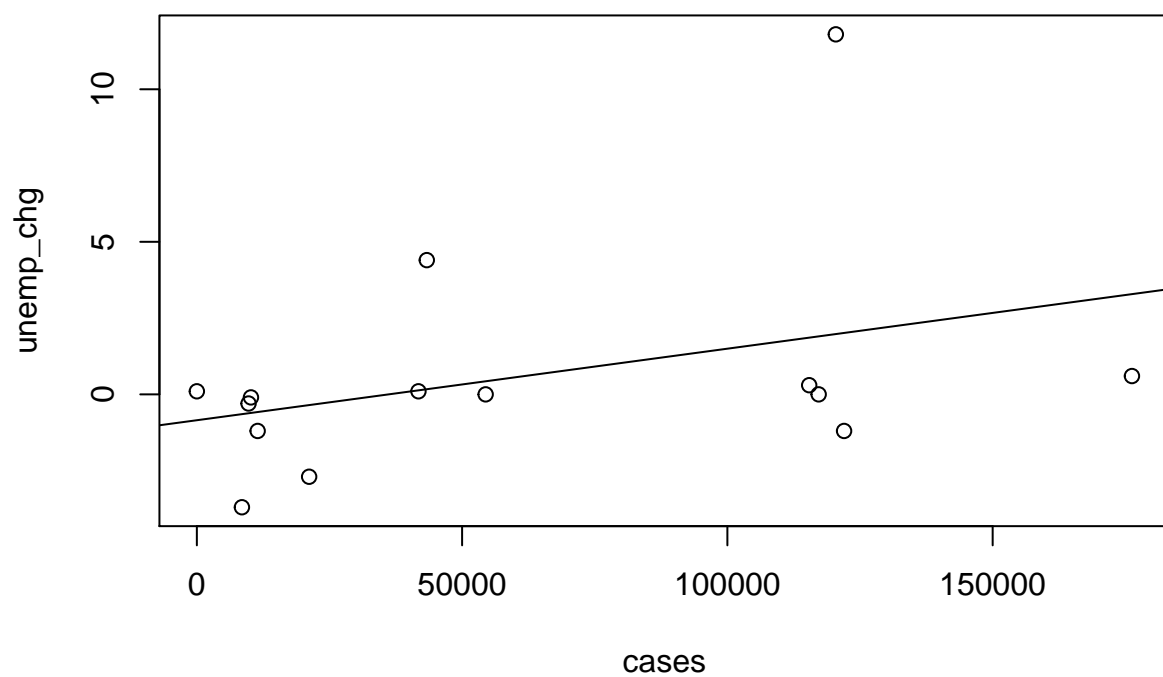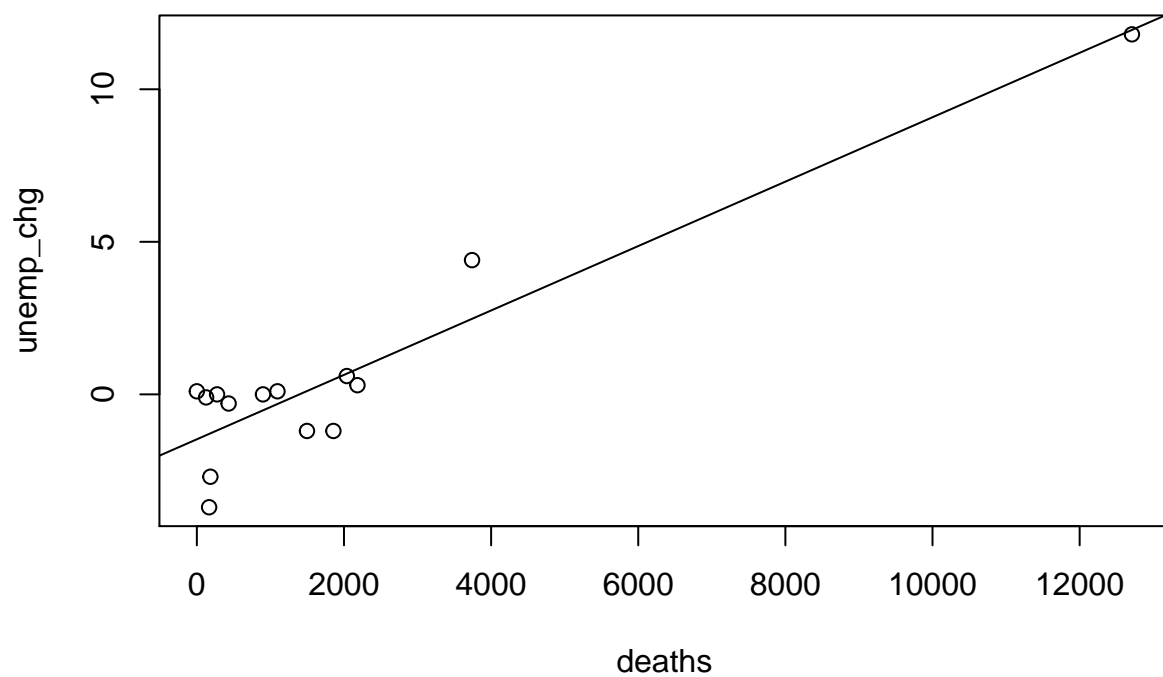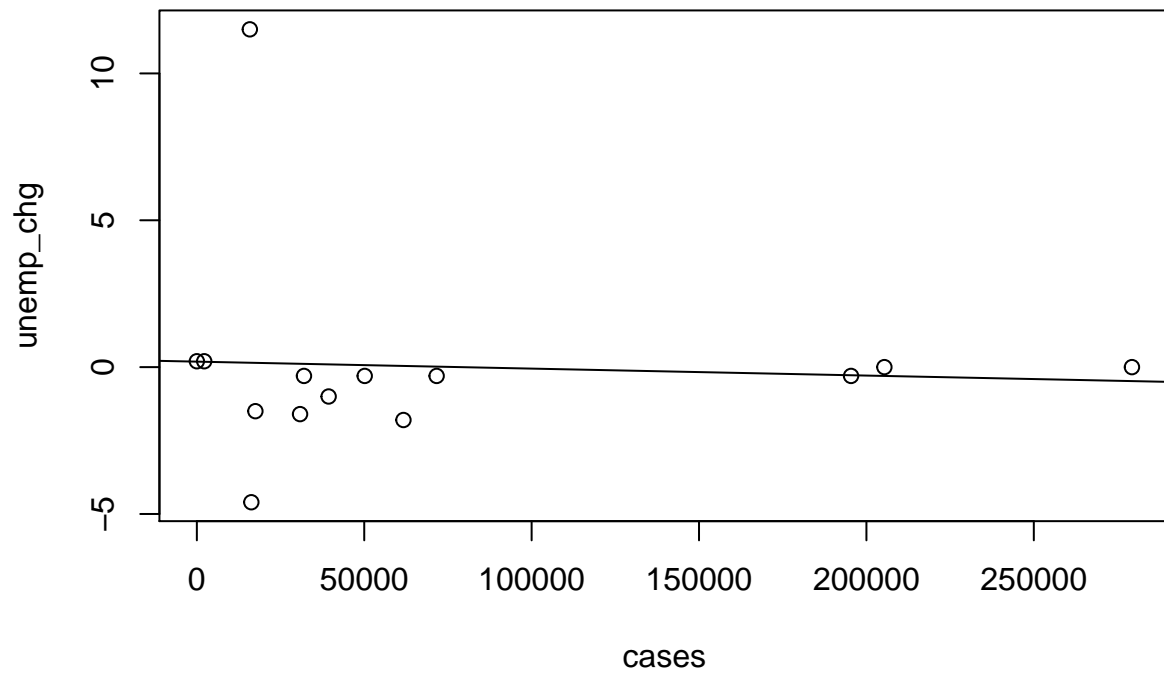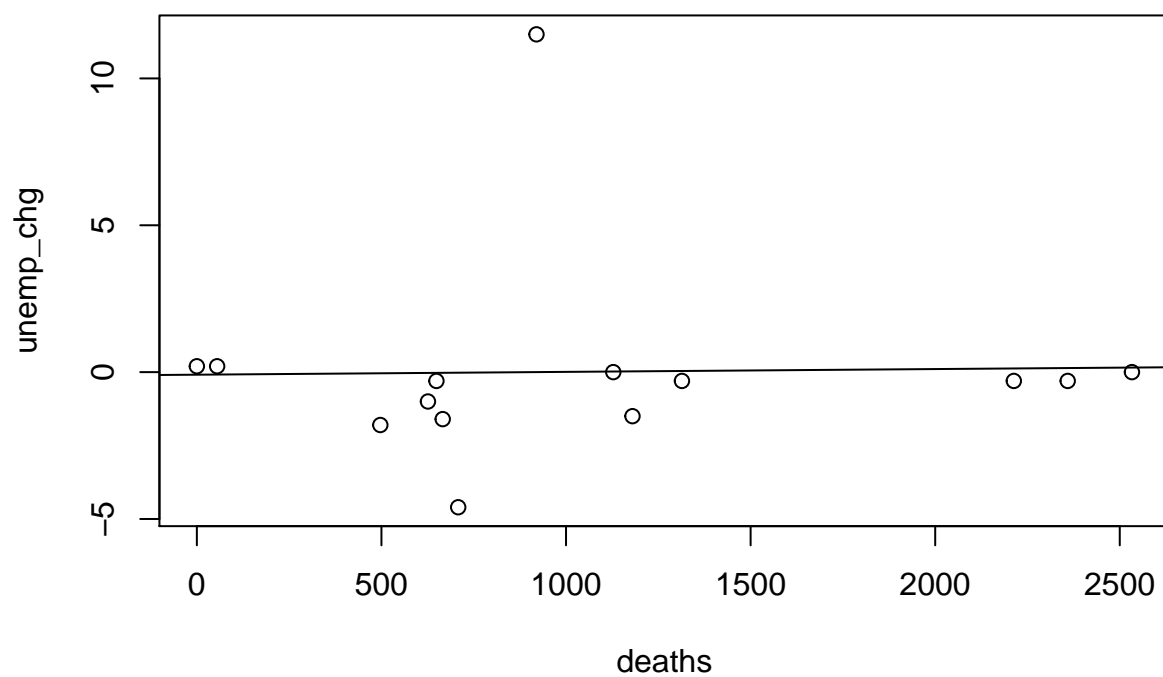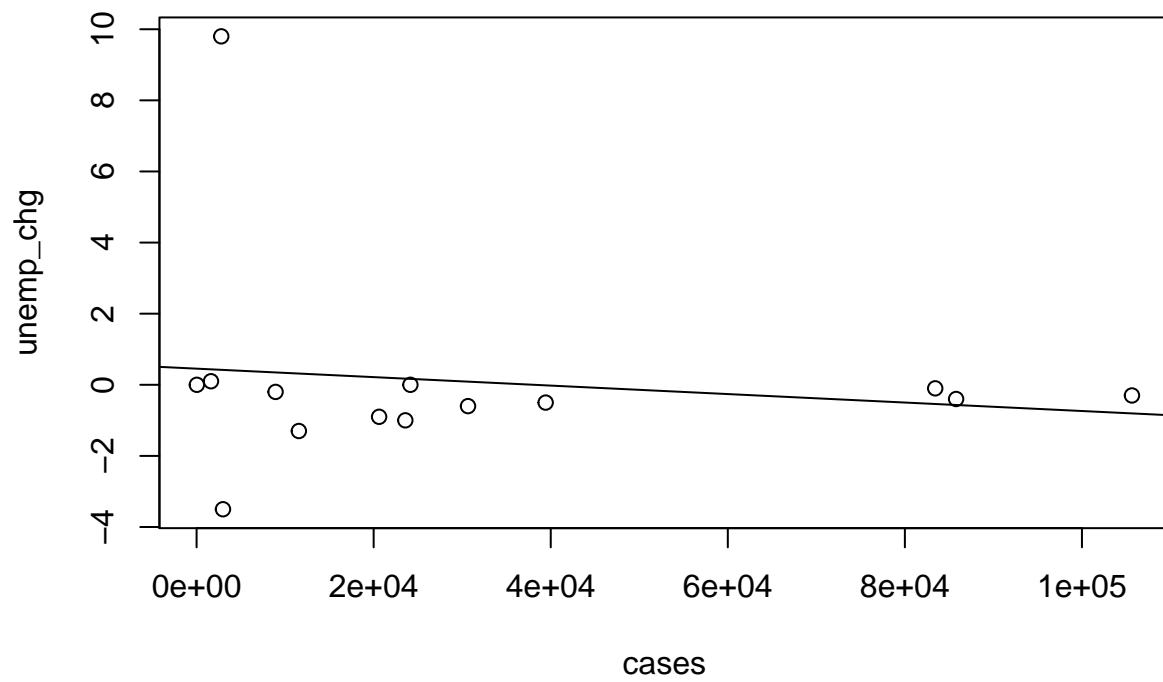# MN unemployment change vs deaths

MO unemployment change vs cases

cases

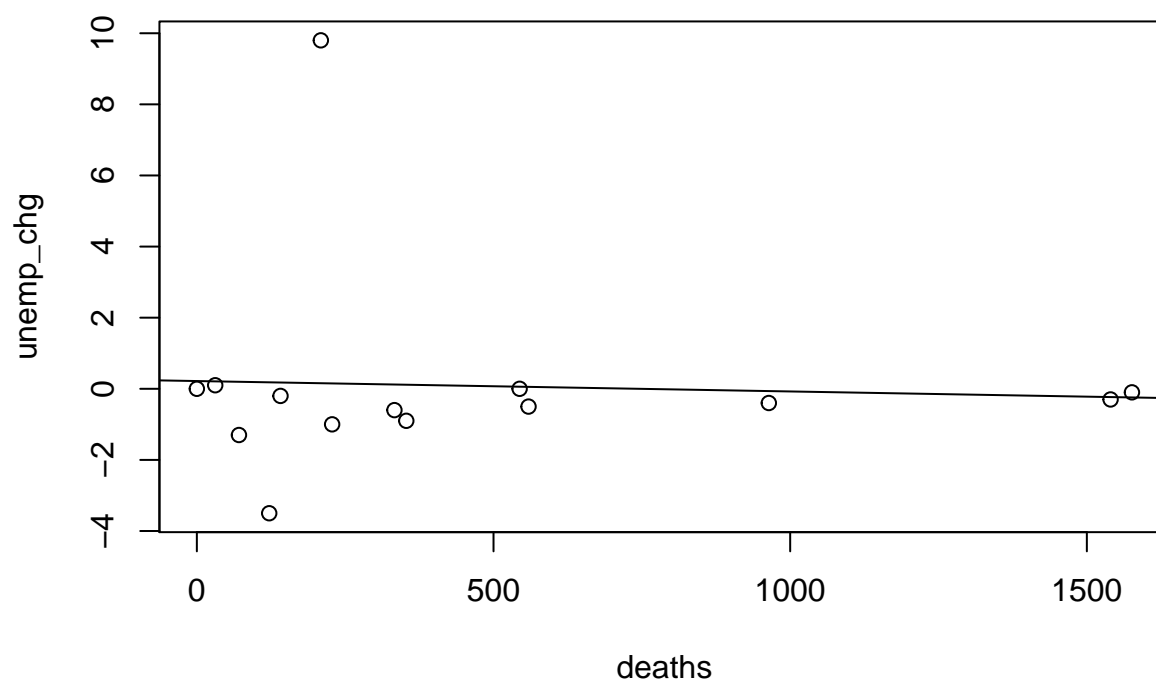# MO unemployment change vs deaths

MS unemployment change vs cases

MS unemployment change vs deaths
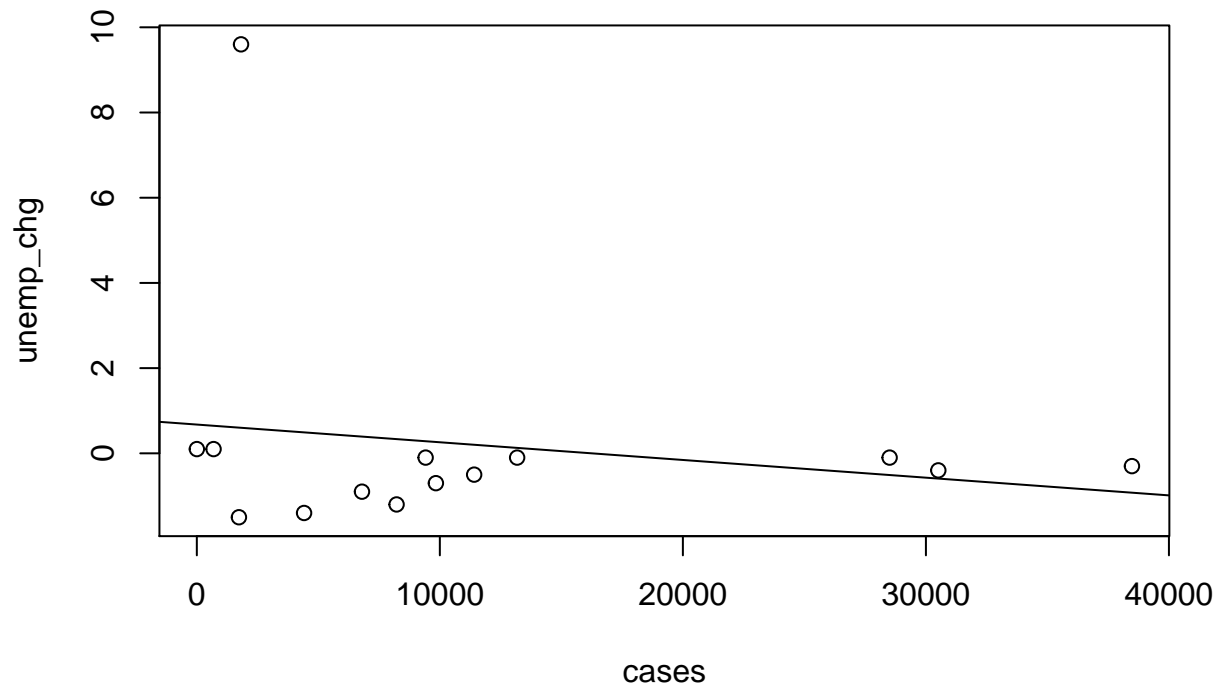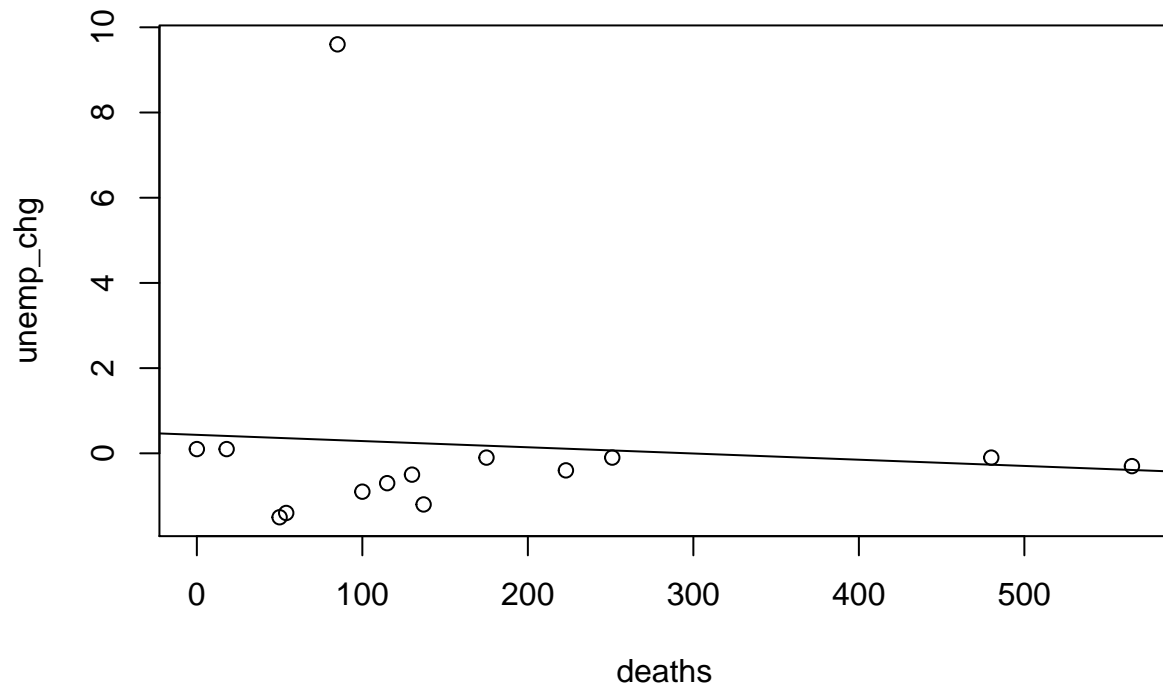
MT unemployment change vs cases

MT unemployment change vs deaths

NC unemployment change vs cases

NC unemployment change vs deaths

ND unemployment change vs cases

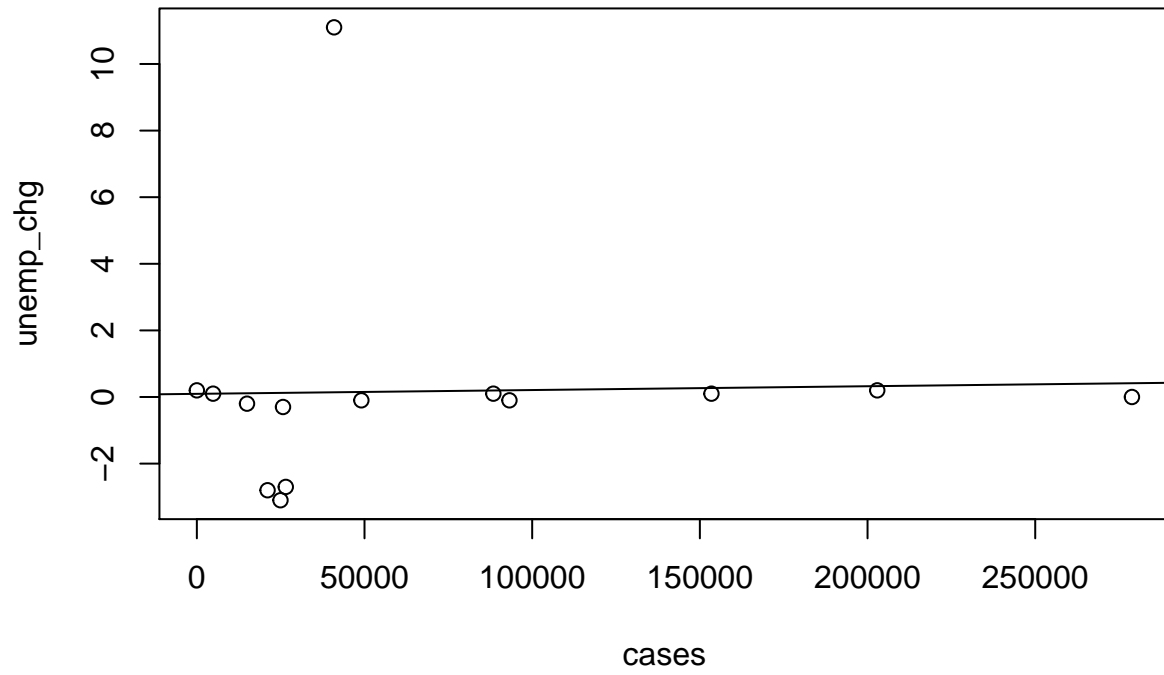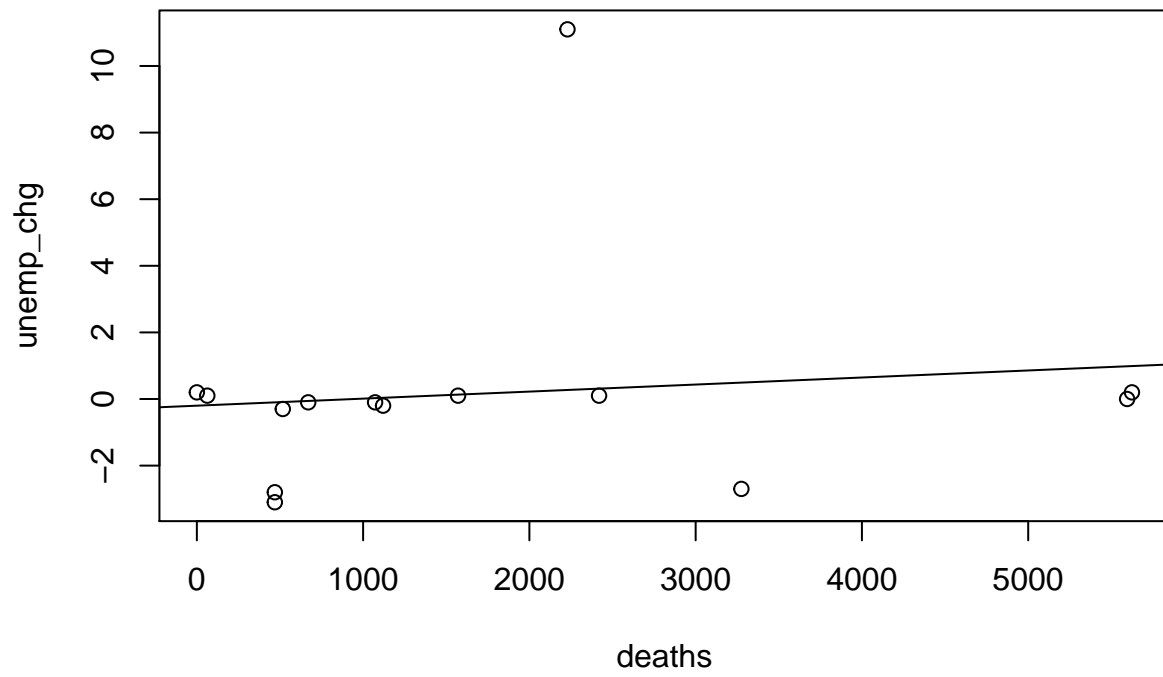# ND unemployment change vs deaths

NE unemployment change vs cases

NE unemployment change vs deaths

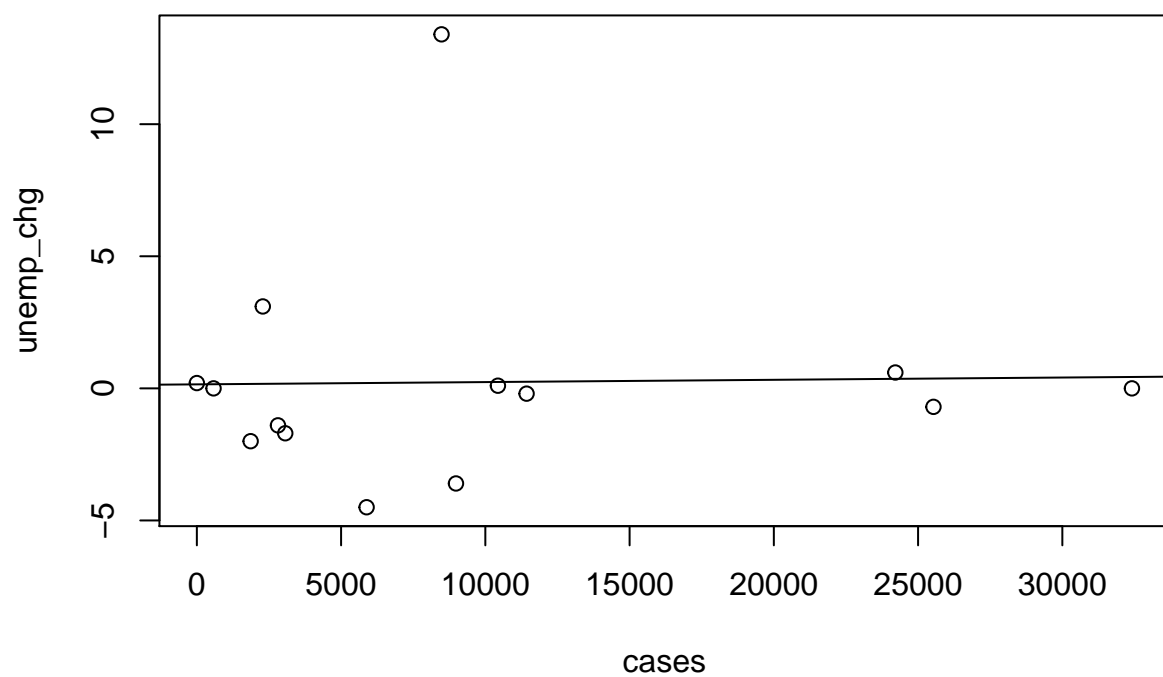# NH unemployment change vs cases

NH unemployment change vs deaths

NJ unemployment change vs cases

NJ unemployment change vs deaths

NM unemployment change vs cases

NM unemployment change vs deaths

NV unemployment change vs cases

NV unemployment change vs deaths

NY unemployment change vs cases

NY unemployment change vs deaths

NYC unemployment change vs cases

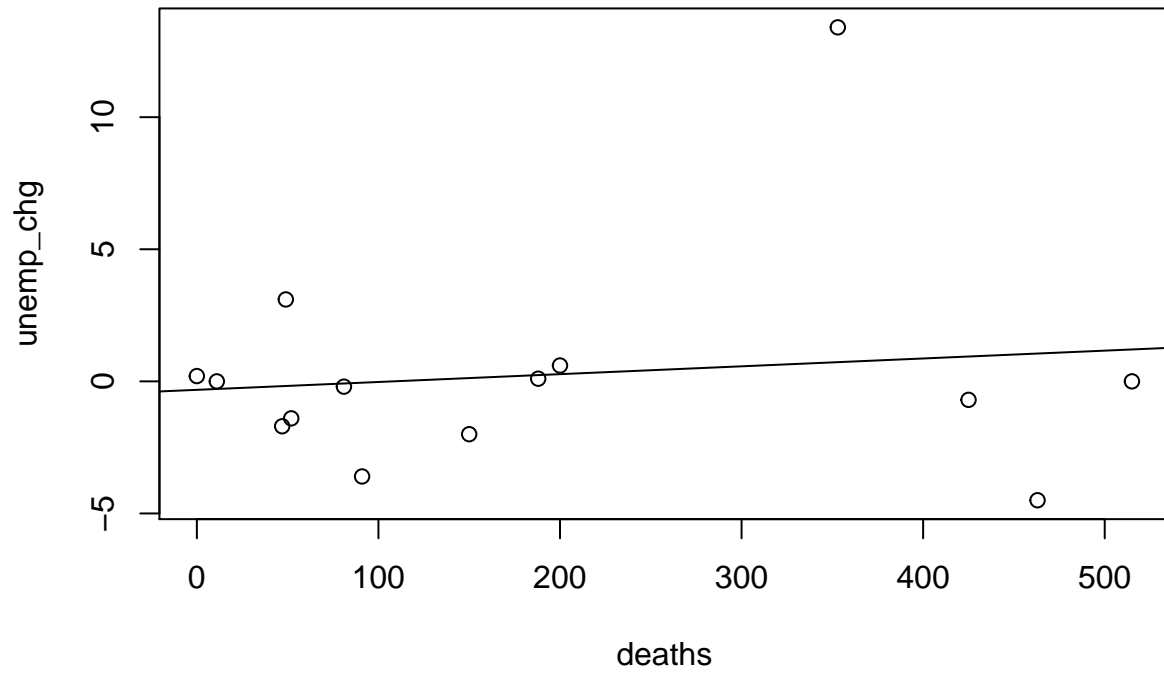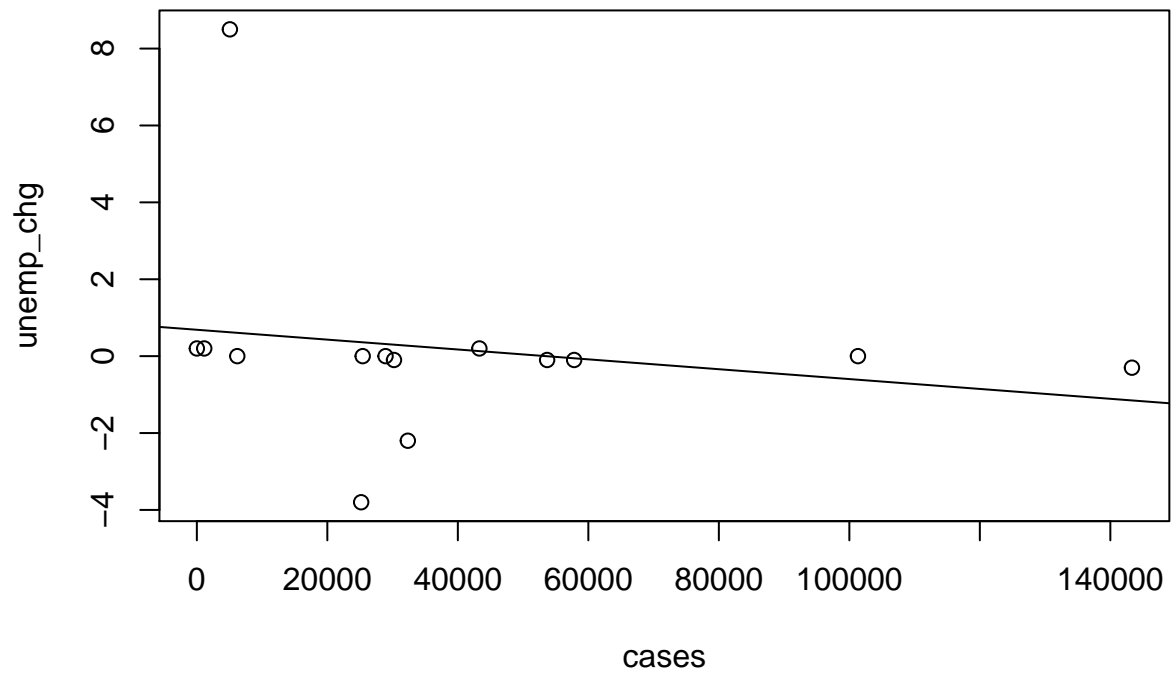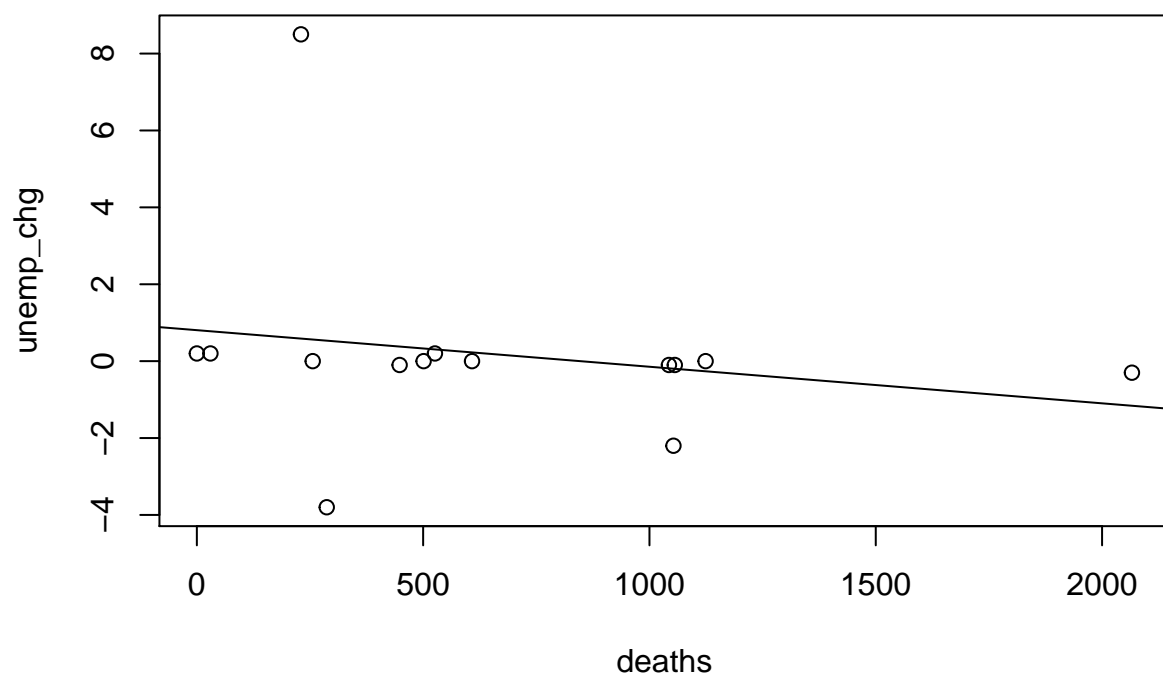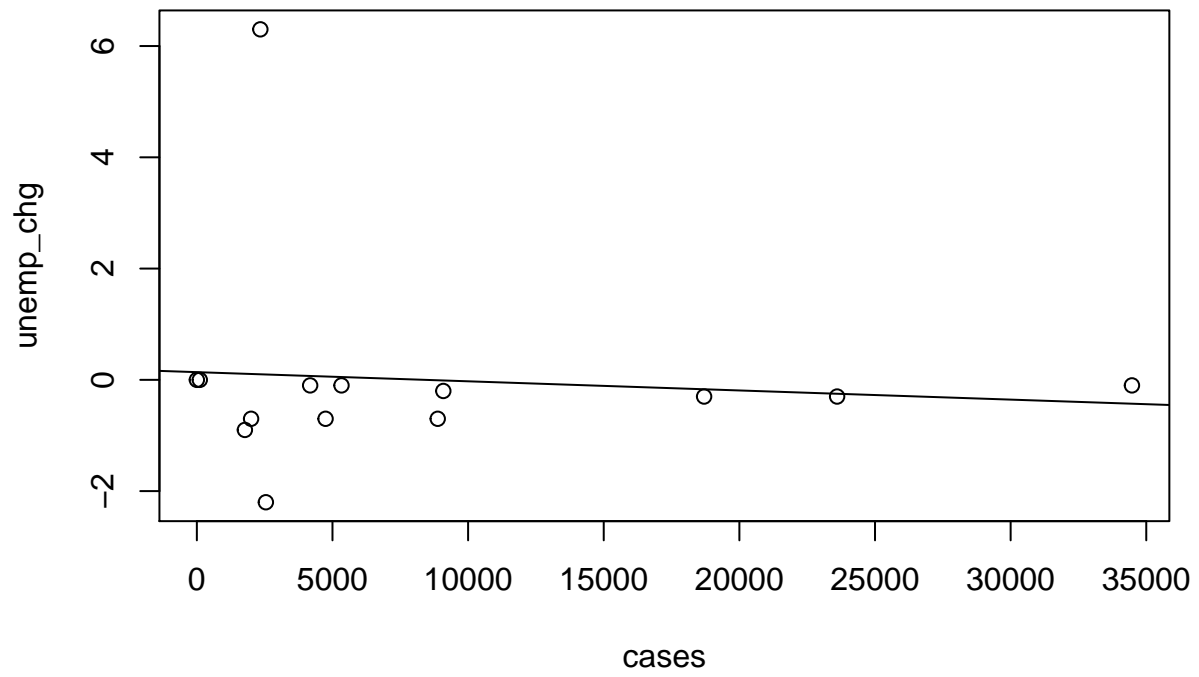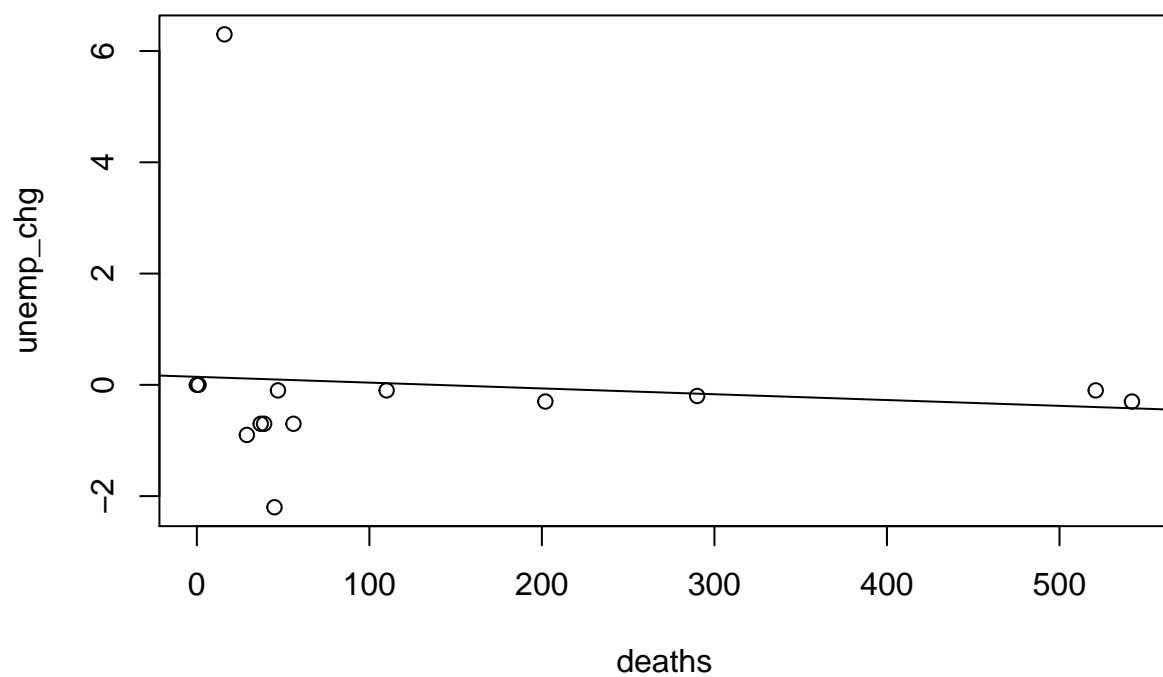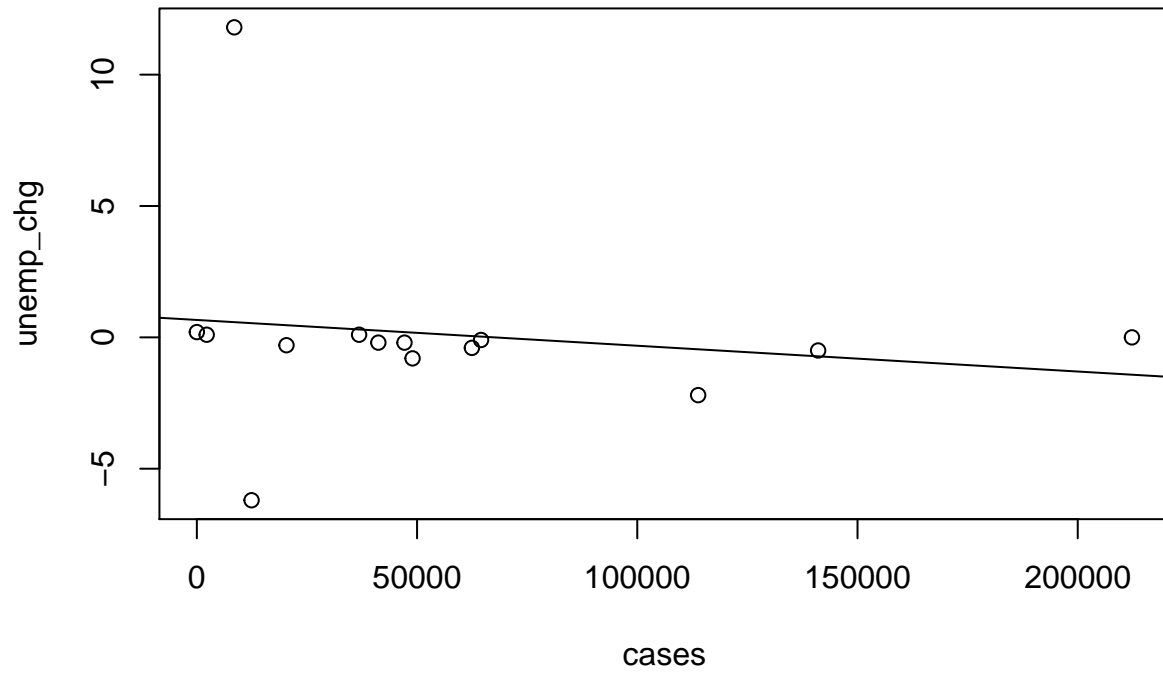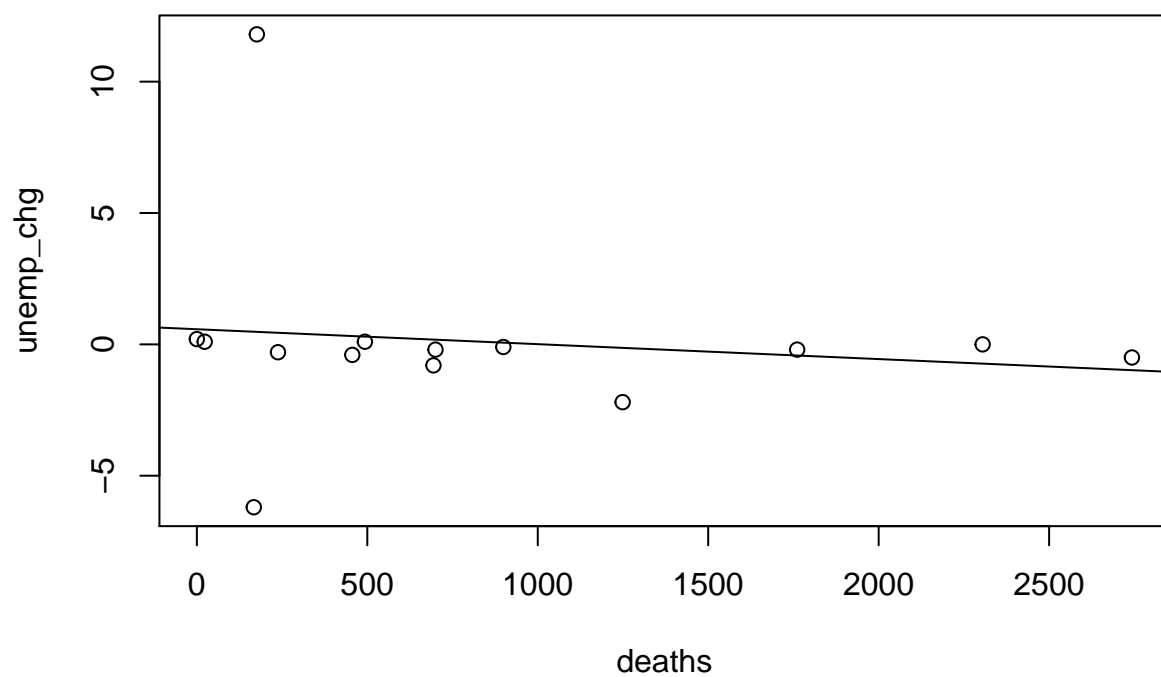# NYC unemployment change vs deaths

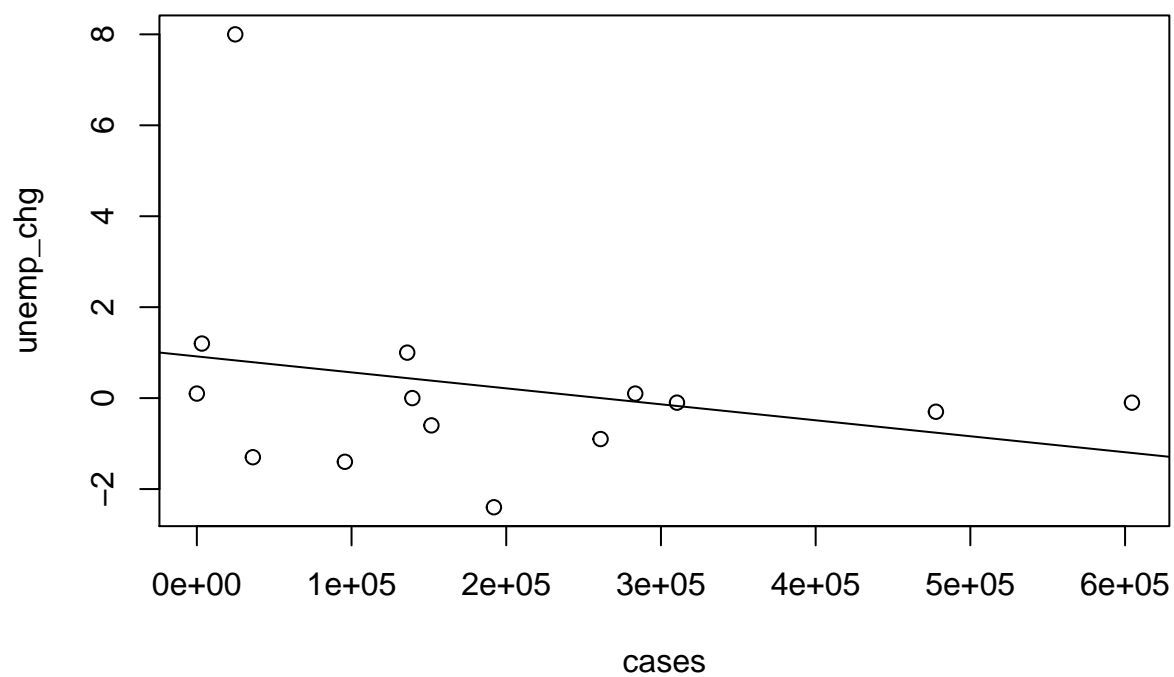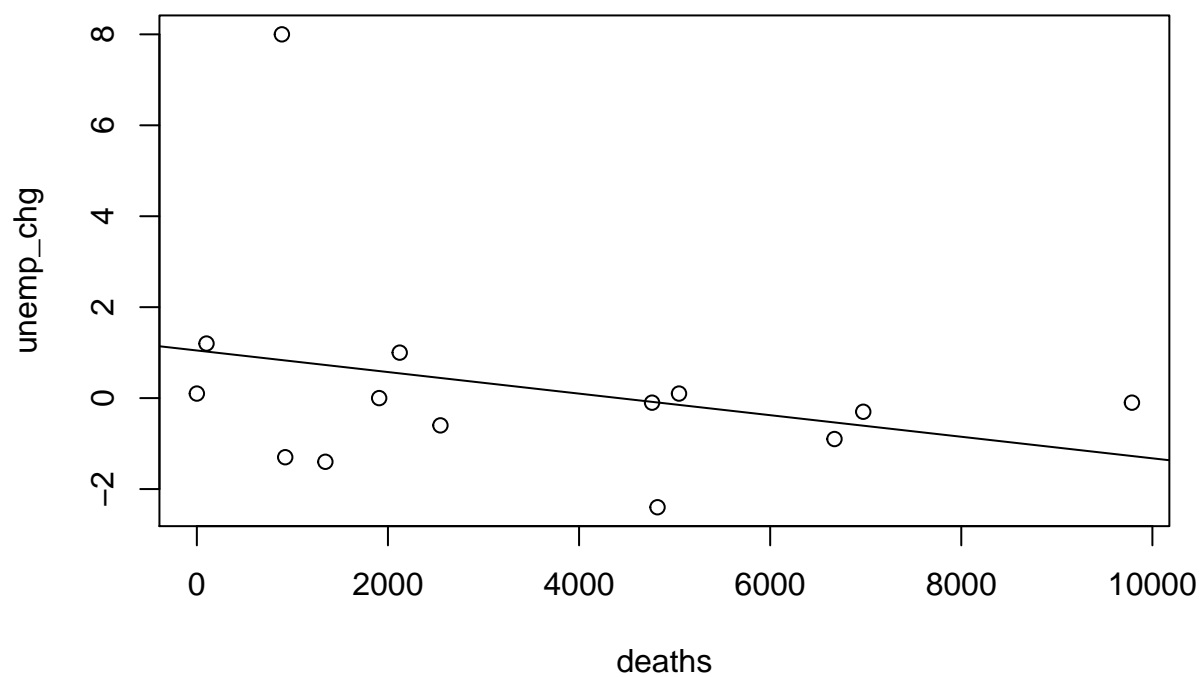# OH unemployment change vs cases

OH unemployment change vs deaths

OK unemployment change vs cases

OK unemployment change vs deaths

OR unemployment change vs cases

OR unemployment change vs deaths

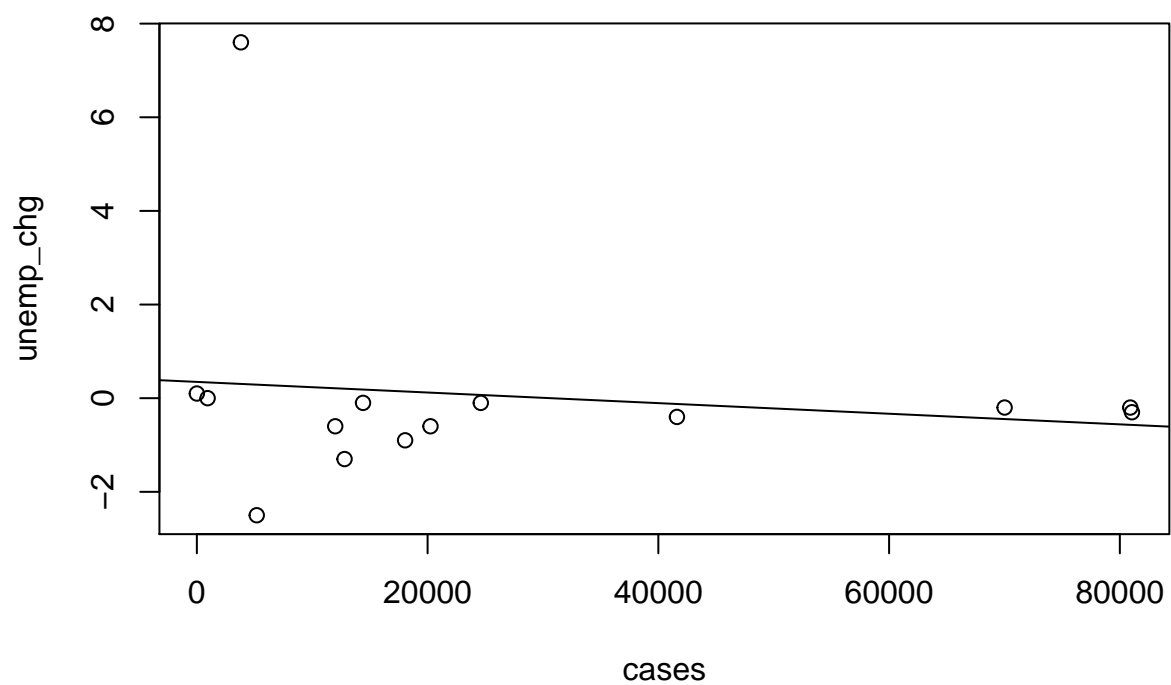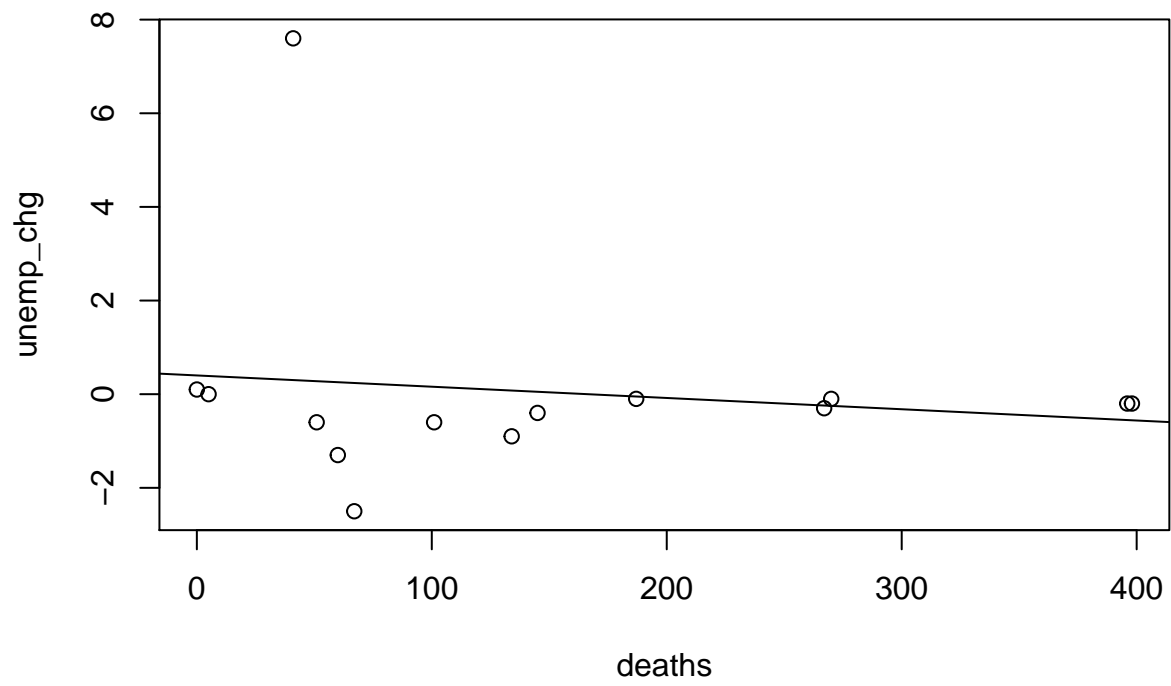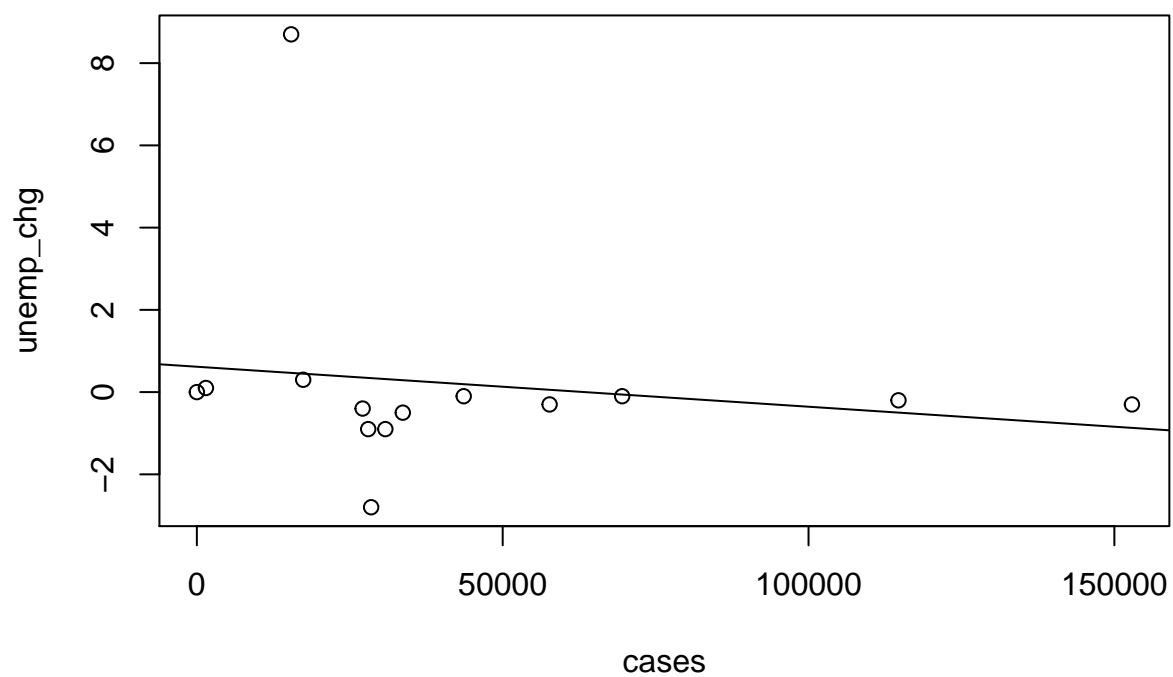**PA unemployment change vs cases**

PA unemployment change vs deaths

RI unemployment change vs cases

RI unemployment change vs deaths

SC unemployment change vs cases

SC unemployment change vs deaths

SD unemployment change vs cases

SD unemployment change vs deaths

TN unemployment change vs cases

TN unemployment change vs deaths

TX unemployment change vs cases

TX unemployment change vs deaths

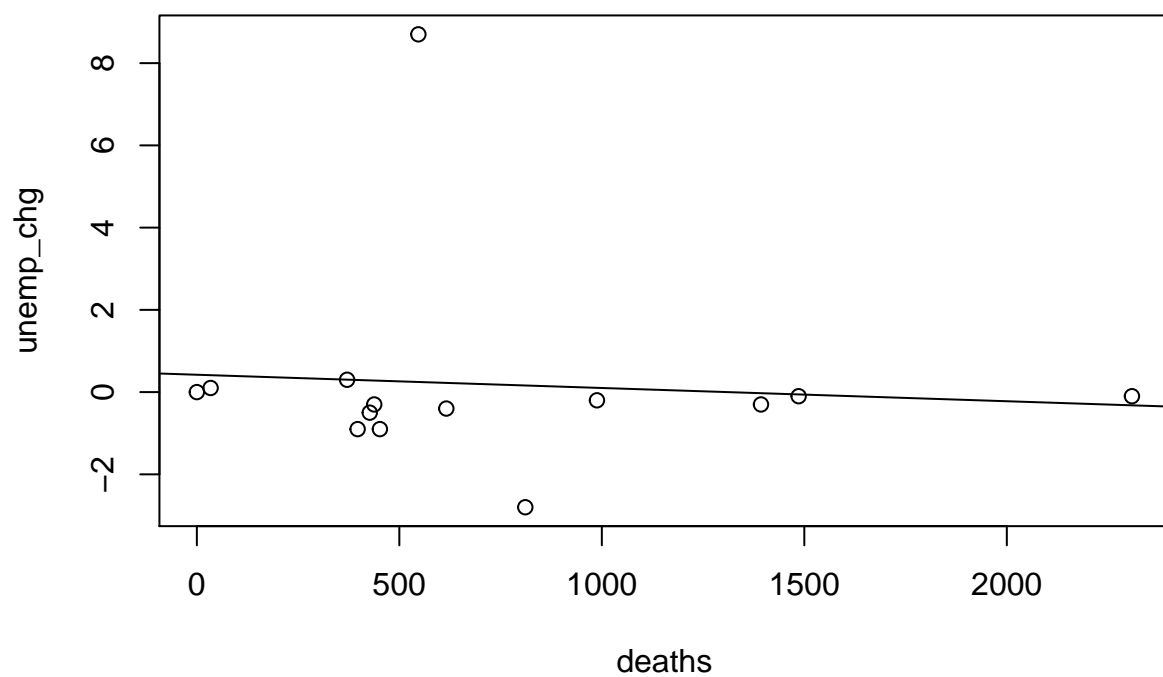# UT unemployment change vs cases

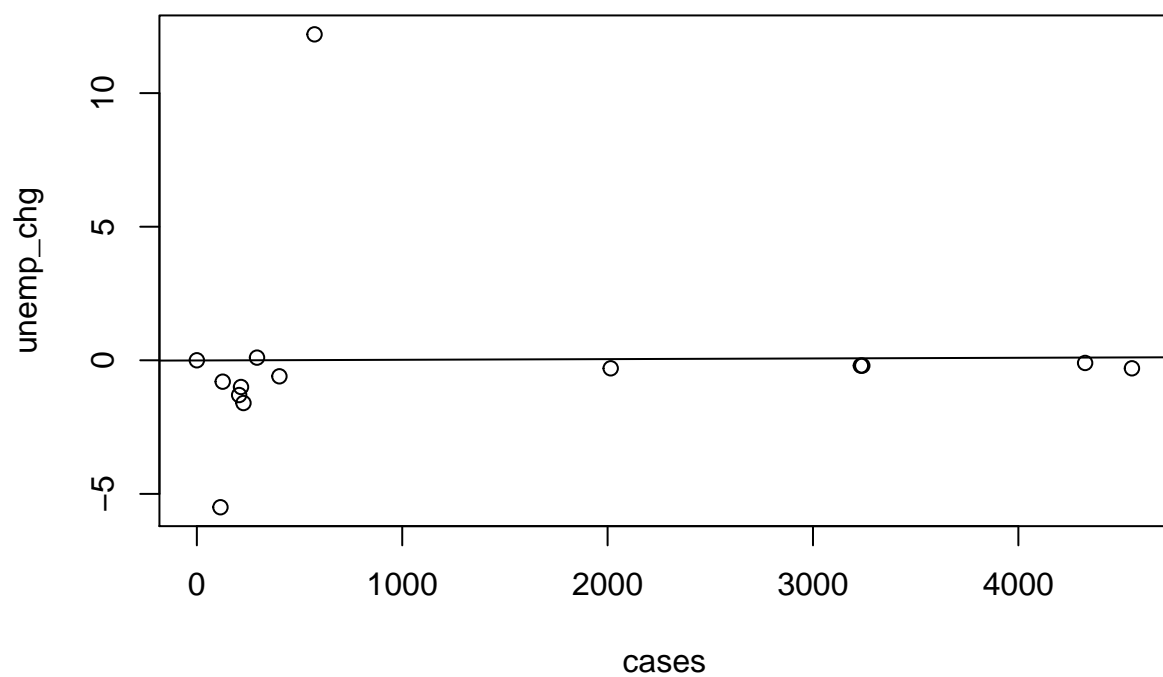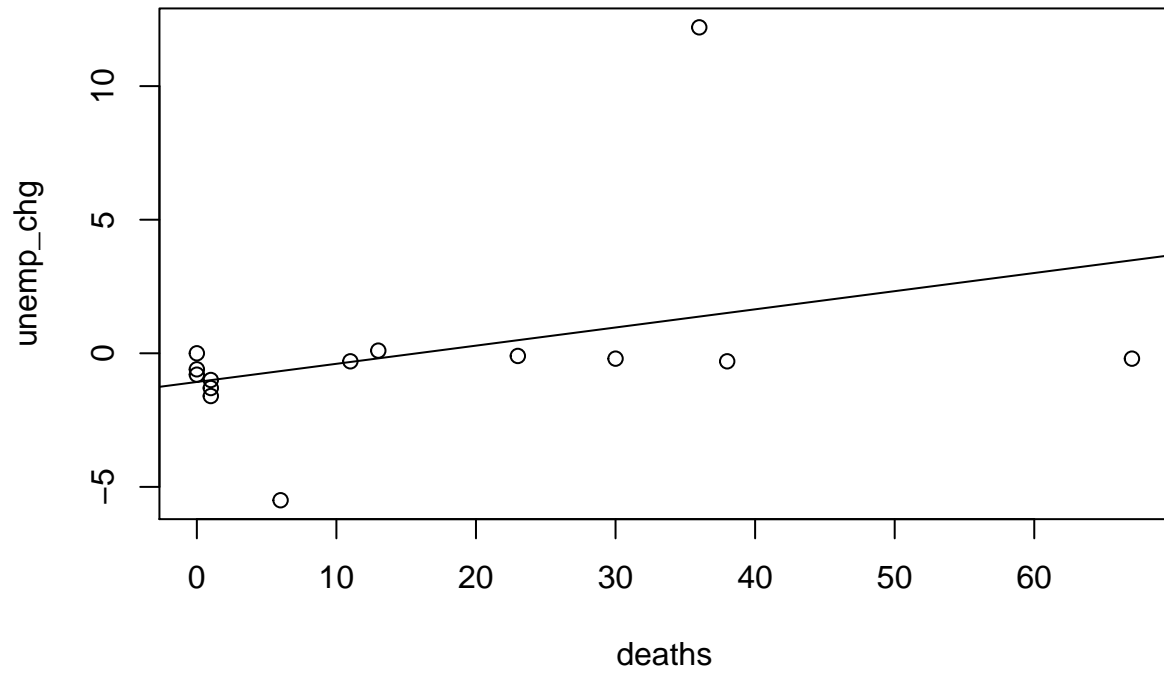# UT unemployment change vs deaths

VA unemployment change vs cases

VA unemployment change vs deaths

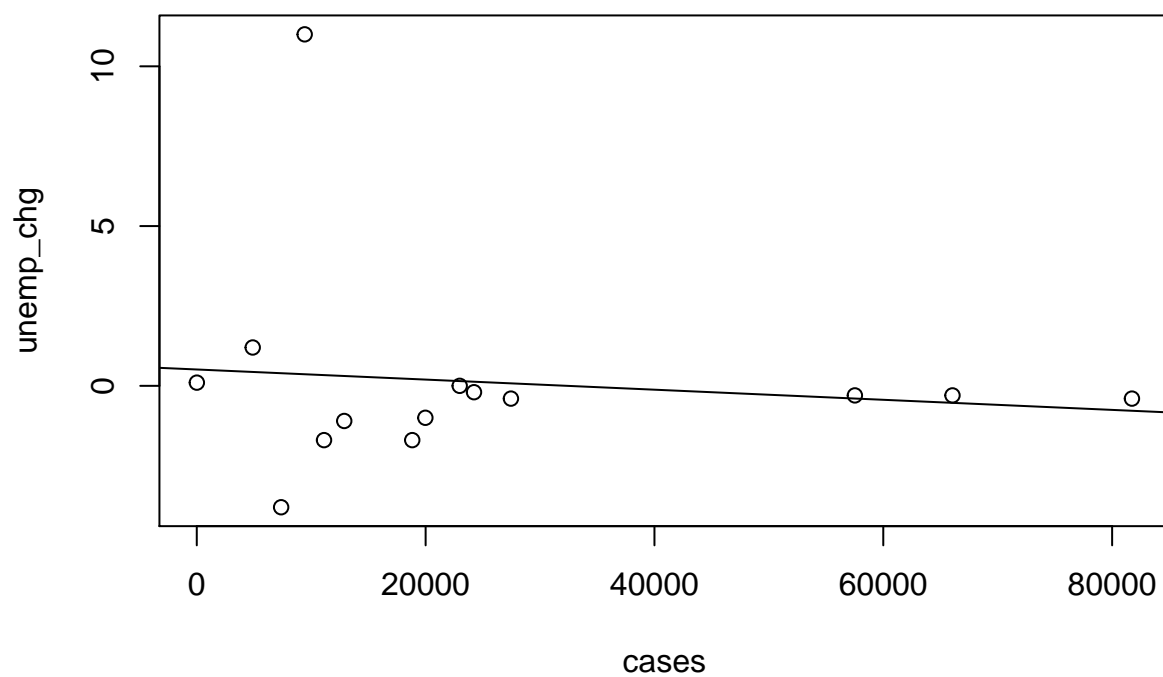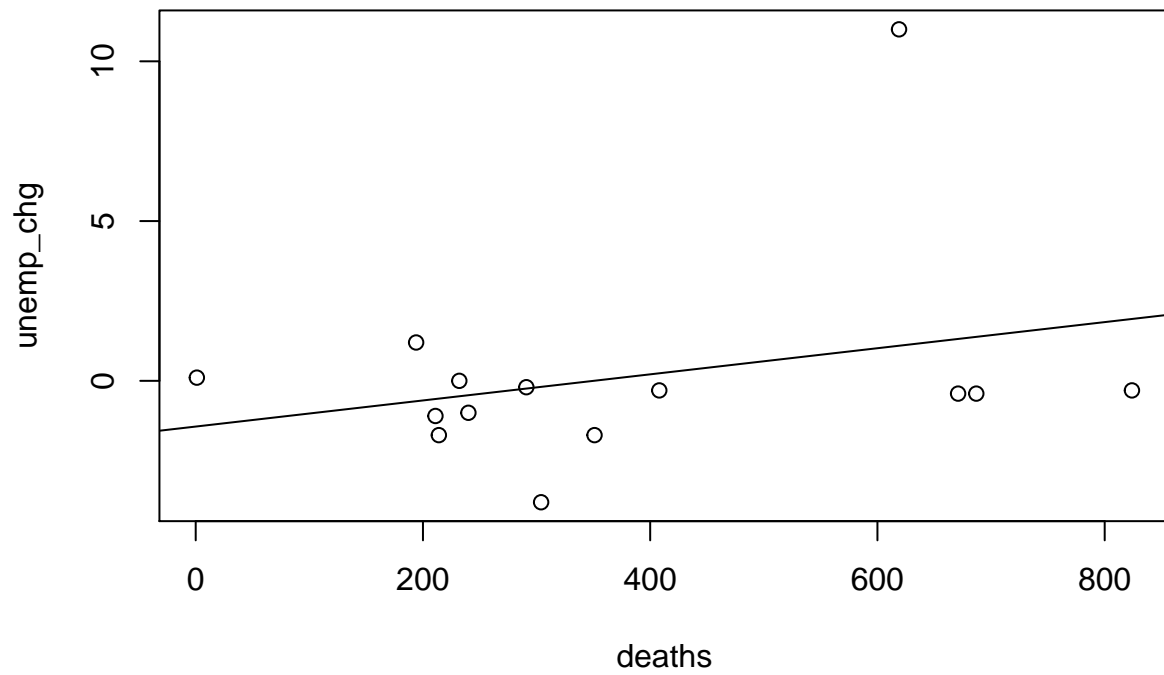# VT unemployment change vs cases

VT unemployment change vs deaths

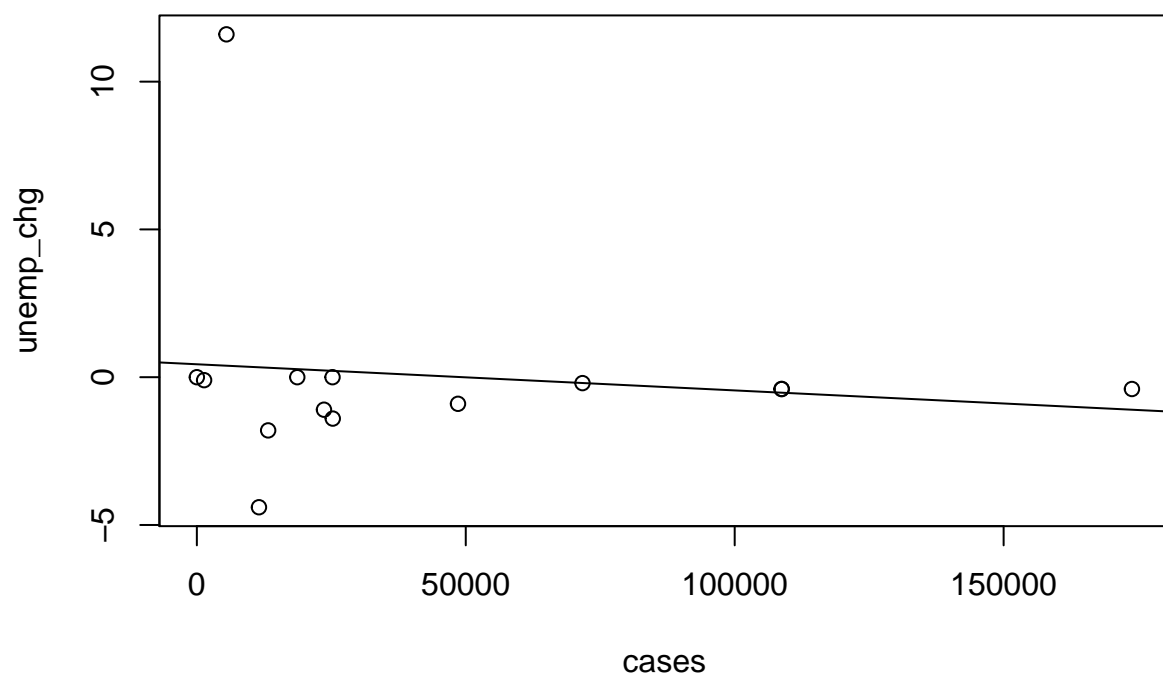# WA unemployment change vs cases

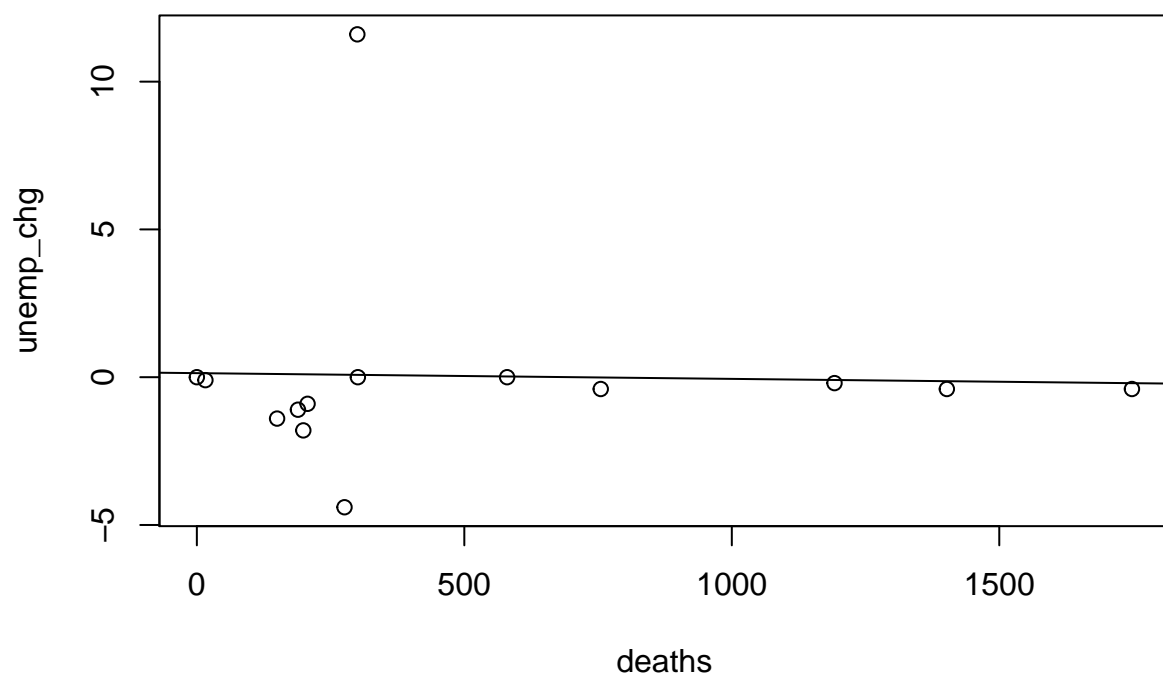# WA unemployment change vs deaths

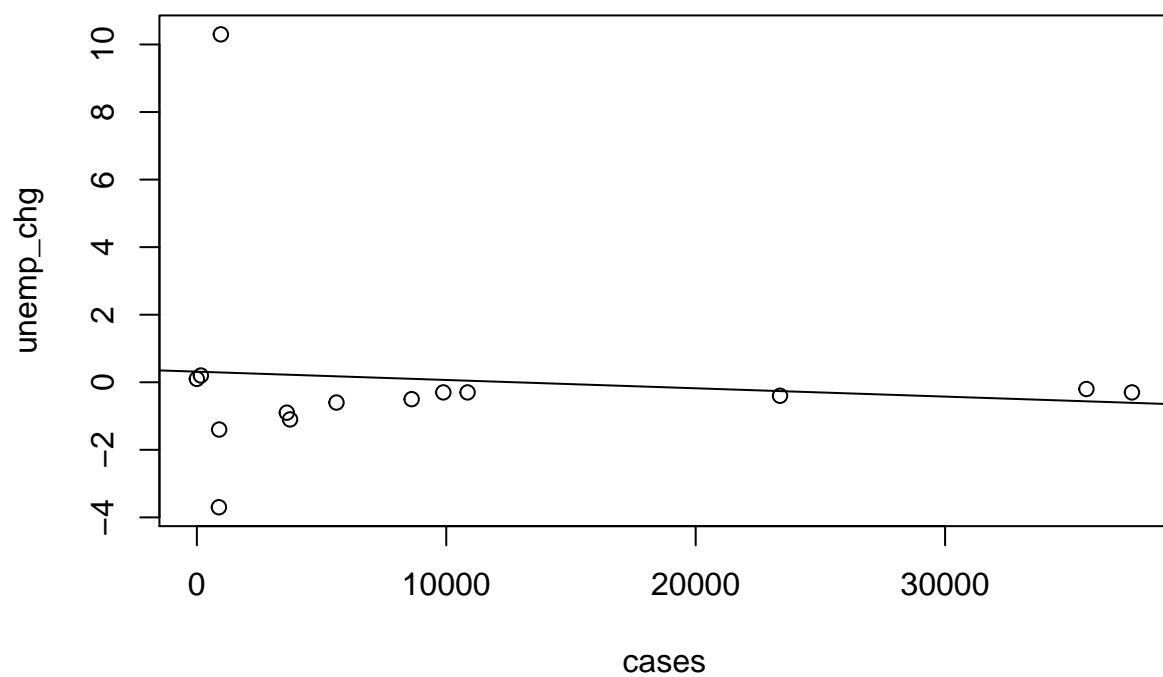# WI unemployment change vs cases
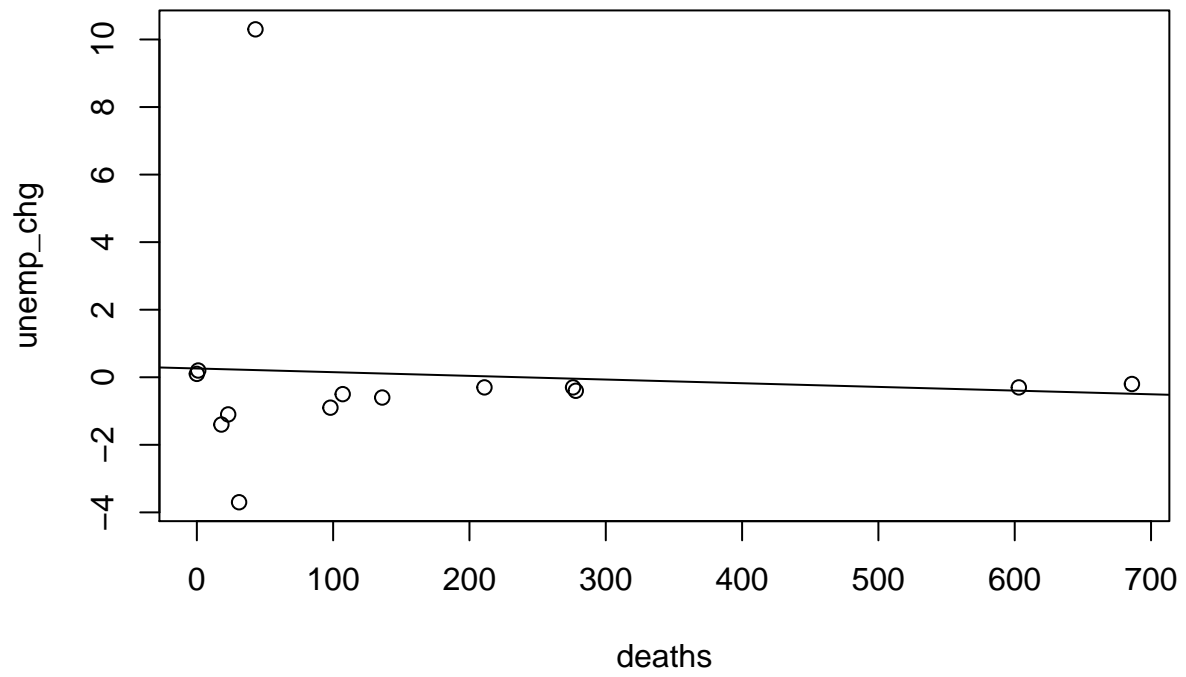
WI unemployment change vs deaths
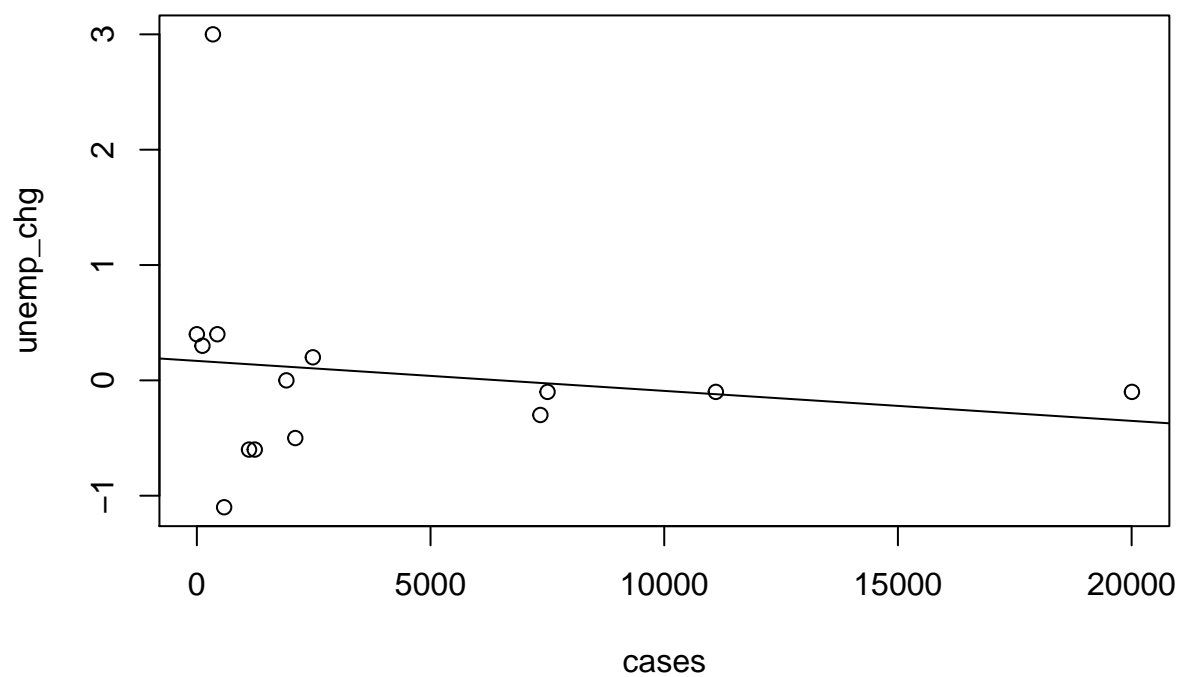
WV unemployment change vs cases

WV unemployment change vs deaths

WY unemployment change vs cases

WY unemployment change vs deaths