Daf_proj

*

9/15/2021

EFFECT OF DISASTERS ON THE STOCK OF INSURANCE COMPANIES

Short Description:-

In this project we are considering four datasets, the first dataset is "Disasters", in this dataset we are considering 5 disasters - Fire,Flood,Hurricane,Severe Storms and Snow. The second,third and fourth datasets are Stocks data for top 3 listed insurance companies - AllState,Progressive and Travellers. So taking stocks data into consideration we are analyzing how disasters effected the stocks prices from year 2000 -2020

Source:-

FEMA(Federal Emergency Management Agency) Disaster Declarations Summary is a summarized dataset describing all federally declared disasters. This dataset lists all official FEMA Disaster Declarations, beginning with the first disaster declaration in 1953 and features all three disaster declaration types: major disaster, emergency, and fire management assistance.

'https://www.fema.gov/api/open/v2/DisasterDeclarationsSummaries.csv'

Yahoo Finance - This is a media property that is part of Yahoo! network.It provides financial news, data and commentary including stock quotes, press releases, financial reports, and original content.

'https://query1.finance.yahoo.com/v7/finance/download/ALL?period1=944006400&period2=1631664000&interval=1d&events=history&includeAdjustedClose=true'

'https://query1.finance.yahoo.com/v7/finance/download/PGR?period1=944006400&period2=1631664000&interval=1d&events=history&includeAdjustedClose=true'

'https://query1.finance.yahoo.com/v7/finance/download/TRV?period1=944006400&period2=1631664000&interval=1d&events=history&includeAdjustedClose=true'

For this project, we need these three libraries, "Tidyverse" and "Janitor" for preprocessing and cleaning the data, "Ggrepel" for plotting the data.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

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

```
library(janitor)
```

```
##
## Attaching package: 'janitor'
```

```
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```
library(ggrepel)
```

Loading Disasters dataset

```
disaster_data <- read_csv('https://www.fema.gov/api/open/v2/DisasterDeclarationsSummaries.csv')
```

```
## Rows: 62591 Columns: 23
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (10): femaDeclarationString, state, declarationType, incidentType, decl...
## dbl   (8): disasterNumber, fyDeclared, ihProgramDeclared, iaProgramDeclared,...
## dttm  (5): declarationDate, incidentBeginDate, incidentEndDate, disasterClos...
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Loading AllState Stock dataset

```
allstate_data <- read_csv('https://query1.finance.yahoo.com/v7/finance/download/ALL?period1=944006400&p
```

```
## Rows: 5482 Columns: 7
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (6): Open, High, Low, Close, Adj Close, Volume
## date (1): Date
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
allstate_data <- subset(allstate_data, select = c("Date","Open","High","Low","Close"))

colnames(allstate_data) <- c("Date", "allstate_open", "allstate_high","allstate_low","allstate_close")

head(allstate_data)
```

```
## # A tibble: 6 x 5
##   Date       allstate_open allstate_high allstate_low allstate_close
##   <date>             <dbl>         <dbl>        <dbl>          <dbl>
## 1 1999-12-01          26.3          26.7         26.1           26.1
## 2 1999-12-02          26.1          26.9         26.1           26.6
## 3 1999-12-03          26.8          28           26.8           27.2
## 4 1999-12-06          27            28           26.8           27
## 5 1999-12-07          26.9          27.9         26.9           27
```

```
## 6 1999-12-08          27            27.7         27            27.2
```

Selected the columns that are needed for the visualizations and formed a subset using those columns. Renamed the columns according to our convenience

Loading Progressive Stock dataset

```
progressive_data <- read_csv('https://query1.finance.yahoo.com/v7/finance/download/PGR?period1=94400640
```

```
## Rows: 5482 Columns: 7
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (6): Open, High, Low, Close, Adj Close, Volume
## date (1): Date
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
progressive_data <- subset(progressive_data, select = c("Date","Open","High","Low","Close"))

colnames(progressive_data) <- c("Date", "progressive_open", "progressive_high","progressive_low","progr

head(progressive_data)
```

```
## # A tibble: 6 x 5
##   Date       progressive_open progressive_high progressive_low progressive_close
##   <date>                <dbl>            <dbl>           <dbl>              <dbl>
## 1 1999-12-01             6.72             6.77            6.64               6.64
## 2 1999-12-02             6.67             6.71            6.55               6.61
## 3 1999-12-03             6.62             6.71            6.55               6.56
## 4 1999-12-06             6.49             6.52            6.38               6.41
## 5 1999-12-07             6.37             6.38            6.26               6.27
## 6 1999-12-08             6.23             6.29            5.99               6
```

Selected the columns that are needed for the visualizations and formed a subset using those columns. Renamed the columns according to our convenience

Loading Travellers Stock dataset

```
travellers_data <- read_csv('https://query1.finance.yahoo.com/v7/finance/download/TRV?period1=9440064008
```

```
## Rows: 5482 Columns: 7
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (6): Open, High, Low, Close, Adj Close, Volume
## date (1): Date
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
travellers_data <- subset(travellers_data, select = c("Date","Open","High","Low","Close"))
```

```
colnames(travellers_data) <- c("Date", "travellers_open", "travellers_high","travellers_low","travellers

head(travellers_data)
```

```
## # A tibble: 6 x 5
##   Date       travellers_open travellers_high travellers_low travellers_close
##   <date>               <dbl>           <dbl>          <dbl>            <dbl>
## 1 1999-12-01            30.4            31.2           30.2             30.9
## 2 1999-12-02            31              31.5           31               31.5
## 3 1999-12-03            32              32.9           31.6             31.8
## 4 1999-12-06            31.9            31.9           30.8             31
## 5 1999-12-07            30.4            31.2           30.4             30.5
## 6 1999-12-08            30.5            31.2           30.4             30.8
```

Selected the columns that are needed for the visualizations and formed a subset using those columns. Renamed the columns according to our convenience

Loading Disaster dataset

```
disaster_data <- subset(disaster_data, select = c("incidentType", "state", "declarationType","fyDeclare

disaster_data <- disaster_data %>%  distinct()

disaster_data$incidentBeginDate = as.Date(disaster_data$incidentBeginDate)

disaster_data <- disaster_data %>% filter(incidentBeginDate>as.Date("2000-01-01"))
```

Selected the columns that are needed for the visualizations and formed a subset using those columns. To remove duplicate values distinct() function is used and filtered the values from the year 2000

Combinning Stock and Disaster datasets

```
disaster_stock_data <- disaster_data %>%
  left_join(allstate_data, by=c("incidentBeginDate" = "Date")) %>%
  left_join(progressive_data, by=c("incidentBeginDate" = "Date"))%>%
  left_join(travellers_data, by=c("incidentBeginDate" = "Date"))%>%
  na.omit((disaster_data_all))

disaster_stock_data
```

```
## # A tibble: 1,189 x 20
##    incidentType    state declarationType fyDeclared declarationDate
##    <chr>           <chr> <chr>                <dbl> <dttm>
##  1 Tornado         KY    DR                    2000 2000-01-10 13:00:00
##  2 Severe Storm(s) NC    DR                    2000 2000-01-31 18:30:00
##  3 Snow            LA    DR                    2000 2000-02-15 17:30:00
##  4 Severe Storm(s) GA    DR                    2000 2000-02-15 15:00:00
##  5 Severe Storm(s) VA    DR                    2000 2000-02-28 14:20:00
##  6 Flood           WV    DR                    2000 2000-02-28 15:00:00
##  7 Severe Storm(s) KY    DR                    2000 2000-02-28 17:30:00
##  8 Severe Storm(s) OH    DR                    2000 2000-03-07 14:45:00
##  9 Severe Storm(s) AL    DR                    2000 2000-03-17 16:00:00
## 10 Tornado         TX    DR                    2000 2000-04-07 13:45:00
```

```
## # ... with 1,179 more rows, and 15 more variables: incidentBeginDate <date>,
## #   incidentEndDate <dttm>, disasterCloseoutDate <dttm>, allstate_open <dbl>,
## #   allstate_high <dbl>, allstate_low <dbl>, allstate_close <dbl>,
## #   progressive_open <dbl>, progressive_high <dbl>, progressive_low <dbl>,
## #   progressive_close <dbl>, travellers_open <dbl>, travellers_high <dbl>,
## #   travellers_low <dbl>, travellers_close <dbl>
```

Joining all the four datasets together on date into a single dataset.

```
summ <- tabyl(disaster_stock_data,incidentType)
summ <- summ %>% filter(n >50)
```

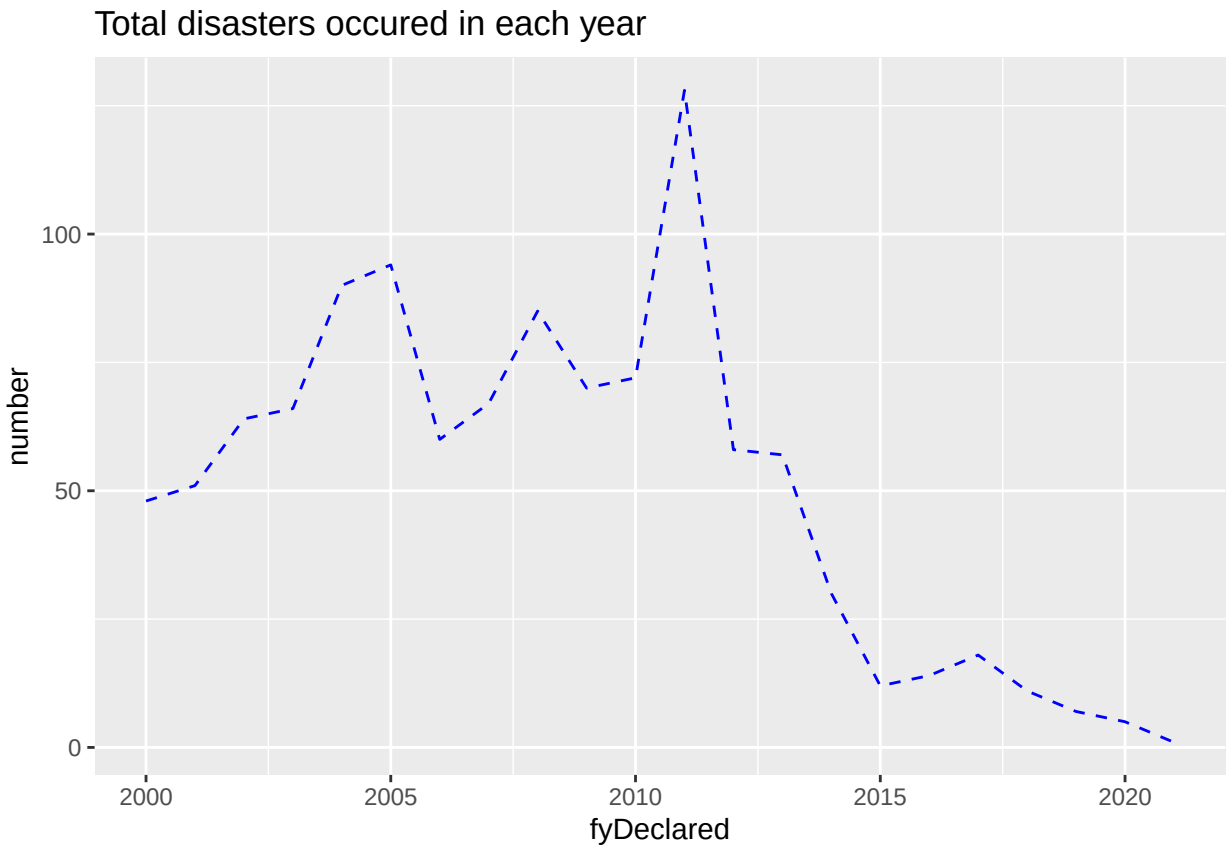
Till now we have loaded the datasets and filtered the data according to our suitability, The Number of Disasters occurred on each year.

```
count_incidents <-tabyl(disaster_stock_data, incidentType, fyDeclared)
count_incidents <- disaster_stock_data %>% filter(incidentType== 'Fire' | incidentType== 'Flood' | inci

count_incidents
```

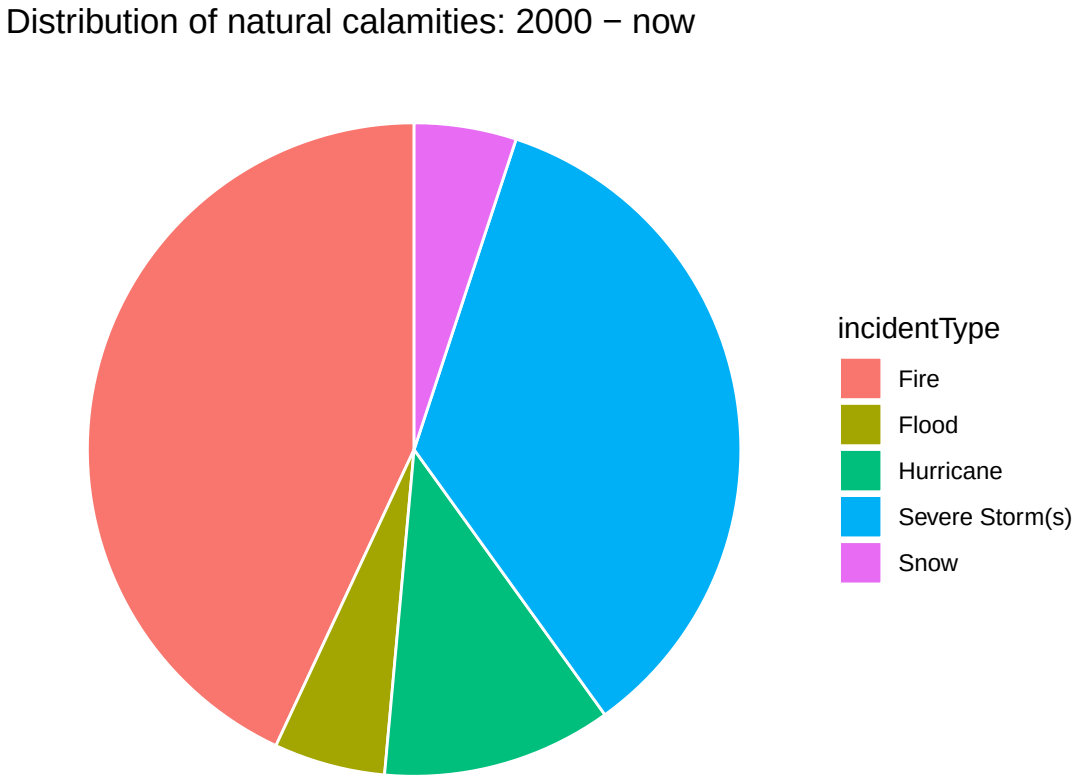
```
## # A tibble: 22 x 2
##    fyDeclared number
##         <dbl>  <int>
##  1       2000     48
##  2       2001     51
##  3       2002     64
##  4       2003     66
##  5       2004     90
##  6       2005     94
##  7       2006     60
##  8       2007     67
##  9       2008     85
## 10       2009     70
## # ... with 12 more rows
```

```
ggplot(data=count_incidents, aes(x=fyDeclared, y = number))+ ggtitle("Total disasters occured in each y
    geom_line(color = "blue", linetype = 2)
```



What is the distribution of the natural calamities in the US from 2000 - Present

```
ggplot(summ, aes(x="", y=n, fill=incidentType ,inherit.aes = FALSE)) +
 geom_bar(stat="identity", width=1, color="white") +
 coord_polar("y", start=0) +
 theme_void() + ggtitle("Distribution of natural calamities: 2000 - now")
```



Impact of natural calamities on the three Insurance Company Stocks

```
disaster_stock_data<- disaster_stock_data %>% mutate(dip_allstate = allstate_close - allstate_open)

disaster_stock_data<- disaster_stock_data %>% mutate(dip_progressive = progressive_close- progressive_o

disaster_stock_data<-disaster_stock_data %>% mutate(dip_travellers = travellers_close - travellers_open

disaster_stock_data <- disaster_stock_data %>% mutate(dip = apply(disaster_stock_data[,c('dip_allstate'

glimpse(disaster_stock_data)
```

```
## Rows: 1,189
## Columns: 24
## $ incidentType         <chr> "Tornado", "Severe Storm(s)", "Snow", "Severe Sto~
## $ state                <chr> "KY", "NC", "LA", "GA", "VA", "WV", "KY", "OH", "~
## $ declarationType      <chr> "DR", "DR", "DR", "DR", "DR", "DR", "DR", "DR", "~
## $ fyDeclared           <dbl> 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2~
## $ declarationDate      <dttm> 2000-01-10 13:00:00, 2000-01-31 18:30:00, 2000-0~
## $ incidentBeginDate    <date> 2000-01-03, 2000-01-24, 2000-01-27, 2000-02-14, ~
## $ incidentEndDate      <dttm> 2000-01-04 00:00:00, 2000-02-01 00:00:00, 2000-0~
## $ disasterCloseoutDate <dttm> 2012-05-10 00:00:00, 2012-06-28 00:00:00, 2016-0~
## $ allstate_open        <dbl> 24.0000, 23.6250, 23.4375, 21.6875, 23.1250, 21.2~
## $ allstate_high        <dbl> 24.1875, 23.8750, 23.9375, 22.0625, 23.6250, 21.5~
## $ allstate_low         <dbl> 23.2500, 22.8125, 23.1875, 21.0625, 22.8125, 20.7~
```

```
## $ allstate_close       <dbl> 23.5000, 22.8125, 23.1875, 21.2500, 22.8750, 21.0~
## $ progressive_open     <dbl> 6.052083, 5.453125, 5.343750, 4.500000, 5.515625,~
## $ progressive_high     <dbl> 6.083333, 5.557292, 5.473958, 4.500000, 5.515625,~
## $ progressive_low      <dbl> 5.890625, 5.416667, 5.322917, 4.312500, 5.328125,~
## $ progressive_close    <dbl> 5.890625, 5.494792, 5.390625, 4.343750, 5.333333,~
## $ travellers_open      <dbl> 33.2500, 31.6250, 32.0000, 24.2500, 32.1250, 24.5~
## $ travellers_high      <dbl> 33.2500, 32.4375, 32.3750, 24.5625, 32.8125, 24.6~
## $ travellers_low       <dbl> 32.6250, 31.3125, 29.8125, 24.0625, 31.2500, 23.1~
## $ travellers_close     <dbl> 33.0000, 32.1250, 30.5625, 24.2500, 31.3750, 23.6~
## $ dip_allstate         <dbl> -0.5000, -0.8125, -0.2500, -0.4375, -0.2500, -0.2~
## $ dip_progressive      <dbl> -0.161458, 0.041667, 0.046875, -0.156250, -0.1822~
## $ dip_travellers       <dbl> -0.2500, 0.5000, -1.4375, 0.0000, -0.7500, -0.937~
## $ dip                  <dbl> -0.500000, -0.812500, -1.437500, -0.437500, -0.75~
```

```
list <- disaster_stock_data %>% filter(incidentType == c('Fire','Flood','Hurricane','Severe Storm(s)','S
```

```
## Warning in incidentType == c("Fire", "Flood", "Hurricane", "Severe Storm(s)", :
## longer object length is not a multiple of shorter object length
```

```
summ <- full_join(summ,list)
```

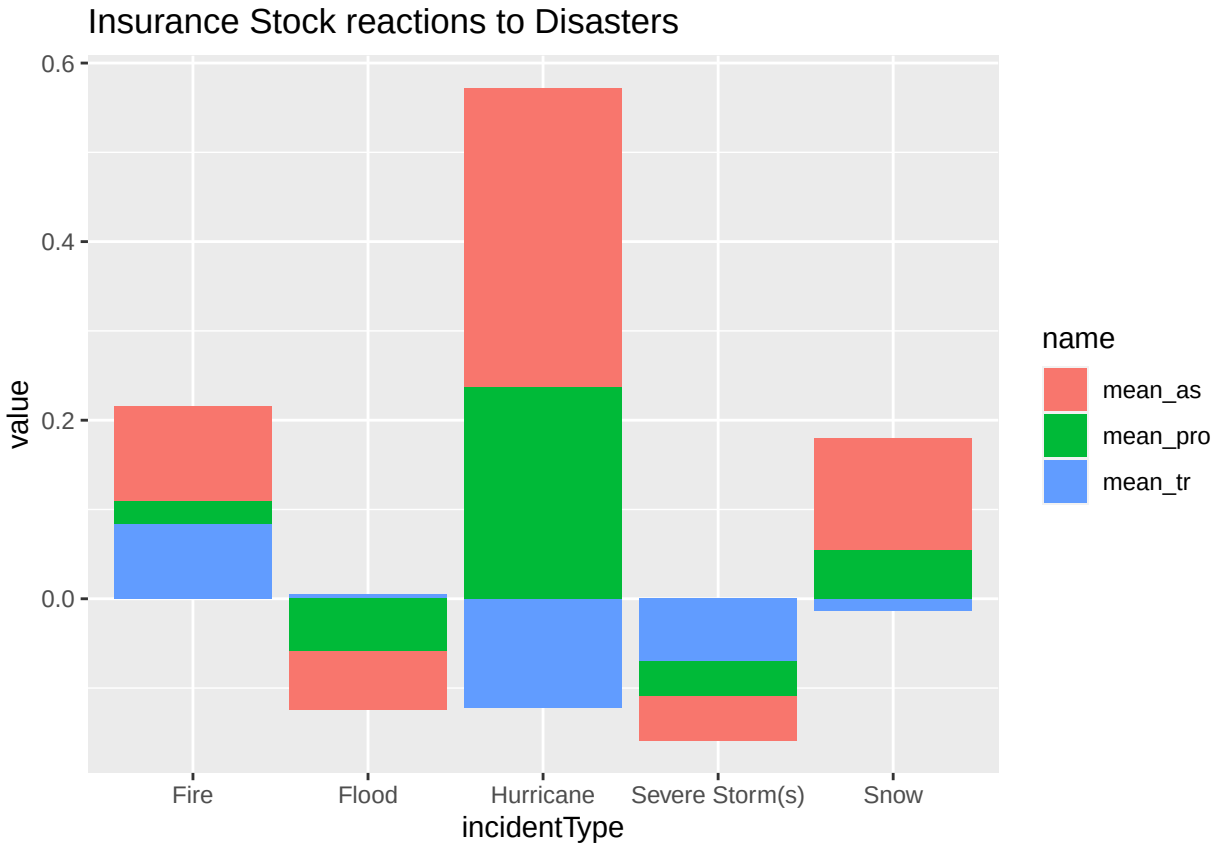
```
## Joining, by = "incidentType"
```

```
summ
```

```
##     incidentType   n    percent     mean_as      mean_tr    mean_pro
##             Fire 477 0.40117746  0.10705260  0.084039295  0.02465141
##            Flood  61 0.05130362 -0.06533247  0.004667067 -0.05855573
##        Hurricane 126 0.10597140  0.33500014 -0.122142393  0.23669582
##  Severe Storm(s) 388 0.32632464 -0.04961045 -0.070779052 -0.03813579
##             Snow  56 0.04709840  0.12596077 -0.013269846  0.05413485
```

```
summ_p <- pivot_longer(summ,cols = c('mean_as','mean_tr','mean_pro'))

ggplot(summ_p, aes(fill=name, y=value, x=incidentType)) +
    geom_bar(position="stack", stat="identity") + ggtitle("Insurance Stock reactions to Disasters")
```



Basic Linear Model of the data

```
model <- lm(formula = n ~ mean_as+mean_tr+ mean_pro, data = summ)

summary(model)
```

```
##
## Call:
## lm(formula = n ~ mean_as + mean_tr + mean_pro, data = summ)
##
## Residuals:
##       1       2       3       4       5
##  162.95 -178.57   13.95  193.62 -191.95
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    203.7      617.1   0.330    0.797
## mean_as       1224.3    17188.6   0.071    0.955
## mean_tr        325.7     7973.4   0.041    0.974
## mean_pro     -1952.1    26169.6  -0.075    0.953
##
## Residual standard error: 364.6 on 1 degrees of freedom
## Multiple R-squared:  0.1437, Adjusted R-squared:  -2.425
## F-statistic: 0.05592 on 3 and 1 DF,  p-value: 0.9758
```

Conclusions:-

The most occurred disaster is Fire. Highest number of disasters occurred in 2011 whereas least number of disasters occurred in 2021 During Fire, Hurricane and Snow we have seen a raise in the Stocks for AllState and Progressive Insurance companies. In the time of flood and severe storms all the stocks of the three Insurance Companies are dropped.

Bias:-

We considered the disasters that occurred more than 50 times. Weekends are not considered for our convenience as the markets are closed on weekends.

Source Code link

https://github.com/ShashiKiran07/daf_proj/blob/main/daf_proj.Rmd