Q6 5.8

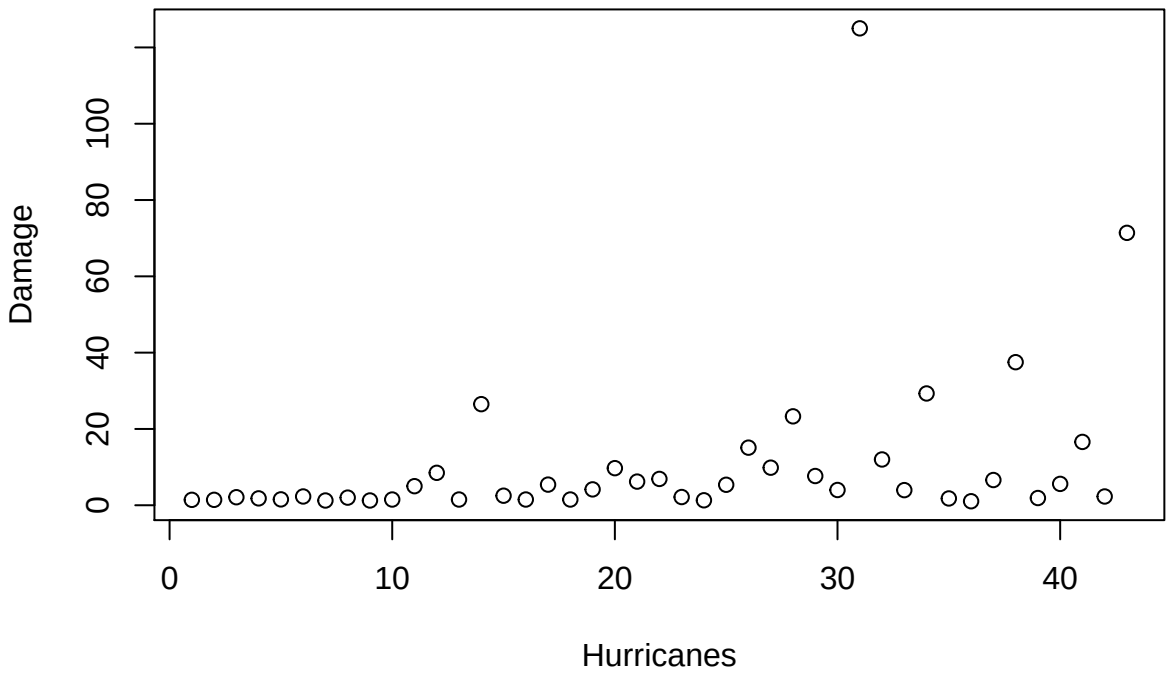
Wentong

5/8/2022

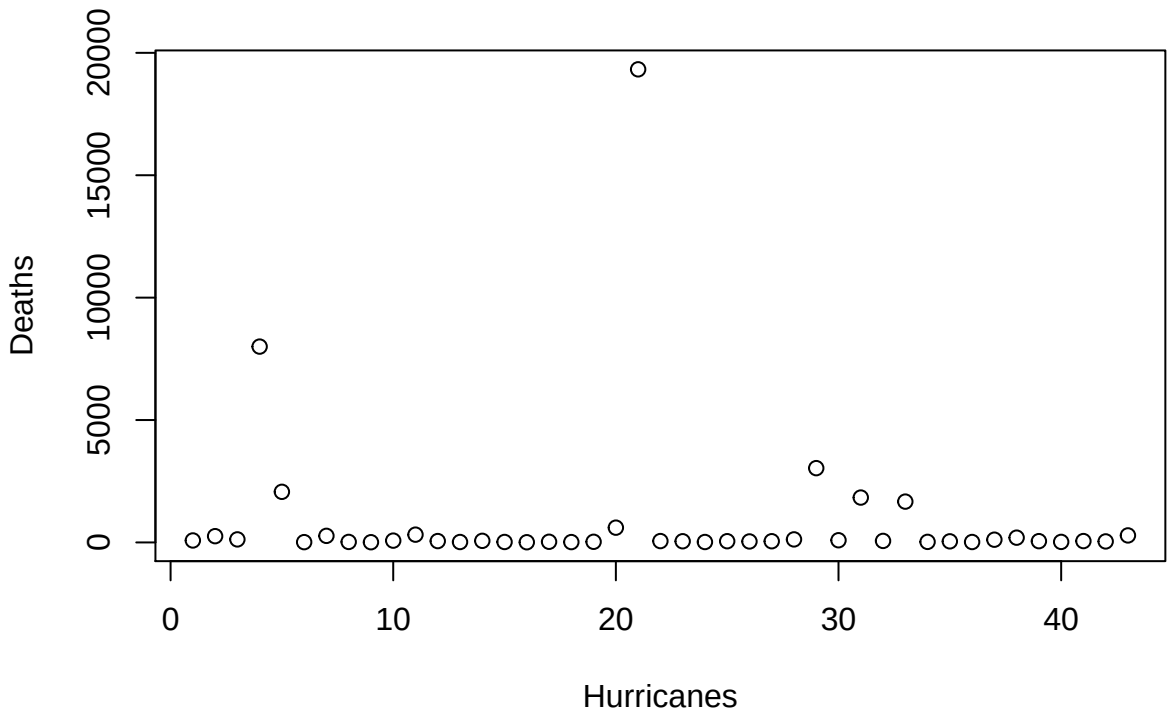
```
load("beta.res.postmean.Rdata")
dat = read.csv("hurricanoutcome2.csv") %>%
  rename(id = HurricanID)

dat$Deaths = gsub(",","",dat$Deaths)
dat$Damage = gsub("\\$", "", dat$Damage)
dat_q6 = dat %>%
  mutate(Damage = as.numeric(Damage),
         Deaths = as.numeric(Deaths),
         Season = as.numeric(Season),
         Maxspeed = as.numeric(Maxspeed),
         Month = as.factor(Month),
         Nature = as.factor(Nature))

plot(dat_q6$Damage, xlab = "Hurricanes", ylab = "Damage")
```




```
plot(dat_q6$Deaths, xlab = "Hurricanes", ylab = "Deaths")
```




```
## Import Data
dat_hur = read.csv("hurrican703.csv") %>%
  janitor::clean_names() %>%
  group_by(id) %>%
  mutate(id = tolower(id),
         wind_early = lag(wind_kt, 1),
         lat_change = lag(latitude, 0) - lag(latitude, 1),
         long_change = lag(longitude, 0) - lag(longitude, 1),
         wind_change = lag(wind_kt, 0) - lag(wind_kt, 1)) %>%
  na.omit() %>%
  as.data.frame()

dat_before = beta.res.postmean %>%
  rename(id = ID) %>%
  mutate(id = tolower(id))

combine_data = merge(dat_before, dat_hur, by = "id") %>%
  rename(beta1 = Wind_prev, beta2 = Lat_change, beta3 = Long_change, beta4 = Wind_change) %>%
  select(id, season, month, nature, intercept, beta1, beta2, beta3, beta4, wind_early, lat_change, long

combine.data2 = combine_data %>%
  select(id, intercept, beta1, beta2, beta3, beta4) %>%
  group_by(id, intercept, beta1, beta2, beta3, beta4) %>%
  summarize(nobs = n())
```

```
## `summarise()` has grouped output by 'id', 'intercept', 'beta1', 'beta2',
## 'beta3'. You can override using the `.groups` argument.
```

```
data_res = merge(combine.data2, dat_q6, by = "id")
data_res = data_res %>%
  mutate(nobs = as.numeric(nobs),
         Maxpressure = as.numeric(Maxpressure),
         Hours = as.numeric(Hours),
         Total.Pop = as.numeric(Total.Pop))

Coef_table = data_res %>%
  select(id, intercept, beta1, beta2, beta3, beta4) %>%
  knitr::kable(digits = 3, col.names = c("ID", "$\\beta_0$", "$\\beta_1$", "$\\beta_2$", "$\\beta_3$",

Coef_table %>% head(10)
```

```
##  [1] "|ID            | $\\beta_0$| $\\beta_1$| $\\beta_2$| $\\beta_3$| $\\beta_4$|"
##  [2] "|:-------------|---------:|---------:|---------:|---------:|---------:|"
##  [3] "|agnes.1972    |     3.951|     0.922|     0.006|    -0.310|     0.545|"
##  [4] "|alex.2010     |     3.799|     0.937|     0.070|    -0.394|     0.540|"
##  [5] "|alicia.1983   |     3.897|     0.904|    -0.075|    -0.399|     0.548|"
##  [6] "|allen.1980    |     3.687|     0.966|     0.131|    -0.546|     0.547|"
##  [7] "|andrew.1992   |     3.676|     0.938|    -0.284|    -0.578|     0.537|"
##  [8] "|betsy.1965    |     3.808|     0.951|    -0.450|    -0.389|     0.424|"
##  [9] "|bob.1991      |     3.629|     0.923|     0.028|    -0.575|     0.438|"
## [10] "|camille.1969  |     3.994|     0.936|     0.073|    -0.573|     0.670|"
```

```
data_damage = data_res %>%
  select(-c(id, Deaths))
str(data_damage)
```

```
## 'data.frame':    43 obs. of  18 variables:
##  $ intercept   : num  3.95 3.8 3.9 3.69 3.68 ...
##  $ beta1       : num  0.922 0.937 0.904 0.966 0.938 ...
##  $ beta2       : num  0.00595 0.06988 -0.07483 0.13064 -0.28433 ...
##  $ beta3       : num  -0.31 -0.394 -0.399 -0.546 -0.578 ...
##  $ beta4       : num  0.545 0.54 0.548 0.547 0.537 ...
##  $ nobs        : num  34 31 24 45 51 68 54 36 29 54 ...
##  $ Season      : num  1972 2010 1983 1980 1992 ...
##  $ Damage      : num  2.1 1.89 2 1.24 26.5 1.42 1.5 1.42 15.1 1.54 ...
##  $ Month       : Factor w/ 6 levels "August","July",..: 3 3 1 1 1 6 1 1 1 6 ...
##  $ Nature      : Factor w/ 3 levels "DS","NR","TS": 3 1 3 2 3 3 3 3 3 3 ...
##  $ Maxspeed    : num  75 95 100 165 150 135 100 150 130 150 ...
##  $ Meanspeed   : num  44.7 51.7 49.2 97.1 73.9 ...
##  $ Maxpressure : num  1001 1007 1011 1010 1015 ...
##  $ Meanpressure: num  423 986 995 916 979 ...
##  $ Hours       : num  210 192 150 276 312 414 330 222 180 330 ...
##  $ Total.Pop   : num  154114 49323 2267825 1387390 819815 ...
##  $ Percent.Poor: num  0 0 0 0.82 0 0 0 0 0.45 0.45 ...
##  $ Percent.USA : num  100 0 100 29.7 71.1 ...
```

```
require(broom)
```

```
## Loading required package: broom
```

```
## Warning: package 'broom' was built under R version 4.0.5
```

```
damage.fit = glm(Damage ~., data = data_damage, family = "gaussian")
damage.tidy = tidy(damage.fit) %>%
```

```
    mutate(
        term = case_when(term == "intercept"~"$\\beta_0$",
                         term == "beta1"~"$\\beta_1$",
                         term == "beta2"~"$\\beta_2$",
                         term == "beta3"~"$\\beta_3$",
                         term == "beta4"~"$\\beta_4$",
                         TRUE ~ term)
    ) %>% na.omit()

df_total = data.frame()
coef_res = function(table){
  for (i in 1:22) {
    df = table[i,]
    if(table[i,5]<0.05){df_total = rbind(df_total, df)
    i = i + 1}
    else{i = i +1}
  }
  return(df_total)
}

coef_res(damage.tidy) %>% knitr::kable(digits = 3)
```

|| || || ||

```
data_deaths = data_res %>%
  select(-c(id, Damage))

deaths.fit = glm(Deaths ~ ., data = data_deaths, family = "poisson", offset = log(Total.Pop))

deaths.tidy = tidy(deaths.fit) %>%
    mutate(
        term = case_when(term == "intercept"~"$\\beta_0$",
                         term == "beta1"~"$\\beta_1$",
                         term == "beta2"~"$\\beta_2$",
                         term == "beta3"~"$\\beta_3$",
                         term == "beta4"~"$\\beta_4$",
                         TRUE ~ term)
    ) %>% na.omit()
coef_res(deaths.tidy)%>% knitr::kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	214.595	12.842	16.710	0.000
β_0	12.755	0.263	48.493	0.000
β_1	139.069	2.314	60.088	0.000
β_2	5.898	0.122	48.512	0.000
β_3	11.301	0.301	37.556	0.000
β_4	-14.510	0.306	-47.425	0.000
nobs	-0.006	0.001	-5.093	0.000
Season	-0.008	0.002	-3.585	0.000
MonthJuly	-2.213	0.157	-14.061	0.000
MonthJune	-0.246	0.091	-2.710	0.007
MonthNovember	-3.042	0.156	-19.543	0.000
MonthOctober	-2.055	0.062	-32.977	0.000

term	estimate	std.error	statistic	p.value
MonthSeptember	-0.691	0.047	-14.612	0.000
NatureNR	1.721	0.131	13.178	0.000
NatureTS	4.010	0.123	32.580	0.000
Maxspeed	0.014	0.001	11.913	0.000
Meanspeed	-0.061	0.003	-19.156	0.000
Maxpressure	-0.376	0.010	-38.641	0.000
Meanpressure	0.008	0.000	41.296	0.000
Total.Pop	0.000	0.000	-8.444	0.000
Percent.Poor	0.050	0.001	55.902	0.000
Percent.USA	-0.019	0.001	-33.874	0.000