Q6 5.8

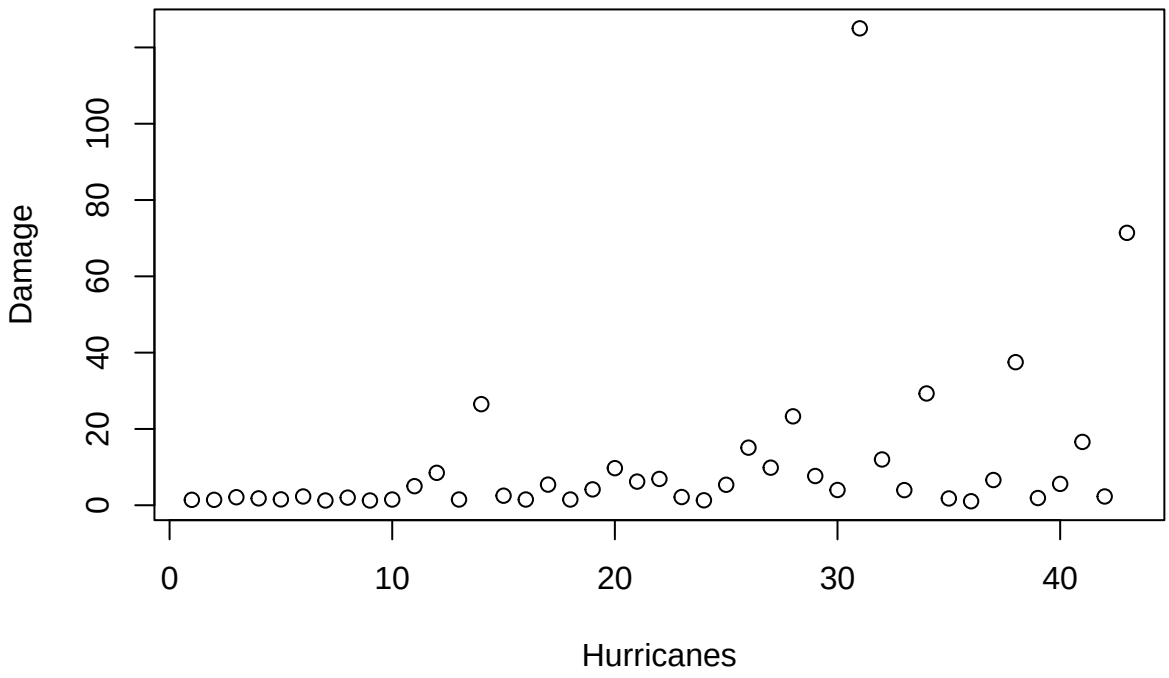
Wentong

5/8/2022

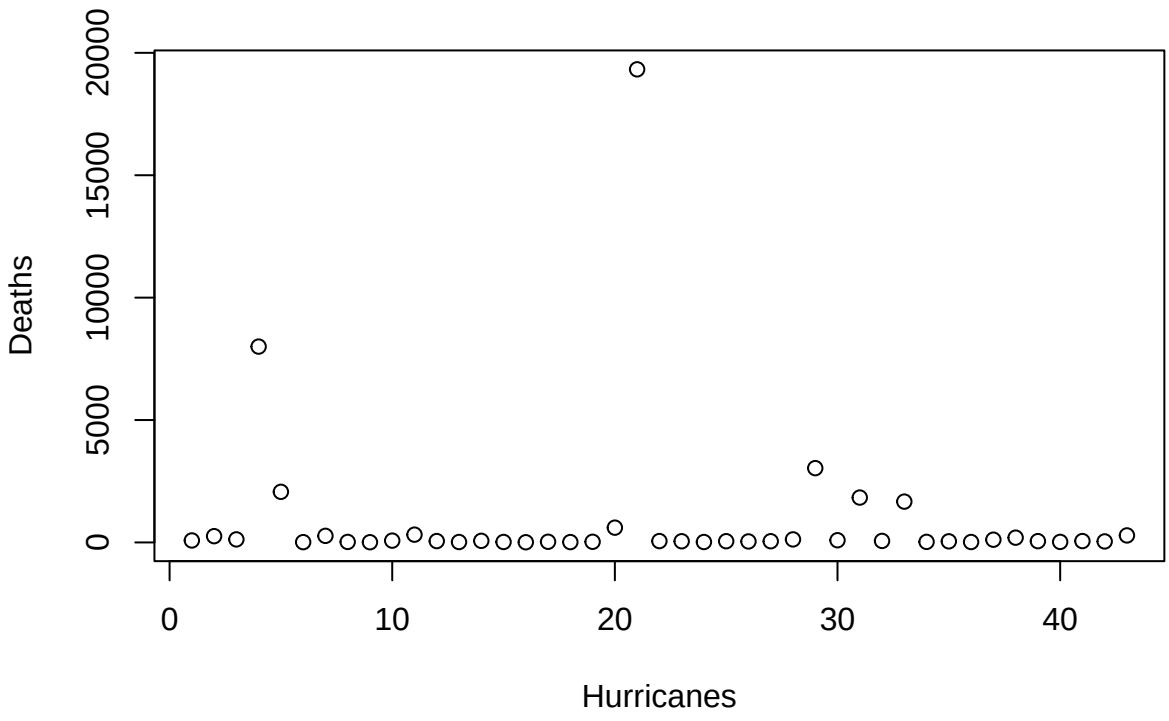
```
load("beta.res.postmean.Rdata")
dat = read.csv("hurricanoutcome2.csv") %>%
  rename(id = HurricanID)

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 us
```

```
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()

Coef_table %>% head(10)
```

```
##  [1] "|id            | intercept|     beta1|      beta2|      beta3|     beta4|"
##  [2] "|:-------------|---------:|---------:|----------:|----------:|---------:|"
##  [3] "|agnes.1972    |  3.950974| 0.9224097|  0.0059532| -0.3103372| 0.5453543|"
##  [4] "|alex.2010     |  3.798737| 0.9370333|  0.0698849| -0.3937358| 0.5400187|"
##  [5] "|alicia.1983   |  3.897408| 0.9036878| -0.0748341| -0.3994486| 0.5477718|"
##  [6] "|allen.1980    |  3.687070| 0.9655304|  0.1306393| -0.5460144| 0.5466129|"
##  [7] "|andrew.1992   |  3.676279| 0.9375384| -0.2843257| -0.5782973| 0.5370158|"
##  [8] "|betsy.1965    |  3.808396| 0.9513766| -0.4500720| -0.3890718| 0.4244575|"
##  [9] "|bob.1991      |  3.629466| 0.9232143|  0.0279527| -0.5751636| 0.4382048|"
## [10] "|camille.1969  |  3.994355| 0.9355674|  0.0729188| -0.5734830| 0.6703910|"
```

```
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
```

```r
damage.fit = glm(Damage ~., data = data_damage, family = "poisson")
```

```
## Warning in dpois(y, mu, log = TRUE): non-integer x = 2.100000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.890000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.240000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 26.500000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.420000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.500000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.420000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 15.100000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.540000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.780000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.050000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.250000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.800000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 6.900000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 4.160000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 9.850000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 2.300000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 9.720000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 6.610000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 8.500000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 37.500000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 16.600000
## Warning in dpois(y, mu, log = TRUE): non-integer x = 2.300000
```

```
## Warning in dpois(y, mu, log = TRUE): non-integer x = 5.370000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.300000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 23.300000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 7.660000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.500000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 5.600000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 2.500000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.500000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 2.150000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 6.200000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 5.410000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 1.500000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 71.400000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 3.960000

## Warning in dpois(y, mu, log = TRUE): non-integer x = 29.300000
```

```
damage.tidy = tidy(damage.fit) %>% 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}
  }
  print(df_total)
}

coef_res(damage.tidy)
```

```
## # A tibble: 17 x 5
##    term             estimate std.error statistic  p.value
##    <chr>               <dbl>     <dbl>     <dbl>    <dbl>
##  1 (Intercept)    -218.       63.8         -3.42 6.34e- 4
##  2 intercept         5.04      0.873        5.78 7.41e- 9
```

```
##  3 beta1            62.8      14.0          4.48 7.48e- 6
##  4 beta2            -1.10      0.424       -2.58 9.81e- 3
##  5 beta3             3.38      0.816        4.14 3.48e- 5
##  6 nobs              0.0492    0.00804      6.12 9.15e-10
##  7 Season            0.0750    0.0126       5.94 2.89e- 9
##  8 MonthJune        -3.42      0.762       -4.48 7.38e- 6
##  9 MonthNovember    -1.90      0.789       -2.41 1.59e- 2
## 10 MonthOctober     -1.29      0.298       -4.33 1.50e- 5
## 11 MonthSeptember   -1.76      0.243       -7.25 4.03e-13
## 12 NatureNR         -4.32      1.13        -3.83 1.27e- 4
## 13 NatureTS         -2.04      0.453       -4.50 6.77e- 6
## 14 Maxspeed          0.0504    0.00676      7.46 8.81e-14
## 15 Meanspeed        -0.0657    0.0154      -4.26 2.02e- 5
## 16 Percent.Poor     -0.0382    0.00586     -6.52 7.05e-11
## 17 Percent.USA      -0.00463   0.00232     -2.00 4.55e- 2
```

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

deaths.fit = glm(Deaths ~ ., data = data_deaths, family = "poisson")

deaths.tidy = tidy(deaths.fit) %>% na.omit()
coef_res(deaths.tidy)
```

```
## # A tibble: 21 x 5
##    term         estimate std.error statistic   p.value
##    <chr>           <dbl>     <dbl>     <dbl>     <dbl>
##  1 (Intercept) 116.       12.6          9.26 2.03e- 20
##  2 intercept    11.7       0.256       45.5  0
##  3 beta1       114.        2.20        51.9  0
##  4 beta2         5.53      0.123       45.1  0
##  5 beta3         8.56      0.285       30.0  7.91e-198
##  6 beta4       -10.5       0.306      -34.3  6.12e-258
##  7 nobs          0.00343   0.00112      3.07 2.12e-  3
##  8 Season        0.00610   0.00209      2.91 3.56e-  3
##  9 MonthJuly    -1.18      0.145       -8.17 3.07e- 16
## 10 MonthJune    -1.29      0.0897     -14.4  5.03e- 47
## # ... with 11 more rows
```