STAT431 - Project

2023-04-14

```
## Removed SDP

### Sex
## 1 - Male
## 2 - Female

### Races
## 1 - Black
## 2 - White
## 3 - Hispanic
## 4 - Asian
## 5 - American Indian
## 6 - Biracial
## 7 - Unknown

### Veteran Status
## 1 - Veteran
## 2 - Not veteran
## 3 - Unknown

### Offense Type - Crime Type
## 1 - Person Crimes
## 2 - Sex Crimes
## 3 - Drug Crimes
## 4 - Property Crimes
## 5 - Other Crimes

### Crime Class - Crime Severity
## 5 - Class X and Murder
## 4 - Class I
## 3 - Class II
## 2 - Class III
## 1 - Class IV

### Regions
## 1 - Region 1
## 2 - Region 2
## 3 - Region 3
## 4 - Region 4
## 5 - Region 5
```

```
library(tidyverse)
library(rjags)
library(readr)
```

```r
incar_data <- read_csv("C:/Users/chand/Desktop/STAT431-Group5/incar_data.csv")

incar_data_two <- incar_data %>%
  select('Sex', 'Race', 'Veteran Status',
         'Offense Type', 'IDHS Region', 'Sentence Time (num)')

incar_data_two <- na.omit(incar_data_two)

colnames(incar_data_two) <- c("Sex", "Race", "VeteranStatus",
                              "OffenseType", "Region", "SentenceTime")

incar_data_two$Sex <- factor(incar_data_two$Sex)
incar_data_two$Race <- factor(incar_data_two$Race)
incar_data_two$VeteranStatus <- factor(incar_data_two$VeteranStatus)
incar_data_two$OffenseType <- factor(incar_data_two$OffenseType)
incar_data_two$Region <- factor(incar_data_two$Region)


myModel <- lm(SentenceTime ~ Sex + Race + VeteranStatus +
                OffenseType + Region, data = incar_data_two)

summary(myModel)
```

```
##
## Call:
## lm(formula = SentenceTime ~ Sex + Race + VeteranStatus + OffenseType +
##     Region, data = incar_data_two)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -26.57 -11.30  -2.98   4.43 578.66
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     27.5702     0.7402  37.246  < 2e-16 ***
## Sex2            -3.9355     0.5337  -7.373 1.71e-13 ***
## Race2           -0.8118     0.2845  -2.854  0.00433 **
## Race3           -0.2324     0.3494  -0.665  0.50603
## Race4            0.2736     1.8663   0.147  0.88343
## Race5            2.4230     3.0830   0.786  0.43192
## Race6           -3.9411     2.4224  -1.627  0.10376
## Race7           -7.0494     2.2648  -3.113  0.00186 **
## VeteranStatus2  -5.9960     0.7249  -8.271  < 2e-16 ***
## VeteranStatus3  -1.2709     0.7415  -1.714  0.08656 .
## OffenseType2    -6.6255     0.3143 -21.078  < 2e-16 ***
## OffenseType3   -12.4381     0.3543 -35.103  < 2e-16 ***
## OffenseType4   -13.3627     0.3834 -34.855  < 2e-16 ***
## OffenseType5   -15.7780     1.3496 -11.691  < 2e-16 ***
## Region2         -3.0665     0.3153  -9.726  < 2e-16 ***
## Region3         -2.3438     0.3670  -6.386 1.73e-10 ***
## Region4         -3.7528     0.4128  -9.091  < 2e-16 ***
## Region5         -2.8357     0.3625  -7.822 5.38e-15 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 18.21 on 27856 degrees of freedom
## Multiple R-squared:  0.1253, Adjusted R-squared:  0.1247
## F-statistic: 234.7 on 17 and 27856 DF,  p-value: < 2.2e-16
```

```
incar_data_three <- incar_data_two %>%
  select("OffenseType", "SentenceTime")

inits <- list(list(a = 0, b = c(0, 0, 0, 0, 0), tausq = 1),
              list(a = 5, b = c(10, 10, 10, 10, 10), tausq = 1),
              list(a = 25, b = c(-10, 10, -10, 10, -10), tausq = 1))

cat('
model {
  for(i in 1:length(SentenceTime)) {
    SentenceTime[i] ~ dnorm(mu[i], tausq)
    mu[i] <- a + b[OffenseType[i]]
  }

  a ~ dnorm(25, 0.001)
  tausq ~ dgamma(0.0001,0.0001)

  for (j in 1:5) {
    b[j] ~ dnorm(0, 0.01)

  }
}
', file = {projBayes1 = tempfile()})

projectBayesModel1 <- jags.model(projBayes1, incar_data_three, inits, n.chains=3)
model1sample1 <- coda.samples(projectBayesModel1, c("b[5]"), n.iter=30000)

summary(window(model1sample1, 28000, 30000))
gelman.plot(model1sample1, autoburnin=FALSE)
gelman.diag(model1sample1, autoburnin=FALSE)
```

```
incar_data_four <- incar_data_two %>%
  select("Race", "Sex", "OffenseType", "Region", "SentenceTime")

inits <- list(list(a = 0,
                   b = c(-10, -10, -10, -10, -10),
                   d = c(0, 1),
                   e = c(10, -5, 4, 2, -8, 9, -10),
                   f = c(-10, -5, 0, 5, 10),
                   tausq = 1),

              list(a = 25,
                   b = c(10, 10, 10, 10, 10),
                   d = c(1, 0),
                   e = c(-4, -8, -10, -6, 9, -6, 9),
                   f = c(10, 5, 0, -5, 10),
                   tausq = 1),

              list(a = 50,
                   b = c(-10, 10, -10, 10, -10),
                   d = c(0, 1),
                   e = c(0, 0, 0, 0, 0, 0, 0),
                   f = c(2, 2, 2, 2, 2),
                   tausq = 1))

cat('
model {
  for(i in 1:length(SentenceTime)) {
    SentenceTime[i] ~ dnorm(mu[i], tausq)
    mu[i] <- a + b[OffenseType[i]] + d[Sex[i]] + e[Race[i]] + f[Region[i]]
  }

  a ~ dnorm(0, 0.0001)
  tausq ~ dgamma(0.0001,0.0001)

  for (j in 1:5) {
    b[j] ~ dnorm(0, 0.01)
  }

  for (k in 1:2) {
    d[k] ~ dnorm(0, 0.01)
  }

  for (l in 1:7) {
    e[l] ~ dnorm(0, 0.01)
  }

  for (w in 1:5) {
    f[w] ~ dnorm(0, 0.01)
  }

}
', file = {projBayes2 = tempfile()})
```

```
projectBayesModel2 <- jags.model(projBayes2, incar_data_four, inits, n.chains=3)
model2sample1 <- coda.samples(projectBayesModel2, c("a", "b", "d", "e", "f"), n.iter=50000)


gelman.diag(model2sample1, autoburnin=FALSE)
```

```
incar_data_four <- incar_data_two %>%
  select("Race", "Sex", "OffenseType", "Region", "SentenceTime")

inits <- list(list(a = 0,
                   b = c(-10, -10, -10, -10, -10),
                   d = c(0, 1),
                   e = c(10, -5, 4, 2, -8, 9, -10),
                   f = c(-10, -5, 0, 5, 10),
                   tausq = 1),

              list(a = 5,
                   b = c(10, 10, 10, 10, 10),
                   d = c(1, 0),
                   e = c(-4, -8, -10, -6, 9, -6, 9),
                   f = c(10, 5, 0, -5, 10),
                   tausq = 1),

              list(a = 25,
                   b = c(-10, 10, -10, 10, -10),
                   d = c(0, 1),
                   e = c(0, 0, 0, 0, 0, 0, 0),
                   f = c(2, 2, 2, 2, 2),
                   tausq = 1))

cat('
model {
  for(i in 1:length(SentenceTime)) {
    SentenceTime[i] ~ dnorm(mu[i], tausq)
    mu[i] <- a + b[OffenseType[i]] + d[Sex[i]] + e[Race[i]] + f[Region[i]]
  }

  tausq ~ dgamma(0.0001,0.0001)

  a ~ dnorm(25, 0.0001)

  for (j in 1:5) {
    b[j] ~ dnorm(0, 0.0001)
  }

  for (k in 1:2) {
    d[k] ~ dnorm(0, 0.0001)
  }

  for (l in 1:7) {
    e[l] ~ dnorm(0, 0.0001)
  }

  for (w in 1:5) {
    f[w] ~ dnorm(0, 0.0001)
  }

}
', file = {projBayes2 = tempfile()})
```

```
projectBayesModel2 <- jags.model(projBayes2, incar_data_four, inits, n.chains=3)
model2sample1 <- coda.samples(projectBayesModel2, c("a", "b", "d","e", "f"), n.iter=25000)
summary(window(model2sample1, 8000, 10000))
```