Open Log and Read in Data

- Use the rm function to remove all active objects in the memory (global environment) and the setwd function to create a working directory and
- Use the sink function to divert the codes and results to a log file and use the read.dta function to load the external Stata data file gsscum7212Teach.dta into R.

```
#Open Log and read in data
setwd("/Users/burrisfaculty/Desktop/DSCode/SOC686")
sink("Shepherd_asgn09.log", split=T)
rm(list=ls(all=TRUE))
mygss <- read.dta("gsscum7212teach.dta", convert.factor=F)
```

Manage Data and Run Logit

- Keep six variables, include mental health (mntlhlth), age (age), sex (sex), race (race), education (educ), and income (inc1k)
- Dichotomize the mental health variable such that the new binary response variable is coded as one (1 = having poor mental health) if the number of days for poor mental health is greater than zero, otherwise it's coded as zero (0 = having excellent mental health). Label this new variable as mntlhlthc2.
- Create dummy variables for sex and race. Note that the race variable has three categories, so please create three dummy variables for race (Alternatively, you can the factor function to turn the race variable into a factor variable and use it directly in the regression). Also you need to be careful and clear about 1) how many of these three dummy variables, all measuring race, are usually used in a regression model and 2) how to interpret the results/corresponding coefficients (e.g., which group is the reference group?). Please also drop missing cases using listwise deletion (any case that has missing information for any of the six variables will be dropped from the sample data).
- Check the descriptive statistics of these variables using the table and the summary function when appropriate. Note that when there is too much output (e.g., tabulation of income), you can present representative information.
- Run an OLS regression of mental health on age, sex (male as the reference category), race (white as the reference category), educa#tion, and income.

```
> #MANAGE DATA AND RUN LOGIT
> #SELECT DATA
> useddta <- subset(mygss,
+                 select=c(mntlhlth, age, sex, race, educ, inc1k))

> #Create dummy variables female (male = 0)
> useddta$female <- as.numeric(useddta$sex==2)

> useddta$male <- as.numeric(useddta$sex == 1)

> #Create Binary Indicator Variables for Multi-Category Nomial Variables
>
> useddta$white <- ifelse(useddta$race == 1, 1, 0)

> useddta$black <- ifelse(useddta$race == 2, 1, 0)

> useddta$other <- ifelse(useddta$race == 3, 1, 0)
```

```
> nmdta <- useddta[complete.cases(useddta),] #no missing data

> #summarize data
> summary(nmdta$mntlhlth)
   Min.     1st Qu.  Median    Mean       3rd Qu.    Max.
   0.000    0.000    0.000     3.991      5.000      30.000

> summary(nmdta$inc1k)
   Min.     1st Qu.  Median    Mean       3rd Qu.    Max.
   0.245    15.921   28.157    39.404     48.475     155.140

> summary(nmdta$age)
   Min.     1st Qu.  Median    Mean       3rd Qu.    Max.
   18.00    31.00    42.00     41.81      51.00      84.00

> table(nmdta$female)

  0         1
  381       369

> table(nmdta$white)

  0         1
  160       590

> #Run OLS
> mntlhlth.model <- lm(formula = mntlhlth ~ age + female + black + other +
educ + inc1k, data = nmdta )

> (summary(mntlhlth.model))

Call:
lm(formula = mntlhlth ~ age + female + black + other + educ +
    inc1k, data = nmdta)

Residuals:
   Min      1Q          Median      3Q          Max
   -6.888   -4.069      -2.818      0.601       27.711

Coefficients:
             Estimate   Std. Error  t value     Pr(>|t|)
(Intercept)  6.911999   1.656997    4.171       3.38e-05 ***
age         -0.009268   0.020120    -0.461      0.64517
female       1.017865   0.528007    1.928       0.05427 .
black       -2.119554   0.790024    -2.683      0.00746 **
other        0.569797   0.955038    0.597       0.55094
educ        -0.188625   0.102889    -1.833      0.06716 .
inc1k       -0.004487   0.007959    -0.564      0.57309
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.105 on 743 degrees of freedom
Multiple R-squared:  0.02043, Adjusted R-squared:  0.01252
F-statistic: 2.582 on 6 and 743 DF,  p-value: 0.01753
```

Task 1 Run Bayesian Linear Regression with Non-Informative Priors

- Use the MCMCregress function from the MCMCpack package to run a Bayesian linear regression of mental health on age, sex (male as the reference category), race (white as the reference category), education, and income using noninformative prior (e.g., N (0, 1e6)). Note that the B0 argument in the MCMCregress function is for specifying priors for precision, the inverse of variance. Use the summary function to summarize the results from this Bayesian linear regression.

```
> #TASK 1 Run Bayesian Linear Reg (Non-Informative Pairs)
> library(MCMCpack)

> mcmc.model = MCMCregress(mntlhlth ~ age + female + black + other + educ + inc1k,
+    data = nmdta, burnin = 1000, mcmc = 10000, thin = 1,seed = 47304,b0 = 0,
B0 = 1,marginal.likelihood = "Chib95")

> summary(mcmc.model)

Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

            Mean        SD  Naive   SE Time-series    SE
(Intercept) 1.843978    0.853808    8.538e-03         8.498e-03
age         0.022910    0.018192    1.819e-04         1.819e-04
female      0.917328    0.462010    4.620e-03         4.620e-03
black       -1.127879   0.610042    6.100e-03         6.231e-03
other       0.586758    0.687921    6.879e-03         6.879e-03
educ        0.078028    0.072803    7.280e-04         7.280e-04
inc1k       -0.009583   0.007784    7.784e-05         7.784e-05
sigma2      51.207177   2.644329    2.644e-02         2.714e-02

2. Quantiles for each variable:

             2.5%       25%         50%         75%         97.5%
(Intercept)  0.15407    1.27122     1.845769    2.427637    3.515911
age         -0.01326    0.01062     0.023028    0.035029    0.058339
female       0.01816    0.60436     0.913995    1.229570    1.836799
black       -2.32048    -1.54254    -1.125286   -0.719287   0.059661
other       -0.76844    0.12268     0.590388    1.047999    1.938521
educ        -0.06344    0.02805     0.078302    0.127620    0.220296
inc1k       -0.02488    -0.01487    -0.009555   -0.004332   0.005538
sigma2      46.27948    49.39755    51.142080   52.924603   56.581265
```

Task 2 Interpret Results Using Credit Intervals

- Compare results from the Bayesian linear regression with those obtained from OLS regression (e.g., are they different/similar, and if so, to what extent?). Use 95% credible intervals to interpret results for selected predictors, such as education and race.

```
> #Task 2 Run Confidence Intervals for Coefficients of Educ and Race
```

```
> confint(mntlhlth.model)
                    2.5 %           97.5 %
(Intercept)         3.65904611      10.16495216
age                 -0.04876665     0.03022966
female              -0.01869809     2.05442787
black               -3.67049817     -0.56860898
other               -1.30509710     2.44469055
educ                -0.39061397     0.01336304
inc1k               -0.02011207     0.01113800

> #Interpret results in document
>
```

For the OLS model:
We are 95% confident that the true coefficient for education is between -0.3906 and 0.0135. Since 0 is included in this interval, it leads us to conclude that education does not have a significant impact on mental health.

We are 95% confident that the true coefficient for black is between -3.6705 and -0.5686. We are 95% confident that the true coefficient for other is between -1.3051 and 2.4447.

For the MCMC model:
We are 95% confident that the true coefficient for education is between -.0634 and .2203.

We are 95% confident that the true coefficient for black is between -2.3205 and 0.0597.

We are 95% confident that the true coefficient for other is between -.7684 and 1.9385.

Comparison:
For the OLS model, we would conclude that there is not a significant effect on mental health for the variables other and education, because 0 is captured in the confidence interval.

For the MCMC model, we would conclude that there is not a significant effect on mental health for all three of the variables, education, black, and other, because 0 is captured within the interval.

Task 3 Run Bayesian Linear Regression with Informative Priors

- Use the MCMCregress function from the MCMCpack package to run a Bayesian linear regression of mental health on age, sex (male as the reference category), race (white as the reference category), education, and income using informative prior (e.g., N (2, 0.1)). Then compare the results from this model with those from the first Bayesian linear regression model.

```
> #Task 3 Run Bayesian Linear Regression (Informative Pairs)
> mcmc.model2 = MCMCregress(mntlhlth ~ age + female + black + other + educ + inc1k,
+           data = nmdta, burnin = 1000, mcmc = 10000, thin = 1,seed = 47304,
+           b0 = 2,
+           B0 = .25,
+           marginal.likelihood = "Chib95")

> summary(mcmc.model2)
```

```
Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

            Mean        SD   Naive    SE Time-series    SE
(Intercept) 4.722149    1.270431    1.270e-02         1.260e-02
age         0.003544    0.019167    1.917e-04         1.917e-04
female      1.118300    0.504697    5.047e-03         5.047e-03
black      -1.465131    0.722499    7.225e-03         7.375e-03
other       1.043492    0.858197    8.582e-03         8.582e-03
educ       -0.081252    0.087277    8.728e-04         8.728e-04
inc1k      -0.006084    0.007843    7.843e-05         7.843e-05
sigma2     50.705524    2.613728    2.614e-02         2.677e-02


2. Quantiles for each variable:

            2.5%        25%         50%         75%          97.5%
(Intercept) 2.21116     3.872749    4.723132    5.5832888    7.208801
age        -0.03457    -0.009474    0.003618    0.0163592    0.040789
female      0.13598     0.775348    1.114231    1.4604423    2.129010
black      -2.88051    -1.951285   -1.463214   -0.9811057   -0.051491
other      -0.66251     0.467246    1.044855    1.6157827    2.741485
educ       -0.25096    -0.140094   -0.081109   -0.0221236    0.090023
inc1k      -0.02138    -0.011390   -0.006088   -0.0007593    0.009153
sigma2     45.86517    48.905398   50.614323   52.4207106   56.030009


> BF = BayesFactor(mcmc.model,mcmc.model2)

> BF
The matrix of Bayes Factors is:
                  mcmc.model        mcmc.model2
mcmc.model        1.000             2.88
mcmc.model2       0.347             1.00

The matrix of the natural log Bayes Factors is:
                  mcmc.model        mcmc.model2
mcmc.model         0.00             1.06
mcmc.model2       -1.06             0.00

 mcmc.model :
   call =
MCMCregress(formula = mntlhlth ~ age + female + black + other +
    educ + inc1k, data = nmdta, burnin = 1000, mcmc = 10000,
    thin = 1, seed = 47304, b0 = 0, B0 = 1, marginal.likelihood = "Chib95")

   log marginal likelihood =  -2563.964


 mcmc.model2 :
   call =
```

```
MCMCregress(formula = mntlhlth ~ age + female + black + other +
    educ + inc1k, data = nmdta, burnin = 1000, mcmc = 10000,
    thin = 1, seed = 47304, b0 = 2, B0 = 0.25, marginal.likelihood =
"Chib95")

   log marginal likelihood =  -2565.021
```

Interpretation:
With a Bayes factor of 2.88, there is very weak evidence in favor of the first model with noninformative priors to the second model with an informative prior of N(2, 0.1).

Task 4 Produce Prediction

- Predict the response value for a 35-year-old white female with college education and sample median income. Please report the mean of the prediction distribution and the 95% credit interval using empirical percentiles.

```
> #Task 4 Make Predictions for 35-year-old white woman with 16 years
education and median income
> library(HDInterval)

> require(rjags)

> library(BEST)

> post.mat = as.matrix(mcmc.model)

> cut.mat = post.mat[,c("(Intercept)", "age",
'female','black','other','educ','inc1k')]

> pred.hyp <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] +
1*post.mat[,'female'] +
+   0*post.mat[,'black'] + 0*post.mat[,'other'] + 16*post.mat[,'educ'] +
median(nmdta$inc1k)*post.mat[,'inc1k']

> summary(pred.hyp)
   Min.     1st Qu.     Median      Mean        3rd Qu.     Max.
   2.735    4.237       4.543       4.542       4.845       6.271

> quantile(pred.hyp, c(.025,.5,.975))
      2.5%        50%         97.5%
      3.670528    4.543092    5.420971
```

The mean of mntlhlth is reported to be 4.542 days. We are 95% confident that the actual value for a 35-year-old white woman with a college degree and median income is between 3.6705 days and 5.4210 days.

Task 5 Calculate Difference in Prediction

- Make the same prediction for an otherwise similar male. Then calculate the difference in the predictions and construct its 95% credible interval. Please answer if there is no difference between these two predictions using results from the credible interval.

```
> #TASK 5 Calculate the difference in Predictions
> pred.hyp2 <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] +
0*post.mat[,'female'] +
+   0*post.mat[,'black'] + 0*post.mat[,'other'] + 16*post.mat[,'educ'] +
median(nmdta$inc1k)*post.mat[,'inc1k']

> summary(pred.hyp2)
      Min.       1st Qu.     Median      Mean        3rd Qu.     Max.
      1.780      3.314       3.629       3.624       3.941       5.316

> quantile(pred.hyp-pred.hyp2, c(.025,.5,.0975))
            2.5%        50%         9.75%
            0.01816011  0.91399525  0.32042325
```

We are 95% confident that difference in predictions of mental health days between a 35-year-old white female with a college degree and median income and a similar male is between 0.0182 and 0.3204. Since 0 is not captured in this confidence interval, there is a significant difference between the two predictions.

Close Out

- Close out the log file

```
> #Close log
> save(nmdta, file = "Assignment_09.rdata")

> sink()
```

R-Script

```
#
source("/Users/burrisfaculty/Desktop/DSCode/SOC686/Shepherd_Lab0
9_SOC686.r", echo=T, max.deparse.length=10000)
library(foreign)


#Open Log and read in data
setwd("/Users/burrisfaculty/Desktop/DSCode/SOC686")
sink("Shepherd_asgn09.log", split=T)
rm(list=ls(all=TRUE))
mygss <- read.dta("gsscum7212teach.dta", convert.factor=F)

#MANAGE DATA AND RUN LOGIT
#SELECT DATA
useddta <- subset(mygss,
                select=c(mntlhlth, age, sex, race, educ, inc1k))

#Create dummy variables female (male = 0)
useddta$female <- as.numeric(useddta$sex==2)
```

```
useddta$male <- as.numeric(useddta$sex == 1)

#Create Binary Indicator Variables for Multi-Category Nomial
Variables

useddta$white <- ifelse(useddta$race == 1, 1, 0)
useddta$black <- ifelse(useddta$race == 2, 1, 0)
useddta$other <- ifelse(useddta$race == 3, 1, 0)

nmdta <- useddta[complete.cases(useddta),] #no missing data

#summarize data
summary(nmdta$mntlhlth)
summary(nmdta$inc1k)
summary(nmdta$age)
table(nmdta$female)
table(nmdta$white)

#Run OLS
mntlhlth.model <- lm(formula = mntlhlth ~ age + female + black +
other + educ + inc1k, data = nmdta )
(summary(mntlhlth.model))

#TASK 1 Run Bayesian Linear Reg (Non-Informative Pairs)
library(MCMCpack)
mcmc.model = MCMCregress(mntlhlth ~ age + female + black + other
+ educ + inc1k,
                         data = nmdta, burnin = 1000, mcmc =
10000, thin = 1,seed = 47304,b0 = 0, B0 = 1,marginal.likelihood
= "Chib95")
summary(mcmc.model)

#Task 2 Run Confidence Intervals for Coefficients of Educ and
Race
confint(mntlhlth.model)

#Interpret results in document

#Task 3 Run Bayesian Linear Regression (Informative Pairs)
mcmc.model2 = MCMCregress(mntlhlth ~ age + female + black +
other + educ + inc1k,
                          data = nmdta, burnin = 1000, mcmc =
10000, thin = 1,seed = 47304,
                          b0 = 2,
                          B0 = .25,
                          marginal.likelihood = "Chib95")
summary(mcmc.model2)
```

```
BF = BayesFactor(mcmc.model,mcmc.model2)
BF
#Task 4 Make Predictions for 35-year-old white woman with 16
years education and median income
library(HDInterval)
require(rjags)
library(BEST)
post.mat = as.matrix(mcmc.model)
cut.mat = post.mat[,c("(Intercept)", "age",
'female','black','other','educ','inc1k')]
pred.hyp <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] +
1*post.mat[,'female'] +
  0*post.mat[,'black'] + 0*post.mat[,'other'] +
16*post.mat[,'educ'] + median(nmdta$inc1k)*post.mat[,'inc1k']
summary(pred.hyp)
quantile(pred.hyp, c(.025,.5,.975))
#TASK 5 Calculate the difference in Predictions
pred.hyp2 <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] +
0*post.mat[,'female'] +
  0*post.mat[,'black'] + 0*post.mat[,'other'] +
16*post.mat[,'educ'] + median(nmdta$inc1k)*post.mat[,'inc1k']
summary(pred.hyp2)
quantile(pred.hyp-pred.hyp2, c(.025,.5,.0975))
#Close log
save(nmdta, file = "Assignment_09.rdata")
sink()
```

Log File

```
> rm(list=ls(all=TRUE))

> mygss <- read.dta("gsscum7212teach.dta", convert.factor=F)

> #MANAGE DATA AND RUN LOGIT
> #SELECT DATA
> useddta <- subset(mygss,
+                 select=c(mntlhlth, age, sex, race, educ, inc1k))

> #Create dummy variables female (male = 0)
> useddta$female <- as.numeric(useddta$sex==2)

> useddta$male <- as.numeric(useddta$sex == 1)

> #Create Binary Indicator Variables for Multi-Category Nomial
Variables
>
> useddta$white <- ifelse(useddta$race == 1, 1, 0)

> useddta$black <- ifelse(useddta$race == 2, 1, 0)
```

```
> useddta$other <- ifelse(useddta$race == 3, 1, 0)

> nmdta <- useddta[complete.cases(useddta),] #no missing data

> #summarize data
> summary(nmdta$mntlhlth)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.000   0.000   0.000   3.991   5.000  30.000

> summary(nmdta$inc1k)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.245  15.921  28.157  39.404  48.475 155.140

> summary(nmdta$age)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  18.00   31.00   42.00   41.81   51.00   84.00

> table(nmdta$female)

  0   1
381 369

> table(nmdta$white)

  0   1
160 590

> #Run OLS
> mntlhlth.model <- lm(formula = mntlhlth ~ age + female + black +
other + educ + inc1k, data = nmdta )

> (summary(mntlhlth.model))

Call:
lm(formula = mntlhlth ~ age + female + black + other + educ +
    inc1k, data = nmdta)

Residuals:
   Min     1Q Median     3Q    Max
-6.888 -4.069 -2.818  0.601 27.711

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.911999   1.656997   4.171 3.38e-05 ***
age         -0.009268   0.020120  -0.461  0.64517
female       1.017865   0.528007   1.928  0.05427 .
black       -2.119554   0.790024  -2.683  0.00746 **
other        0.569797   0.955038   0.597  0.55094
educ        -0.188625   0.102889  -1.833  0.06716 .
inc1k       -0.004487   0.007959  -0.564  0.57309
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.105 on 743 degrees of freedom
Multiple R-squared:  0.02043,    Adjusted R-squared:  0.01252
F-statistic: 2.582 on 6 and 743 DF,  p-value: 0.01753


> #TASK 1 Run Bayesian Linear Reg (Non-Informative Pairs)
> library(MCMCpack)

> mcmc.model = MCMCregress(mntlhlth ~ age + female + black + other + 
educ + inc1k,
+                          data = nmdta, burnin = 1000, mcmc = 10000, 
thin = 1,seed = 47304,b0 = 0, B0 = 1,marginal.likelihood = "Chib95")

> summary(mcmc.model)

Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

                 Mean       SD  Naive SE Time-series SE
(Intercept)  1.843978 0.853808 8.538e-03      8.498e-03
age          0.022910 0.018192 1.819e-04      1.819e-04
female       0.917328 0.462010 4.620e-03      4.620e-03
black       -1.127879 0.610042 6.100e-03      6.231e-03
other        0.586758 0.687921 6.879e-03      6.879e-03
educ         0.078028 0.072803 7.280e-04      7.280e-04
inc1k       -0.009583 0.007784 7.784e-05      7.784e-05
sigma2      51.207177 2.644329 2.644e-02      2.714e-02

2. Quantiles for each variable:

                2.5%      25%       50%       75%     97.5%
(Intercept)  0.15407  1.27122  1.845769  2.427637  3.515911
age         -0.01326  0.01062  0.023028  0.035029  0.058339
female       0.01816  0.60436  0.913995  1.229570  1.836799
black       -2.32048 -1.54254 -1.125286 -0.719287  0.059661
other       -0.76844  0.12268  0.590388  1.047999  1.938521
educ        -0.06344  0.02805  0.078302  0.127620  0.220296
inc1k       -0.02488 -0.01487 -0.009555 -0.004332  0.005538
sigma2      46.27948 49.39755 51.142080 52.924603 56.581265


> #Task 2 Run Confidence Intervals for Coefficients of Educ and Race
> confint(mntlhlth.model)
                  2.5 %      97.5 %
(Intercept)  3.65904611 10.16495216
```

```
age         -0.04876665  0.03022966
female      -0.01869809  2.05442787
black       -3.67049817 -0.56860898
other       -1.30509710  2.44469055
educ        -0.39061397  0.01336304
inc1k       -0.02011207  0.01113800

> #Interpret results in document
>
> #Task 3 Run Bayesian Linear Regression (Informative Pairs)
> mcmc.model2 = MCMCregress(mntlhlth ~ age + female + black + other +
educ + inc1k,
+                            data = nmdta, burnin = 1000, mcmc = 10000,
thin = 1,seed = 47304,
+                            b0 = 2,
+                            B0 = .25,
+                            marginal.likelihood = "Chib95")

> summary(mcmc.model2)

Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

                 Mean       SD  Naive SE Time-series SE
(Intercept)  4.722149 1.270431 1.270e-02      1.260e-02
age          0.003544 0.019167 1.917e-04      1.917e-04
female       1.118300 0.504697 5.047e-03      5.047e-03
black       -1.465131 0.722499 7.225e-03      7.375e-03
other        1.043492 0.858197 8.582e-03      8.582e-03
educ        -0.081252 0.087277 8.728e-04      8.728e-04
inc1k       -0.006084 0.007843 7.843e-05      7.843e-05
sigma2      50.705524 2.613728 2.614e-02      2.677e-02

2. Quantiles for each variable:

                2.5%       25%       50%        75%     97.5%
(Intercept)  2.21116  3.872749  4.723132  5.5832888  7.208801
age         -0.03457 -0.009474  0.003618  0.0163592  0.040789
female       0.13598  0.775348  1.114231  1.4604423  2.129010
black       -2.88051 -1.951285 -1.463214 -0.9811057 -0.051491
other       -0.66251  0.467246  1.044855  1.6157827  2.741485
educ        -0.25096 -0.140094 -0.081109 -0.0221236  0.090023
inc1k       -0.02138 -0.011390 -0.006088 -0.0007593  0.009153
sigma2      45.86517 48.905398 50.614323 52.4207106 56.030009


> BF = BayesFactor(mcmc.model,mcmc.model2)
```

```
> BF
The matrix of Bayes Factors is:
            mcmc.model mcmc.model2
mcmc.model       1.000        2.88
mcmc.model2      0.347        1.00

The matrix of the natural log Bayes Factors is:
            mcmc.model mcmc.model2
mcmc.model        0.00        1.06
mcmc.model2      -1.06        0.00

 mcmc.model :
   call =
MCMCregress(formula = mntlhlth ~ age + female + black + other +
    educ + inc1k, data = nmdta, burnin = 1000, mcmc = 10000,
    thin = 1, seed = 47304, b0 = 0, B0 = 1, marginal.likelihood =
"Chib95")

   log marginal likelihood =  -2563.964


 mcmc.model2 :
   call =
MCMCregress(formula = mntlhlth ~ age + female + black + other +
    educ + inc1k, data = nmdta, burnin = 1000, mcmc = 10000,
    thin = 1, seed = 47304, b0 = 2, B0 = 0.25, marginal.likelihood =
"Chib95")

   log marginal likelihood =  -2565.021


> #Task 4 Make Predictions for 35-year-old white woman with 16 years
education and median income
> library(HDInterval)

> require(rjags)

> library(BEST)

> post.mat = as.matrix(mcmc.model)

> cut.mat = post.mat[,c("(Intercept)", "age",
'female','black','other','educ','inc1k')]

> pred.hyp <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] +
1*post.mat[,'female'] +
+   0*post.mat[,'black'] + 0*post.mat[,'other'] + 16*post.mat[,'educ']
+ median(nmdta$inc1k)*post.mat[,'inc1k']

> summary(pred.hyp)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
  2.735   4.237   4.543   4.542   4.845   6.271

> quantile(pred.hyp, c(.025,.5,.975))
    2.5%      50%    97.5% 
3.670528 4.543092 5.420971 

> #TASK 5 Calculate the difference in Predictions
> pred.hyp2 <- 1*post.mat[,"(Intercept)"]+ 35*post.mat[,'age'] + 
0*post.mat[,'female'] +
+   0*post.mat[,'black'] + 0*post.mat[,'other'] + 16*post.mat[,'educ'] 
+ median(nmdta$inc1k)*post.mat[,'inc1k']

> summary(pred.hyp2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  1.780   3.314   3.629   3.624   3.941   5.316 

> quantile(pred.hyp-pred.hyp2, c(.025,.5,.0975))
      2.5%        50%      9.75% 
0.01816011 0.91399525 0.32042325 

> #Close log
> save(nmdta, file = "Assignment_09.rdata")

> sink()
```