

Question 2

a) R code:

```
# Question 2
setwd("~/Desktop/STAT 371/a1")
name=read.table("A1_variates.txt",header = TRUE)
attach(name)

# a)
bv <- c(-250000, 50, -200, 12000, 200000) #creates the Beta vector
set.seed(20884580)
n <- length(size)
sigma <- 15000
ei <- rnorm(n,0,sigma)
y <- bv[1]+ bv[2]*size + bv[3]*age + bv[4]*employees + bv[5]*col + ei
y
```

output:

```
[1] 198985.60 223391.14 64428.97 217337.18 165669.78 21113.32 23529.00 80104.40 156791.59 120950.98 24754.71
[12] 32663.89 24960.01 109244.98 91352.35 135724.68 42175.64 191351.17 207042.44 164114.27 153564.18 181715.28
[23] 222266.42 -27350.39
```

b) R code:

```
# b)
A1sim.lm = lm(y~size+age+employees+col)
summary(A1sim.lm)
```

output:

```
Call:
lm(formula = y ~ size + age + employees + col)

Residuals:
    Min       1Q   Median       3Q      Max
-32285   -5884    -211    9282   29974

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -320631.98   74983.15  -4.276  0.000408 ***
size          106.75     21.96    4.860  0.000109 ***
age         -137.11     505.05   -0.271  0.788954
employees    6846.56    2134.41    3.208  0.004633 **
col         244468.23   70429.60    3.471  0.002558 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14810 on 19 degrees of freedom
Multiple R-squared:  0.9699,    Adjusted R-squared:  0.9635
F-statistic: 152.8 on 4 and 19 DF,  p-value: 3.644e-14
```