M07_activity

Hannah Valenty

2024-07-23

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
df <- read.csv('data/expectancy.csv')

df3<-select(df,Life.expectancy, Adult.Mortality,
infant.deaths,HIV.AIDS,BMI, GDP,Schooling)%>%
na.omit()
```

Task 1.A

```
des_matrix <- as.matrix(cbind(rep(1, 151), df3$Adult.Mortality,
                                  df3$infant.deaths, df3$HIV.AIDS,
                                  df3$BMI, df3$GDP, df3$Schooling))
head(des_matrix)
```

```
##      [,1] [,2] [,3] [,4] [,5]       [,6] [,7]
## [1,]    1  263   62  0.1 19.1   584.2592 10.1
## [2,]    1   74    0  0.1 58.0  3954.2278 14.2
## [3,]    1   19   21  0.1 59.5  4132.7629 14.4
## [4,]    1  335   66  1.9 23.3  3695.7937 11.4
## [5,]    1   13    0  0.2 47.7 13566.9541 13.9
## [6,]    1  116    8  0.1 62.8 13467.1236 17.3
```

Task 1.B

```
xtx <- t(des_matrix)%*%des_matrix
inverse_xtx <- solve(xtx)
xty <- t(des_matrix)%*%df3$Life.expectancy
beta <- inverse_xtx %*% xty
beta
```

```
##               [,1]
## [1,]  5.590038e+01
## [2,] -2.916393e-02
## [3,] -1.791052e-03
## [4,] -9.045282e-01
## [5,] -9.302135e-04
## [6,]  3.851518e-05
## [7,]  1.577959e+00
```

Task 2

```
mod_df3 <- lm(Life.expectancy~Adult.Mortality + infant.deaths + HIV.AIDS +
                BMI + GDP + Schooling, data=df3)
coef(mod_df3)
```

```
##     (Intercept) Adult.Mortality   infant.deaths        HIV.AIDS             BMI
##    5.590038e+01   -2.916393e-02   -1.791052e-03   -9.045282e-01   -9.302135e-04
##             GDP       Schooling
##    3.851518e-05    1.577959e+00
```

Task 3

```
boot_estimate <- replicate(10000,{
boot_sam <- df3[sample(1:nrow(df3), nrow(df3),
                    replace=TRUE),]
boot_model <- lm(Life.expectancy~Adult.Mortality + infant.deaths + HIV.AIDS +
                BMI + GDP + Schooling, data=df3)
coef(boot_model)
})

estm <- data.frame(t(boot_estimate))

summarize(estm, mean_b0=mean(X.Intercept.), mean_b1=mean(Adult.Mortality),
          mean_b2=mean(infant.deaths), mean_b3=mean(HIV.AIDS),
          mean_b4=mean(BMI), mean_b5=mean(GDP), mean_b6=mean(Schooling))
```

```
##    mean_b0     mean_b1      mean_b2    mean_b3       mean_b4      mean_b5
## 1 55.90038 -0.02916393 -0.001791052 -0.9045282 -0.0009302135 3.851518e-05
##    mean_b6
## 1 1.577959
```