IPS9 in R: Multiple regression (Chapter 11)

Margaret Chien and Nicholas Horton (nhorton@amherst.edu)

July 24, 2018

Introduction and background

These documents are intended to help describe how to undertake analyses introduced as examples in the Ninth Edition of *Introduction to the Practice of Statistics* (2017) by Moore, McCabe, and Craig.

More information about the book can be found here. The data used in these documents can be found under Data Sets in the Student Site. This file as well as the associated R Markdown reproducible analysis source file used to create it can be found at https://nhorton.people.amherst.edu/ips9/.

This work leverages initiatives undertaken by Project MOSAIC (http://www.mosaic-web.org), an NSF-funded effort to improve the teaching of statistics, calculus, science and computing in the undergraduate curriculum. In particular, we utilize the `mosaic` package, which was written to simplify the use of R for introductory statistics courses. A short summary of the R needed to teach introductory statistics can be found in the mosaic package vignettes (http://cran.r-project.org/web/packages/mosaic). A paper describing the mosaic approach was published in the *R Journal*: https://journal.r-project.org/archive/2017/RJ-2017-024.

Chapter 11: Multiple Regression

This file replicates the analyses from Chapter 11: Multiple Regression.

First, load the packages that will be needed for this document:

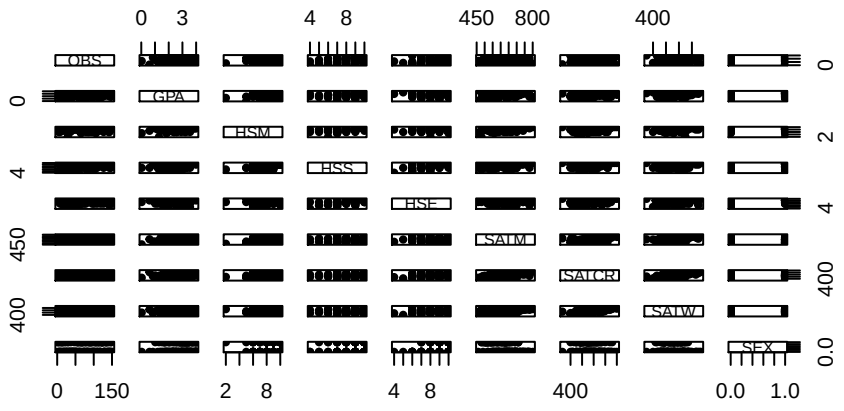
```
library(mosaic)
library(readr)
```

Section 11.1: Inference for multiple regression

```
GPA <- read_csv("https://nhorton.people.amherst.edu/ips9/data/chapter11/EG11-01GPA.csv")
### Figure 11.4, page 621
options(digits = 2)
cor(GPA)
```

```
##          OBS    GPA    HSM   HSS  HSE   SATM SATCR   SATW    SEX
## OBS    1.000 -0.018  0.059 0.026 0.12 -0.083 -0.04 -0.057  0.093
## GPA   -0.018  1.000  0.420 0.443 0.36  0.330  0.25  0.223  0.089
## HSM    0.059  0.420  1.000 0.670 0.48  0.325  0.15  0.072 -0.034
## HSS    0.026  0.443  0.670 1.000 0.70  0.215  0.22  0.161  0.096
## HSE    0.117  0.359  0.485 0.695 1.00  0.134  0.26  0.185  0.182
## SATM  -0.083  0.330  0.325 0.215 0.13  1.000  0.58  0.551 -0.408
## SATCR -0.040  0.251  0.150 0.215 0.26  0.579  1.00  0.734 -0.151
## SATW  -0.057  0.223  0.072 0.161 0.19  0.551  0.73  1.000 -0.098
## SEX    0.093  0.089 -0.034 0.096 0.18 -0.408 -0.15 -0.098  1.000
```

```
### Figure 11.5
pairs(~ ., data = GPA,
      pch=20, main="GPA Scatterplot Matrix")
```



Section 11.2: A case study