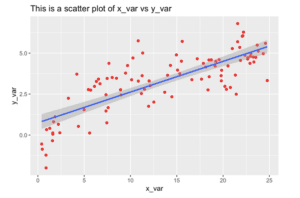
This involves downloading an Excel file from the internet and importing it into the RStudio environment. They will then write code to give summary statistics and the five-number summary.



The Sharp Sight website has a nice explanation of using ggplot for creating scatter plots.

# Follow the usual procedures for code submission.

library(readxl) # Download this package if you dont have it already

# Go to import dataset, then "From Excel" and find you file "backpack" on you computer.

backpack <- read\_excel("//ais-main/users/kevin-smith/Desktop/backpack.xls")

View(backpack)

# you need to make it a data.frame first with this command.

as.data.frame(backpack)

# don't forget to load ggplot, dplyr and/or tidyverse.

summary(backpack) # gives summary statistics

fivenum(backpack$boyweight) # because this is a data.frame, I need to use the "$" to indicate which column to analyse.

fivenum(backpack$packweight)

library(ggplot2)

scatter145=ggplot(data=backpack, aes(boyweight,packweight)) +

geom\_point()

scatter145 # this just give a basic scatterplot

# now we add some color

scatter145b=ggplot(data=backpack, aes(boyweight,packweight,colour=body.wt)) +

geom\_point()

scatter145b

# Next we add a small color box.

scatter145c = scatter145b+ geom\_point(size=2) +

xlab("Body Weight (lb)") +

ylab("Pack weight (lb)") +

ggtitle("Backpack Weight")

scatter145c

# Here we will add the confidence intervals and regression line with the "lm" command

# lm means linear model

scatter145d=scatter145c+ geom\_point(size=3) +

xlab("Body Weight (lb)") +

ylab("Pack weight (lb)")+

ggtitle("Backpack Weight")+

geom\_smooth(method = "lm")

scatter145d

Go here to read the article on creating another scatter plot.