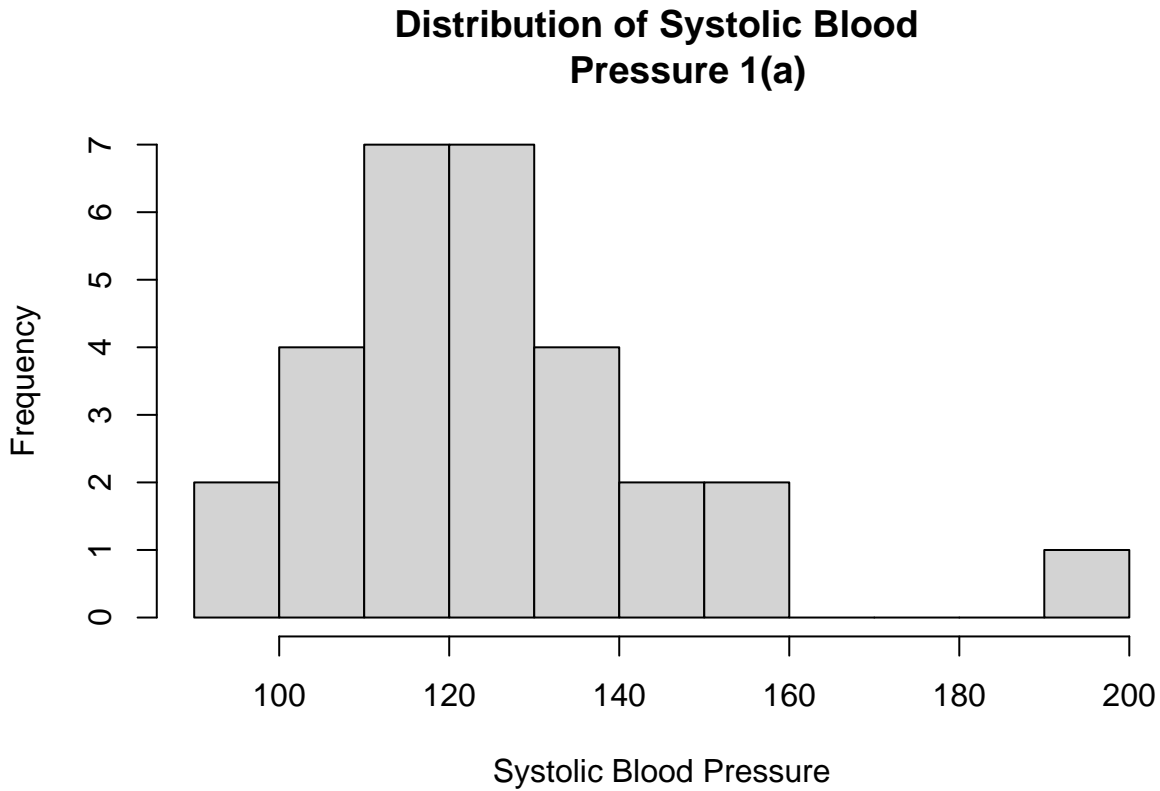# STA 138 Discussion 1 Solutions

Winter 2023 Andrew Farris

1. (Note that solutions here illustrate both base R and ggplots; use of either or both is fine!)
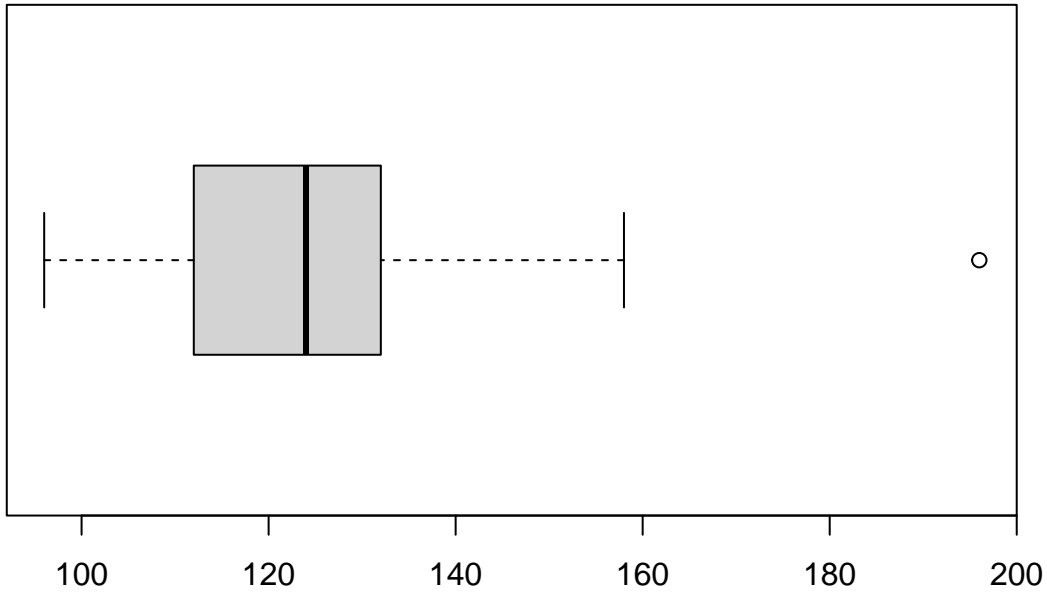
(a)

Distribution of Systolic Blood Pressure 1(a)

(b) With the exception of a single relative outlier, the data appears fairly symmetric.

(c)

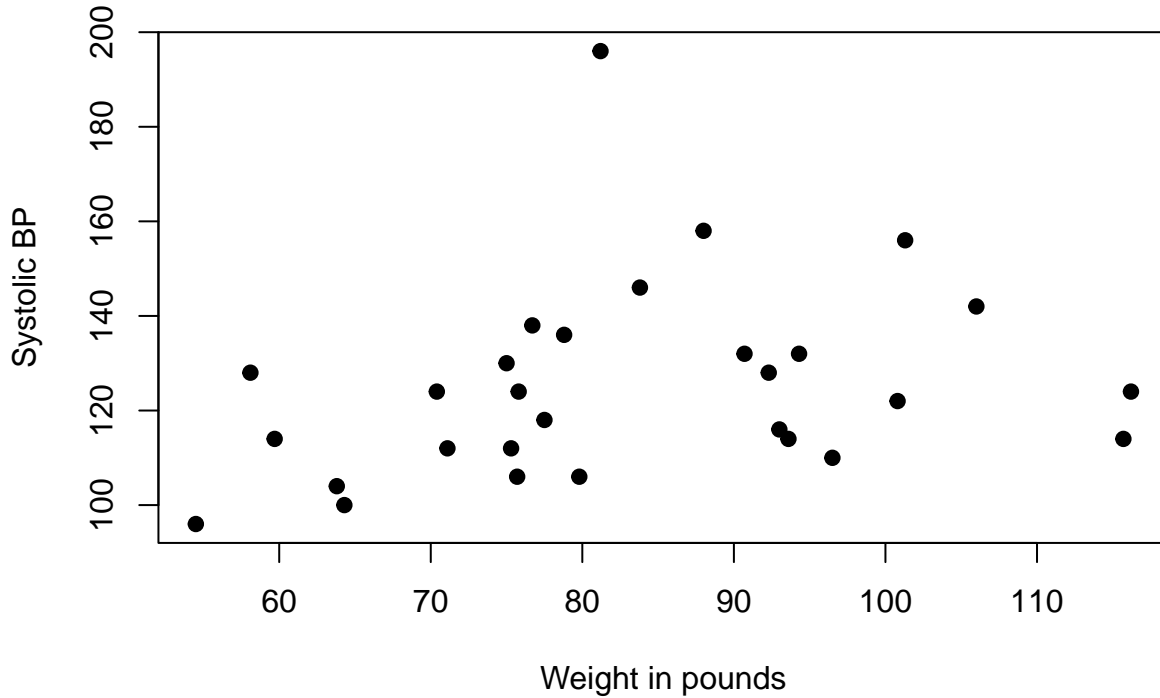**Distribution of Systolic Blood Pressure 1(c)**

(d) There is one clear, unusually large observation apparent from the boxplot.
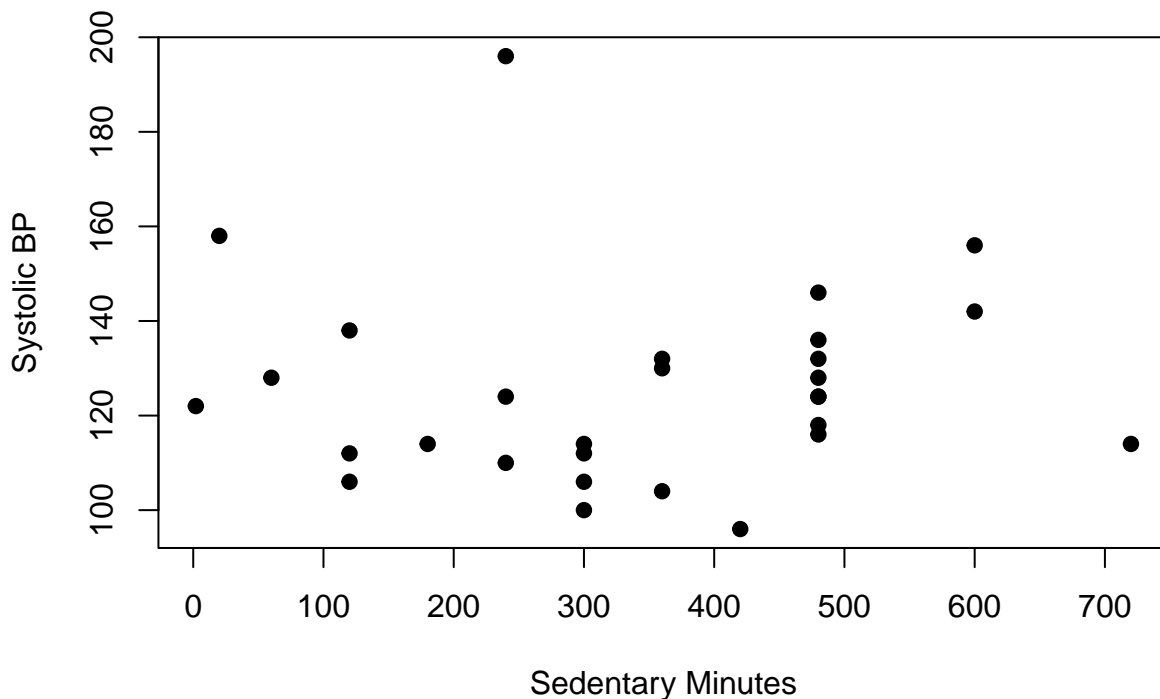
2.

(a)

2(a):Relationship between Systolic Blood Pressure and Weight

(b) There seems to be a slight positive trend; as weight increases, the average systolic blood pressure appears to increase slightly as well.
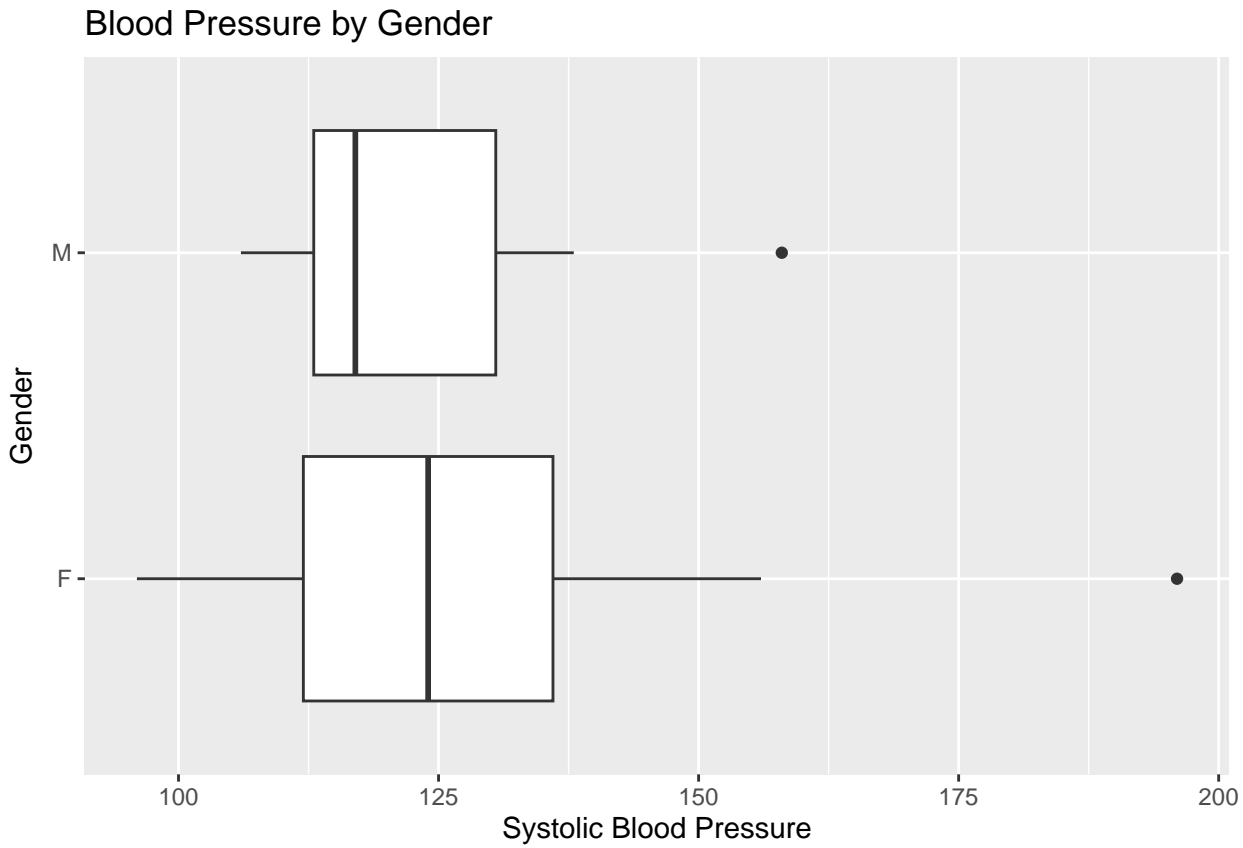
(c)

2(c):Relationship between Systolic Blood Pressure and Sedentary Min

(d) There does not seem to be a notable trend here.

3.

(a)

Blood Pressure by Gender
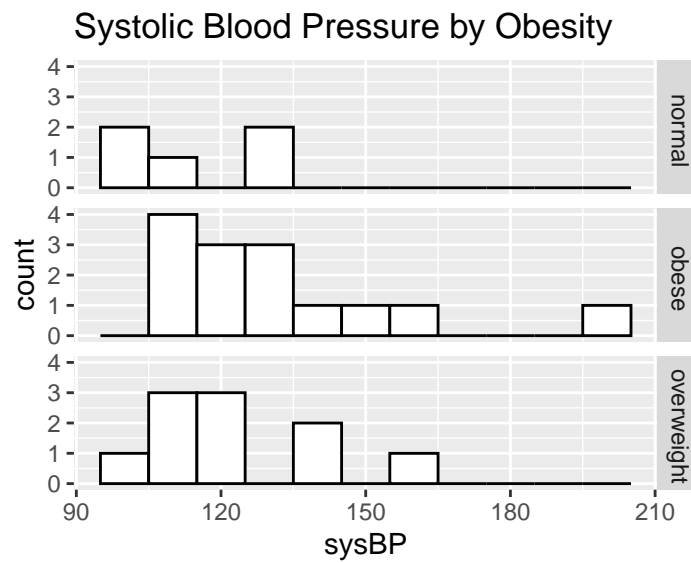

(b) The histogram for males and females appear to completely overlap, so they show no significant difference.

(c)

Systolic Blood Pressure by Obesity

(d) The 'Normal' weight category appears to have lower systolic blood pressures on average than 'obese' and 'overweight,' but there does not appear to be much of a difference between 'overweight' and 'obese.'

4.

(a) The average systolic blood pressure is: 125.4483.

(b) A table of results follows:

	marriage	ave sysBP
1	divorced	133.50
2	married	124.77
3	nevermarried	115.50
4	other	126.00
5	widowed	132.75

(c) A table of results follows:

	marriage	sd sysBP
1	divorced	22.85
2	married	21.34
3	nevermarried	19.27
4	other	22.21
5	widowed	8.81

(d) A table of results follows:

	marriage	number
1	divorced	12
2	married	52
3	nevermarried	16
4	other	12
5	widowed	8

(e) The five number summary is: 54.5 (min), 75 (Q2), 79.800003 (Median), 93.599998 (Q3), 116.2 (Max)

Code Appendix

```r
library(ggplot2)
patients101 = read.csv("patients101.csv")
#Problem 1 (a)
hist(patients101$sysBP, breaks=10, xlab = "Systolic Blood Pressure",main = "Distribution of Systolic Blood
    Pressure 1(a)")
#ggplot(patients101, aes(x = sysBP)) + geom_histogram(binwidth = 5,color = "black",fill = "white")+
#  xlab("Systolic Blood Pressure")+ggtitle("1(a): Distribution of systolic blood pressure")
#Problem 1 (c)
boxplot(patients101$sysBP, horizontal = TRUE,main = "Distribution of Systolic Blood Pressure 1(c)")
#ggplot(patients101, aes(y=sysBP, x = factor("")))+ geom_boxplot() + ylab("Systolic Blood Pressure") + xlab("
#Problem 2 (a)
plot(patients101$weight, patients101$sysBP, main = "2(a):Relationship between Systolic Blood
    Pressure and Weight",xlab = "Weight in pounds",ylab = "Systolic BP",pch = 19)
#qplot(weight, sysBP, data = patients101) + ggtitle("2(a):Relationship between Systolic Blood Pressure and Wei
#Problem 2 (c)
plot(patients101$sedmins, patients101$sysBP, main = "2(c):Relationship between Systolic Blood Pressure and Sed
#qplot(sedmins, sysBP, data = patients101) + ggtitle("2(c):Relationship between Systolic Blood Pressure and Se
#Problem 3 (a)
#boxplot(sysBP ~ gender, data = patients101, main = "Blood Pressure by Gender",horizontal = TRUE)
ggplot(patients101,aes(y = sysBP,x =gender)) + geom_boxplot() + ylab("Systolic Blood Pressure") + xlab("Gender
#Problem 3 (c)
ggplot(patients101, aes(x = sysBP)) + geom_histogram(binwidth = 10,color = "black",fill = "white") + facet_gri
#Problem 4
#part(a)
aveBP = round(mean(patients101$sysBP),4)
#part(b)
all.ave = aggregate(sysBP ~ marriage, data = patients101, mean)
#part(c)
all.sd = aggregate(sysBP ~ marriage, data = patients101, sd)
#part(d)
all.n= aggregate(sysBP ~ marriage, data = patients101, length)
#part(e)
fns.weight = fivenum(patients101$weight)
```