Homework 4
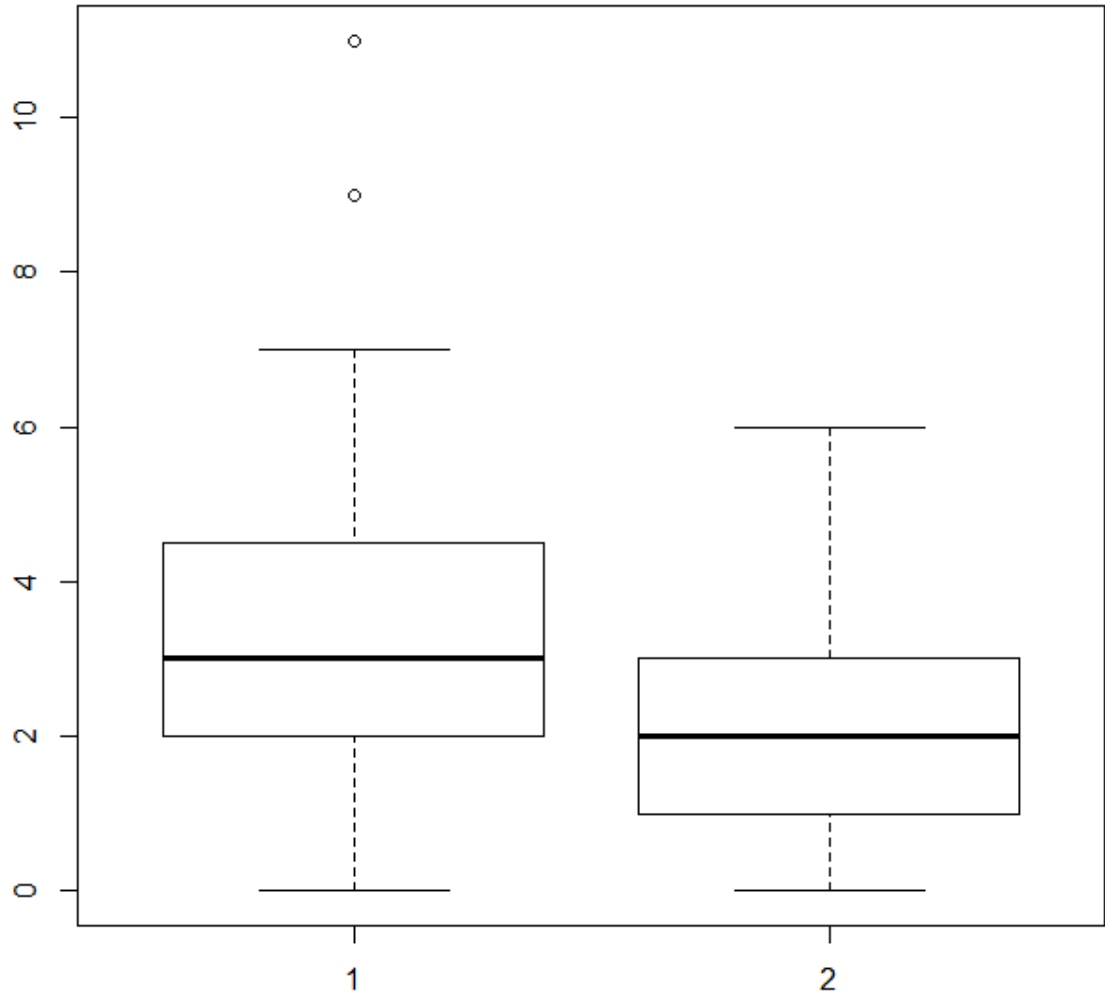
1. (a)

(b)Yes. There are 2 outliers in the data of women's goals.

(c) Empirical rule works well for men's data. For women's data, we have only one observation that is bigger than the 3-sigma cutoff, but the empirical rule still works here.

(d)

Men	Proportion	Chebycheffs Proportion
Within 1 Sd	0.671875	At least 0
Within 2 Sd	0.96875	At least 0.75
Within 3 Sd	1	At least 0.89

Women	Proportion	Chebycheffs Proportion
Within 1 Sd	0.75	At least 0
Within 2 Sd	0.9375	At least 0.75
Within 3 Sd	0.96875	At least 0.89

So from the table we can see that Chebycheffs rule works here. A little bit conservative though.

2.

(a)

	Cigarette use yes	Cigarette use no
Alcohol use Yes	74.35%	25.65%
Alcohol use No	14.07%	85.93%

(b)

	Cigarette use yes	Cigarette use no
Alcohol use Yes	96.92%	64.02%
Alcohol use No	3.08%	35.98%

(c) if there is no association. I would expect that each row in (a) has the same percentage, and each column in (b) has the same percentage.

(d) They are dependent.

The odds ratio is 1449/500/(46/281)=17.7. This high ratio implies there exists some association between them.

3.

	Case	control
Contraception	62	261
Non	386	1467

Odds ratio is a good measure of association. Odds ratio=62/261/(386/1467)=0.9028.

Others may just focus on a row or column, therefore can not find the relationship between the two variables.