Professor Notes About the "Univariate EDA - Cat" Homework

- In these questions I noted what I considered to be the major characteristics of the results. Your major characteristics may differ somewhat from mine. However, you should NOT just relist all of the information in the tables.
- Note how each table and figure is labeled and referred to in the answers. Also note that if abbreviations are used for levels of the cateorical variable that you should present some sort of key for what those abbreviations represent. In my case, I presented the key in the table and figure labels.
- You should use no more than one decimal place for percentages (or three decimals for proportions).

Chemical Waste Disposal

Most respondents thought that the Environmental Protection Agency was the agency that disposed of chemical wastes followed by the Occupational Safety and Health Administration (Table 1, Figure 1). Relatively few people thought the Department of Transportation or National Institutes of Health disposed of chemical wastes.

Table 1. Percentages table for perceived choice of agency that disposes of chemical waste. Abbreviations are: d=Department of Transportation, e=Environmental Protection Agency, n=National Institutes of Health, o=Occupational Safety and Health Administration, and u=unanswered.

```
agency
   d    e    n    o    u
10.2 47.7 12.5 28.4  1.1
```

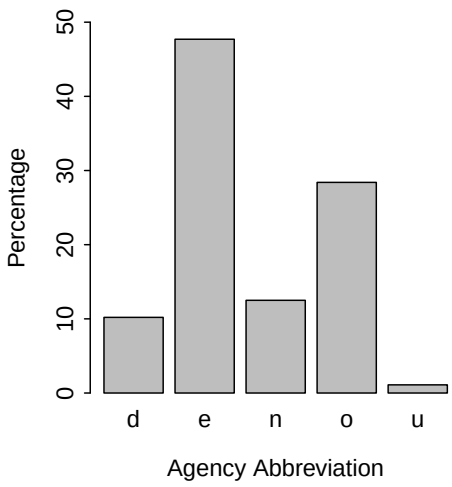

Figure 1. Bar chart of percentages by perceived choice of agency that disposes of chemical waste. Abbreviations are in Table 1.

Water Usage

Most students left the water on for between 6 and 15 minutes, with students approximately evenly split between 6-10 and 11-15 minutes Table 2. Few students left the water on for less than 5 minutes before entering the shower.

Table 2. Percentage table for how much time students at Rice University leave the water on before entering the shower. Note that A=0-5 mins, B=6-10 mins, C=11-15 mins, and D=more than 15 mins.

```
water
   A    B    C    D
 5.4 37.7 40.0 16.9
```

R Appendix.

```
library(NCStats)
setwd('C:/aaaWork/Books/IntroStats/HW/')
d <- read.csv("Agency.csv")
( a.tbl <- xtabs(~agency,data=d) )
( p.tbl <- percTable(a.tbl) )
barplot(p.tbl,xlab="Agency Abbreviation",ylab="Percentage",ylim=c(0,50))

df <- read.csv("WaterUsage.csv")
freq <- xtabs(~water,data=df)
perc <- percTable(freq,digits=1)
```