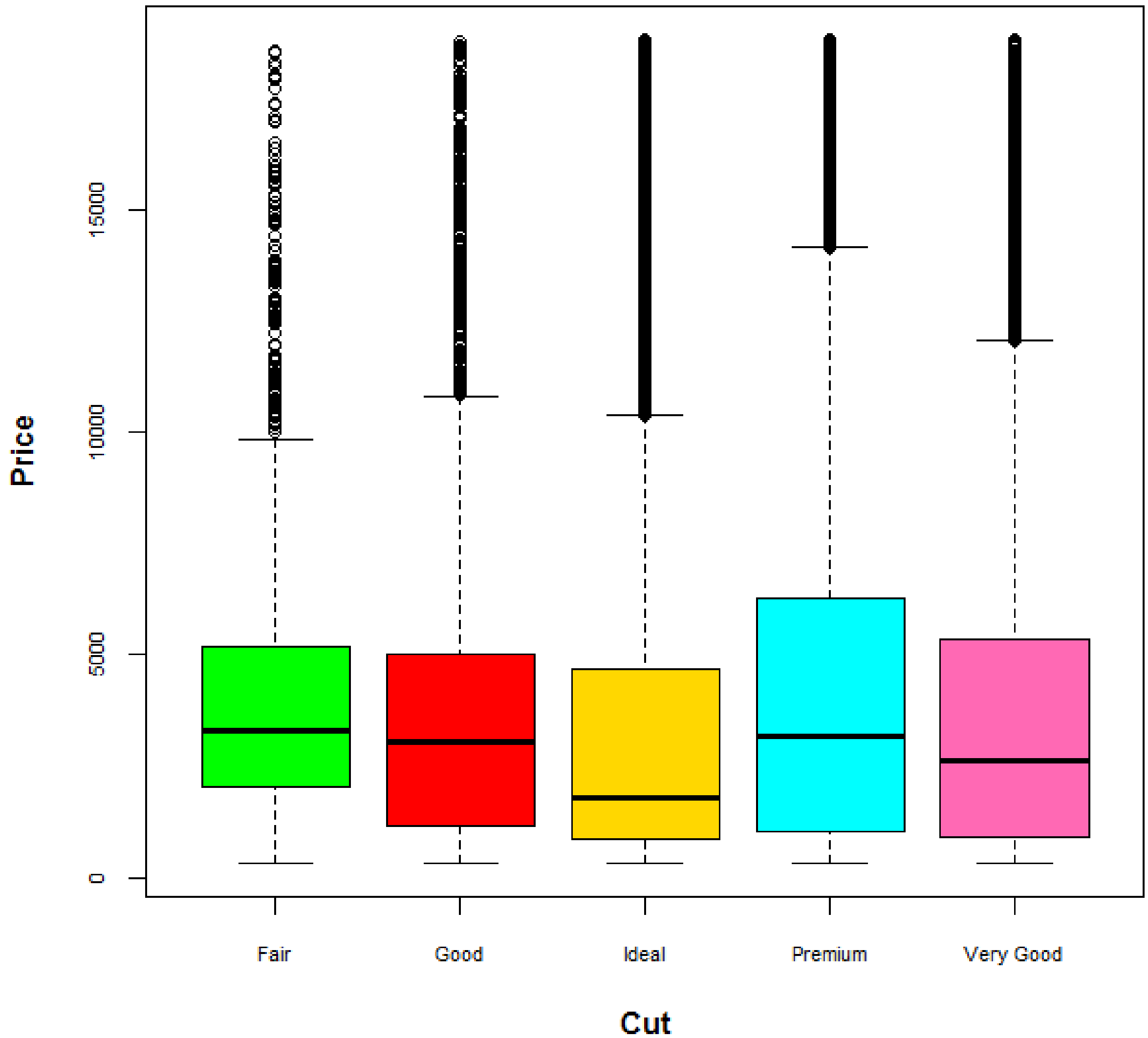
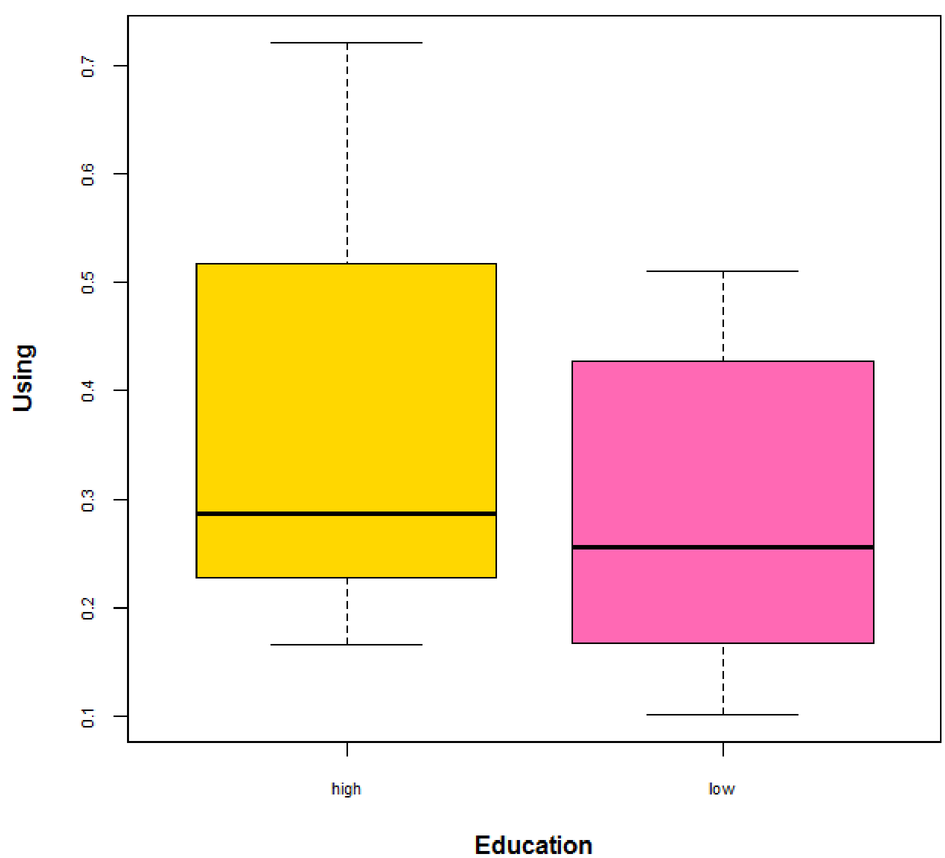
```
> #setwd("H:/EEB697/homework3-master/homework3-master")
################################### Q1 #####################################

> diamonds<-read.csv("diamond.csv")
> head(diamonds)
  price       cut carat
1   326     Ideal  0.23
2   326   Premium  0.21
3   327      Good  0.23
4   334   Premium  0.29
5   335      Good  0.31
6   336 Very Good  0.24
> cut_levels <- diamonds$cut["Ideal"]
> Fair_cut <- diamonds[diamonds[, "cut"] == "Fair",]
> Good_cut <- diamonds[diamonds[, "cut"] == "Good",]
> Ideal_cut <- diamonds[diamonds[, "cut"] == "Ideal",]
> Premium_cut <- diamonds[diamonds[, "cut"] == "Premium",]
> Very_Good_cut <- diamonds[diamonds[, "cut"] == "Very Good",]
>
> dev.new()
NULL
> plot(diamonds$price~diamonds$cut,
+      col=c("green","red","gold","cyan","hotpink"),
+      cex.axis=.7,xlab="",ylab="")
> # x axis
> mtext(text = expression(bold("Cut")),
+       side = 1,line =3)
> # y axis
> mtext(text = expression(bold("Price")),
+       side = 2, #side 2 = left
+       line = 3)
>
>
```




```
# As the weight of the diamond also affects the price, we use price per carat
# for our data analysis
> Dimonds_m1 <-
glm(diamonds$price/diamonds$carat~diamonds$cut,family="poisson")
There were 50 or more warnings (use warnings() to see the first 50)
> coef_Dimonds_m1<-coef(Dimonds_m1)
>
> Fair_cut_effect <- exp(coef_Dimonds_m1[1])
> print(Fair_cut_effect)
(Intercept)
   3767.256
> # This value is the base price for diamond per carat
>
> Good_cut_effect <- exp(coef_Dimonds_m1[2])
```

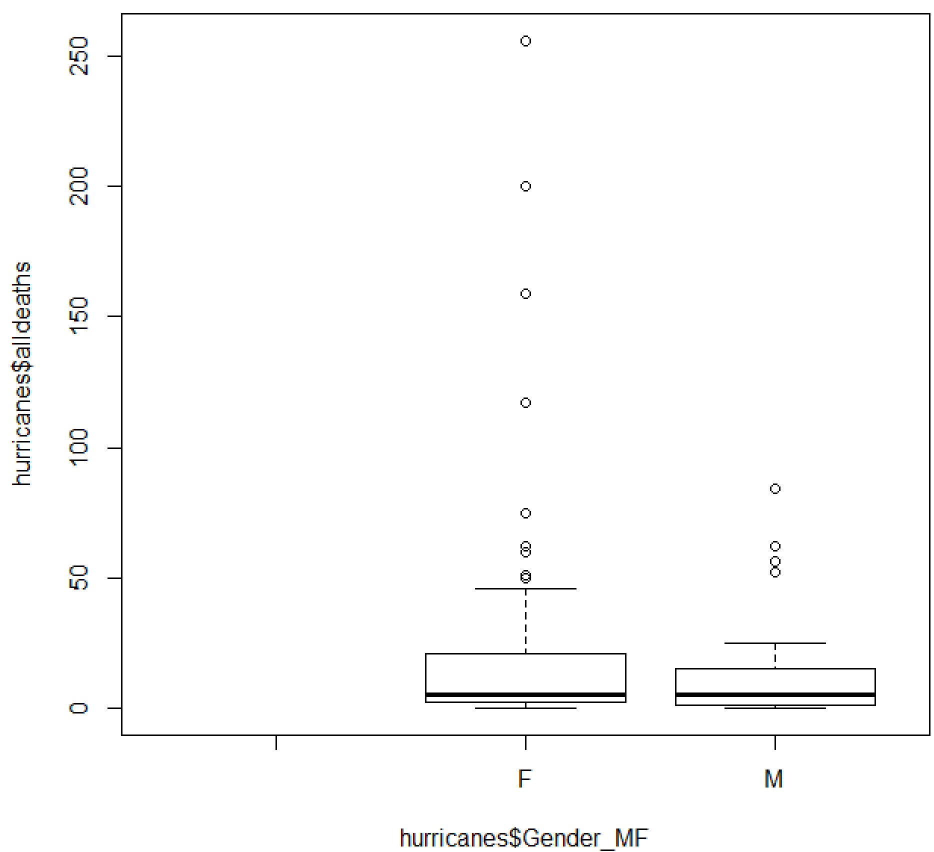
```
> print(Good_cut_effect)
diamonds$cutGood
        1.024626
> # It shows 2.4 percent increase in price as cut is good
>
> VeryGood_cut_effect <- exp(coef_Dimonds_m1[5])
> print(VeryGood_cut_effect)
diamonds$cutVery Good
             1.065531
> # It shows 6 percent increase in price as cut is very good
>
> Premium_cut_effect <- exp(coef_Dimonds_m1[4])
> print(Premium_cut_effect)
diamonds$cutPremium
            1.12095
> # It shows 12 percent increase in price as cut is Premium
>
> Ideal_cut_effect <- exp(coef_Dimonds_m1[3])
> print(Ideal_cut_effect)
diamonds$cutIdeal
         1.040466
> # It shows 4 percent increase in price as cut is Ideal
>
> confint(Dimonds_m1)
Waiting for profiling to be done...
                          2.5 %     97.5 %
(Intercept)           8.23330614 8.23489781
diamonds$cutGood      0.02341318 0.02524207
diamonds$cutIdeal     0.03884446 0.04049228
diamonds$cutPremium   0.11334041 0.11501291
diamonds$cutVery Good 0.06262939 0.06431765

 # The confidence seems to be reasonable as it does not cross the zero and
its range is not large
```

```
################################### Q2 ##################################
> cuse <- read.csv("contraception.csv")
> head(cuse)
    age education notUsing using Total
1   <25       low       53     6    59
2   <25       low       10     4    14
3   <25      high      212    52   264
4   <25      high       50    10    60
5 25-29       low       60    14    74
6 25-29       low       19    10    29
>
> dev.new()
NULL
> plot(cuse$using/cuse$Total~cuse$education,
+      col=c("gold","hotpink"),cex.axis=.7,
+      xlab="",ylab="")
> # x axis
> mtext(text = expression(bold("Education")),
+       side = 1,line =3)
> # y axis
> mtext(text = expression(bold("Using")),
+       side = 2, #side 2 = left
+       line = 3)
```



```
> response <- cbind(cuse$using,cuse$notUsing)
> predictor <- cuse$education
> m1 <- glm(response~predictor,family="binomial")
> coef(m1)
 (Intercept) predictorlow 
 -0.81020374   0.09248529 
> coef(m1)[2]/4
predictorlow 
  0.02312132 
> confint(m1)
Waiting for profiling to be done...
                  2.5 %     97.5 %
(Intercept)  -0.9460962 -0.6766394
predictorlow -0.1239481  0.3078275
> # Higher education has Maximum 2.3% effect on using contraception. However,
> # Intercept in confidence crosses zero which means education doesn't have
> # significant effect on contraception and we can reject the Null
Hypothesis.
```

```
#################################### Q3 ####################################
> hurricanes <- read.csv("Hurricane Dataset.csv")
> head(hurricanes)
Year    Name  MasFem MinPressure_before Minpressure_Updated.2014 Gender_MF
Category alldeaths  NDAM
1 1950     Easy 6.77778        958         960         F       3        2  1590
2 1950     King 1.38889        955         955         M       3        4  5350
3 1952     Able 3.83333        985         985         M       1        3   150
4 1953  Barbara 9.83333        987         987         F       1        1    58
5 1953 Florence 8.33333        985         985         F       1        0    15
6 1954    Carol 8.11111        9           960         F       3       60 19321
  Elapsed.Yrs Source  ZMasFem ZMinPressure_A    ZNDAM
1          63    MWR -0.00094       -0.35636 -0.43913
2          63    MWR -1.67076       -0.51125 -0.14843
3          61    MWR -0.91331        1.03765 -0.55047
4          60    MWR  0.94587        1.14091 -0.55758
5          60    MWR  0.48108        1.03765 -0.56090
6          59    MWR  0.41222       -0.25310  0.93174
>
> dev.new()
NULL
> plot(hurricanes$alldeaths~hurricanes$Gender_MF)
>
```



```
> response <- hurricanes$alldeaths
> predictor <- hurricanes$Gender_MF
> glm_gender <- glm(response~predictor,family="poisson")
> gender_m1 <- coef(glm_gender)
> slope <- exp(gender_m1[2])
> print(slope)
predictorM
 0.5990948
> intercept <- exp(gender_m1[1])
> print(intercept)
(Intercept)
   23.75806
> confint(glm_gender)
Waiting for profiling to be done...
                 2.5 %     97.5 %
(Intercept)  3.1164152  3.2185581
predictorM  -0.6211542 -0.4056501
> effect<- exp(gender_m1[1])-exp(gender_m1[1]+gender_m1[2]*1)
> print(effect)
(Intercept)
   9.524731
> # Here our analysis with poisson glm says that 9 people more die in
> # hurricanes with female name. However, they used negative binomial for
> #their research which seems to be reasonable.
```

```
######################################## Q4 #############################
# Goodman RC, Phillips OL, Baker TR (2013) Data from:
# The importance of crown dimensions to improve tropical tree biomass
# estimates.
# Dryad Digital Repository. doi:10.5061/dryad.p281g

# We want to investigate the effect of Monopodial architectural type on total
# biomass estimation

# Abbreviation Description:

# ρ wood density

# ρ source this study or Global Wood Density Database (GWDD; citations
# below)
# Chave J, Coomes DA, Jansen S, Lewis SL, Swenson NG, Zanne AE (2009) Towards
# a worldwide wood economics spectrum. Ecology Letters 12(4): 351-366.
# doi:10.1111/j.1461-0248.2009.01285.x

# Zanne AE, Lopez-Gonzalez G, Coomes DA, Ilic J, Jansen S, Lewis SL, Miller
# RB,
# Swenson NG, Wiemann MC, Chave J (2009) Data from: Towards a worldwide wood
# economics spectrum. Dryad Digital Repository. doi:10.5061/dryad.234

# DRH   Diameter at reference height
# POM   Point of measurement for diameter
# HFMB  Height of first major branch
# HTotal        Total height
# AvgCR Average crown radius
# CEA   Crown ellipse area
# Mono  Monopodial architectural type (1=yes, 0=no)
# Crown mass    Dry mass of the crown (everything above first major branch)
# Stem mass     Dry mass of stump and stem
# Total AGB     Total aboveground dry mass

# Null hypothesis : trees with a monopodial architectural type are
# estimated to have 21-44 % less mass than trees with other growth patterns.


> install.packages("xlsx")
> library("xlsx")
>
> Tree_Biomass <- read.xlsx("Tree Biomass Data.xlsx",3, header=TRUE)
> head(Tree_Biomass)
Number   Species    Family Ï...g.cm3.  Ï..source DRH...cm. POM...m. HFMB...m.
1   1    Theobroma cacao  Malvaceae   0.4741 this study  10.6   1.3    5.60
2   2    Acacia loretensis  Fabaceae  0.6007 this study  15.1   1.3    3.60
3   3    Drypetes amazonica Putranjivaceae  0.7103 this study  16.0  1.3 4.85
4   4 Pourouma cecropiifolia  Urticaceae    0.3557 GWDD  18.9   1.3    10.30
5   5 Ocotea javitensis  Lauraceae     0.5117 this study  21.5   1.3    13.60
6   6 Pseudolmedia laevis  Moraceae    0.6185   GWDD     23.0   1.3    14.90
HTotal...m. AvgCR...m. CEA...m2. Mono Crown.mass...kg. Stem.mass...kg.
Total.AGB...kg.
1  10.6  2.6   21.4    M      14.8          27.4          42.2
2  18.0  2.7   20.7    U      101.7         43.6          145.3
3  16.5  2.7   21.9    M      93.3          54.7          147.9
4  16.2  3.3   34.2    U      63.2          75.0          138.2
```

```
5  18.0  2.8    23.4    M       32.5          132.6            165.2
6  25.1  3.7    41.9    M       149.3         271.5            420.8
>
> dev.new()
NULL
> plot(Tree_Biomass$Total.AGB...kg~Tree_Biomass$Mono,col=c("gold","hotpink"))
>
> glm_mass<- glm(Tree_Biomass$Total.AGB...kg~Tree_Biomass$Mono,family =
"poisson")
There were 48 warnings (use warnings() to see them)
> Intercept_mass <- coef(glm_mass)[1]
> Slope_mass <- coef(glm_mass)[2]
> exp(Intercept_mass)-exp(Intercept_mass+Slope_mass)
(Intercept)
  -8658.059
> # It says 8658 kg less mass for Monopodial architectural type trees
> # than other growth patterns.
> confint(glm_mass)
Waiting for profiling to be done...
                      2.5 %   97.5 %
(Intercept)        7.574329 7.600933
Tree_Biomass$MonoU 1.670330 1.697605
There were 50 or more warnings (use warnings() to see the first 50)
> # the confidence interval is reasonable.
```