# Exercise 3: Exploring graphs using R

Grading system: Total obtainable points are 5.

**Task 1** ( 2.5 points): Make your own data to explore graph types.

-Dot plot

-Pie chart

-Bar plot

-Box plot

-Histogram

-Scatter plot

**Task 2** ( 2.5 points): Import data and review previous tasks.

-Import Data into R, and print them.

-Based on the imported data, write code to show descriptive statistics of the imported data

-Based on the imported data, write code to find the probability of two events happening. -Based on the imported data, write code to find out the probability of either of two events happening.

-Based on the imported data, write code to make 4 relevant and different types of graphs.

**TASK 1**

The type of graph you draw will depend on whether the variable you are plotting is continuous or discrete.

Consider you have a sample of OU students,

If you have a discrete variable on x axis( eg male/female, subject major ), you would make a bar plot or a box plot or a pie chart.

If you have a continuous variable on x axis( eg height, weight), you would make a line or dot plot.

If you have two continuous variables, you would make a scatter plot.

If you want a frequency distribution or a probability distribution of a discrete or a continuous variable, you would plot a histogram.

# Create the data for the chart.

u <- c(5,3,7,3,55,6,7,8,9)

v <- c(7,12,28,3,41,41,7,12,22)

# Make the plots

plot(v)

barplot(v)

pie(v)

hist(v)

boxplot(v)

plot(v,u)

#Add labels to the charts and do further manipulation in further tasks

**Task 1.1:Dot plot**

# Create the data for the chart.  
v <- c(7,12,28,3,41)  
  
# Plot the bar chart.   
plot(v)  
  
# Plot the bar chart.  
plot(v,type = "o", col = "red", xlab = "Month", ylab = "Rain fall",  
 main = "Rain fall chart")

**Task 1.2: Pie chart**

# Create data for the graph.  
x <- c(21, 62, 10, 53)  
labels <- c("Biology", "Chemistry", "Physics", "Math")  
  
# Plot the chart.  
pie(x,labels)  
  
# Create data for the graph.  
x <- c(21, 62, 10, 53)  
labels <- c("Biology", "Chemistry", "Physics", "Math")  
  
# Plot the chart with title and rainbow color pallet.  
pie(x, labels, main = "Class number pie chart", col = rainbow(length(x)))  
  
**Task 1.3: Barplot**

# Create the data for the chart.  
H <- c(7,12,28,3,41)  
M <- c("Biology", "Chemistry", "Physics", "Math",”Art”)  
  
# Plot the bar chart.  
barplot(H,names.arg = M,xlab = "Class",ylab = "Size",col = "blue",  
main = "Size of class chart",border = "red")  
  
  
**Task 1.4: Histogram**

# Create data for the graph.  
v <- c(9,13,21,8,36,22,12,41,31,33,19)  
  
# Create the histogram.  
hist(v,xlab = "Weight",col = "yellow",border = "blue")  
  
**Task 1.5: Boxplots**

# Plot the chart.  
boxplot(iris)

StudentName <- c("Tom","Dick","Harry","Mary","Jane","Jose","Eduardo","Chris","Anuj")

Age <- c(5,3,7,3,55,6,7,8,9)

Weight <- c(7,12,28,3,41,41,7,12,22)

UniversityOfOklahoma <- data.frame(Age,Weight,StudentName)

boxplot(UniversityOfOklahoma)

UniversityOfOklahoma <- data.frame(Age,Weight)

boxplot(UniversityOfOklahoma)

**Task 1.6 Scatter plot**

print(Students$Age)

print(Students$Weight)

plot(Students$Age,Students$Weight)

plot(Students$Weight,Students$Age)

**TASK 2**

Go to <http://www.data.gov> and explore the datasets there. Students are encouraged to go to the healthdata section and search for natality measures or use this link for uniformity and ease of instruction.

<https://www.healthdata.gov/dataset/nchs-natality-measures-females-race-and-hispanic-origin-united-states>

Some code that might help your brain juices flowing.

DO NOT USE THIS EXACT SAME CODE IN YOUR REPORT

# to print the data

print(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States)

# probability of children being born being Native American in 2015  
> 44299/3978497  
[1] 0.01113461

# probability of children being born being Native American or Pacific Islander in 2015  
> (44299/3978497+281264/3978497)  
[1] 0.08183065

# probability of children being born being Native American and Pacific Islander in 2015  
> (44299/3978497)\*(281264/3978497)  
[1] 0.0007871727

#plotting live births for all races in all years

plot(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`Live Births`[1:56])

# same plot as before, but with years a x axis

plot(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$Year[1:56],NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`Live Births`[1:56])

#scatter plot of birth rate vs fertility rate

plot(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`BirthRate`[1:56],NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`Fertility Rate`[1:56])

# histogram of live births for all years for all races

hist(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`Live Births`[1:56])

# pie chart for live births in 2011,2012,2013,2014,2015 for all races

pie(NCHS\_Natality\_Measures\_for\_Females\_by\_Race\_and\_Hispanic\_Origin\_United\_States$`Live Births`[1:5])