

## Statistics 251: Lab 1 Handout

First Name: Musa Last Name: Rasheed

Student Number: 25618232

Students should write the name they are registered with. Do not put nicknames/short forms on the handout. Please write down your answers neatly and do show your work.

### Part 1

Write down the R code that you used:

```
data = seq(1, 57, by=4)
matrix(data, nrow = 3, ncol = 5)
```

```
1 13 25 37 49
5 17 29 41 53
9 21 33 45 57
```

### Part 2

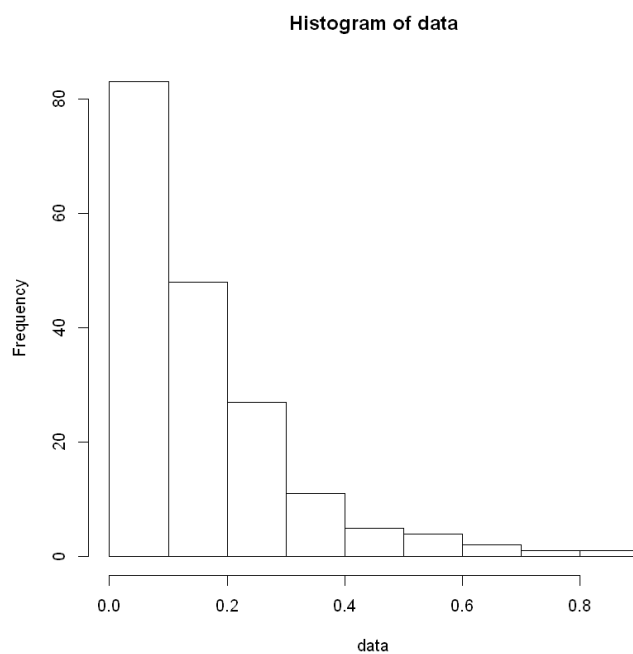
1. How many variables are there in the dataset? 5
  - a. Quake, Richter, Distance, Soil, Accel
2. Write down the R code that you used to calculate the mean and variance:

```
Mean: mean(Earthquake[3])
Variance:
n = length(Earthquake[3])
avg = mean(Earthquake[3])
xi = Earthquake[3]
print((1/(n-1))*sum((xi-avg)^2))
```

Write down the mean: [45.60](#). Write down the result from using  $var(x)$ : [3865.117](#).

3. Write down the R code that you used to generate the histogram.

```
data = Earthquake[5]
hist(data)
```



Describe the shape of the histogram.

It is a right skewed distribution whose mean is bigger than its median. The distribution is not very flat

Five-number summary: [min = 0.003](#), [Q1 = 0.04425](#), [median = 0.113](#), [Q3 = 0.21925](#), [max=0.81](#)

```
Summary(data)
```

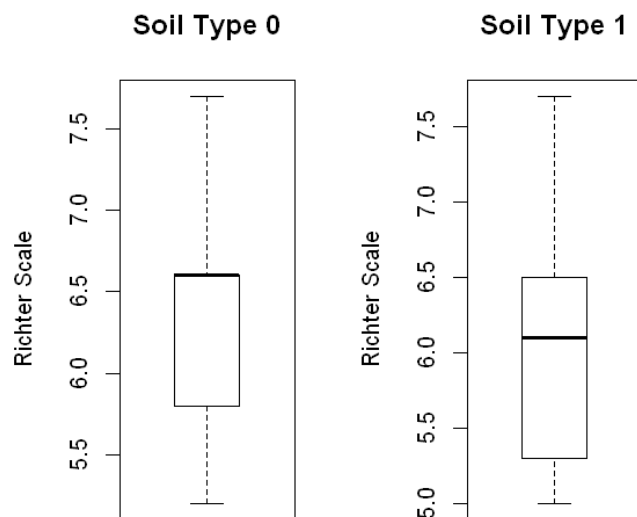
1	summary(data)					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	0.00300	0.04425	0.11300	0.15422	0.21925	0.81000

4. Write down the R code that you used to generate the side-by-side boxplots.

```
options(repr.plot.width=5, repr.plot.height=5)
```

```
soil0 = Earthquake[Earthquake[4]==0,2] #Soil type is col 4  
soil1 = Earthquake[Earthquake[4]==1,2] #Richter is col 2
```

```
par(mfrow = c(1,2))#Makes space for 2 plots (2 row 1 col)  
boxplot(soil0, main = "Soil Type 0", ylab = "Richter  
Scale")  
boxplot(soil1, main = "Soil Type 1", ylab = "Richter  
Scale")
```



Does the *Richter* intensity differ by *soil* type? Why?

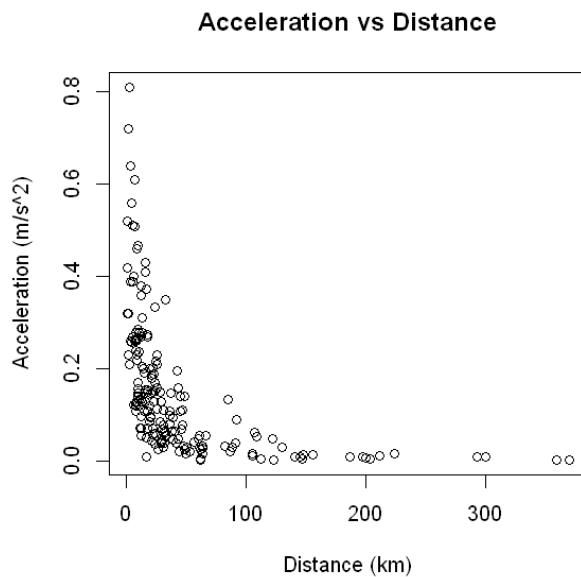
Yes (although this result could be affected by the fact that there was a LOT more data for soil type 1 than type 0.) It appears generally, soil type 0 has more intense earthquakes than soil type 1. Not only is the median shifted upwards for type 0, but so is quartile 1, meaning even lower intensity earthquakes are more powerful. This leads to the conclusion that type 0 soil has more intense earthquakes than type 1.

5. Which type of plot did you use? Scatter Plot

Write down the R code that you used to generate the plot.

```
d = Earthquake[3]
a = Earthquake[5]

plot(d,a,main = "Acceleration vs Distance", xlab = "Distance (km)", ylab =
"Acceleration (m/s^2)")
```



Describe the relationship.

It seems to be some sort of decay plot. As the distance increases, the acceleration decreases.

$$a = e^{-kd}$$

Where  $k$  is some constant.