

1. Read the below stem and leaf plot, enter in the data and make a histogram

The decimal point is 1 digit(s) to the right of the |

```
8 | 028
9 | 115578
10 | 1669
11 | 01
```

2. To understand a skewed data set we use the `log` transform. That is take logarithm of the data set. Take `exec.pay` dataset in `UsingR` package. Find, the mean, median and quartiles of `exec.pay` and decide if it is skewed. Plot the histogram and boxplot of the data set. Then consider the log-transform

```
> y= log(1+exec.pay,10)
```

Find, the mean, median and quartiles of `y`. Plot the histogram and boxplot of `y`. How do they compare with `exec.pay` ?

3. Fit a density estimate to the dataset `pi2000` in package `UsingR`
4. Suppose `x` is a vector. Describe what each of the below commands do.

```
> length(x)
> x[2]
> x[-2]
> x[1:5]
> x(length(x) -5 : length(x))
> x[c(1,3,5)]
> x[x>3]
> x[x<-2 | x>2]
> which(x == max(x))
```

5. One can generate random data in R. Using the in-built help menu, describe what the below command is producing.

```
> x = rnorm(100)
```

Run the above two times and create two different histograms for two different `x`. In each case perform a density plot on them.

6. Plot the histogram and density plots for the `faithful` dataset.

```
> data(faithful)
> attach(faithful)
> # make eruptions visible
> hist(eruptions,15,prob=T)
> lines(density(eruptions))
> lines(density(eruptions,bw="SJ"),col="red")
```

- (a) Describe what each of the commands are doing.
- (b) Describe what happens if in each of the last lines we change `bw="SJ"` to `bw="0.01"`, or `bw="0.1"`, or `bw="0.1"` and execute the above.

7. You can create your own functions in R. These are created using the `function` command. For example, we can design our own function to calculate mean.

```
> MYMEAN = function(x) { sum(x)/length(x) }
```

Then you can say

```
> x = c(1,2,3)
> MYMEAN = function(x) { sum(x)/length(x) }
> MYMEAN(x)
```

```
[1] 2
```

A function in R is another object, with the class `function`. It typically will return the last value computed in the body. Can you write a function to calculate the `mode` of a dataset `x` ?