Stem and leaf plots

5 marks

Stem and leaf plots. Install the package `aplpack`; the function `stem.leaf(...)` from that package will be used to construct the displays. You will work with the following data:

```
set.seed(1234567) # Use this so we all get the same values
x30 <- rnorm(30)
x100 <- rnorm(100)
x500 <- rnorm(500)
x1000 <- rnorm(1000)
```

Stem and leaf plots are then generated as follows:

```
library(aplpack)
stem.leaf(x30)
stem.leaf(x100)
stem.leaf(x500)
stem.leaf(x1000)
```

a. **(3 marks)** What characteristics of the data can you see (or easily determine) from the stem and leaf display? DO NOT hand these displays in!

   With the number of points increasing, although there are some variations, we can still see a clear pattern that all graphs form a bell curve, which is generated from normal distribution. From observation, the center of the bell curve is approximately 0. However, the number of subdivision gets larger and larger.

b. **(2 marks)** How does the stem and leaf adapt to increasing sample size? Suppose a sample of $n = 10,000$ were to be displayed on a single sheet of paper (using the default values for `stem.leaf()`). What would be the problems, if any, with that?

   When we increase the sample size, the stem and leaf will adapt itself by increasing the number of categories (or the width of the graph) and number of data in every category (or the height of the graph). Suppose we have a sample of n=10000, we will encounter a problem that the width and the height of the graph get really large. We cannot see a clear pattern from the graph in a paper