Worksheet5

Sunny Lee

3/1/2021

1)

a) Splitting gpa between 2.7 and 3.1:

```
R1 = c(90)
R2 = c(120, 75)
avgR1 = mean(R1)
avgR2 = mean(R2)
RSS1 = sum((R1 - avgR1)^2)
RSS2 =sum((R2 - avgR2)^2)

RSStotal = RSS1 + RSS2
RSStotal
```

```
## [1] 1012.5
```

b) Splitting the gpa between 3.1 and 3.4:

```
R1 = c(90, 75)
R2 = c(120)
avgR1 = mean(R1)
avgR2 = mean(R2)
RSS1 = sum((R1 - avgR1)^2)
RSS2 =sum((R2 - avgR2)^2)

RSStotal = RSS1 + RSS2
RSStotal
```

```
## [1] 112.5
```

c) Splitting the experience between 5 and 11:

```
R1 = c(75)
R2 = c(90, 120)
avgR1 = mean(R1)
avgR2 = mean(R2)
RSS1 = sum((R1 - avgR1)^2)
RSS2 =sum((R2 - avgR2)^2)

RSStotal = RSS1 + RSS2
RSStotal
```

```
## [1] 450
```

Splitting the experience between 11 and 12:

```
R1 = c(90, 75)
R2 = c(120)
```

```r
avgR1 = mean(R1)
avgR2 = mean(R2)
RSS1 = sum((R1 - avgR1)^2)
RSS2 =sum((R2 - avgR2)^2)

RSStotal = RSS1 + RSS2
RSStotal
```

```
## [1] 112.5
```

After calculating all of the RSS values for each split, we find that splitting the gpa between 3.1 and 3.4 or splitting the experience between 11 and 12 will result in the lowest RSS value. Thus, we should pick one of these splits as our first binary split.