DATA ANALYTICS WITH R, EXCEL AND TABLAEU

ASSIGNMENT 8.1 ANSWERS

By ASHISH S SHANBHAG

ashishshanbhag108@gmail.com

Question no:

5)

1a) Calculate the average salary by gender and smoking status

Ans > x <- tapply(salary, list(gender = gender), mean)

```
> x
gender
Female
Male
16.46353 17.93035
```

1b) Which gender has the highest mean salary?

Ans

```
> by(salary, gender, mean, na.rm = TRUE)
gender: Female
[1] 16.46353

gender: Male
[1] 17.93035
```

We can answer this by looking above. For these data, the gender with the highest mean salary is Male.

1c) Report the highest mean salary.
Ans
mean(salary[gender == Male])
for example. For these data, the highest mean salary is
> x[which(x == max(x))]
Male
17.93035

1d) Compare the spreads for the genders by calculating the standard deviation of salary by gender.

Ans

```
> y <- tapply(salary, list(gender = gender), sd)

> y
gender
Female
Male
2.122113 1.077183
```
For these data, the the largest standard deviation is approximately 2.12 which was attained by the Female gender