

# Class 12 HW

Lilia Jimenez (PID:A16262599)

##Section 4: Population Scale Analysis

How many samples do we have?

```
expr<-read.table("rs8067378_ENSG00000172057.6.txt")
head(expr)
```

	sample	geno	exp
1	HG00367	A/G	28.96038
2	NA20768	A/G	20.24449
3	HG00361	A/A	31.32628
4	HG00135	A/A	34.11169
5	NA18870	G/G	18.25141
6	NA11993	A/A	32.89721

```
nrow(expr)
```

[1] 462

Q13:How many of each genotype?

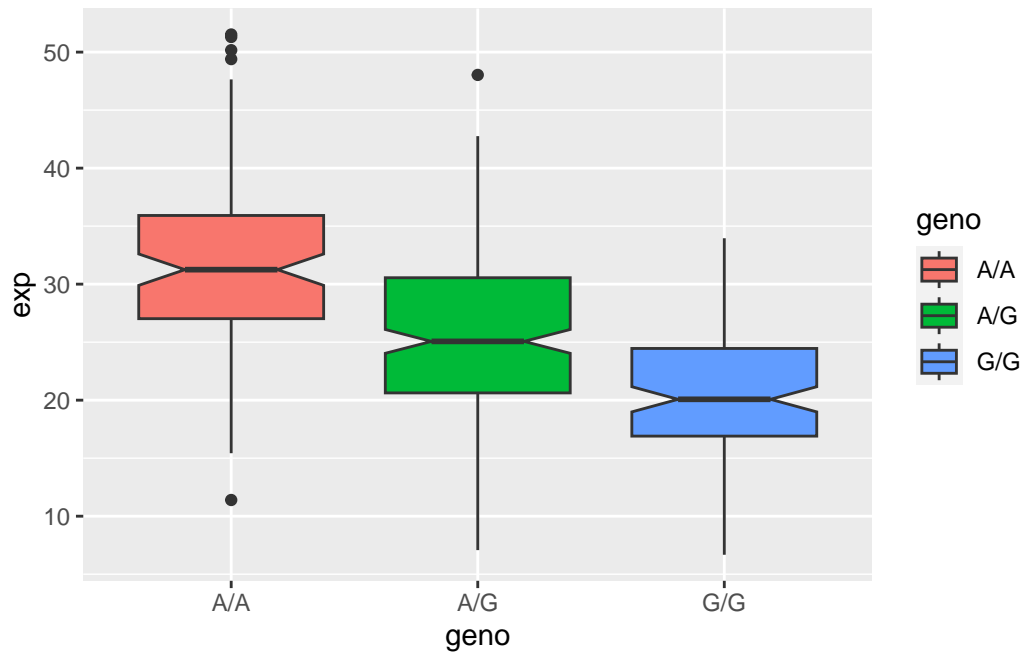
```
table(expr$geno)
```

A/A	A/G	G/G
108	233	121

```
library(ggplot2)
```

Q.14:Lets make a boxplot

```
ggplot(expr)+aes(x=geno,y=exp, fill=geno)+
  geom_boxplot(notch = TRUE)
```



q14. cont:what could you infer from the relative expression value between A/A and G/G displayed in this plot? Does the SNP effect the expression of ORMDL3?

A/A resulted in higher expression and G/G had lower expression. This tells us that this SNP does effect the expression of ORMDL3.