

# Class12: Population Analysis

Claire Lua A16922295

Q13. Read this file into R and determine the sample size for each genotype and their corresponding median expression levels for each of these genotypes.

Hint: The `read.table()`, `summary()` and `boxplot()` functions will likely be useful here. There is an example R script online to be used ONLY if you are struggling in vein. Note that you can find the medium value from saving the output of the `boxplot()` function to an R object and examining this object. There is also the `medium()` and `summary()` function that you can use to check your understanding.

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

	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

```
table(ppl$geno)
```

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

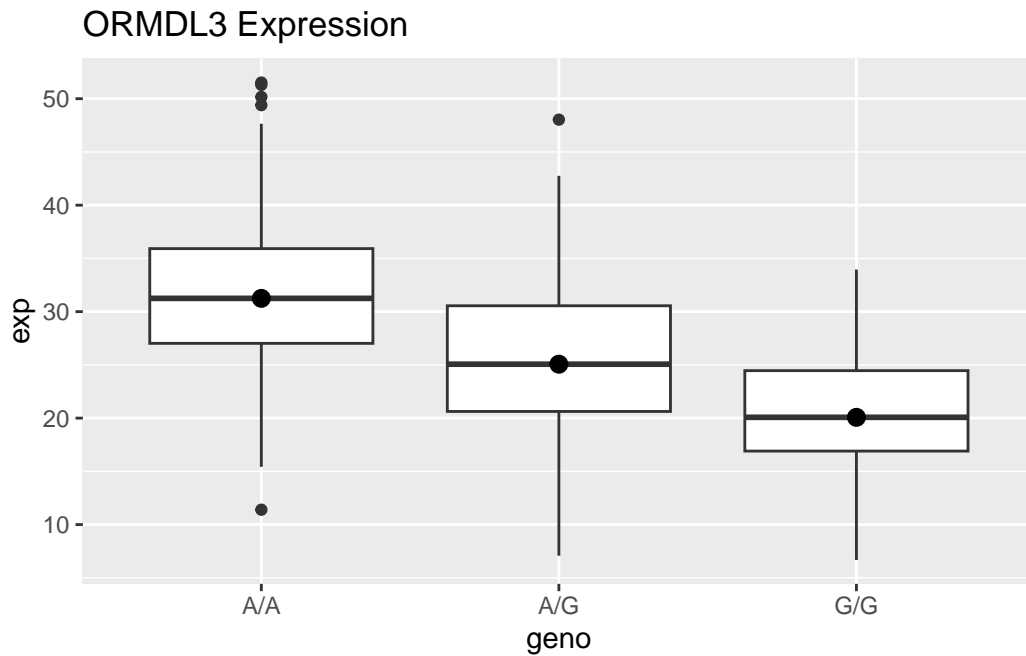
Q14. Generate a boxplot with a box per genotype, what could you infer from the relative expression value between A/A and G/G displayed in this plot? Does the SNP effect the expression of ORMDL3?

Hint: An example boxplot is provided overleaf – yours does not need to be as polished as this one.

```
library(ggplot2)
```

```
plot <- ggplot(ppl, aes(geno, exp)) +  
  geom_boxplot() +  
  stat_summary(fun = "median") +  
  labs(title = "ORMDL3 Expression", "Geno", "Exp")  
  
plot
```

Warning: Removed 3 rows containing missing values or values outside the scale range (``geom_segment()``).



```
summary(plot)
```

```
data: sample, geno, exp [462x3]  
mapping: x = ~geno, y = ~exp  
faceting: <ggproto object: Class FacetNull, Facet, gg>  
compute_layout: function
```

```

draw_back: function
draw_front: function
draw_labels: function
draw_panels: function
finish_data: function
init_scales: function
map_data: function
params: list
setup_data: function
setup_params: function
shrink: TRUE
train_scales: function
vars: function
super: <ggproto object: Class FacetNull, Facet, gg>
-----
geom_boxplot: outliers = TRUE, outlier.colour = NULL, outlier.fill = NULL, outlier.shape = 1
stat_boxplot: na.rm = FALSE, orientation = NA
position_dodge2

geom_pointrange: na.rm = FALSE, orientation = NA
stat_summary: fun.data = NULL, fun = median, fun.max = NULL, fun.min = NULL, fun.args = list
position_identity

```