

Class12

Nicholas Thiphakhinleo

2024-11-10

Proportion of G/G in MXL Population

```
mxl <- read.csv("373531-SampleGenotypes-Homo_sapiens_Variation_Sample_rs8067378.csv")
head(mxl)
```

	Sample..Male.Female.Unknown.	Genotype..forward.strand.	Population.s.	Father
1		NA19648 (F)	A A ALL, AMR, MXL	-
2		NA19649 (M)	G G ALL, AMR, MXL	-
3		NA19651 (F)	A A ALL, AMR, MXL	-
4		NA19652 (M)	G G ALL, AMR, MXL	-
5		NA19654 (F)	G G ALL, AMR, MXL	-
6		NA19655 (M)	A G ALL, AMR, MXL	-
	Mother			
1	-			
2	-			
3	-			
4	-			
5	-			
6	-			

```
table(mxl$Genotype..forward.strand)/nrow(mxl) *100
```

A A	A G	G A	G G
34.3750	32.8125	18.7500	14.0625

Section 4

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
```

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

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

Boxplot

```
library(ggplot2)
ggplot(expr) + aes(x=geno, y=exp, fill=geno) +
  geom_boxplot(notch=T)
```

