

Class 17

Ava Limtiaco (PID: A18007672)

```
library(tximport)
```

```
folders <- dir(pattern="SRR21568*")  
print(folders)
```

```
[1] "SRR2156848_quant" "SRR2156851_quant"
```

```
list.files(pattern="_quant")
```

```
[1] "SRR2156848_quant" "SRR2156851_quant"
```

```
list.files("SRR2156848_quant", full.names = TRUE)
```

```
[1] "SRR2156848_quant/abundance.h5" "SRR2156848_quant/abundance.tsv"  
[3] "SRR2156848_quant/run_info.json"
```

```
library (tximport)  
folders <- dir(pattern="SRR21568*")  
samples <- sub("_quant", "", folders)  
files <- file.path(folders, "abundance.h5")  
names(files) <- samples
```

```
txi.kallisto <- tximport(files, type="kallisto",txOut=TRUE)
```

1 2

```
head(txi.kallisto$counts)
```

	SRR2156848	SRR2156851
ENST00000539570	0	0
ENST00000576455	0	0
ENST00000510508	0	0
ENST00000474471	0	0
ENST00000381700	0	0
ENST00000445946	0	0

```
colSums(txi.kallisto$counts)
```

SRR2156848	SRR2156851
2563611	2111474

```
sum(rowSums(txi.kallisto$counts)>0)
```

```
[1] 76017
```

```
to.keep <- rowSums(txi.kallisto$counts) > 0  
kset.nonzero <- txi.kallisto$counts[to.keep,]
```

```
library(tximport)  
samples <- c("SRR600956", "SRR2156848_quant", "SRR2156851_quant")  
files <- file.path(samples, "abundance.tsv")  
names(files) <- samples  
x <- txi.kallisto$abundance
```

```
x <- txi.kallisto$abundance  
x <- x[rowSums(x) > 0, ]  
x <- x[apply(x, 1, function(row) sd(row) > 0), ]  
  
dim(x)
```

```
[1] 76017      2
```

```
pca <- prcomp(t(x), scale. = TRUE)
```

```
pca <- prcomp(t(x), scale=TRUE)
```

```
summary(pca)
```

Importance of components:

	PC1	PC2
Standard deviation	275.7	1.273e-10
Proportion of Variance	1.0	0.000e+00
Cumulative Proportion	1.0	1.000e+00

```
plot(pca$x[,1], pca$x[,2],  
     col=c("blue", "blue", "red", "red"),  
     xlab="PC1", ylab="PC2", pch=16)
```

