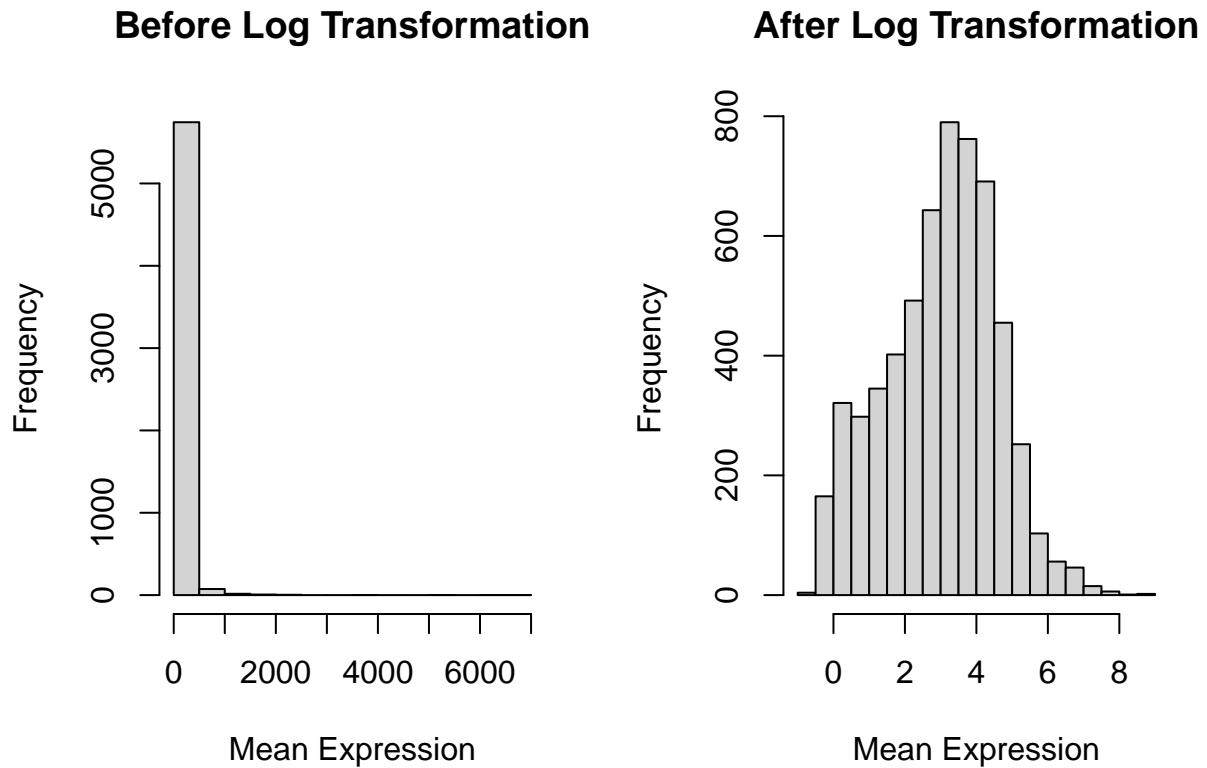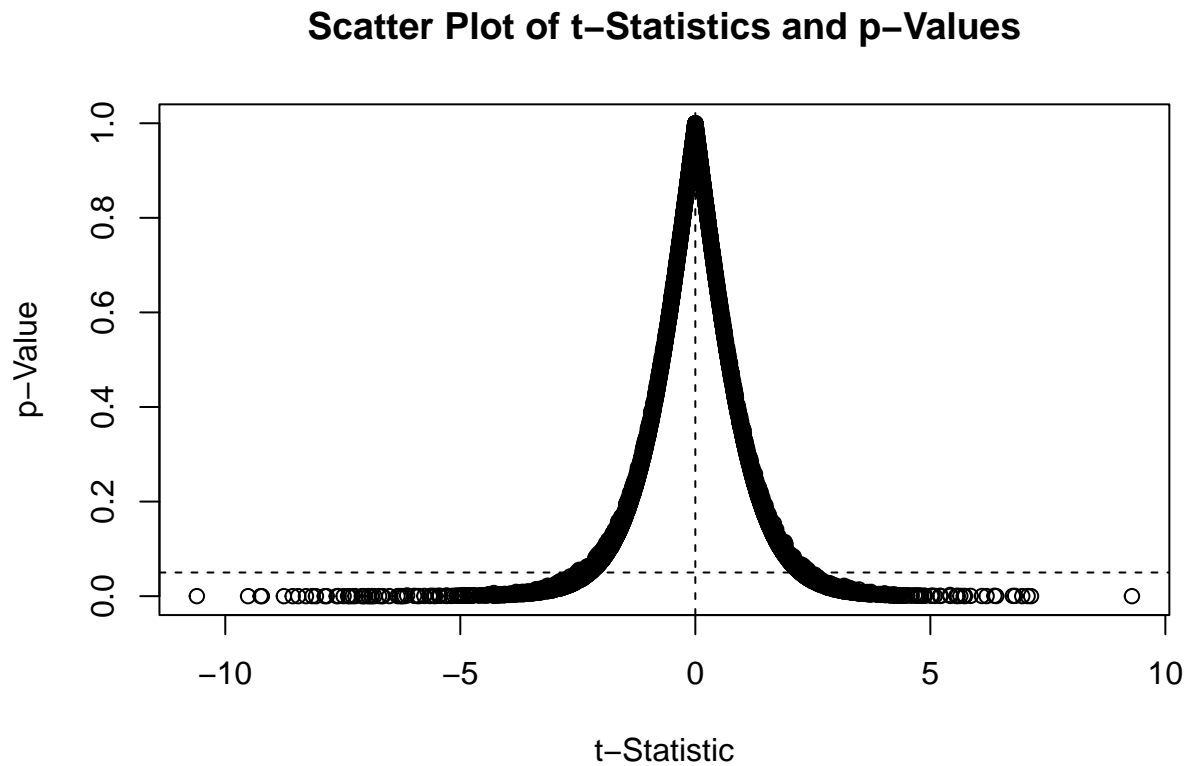```
Warning in scan(file = file, what = what, sep = sep, quote = quote, dec = dec,
: EOF within quoted string

Warning in scan(file = file, what = what, sep = sep, quote = quote, dec = dec,
: number of items read is not a multiple of the number of columns
```

# Question 1

**Before Log Transformation**     **After Log Transformation**

# Question 2

**Scatter Plot of t–Statistics and p–Values**



# Question 3

```
t_stat   p_val
5031 -9.221930 3.322445e-06
2332 9.289188 3.398338e-06
478 -9.515228 4.127453e-06
2961 -8.566569 6.875124e-06
747 -8.448239 7.581213e-06
589 -8.283231 9.659293e-06
5001 -8.753478 1.002803e-05
632 -8.071365 1.111684e-05
1838 -7.884366 1.393293e-05
636 -7.630266 1.803736e-05
3696 -7.840913 1.948641e-05
2657 -7.482368 2.112308e-05
740 -7.353670 2.458047e-05
1210 7.142123 3.166667e-05
4548 -7.094669 3.344832e-05
4255 -7.387284 3.406065e-05
2536 -9.248079 4.044308e-05
2538 -6.957720 4.211609e-05
5136 -6.894082 4.247315e-05
```
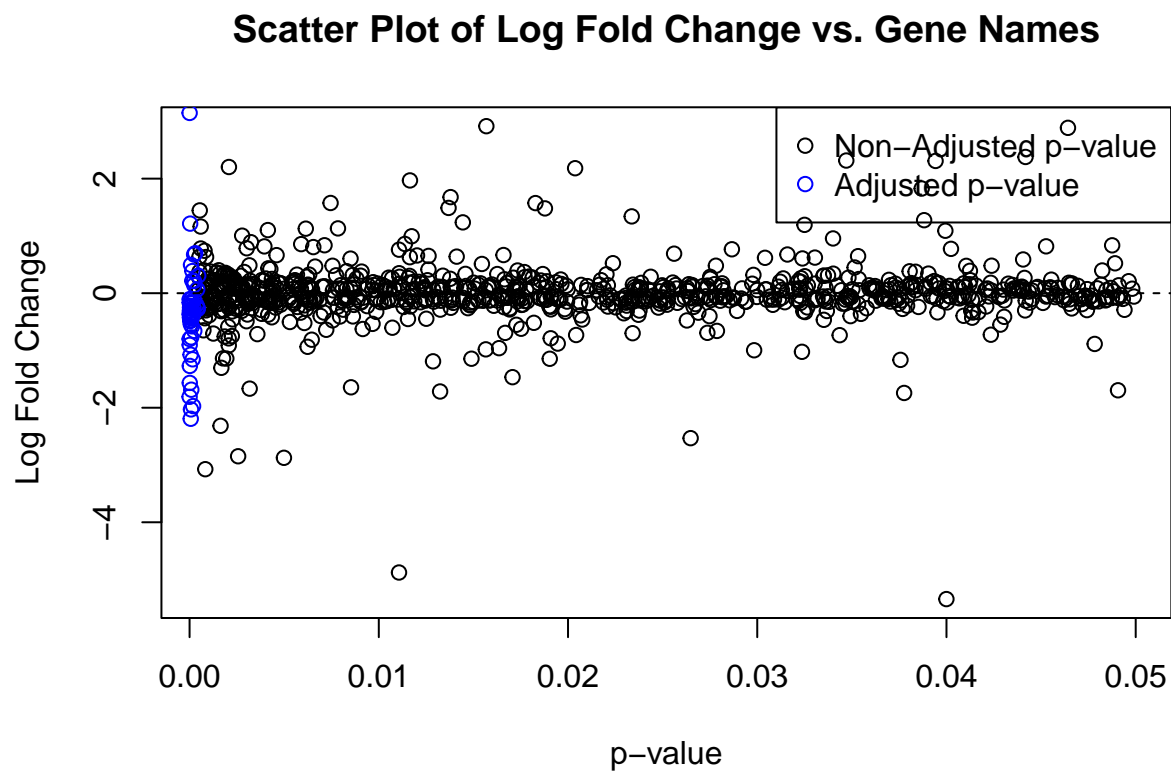
```
4380 -7.233021 5.282587e-05
```

# Question 4

```
## Warning: NaNs produced
```

```
## Warning: NaNs produced
```

```
## Warning: NaNs produced
```

```
## Warning: NaNs produced
```

Scatter Plot of Log Fold Change vs. Gene Names

The graph above shows the log2 fold change in mean expression for each gene when comparing pregnant and non pregnant males. We performed a Benjamini & Hochberg p value correction on the data which took us from 925 samples with a p<0.05 to 56, these 56 are highlighted in blue. Thus, approximately 94% of our significantly differentially expressed genes are false positives. Comparing the adjusted and non adjusted p-values in terms of log fold change between pregnant and non pregnant males, you can see that stronger correlation (p-value) is skewed more towards a decrease in expression (negative log fold change).