

Final Project

2025-07-14

```
genes<- read.csv("~/Desktop/QBS 103 material/Final project/QBS103_GSE157103_genes.csv")
matrix<- read.csv("~/Desktop/QBS 103 material/Final project/QBS103_GSE157103_series_matrix-1.csv")
```

```
genes_trans <- as.data.frame(t(genes))
colnames(genes_trans) <- as.character(genes_trans[1,])
genes_trans <- genes_trans[-1,]
```

```
genes_trans$participant_id <- rownames(genes_trans)
```

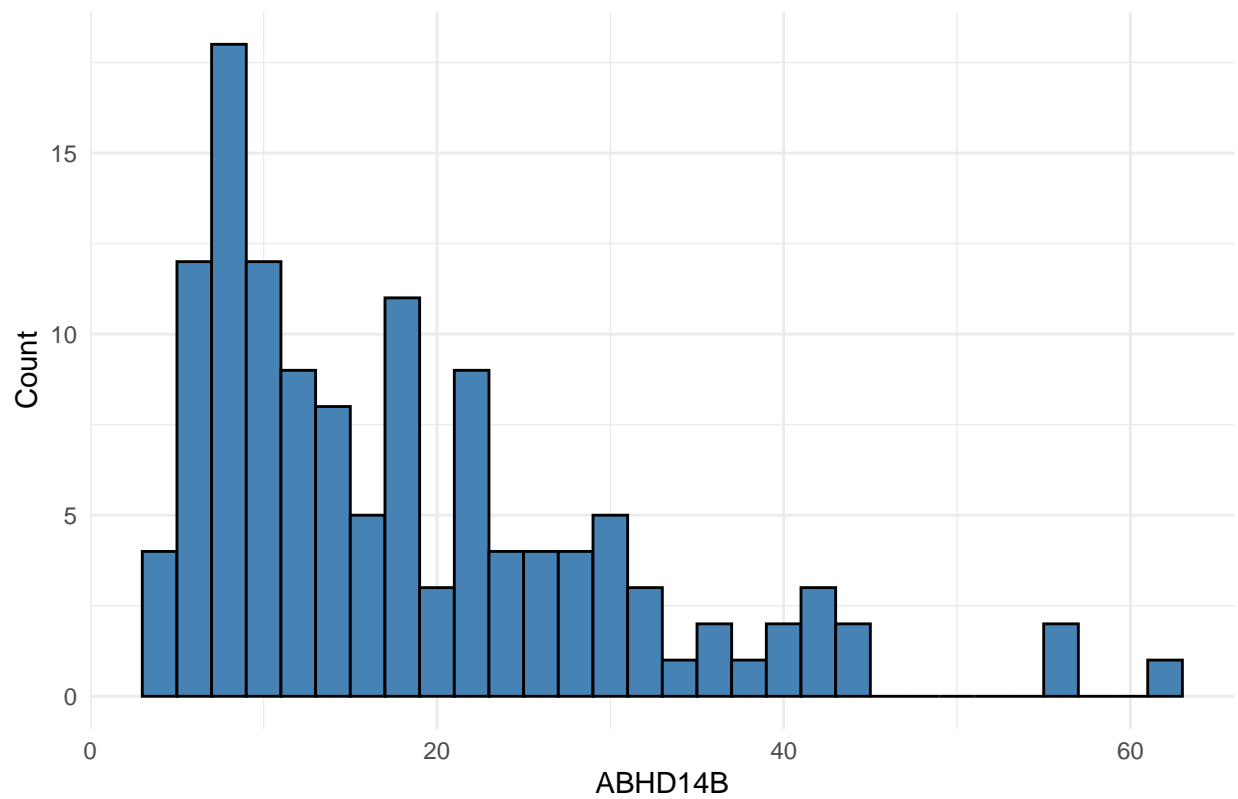
```
rownames(matrix)<- matrix$participant_id
merged_df <- merge(genes_trans, matrix , by = "row.names")
```

```
colnames(merged_df)[1] <- "participant_id"
```

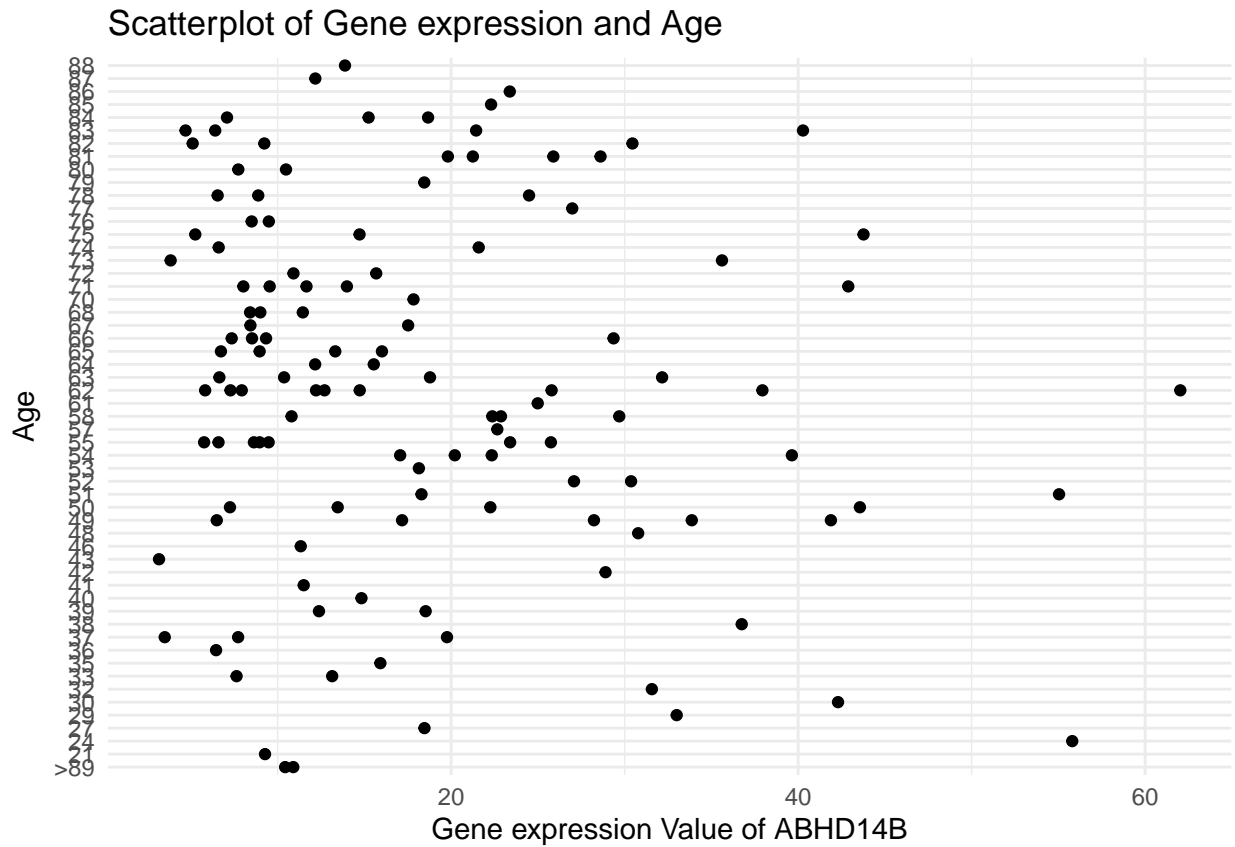
```
selected_cov <- c("age", "sex", "disease_status", "ABHD14B")
subset_merged <- merged_df[,selected_cov]
```

```
library(ggplot2)
subset_merged$ABHD14B <- as.numeric(subset_merged$ABHD14B)
ggplot(subset_merged, aes(x = ABHD14B)) +
  geom_histogram(binwidth = 2.0, fill = "steelblue", color = "black") +
  labs(title = "Gene Expression of ABHD14B", x = "ABHD14B", y = "Count") +
  theme_minimal()
```

Gene Expression of ABHD14B



```
ggplot(subset_merged, aes(x = ABHD14B, y = age)) +  
  geom_point() +  
  labs(title = "Scatterplot of Gene expression and Age", x = "Gene expression Value of ABHD14B", y = "Age") +  
  theme_minimal()
```



```
ggplot(subset_merged, aes(x = sex, y = ABHD14B, fill = disease_status)) +
  geom_boxplot() +
  labs(title = "Gene Expression by Sex and Disease Status",
       x = "Sex", y = "Gene Expression") +
  theme_minimal()
```

