

# Genes and Positive Association

Janav Sama, Student  
Northeastern University

**Research Question:** What are the top 5 genes that have a large number of positive associations?

```
select gene, # the gene symbol  
       gene_name, # gene full name  
       count(*) as num_records # unique associations  
from gad # name of the database  
where association = 'Y' # setting condition  
group by gene, gene_name # to find out the most affected genes  
having num_records >= 100 # setting the minimum for sorting  
order by num_records desc # setting the order to most occurring gene  
limit 5; # setting a limit to get the most occurring gene
```

Figure 1: SQL Query Executed

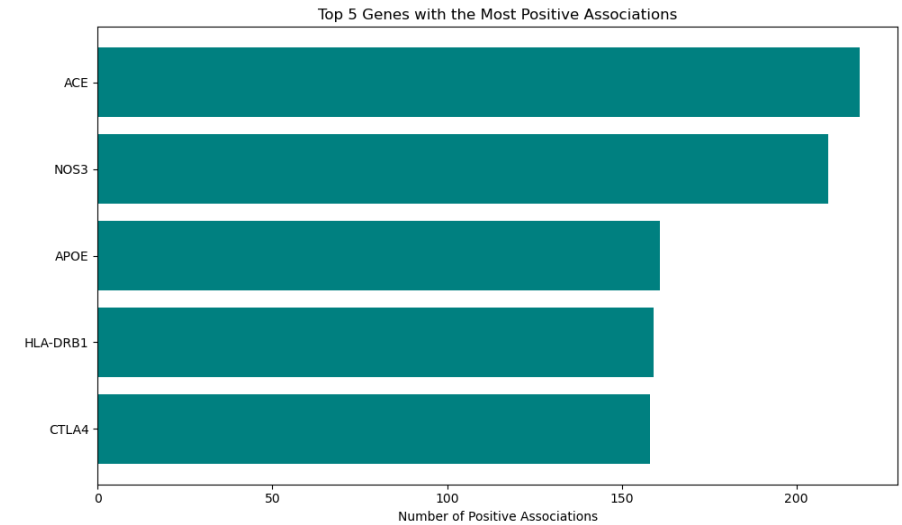


Figure 2: Bar chart representing output

The bar chart highlights the top five genes with the highest number of positive disease associations. ACE and NOS3, which are primarily linked to cardiovascular diseases, show the most associations. APOE, ranks third, showing its pivotal role in neurodegenerative conditions. HLA-DRB1 and CTLA4, essential in immune regulation, demonstrate significant associations with autoimmune diseases. Understanding the frequency of their positive associations can help prioritize future research into potential treatments, particularly for conditions where these genes play a critical role.

**Citation:** Becker et al. (2004). The Genetic Association Database. *Nature Genetics*, 36(5):431-2.

