Michael Lee, Jonathan Wolf

CSE 446

TCGA-Project

Goal One:

Get Our Data!

What we want is of the form:

-Random sample of patients from multiple cohorts

Expression Gene X | Person Y | Age | Nearest mutation to gene (promoter) X | age | smoking| status

-We need to query a random sample of people and genes, we will have multiple patients potentially the same amount.