**Final Project Proposal**

**UNCC Data Analytics & Visualization Bootcamp**

|  |  |
| --- | --- |
|  |  |
| **Project Title:** | Final Project: Cancer Statistics |
|  |  |
| **Team members:** | Zackery Rogers  Luis Rojas  Hector Ladero Rivas  Akanksha Kukhraniya  Mariam Ahmad |
|  |  |
| **Project Description:** | Taking Cancer data and showing which states have the most cancer cases and the types of cancer present, along with age, testing, and race. |
|  |  |
| **Dataset(s) to be used:** | Source: CDC SEER,  Kaggle    Dataset Name: Breast-Cancer Wisconsin (Diagnostic Data Set),  State Data, State Cluster, Cancer Type\_StateCount  File Type: .csv,  .csv |
|  |  |
| **Rough Breakdown of Tasks:** | Topic idea generation  Documentation of project proposal  Manage GitHub  Perform ETL process to clean and load data into Database  Create ML model by using Scikit Learn  Import ML libraries  Tableau to show visuals  Create Flask and Heroku  Create HTML using JavaScript |