SUMMARY

CLEANING THE DATA & PREPARING THE DATASET

WE USED PYTHON FOR THE ANALYSIS. PACKAGES LIKE SEABORN, MATPLOTLIB, NUMPY , PANDAS WERE UTILIZED

EXPLORING THE DATA

BY ADDING ADDITIONAL DATASETS WITH AREA AND DEMOGRAPHY BASED INFO

DEFINING THE RELATIONSHIP BETWEEN VARIABLES

BASED ON THE EDA RESULTS WE FOUND SOME INTERESTING VARIABLES & EXPLORED MORE THROUGH VISUALIZATION

PREDICTION MODEL & SUGGESTION

WE USED DECISION TREE MODEL TO PREDICT THE APPROVAL PROBABILITY OF LOANS

SUMMARY

DATASETS USED

WE USED USED APPROVED LOANS, REMOVED LOANS, NAICS, HOUSEHOLD INCOME BASED ON CONGRESSIONAL DISTRICTS, POLITICAL AFFILIATION OF CONGRESSIONAL DISTRICTS IN GEORGIA DISTRICTS, RACES IN GEORGIA CONGRESSIONAL DISTRICTS.

ADDITIONAL WEBSITES

- https://www.govtrack.us/congress/members/GA/6
- https://www.federalpay.org/paycheck-protection-program/michael-s-williams-atlanta-ga
- https://datausa.io/profile/geo/warren-county-ga