DATA*6300 : FINAL PROJECT

DISASTER RELIEF FUND ALLOCATION ANALYSIS: IDENTIFYING EXCESSIVE GRANTS TO HOMEOWNERS

By:

Poojan Vadaliya (1281587)

CONTENT

Overview

Road Home Program was introduced to assist the home owners affected by hurricane Katrina and Rita.

01

Aiding Louisiana homeowners

- Extensive property damage beyond insurance and FEMA assistance.

02

TYPES OF GRANTS OFFERED

- Compensation Grants
- Additional Compensation Grants
- Elevation grants
- Mitigation grants

03

DATASET OVERVIEW

Louisiana Division of Administration

- Compensation Grants
- Records of property owners who received grants.
- Pre-storm value, Repair cost, Insurance amounts

04

DISCREPANCIES IN GRANT AMOUNTS

- Factors causing inconsistencies
  - Additional insurance payments
  - Fraud cases
- Impact on actual funds owed by property owners to the state.

PROBLEM STATEMENT

01

FINDING HOMEOWNERS OR CONTRACTORS WHO TOOK EXCESSIVE GRANTS.

BINARY CLASSIFICATION PROBLEM

02

FINDING THE NUMBERS THAT COULD HAVE BEEN SAVED.

EXPLORATORY DATA ANALYSIS

Structure Demographics

- Structure Type
- GIS State
- GIS City
- GIS Zip
- Parish
- NOLA Planning Demographics

Grant Statistics

- Total CG Amount
- Total ACG Amount
- Total Elevation Amount
- Total IMM Amount
- Total Closing Amount
- ARS File (Yes/No)

DATA DESCRIPTION

Closing Options

- Closing Damage Assessment
- PSV at Closing
- Closed File - Option 1
- Closed File - Option 2/3
- Closed with Approved Unmet Needs

Damage Assessment

- Current Damage Assessment
- Current Damage Assessment-Type1
- Current Damage Assessment-Type 2
- Damage Type 1 or 2
- Current PSV

Census Demographics

- Census Blocks
- Block Groups
- Census Tracks

Samples : 130053

Columns : 36

Preprocessing Steps

Handling Missing Values

Elimination Of NOLA Files

01

02

Removing Unnecessary Feature Columns

Encoding the Categorical Data

One Hot Encode
Binary Encode
Frequancy Encode

03

04

Handling Numerical Columns

Normalising Numerical Columns
Identifying Outliers

Handling Class Imbalance

Y : 8925
N : 121128

→

Y : 17850
N : 17850

05

Structure Demographics

- Structure Type
- GIS State
- GIS City
- GIS Zip
- Parish
- NOLA Planning Demographics

Grant Statistics

- Total CG Amount
- Total ACG Amount
- Total Elevation Amount
- Total IMM Amount
- Total Closing Amount
- ARS File (Yes/No)

Census Demographics

- Census Blocks
- Block Groups
- Census Tracks

Closing Options

- Closing Damage Assessment
- PSV at Closing
- Closed File - Option 1
- Closed File - Option 2/3
- Closed with Approved Unmet Needs

Damage Assessment

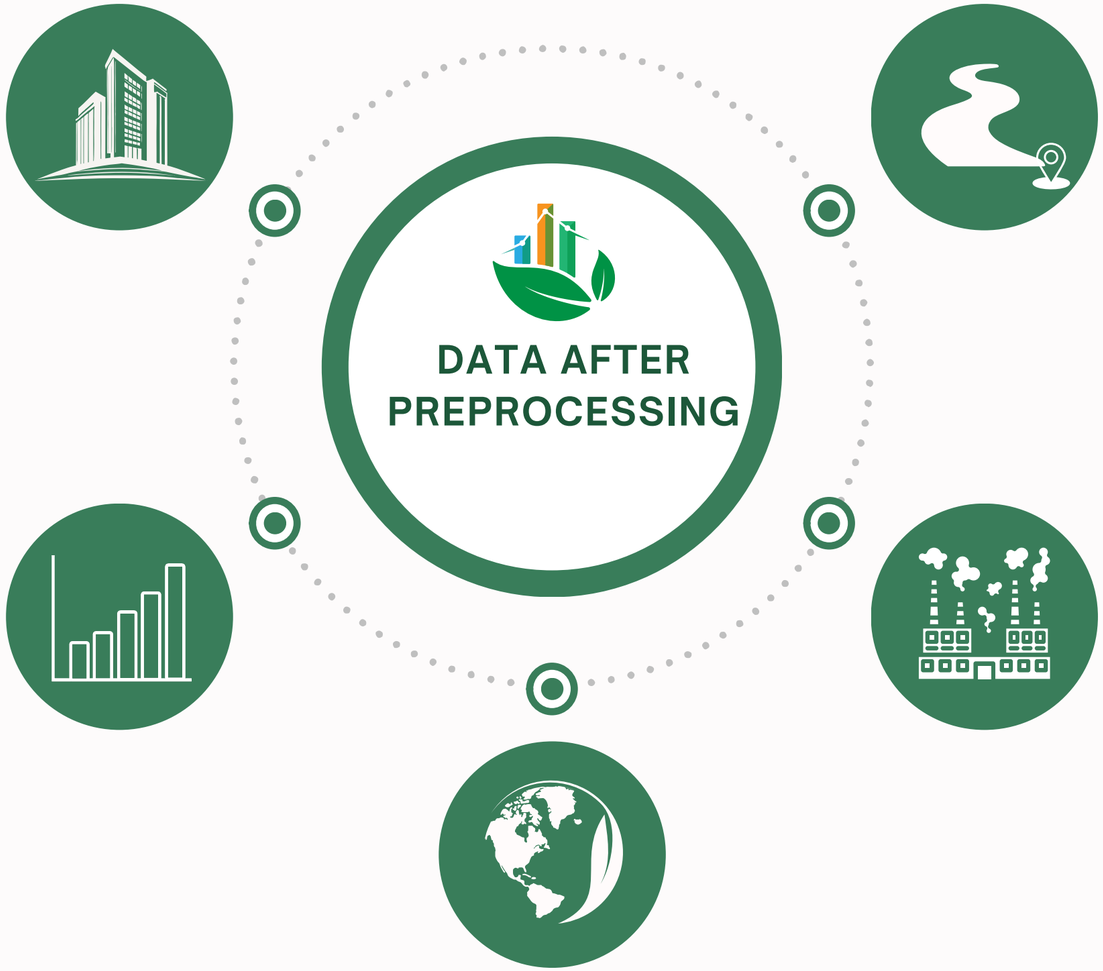
- Current Damage Assessment
- Current Damage Assessment-Type1
- Current Damage Assessment-Type 2
- Damage Type 1 or 2
- Current PSV

Structure Demographics

- Structure Type
- GIS State
- GIS City
- GIS Zip
- Parish
- NOLA Planning Demographics

Grant Statistics

- Total CG Amount
- Total ACG Amount
- Total Elevation Amount
- Total IMM Amount
- Total Closing Amount
- ARS File (Yes/No)



Census Demographics

- Census Blocks
- Block Groups
- Census Tracks

Closing Options

- Closing Damage Assessment
- PSV at Closing
- Closed File - Option 1
- Closed File - Option 2/3
- Closed with Approved Unmet Needs
- Closing total DOB

- **Difference between PSV and Current Damage Assessment**

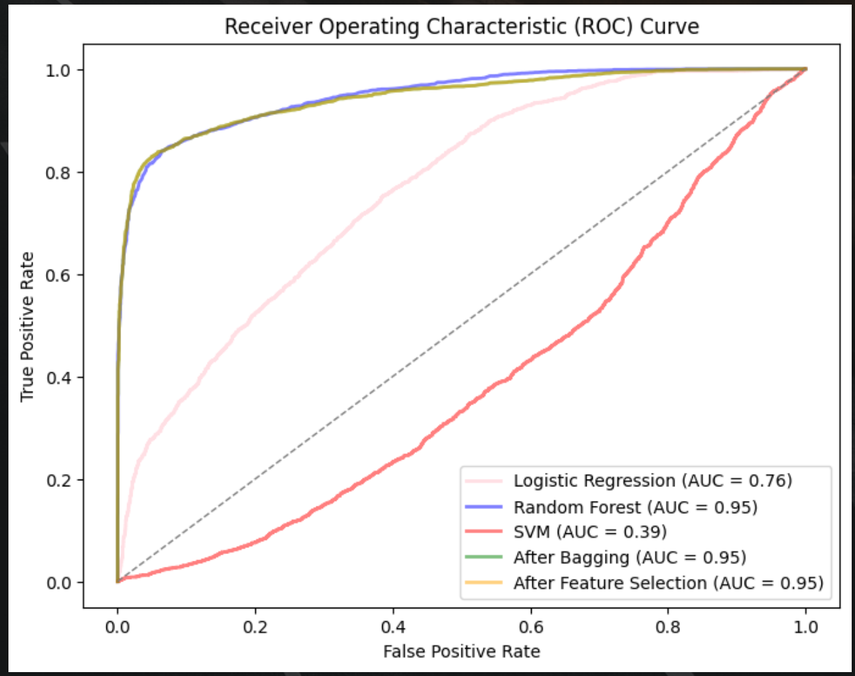
Damage Assessment

- Current Damage Assessment
- Current Damage Assessment-Type1
- Current Damage Assessment-Type 2
- Damage Type 1 or 2
- Current Pre Sale Value (PSV)

MODELS

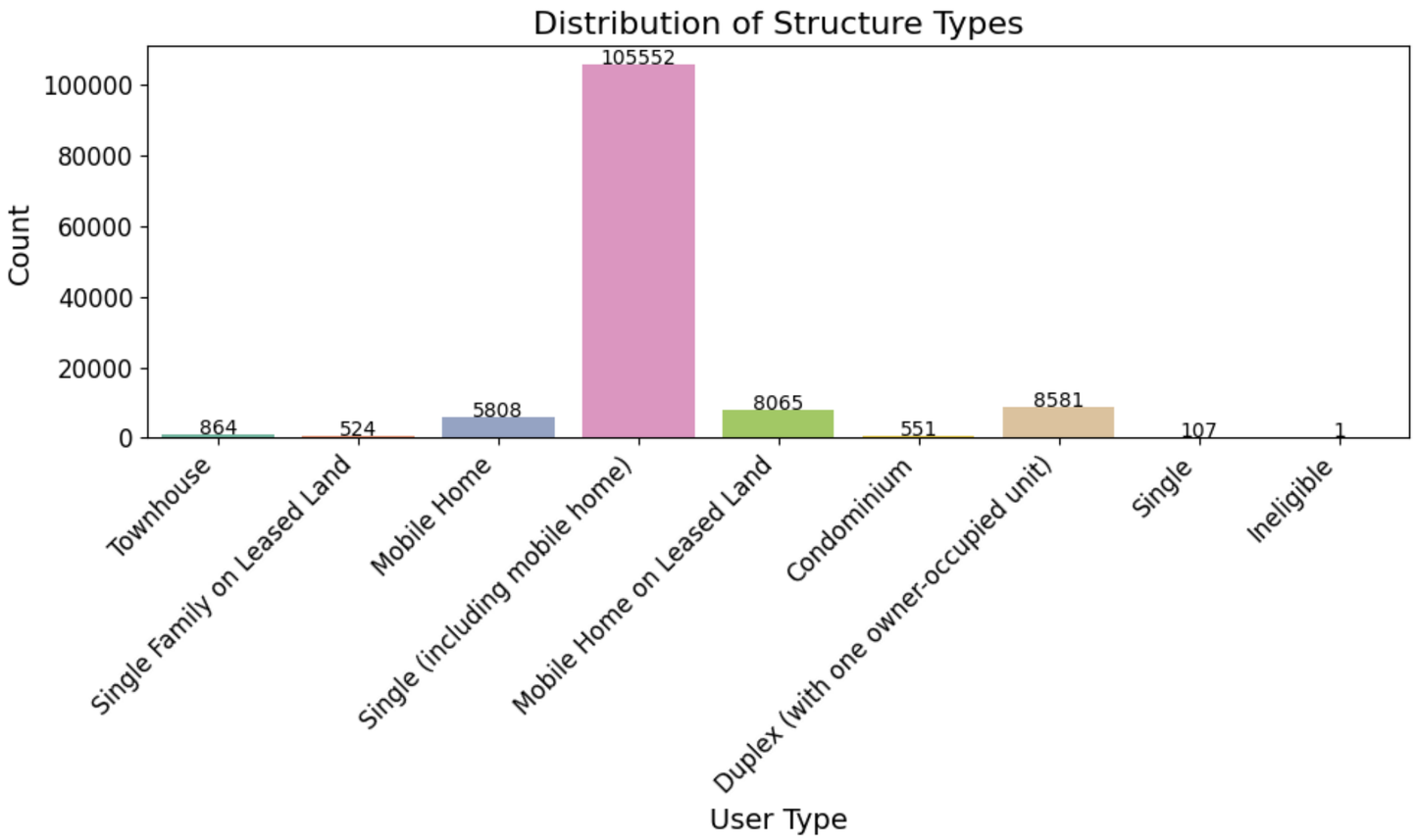
SELECTION AND EVALUATION

Model Name		Precision	Recall	F1-Score	Test Accuracy
Logistic Regression	0	0.76	0.52	0.62	67.89%
	1	0.64	0.83	0.72	
Random Forest	0	0.89	0.81	0.85	85.52%
	1	0.83	0.90	0.86	
Support Vector Machines	0	0.50	0.97	0.66	49.92%
	1	0.51	0.03	0.06	
After Fine Tuning					
Forward Subset Selection	0	0.88	0.85	0.87	86.92%
	1	0.86	0.89	0.87	
Bagging	**0**	**0.90**	**0.87**	**0.87**	**87.06%**
	1	**0.85**	**0.90**	**0.87**	



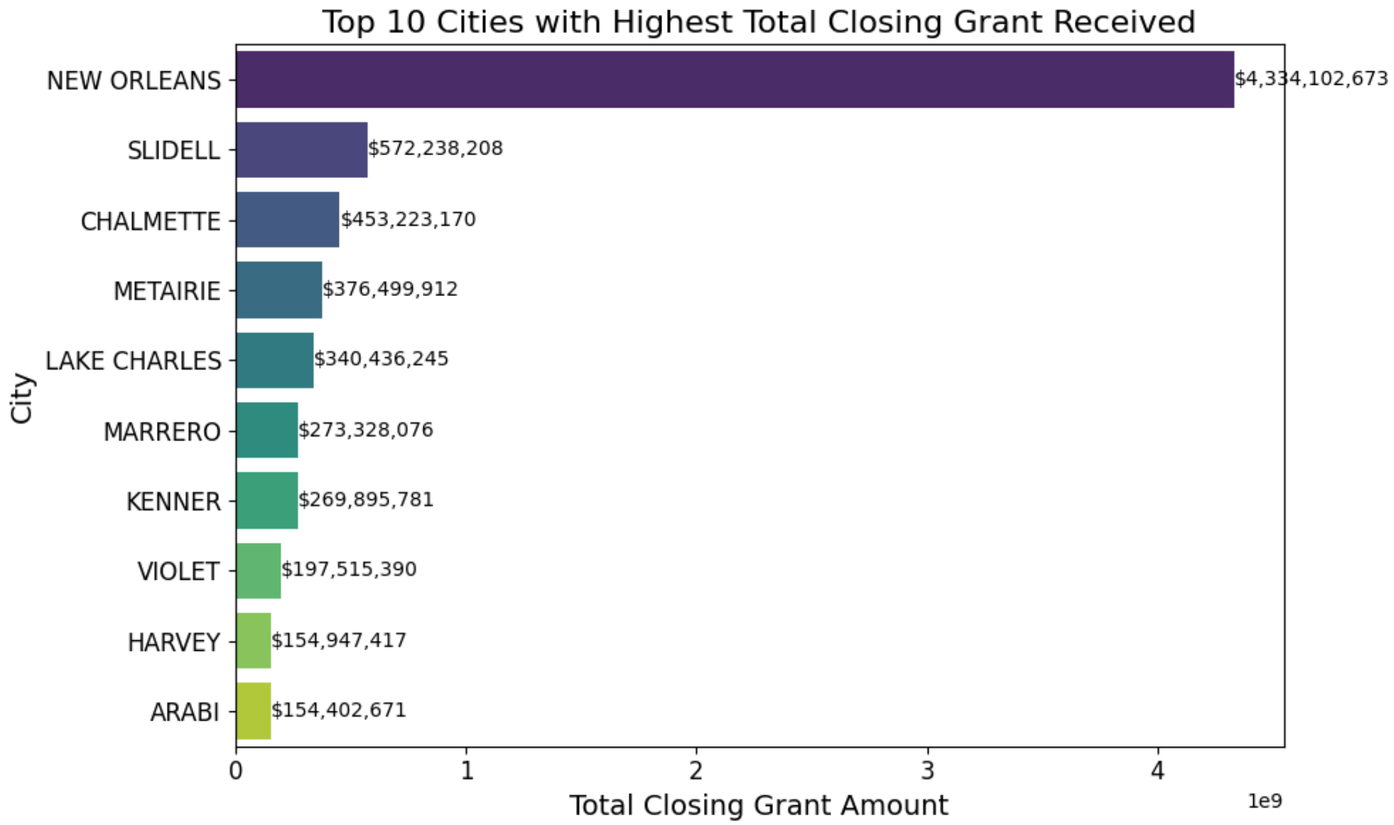
ANALYSED CoNCLUSIONS

BIFURCATION OF SAMPLES ACCORDING TO TYPE OF STRUCTURE



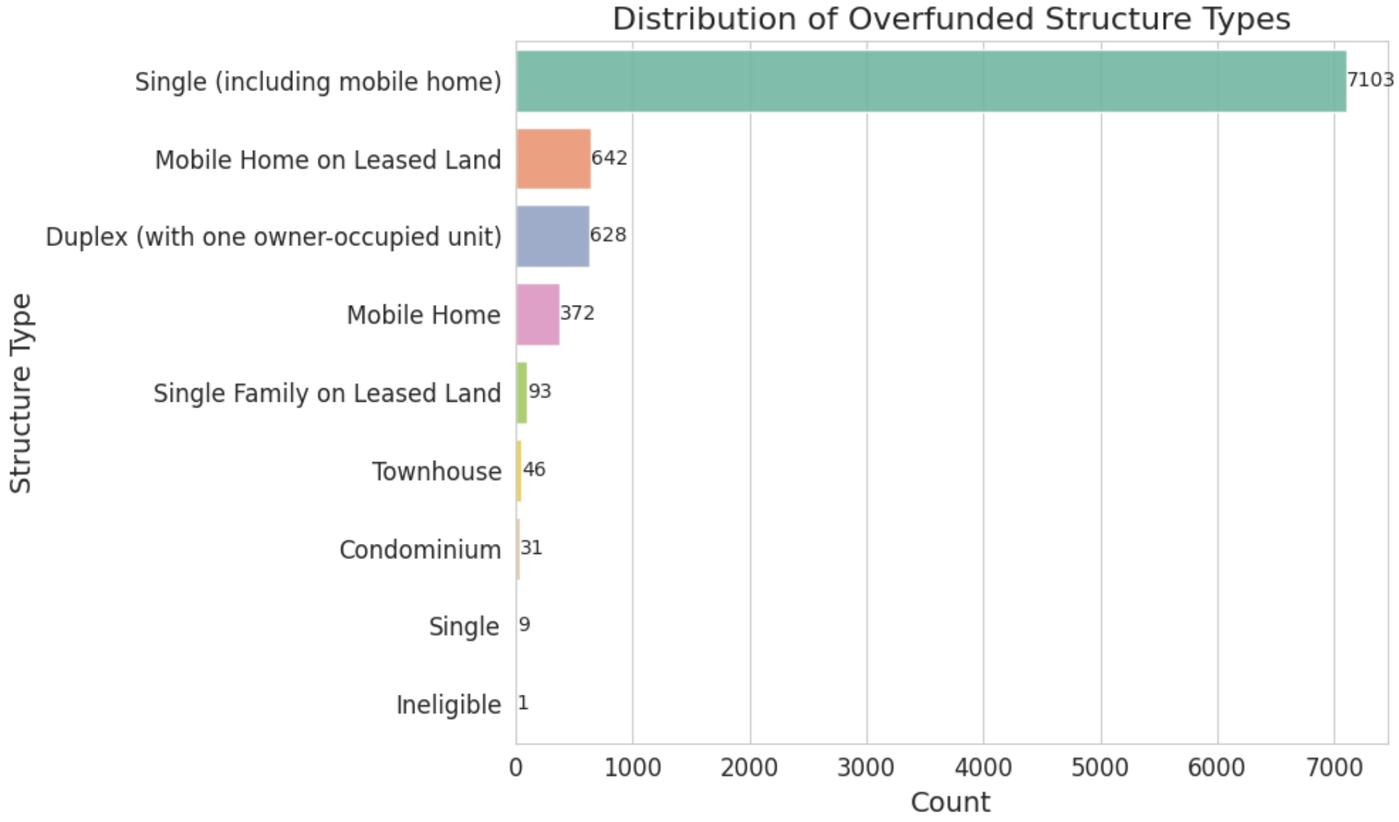
ANALYSED CONCLUSIONS

TOP 10 CITIES WITH MAXIMUM GRANTS



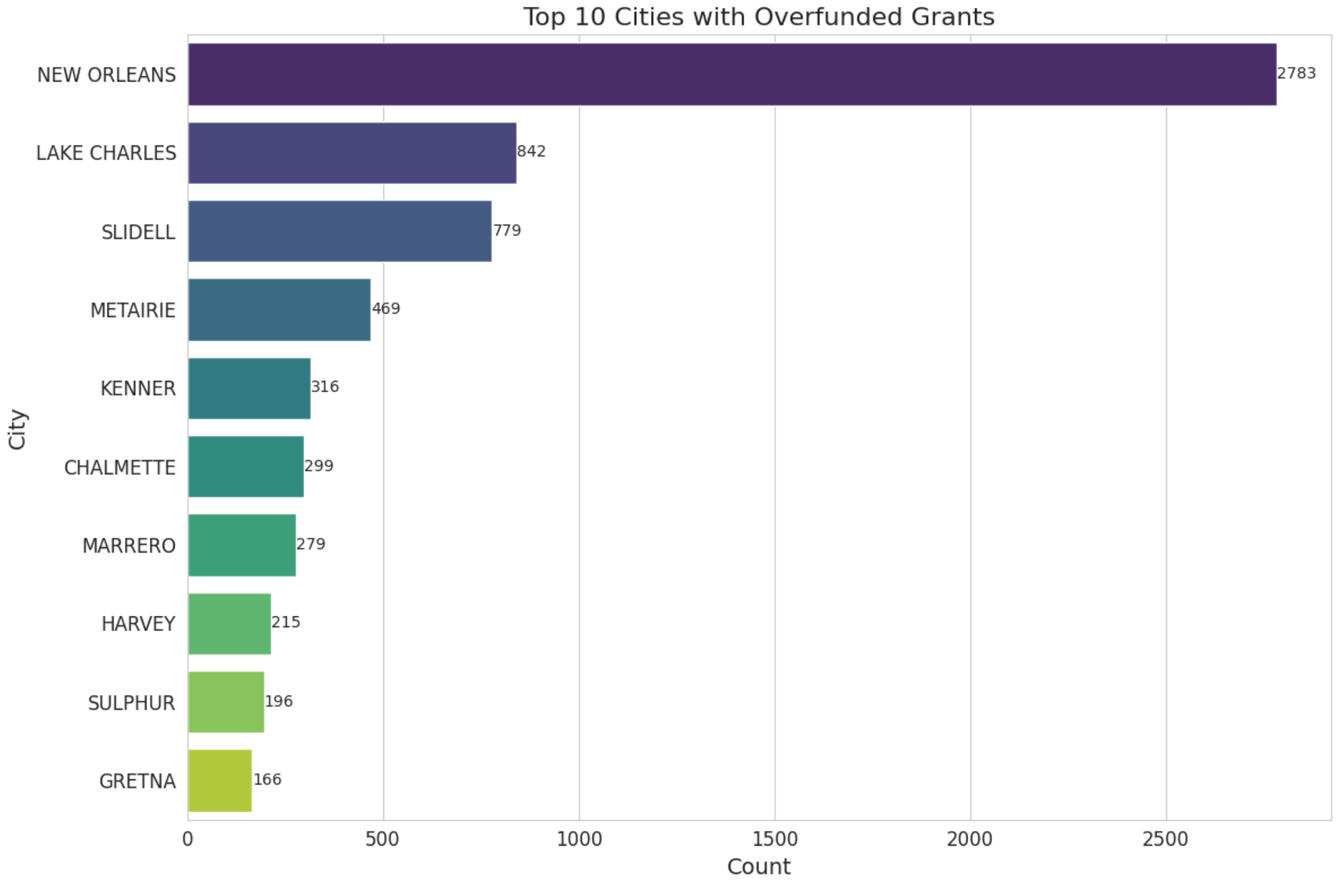
ANALYSED CONCLUSIONS

COUNT OF STRUCTURE TYPES THAT WERE FUNDED MORE THAN REQUIRED



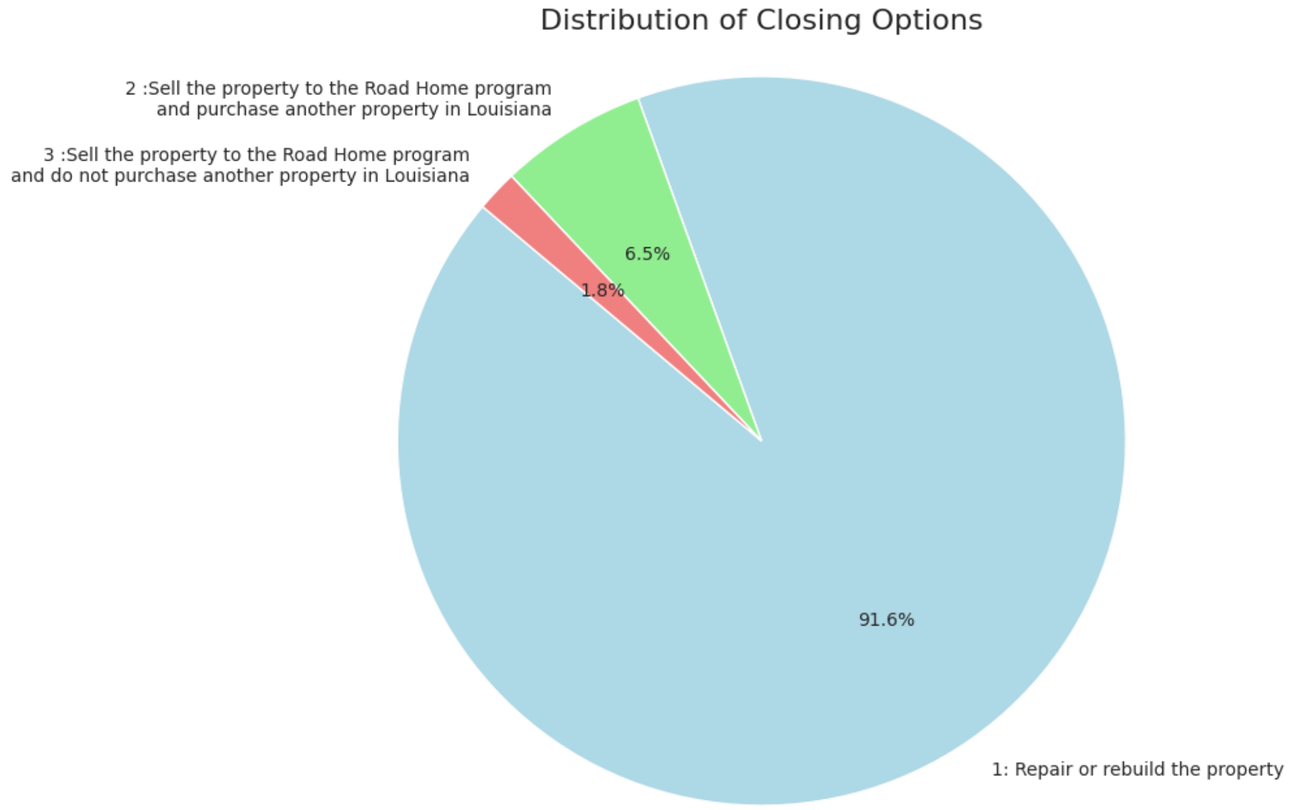
ANALYSED CONCLUSIONS

HOMEOWNERS /CONTRACTORS THAT WERE GRANTED MORE FUNDS THAN REQUIRED CITYWISE



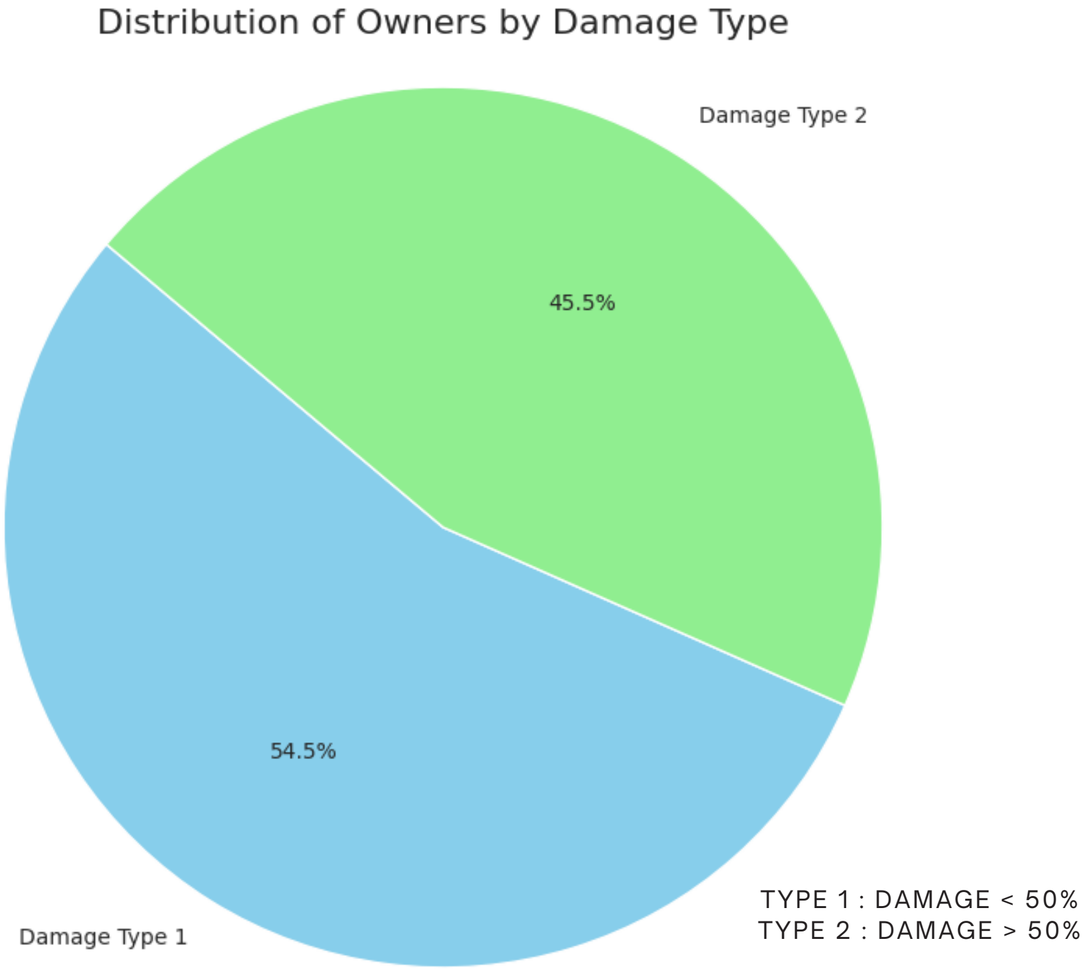
ANALYSED CONCLUSIONS

OPTIONS FOR GRANT ALLOCATION WHICH WERE SELECTED AMONG THE GRANT OWNERS.



ANALYSED CONCLUSIONS

DAMAGE TYPE OF THE STRUCTURES ASSESSED BY THE OWNERS



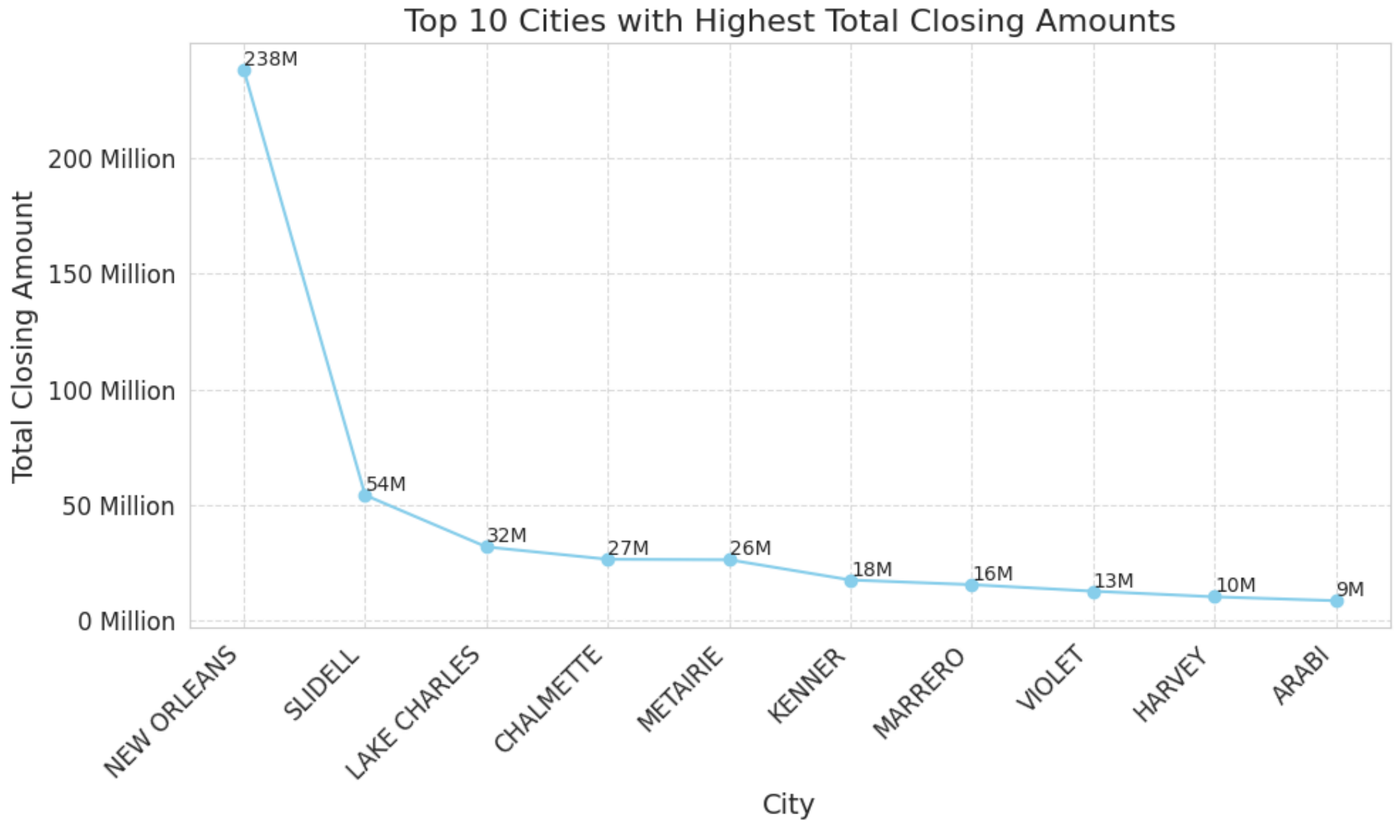
TYPE 1 : DAMAGE < 50%
TYPE 2 : DAMAGE > 50%

ANALYSED CONCLUSIONS

FINDING OUT THE TOTAL MONEY DISBURSED THAT WERE CLASSIFIED AS FRAUDULENT DISTRIBUTION

TOTAL MONEY THAT COULD HAVE BEEN SAVED

572 MILLION



CONCLUSION

- There was a **heavy class imbalance** observed in the dataset.
- **Fine-tuning Random Forest** helped to get the most accuracy among all other approaches with **validation - accuracy of 87.06%.**
- **New Orleans** was the city, with a maximum grant allocation of around **4.3 Billion**.
- A total of **572 Million USD** could have been saved if the fraudulent fund allocations had been found earlier.

THANK YOU

ANY QUESTIONS?