CPCS Findings

Kevin, Vedika, Sam, and PJ

Generalized graphing and visualization
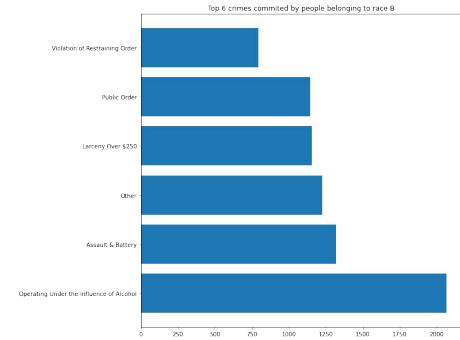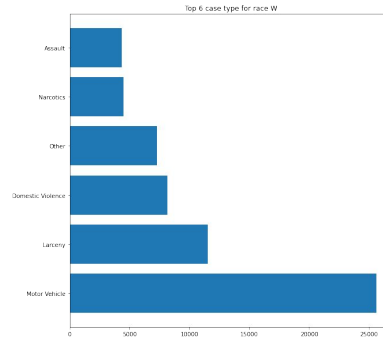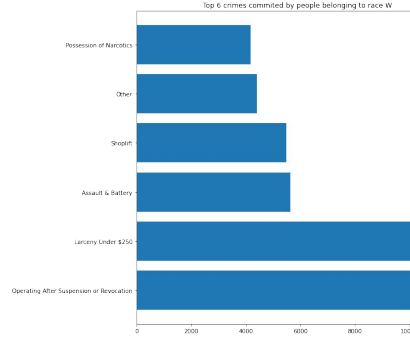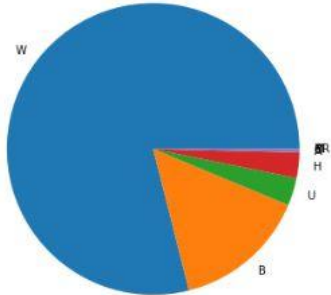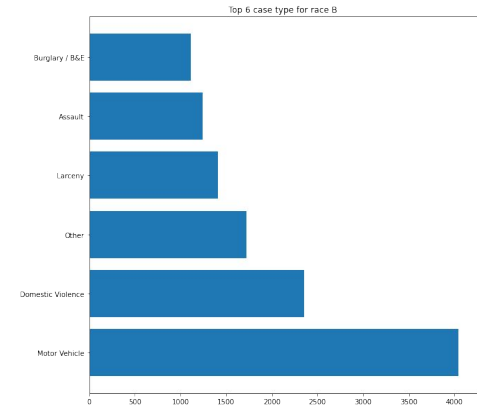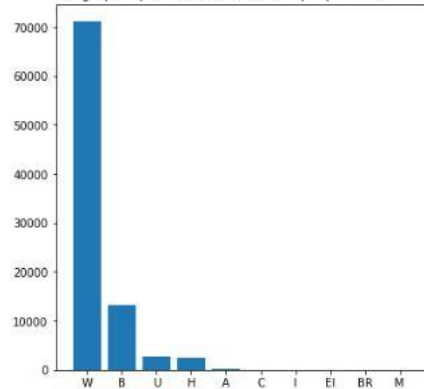
Questions that we asked

- Is there a correlation between race and prosecution frequency
- Do different groups of people have any noticeable discrepancy when committing the same crimes
- Do different groups of people commit different crimes
- What is size of each data set and how much data will I need to get from each one

Potential obstacles/limitation of the project

- Being able to meaningfully quantify the groups of people who are being discriminated against
- The potential inconsistency/discrepancy between the race of people vs their ethnicity

Berkshire and Northwestern Data collection

- Which data had correlation and how could this be expanded on
- What crime, dispositions, and sentences were most common among races (by percentage)
- Which races had the highest percentage of which crimes

Berkshire Correlation matrix:

|	Race	Gender	Case Type	Crime Type	Disp Type	Disp Desc	Sentence Translation
Race	1.000000	-0.043269	0.020531	0.024806	0.053840	0.021522	0.041500
Gender	-0.043269	1.000000	-0.109267	-0.048112	0.003908	-0.055263	-0.026162
Case Type	0.020531	-0.109267	1.000000	0.535741	0.021329	0.111777	0.044718
Crime Type	0.024806	-0.048112	0.535741	1.000000	0.016928	0.089496	0.065391
Disp Type	0.053840	0.003908	0.021329	0.016928	1.000000	0.385834	0.284710
Disp Desc	0.021522	-0.055263	0.111777	0.089496	0.385834	1.000000	0.288994
Sentence Translation	0.041500	-0.026162	0.044718	0.065391	0.284710	0.288994	1.000000

Northwestern Correlation matrix:

|	Race:	Gender:	Crime Type:	Disposition:	Sentence Type:
Race:	1.000000	-0.008368	-0.002590	0.012544	0.011472
Gender:	-0.008368	1.000000	-0.039532	-0.009664	0.062917
Crime Type:	-0.002590	-0.039532	1.000000	0.065827	0.024632
Disposition:	0.012544	-0.009664	0.065827	1.000000	0.139413
Sentence Type:	0.011472	0.062917	0.024632	0.139413	1.000000

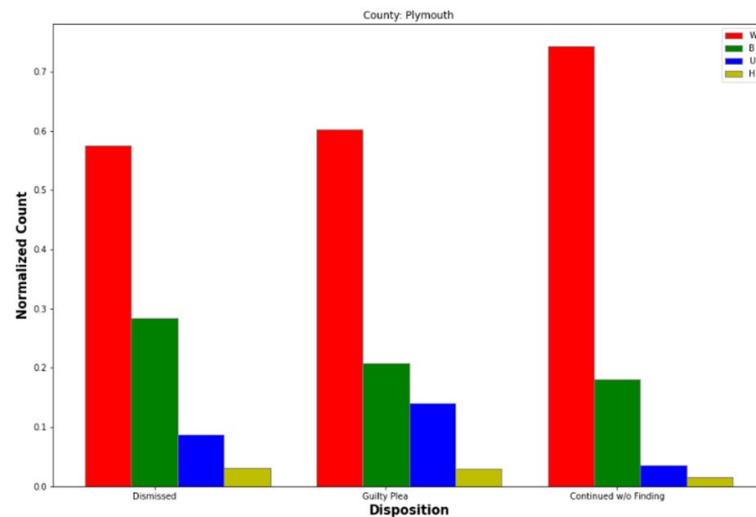# Visual Representation of Crime–Race Relation
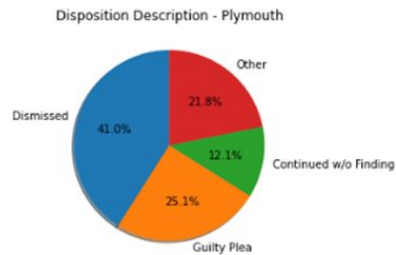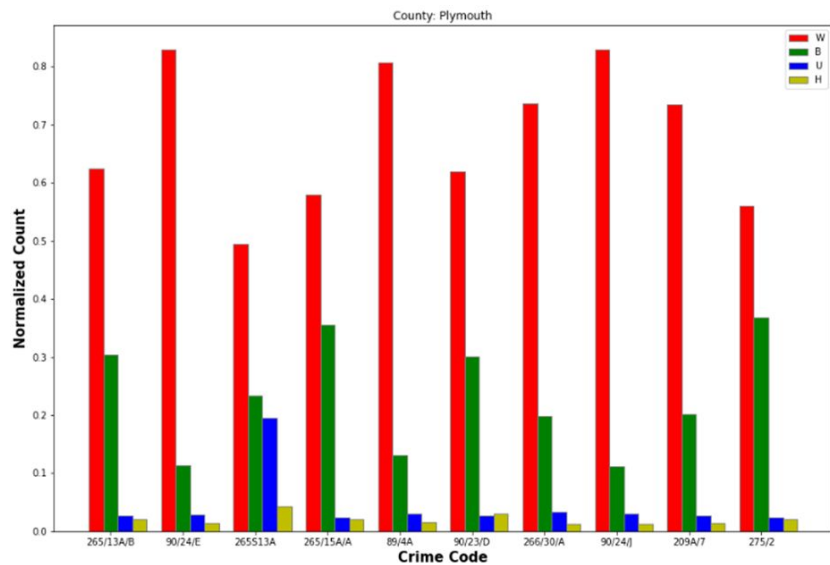

Berkshire Bar Graph Data:
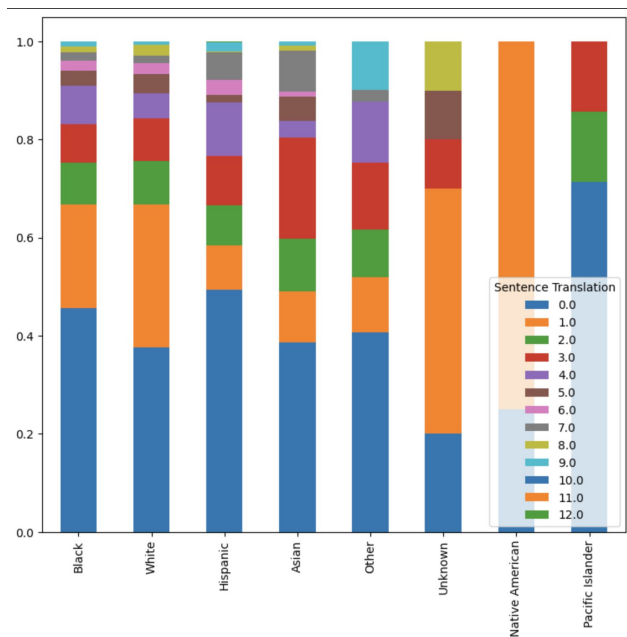
Northwestern Bar Graph Data:
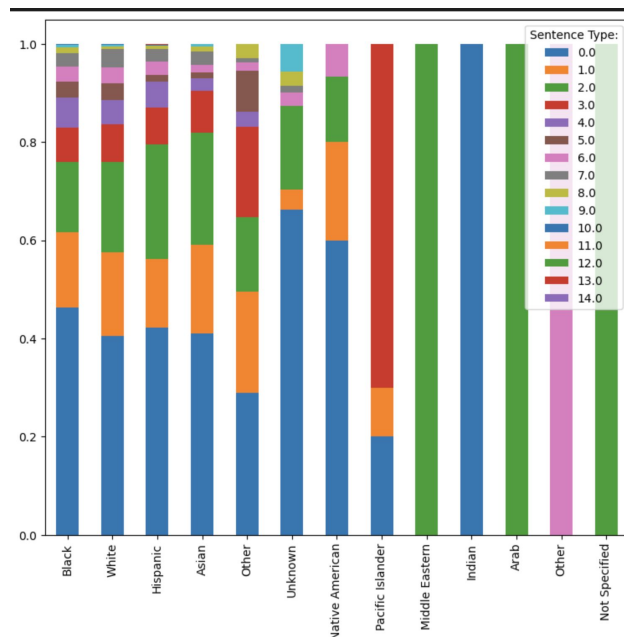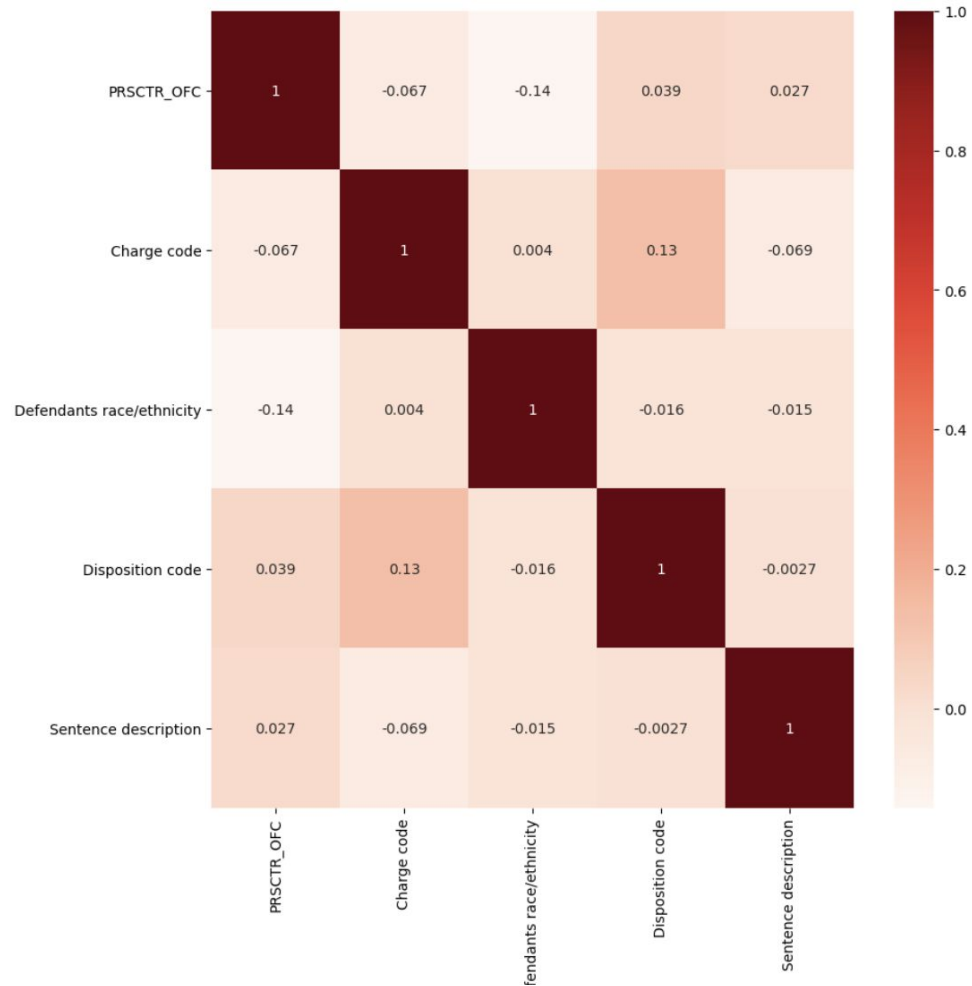
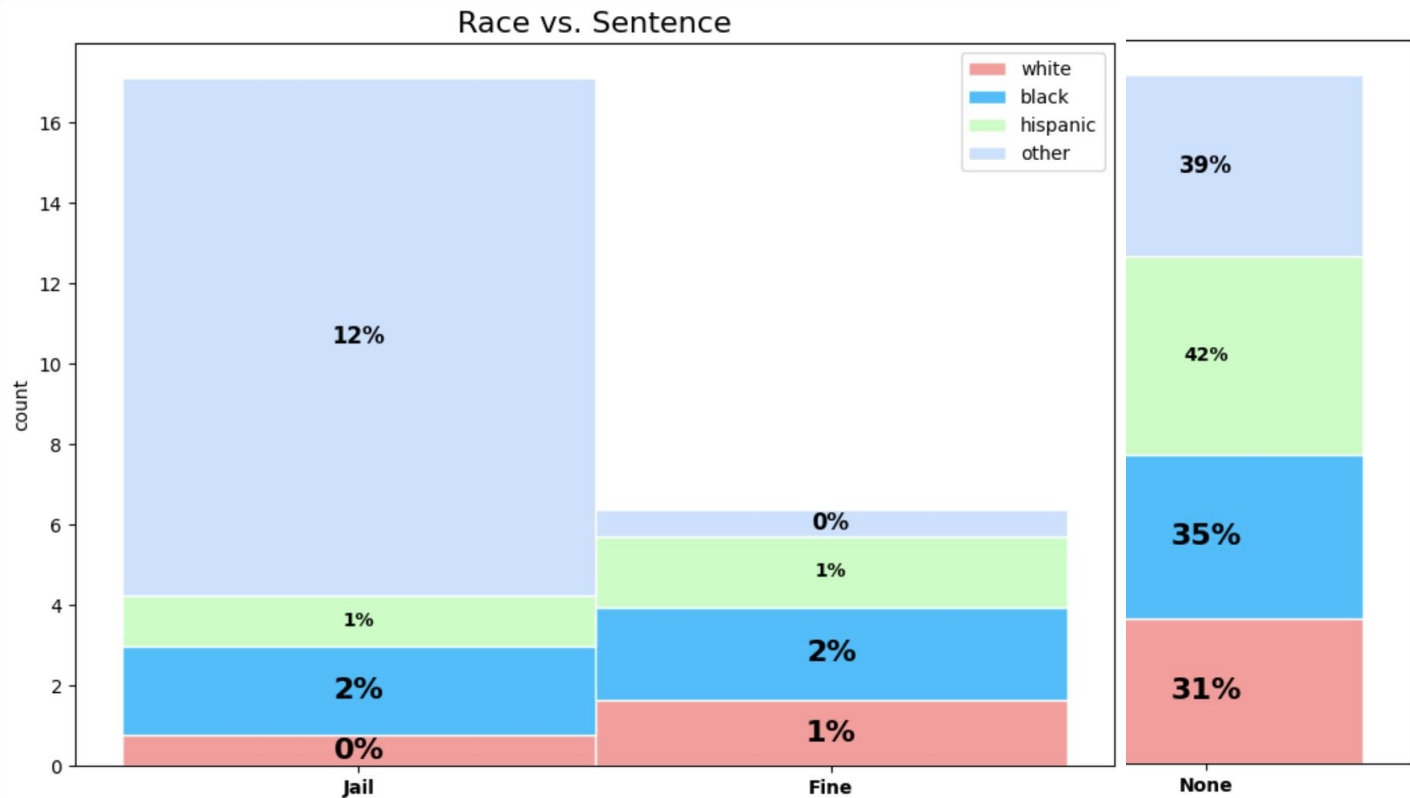# Essex Data Insights
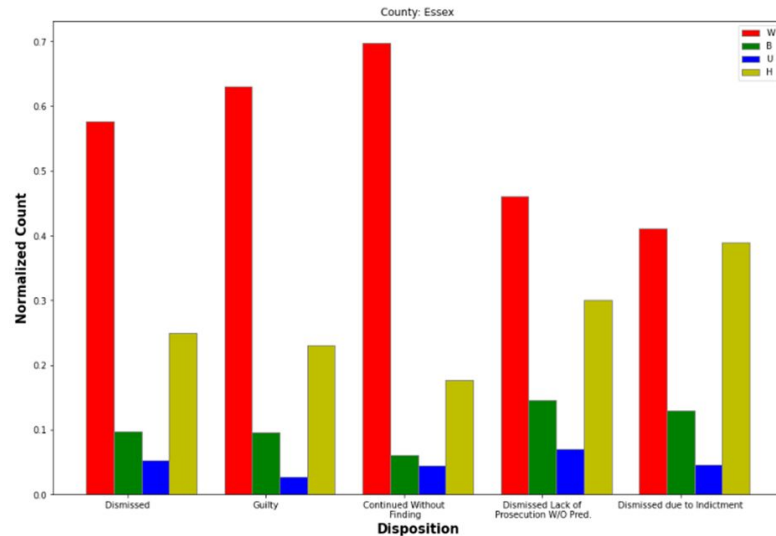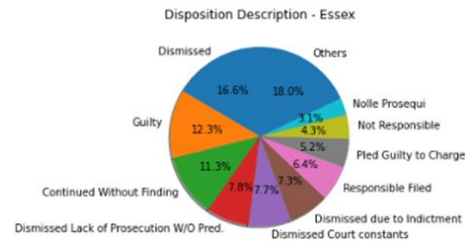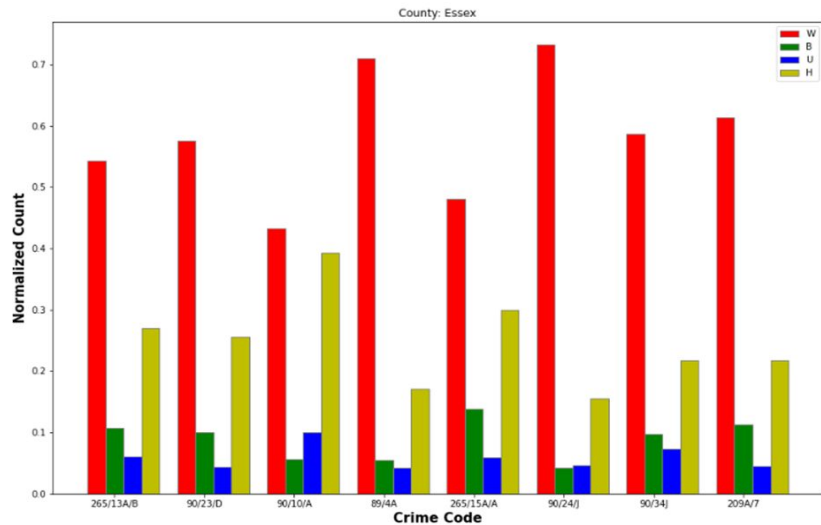
Plymouth Data Insights

Worcester

- Heatmap showing correlation between the features
- Most significant correlations with Defendant's race/ethnicity
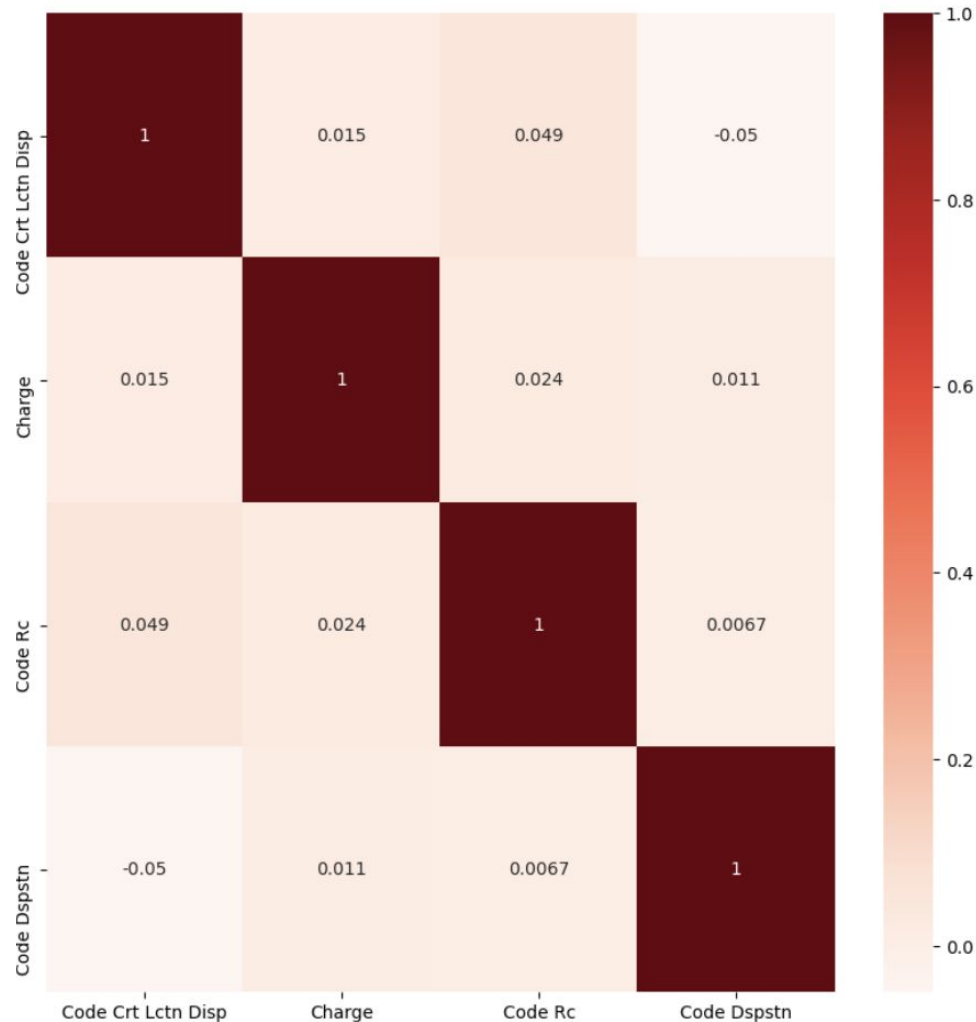 - Disposition Code
  - Sentence Description

Worcester:
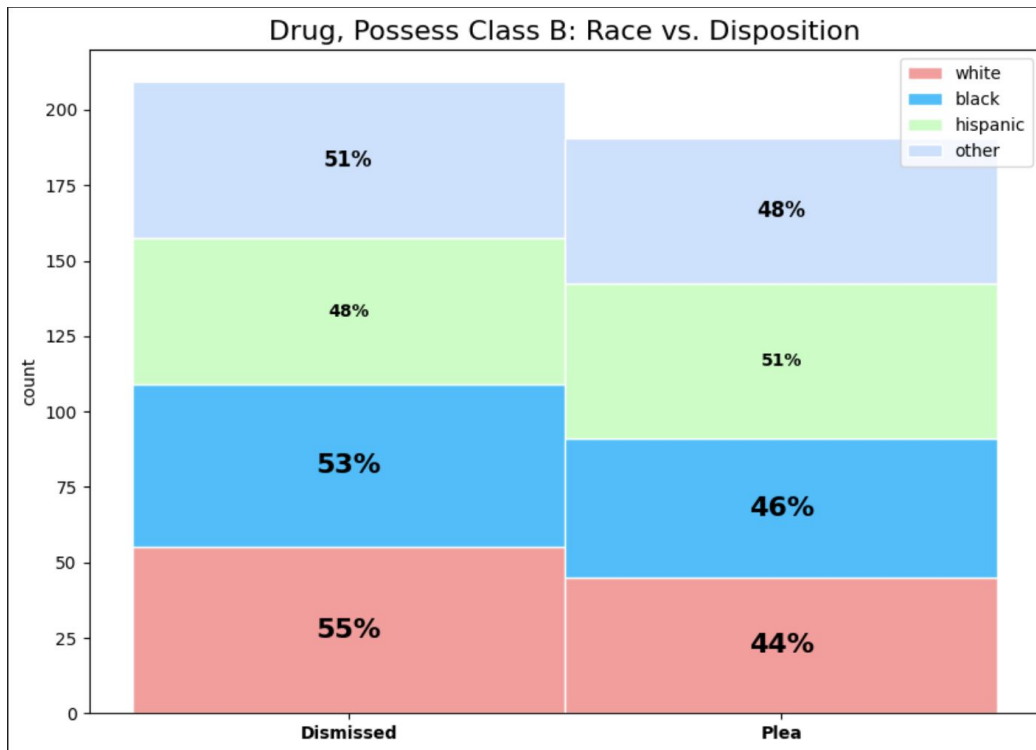# DRUG, POSSESS CLASS B c94C ?34

Race vs. Sentence

Norfolk

- Heatmap showing correlation between the features
- Most significant correlations with Defendant's Race (Code Rc)
  - Court Location (Code Crt Lctn Disp)
  - Charge

Norfolk:
DRUG, POSSESS CLASS B c94C ?34

Drug, Possess Class B: Race vs. Disposition

Thank you