Presentation 1

Links to presentation(s) and code(s) on GitHub

Presentation:
https://github.com/arvindkrishna87/STAT390_LegalAid_Fall2025/blob/main/Presentations/Data%20integrity%20checks/CARAllCallsIntegrity_7Oct_EDATeam2.pdf

Code:

https://github.com/arvindkrishna87/STAT390_LegalAid_Fall2025/blob/main/Code/Data%20integrity%20checks/AllInboundCalls_6Oct_CarolineCorr.R

What did you do?

I wrote the code for the inbound call analysis across both datasets (CAR

and all calls)

How does it help the project?

This helped the project by providing the basis for our analysis and visualization of data integrity. Using this comparison between both datasets as an indicator, we will be able to gauge how well the datasets align with one another and identify potential data quality issues.

Issues faced (if any)

We faced a lot of issues trying to read the data in at first, and we are still struggling to understand a lot of the nuances of both datasets.

Attempts to resolve issues (if any)

We luckily found some code from spring quarter of how to import the datasets in R, and we also consulted ChatGPT for other issues we had during the import process and for debugging.

Issues resolved (if any)

We have now successfully loaded in both datasets completely, but we still have a long way to go with cleaning and understanding the data.

Next steps

It was brought to my attention that I did not take into account the 6 phone lines when filtering the all calls data, so moving forward I will be able to apply that and get a better idea of how inbound calls actually match up between both datasets.

References (Mention if you built up on someone else's work)

I used the R data import code from spring quarter to get started loading in the CAR data.

Link:
https://github.com/arvindkrishna87/STAT390_SP25_LegalAid/blob/main/Presentations/explore.R