COP the DATA

Jett Shepard, Sabrina Baccam, Melissa Kinney,
Gabri Cubero, Juanes Zapata

BLUEBONNET DATA

Background:

Client: COP the DATA

Cop the Data is a nonprofit from San Antonio, TX that aims to hold officers accountable, deter misconduct, and assist in bettering hiring conditions that are in result of a lack of police transparency laws and police department collaboration across Texas. To do this, they have made a centralized hub of visuals and data to view misconduct information that is processed by each of the police precincts.

To prevent “wandering officers” (i.e. police professionals that were either suspended or fired and then move to another police precinct that are uninformed of their previous misconducts), Bluebonnet Data has aided in dashboard and database support to house disciplinary document information and automate processes.

BLUEBONNET DATA

Previous Deliverables

Previously, a PowerBI dashboard has been created and updated in a format that the client has approved located here.

This interactive dashboard has been populated with data from the San Antonio Police Department (SAPD) and Bexar County Sheriff's Office (BCSO) from the previous cohort.

Image credited to copthedata.org via https://copthedata.com/officer-suspensions-dashboard/

Updated Dashboard

Since then, our team has worked on new document intelligence models to read police files to update connected technologies with data from them.

This dashboard has been re-populated with data from the San Marcos Police Department (SMPD), and we intend to update it with the Austin Police Department's (APD) data once a flow is created and connected.

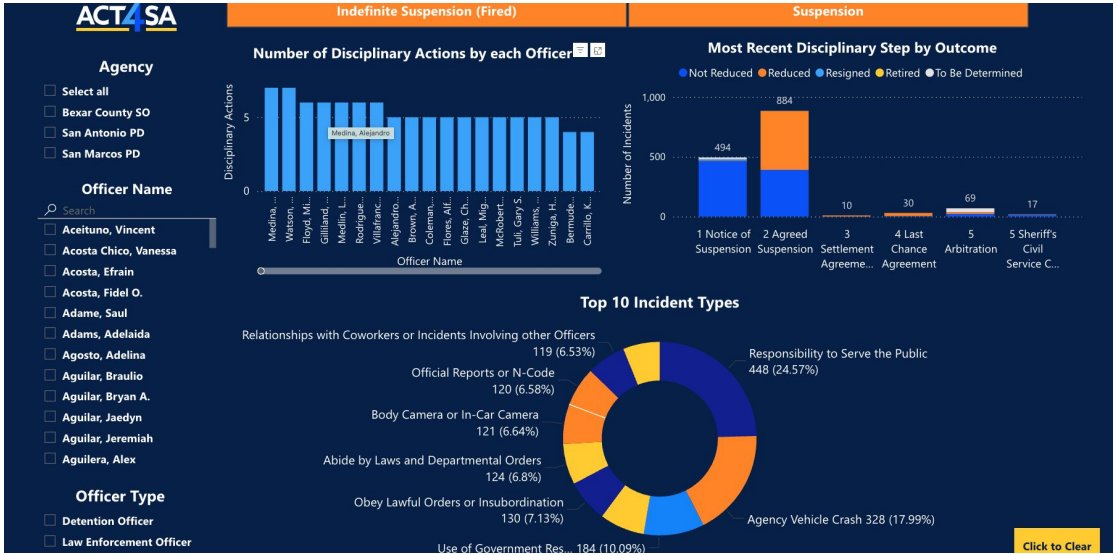

Image credited to copthedata.org via https://copthedata.com/officer-suspensions-dashboard/

Project Overview:

Bluebonnet Data's Purpose:

Bluebonnet Data aims to help populate this database and dashboard by continuing the work from before, which includes:

- Including data from different police departments and different document types
- Fixing old broken workflows and/or composing new ones
- Create, train, compose, and test new document intelligence models per document type
- Updating the Power BI dashboard and ensuring connection with associated technologies from the Microsoft Suite

BLUEBONNET DATA

Overview of the Technical Specifications

- **Data Sources:**
  - PDFs of different documents types including:
    - APD's Notice of Suspension (NoS) documents
    - APD's Agreed Suspension (AS) documents
    - SMPD's Notice of Suspension (NoS) documents
    - SMPD's Last Chance Agreement (LCA) documents

- **Tools:**

Power Apps

Data in the form of PDFs are to be inputted through a Power App

Power Automate

The data will be processed through workflows created and connected by Power Automate

Azure SQL Database

Data will then populate in the SQL database within Azure so the dashboard can read from the records

Power BI

Data is then presented via a configured Power BI dashboard

BLUEBONNET DATA

Progress Steps to Complete the Technical Goals

Completed:

- Created and iteratively trained document intelligence models for each of the document types
- Composed trained models into one composed model for each police department
  - Models can be retrained if needed
  - Composed models can be recomposed if more document types for the police department are considered
- Updated old data points in the SQL database to test new flows
- Created a Power App to process documents and connect to the workflows previously created through Power Automate
- Documentation for future cohorts including both written and video information

In Progress:

- Testing all composed models for sufficient accuracy/confidence scores for each fields

BLUEBONNET DATA

Next Steps and Blockers

Next Steps:

- All composed document intelligence models should be properly tested so that the Azure database can test out the existing flows
- Continue adding more police departments' document data
- New flow for APD - NoS model will need to be created in Power Automate and connected in the Azure SQL database

Blockers:

- Lost two cohort members and could not receive specific import keys from one until after the cohort (which prevented us from testing one model's accuracy)
- The SMPD - LCA document type model has a very limited number of documents available to train on, thus any low confidence scores from the AI's predictions could not be fixed by adding more documents to the training group and re-evaluating field values
- Our point-of-contact for our client has left their position in the middle of the cohort, so access to specific technologies were blocked and limited until they could get back from traveling

BLUEBONNET DATA