Title: Predictive EMS Dispatching with Fire Dispatch Data

Team Members: Kieren Gill, Brad Zhang

Project Description:
In a lively city like New York City, the sound of sirens are abundant. However, with so many people, cars, and traffic, based on a call from Fire Dispatch, what type of incidents will prompt EMS dispatching? How quick will EMS arrive at the scene? Our application aims to clean, process, and analyze data from both EMS and Fire Dispatch Data in New York City. Through data analytics and visualizations, we hope to gain insight into response times based on boroughs, incident types, alarm levels, and many more. To attempt to take things further, we hope to use machine learning models (multiple regression) to predict the likelihood that EMS should be dispatched based on the nature of calls received from fire stations to see if there is any overlap.

What insight will you derive from the data?
We hope that our data will be able to predict when to dispatch EMS based on calls received by the fire department - reducing wait times for first responders to assist victims of fires/etc.

Who will benefit from your analytics?
The victims benefit, and the first responders benefit because they can provide treatment quickly before further complications arise.

Describe how you will check the goodness of the analytic. i.e., why do you believe that the results are accurate and can be trusted? Include references to research papers.

After obtaining the analytics, we wish to perform a machine learning model in order to predict the likelihood of dispatching EMS based on reports made to the Fire Dispatch. Thus, to check the goodness of our model as well as being able to conclude that our results are accurate and can be trusted, we will use the train-test split concept within Machine Learning. In the research paper below, it covers the idea of splitting our cleaned dataset into a train and a test set. Once our model has been properly trained, we may compare the validity and performance with our testing set. This will produce a confusion matrix with our accuracy scores based on the effectiveness of our model.

Research paper: https://link.springer.com/article/10.1007/s10664-020-09881-0

Data sources (one per each team member). Include the source, link and schema of the datasource. *Please do not use Kaggle* Schema could be: Field Name, Data type, Brief description

Data Source 1: Fire Incident Dispatch
Source: NYC OpenData
https://data.cityofnewyork.us/Public-Safety/EMS-Incident-Dispatch-Data/76xm-jjuj
Schema:

Fire Incident Dispatch Data	
STARFIRE_INCIDENT_ID	INT
INCIDENT_DATETIME	DATETIME
ALARM_BOX_BOROUGH	STR
ALARM_BOX_NUMBER	INT
ALARM_BOX_LOCATION	STR
INCIDENT_BOROUGH	STR
ZIPCODE	STR
POLICEPRECINCT	INT
CITYCOUNCILDISTRICT	INT
COMMUNITYDISTRICT	INT
COMMUNITYSCHOOLDISTRICT	INT
CONGRESSIONALDISTRICT	INT
ALARM_SOURCE_DESCRIPTION_TX	STR
ALARM_LEVEL_INDEX_DESCRIPTION	STR
HIGHEST_ALARM_LEVEL	STR
INCIDENT_CLASSIFICATION	STR
INCIDENT_CLASSIFICATION_GROUP	STR
DISPATCH_RESPONSE_SECONDS_QY	INT
FIRST_ASSIGNMENT_DATETIME	DATETIME
FIRST_ACTIVATION_DATETIME	DATETIME
FIRST_ON_SCENE_DATETIME	DATETIME
INCIDENT_CLOSE_DATETIME	DATETIME
VALID_DISPATCH_RSPNS_TIME_INDC	STR
VALID_INCIDENT_RSPNS_TIME_INDC	STR
INCIDENT_RESPONSE_SECONDS_QY	INT
INCIDENT_TRAVEL_TM_SECONDS_QY	INT
ENGINES_ASSIGNED_QUANTITY	INT
LADDERS_ASSIGNED_QUANTITY	INT
OTHER_UNITS_ASSIGNED_QUANTITY	INT

Data Source 2: EMS Incident Dispatch

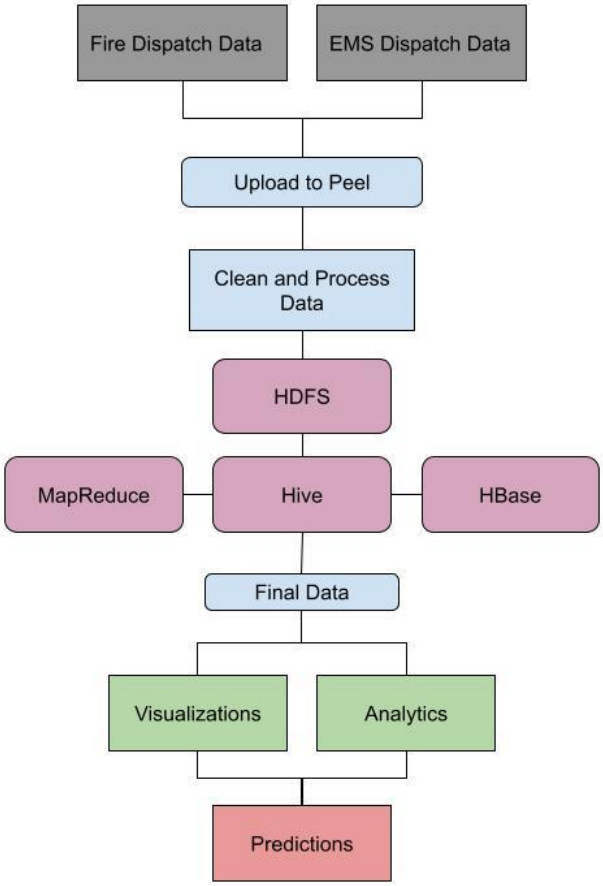
Source: NYC OpenData

https://data.cityofnewyork.us/Public-Safety/Fire-Incident-Dispatch-Data/8m42-w767

Schema:

EMS Incident Dispatch Data	
CAD_INCIDENT_ID	Number
INCIDENT_DATETIME	Date & Time
INITIAL_CALL_TYPE	Plain Text
INITIAL_SEVERITY_LEVEL_CODE	Number
FINAL_CALL_TYPE	Plain Text
FINAL_SEVERITY_LEVEL_CODE	Number
FIRST_ASSIGNMENT_DATETIME	Date & Time
VALID_DISPATCH_RSPNS_TIME_INDC	Plain Text
DISPATCH_RESPONSE_SECONDS_QY	Number
FIRST_ACTIVATION_DATETIME	Date & Time
FIRST_ON_SCENE_DATETIME	Date & Time
VALID_INCIDENT_RSPNS_TIME_INDC	Plain Text
INCIDENT_RESPONSE_SECONDS_QY	Number
INCIDENT_TRAVEL_TM_SECONDS_QY	Number
FIRST_TO_HOSP_DATETIME	Date & Time
FIRST_HOSP_ARRIVAL_DATETIME	Date & Time
INCIDENT_CLOSE_DATETIME	Date & Time
HELD_INDICATOR	Plain Text
INCIDENT_DISPOSITION_CODE	Number
BOROUGH	Plain Text
INCIDENT_DISPATCH_AREA	Plain Text
ZIPCODE	Plain Text
POLICEPRECINCT	Number
CITYCOUNCILDISTRICT	Number
COMMUNITYDISTRICT	Number
COMMUNITYSCHOOLDISTRICT	Number
CONGRESSIONALDISTRICT	Number
REOPEN_INDICATOR	Plain Text
SPECIAL_EVENT_INDICATOR	Plain Text
STANDBY_INDICATOR	Plain Text
TRANSFER_INDICATOR	Plain Text

2. Draw initial design diagram(s) using PowerPoint, Keynote, Google Drawing, etc. to describe your project.



3. Create a Task List

Create a list of tasks for your analytics project - create stories in Jira (https://www.atlassian.com/software/jira). Assign team members to tasks, and assign a due date to each task. This can be just a simple story with a titled tasks, assigned a team member name, and target completion date. Provide screenshots of your teams sprint, even if you are one person.