**HTU Upskilling Program**

**Data Science Track**

**Capstone Project: 911 Type of call Prediction.**

Trainee Name : Kifah Mryan

Date: 31-1-2023

* **Table of contents :**

|  |  |  |
| --- | --- | --- |
| **#** | **Content** | **page** |
|  |  |  |
| **►** | **Cover page……………………………………….....................................** | **1** |
| **►** | **Table of contents…………………………………………………………** | **2** |
| **1.** | [Introduction](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch1)**………………………………………………………………** | **3** |
| **2.** | [Dataset](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch2)**…………………………………………………………………....** | **3** |
| **3.** | [Creating New Features](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch3)**…………………………………………………...** | **4** |
| **4.** | [Exploratory Data Analysis (EDA)](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch4)**………………………………………….** | **4** |
| **5.** | [Overall 911 Emergency Calls](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch5)**……………………………………………** | **5** |
| **6.** | [Feature Engineering](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch6)**……………………………………………………..** | **9** |
| **7.** | [Heatmap](https://www.kaggle.com/code/tomasmantero/emergency-911-calls-exploratory-data-analysis#ch7)**…………………………………………………………………** | **9** |
| **8.** | **Git hub Source…………………………………………………………...** | **11** |
| **9.** | Results………………………………………………………………………… | **11** |

1. **Introduction**

One of the objectives of this notebook is to**show step-by-step how to analyse and visualize the dataset to better understand 911 calls and what originates them.**Moreover, is to explain most of the concepts used so that you understand why we are using them.

Given data about 911callsforMontgomery County, PA, let's try to predict the **type** of a given call.

We will use a multi-input TensorFlow neural network to make our predictions.

1. **Dataset: Kaggle**

https://www.kaggle.com/datasets/mchirico/montcoalert

* This dataset contains emergency calls from Montgomery County, PA.
* It includes calls from 2015 to 2020.

**Montgomery County**

Montgomery County, locally also referred to as Montco, is a county located in the Commonwealth of Pennsylvania. As of the 2010 census, the population was 799,874, making it the third-most populous county in Pennsylvania, after Philadelphia and Allegheny Counties. The county seat is Norristown. Montgomery County is very diverse, ranging from farms and open land in Upper Hanover to densely populated rowhouse streets in Cheltenham.

**911 Calls**

Created by Congress in 2004 as the 911 Implementation and Coordination Office (ICO), the National 911 Program is housed within the National Highway Traffic Safety Administration at the U.S. Department of Transportation and is a joint program with the National Telecommunication and Information Administration in the Department of Commerce.

For more information: <http://www.911.gov/about_national_911program.html>

Feature Columns:

* **lat:** String variable, Latitude
* **lng:** String variable, Longitude
* **desc:** String variable, Description of the Emergency Call
* **zip:** String variable, ZIP Code
* **title:** String variable, Title of Emergency
* **timeStamp:** String variable, Date and time of the call, YYYY-MM-DD HH:MM:SS
* **twp:** String variable, Township
* **addr:** String variable, General Address
* **e:** String variable, Dummy variable, Index column (always 1)

1. **Creating new features**

### **Reason feature**

In the titles column, there are "Reasons/Departments" specified before the title code. These are **EMS, Fire, and Traffic.** We are going to use .apply() with a custom lambda expression to create a new column called "Reason" that contains this string value.

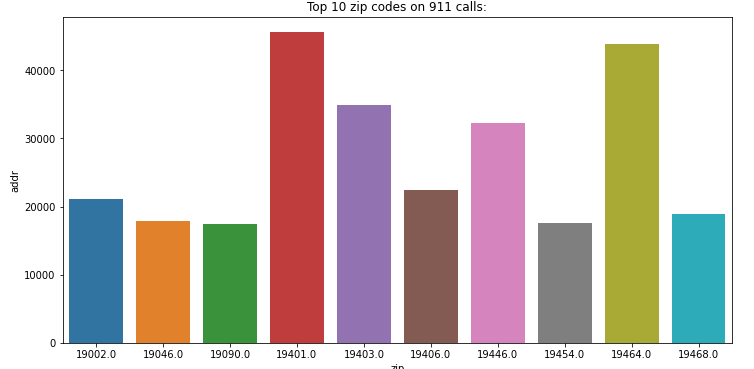
For example, if the title column value is EMS: BACK PAINS/INJURY, the Reason column value would beEMS.

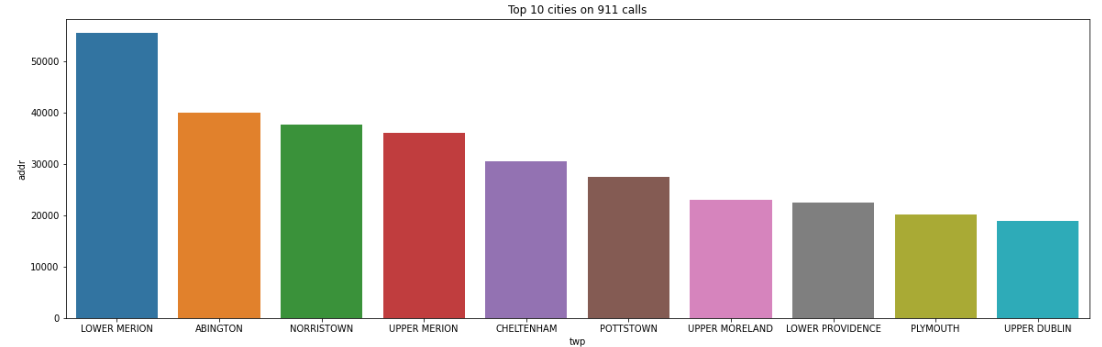
1. **Exploratory Data Analysis (EDA)**

**What is the most common Reason for a 911 call based on this new column?**

* The number one reason for 911 calls is Emergency Medical Services.
* Almost half of the reasons are for EMS.

**Visualization: python, matplot, seaborn.**

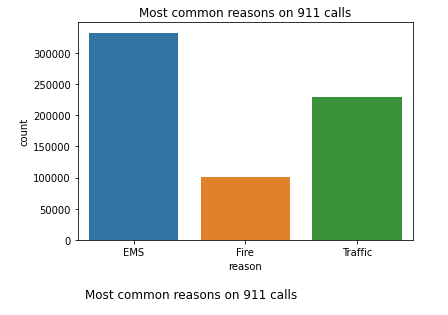




1. **Overall 911 Emergency Calls**

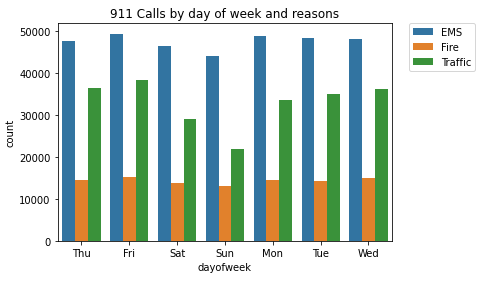
The above bar chart shows the **top 10 emergency calls** from all the categories.

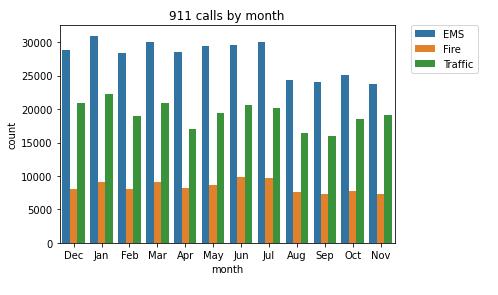
* Vehicle accidents are the number one reason people call 911.
* Disabled vehicles and fire alarms are in second and third place.

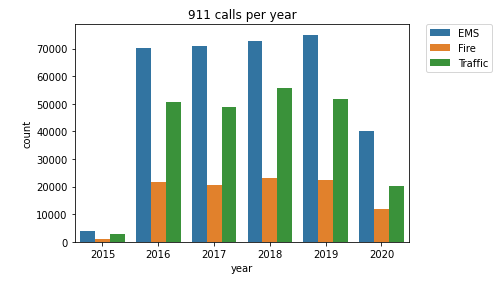


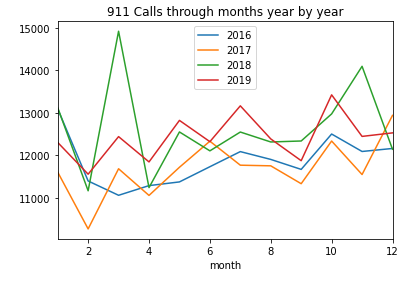
### Weekly, monthly and yearly calls

* It looks like Friday is the day with more calls during the week.
* Regarding the monthly calls, looks like during the first semester there are more calls.

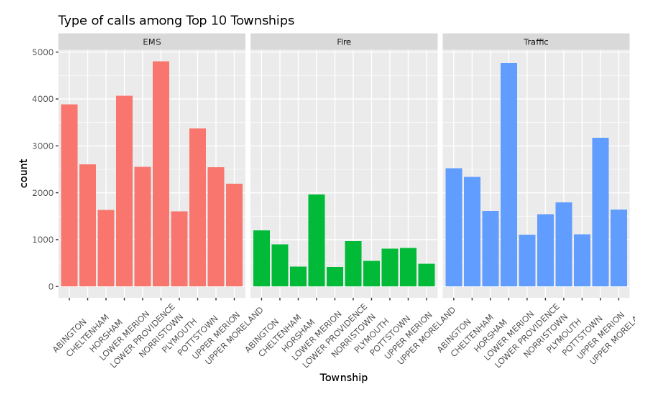








### **How does the type vary among the top 10 townships?**

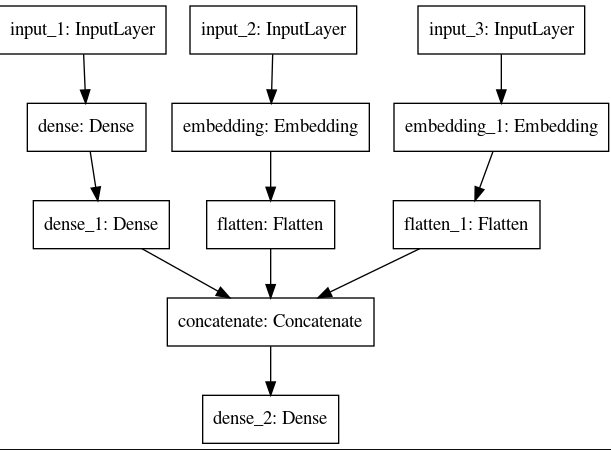


* Lower and Upper Merion have the highest number of Traffic calls
* Norristown has the highest number of EMS calls
* Lower Merion also has the highest number of Fire calls.

1. **Feature Engineering**

* **attribute information**

1. I’ve used a multi-input TensorFlow neural network to make our predictions.

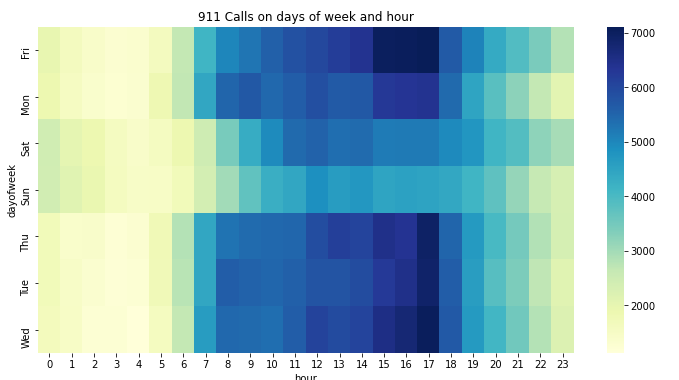


1. I will try to identify:

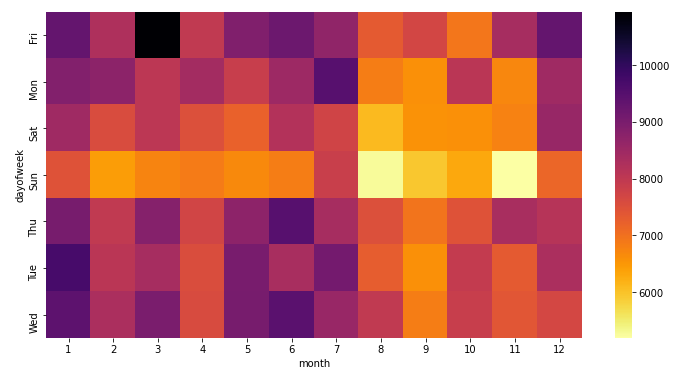
* Locations from which 911 calls are most frequent
* Time daily, monthly, and weekly patterns of 911 calls
* Major Causes of 911 calls

1. **Heat Maps:**

Now let's move on to creating heatmaps with seaborn and our data. We'll first need to restructure the data frame so that the columns become the Hours and the Index becomes the Day of the Week. There are lots of ways to do this, but I would recommend trying to combine groupby with an unstack method.



We can notice that most of the calls happen between 3 pm and 6 pm, mainly at 5 pm. The most common day is Friday, we can also say that at the weekends there fewer calls than on the other days.

We can notice that most of the calls happen on Fridays in march. We can also notice a great amount on Tuesdays in January.

1. **Git hub Source:**

[**https://github.com/kifah-mryan/911-call-type-predection.git**](https://github.com/kifah-mryan/911-call-type-predection.git)

1. **Results:**

Model loss: 0.00687

Model accuracy: 99.95%