**Visualizing Operational Efficiency in Animal Shelters**

*Milestone 1 – Gather the required data and check it’s usability*

(At the end of the page we have enlisted the reference links)

I shall begin laying the content for Milestone 1 using a famous quote by

***Edward Deming i.e. - Without data, all we have is an opinion.***

Data is considered to be the fuel for any data related project let it be Data Analysis, Data Visualization, Machine Learning, Deep Learning etc.. The first milestone thus stated is to gather data for answering the three primary questions which are as follows:

1. Over time, has the adoption rate of animals from the shelter increased or decreased?
2. Which aspect of shelter operations demands the highest financial investment?
3. How many animals are brought into the shelter on average, both yearly and monthly?

***Data Sources:***

**Parent Source**

https://www.austintexas.gov/austin-animal-center

**Animal Shelter Data**

https://data.austintexas.gov/browse?q=austin%20animal%20center&sortBy=relevance&utf8=✓

**Intakes**

https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Intakes/wter-evkm/about\_data

**Outcomes**

https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238/about\_data

For the above project we have used two main sheets from the Austin Animal Shelter Data i.e. Intakes and Outcomes. Both these sheets have been uploaded to the [GitHub Repo](https://github.com/Manasa2009/Interactive-Data-FA24/tree/2c31c744b1817d064f56f7415f0bf6156dcaad93/Final%20Project) as well.

***Data Structure:***

**Animal Intakes** data basically refers to records and statistics collected on animals taken in by shelters or rescue organizations or animal control facilities.

The following data typically includes information such as

* Animal ID
* Name
* Date Time
* Month Year
* Found Location
* Intake Type
* Intake Condition
* Animal Type
* Sex upon Intake
* Age upon Intake
* Breed
* Colour

**Animal Outcomes** data basically refers to the records detailing the eventual outcomes or dispositions of animals in shelters or rescue organizations.

The following data typically includes information such as

* Animal ID
* Name
* Date Time
* Month Year
* Date of Birth
* Outcome Type
* Outcome Subtype
* Animal Type
* Sex upon Outcome
* Age upon Outcome
* Breed
* Colour

***Data Quality:***

While I was initially taking a look into the .csv files there seems to be a lot of missing data, there would be a huge necessity to look for ways on how to handle the missing values either by eliminating the data or by imputing the data. Secondly there are few columns that can be eliminated as well such as Name, Date Time. Name is not needed for our analysis as we won’t go case by case, The highest granularity that we are thinking to dive into is Month and Year so Date Time column is not required for this project.

Out of the three primary questions one of them can’t be answered i.e. Which aspect of shelter operations demands the highest financial investment? Instead of that question we can frame few other questions to visualize such as:

* Which intake type is contributing to the influx of animals into the shelter?
* Is there any specific location from which the shelter is getting excess of intakes?
* How many animals are being euthanised on average in a year?

As we commence with the start of the project, we will be able to frame more questions and dive deeper into the dataset.