Data Science Capstone Project Ideas

The following are the three ideas I have for the Capstone project.

1. Chicago taxi data
2. Traffic collision data
3. Air Traffic landing statistics
4. **Chicago Taxi trips**

This dataset from [https://catalog.data.gov](https://catalog.data.gov/dataset/taxi-trips) contains the various details about the millions of taxi rides in the year 2015, traveling through the streets of Chicago. Each row in the dataset is a distinct taxi ride and has the following information

* Which taxi
* Trip start and end time
* Length of the trip in time and distance
* Pick up point and destination of the trip
* Cost of the trip
* Type of payment
* Taxi company

Problem Statement: Perform trend analysis on taxi trips based on the metrics like locality, cost, duration, etc.

1. **Traffic collision data**

This dataset is from [https://catalog.data.gov](https://catalog.data.gov/dataset/traffic-collision-data-from-2010-to-present), containing the traffic collision incidents happened in Los Angeles from 2010 till Present. It contains the following data

* The DR number
* Date and time when the incident happened and reported
* Area it occurred and reported
* Crime codes
* Victim details
* Location details etc.

Problem Statement: To identify collision prone locations, analyze the different age group involved in collisions, to establish preventive measures for reducing collisions based on the root cause.

1. **Air Traffic landing statistics**

The dataset from [https://data.sfgov.org](https://data.sfgov.org/Transportation/Air-Traffic-Landings-Statistics/fpux-q53t) contains data from San Francisco

International airport.

* Passenger dataset
* Cargo dataset
* Aircraft landing dataset

Problem Statement: Perform comparative analysis on air traffic trend.