**Problem Statement**

**Organization Name:** UMD Department of Transportation Services (DOTS)

**Dataset Name:** Campus Traffic Count Sensor Data**A picture containing transport

Description automatically generated**

**Difficulty:** (

Level 4: Participants with advanced data analysis skills.

The problem statement is open-ended and require multitudes of analytical perspectives and visualizations. Statistical modeling is highly recommended.

# Background

Over the summer in 2019, DOTS installed Numina sensors at 5 intersections on campus to count bikes, pedestrians, cars, buses, trucks and identify path of travel. The locations are as follows:

* Baltimore Avenue & Regents Drive (South Gate): 2 sensors on 2 poles
* Campus Drive & Paint Branch Drive: 2 sensors on 1 pole
* Stadium Drive & Regents Drive: 4 sensors on 2 poles
* University Boulevard & Paint Branch Drive: 1 sensor on 1 pole
* Campus Drive & Presidential Avenue: 1 sensor on 1 pole

The purpose of collecting traffic count sensor data is to improve DOTS’ ability to monitor traffic and understand travel behavior by pedestrians, bicyclists, and drivers over time.

# Questions

DOTS would like each team to explore this dataset in the following aspects:

* Summarize (visualize) the time trend for traffic by each mode. What is the temporal dynamic of traffic at each location?
* Explain the dynamic you observe: What factors may affect the traffic on campus? We think it’s worth exploring the impact of weather conditions, class schedules, noticeable events on campus (e.g. Game Day), and other factors that may cause traffic to fluctuate.
* After the initial explorations, you can choose either one of the following three topics to continue developing your story:
  + As part of the Campus Climate Plan, DOTS aims to reduce the number of single-occupancy vehicles. Using data analysis and visualization, come up with a few ideas to improve transportation planning and traffic management on campus
  + Build a traffic count predictive model using data science techniques and explain the usefulness of your model.
  + Use traffic and count data to measure the campus’ social, economic, and health benefits to walking, biking, and taking transit.

Use your data analytical skills and presentation skills to uncover and tell a story about the traffic on our campus. May the best analysts win the Data Challenge!

# Data Considerations

This simplified dataset consists of traffic count information by different modes (walk, bike, car, bus, and truck) at a 15-minute interval in a 2-week period (Oct 28 – Nov 10) for all 10 sensors on campus. The dataset has 7 columns and 13,480 rows. The 7 variables include a time stamp (“time”), a location identifier (“location”), and traffic count by each of the five modes.

Campus Traffic Monitor Using Data Mining Techniques (Python, R)

Explored and visualized traffic count data file of 7 columns and 14K rows to summarize the time trend for traffic by each mode in Python

Built a traffic count predictive model to explore the impact of factors that may cause traffic to fluctuate in R, and explained the usefulness to professionals