**TOPIC PROPOSAL**

**Team members**: xxx, xxx, xxxx, xxxx

**Topic:** DCConstruction Prediction

Our topic is identifying which areas of Washington, DC are likely to have construction.

Our data set is from opendata.dc.gov, and contains approximately 7,105 rows and 31 columns, including variables on construction permits like issue date, effective date, expiration date, latitude and longitude.

We intend to use various machine learning models to predict construction zones.

ATTRIBUTES:

1. Application Date 2. ISEXCAVATION 3. ISFIXTURE

**SMART QUESTION:**

1. Based on current construction permits, can we identify the level of construction in areas of Washington, DC?

**MODELS:** Logistic regression, Decision tree, Random forest, GIS Spatial Analysis

**DATASET SOURCE:** This is a CSV File from Opendata DC

Link: https://github.com/ninae555/data\_mining\_project/blob/main/Construction\_Permits\_in\_2023.csv

**GITHUB REPO Link:** https://github.com/ninae555/data\_mining\_project