**Proposal for Capstone Project**

**Title**: Predicting the time required for obtaining a building permit in NYC

**Problem**: Obtaining a construction permit can be a major hurdle for new developments in New York City. The process is highly uncertain and can take from several weeks to a year depending on numerous factors, including the zone of construction, project type, total construction cost, and many other factors. We propose to use data from New York City open data on building permit approvals to explore these factors and build a machine learning model for predicting the likelihood of approval of building permits as well as the estimated time required for the approval.

**Who might care**? Construction companies, developers, building contractors, architects, real estate investors, city administration (department of buildings), and other stakeholders can use such a model to predict the likelihood of the approval of the building permit and to estimate the time required for the approval. They can then incorporate the time estimate in their project management processes, investment return models, as well as total cost estimates. For the city administration, having an accurate estimate of the approval process of the applications in the pipeline can help to allocate resources more efficiently.

**Data**: The building permit approval data will be acquired from the New York City open data portal. This data contains information about each project approval, including detailed information about the parameters of the proposed building, the location, owner information, project cost estimates, and many more.

**Modeling approach**: the model and the approach to be decided.

**Possible limitations**: The prediction of the model will depend on the data from previous approvals. Any changes that might affect the organizational structure of the department of buildings, zoning rules or the technology used in new buildings might impact the accuracy of the predictive model.

**Deliverables**:

1. Codes (notebooks) for:

1. data acquisition
2. data cleaning
3. data exploration analysis
4. machine learning model development

2. Report on the capstone project

3. Presentation on the capstone project