Impact of Local Economic Conditions on NYC Taxi Usage and Fares

Problem Statement

The project aims to investigate how local unemployment rates and economic conditions impact the usage of taxis and the fares they can charge in New York City. Additionally, it will explore the risk premiums required by drivers to operate in economically disadvantaged areas.

Use Case

Understanding the relationship between local economic conditions and taxi usage can provide valuable insights for policymakers, taxi companies, and drivers. It can help in making informed decisions regarding fare adjustments, service availability, and support measures for drivers in economically distressed areas.

Criteria for Success

* Identification of significant correlations between local unemployment rates and taxi usage.
* Analysis of fare variations in relation to local economic conditions.
* Determination of risk premiums required by drivers in high-risk areas.
* Development of a predictive model to forecast taxi usage and fare adjustments based on economic indicators.

Scope of Solution Space

* Analysis of historical NYC taxi trip data.
* Integration of local unemployment rates and other economic indicators.
* Geospatial analysis to identify high-risk areas.
* Development of statistical and machine learning models to analyze and predict trends.

Data Sources

* [NYC Taxi Data: Available from NYC Open Data1](https://data.cityofnewyork.us/Transportation/2020-Yellow-Taxi-Trip-Data/kxp8-n2sj).
* Local Unemployment Rates: Obtainable from the Bureau of Labor Statistics or NYC government reports.(available us bureau of labor statistics

Architecture

* Create a pipeline that will download chunks of the data clean it and store in the cloud
* Use Apache spark for data analysis
* Use Collab to train model in the cloud
* Use Tableau for visualization

Deliverables

* Github Repository
* Slide deck
* Readme File
* Model that predicts rate of taxi rides
* Visualizations