**IDEA 1: Analyzing the Economic Impact of Subway Delays on Commuters in NYC**

Data 400 – Mini Project

Imagine you are a daily commuter in New York City, relying on the subway to get to work on time. However, frequent delays and unexpected disruptions add extra minutes—or even hours - to your journey. Public transportation is essential for economic activity, ensuring that workers arrive on time and businesses operate efficiently. However, subway delays impose hidden economic costs on commuters, including lost wages, increased travel expenses, and reduced productivity.

As a data analyst at MTA Distributors, my role is to analyze how subway delays impact the economic efficiency of commuters by examining additional travel time, delays during peak vs. off-peak periods, and overall journey time performance. My research question is: **“What is the economic cost of subway delays on commuters in NYC, and how does it vary by time of day and subway line?”**

1. **Data Source**

The dataset used in this study comes from Data.Gov website - MTA Subway Customer Journey-Focused Metrics dataset. This is the [link](https://catalog.data.gov/dataset/mta-subway-customer-journey-focused-metrics-beginning-2020) to the dataset.

It captures 2,856 records spanning various subway lines, time periods, and passenger experiences. This dataset provides a comprehensive view of subway performance, commuter delays, and their potential economic impact.

The dataset includes several key sections:

* Subway System Information: Division (A or B), Subway Line, and Monthly Data Collection Period.
* Passenger Flow Metrics: Total number of passengers using each subway line in peak and off-peak hours.
* Delay Indicators: Additional time spent on platforms and trains due to service delays.
* Performance Measures:

+ Percentage of trips delayed by more than five minutes.

+ Overall customer journey time performance, measuring how efficiently subway systems operate.

* Economic Impact Proxies:

+ Aggregate lost time per passenger, which can be converted into estimated economic losses using NYC’s average hourly wage.

This dataset is well-structured, with no missing values in key variables. However, additional assumptions may be required to translate lost time into economic impact, such as estimating wage levels and productivity losses.

1. **Data Analysis**

To answer the research question, this study will conduct a comprehensive exploratory data analysis (EDA) alongside economic impact estimation. The goal is to identify the subway lines and time periods that contribute the most to commuter productivity losses due to delays.

* Estimating Economic Loss Due to Subway Delays

+ Calculate total lost time per passenger by summing additional platform time and additional train time.

+ Estimate economic loss using NYC’s average hourly wage:

Economic Loss = Total Lost Time × Num Passengers × Hourly Wage

+ Compare economic losses across subway lines to identify the most costly delays.

* Peak vs. Off-Peak Delay Analysis

+ Compare over\_five\_mins\_perc (percentage of trips delayed over five minutes) during peak vs. off-peak hours.

+ Identify whether economic losses are higher during peak commuting hours, affecting worker productivity.

* Subway Line Performance Analysis

+ Rank subway lines by customer journey time performance to determine the least efficient routes.

+ Use bar charts and heatmaps to visualize which subway lines experience the most severe delays and economic losses.

1. **Implications for Stakeholders**

This study provides key insights into the economic impact of subway delays, helping transit authorities, policymakers, businesses, and commuters make informed decisions.

* City Transit Authorities & Policymakers

+ Prioritize improvements on subway lines with the highest economic losses.

+ Optimize peak-hour scheduling to reduce commuter productivity losses.

+ Support transit funding by quantifying the financial impact of delays.

* Employers & Businesses

+ Assess how delays affect employee productivity and adjust work policies accordingly.

+ Promote flexible work arrangements to mitigate lost time due to commuting inefficiencies.

* Commuters & NYC Residents

+ Help commuters make informed transit decisions based on delay patterns.

+ Advocate for real-time service updates and transparency from the MTA.

1. **Ethical, Legal, and Societal Implications**

This study analyzes subway performance data to quantify the economic cost of delays, ensuring ethical and responsible data use. Key considerations include:

* Data Privacy & Ethical Use

+ The dataset is aggregated and anonymized, containing no personal commuter data.

+ Findings will be used to improve transit efficiency rather than penalize specific subway lines.

* Equity & Social Impact

+ Lower-income commuters who rely on public transit may be disproportionately affected by delays, reinforcing economic inequality.

+ Results should be used to advocate for fairer transit policies and infrastructure investments in underserved areas.

* Legal & Policy Considerations

+ Findings can support funding proposals for subway improvements under public transit policies.

+ Transparency in reporting delay-related economic losses can push for greater MTA accountability.