NYC Subway: Interlined, or Deinterlined?

(COMP3125 Individual Project)

\*Note: Do not used sub-title

Matthew Santorsa  
*dept. name of organization*

*Abstract*—This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. *\*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract*. (*provide a short abstract*)

Keywords—Public Transit, New York City, MTA (provide 3-5 keywords)

# Introduction (*Heading 1*)

The New York City subway is an interesting beast, created back in the early 20th century to transport commuters throughout the boroughs, to help clear congestion within the city. Several lines were created, primarily running on shared tracks through downtown Manhattan. This is commonly known as interlining, and it has its problems, such as delaying trains if one breaks down within a station, or if the track along the line fails. A single delay on one train line can fan out to delay all the other lines. This brings up the question, how does interlining affect the NYC subway? Does it significantly change the average margins of time that it takes for a train to take its route, end to end? Was interlining the right way for the Metropolitan Transportation Authority (MTA) to go?

# Datasets

## Source of dataset

My dataset originates from the MTA itself, on the NY State Govt website. This makes it credible, as it is maintained by the state government, and the MTA itself. They are produced from statistics generated from subway metrics, and originally were produced at the start of 2019, as that is the timeframe of available data from when I began this project.

## Dataset Characteristics

The dataset’s format is a CSV file, and it contains the following columns: Month, Scheduled Day Type, Time Period, Line, Direction, Stop Path ID, Average Actual Runtime, 25th Percentile Runtime, 50th Percentile Runtime, 75th Percentile Runtime, Actual Trains, Distance, Average Speed, Average Scheduled Runtime, Scheduled Trains, Origin Station ID, Destination Station ID, Origin Station Name, Destination Station Name, and Number of Stops. When cleaning the Data, I removed all of the columns except for Month, Line, the percentile columns, average runtime, and the destination and origin station names. After I separated the lines, I ended up dropping the destination and origin station names as well, to help clean up the data further. Finally, I created a datetime column, and a time\_numeric column to allow for the proper handling of yearly trends over the 6 years since the dataset started. For the dataset size, it is roughly 112 thousand rows of data, pre cleaning.

# Methodology

In this part, you should give an introduction of the methods/model. First, what’s the method/model. What’s the assumption of this method/model. What’s the advantage/disadvantage of this method/model. Why did you choose it. What Python module or function do you apply to apply this method/model. Any optional input/extra work did you adjust to make the results better. If you have multiple methods, feel free to use subsection A., B. to separate them.

## Average Lines

To find the fastest line, I utilized the average runtime column as my main data type. This is because it is an aggregate of the entire months’ data, and is not a quartile like the other 3 available runtimes. The only downside is that I am taking an average of an average, which may be less accurate than if I had the original runtimes, but the MTA does not provide this data to the public. Utilizing Pandas, I was able to complete my data cleaning, and then eventually arrive at my results.

## Linear Regression

For my graphs, I utilized matplotlib, as well as sklearn linear model for my linear regression, on top of seaborn to make my graphs look more readable. Of course, Pandas was still in use for the datasets, as it is the backbone of my project.

# Results

In this section, present your findings using an appropriate method, such as equations, numerical summaries, or visualizations like charts and graphs. Clearly explain all results and provide guidance on how to interpret them. If any unexpected results arise, discuss possible reasons or contributing factors. To improve clarity and organization, consider using subsections (e.g., A, B) to separate different aspects of your results

Example: After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Result A

Example: XXX

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Results B

Example: Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

## Results C

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

# Discussion

Every method/project has its shortage or weakness. Please discuss the unsatisfied results in your project. And discuss the feasible suggestions of future work to revise/improve your result.

Example: xxx

# Conclusion

In this part, you should summarize your project. What important results did you find for your topic and what’s the effect of this result on the real-world?

Example: xxx

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

Use the IEEE format for the citation. The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...” Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

1. [1] M. T. Authority, “MTA subway end-to-end running times: Beginning 2019: State of New York,” MTA Subway End-to-End Running Times: Beginning 2019 | State of New York, https://data.ny.gov/Transportation/MTA-Subway-End-to-End-Running-Times-Beginning-2019/sp9g-mzjh/about\_data (accessed Apr. 9, 2025).

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**