Task 2.1: Create a Simulator

The goal of this task was to create a discrete-event simulator to model the operation of the bike share system. We implemented our model to mimic real life interactions over the span of 24 hours. In our model, riders arrive randomly at stations to demonstrate the randomness of customer arrival in real life. After arriving they pick a station based on a predetermined probability. If their bike is available, they begin their ride. If not, they wait. The duration for the ride is characterized by several parameters, and the simulation ends once the bike is returned.

To ensure accuracy, we ran several unit tests along with integration testing to ensure that all the individual components interacted seamlessly and accurately. We also went ahead and compared our results with other data to ensure that our model was logical.

Task 2.2: A Baseline Experiment

Using the model from 2.1, we ran an experiment in which each station had 10 bikes. We aimed to calculate the success rate of rentals for 3500 riders with 2.38 riders per minute and log-normal distribution parameters `$\mu = 2.78$` and `$\sigma = 0.619$ for the duration of the ride. We then set the simulation to run 30 times for the 3500 riders in the given day. Each run gave us a better understanding that allowed us to calculate a 90% confidence interval. We then calculated the means and standard deviations for each metric. In our simulation we saw that the overall success rate ranged from 97.91 to 98.28% whereas the average wait time was between 7.76 and 9.67 minutes. To assess the uncertainty, these 90% intervals were calculated using a t-distribution to ensure that our results were reliable. By combining our modeling with this experiment, we were able to gain invaluable insights that could help optimize bike rentals and ensure biker satisfaction.

Task 3: Idealized Experiment

We began our simulation with the assumption that each station starts with one bike. To better match supply with demand, we increased the initial count of bikes at stations with waiting riders by one after each simulation trial. The simulation continued until no riders were left waiting, ensuring that everyone could successfully rent a bike.

However, since the start and end points for each rental were randomly assigned in each trial, a successful simulation did not guarantee success in subsequent runs. To address this, we implemented a condition: the simulation would conclude only after achieving a 100% success rate in nine consecutive trials. This approach suggests a confidence level of over 90% in the stability of our results.

<u>The output from one of our simulations is presented below, showing the minimum number of bikes needed at each station to meet demand completely:</u>

- South Waterfront Walkway - Sinatra Dr & 1 St: 36
- Grove St PATH: 32
- Hoboken Terminal - Hudson St & Hudson Pl: 37
- Hoboken Terminal - River St & Hudson Pl: 37
- Newport Pkwy: 30
- City Hall - Washington St & 1 St: 31
- Newport PATH: 30
- 12 St & Sinatra Dr N: 26
- Hoboken Ave at Monmouth St: 30
- Marin Light Rail: 27
- Hamilton Park: 32
- 14 St Ferry - 14 St & Shipyard Ln: 26
- Liberty Light Rail: 25
- Columbus Dr at Exchange Pl: 26
- Harborside: 23
- 11 St & Washington St: 27
- Washington St: 24
- Sip Ave: 26
- Hudson St & 4 St: 24
- 8 St & Washington St: 24
- Madison St & 1 St: 24
- City Hall: 24
- Warren St: 25
- Newark Ave: 22
- Columbus Park - Clinton St & 9 St: 22
- Grand St & 14 St: 23
- Church Sq Park - 5 St & Park Ave: 23
- Columbus Drive: 21
- Van Vorst Park: 20
- Clinton St & Newark St: 24
- Grand St: 19
- Paulus Hook: 23
- Manila & 1st: 21
- 9 St HBLR - Jackson St & 8 St: 21
- Bloomfield St & 15 St: 20

- 4 St & Grand St: 22
- 7 St & Monroe St: 21
- JC Medical Center: 20
- Clinton St & 7 St: 20
- Willow Ave & 12 St: 20
- Morris Canal: 19
- McGinley Square: 24
- Brunswick & 6th: 20
- Jersey & 3rd: 19
- Brunswick St: 19
- Baldwin at Montgomery: 20
- Adams St & 2 St: 19
- Southwest Park - Jackson St & Observer Hwy: 18
- Marshall St & 2 St: 19
- Journal Square: 20
- Madison St & 10 St: 18
- 6 St & Grand St: 20
- Dixon Mills: 15
- Lafayette Park: 16
- Riverview Park: 19
- Stevens - River Ter & 6 St: 18
- Mama Johnson Field - 4 St & Jackson St: 16
- Pershing Field: 18
- Hilltop: 20
- Jersey & 6th St: 16
- Essex Light Rail: 16
- Monmouth and 6th: 15
- Oakland Ave: 21
- Adams St & 11 St: 16
- Bergen Ave: 17
- Fairmount Ave: 18
- Montgomery St: 15
- Christ Hospital: 18
- Astor Place: 17
- Heights Elevator: 17
- Lincoln Park: 14
- Leonard Gordon Park: 19
- Communipaw & Berry Lane: 13
- 5 Corners Library: 13
- Glenwood Ave: 14

- Union St: 12
- Dey St: 11
- Jackson Square: 10
- Bergen Ave & Stegman St: 7
- Grant Ave & MLK Dr: 6
- JCBS Depot: 2

This list highlights the variability in bike requirements, ranging from 2 to 37 bikes per station. Our methodology ensures over a 90% success rate after 10 iterations. Should we opt for a higher threshold, such as 99%, the simulation might yield even more ideal scenarios, potentially adjusting the final bike count requirements upwards to ensure availability under nearly all conditions.

Task 4: Checking Assumptions

The interarrival times were computed by calculating the time differences between consecutive trip start timestamps, which were then converted into minutes for ease of analysis. Descriptive statistics provided an initial overview, revealing the mean, median, standard deviation, and other relevant metrics. To visually assess the fit of an exponential distribution, we plotted the histogram of the interarrival times overlaid with a theoretical exponential curve based on the calculated mean interarrival time. Additionally, a Kolmogorov-Smirnov test was performed to statistically compare the empirical distribution of the interarrival times to a theoretical exponential distribution. The test results, including the p-value and KS statistic, provided quantitative evidence on the adequacy of the exponential distribution model.

Furthermore, the assumptions of stationarity and independence intrinsic to our simulation model were scrutinized. We applied rolling statistics and plotted these over time to inspect for any trends or systematic changes that would suggest non-stationarity. Autocorrelation tests were employed to assess the independence of the interarrival data points. The results from these tests helped confirm whether the underlying assumptions of our simulation model held true, ensuring its validity and reliability for simulating the NYC Bike Share system. Through this comprehensive analysis, we have reinforced the robustness of our model parameters, providing a solid foundation for further experiments and recommendations to the city planners.