

Final Project Scope and Schedule

MSDS 434

Winter 2020

By
Christine Byron

Citi Bike Customers: Real-Time Recommendations

Background

As our organization scales, we continuously seek to leverage artificial intelligence and machine learning for gaining a competitive advantage. With New York City plans to expand the Citi Bike rideshare program, our dedication to a complete rider experience is more important than ever. This means not only increasing our availability of bikes and number of stations, but also further developing our technology.

Recommendation systems have been identified as having the potential to change the way our application communicates and interacts with our riders, while also allowing our company to maximize our ROI with the information gathered on each usage and purchase. In using available information to predict how a single customer will buy our offering, we can enhance their purchasing experience, while also retaining customers and increasing overall sales. Additionally, as data is gathered it can be systematically stored to be used for future rider interactions, as well as strategic decisions across the organization.

Problem Statement.

While Citi Bike has proven to be more successful than anyone could have imagine, there is still plenty of untapped market potential within the nation's most populated city. As usage grows, we aspire to continuously appeal to new riders. Our research has indicated that there is plenty of opportunity to expand within the daily ridership group. In our assessment of how we currently interact with riders, various insufficiencies became apparent:

- Inadequate personalization for the Customer User Type: the current application is neglecting features that dynamically tailor content based on the needs of each user within this user type
- Limited metrics collected for business value purposes: the data gathered from the current application neglects key demographic data, as well as information from previous purchases, product ratings, and user behavior

Given that historically the Customer user type has only accounted for 0.11%, we feel that improved technology can provide a tremendous value to single pass riders. By incorporating recommendations to help personalize the experience as much as possible, the potential benefits of customer retention and improved engagement can be realized.

Project Scope

Using the agile methodology, this project will develop a real-time recommendation application to increase the user interaction and enrich shopping potential for our Customer user type. In this 10 week MVP, we will utilize demographic data from customers, as well as information from previous purchases and user behavior to predict which pass our daily rider should select.

Data Understanding

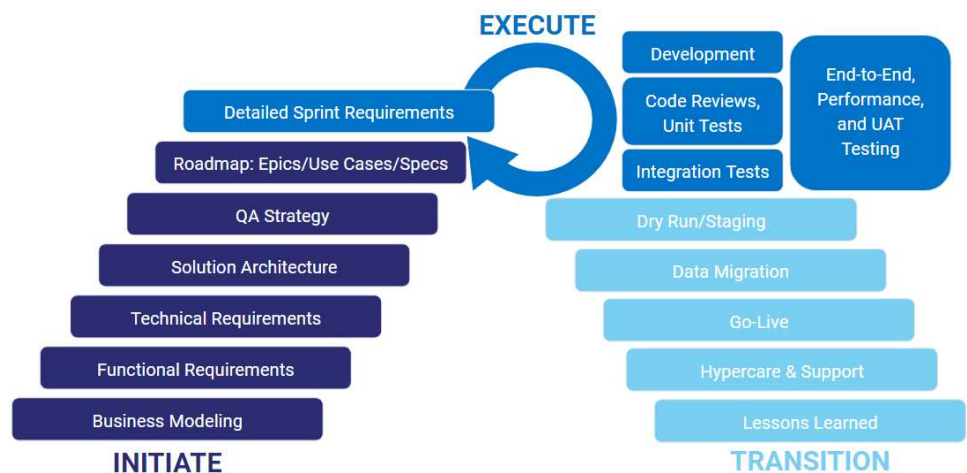
This project will focus on the Citi Bike daily ridership and membership data hosted in Google BigQuery. This dataset consists of the trips collected since its launch in September 2013, and is updated daily. It is important to note that the data has been processed by Citi Bike to remove all trips that are taken by staff to service and inspect the system, as well as any trips below 60 seconds in length, which are considered false starts.

The data includes:

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long
- Bike ID
- User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member)
- Gender (Zero = unknown; 1 = male; 2 = female)
- Year of Birth

Assumptions

This vision is dependent on thoughtful planning and agile execution. Using a defined framework, the project team will be committed to a successful and collaborative delivery through the initiate, execute, and transition phases of the project.



Applying defined to-be requirements, proposed solution, impacts and cross-team coordination, as well as thorough analysis, project management will ensure that the plan is efficiently executed. They will do this while also managing timelines, budgets, issues & risks. Execution of those requirements will be completed by engineering scrum teams using an agreed upon environment strategy and continuous development. Designated environments will host build and testing activities through the application development lifecycle.

10 Week Project Plan

Application/feature demonstrations are invaluable for keeping stakeholders up to speed with the progress of product development. We encourage questions and sharing of feedback to Product Management and the Scrum team to help maximise the ultimate value of the product.

Over the next 10 weeks, the team will commit to producing the following demos:

Week	Milestone: High Level Goal	Demo Deliverables
1	Project Kick Off	Review Project Charter, Problem Statement, Scope, and Timeline
2	Environment Creation	Execute the project environment strategy
3	Continuous Deployment	Demonstrate approach for integrating code into a shared repository
4	Setup GCP	Create Google Cloud Project and produce skeleton for the application
5	Setup BigQuery	Enable BigQuery in the GCP to run queries, load data, and export data
6	Modeling & Prediction	Evaluate and transform variables with BigQuery ML
7	Setup AutoML	Produce predictions using AutoML Tables
8	Setup Stackdriver	Show performance and diagnostics monitoring capabilities
9	Integrate API & Manage Billing	Verify integration of the application with Google Cloud Platform Billing API
10	Deploy MVP	Reveal final analytics application and release to production