Cyclistic bike-share analysis case study Timeline

## Project Start

Owner: Dawid Sosinski

Date : 19/12/2021

Status: **Complete**

Notes:

* Project start meeting took place on 19/12/2021
* All parties involved were present
* All agreed to open the project
* Final project delivery date was agreed - No later than 28/02/2022

Complete Date:19/12/2021

Sign Off Report: [Report](https://docs.google.com/document/d/1TcvmUB6T-_At-F1LhQ3CY0KGTqXK1kA7KK8iFdu6u8M/edit#heading=h.paiaet1lpazn)

## Project requirements

Owner: Dawid Sosinski

Estimation: 1 week

Status: **Complete**

Notes:

* Meeting with Lily Moreno has taken place on 22/12/2021
* All requirements were comunitacted and saved in this document
* Final delivery date was confirmed
* Weekly status update meeting was arranged
* Project plan was shared with Lily Moreno and Cyclistic marketing analytics team
* Any further changes to the scope of the project need to be presented in this document as comments and approved during weekly status meetings

Complete Date: 27/12/2021

Sign Off Report: [Report](https://docs.google.com/document/d/1TcvmUB6T-_At-F1LhQ3CY0KGTqXK1kA7KK8iFdu6u8M/edit#heading=h.hbs4uginn00j) - section Problem Statement, Goal and Key Results.

## Data Preparation

Owner: Dawid Sosinski

Estimation: 1 week

Status: **Complete**

Notes:

* Data was downloaded from [here](https://divvy-tripdata.s3.amazonaws.com/index.html)
* Data was organized into separate folders according to the formats
* Email was sent to validate data source legitimacy
* Data formats were assessed
* Decision regarding data to analyse was made
* Risk assessment was completed

Complete Date: 5/01/2021

Sign Off Report: [Report](https://docs.google.com/document/d/1TcvmUB6T-_At-F1LhQ3CY0KGTqXK1kA7KK8iFdu6u8M/edit#heading=h.veqs8ks90tid) - Section Data

## Data Analysis

Owner: Dawid Sosinski

Estimation: 1 week

Status: **In progress**

Notes:

* Tech assessment has been made
* Merging all data files into one file using R has been done
* Data cleaning has need completed
* Due to large amount of data, 3.7GB, it was decided to use AWS Athena for SQL processing
* Data exported to parquet file, uploaded to AWS S3. AWS Crawler executed.
* Data available in AWS Athena
* SQL queries executed; graphs created in GSheets

Complete Date: 12/01/2022

Sign Off Report: [Report](https://docs.google.com/document/d/1TcvmUB6T-_At-F1LhQ3CY0KGTqXK1kA7KK8iFdu6u8M/edit#) - Section Data Analysis

## Final Report

Owner: Dawid Sosinski

Estimation: 1 week

Notes:

Complete Date:

Sign Off Report:

## Presentation to the Executive Team

Owner: Lily Moreno

Date: 28/02/2022

Notes:

Complete Date:

Sign Off Report: