# ADS-507 – Initial Final Team Project Proposal

Fill out this form and submit it by the end of Module 3 in Canvas.

Team Number: \_\_\_\_\_2\_\_\_\_\_\_\_\_

Team Leader/Representative: Marvin Moran

Full names of team members:

1. Marvin Moran

2. Muris Saab

3. Ravita Kartawinata

Title of your Production-Ready Data Pipeline project: Cocktail data pipeline

Short description of your project and objectives:

We will utilize 4 datasets to construct a data pipeline for non-alcohol and alcohol beverages with ETL approach that would store a final dataset which will be used to perform EDA.

Name of your selected datasets:

1. Beverages by name

2. Beverages by ingredients

3. Beverages by category

4. Beverages by glass type

Description of your selected datasets (data source, format, size of dataset, etc.):

<https://www.thecocktaildb.com/api.php>

Beverages by name : Table

Description automatically generated with low confidence

Beverages by ingredients: A picture containing table

Description automatically generated

Beverages by category : Graphical user interface, text, application

Description automatically generated

Beverages by glass type: Graphical user interface, text, application

Description automatically generated

Please provide the link for your GitHub repository here: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

How many times have your members met in the last two weeks? 2 times

List the specific contributions that each team member is providing for the Final Team Project in the table below.

* **NOTE:** ALL students on the team should contribute equally to the Final Team Project.

| Team Member 1 (Marvin) | Team Member 2 (Muris) | Team Member 3 (if applicable) (Ravita) |
| --- | --- | --- |
| * Team Leader * Identifying datasets to use * Extract, Transform and Load Data * Generate design document and presentation | * Identifying datasets to use * Extract, Transform and Load Data * Generate design document and presentation | * Identifying datasets to use * Extract, Transform and Load Data * Develop a ground work of ipynb |

Comments/Roadblocks: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

<https://www.kaggle.com/datasets/mohamedharris/supermart-grocery-sales-retail-analytics-dataset>

<https://www.kaggle.com/datasets/abhishekrp1517/online-retail-transactions-dataset>

CSV Files

A picture containing box and whisker chart

Description automatically generated

API – json files