**Capstone Project**

Case Study -

U F H

Its an ecommerce company dealing in organic groceries.

They are still an early stage startup. They operate in 20 Cities in India at this moment accross 6 States. They promise to deliver the order within 48 hours after order placement. They offer 50 different type of products in different SKUs.

At this moment they recieve between 500-1000 orders per day across India.

They soon plan to expand their capabilities to serve between 10000 to 20000 orders per day.

Work Flow-

Diagram

Description automatically generated

1. Create MySQL cloud SQL instance on GCP
2. Create an OLTP schema using MySQL(Create an SQL script containing the queries for table creation and import it into your SQL instance.)
3. Generate fake data using faker library and store in dataframe.
4. Insert the fake data generated into OLTP schema using sqlalchemy and pymysql.
5. Create star schema defined in table in Big Query(Create an SQL script containing the queries for table creation and run it in bigquery editor)
6. Build the ETL pipeline to extract data from MySQL using pandas\_gbq and Sqlalchemy then transform it into star schema table then load them into bigquery using the same libraries.
7. Run incremental ETL to load 5000 more records and populate them into oltp and then in bigquery.
8. Perform analysis on the transferred data using BigQuery.