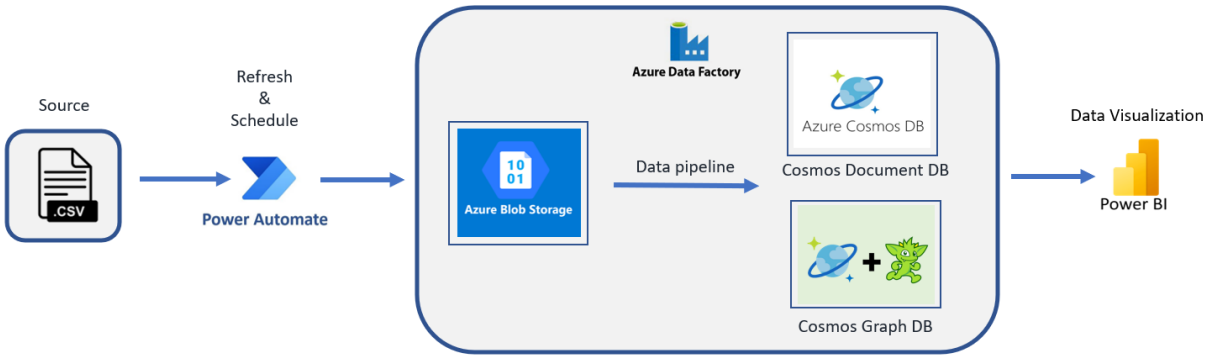
P2 - OVERVIEW OF PROJECT DATA ARCHITECTURE AND DATA MODELS

TOPIC: E-COMMERCE DATA ANALYSIS

TEAM: 3

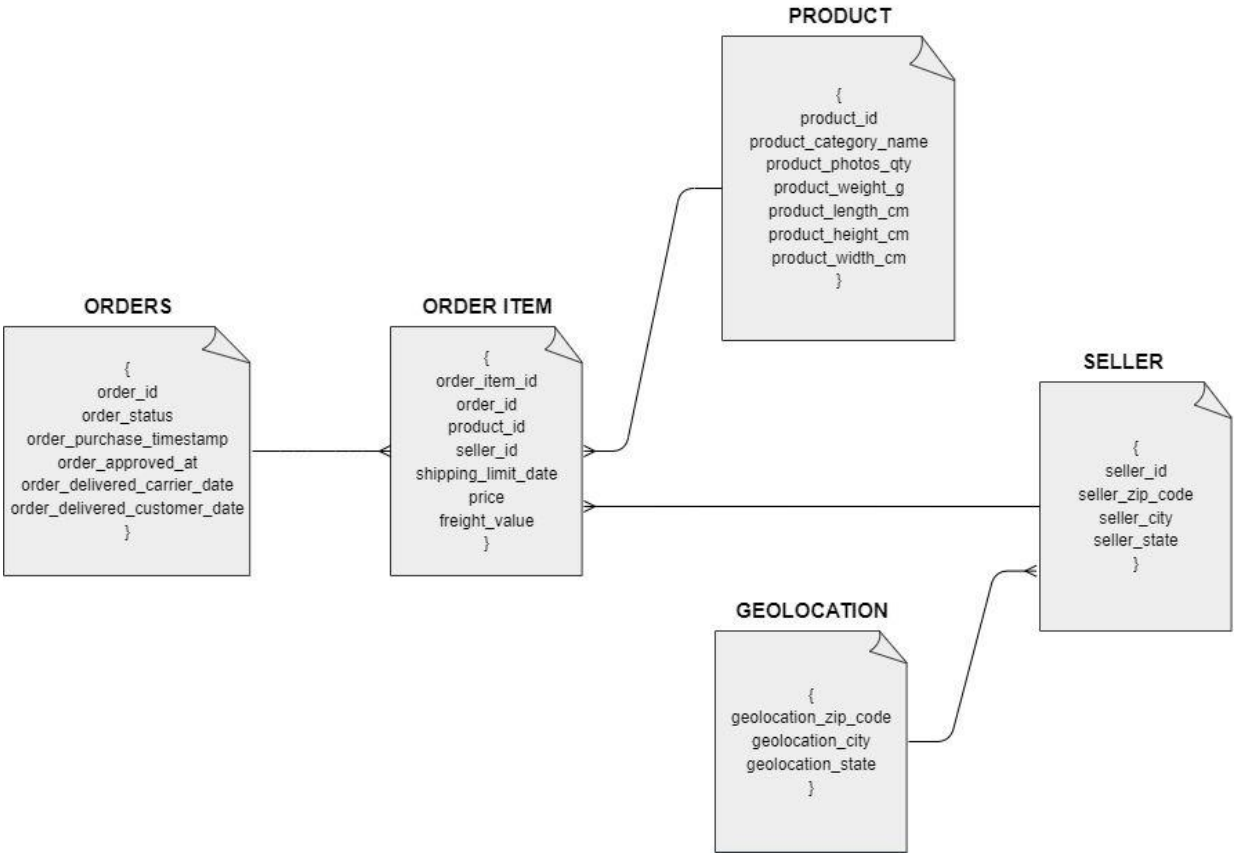
Abbas Furniturewala - NUID 002193272
Shubham Idekar - NUID 002776415
Shrutika Salian - NUID 002142365
Harshit Parikh - NUID 001044838

DATA ARCHITECTURE DIAGRAM:



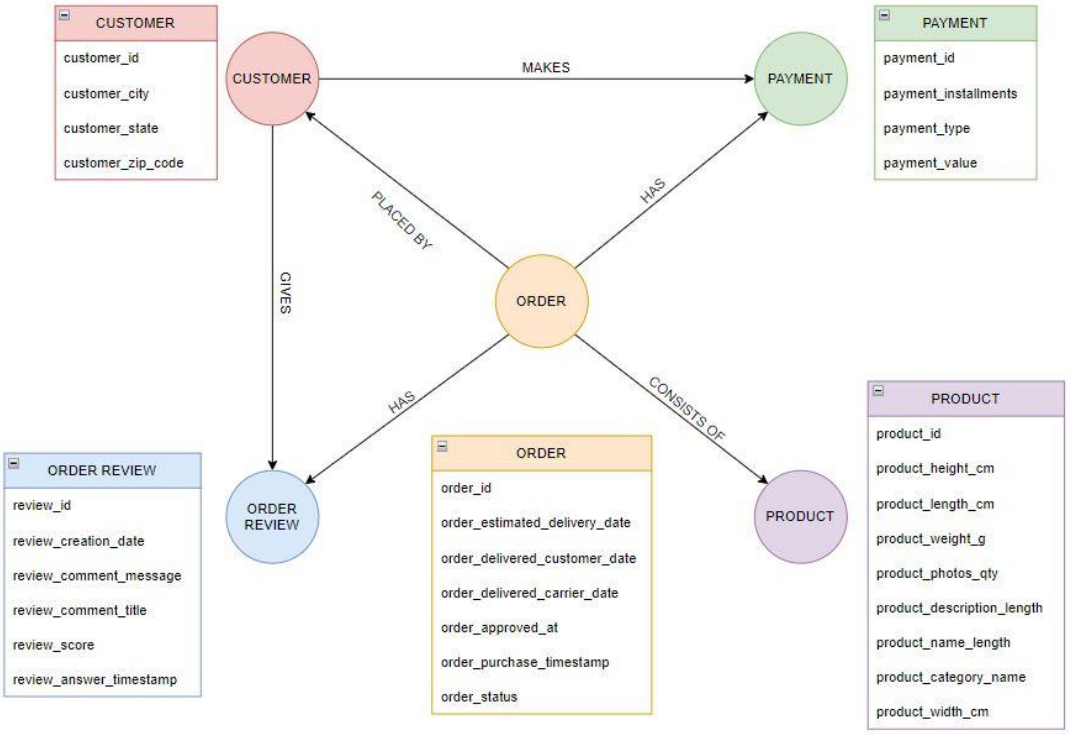
1. Data will be stored in a folder in the form of .csv files.
2. Leverage ***Power Automate*** flow to schedule a fetch and refresh of data from folder and place it in ***Azure Blob Storage***.
3. Build pipelines for data transformation and migrate data from Azure Blob Storage to ***Cosmos Document DB*** using NoSQL API and in ***Cosmos Graph DB*** using Gremlin API using ***Azure Data Factory***.
4. Visualize and create interactive dashboards on ***PowerBI*** for e-commerce data analysis, gaining insights of Brazil e-commerce market trends.

DOCUMENT DATA MODEL



1. The above ER diagram explains our Document Data model having 5 documents respective of Orders, Order Item, Geolocation, Seller and Products to create flexible schema for necessary business requirements.
2. We can use it such that the platform can support rich querying capabilities, such as searching for products by keyword, filtering by product attributes, and sorting by most sold.
3. This model will help us determine reports on sales trends, popular products, and top-performing sellers, helping to guide business decisions and marketing strategies.

GRAPH DATA MODEL



1. The above graph explains our graph data model for Cosmos graph DB, having 5 nodes which describe our 5 entities: Customer, Order, Product, Order Review, and Payment.
2. Node Order is at the center of our graph data model as it is the center and link between all our other nodes: Customer, Product, Order Review and Payment
3. Relationships exist between nodes and are as follows:
    - Order is placed by a Customer
    - Order consists of Product
    - Customer makes Payment
    - Customer gives Order Review
    - Order has Payment
    - Order has Order Review