Data Modeling and Architecture

Agenda

- Problem Definition
- Learning Journey and Outcomes
- Data Pipeline and Architecture
- Data Understanding
- Data Definition
- Stored Procedures and Functions
- Views
- Quality Checks
- Summary

Problem Definition

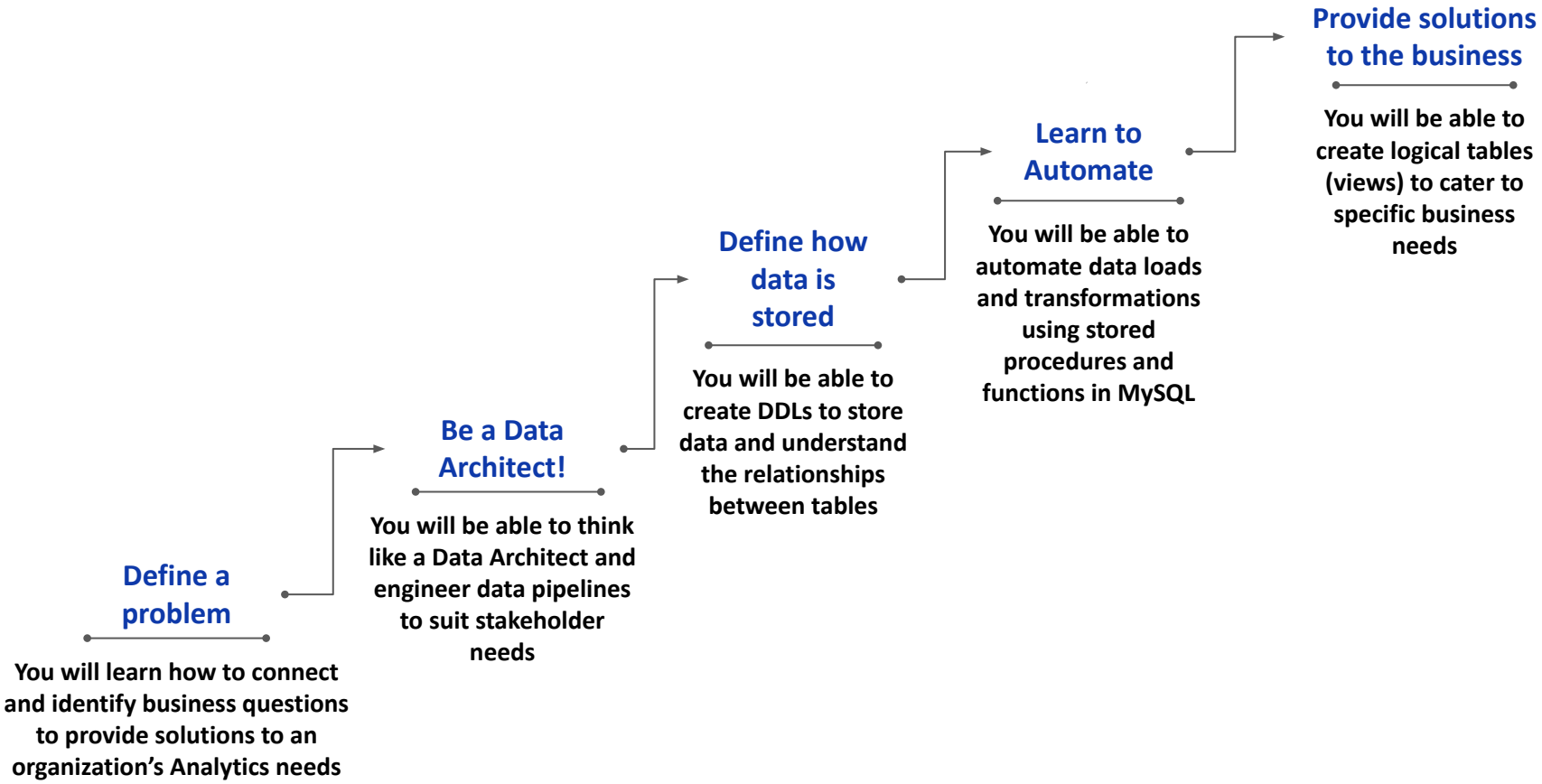
Business Context:

GL-Eats is a food delivery app that delivers orders in ~100 cities across multiple countries. Over the last few weeks, their revenue has dipped, and multiple Teams in GL-eats like Marketing, Delivery Operations, and Revenue Management are rushing to find answers to a lot of questions to solve this problem.

Our Objective:

As part of the centralized **Data and Analytics Team**, our goal is to **Architect a data pipeline in MySQL** to cater to all the questions asked by the various teams, by creating and maintaining a database through which we can ingest our data generated every week and finally create usable tables which can be readily queried by the business stakeholders to find answers to their questions.

Learning Journey and Outcomes



Here are the business questions asked...

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to **Restaurants**

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to Delivery

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to Delivery and Restaurants

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to Delivery and Orders

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to Orders and Restaurants

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

Do high orders come only from highly rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

What is the cancellation rate of a customer?

What is the WoW% of Revenue?

Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

Questions related to Orders and Customers

How many Orders do we get each week? What is the trend across weeks?

Which Cities are our Vendor Restaurants located? What is the distribution?

Are there specific cities that have a high rate of cancellation?

Which are my high revenue generating restaurants?

Are our Delivery Rating decreasing across weeks?

What is the percentage of customers not satisfied with the service?

What is the average number of Orders an Employee delivers?

Are we giving higher Discounts for higher Order Costs?

What is the average Order Cost each week? Is it trending upwards or downwards?

How many Employees are overworked and delivering above the acceptable limit?

What the distribution of Customers by Order Frequency across?

Do Customers who give a low Delivery Rating in the previous week, tend to order lesser next week?

Are cancellations more likely from low rated restaurants?

What is the distribution of Restaurants across Cuisines?

Do Customers who get a higher Discount tend to give a higher Delivery Rating?

How many Customers have used our services so far, across Countries?

Do high orders come only from highly rated restaurants?

What the average number of visits to a restaurant by an employee?

What is the average rating of a customer?

What is the average Order Quantity each week? Is it trending upwards or downwards?

How has employee performance been by week?

What is the distribution of employees in a locality?

What is the cancellation rate by employee?

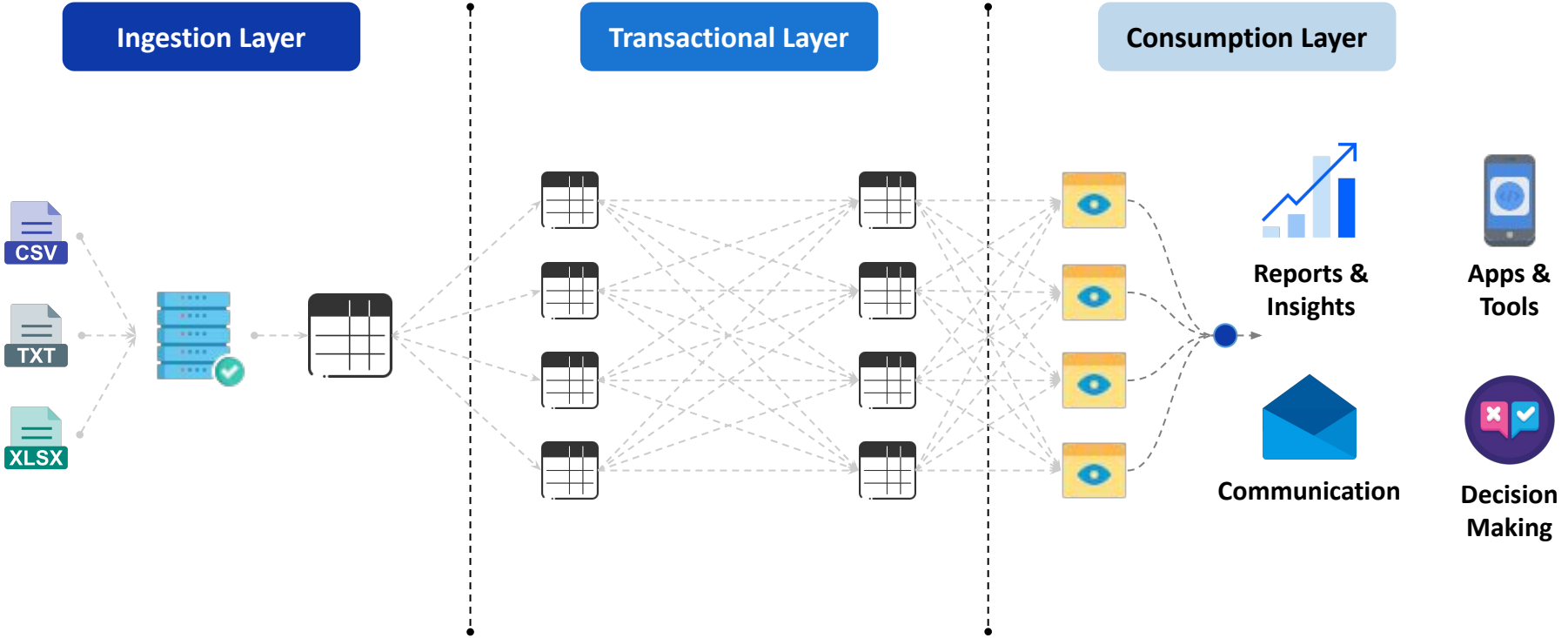
What is the cancellation rate of a customer?

What is the WoW% of Revenue?

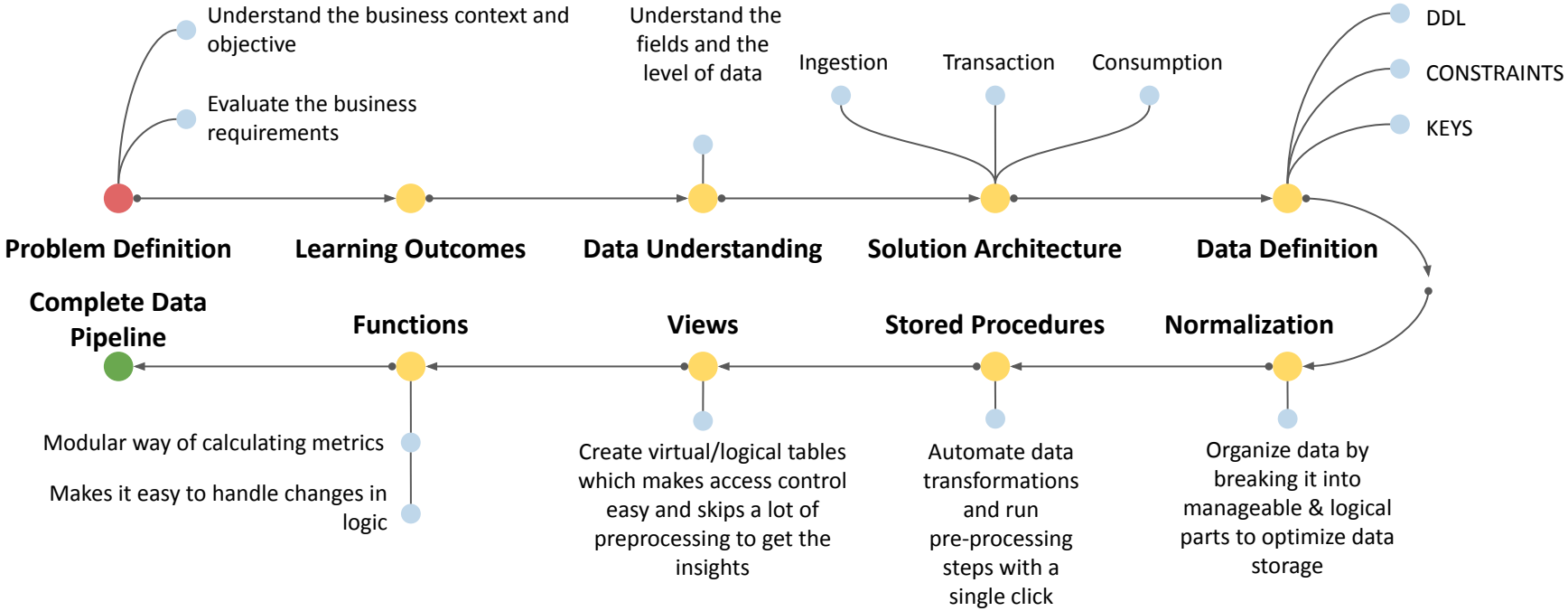
Is there a correlation between restaurant ratings and delivery ratings?

Who are the high and low performing employees?

What does a typical Data Pipeline look like?



Summary


Understand the business context and objective
Evaluate the business requirements
Understand the fields and the level of data
Ingestion
Transaction
Consumption
DDL
CONSTRAINTS
KEYS
Problem Definition
Learning Outcomes
Data Understanding
Solution Architecture
Data Definition
Complete Data Pipeline
Functions
Views
Stored Procedures
Normalization
Modular way of calculating metrics
Makes it easy to handle changes in logic
Create virtual/logical tables which makes access control easy and skips a lot of preprocessing to get the insights
Automate data transformations and run pre-processing steps with a single click
Organize data by breaking it into manageable & logical parts to optimize data storage

Great Learning

Happy Learning !