



Southern Alberta  
Institute of Technology



DESIGN DOCUMENT

# **AMAZON SALES ANALYSIS**

Presented By  
**SAGAR SINGH**

Under Guidance  
**DR. JUNAID QAZI**

## Introduction

### Purpose

Discuss the aim to understand sales performance, product preferences, and geographical sales distribution to inform future business strategies for amazon and understand trends from the dataset.

### Scope

Collection, processing, and visualization of Amazon Sales Data, analysis and visualization of sales data by category, geographical distribution, and revenue trends.

## System Overview

### System Architecture

The system architecture will consist of the following components:

- Data Sourcing: Kaggle dataset ["Amazon Sales Dataset"](#)
- Data Modelling: Entity-Relationship Diagram (ERD) and Star schema design
- Visualization: Power BI for interactive dashboard development

### System Components

The major components of the system are:

- User Interface: Dashboard for data analysis

## Data Management

### Data Sourcing

The data source for this project is the Kaggle dataset ["Amazon Sales Data"](#).

### ETL Process

The ETL process will include the following steps:

- Extract: Download the dataset from Kaggle and load it as CSV file into a Power BI
- Transform: Cleaning the data, handling missing values and removing duplicates from various columns
- Load: Load the transformed data into a Power BI

## Data Modelling

### Entity-Relationship Diagram

The ERD will consist of the following entities:

- Amazon Sales Report
- Category lookup
- Location Lookup
- Order lookup
- Shipment lookup

### Schema Design

The database schema will consist of the following tables:

Based on the diagram in the image, the database schema appears to be designed for an e-commerce or sales reporting system involving Amazon sales data. Here's a star schema similar to the one you've described:

#### Fact Table

Sales\_Facts

- ASIN
- Shipment\_ID
- Location\_ID
- Amount
- B2B
- Cost
- Currency
- Date
- Order\_ID
- Promotion\_IDs
- Cancellation Rate
- Cancelled Orders
- Completed Orders
- Net Profit
- Profit
- ROI
- Sales Conversion Rate
- Total Cost
- Total Orders
- Total Revenue

### **Dimension Tables**

- Category Lookup
- ASIN (Primary Key)
- Category
- Quantity
- Size

### **Shipment Lookup**

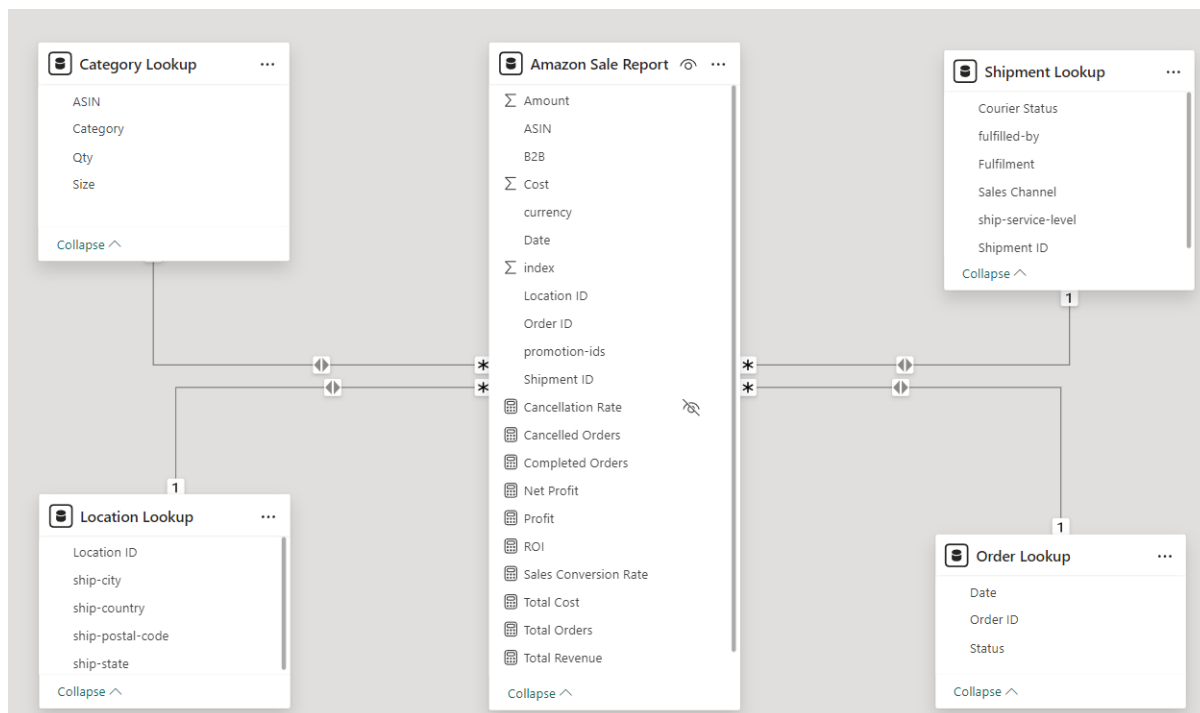
- Shipment\_ID (Primary Key)
- Courier Status
- Fulfilled\_by
- Fulfilment
- Sales Channel
- Ship\_Service\_Level

### **Location Lookup**

- Location\_ID (Primary Key)
- Ship\_City
- Ship\_Country
- Ship\_Postal\_Code
- Ship\_State

### **Order Lookup**

- Order\_ID (Primary Key)
- Date
- Status

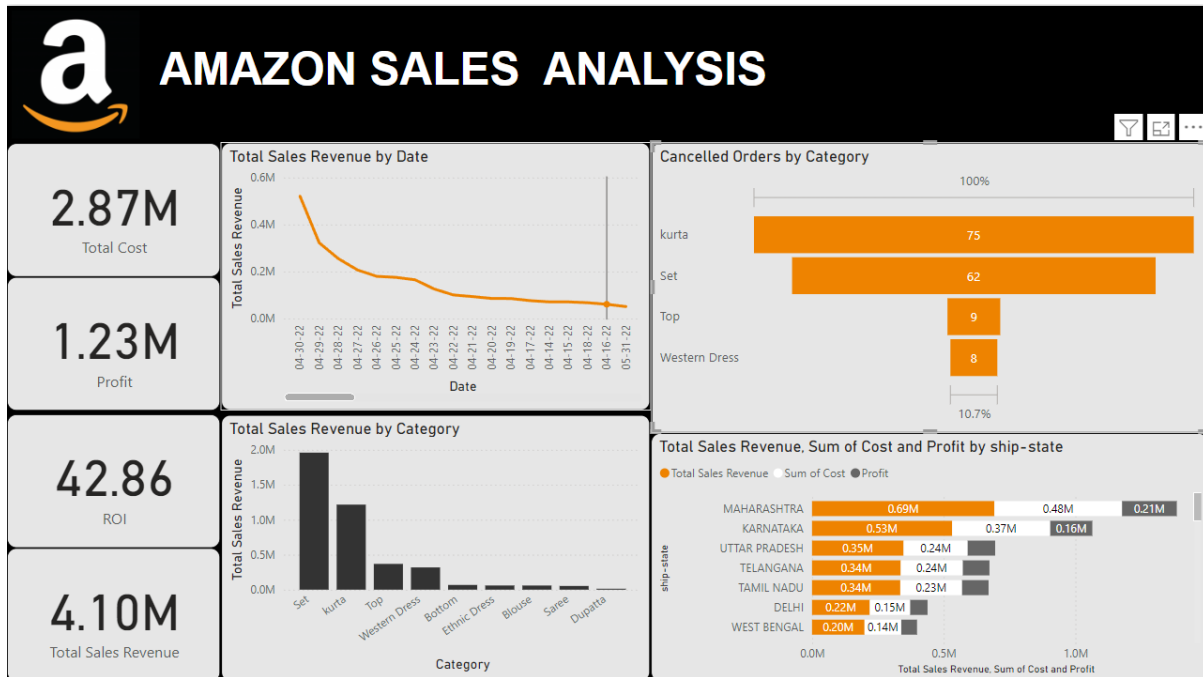


## Visualization & User Interface

### Layout

The dashboard will consist of the following sections:

- Introduction: Overview of the project and dataset
- Describes the layout of the dashboard with sections for total revenue, ROI by category, and sales by state.
- Users can filter data by date, category, or state.
- Data Visualization Tool: Power BI



## Features & Functionalities

The dashboard will offer the following features and functionalities:

- Interactive Elements: Filtering & sorting
- Navigation: Easy navigation between visualizations

## Conclusion

Describes the structure and extent of the Amazon Sales data. Give users access to an informative and interactive dashboard so they can explore and analyze it.