

Logistics Optimization for Delivery Routes – Amazon

Project Overview:

Amazon, a global e-commerce leader, handles millions of daily shipments across various regions. The logistics network is vast, comprising warehouses, fulfillment centers, and last-mile delivery partners. With increasing order volumes, delays and inefficiencies in delivery routes can significantly affect customer satisfaction and operational costs.

Currently, Amazon's logistics team faces challenges in:

- Identifying delays and their root causes (e.g., route congestion, warehouse delays).
- Optimizing delivery routes for faster fulfillment.
- Improving shipment efficiency through data-driven insights.

The logistics data, stored in structured relational databases, can be analyzed using SQL to extract actionable insights that can improve route planning, minimize delivery delays, and enhance overall operational performance.

Project Objective:

Build a SQL-driven Logistics analytics system to analyze delays, optimize routes, and enhance shipment efficiency by leveraging queries, aggregations. The project aims to answer key business questions, uncover inefficiencies, and recommend actionable improvements based on data analysis.

(If you're unfamiliar with any concepts or domain-specific terms, feel free to take the help of Google or AI tools like ChatGPT or Gemini to understand them better and apply them effectively in your project.)

Dataset Description

The dataset will include the following key tables:

(Please download the hyperlinked datasets)

1. Orders Table

Column Name	Description
Order_ID	Unique identifier for each order.
Customer_ID	Unique identifier for the customer placing the order.
Warehouse_ID	Identifier of the warehouse fulfilling the order.
Route_ID	Identifier for the delivery route taken for the order.
Order_Date	Date when the order was placed.
Expected_Delivery_Date	Estimated date of delivery as per SLA (Service Level Agreement).
Actual_Delivery_Date	Actual date when the delivery was completed.
Delivery_Status	Status of delivery – whether it was “On Time” or “Delayed”.

2. Routes Table

Column Name	Description
Route_ID	Unique identifier for each delivery route.
Start_Location	Starting location of the delivery route (typically a warehouse or fulfillment center).
End_Location	Final destination of the delivery route (city/region).
Distance_KM	Total distance of the route in kilometers.
Average_Travel_Time_Min	Average travel time for the route in minutes (without delays).
Traffic_Delay_Min	Additional time (in minutes) caused by traffic conditions.
Delivery_Status	Status of delivery – whether it was “On Time” or “Delayed”.

3. Warehouses Table

Column Name	Description
Warehouse_ID	Unique identifier for each warehouse.
Location	City or location of the warehouse.
Processing_Time_Min	Average time (in minutes) taken to process and prepare shipments for dispatch.
Dispatch_Time	Average time of the day when shipments are dispatched from the warehouse.
Traffic_Delay_Min	Additional time (in minutes) caused by traffic conditions.
Delivery_Status	Status of delivery – whether it was “On Time” or “Delayed”.

4. Delivery Agents Table

Column Name	Description
Agent_ID	Unique identifier for each delivery agent.
Route_ID	Route assigned to the delivery agent.
Shift_Hours	Number of hours in the agent's working shift.
Avg_Speed_KM_HR	Average delivery speed of the agent in kilometers per hour.
On_Time_Percentage	Percentage of deliveries made on or before the expected delivery time.

5. Shipment Tracking Table

Column Name	Description
Shipment_ID	Unique identifier for each shipment checkpoint record.
Order_ID	Identifier of the order being tracked.
Checkpoint	Name or number of the checkpoint (e.g., warehouse, in-transit hub).
Checkpoint_Time	Timestamp when the shipment crossed the checkpoint.
Delay_Reason	Reason for delay at this checkpoint (e.g., Traffic, Weather, Sorting Delay, None).

Tasks to be Performed

(Please refrain from using AI to perform the tasks mentioned below, as it will only provide generic solutions.)

Task 1: Data Cleaning & Preparation (10 Marks)

- Identify and delete duplicate Order_ID records.
- Replace null Traffic_Delay_Min with the average delay for that route.
- Convert all date columns into YYYY-MM-DD format using SQL functions.
- Ensure that no Actual_Delivery_Date is before Order_Date (flag such records).

Task 2: Delivery Delay Analysis (15 Marks)

- Calculate delivery delay (in days) for each order
- Find Top 10 delayed routes based on average delay days.
- Use window functions to rank all orders by delay within each warehouse.

Task 3: Route Optimization Insights (10 Marks)

- For each route, calculate:
 - Average delivery time (in days).
 - Average traffic delay.

- Distance-to-time efficiency ratio: Distance_KM / Average_Travel_Time_Min.
- Identify 3 routes with the worst efficiency ratio.
- Find routes with >20% delayed shipments.
- Recommend potential routes for optimization.

Task 4: Warehouse Performance (10 Marks)

- Find the top 3 warehouses with the highest average processing time.
- Calculate total vs. delayed shipments for each warehouse.
- Use CTEs to find bottleneck warehouses where processing time > global average.
- Rank warehouses based on on-time delivery percentage.

Task 5: Delivery Agent Performance (10 Marks)

- Rank agents (per route) by on-time delivery percentage
- Find agents with on-time % < 80%.
- Compare average speed of top 5 vs bottom 5 agents using subqueries.

Task 6: Shipment Tracking Analytics (15 Marks)

- For each order, list the last checkpoint and time.
- Find the most common delay reasons (excluding None).
- Identify orders with >2 delayed checkpoints

Task 7: Advanced KPI Reporting (10 Marks)

Calculate KPIs using SQL queries:

Average Delivery Delay per Region (Start_Location).

On-Time Delivery % = (Total On-Time Deliveries / Total Deliveries) * 100.

Average Traffic Delay per Route.

Task 8: PPT Presentation and Video Submission(20 Marks)

PPT Submission (10 Marks)

- Present your analysis and findings by copying all the queries and result tables from the previous steps into a PowerPoint presentation.
- Copy and Paste SQL queries and its corresponding results for the tasks 1 to 7.
- Ensure that the tables are formatted clearly, and the queries are concise. Use charts, graphs, or tables to make your data more digestible.
- Record a video explaining the project for a maximum of 5 mins. This should include project understanding analysis and explanation of outcomes. Upload in drive and share the drivelink in the ppt.(10 marks)

(The summary should be in your own words and must not be generated using AI. Please don't write a script and read it aloud. Also, screenshare and show the key findings. Marks will be deducted for failing to do so.)

Note:

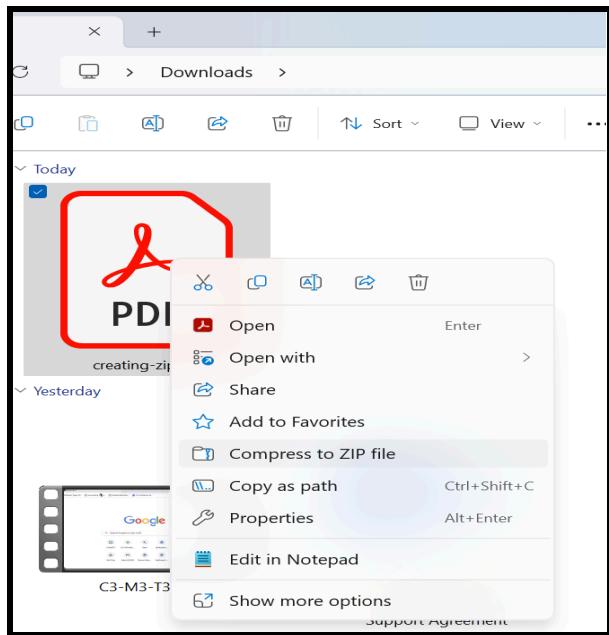
- Submit your PPT and SQL script after zipping those in a folder.
- Plagiarism will result in a penalty, including possible project disqualification.
- The project will be evaluated based on the quality of analysis and visualizations, depth of insights, feasibility of recommendations, clarity of explanations, and adherence to instructions and deliverables.
- If the student sets their own criteria, they need to clearly mention and explain it. Marks will be given according to the specified criteria if they are acceptable.
- Remember to keep the video length less than 5 minutes with your face clearly visible.

Submission Guidelines:

- Save the PPT, SQL Script in a folder and then convert it into a zipped (.zip) folder. **(Please note, the drivelink for the video created should also be added in the Ppt itself.)**
- Upload the zipped folder on your respective dashboard.
- Failure to comply with submission guidelines will result in no grading/0 marks.

How to ZIP a PDF file:

- Put all of the documents you want to compress (or just one) into a new folder.
- Right click on that folder.



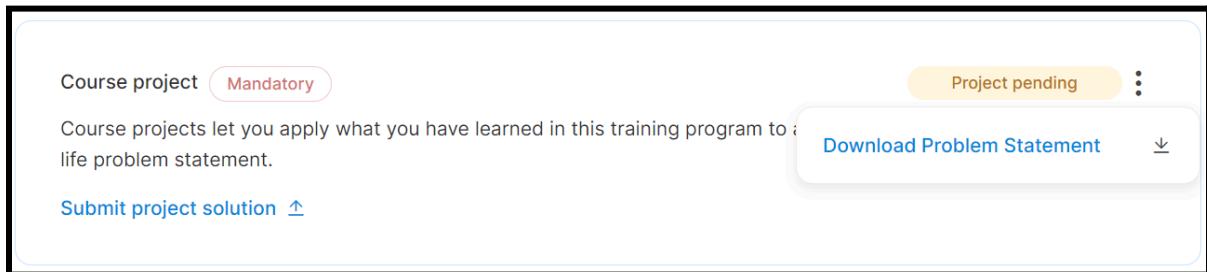
- Select the “Compress to ZIP file” option and then click “Compressed (Zipped) folder.”
- A new .ZIP file will be created that contains your document(s).

In order to submit the projects please follow the following steps:

1. Click on “ Your progress [View details](#)” after logging into your dashboard.

A screenshot of a digital marketing course dashboard. At the top, it says 'Welcome Ayush' and 'Digital Marketing Course' with a placement guarantee deadline of '29 Sep, 2024'. Below this, there are three progress metrics: 'Your progress' (17%, 'Keep going!'), 'Current score' (77%, 'Good work. You are on track to graduate'), and 'Session Attendance' (76%, 'Amazing, you're a pro. Keep it up!'). Under the 'Ongoing' section, it lists 'Course 3 : Search Engine Marketing' and 'Module 4 : Display Ads - GDNGoogle Display Network'. It also shows a 'Continue Learning >' button. On the right, there's a 'Sessions' section with 'View all Sessions' and 'You have 0 sessions this week'. At the bottom, it shows 'Today (4 Sep, 2024)' and '0 Sessions'.

2. Next, click on the tab for the specific child course for which you want to download the problem statement. Then, scroll down to find the "**Course Project**" section.
3. Now, click on the three dots on the right-hand side of the "Course Project" tab to select "**Download Problem Statement**".



4. Please follow the guidelines (screenshot is shared below) provided in the project to ensure correct submissions. Then, click on "**Upload Project Solution**" to submit your work.

