**Data Analytics May Placement Project**

***Part A-The SQL Analysis case study-***

1. **Identify Peak Purchase Times and Their Impact on Delays:** This query determines the peak times for ticket purchases and analyzes if there is any correlation with journey delays.
2. **Analyze Journey Patterns of Frequent Travelers:** This query identifies frequent travelers (those who made more than three purchases) and analyzes their most common journey patterns.
3. **Revenue Loss Due to Delays with Refund Requests:** This query calculates the total revenue loss due to delayed journeys for which a refund request was made.
4. **Impact of Railcards on Ticket Prices and Journey Delays:** This query analyzes the average ticket price and delay rate for journeys purchased with and without railcards.
5. **Journey Performance by Departure and Arrival Stations:** This query evaluates the performance of journeys by calculating the average delay time for each pair of departure and arrival stations.
6. **Revenue and Delay Analysis by Railcard and Station**

This query combines revenue analysis with delay statistics, providing insights into journeys' performance and revenue impact involving different railcards and stations.

1. **Journey Delay Impact Analysis by Hour of Day**

This query analyzes how delays vary across different hours of the day, calculating the average delay in minutes for each hour and identifying the peak hours for delays.

**Important Notes-**

1. **Create a function for data cleaning if required**
2. **Use bulk insert (in MSSQL) /load\_into\_file (in MySQL) to upload the data.**
3. **Write the interpretation of every query result after the query in SQL.**

**Part B-For Visual reports in Power BI/Tableau**

1. **Customer Loyalty and Retention Analysis**

Analyze customer loyalty and retention by identifying repeat customers, measuring their purchase frequency, and evaluating their impact on revenue over time.

1. **Impact of Journey Time on Customer Satisfaction**

Analyze how the actual journey time (from departure to actual arrival) impacts customer satisfaction, assuming delays reduce satisfaction levels (where satisfaction decreases as delay increases (e.g., -1 satisfaction point per 10 minutes of delay).

1. **Profitability Analysis Based on Ticket Type and Class**

Perform a profitability analysis based on ticket type and class, with dynamic filtering to show insights for different periods or stations (consider 55% of the ticket price is the cost of a ticket.

1. **Frequent Traveler Analysis with Dynamic Segmentation**

Identify and analyze frequent travelers dynamically segmented by the number of journeys they made.

1. **Analyzing Delays to Optimize Train Schedule**

Optimize train schedules to minimize delays. To achieve this, you need to analyze the average delay for trains by the hour of the day, considering that delays can vary based on the time of departure.

**Instructions**

* The dataset will be given through a drive link mentioned below
* You have to submit the project by **Sunday 16th June 2024 by EOD.**
* Kindly submit your ‘XYZ.sql/xyz.docx’ file and ‘XYZ.PowerBI/XYZ.Tableu’ to sagar[@learnbay.co](mailto:ashishsom@learnbay.co) within the timeframe, submission of the project after the due date will be considered disqualified. Late submission will be considered with a valid reason.
* After submission of the project you’ll get a link to book a time for the project presentation on **17th June 2024**
* If you missed your date of presentation, you will get the chance to present this current project in the next project slot.

**Selection Criteria/feedback**

* Selection of candidates will be based on their approach to working **on the Project,  presentation skills(Storytelling skills), and subject knowledge points(a mock round)(Questions related to SQL & Power BI/Tableau).**
* **Note: you need to score 80% to clear this round.**
* Once the presentation is done every candidate will get their feedback during the session and outcome and score via mail with the status of whether they are selected or not.
* Selected candidates’ data will be shared with the placement team for 1 on 1 resume session.
* Candidates who are not selected in this process will be carried forward to the next project.
* Kindly do not book multiple slots, if found it shall consider as cancelled. If any change in the slot date and time kindly inform or cancel the previous slot.
* If you are absent or unable to present on the day of the presentation, in that case getting another will be subject to availability.