**SQL DATA ANALYSIS COURSE DATA DOCUMENTATION**

**DATA UNDERSTANDING AND OBSERVATION**

The dataset is from an SQL data analysis course, in which the data was collated based on the students who registered for the course. My first observation about this dataset is that it includes data on the people who registered for this course over the years. Each column in the dataset represents their ID, start time, completion time, Name, Country/Location, Age group, Occupation/Industry, SQL Experience Level, Preferred learning style, Preferred SQL Database, Availability (Days), Availability (Time), and their Topic of Interest.

**Project Objective**

The primary objective of the SQL for Data Analysis course is to equip participants with practical skills in utilizing SQL queries to analyze and derive insights from real-world datasets. Through hands-on experience, participants will develop proficiency in data manipulation, querying, and exploration using SQL. This course aims to prepare individuals for roles in data analysis across diverse industries.

**Data Sources**

SQL Data Analysis Course Data: The main dataset utilized for this analysis is the "SQL-Course-Data.xlsx," which comprises comprehensive information on students who registered for the course.

**Tools**

Excel - Data Cleaning and Data Analysis

PowerBI - Creating reports

**Data Cleaning/Preparation**

* Loading and inspecting the data.
* Used Power Query for my cleaning process
* Checked for duplicates and none was found.
* The data consists of 13 columns and 1617 row counts including the header.
* There are blanks in the topic of Interest column, which has 85% empty cells and 15% Valid, I dropped the column so it doesn’t affect my analysis.

**QUESTIONS BASED ON MY ANALYSIS**

After reviewing the dataset, these questions were considered for my analysis:

1. **Demographic Analysis:**

* How do student age groups vary?
* Which countries/locations are most represented?
* What are the main student occupations?
* What learning styles do students prefer?

1. **Enrollment Patterns:**

* Are there peak months/seasons for enrollment?

1. **Course Completion Rates:**

* What are the overall completion rates?

1. **SQL Experience Level:**

* How are students distributed by SQL experience?
* Are completion rates affected by experience level?

1. **Preferred SQL Database:**

* Which databases do students prefer?
* Are preferred databases linked to course performance?

1. **Time Availability:**

* What are common study times/days?
* Is there a study habit correlation with performance?

**RESULTS**

Results from the analysis of the SQL data analysis course dataset provide valuable insights into various aspects of student demographics, enrollment patterns, course completion rates, SQL experience levels, and time availability:

**Demographic Analysis:**

* The age group of 25-29 exhibits the highest enrollment numbers, indicating a significant interest among individuals in this age range.
* A total of 1616 students have enrolled in the course, showcasing a substantial level of participation.
* Nigeria emerges as the most common country represented among students, suggesting a regional interest in the course.
* The preferred learning styles among students include video tutorials, hands-on exercises, and live demonstrations, with 346 students opting for this approach.
* Beginners predominantly opt for live demonstrations, with approximately 278 responses indicating a preference for this learning style.
* The predominant occupation among enrolled students is "Data Analyst," indicating a strong interest from professionals in this field.

**Enrollment Pattern:**

* The peak period for enrollment occurs mostly during the holiday season, particularly in December, indicating a trend of increased interest and participation during this time.

**Course Completion:**

* The course boasts a remarkable completion rate of 100%, highlighting the effectiveness of the course content, structure, and instructor guidance in facilitating student success.

**SQL Experience Level**:

* The majority of students exhibit a beginner-level proficiency in SQL, with 1279 individuals falling into this category, followed by 327 intermediate-level students, and only 10 advanced-level students.
* MySQL emerges as the preferred database among students, indicating a preference for this platform in SQL-related tasks.

**Time Availability:**

* Evening hours are identified as the most available time for students to engage in course-related activities.
* Weekends emerge as the most available days for students, indicating a preference for dedicating time to learning during leisure periods**.**

**RECOMMENDATION**

* Develop course content that directly addresses the needs and preferences of the majority demographics, such as focusing on tasks relevant to "Data Analyst" occupations and incorporating preferred learning styles like video tutorials, hands-on exercises, and live demonstrations.
* Ensure comprehensive coverage of MySQL topics, as it is the preferred database among students.
* Focus marketing efforts towards the age group of 25-29, which represents the highest enrollment, and consider tailoring strategies to attract students from Nigeria, the most represented country.
* Capitalize on peak enrollment periods, particularly during December, by offering promotions or discounts to attract more participants.
* Provide strong support and resources for beginner students, who make up the majority of enrollments, to enhance their learning experience and increase retention rates.
* Schedule live sessions, workshops, or support sessions during evenings and weekends to accommodate the availability patterns of students, maximizing engagement and participation.
* Encourage students to provide feedback throughout the course and use this feedback to continually improve course content, delivery methods, and support mechanisms to better meet student needs and preferences.
* Continuously track course completion rates and analyze factors influencing completion, such as SQL experience level, to identify areas for improvement and intervention.