Library Database — Report.

1. Brief Database Description

The project implements a Library Management database named "library_db" with the following logical tables: Authors, Books, Members, Departments, LibraryStaff, Borrow_History, and BookOrders. The schema includes enum types for membership, gender, and fulfilment status and uses foreign keys to maintain referential integrity.

2. Visualization Insights (summary)

Top authors by number of books:
The bar chart highlights authors contributing the largest number of titles; these authors are likely central to the collection and acquisition strategy.
Member distribution by membership type:
The pie chart shows the proportion of students, standard, and premium members which can inform membership-targeted services.
Monthly borrowing trends:
The line chart reveals seasonal patterns and months with higher borrowing volume which can guide staffing and promotions.
Available copies by genre:
The horizontal bar chart exposes genres with low availability where additional copies might be needed.
Book orders by supplier & fulfilment status:
The heatmap/stacked bar identifies top suppliers and how quickly orders are fulfilled.

3. Challenges Faced and Solutions.

While creating and editing this project the challenges faced are:
i. Creating advanced queries and to solve this postgreSql documentation was consulted and some other documentation.
ii. Had challenges running the pandas equivalent of the queries and the visualizations on a .py file due to inability of powershell to run the dependency libraries needed. So I reverted to using an .ipynb file so as to be able to run each cells of the notebook on the dependency libraries needed in the activated python environment and as well generated outputs for each cell and save visuals of each visualization in png format.

4. SQL vs Pandas comparison

SQL is well-suited for set-based aggregation, leverages indexes, and runs close to the data. It is reproducible in DB clients and scales well. Pandas offers flexibility for complex transformations, tighter integration with plotting libraries, and a convenient environment for exploratory analysis when the dataset fits into memory.