Apnaklub Assignment - Data Analyst (Intern)

[Apnaklub](#) is a B2B wholesale platform empowering India's retailers and kiranas by connecting them directly to brands and wholesalers. We provide a seamless, tech-driven solution for small retailers to source inventory at competitive prices, helping them optimise their supply chain and improve profitability. Our mission is to simplify wholesale distribution and create a stronger, more inclusive retail ecosystem across the country.

Your Task:

- Analyse the dataset and generate insights that you believe would be valuable for the business
- Feel free to explore customer behaviour, sales trends, SKU performance, or any other patterns you discover
- Document the tasks you've chosen to undertake and explain why they are relevant
- Use Python and SQL for the analysis, providing code, comments, and clear explanations of your process
- Assume that if the same dataset is present in database, share few same queries that you would run to extract data
- Present your findings in a well-organised format, including visualisations where appropriate

Assessment Criteria:

- Problem formulation and reasoning for chosen tasks
- Proficiency in Python and SQL
- Quality and depth of insights
- Code efficiency and documentation
- Presentation of results (visuals and explanations)

Submission Docs:

- Python notebook
- PDF/PPT

About the data:

You are provided with a order-dataset containing the following columns:
- `user_id`: Unique identifier for a particular user
- `order_id`: Unique identifier for an order placed by a particular `user_id`
- `Order_date`: The date on which the order was placed
- `sku_id:` Unique identifier for an product which was in a particular order
- `quantity:` Quantity of the `sku_id` that was placed in the `order_id`
- `placed_gmv:` Total value of of the `sku_id` that was purchase (Note, this is not selling price, but rather (selling price x quantity)