A blue and white logo

AI-generated content may be incorrect.Data Science Assignment

Internship Program 2025

Prepared by: Kiran Pokharel

Date: 31.03.2025

# First Task

Below are the SQL queries I have created using the tables below for Users, Sales and Purchases.

## Query 1: Number of users registered in Finland in the last 30 days

SELECT COUNT(\*) AS users\_registered\_last\_30\_days

FROM Users

WHERE User\_country = 'Finland'

AND User\_registration\_timestamp\_utc >= (CURRENT\_DATE - INTERVAL '30 DAYS');

## Query 2: Number of users with at least one purchase in the past 30 days with more than one unique product

SELECT COUNT(DISTINCT s.User\_id) AS users\_with\_multiple\_products

FROM Sales s

JOIN Purchases p ON s.Purchase\_id = p.Purchase\_id

WHERE s.Timestamp\_utc >= (CURRENT\_DATE - INTERVAL '30 DAYS')

GROUP BY s.User\_id

HAVING COUNT(DISTINCT p.Product\_id) > 1;

## Query 3: Retrieve the most recent price for each purchased product

SELECT DISTINCT ON (Product\_id)

Product\_id,

Price,

Timestamp\_utc AS last\_purchase\_date

FROM Purchases p

JOIN Sales s ON p.Purchase\_id = s.Purchase\_id

ORDER BY Product\_id, Timestamp\_utc DESC;

# Second Task (Data Analysis: Key Insights & Recommended Metrics)

I conducted this data analysis assignment using Noppe, an interactive Python notebook environment. I utilized the pandas library for data manipulation and analysis. Visualizations were generated using matplotlib and seaborn, and images were saved directly from the notebook using the savefig function, for high-quality visuals for presentation.

**Key Insights from Visualizations:**

## User Purchase Behaviour:

As shown in figure 1 below, the distribution plot of user purchases clearly shows most users making few or no purchases, underscoring the importance of enhancing user activation and ongoing engagement strategies.

A graph of a distribution

AI-generated content may be incorrect.

Figure Distribution of Total Purchase per User

## 2.2 Geographic Focus:

The bar chart shown in figure 2 below visualizes user registration by country. It highlights significant user activity in Finland and other Nordic countries and presents clear opportunities for targeted regional marketing initiatives.

A graph of blue bars

AI-generated content may be incorrect.

Figure Top 10 Countries by Registration

## 2.3 Device Preferences:

A bar chart illustrating preferred devices as shown in figure 3 reveals a strong preference among users for mobile devices, primarily iOS and Android, emphasizing the need to prioritize mobile platforms in development and marketing.

A bar graph with text

AI-generated content may be incorrect.

Figure Preferred Devices for Purchases

## 2.4 Spending Patterns:

The histogram of spending patterns shown in figure 4 demonstrates that most users have low to moderate spending levels, pointing towards the potential of increasing average order values through targeted promotions and personalized offers.

A graph of a distribution of total spending

AI-generated content may be incorrect.

Figure Distribution of total spending

## 2.5 Recommended Metrics

I have made a list of business metrics that are recommended to monitor regularly. These metrics are chosen to help understand and improve business performance by providing clear, measurable insights into customer behavior and business operations. They are designed to help a company like Wolt evaluate the effectiveness of their strategies related to user acquisition, retention, customer spending behavior, and device preferences. Each metric is briefly explained to indicate its importance and how it can guide strategic decision-making.

**User Activation Rate:**

Measure the percentage of users who complete their first purchase post-registration to evaluate onboarding effectiveness.

**Customer Retention Rate:**

Track repeat purchasing behavior to gauge customer loyalty and long-term engagement.

**Average Order Value (AOV):**

Monitor spending per order to evaluate the success of sales strategies and promotions.

**Device Usage Distribution:**

Analyze the distribution of devices used for purchases to inform strategic decisions on platform investment and user experience improvements.