

ASAP IDEA FEST

Monthly Ideathon Series- DECEMBER 2025

1. Team Information

- **Team Name:** DataVerse
- **College/University/Organization:** makBig
- **Department/Stream:** Data Science

2. **Theme of the Month: Using Data Analytics to Improve E-Commerce Sales and User Experience**

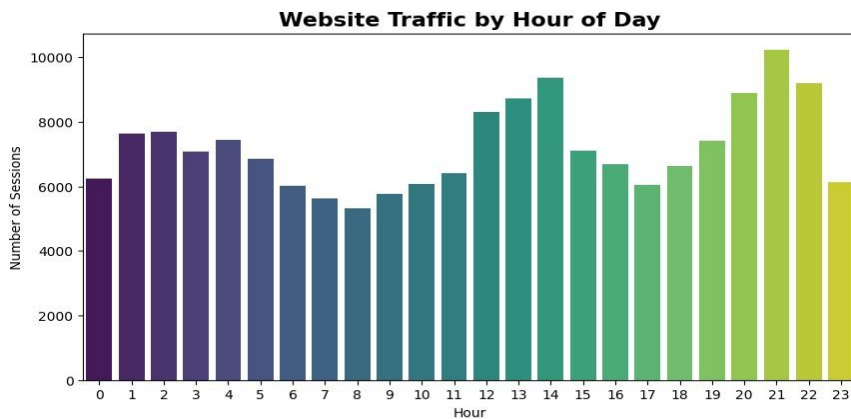
3. Problem statement : This dataset is an e-commerce website log created to help data analysts practice **Exploratory Data Analysis (EDA)** and **data visualization**.

The dataset contains information such as:

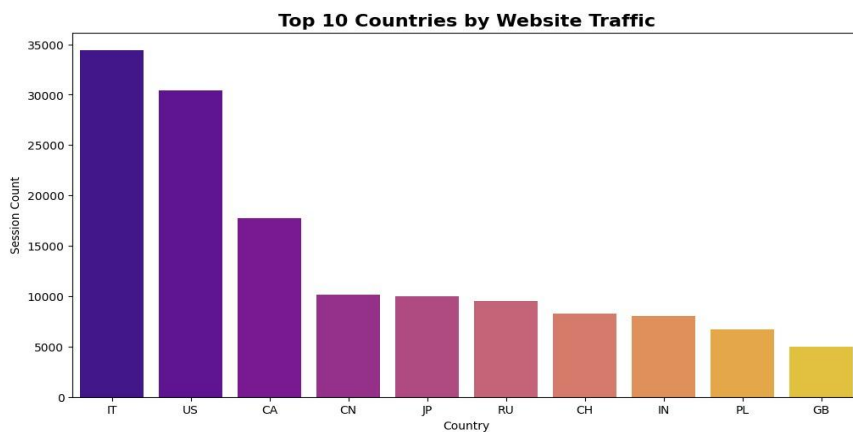
- When the website was accessed
- IP address of the visitor
- Country from which the website was accessed
- Language used on the website
- Amount of sales made by that IP address

Your task is to analyze this log data and generate meaningful insights about website traffic, user behavior, and sales patterns using EDA and data visualization techniques

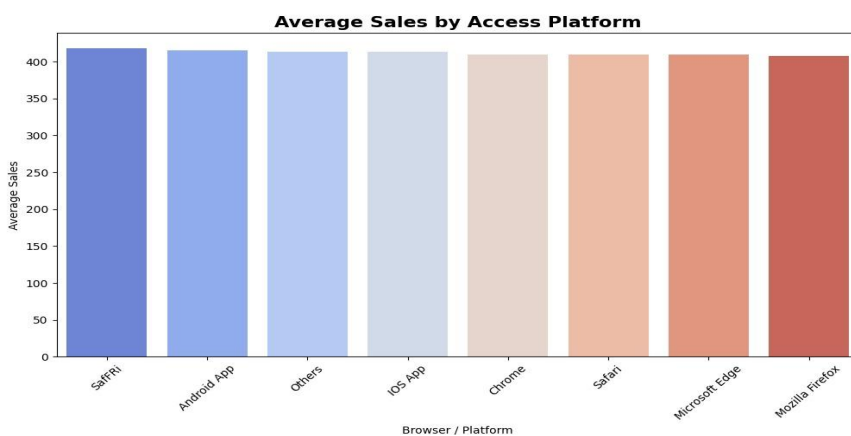
4. Solution



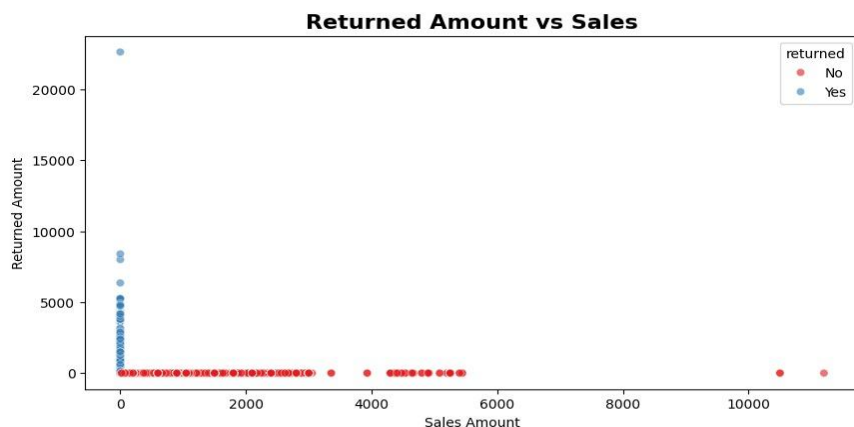
- Peak user activity occurs between 9 PM and 11 PM, with the highest traffic at 10 PM.
- Midday engagement rises around 1–2 PM, indicating lunchtime browsing behavior.
- Lowest activity appears between 5 AM and 8 AM, when users are generally inactive.
- Overall traffic gradually increases from morning to evening before reaching its peak at night.



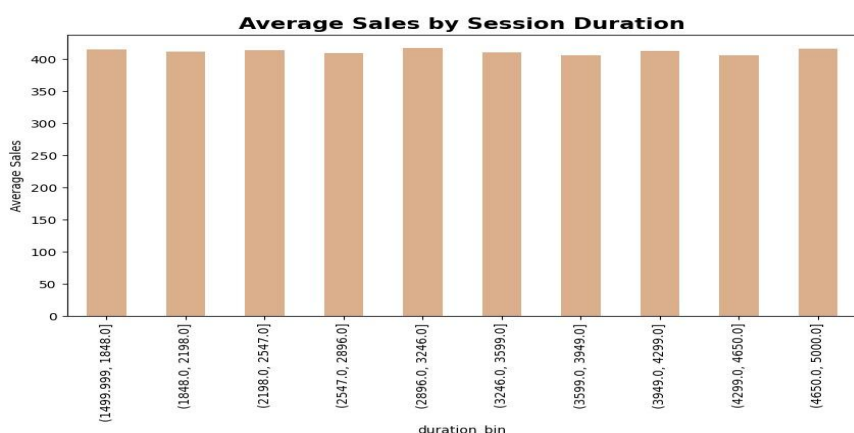
- Italy (IT) is the top traffic source, contributing the highest number of sessions (~34,000).
- The US and Canada follow, indicating strong engagement from North America.
- Countries like China (CN), Japan (JP), and Russia (RU) show moderate traffic levels.
- India (IN), Poland (PL), and the UK (GB) appear in the lower end of the top 10 list.
- Traffic distribution is highly uneven, with the top 3 countries contributing the majority of sessions.



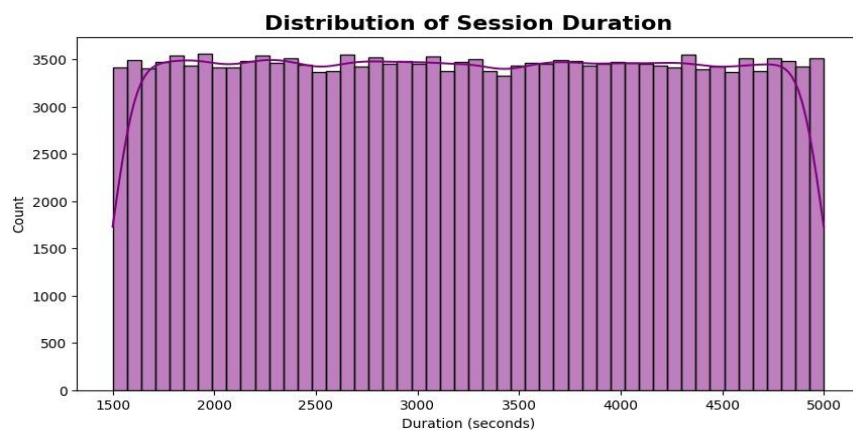
- Average sales are almost identical across all platforms (~410–420).
- Safari and the Android App show only slightly higher averages, but not significantly.
- Major platforms (Chrome, Safari, Edge, Firefox) perform nearly the same in sales.
- Overall, platform type has no meaningful impact on customer spending.



- Most returned transactions (blue points) occur at very low sales amounts, mainly near ₹0–₹200 — meaning customers mostly return low-value purchases.
- Returned amounts are sometimes extremely high (up to 20,000+) even when sales amounts are small, indicating possible data quality issues or aggregated refund values.
- High-value sales (₹3000–₹10,000+) are almost never returned — all of them appear as “No return.”
- There is no clear correlation between sales amount and returned amount, showing that purchase value does not predict return behavior.



- Average sales remain nearly the same across all session duration bins (~405–420), showing very little variation.
- Longer browsing sessions do not lead to higher sales, meaning session duration has weak or no influence on user spending.
- Even short duration and long duration users spend almost equally, indicating uniform buyer behavior.
- This suggests that purchase decisions are not strongly tied to how long a user stays on the website.



- Session durations are evenly spread across the range of ~1500 to 5000 seconds.
- Most duration ranges appear equally frequent, showing no major peaks or dips.
- There are no strong outliers, indicating stable browsing patterns.
- Session duration does not vary much by user type or platform, showing consistent usage behavior.

