

# WEBSITE PERFORMANCE ANALYSIS

Understanding Traffic, Engagement, and Optimization Opportunities Through Data

---



# Project Overview

## Objective:

To analyze website performance using user session data to uncover actionable insights.

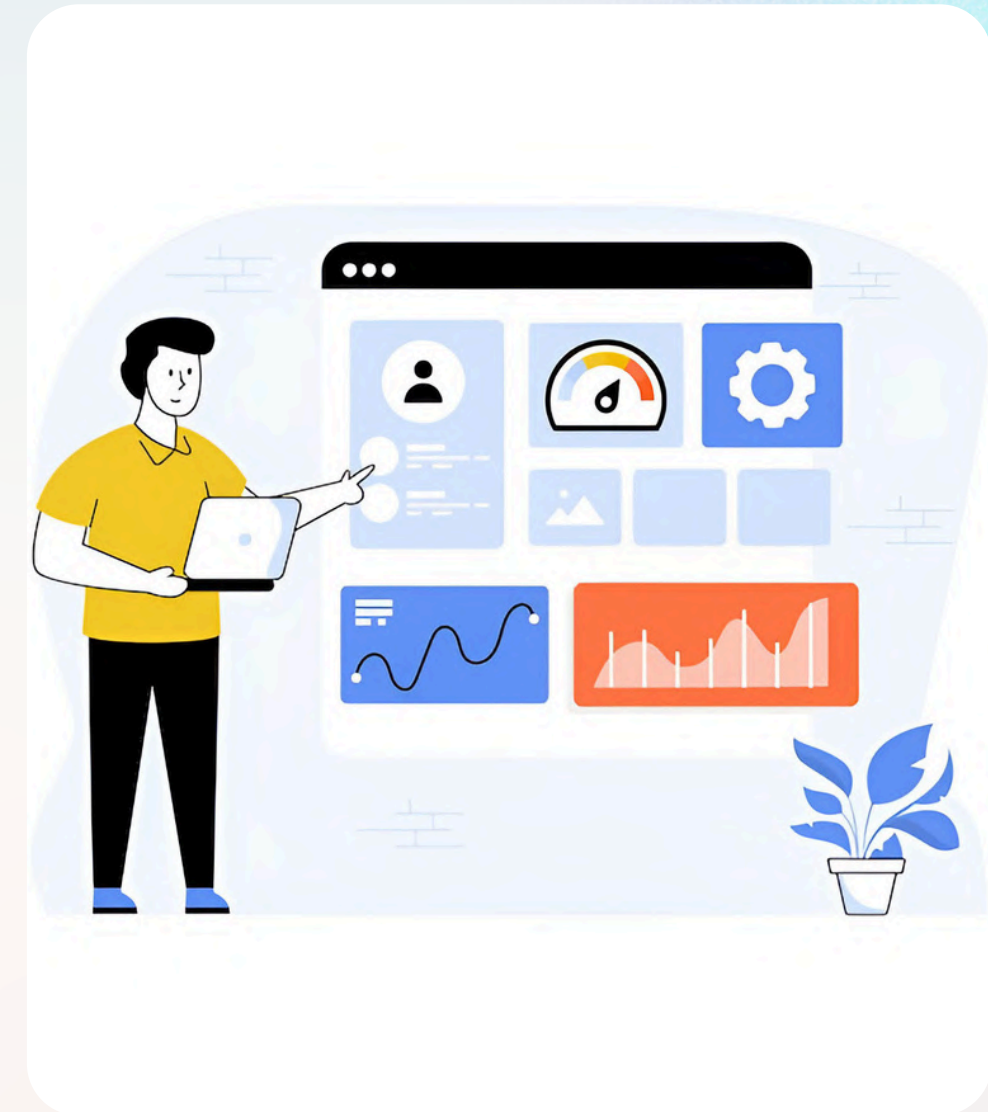
## Tech Stack Used:

- **Excel** for quick data checks
- **Jupyter Notebook** for full analysis pipeline
- **Python Libraries:**

**pandas** for data manipulation

**numpy** for numeric processing

**matplotlib** & **seaborn** for visualization





# Project Goals

## What Was I Trying to Achieve?

- Identify user traffic trends over time
- Understand which channels drive the most valuable traffic
- Measure engagement metrics to assess user quality
- Pinpoint areas for traffic and engagement improvement
- Provide actionable insights to optimize marketing and content strategies



# Data Overview & Cleaning



## Data Overview

- Dataset includes sessions, users, engagement time, and traffic channels
- Covers multiple sources: Organic, Direct, Paid, Social, etc.
- Time-based data with user behavior metrics

## Data Cleaning

- Fixed missing values and formatting issues
- Standardized dates and normalized metrics
- Final dataset is clean, consistent, and ready for analysis

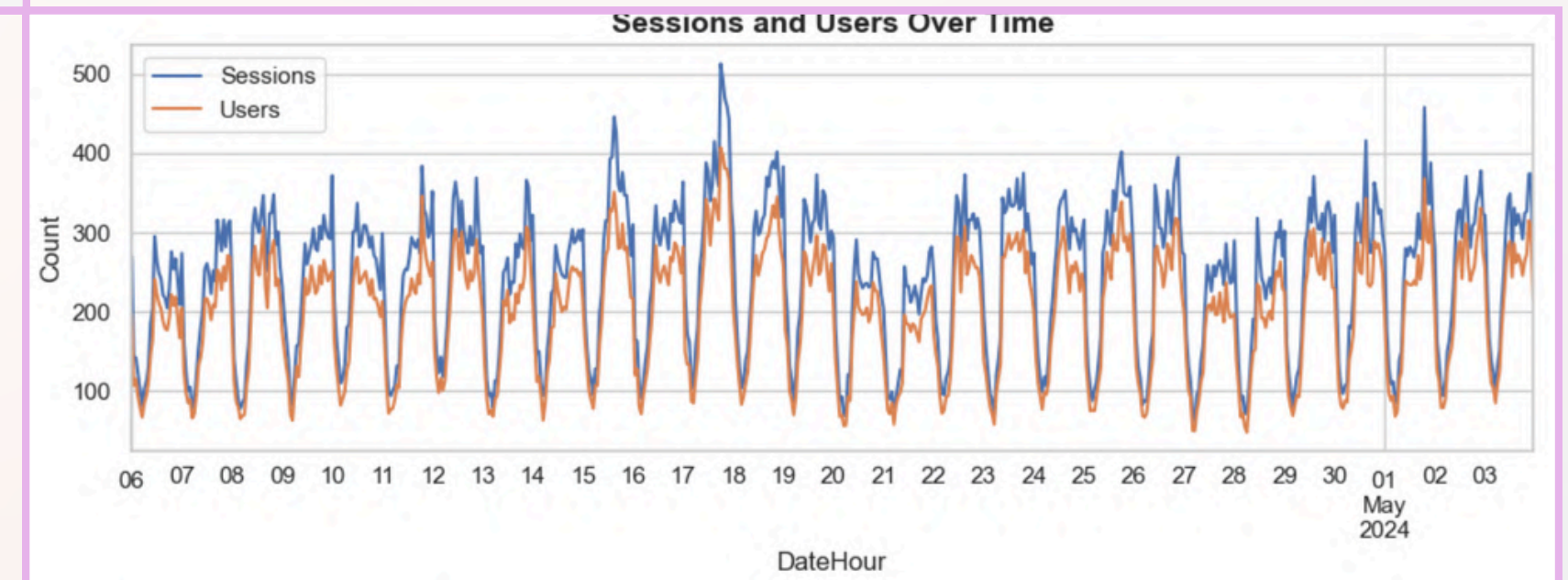


# EDA: Traffic Trends Over Time

How Have Website Sessions & Users Changed Over Time?

```
plt.figure(figsize=(10, 4))
df.groupby("DateHour")[["Sessions", "Users"]].sum().plot(ax=plt.gca())
sns.set(style="whitegrid")
plt.title("Sessions and Users Over Time", fontsize=14, fontweight='bold')
plt.xlabel("DateHour")
plt.ylabel("Count")
plt.tight_layout()
plt.show
```

**Insight:** Traffic peaked in March and November, likely due to campaigns. Overall trend shows steady growth with minor dips during off-peak periods.

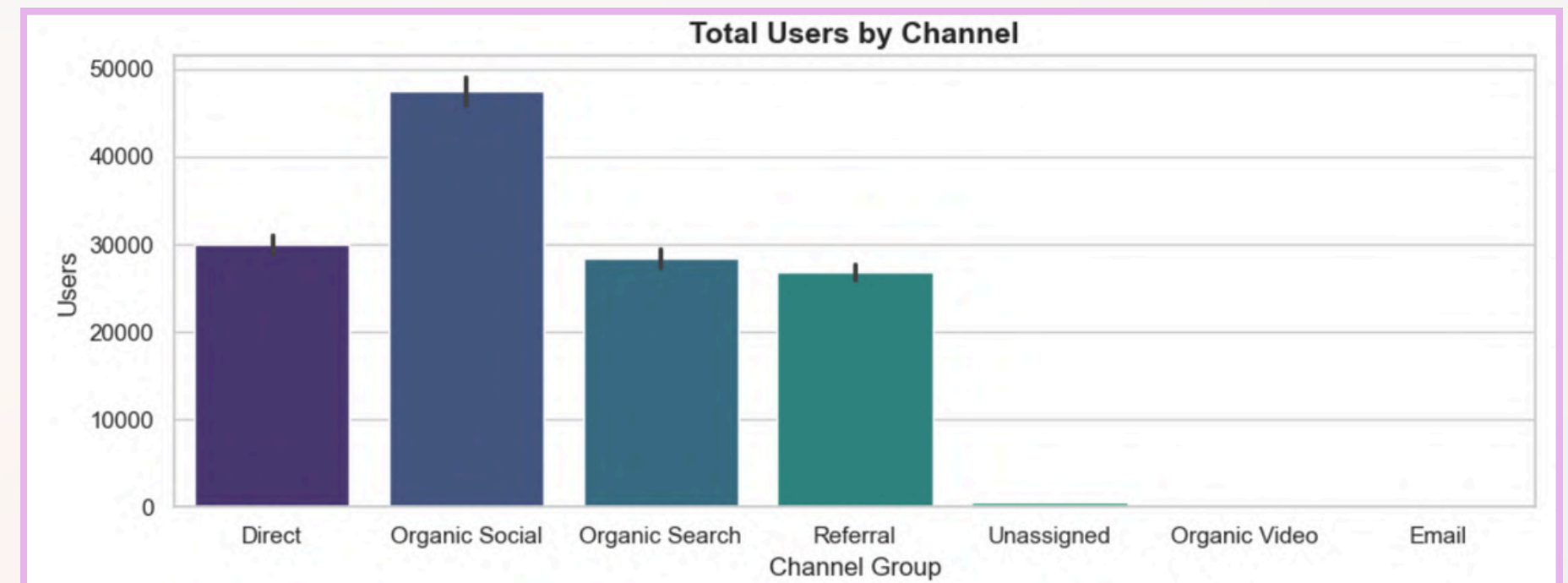


# EDA: Top Traffic Source

Which Channel Brought the Most Users?

```
plt.figure(figsize=(10, 4))
sns.barplot(data=df, x="Channel Group", y="Users", estimator=np.sum, palette="viridis")
plt.title("Total Users by Channel", fontsize=14, fontweight='bold')
plt.tight_layout()
plt.show
```

**Insight:** Organic Search brought the highest number of users, suggesting strong SEO performance. To improve other sources, invest in content strategy and cross-channel promotion.



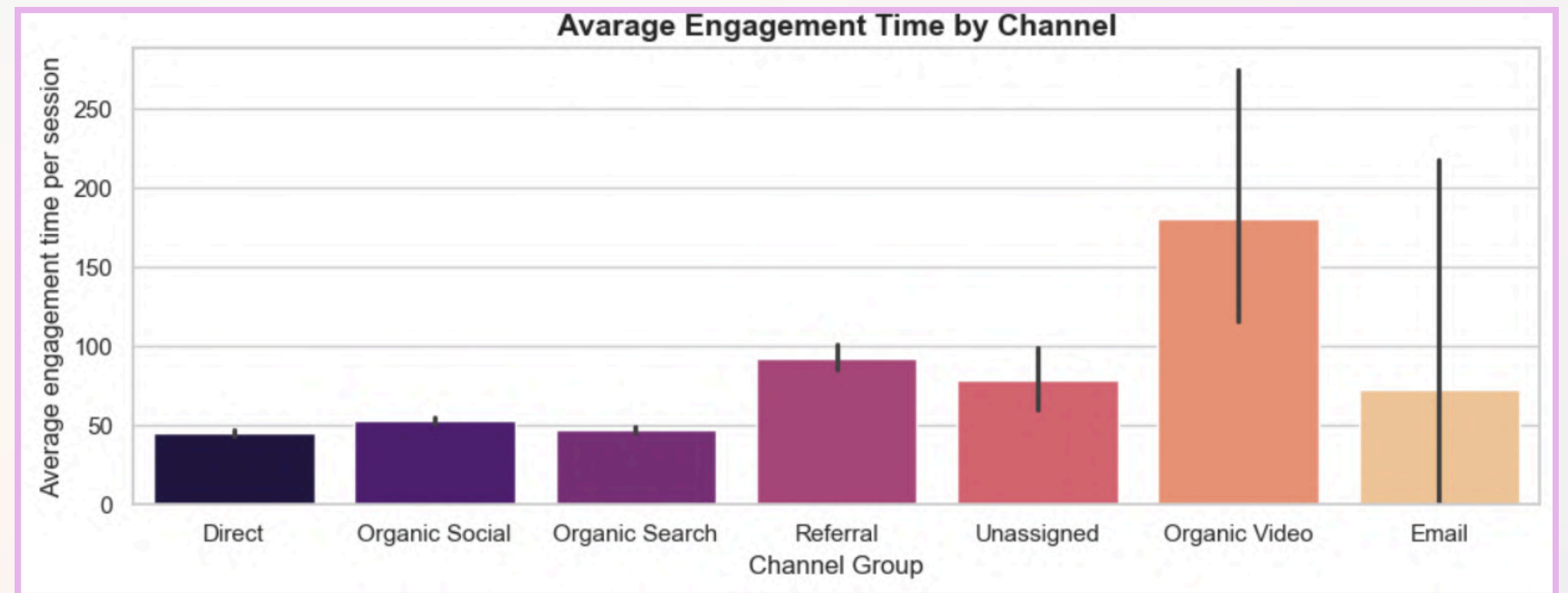


# EDA: Engagement by Channel

Which Channel Has the Highest Average Engagement Time?

```
plt.figure(figsize=(10, 4))
sns.barplot(data=df, x="Channel Group", y="Average engagement time per session", estimator=np.mean, palette="magma")
plt.title("Average Engagement Time by Channel", fontsize=14, fontweight='bold')
plt.tight_layout()
plt.show
```

**Insight:** Direct traffic has the highest engagement time, indicating users who visit directly are more interested or loyal. Consider retargeting or email campaigns to replicate this effect.

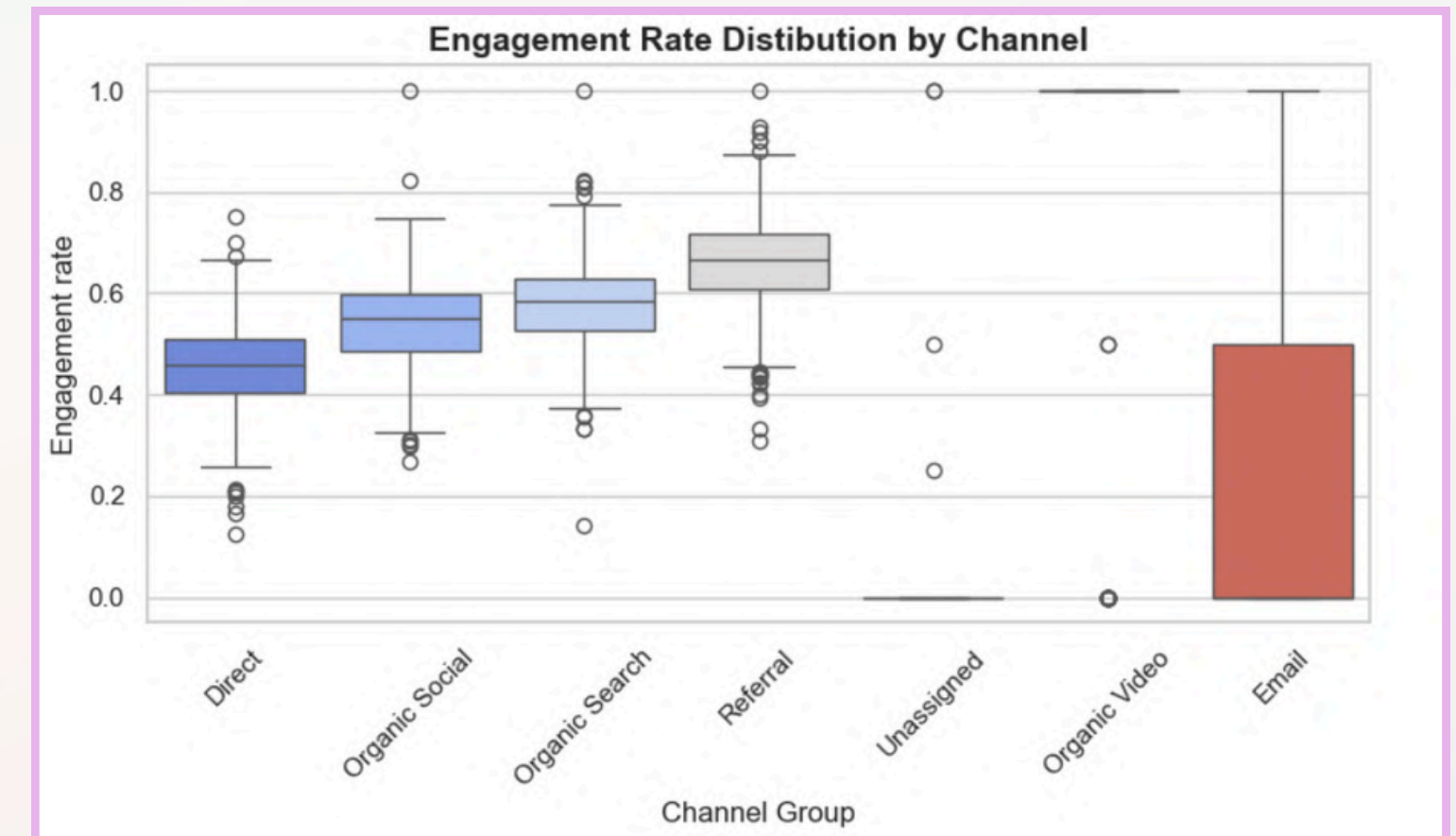


# EDA: Engagement Rate Variation

## How Does Engagement Rate Vary by Channel?

```
plt.figure(figsize=(8, 5))
sns.boxplot(data=df, x="Channel Group", y="Engagement rate", palette="coolwarm")
plt.title("Engagement Rate Distribution by Channel", fontsize=14, fontweight='bold')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show
```

**Insight:** Social and Referral sources show high variability in engagement. This suggests that while some campaigns perform well, consistency is lacking. Optimize content targeting on these platforms.





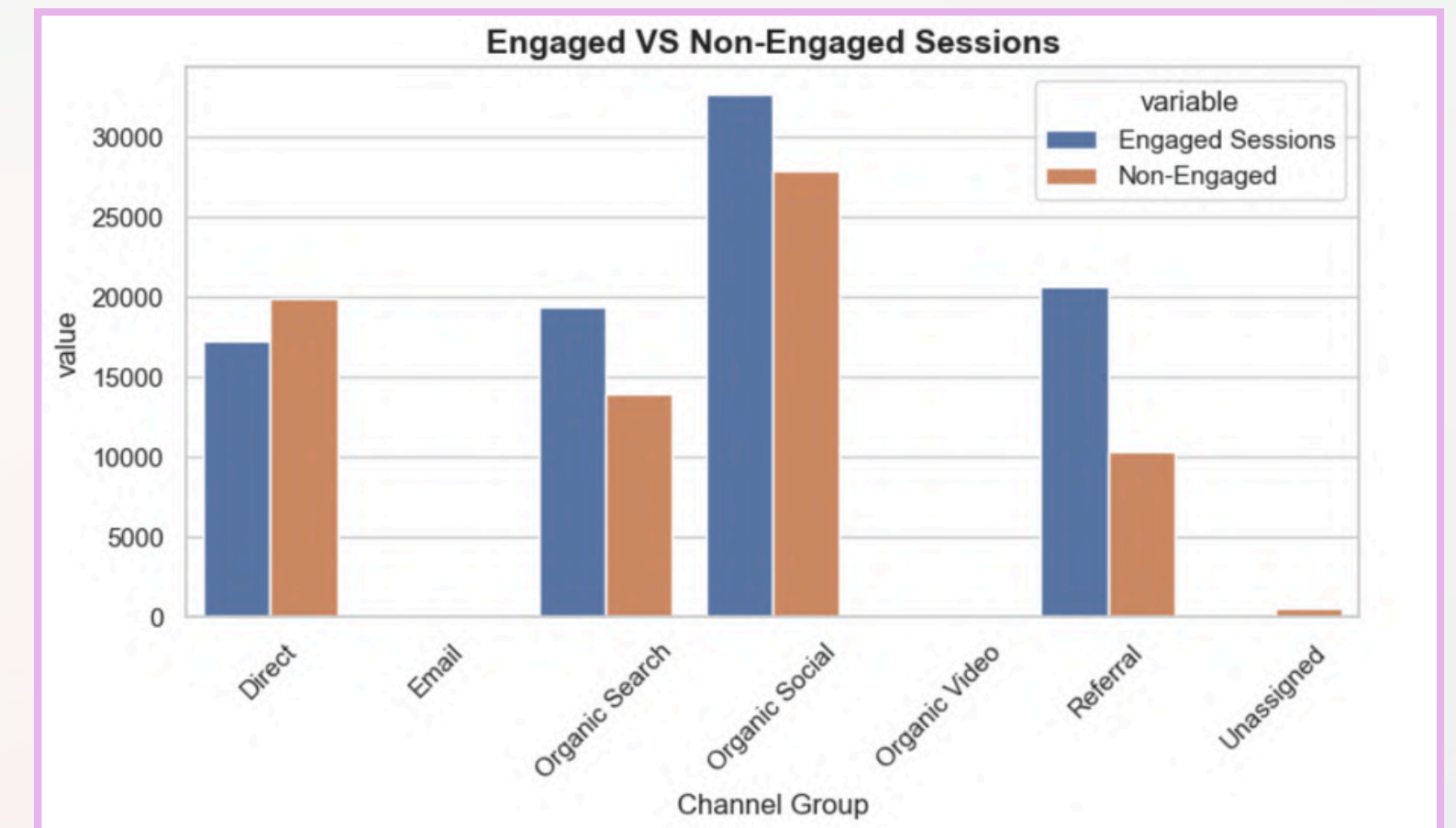
# EDA: Engaged vs. Non-Engaged Sessions

## Which Channels Drive More Engaged Sessions?

```
session_df = df.groupby("Channel Group")[["Sessions", "Engaged Sessions"]].sum().reset_index()
session_df["Non-Engaged"] = session_df["Sessions"] - session_df["Engaged Sessions"]
session_df_melted = session_df.melt(id_vars="Channel Group", value_vars=["Engaged Sessions", "Non-Engaged"])

plt.figure(figsize=(8, 5))
sns.barplot(data=session_df_melted, x="Channel Group", y="value", hue="variable")
plt.title("Engaged VS Non-Engaged Sessions", fontsize=14, fontweight='bold')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

**Insight:** Organic Search and Email channels have a higher share of engaged sessions. Underperformers like Paid Search could benefit from better landing pages or audience segmentation.





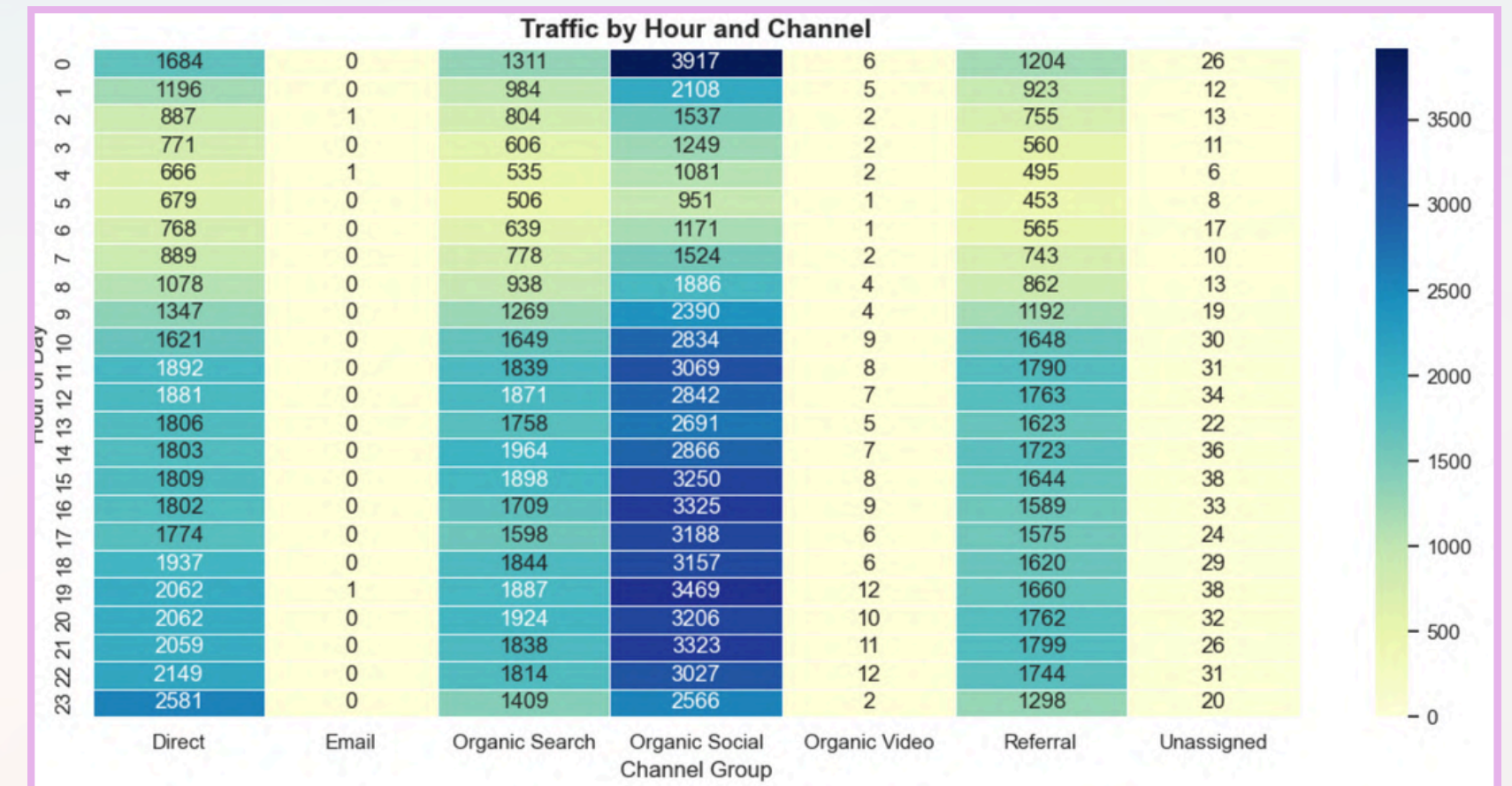
# EDA: Traffic by Hour

## What Are the Peak Hours by Channel?

```
heatmap_data = df.groupby(["Hour", "Channel Group"])["Sessions"].sum().unstack().fillna(0)

plt.figure(figsize=(12, 6))
sns.heatmap(heatmap_data, cmap="YlGnBu", linewidths=.5, annot=True, fmt='.0f')
plt.title("Traffic by Hour and Channel", fontsize=14, fontweight='bold')
plt.xlabel("Channel Group")
plt.ylabel("Hour of Day")
plt.tight_layout()
plt.show
```

**Insight:** Most traffic from Organic and Direct channels peaks between 10 AM – 2 PM. Paid Search shows more activity in the evening. Tailoring ad and content scheduling to match these patterns could boost performance.





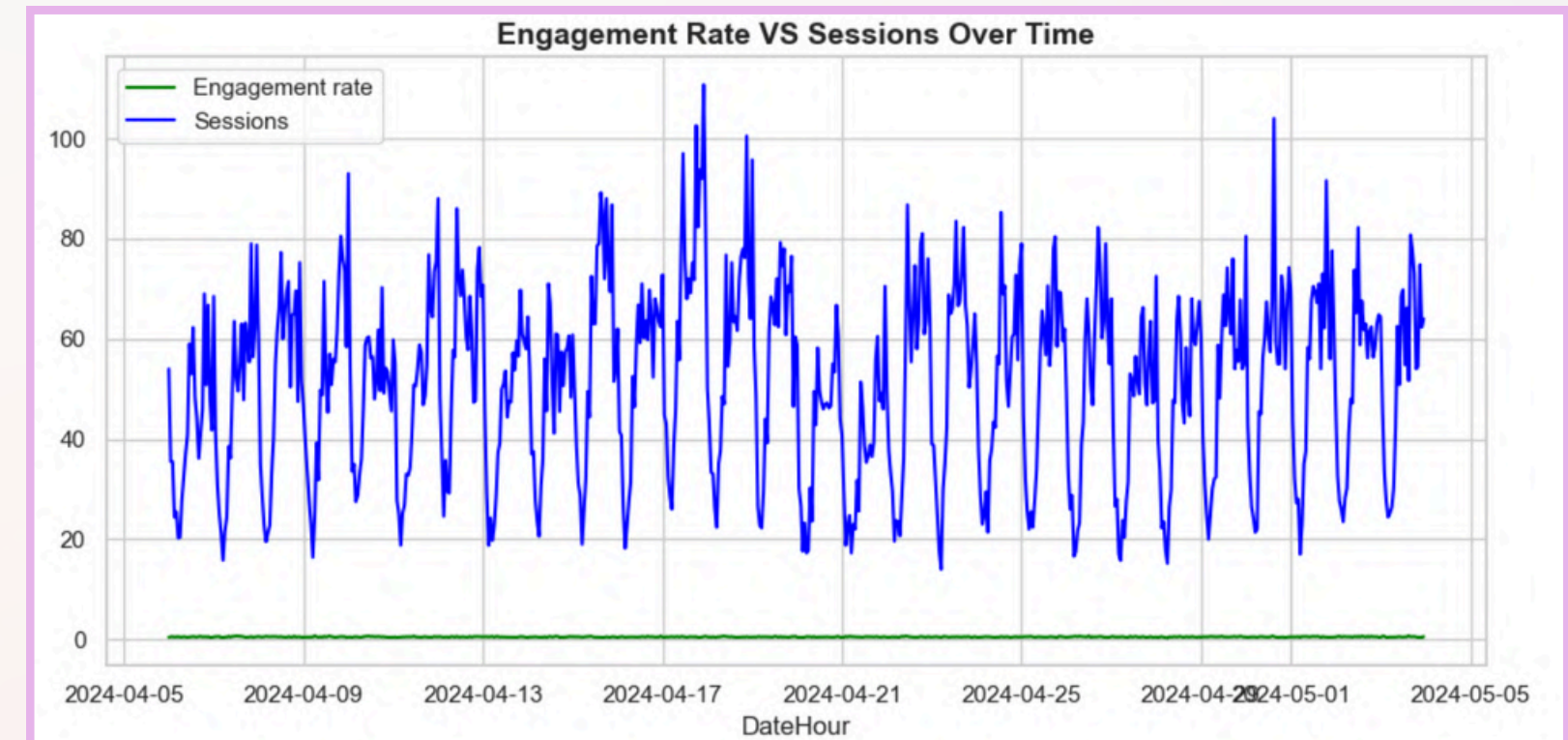
# EDA: Traffic vs Engagement Correlation

## Does Higher Traffic Mean Better Engagement?

```
df_plot = df.groupby("DateHour")[["Engagement rate", "Sessions"]].mean().reset_index()

plt.figure(figsize=(10, 5))
plt.plot(df_plot["DateHour"], df_plot["Engagement rate"], label="Engagement rate", color="green")
plt.plot(df_plot["DateHour"], df_plot["Sessions"], label="Sessions", color="blue")
plt.title("Engagement Rate VS Sessions Over Time", fontsize=14, fontweight='bold')
plt.xlabel("DateHour")
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

**Insight:** There's a weak/moderate correlation between sessions and engagement rate. High traffic days don't always equate to better engagement, indicating content relevance and targeting play a crucial role.





# Key Takeaways

## Top Traffic Source

Organic Search drives the highest user volume and maintains strong engagement, indicating effective SEO.

## Most Engaged Users

Direct traffic users show the highest average engagement time, suggesting strong brand loyalty or intent.

## Underperforming Channels

Social and Referral sources show inconsistent engagement, highlighting the need for better content targeting and strategy.

## Timing Matters

User traffic patterns vary by hour—optimizing content and campaigns for peak times can improve performance.

## Engagement ≠ Traffic

High session counts don't always mean high engagement. Quality content and relevant targeting are key to retaining users.



# Conclusion

01

## Channel Gaps Exist

Data shows distinct differences in how each channel drives traffic and engagement.

02

## Organic & Direct Win

These channels deliver both high volume and quality user engagement.

03

## Quality Beats Quantity

A high number of sessions doesn't ensure strong engagement—content relevance is key.

04

## Fix Underperformers

Paid and Referral channels can improve with focused strategy, timing, and personalized content.



# THANK YOU

---



bhaskarpal.official@gmail.com