Assignment || Growth Data Analyst || **WP**POOL

Section 1:

WPPOOL primarily focuses on developing and providing unique WordPress plugins that simplify business operations for websites. WPPOOL faced many challenges in its early days but quickly learned and grew into a successful **WordPress** product company. Some of the popular products from WPPOOL include **WP Dark Mode, Google Sheets to WP Table Live Sync, Easy Video Reviews, Jitsi Meet, Zero BS Accounting, and KeenShot, among others.**

WP POOL products are developed with customers' pain points in mind. WPPOOL's customers primarily include WordPress users, web developers, designers, bloggers, and business owners who want to enhance their website's functionality and user experience.

WPPOOL excels in providing highly optimized, fast, and bug-free WordPress products that solve customers' pain points. All products are coded with high optimization and they are regularly updated. Their amazing products provide quality and creative solutions for customers' WordPress websites.

Section 2:

🎯 Tasks:

1: Data Exploration & Cleaning:
A: Load the dataset and check for missing values, duplicates, or inconsistencies

I have performed data loading and checked for missing values, duplicates, and inconsistencies using two methods:

Power BI (Power Query): I imported the dataset, viewed the column distribution, and checked the column profile.

Python: Used the Pandas library to check for missing values, duplicates, and inconsistencies.

Data Profile:

Variable Name	Data Type	# Missing Values	# Unique Values	Min	Max	Mode
user_id	INT	0	20000	1	20000	
install_date	DATE	0	366	1/1/2023	1/1/2024	8/13/2023
last_active_date	DATE	0	357	1/3/2023	1/1/2024	1/1/2024
subscription_type	STRING	0	2			
country	STRING	0	7			
total_sessions	INT	0	300	1	300	78
page_views	INT	0	856	1	1500	60
download_clicks	INT/BINARY	0	2			
activation_status	INT/BINARY	0	2	0	364	
days_active	INT	0	361			0
pro_upgrade_date	DATE	15971	337	1/8/2023	1/1/2024	12/31/2023
plan_type	STRING	15971	4			
monthly_revenue	INT	0	4	0	99	0
churned	INT/BINARY	0	2	0	1	0

Comments:
1. We Have 4 types of variables, (INT, DATE, STRING, BINARY)

2: We Have presence in 7 countries

3. We have missing values in two columns (Pro_upgrade_date and plan_type). However, it's fine because these values are intended to be blank for free subscription-type users. The blank value in the plan_type column will be replaced with **NO_plan** using Power BI Power Query, as it will look better in the final report.

4: We can handle missing values

B: Explain how you handled missing data and why you chose that method.

I use Power BI (Power Query) because it makes it easy to handle missing data with a simple interface. I use Python pandas to double-check the data and ensure everything is correct.

C: Provide a summary of the dataset, including the distribution of Free vs. Pro users.

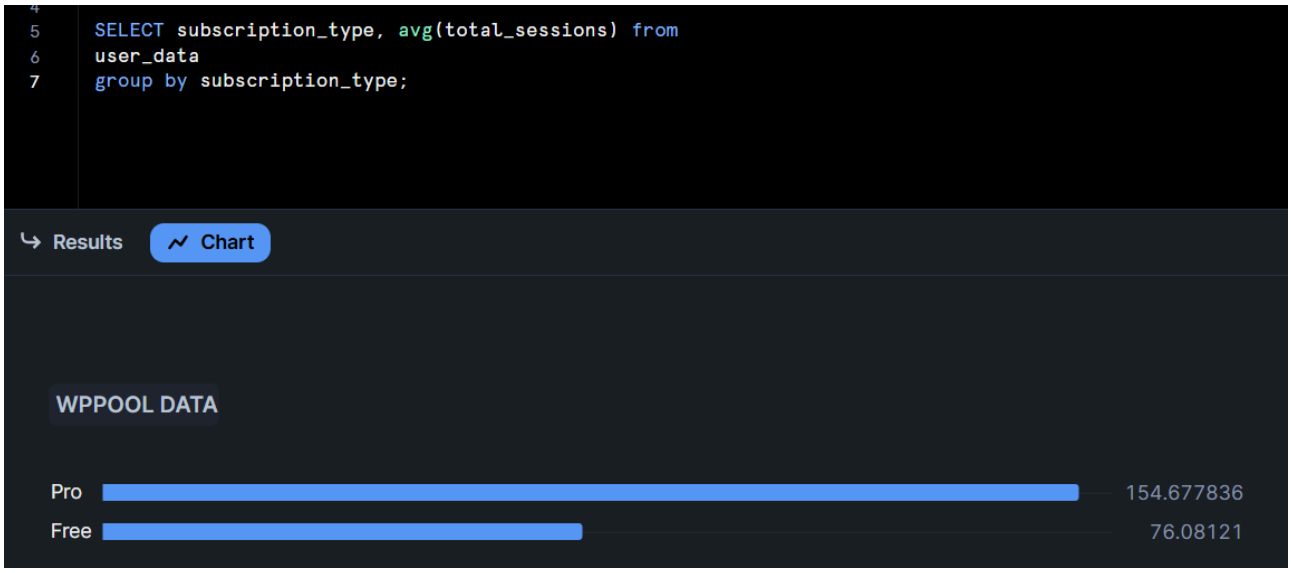
The summary of the dataset is provided above.,The distribution of Free vs. Pro user:

Free user :15971
Pro user : 4029
Internal summary:

I have done this SQL task using Snowflake because it allows chat to be displayed together.

2 :User Engagement Analysis:

A: Identify the average number of sessions for Free vs. Pro users (Optional: SQL query)



Pro user average number of sessions is : 155
Free user average number of sessions is : 76

B: Find the top 5 most active users based on total sessions (Optional: SQL query).

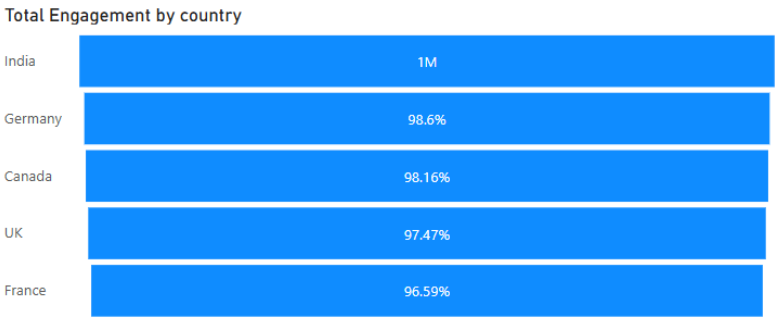
```sql
// Find the top 5 most active users based on total sessions (Optional: SQL query).

SELECT user_id, total_sessions

from user_data
order by total_sessions  desc,user_id
limit 5;
```

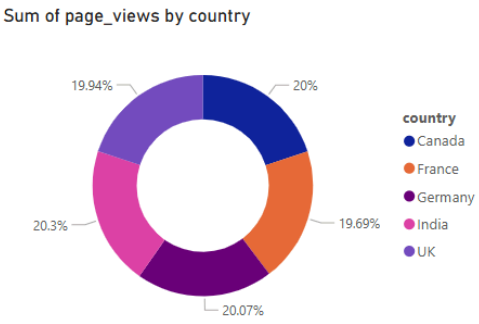
Results | Chart

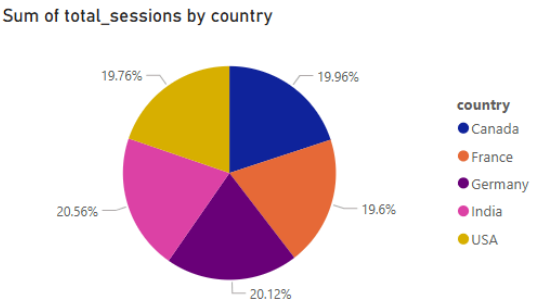
USER_ID	TOTAL_SESSIONS
189	300
822	300
1572	300
2457	300
3349	300

I have provided the top 5 active users with the highest 'Total Sessions' (300) in ascending order of user_id.

C: Identify the top 5 countries with the highest engagement.







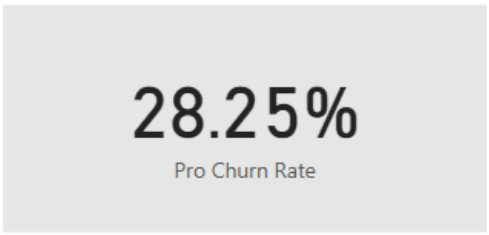
Top 5 countries with the highest engagement: Canada, France, Germany, India, UK.

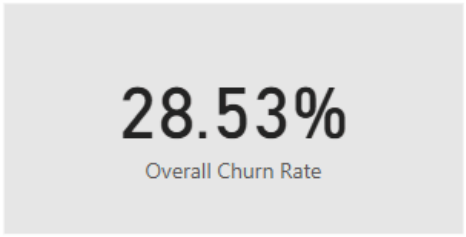
The total engagement is calculated using this DAX:

```
Engagement =
sum('wppool_growth_data'[total_sessions])+SUM(wppool_growth_data[page_views])
```

3:Churn Analysis

A: Calculate the overall churn rate for Free vs. Pro users (Optional: SQL query).






```
19
20
21
22
23  select subscription_type,
24  count(case
25          when churned=1 then user_id end) *100/ count(user_id) as churn_rate
26  from user_data
27  group by subscription_type;
28
29
```

Results | Chart

	SUBSCRIPTION_TYPE	CHURN_RATE
1	Free	28.595579
2	Pro	28.245222

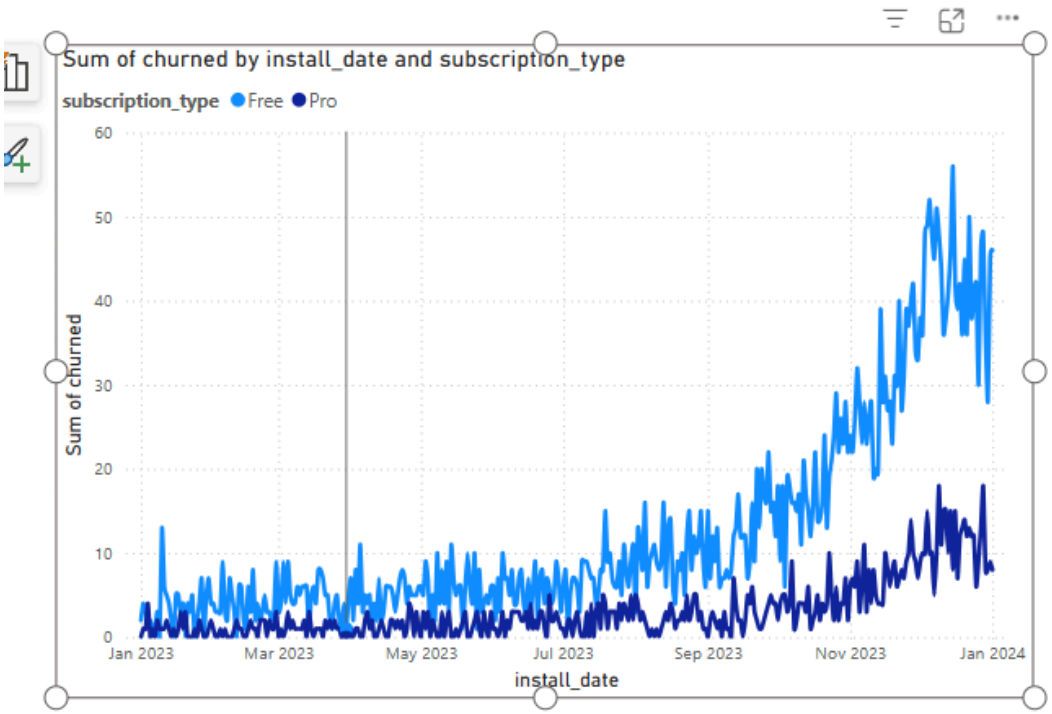
The overall churn rate, including Free and Pro users, is 28.53%

C: Compare churn trends between Free and Pro users.

```sql
SELECT
    subscription_type,
    sum(churned) as total_churned,// 0 sum kuno kaj asba na
    CONCAT(
        COUNT(CASE WHEN churned = 1 THEN user_id END) * 100 / COUNT(user_id),
        '%'
    ) AS churn_rate
FROM
    user_data
GROUP BY
    subscription_type;
```

Results | Chart

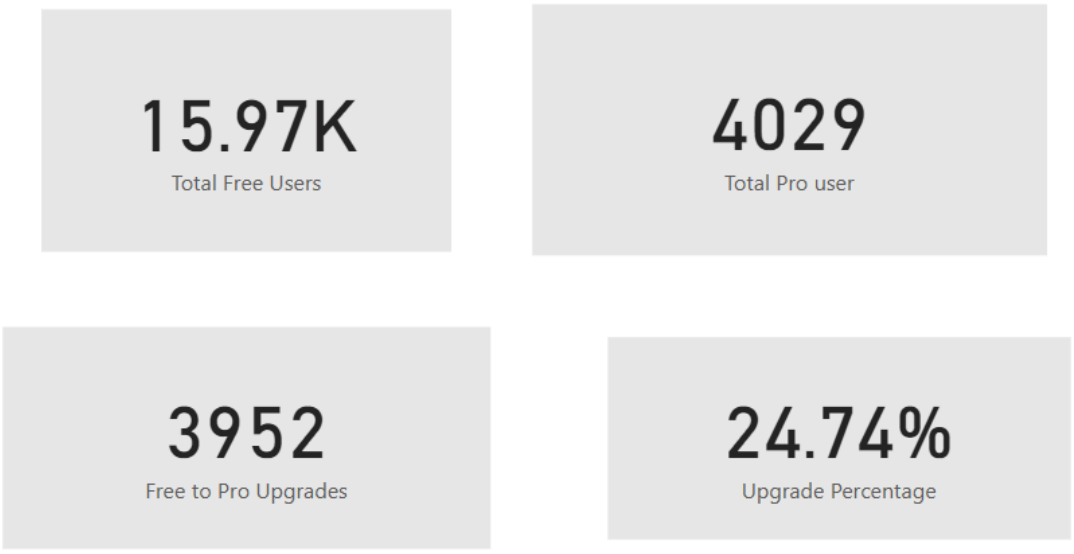
SUBSCRIPTION_TYPE	TOTAL_CHURNED	CHURN_RATE
Free	4567	28.595579%
Pro	1138	28.245222%



There is slightly more churn among Free users compared to Pro users, although the difference is small.

4:Revenue & Upgrade Trends

A: What percentage of users upgraded from Free to Pro?



Using this dax:

```
Free to Pro Upgrades =
CALCULATE(
    COUNTROWS(wppool_growth_data),
    wppool_growth_data[subscription_type] = "Pro" &&
NOT(ISBLANK(wppool_growth_data[pro_upgrade_date])) &&
    wppool_growth_data[install_date] < wppool_growth_data[pro_upgrade_date]
)
```

Explain this dax: Some users may have taken the Pro subscription directly,If the install_date is earlier than the pro_upgrade_date, it means the user was initially on the Free subscription and later upgraded to the Pro subscription.

```
Upgrade Percentage =
DIVIDE(
    [Free to Pro Upgrades],
    [Total Free Users],
    0
)
```

Than divide the (Free to Pro upgrades/total free users)

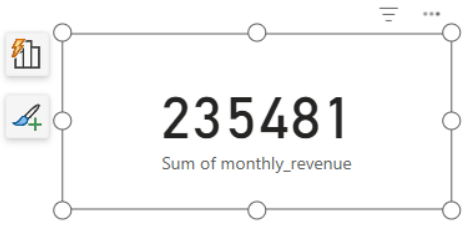
24.74 percentage of users upgraded from Free to Pro

B: Calculate the total monthly revenue from Pro users (Optional: Google Colab) Using the jupyter notebook:

Calculate the total monthly revenue from Pro users

```
[6]: total_monthly_revenue = df[df['subscription_type'] == 'Pro']['monthly_revenue'].sum()
     total_monthly_revenue
```

```
:[6]: 235481
```



Free users do not generate revenue

C: Which Pro plan (Basic, Standard, or Enterprise) contributes the most revenue?

Which Pro plan (Basic, Standard, or Enterprise) contributes the most revenue?

```
]: result = df.dropna(subset=['plan_type']).groupby('plan_type', as_index=False)['monthly_revenue'].sum()
   result
```

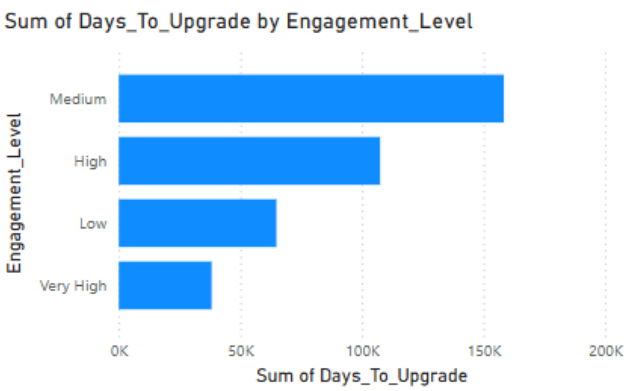
]:

	plan_type	monthly_revenue
0	Basic	80339
1	Enterprise	78764
2	Standard	76378

Basic contributes the most revenue, generating 80,339

D: Analyze how long it takes for Free users to upgrade based on country and engagement level.

country	avg_days_to_update
UK	87.90
USA	90.24
France	91.16
Germany	91.53
India	92.58
Canada	93.57
Australia	93.68



UK upgrades fastest (87.90 days), Australia slowest (93.68 days).
Medium engagement users take the most time to upgrade.
Very Higher engagement leads to quicker upgrades.

How to solve this problem : Days_To_Upgrade is calculated for users who have upgraded to the Pro version. It is determined by the difference in days between the install_date and pro_upgrade_date

```
Days_To_Upgrade =
IF(
    NOT(ISBLANK(wppool_growth_data[pro_upgrade_date])),
    DATEDIFF(wppool_growth_data[install_date],
wppool_growth_data[pro_upgrade_date], DAY),
    BLANK()
)
```

If the pro_upgrade_date is blank (indicating no upgrade), the value is blank.

The report shows the average number of days it takes for users to upgrade to Pro.

Engagement Levels Defined:

```
Engagement_Level =
IF([total_sessions] >= 200 && [page_views] >= 1000, "Very High",
```

```
IF([total_sessions] >= 100 && [page_views] >= 500, "High",
IF([total_sessions] >= 50 && [page_views] >= 100, "Medium", "Low")))
```

Then, I created an engagement report

5: Actionable Growth Recommendations

A: Suggest three strategies to reduce churn**.**

High-Quality Products Sell: Ensure Customer Satisfaction and Long-term loyalty

Messages on user activity: Send notifications based on user activity and preferences to re-engage users and remind them of the benefits of continuing with the service.

RFM analysis: Special offers and Provide discounts to the best customers based on RFM analysis.

Support System: Offering a better customer support system. using multiple channels to solve user problems.

B : Propose two ways to increase Free-to-Pro conversions.

Engage Users: Looking at the engagement levels, I see that users with **very high engagement** quickly upgraded to Pro, and some low engagement users also upgraded to Pro. By targeting them with various offers, we can encourage free users to upgrade to Pro.

Benefits: We need to inform free users about the various benefits and features of upgrading to Pro

Additionally, the market should be operated based on the country.

C : Identify potential market expansion opportunities based on country trends

Target high-growth markets: Identify countries where the demand for your product or service is increasing.The report shows that there is high demand for WordPress in the US, UK, and India.Our report shows that in these countries, there has been an increase in the number of users upgrading to Pro.

6: Conversion Rate Optimization (CRO)

A: If WPPOOL increases the landing page conversion rate by 10%, what would be the estimated impact on Pro upgrades?

If WPPOOL increases their landing page conversion rate by 10%, the estimated impact on Pro upgrades would be:

We see, **Current conversion rate**: 24.74%

New Conversion Rate = Current Conversion Rate + (10% increase in the current rate).

New Conversion Rate is : **27.21%**

Let's assume WPPOOL receives **1,000 visitors** per month to the landing page

Pro upgrades before increase=1,000×24.74%=247.4 users

Pro upgrades after increase=1,000×27.21%=272.1 users

If WPPOOL increases the landing page conversion rate by 10%, they could get 24.7 more Pro upgrades for every 1,000 visitors.

B: (Optional: Google Colab) Run a simple A/B test simulation (e.g., using a chi-square test) to evaluate conversion optimization.

```
: from scipy.stats import chi2_contingency
  import numpy as np

  # Data input
  group_a = np.array([3952, 20000])   # Current Conversion Rate landing page
  group_b = np.array([4350, 19602])   # Increase Conversion Rate by 10%

  # Chi-Square test
  chi2, p_value, _, _ = chi2_contingency([group_a, group_b])

  print(f"P-value: {p_value}")
  if p_value < 0.05:
      print("The difference is statistically significant.")
  else:
      print("The difference is not statistically significant.")
```

```
P-value: 1.650391535367903e-06
The difference is statistically significant.
```

The new landing page has achieved statistical significance.

C: Suggest three A/B test ideas that could help improve the conversion rate, and explain how you would measure their success.

1: Page optimizations
2: Search results Optimization
3: Email Marketing
4: CTA Optimization

7: Growth Strategy & KPI Recommendations

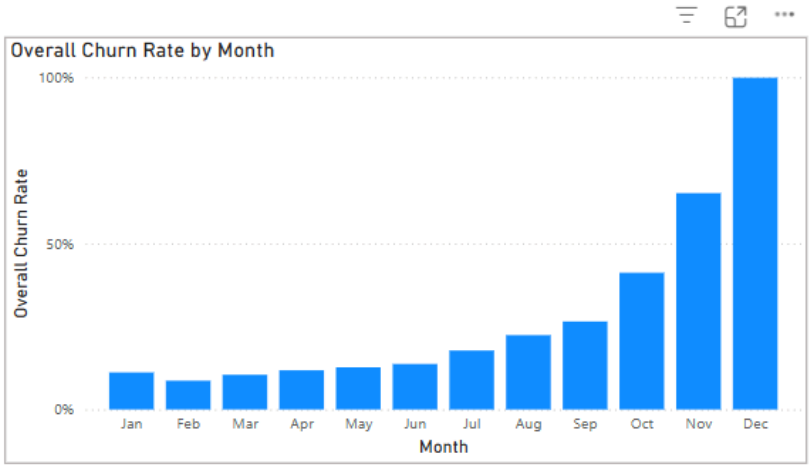
A: Identify 3 key performance indicators (KPIs) WPPOOL should track.

1: **Conversion Rate (Free to Pro)**:

Why It's Important:The conversion rate is important because it shows how well WPPOOL is turning free users into paying customers. A higher conversion rate means that users see value in the Pro version and are willing to pay for extra features.

2:Customer Churn Rate:

Why It's Important:Churn is an important measure of customer dissatisfaction.Reducing churn is important for keeping a strong and loyal user base. It helps identify months when user experience or customer support need improvement.

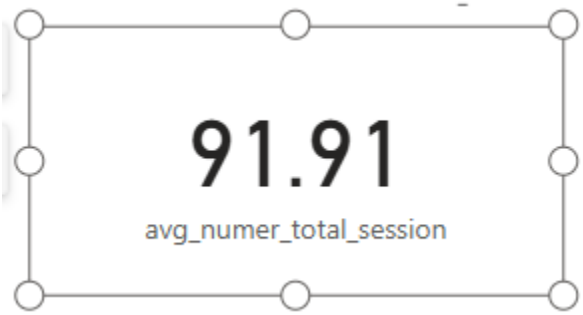


28.53%

Overall Churn Rate

3: Average Total Sessions:

The average number of total sessions shows how often users use the WPPOOL plugin. More sessions mean higher user engagement, satisfaction, and a greater chance of conversions.



4: Total Revenue and MOM Growth%

Total Revenue

235K

Month	Total Revenue	Previous Month Revenue	MoM Growth (%)
Jan	19792		
Feb	19906	19450	2.34%
Mar	18985	19906	-4.63%
Apr	21291	18985	12.15%
May	19003	21291	-10.75%
Jun	19765	19003	4.01%
Jul	20141	19765	1.90%
Aug	19873	20141	-1.33%
Sep	17563	19873	-11.62%
Oct	19165	17563	9.12%
Nov	20139	19165	5.08%
Dec	19858	20139	-1.40%
Total	**235481**		

Why It's Important:

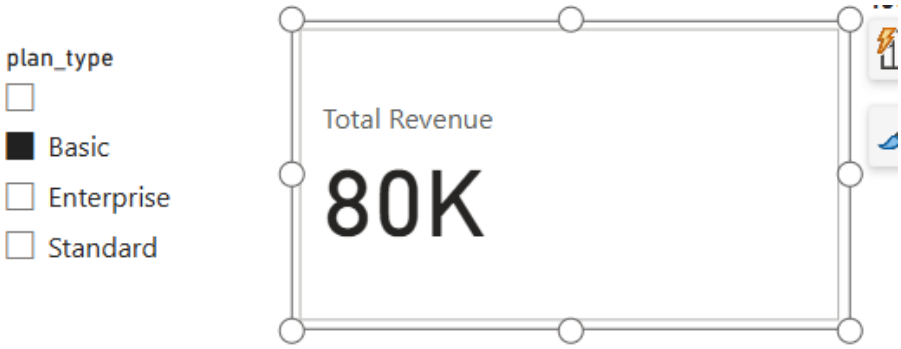
It measures the total income generated from Pro users, showing the financial health and success of the business.

It tracks the percentage change in revenue from month to month, helping to evaluate growth trends and identify patterns for improvement.

2: Suggest 2 actionable growth strategies WPPOOL can implement based on your analysis.

Plan-Based Upselling Strategy:

WPPOOL has three plans—Basic, Standard, and Premium. By suggesting plan upgrades to the right users at the right time, WPPOOL can increase its revenue. It's crucial to recommend the right plan to users based on their usage benefits and **target specific** countries to increase revenue. A large revenue comes from basic plan users. To retain them, offer various incentives.



User Engagement & Retention Campaign:

Many users install WPPOOL but stop using it over time. An interactive campaign can help keep them active.It could be a **custom feature suggestion**, where free users are encouraged to upgrade to the Pro version. Because increasing Pro users will also

increase revenue.Regular updates based on user feedback.WPPOOL can implement personalized email and other social media campaigns based on user behavior.

Why is it important:

It will increase retention and reduce churn.

It will improve user engagement and conversion.

C: How would you measure the success of these strategies?

the success of these strategies,WPPOOL can track the following metrics

Conversion Rate (Free to Pro),Revenue Growth,User Engagement,Churn Rate Reduction,

By monitoring these KPIs, WPPOOL can determine how effectively the strategies are driving growth and user retention.

We have identified all the KPIs in this project.

8: Data Storytelling & Visualization

Creating data visualizations in Power BI:

https://app.powerbi.com/view?r=eyJrIjoiMjJjOGU4MjAtNTgwZC00MmUxLTg1YWYtNjM5NTdjZTJkNWUzIiwidCI6IjM0YmVhMGY1LTRlNmMtNDdjZC05NjFmLTIxMjA1ZGYxMjQ1MSIsImMiOjEwfQ%3D%3D

Due to a lack of time, I couldn't do Data Storytelling.

SQL Query

https://app.snowflake.com/bjqfstl/yg45765/wHanEoqtePJ#query

Python and MySql code :

https://drive.google.com/drive/folders/1gsTSz0FBT7Su1Cj-kgoURLxdF-wGvbZP?usp=sharing

Doing it all,
In all new ways

Section 3

A: What are your 3 most favorite Books?

Ans: মুক্ত বাতাসের খোঁজে ,বেলা ফুরাবার আগে,Atomic Habits

B : What are your 3 most favorite TV shows?
Ans: Shark tank, Ertuğrul Gazi,Salahuddin Ayyubi

C: What motivates you the most in a workplace?
Ans: **"You cannot change your future, but you can change your habits, and surely your habits will change your future." – Dr. A.P.J. Abdul Kalam.**
This quote inspires me the most at work. I believe each day is a chance to improve, so I try to make the best use of it. Listening to or reciting the Quran also gives me peace, strength, and guidance. It helps me stay focused, positive, and motivated to do my best in the workplace.