- **Business Analysis Project – Customer Churn & Segmentation**

This repository contains the deliverables for my Business Analysis Internship Project with SaiKet Systems.

The project focuses on analyzing customer churn in a telecommunications company and developing visualizations & insights to identify at-risk customers.

1 🏗 Project Overview

Objective: Analyze customer churn behavior, clean and prepare data, perform exploratory analysis, and create visualizations for customer segmentation.

Dataset:`Telco_Customer_Churn_Cleaned_dataset.CSV`

Tools Used: Python, Pandas, Matplotlib, Seaborn, Plotly

2. ✓ **Completed Tasks**

As per the internship requirement, I have **completed 4 out of 5 tasks:**

- **Task 1 – Understand the Dataset**

Steps:

- Loaded dataset with Pandas.
- Displayed first 10 rows.
- Checked data types of each column.
- Verified missing values.
- Skills: Data loading, data inspection.
- Output: Dataset summary and head preview.

- **Task 2 – Data Cleaning**

- Steps:

- Handled missing values (imputation/drop where required).
- Removed duplicate records.
- Standardized column names (converted to lowercase, replaced spaces with underscores).
- Skills: Data cleaning, preprocessing.
- Output: Clean, analysis-ready dataset.

- **Task 3 – Exploratory Data Analysis (EDA)**

-Steps:

- Generated summary statistics (mean, median, mode).
- Created histograms & boxplots for numerical columns.
- Analyzed churn proportions (churn vs. non-churn).
- Skills: Statistical analysis, data visualization.
- Outputs:
- Distribution plots of customer attributes.
- Churn vs. Non-churn visualization.

- **Task 4 – Customer Segmentation Visualization**
  - Steps:
  - Created **tenure groups**: `0–12 months`, `13–36 months`, `37+ months`.
  - Built D**onut chart** showing distribution across tenure groups.
  - Built **Clustered Bar Chart** comparing average monthly charges across tenure groups.
  - Added annotations to highlight significant trends:
  - New customers (`0–12 months`) pay higher average charges.
  - Long-term customers (`37+ months`) generally have stable/lower charges.
  - Skills: Segmentation, advanced visualization (Pie, Donut, Bar).
  - Outputs:
  - Donut chart of tenure distribution.
  - Bar chart of average monthly charges by tenure.

Notes

- All work is original and follows SaiKet Systems internship guidelines.
- Outputs are saved in CSV for reproducibility.
- Visualizations are interactive (Plotly) for better presentation.

👨‍💻 Author

Puneet Tiwari

📧 Email: [tiwaripuneet9695@gmail.com]

📧 Email: [puneettiwari.tech@gmail.com] (new)

“ I successfully completed this internship project with **SaiKet Systems, gaining hands-on experience in business analysis, data cleaning, visualization, and customer churn analysis. ”**

Thankyou @Saiket Systems

WhatsApp | Experience | Puneet Tiwa | Media file types support | Puneet Tiwari.pdf | screenshot shortcut key

File C:/Users/poorn/Downloads/Puneet%20Tiwari.pdf

Puneet Tiwari.pdf 1 / 1 47%

1

SaiKet Systems
Innovate. Elevate. Excel.

Date: 05-08-2025

SKS/A2/C40918

INTERNSHIP OFFER LETTER

Dear **Puneet Tiwari**

Congratulations on being selected for the position of **Business Analysis Intern** at Saiket Systems. Your impressive qualifications, analytical skills, and enthusiasm for Business Analysis have stood out, and we are eager to have you as a valuable addition to our team.

As a Business Analysis Intern at Saiket Systems, you will embark on a journey to explore the realm of data-driven insights and decision-making. Your role will involve working on a diverse range of projects that require your expertise in Business Analysis and interpretation. You will be instrumental in uncovering meaningful patterns and trends within our data, which will drive our company's strategies and initiatives forward.

We would like to take this opportunity to congratulate you once again and welcome you to Saiket Systems. We are enthusiastic about the prospect of working with you and are eager to start this journey together.

Best regards,

Team Saiket Systems

support@saiket.in www.saiket.in

Hot days ahead 28°C

Search

ENG IN 23:32 13-08-2025

WhatsApp | Experience | Puneet Tiwa | Media file types support | Puneet Tiwari.pdf | screenshot shortcut key

File C:/Users/poorn/Downloads/Puneet%20Tiwari.pdf

Puneet Tiwari.pdf 1 / 1 47%

1

SaiKet Systems
Innovate. Elevate. Excel.

Date: 05-08-2025

SKS/A2/C40918

INTERNSHIP OFFER LETTER

Dear **Puneet Tiwari**

Congratulations on being selected for the position of **Business Analysis Intern** at Saiket Systems. Your impressive qualifications, analytical skills, and enthusiasm for Business Analysis have stood out, and we are eager to have you as a valuable addition to our team.

As a Business Analysis Intern at Saiket Systems, you will embark on a journey to explore the realm of data-driven insights and decision-making. Your role will involve working on a diverse range of projects that require your expertise in Business Analysis and interpretation. You will be instrumental in uncovering meaningful patterns and trends within our data, which will drive our company's strategies and initiatives forward.

We would like to take this opportunity to congratulate you once again and welcome you to Saiket Systems. We are enthusiastic about the prospect of working with you and are eager to start this journey together.

Best regards,

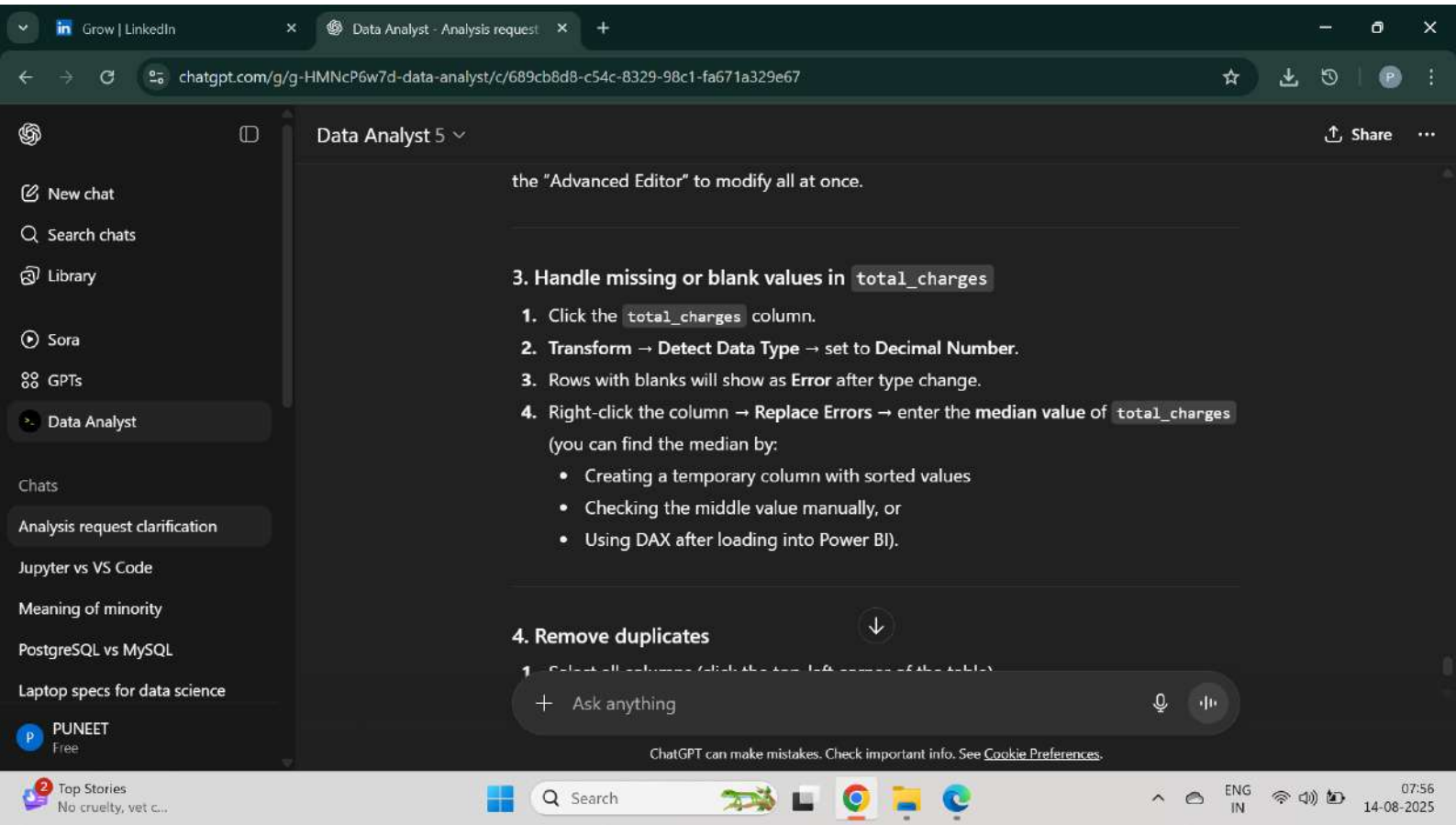
Team SaiKet Systems

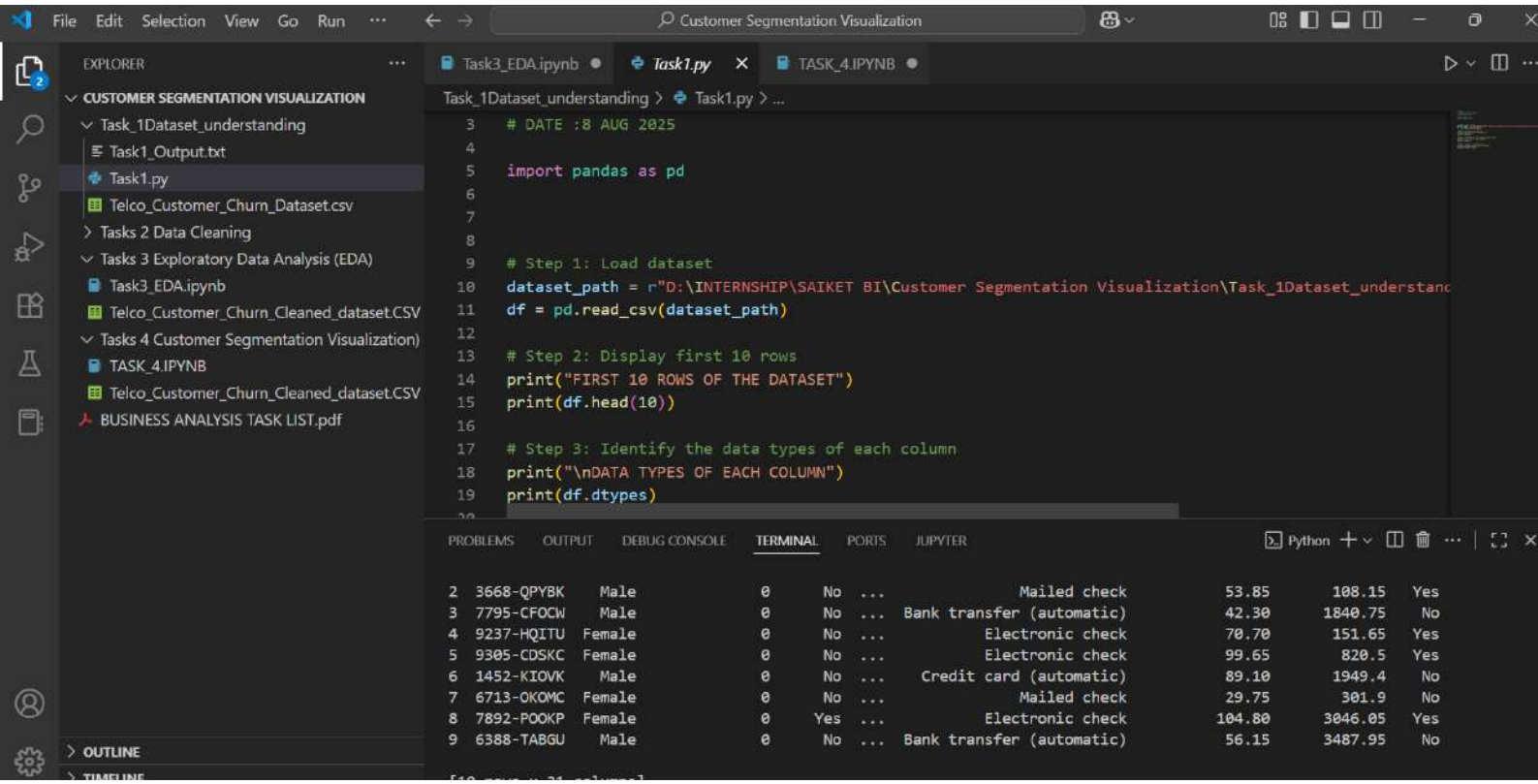
support@saiket.in www.saiket.in

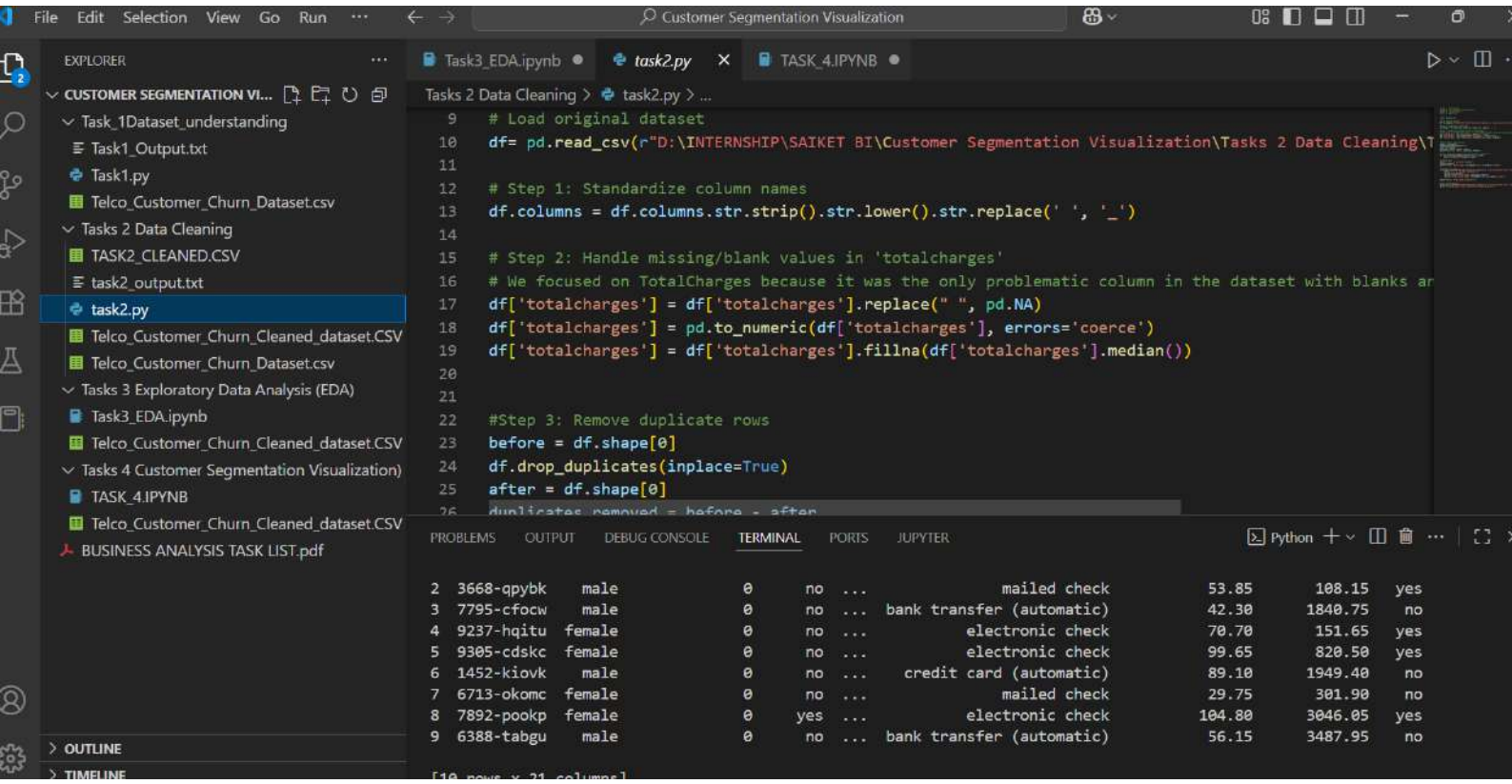
Hot days ahead 28°C

Search

ENG IN 23:32 13-08-2025


Grow | LinkedIn
Data Analyst - Analysis request
chatgpt.com/g/g-HMNcP6w7d-data-analyst/c/689cb8d8-c54c-8329-98c1-fa671a329e67
Data Analyst 5
Share
New chat
Search chats
Library
Sora
GPTs
Data Analyst
Chats
Analysis request clarification
Jupyter vs VS Code
Meaning of minority
PostgreSQL vs MySQL
Laptop specs for data science
PUNEET
Free
the "Advanced Editor" to modify all at once.
3. Handle missing or blank values in total_charges
1. Click the total_charges column.
2. Transform → Detect Data Type → set to Decimal Number.
3. Rows with blanks will show as Error after type change.
4. Right-click the column → Replace Errors → enter the median value of total_charges (you can find the median by:
• Creating a temporary column with sorted values
• Checking the middle value manually, or
• Using DAX after loading into Power BI).
4. Remove duplicates
Ask anything
ChatGPT can make mistakes. Check important info. See Cookie Preferences.
Top Stories
No cruelty, vet c...
Search
ENG
IN
07:56
14-08-2025

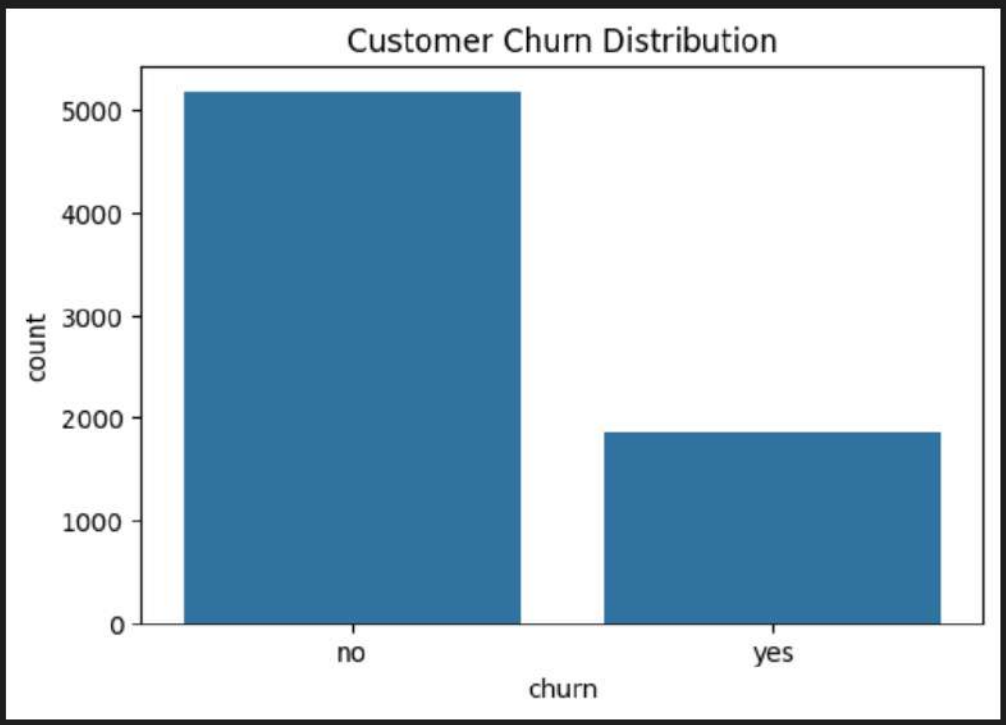

File Edit Selection View Go Run
Customer Segmentation Visualization
EXPLORER
CUSTOMER SEGMENTATION VISUALIZATION
Task_1Dataset_understanding
Task1_Output.txt
Task1.py
Telco_Customer_Churn_Dataset.csv
Tasks 2 Data Cleaning
Tasks 3 Exploratory Data Analysis (EDA)
Task3_EDA.ipynb
Telco_Customer_Churn_Cleaned_dataset.CSV
Tasks 4 Customer Segmentation Visualization)
TASK_4.IPYNB
Telco_Customer_Churn_Cleaned_dataset.CSV
BUSINESS ANALYSIS TASK LIST.pdf
OUTLINE
Task3_EDA.ipynb
Task1.py
TASK_4.IPYNB
Task_1Dataset_understanding > Task1.py > ...
# DATE :8 AUG 2025
import pandas as pd
# Step 1: Load dataset
dataset_path = r"D:\INTERNSHIP\SAIKET BI\Customer Segmentation Visualization\Task_1Dataset_understanc
df = pd.read_csv(dataset_path)
# Step 2: Display first 10 rows
print("FIRST 10 ROWS OF THE DATASET")
print(df.head(10))
# Step 3: Identify the data types of each column
print("\nDATA TYPES OF EACH COLUMN")
print(df.dtypes)
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS JUPYTER
Python
2 3668-QPYBK Male 0 No ... Mailed check 53.85 108.15 Yes
3 7795-CFOCW Male 0 No ... Bank transfer (automatic) 42.30 1840.75 No
4 9237-HQITU Female 0 No ... Electronic check 70.70 151.65 Yes
5 9305-CDSKC Female 0 No ... Electronic check 99.65 820.5 Yes
6 1452-KIOVK Male 0 No ... Credit card (automatic) 89.10 1949.4 No
7 6713-OKOMC Female 0 No ... Mailed check 29.75 301.9 No
8 7892-POOKP Female 0 Yes ... Electronic check 104.80 3046.05 Yes
9 6388-TABGU Male 0 No ... Bank transfer (automatic) 56.15 3487.95 No


File Edit Selection View Go Run
Customer Segmentation Visualization
EXPLORER
CUSTOMER SEGMENTATION VI...
Task_1Dataset_understanding
Task1_Output.txt
Task1.py
Telco_Customer_Churn_Dataset.csv
Tasks 2 Data Cleaning
TASK2_CLEANED.CSV
task2_output.txt
task2.py
Telco_Customer_Churn_Cleaned_dataset.CSV
Telco_Customer_Churn_Dataset.csv
Tasks 3 Exploratory Data Analysis (EDA)
Task3_EDA.ipynb
Telco_Customer_Churn_Cleaned_dataset.CSV
Tasks 4 Customer Segmentation Visualization)
TASK_4.IPYNB
Telco_Customer_Churn_Cleaned_dataset.CSV
BUSINESS ANALYSIS TASK LIST.pdf
OUTLINE
TIMELINE
Task3_EDA.ipynb
task2.py
TASK_4.IPYNB
Tasks 2 Data Cleaning > task2.py > ...
9 # Load original dataset
10 df= pd.read_csv(r"D:\INTERNSHIP\SAIKET BI\Customer Segmentation Visualization\Tasks 2 Data Cleaning\T
11
12 # Step 1: Standardize column names
13 df.columns = df.columns.str.strip().str.lower().str.replace(' ', '_')
14
15 # Step 2: Handle missing/blank values in 'totalcharges'
16 # We focused on TotalCharges because it was the only problematic column in the dataset with blanks ar
17 df['totalcharges'] = df['totalcharges'].replace(" ", pd.NA)
18 df['totalcharges'] = pd.to_numeric(df['totalcharges'], errors='coerce')
19 df['totalcharges'] = df['totalcharges'].fillna(df['totalcharges'].median())
20
21
22 #Step 3: Remove duplicate rows
23 before = df.shape[0]
24 df.drop_duplicates(inplace=True)
25 after = df.shape[0]
26 duplicates_removed = before - after
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS JUPYTER
Python
2 3668-qpybk male 0 no ... mailed check 53.85 108.15 yes
3 7795-cfocw male 0 no ... bank transfer (automatic) 42.30 1840.75 no
4 9237-hqitu female 0 no ... electronic check 70.70 151.65 yes
5 9305-cdskc female 0 no ... electronic check 99.65 820.50 yes
6 1452-kiovk male 0 no ... credit card (automatic) 89.10 1949.40 no
7 6713-okomc female 0 no ... mailed check 29.75 301.90 no
8 7892-pookp female 0 yes ... electronic check 104.80 3046.05 yes
9 6388-tabgu male 0 no ... bank transfer (automatic) 56.15 3487.95 no
[10 rows x 21 columns]

Task3_EDA.ipynb ● task2.py TASK_4.IPYNB ●

Tasks 3 Exploratory Data Analysis (EDA) > Task3_EDA.ipynb > Cell 1 – Import libraries & load dataset > # Correlation matrix (numeric feat

Generate + Code + Markdown | Run All Restart Execute Group 1 Execute Group 2 ··· Pyth

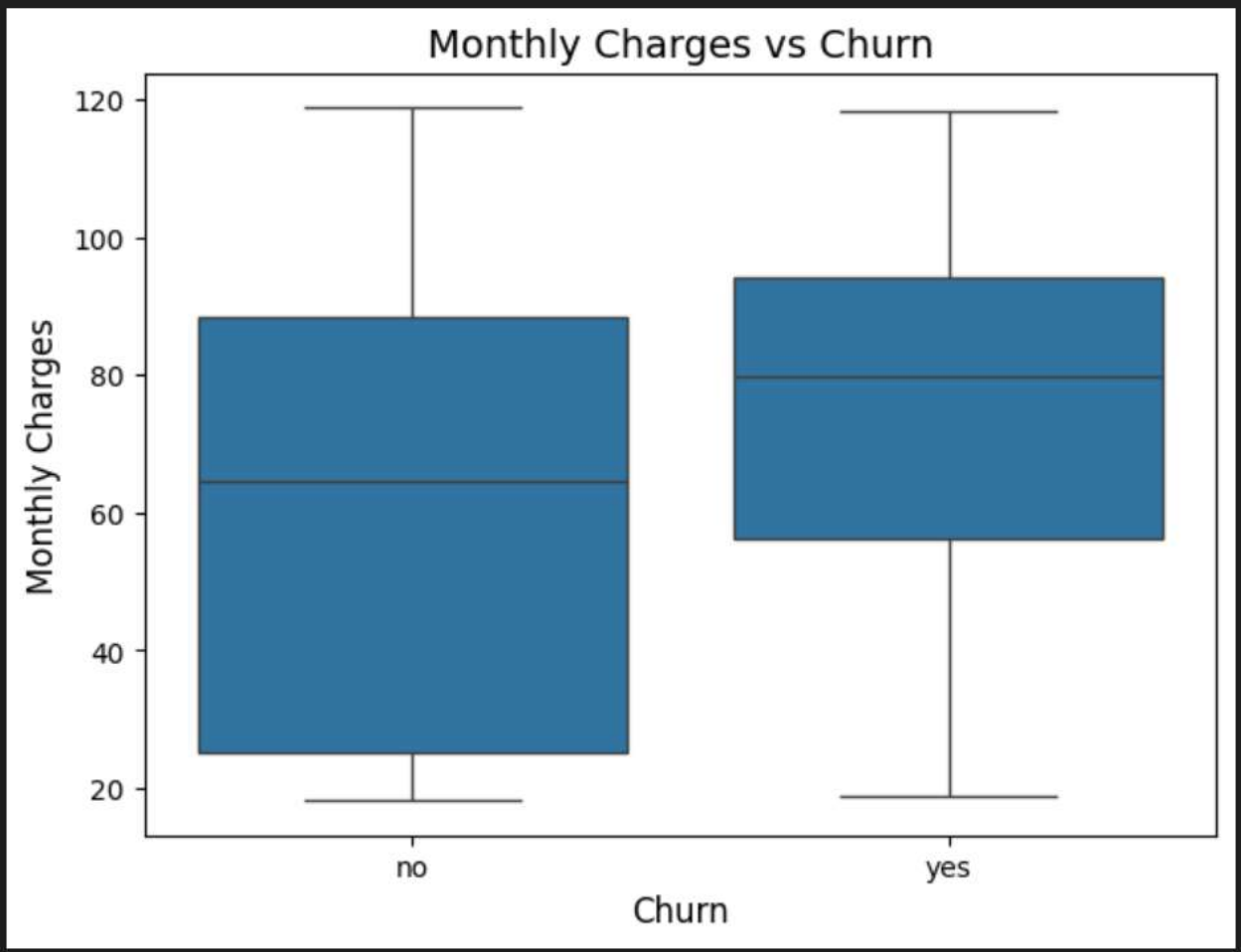


- Cell 3 – (Churn distribution)

asks 3 Exploratory Data Analysis (EDA) > Task3_EDA.ipynb > Cell 1 – Import libraries & load dataset > Cell 3 – (Churn distribution)

Generate + Code + Markdown | Run All Restart Execute Group 1 Execute Group 2 ...

6]



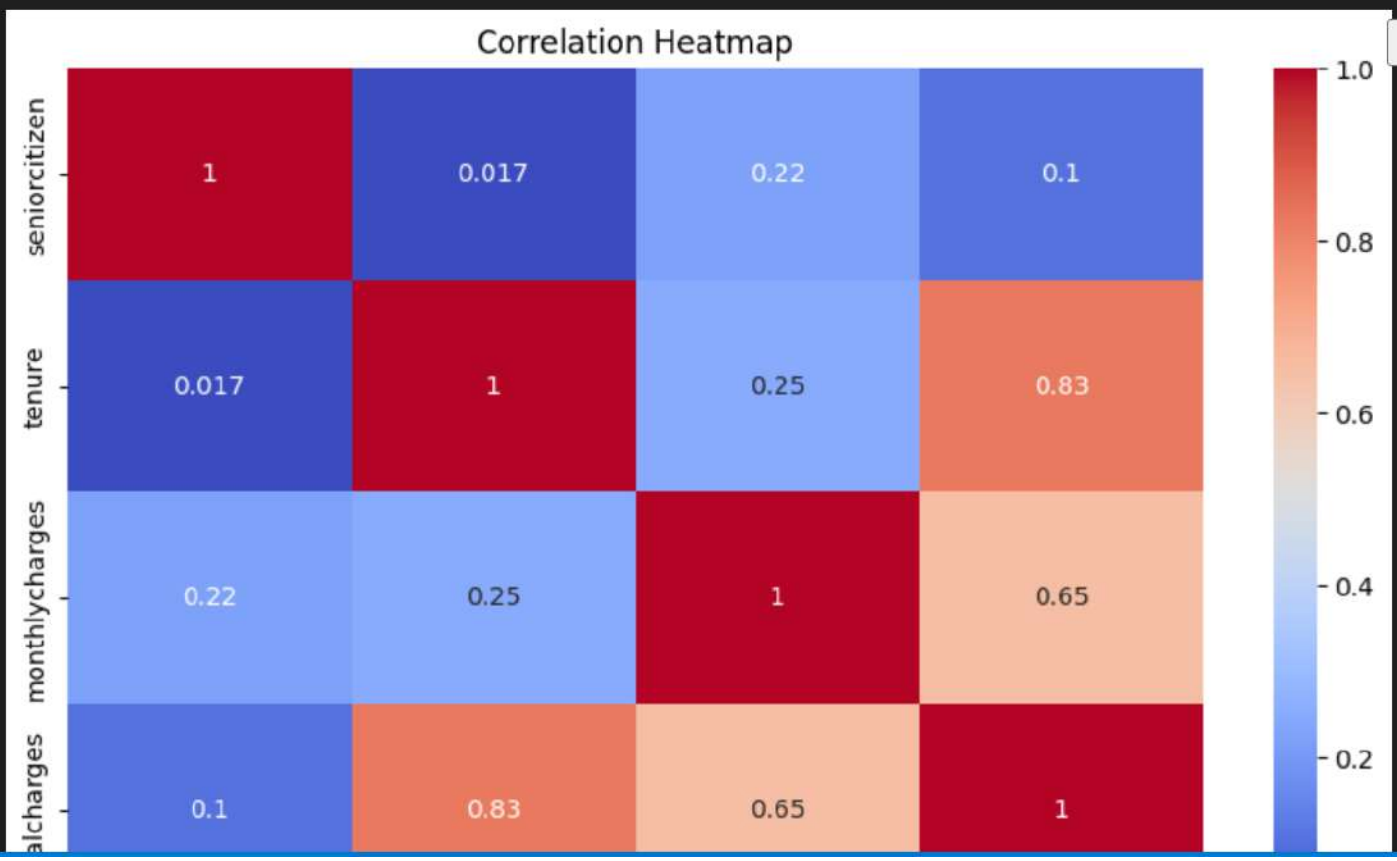
Task3_EDA.ipynb ● task2.py TASK_4.IPYNB ●

Tasks 3 Exploratory Data Analysis (EDA) > Task3_EDA.ipynb > Cell 1 – Import libraries & load dataset > Cell 3 – (Churn distribution)

Generate + Code + Markdown | Run All Restart Execute Group 1 Execute Group 2 ···

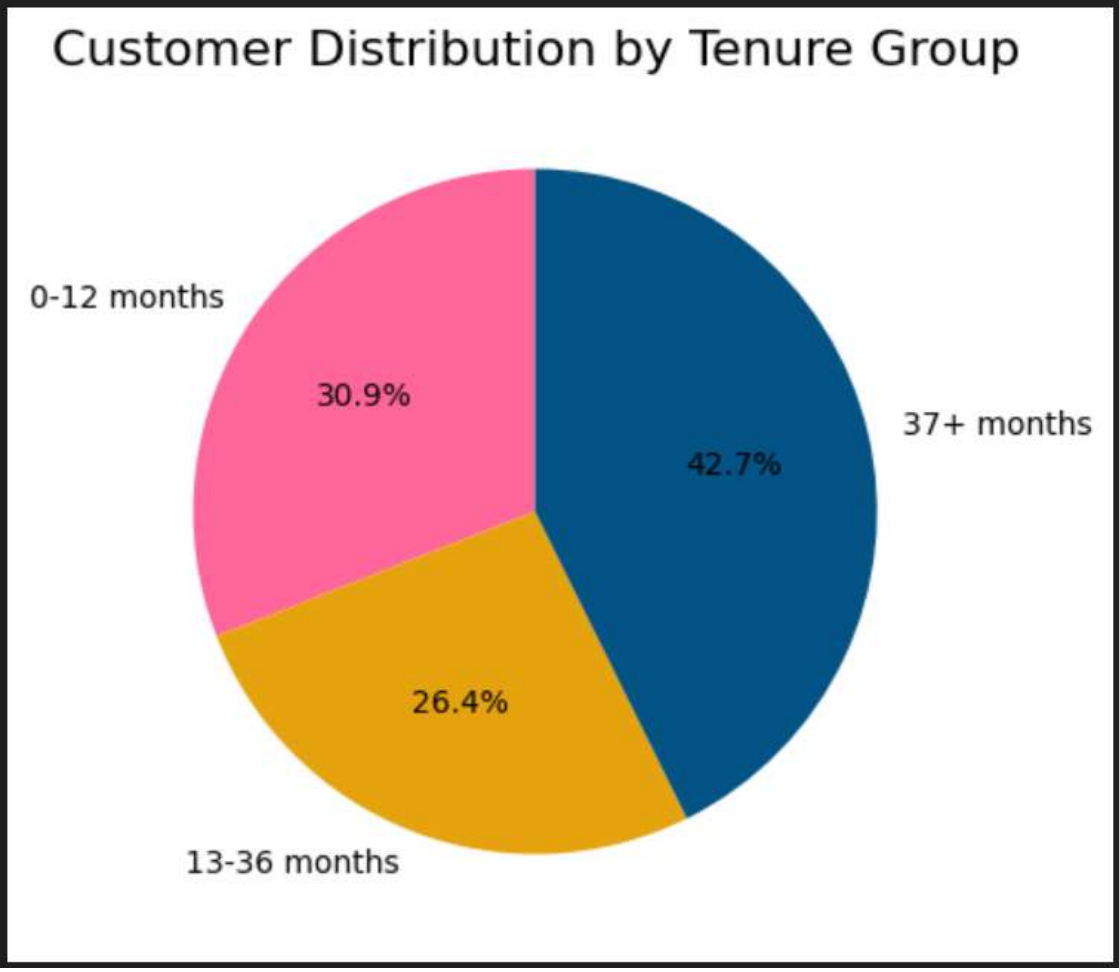
```
plt.show()
```

[7]



```
plt.show()
```

[3]


Customer Distribution by Tenure Group
0-12 months
30.9%
37+ months
42.7%
26.4%
13-36 months

Customer Segmentation Visualization

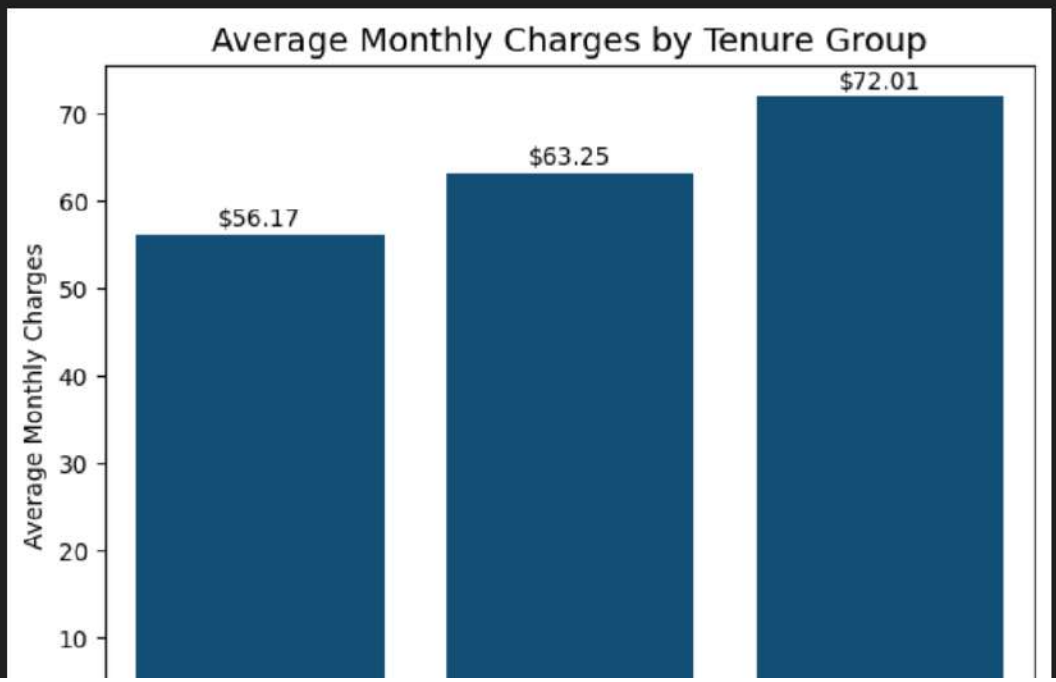
Task3_EDA.ipynb ● task2.py TASK_4.IPYNB ●

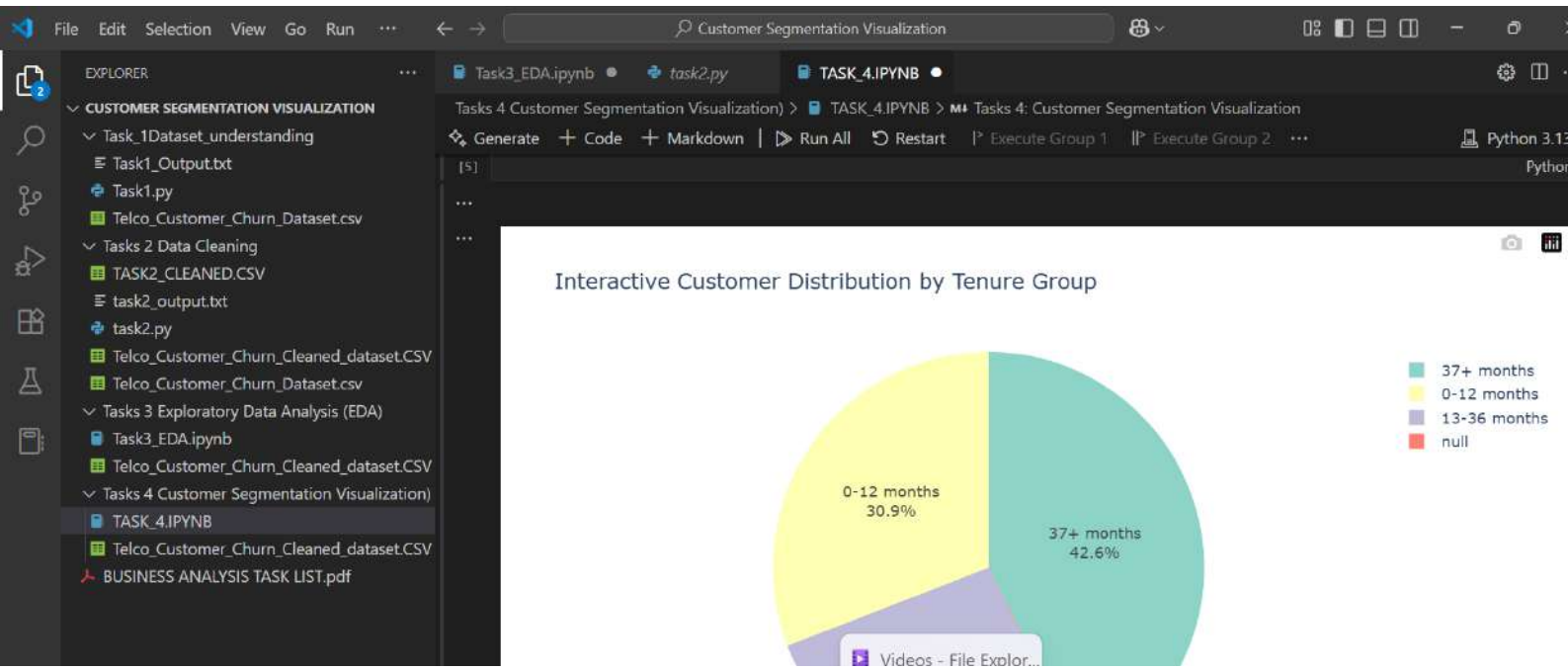
Tasks 4 Customer Segmentation Visualization) > TASK_4.IPYNB > M↓ Tasks 4: Customer Segmentation Visualization

Generate + Code + Markdown | Run All Restart Execute Group 1 Execute Group 2 ··· Python 3

[4]

```
C:\Users\poorn\AppData\Local\Temp\ipykernel_12168\624412221.py:2: FutureWarning: The default of observed=Fals
  avg_charges = df.groupby('tenure_group')['monthlycharges'].mean().reset_index()
```




File Edit Selection View Go Run
Customer Segmentation Visualization
EXPLORER
CUSTOMER SEGMENTATION VISUALIZATION
Task_1Dataset_understanding
Task1_Output.txt
Task1.py
Telco_Customer_Churn_Dataset.csv
Tasks 2 Data Cleaning
TASK2_CLEANED.CSV
task2_output.txt
task2.py
Telco_Customer_Churn_Cleaned_dataset.CSV
Telco_Customer_Churn_Dataset.csv
Tasks 3 Exploratory Data Analysis (EDA)
Task3_EDA.ipynb
Telco_Customer_Churn_Cleaned_dataset.CSV
Tasks 4 Customer Segmentation Visualization)
TASK_4.IPYNB
Telco_Customer_Churn_Cleaned_dataset.CSV
BUSINESS ANALYSIS TASK LIST.pdf
Task3_EDA.ipynb
task2.py
TASK_4.IPYNB
Tasks 4 Customer Segmentation Visualization) > TASK_4.IPYNB > Tasks 4: Customer Segmentation Visualization
Generate Code Markdown Run All Restart Execute Group 1 Execute Group 2
Python 3.13
[5]
Interactive Customer Distribution by Tenure Group
37+ months
0-12 months
13-36 months
null
0-12 months
30.9%
37+ months
42.6%
Videos - File Explor...