```python
# installing the essential libraries
!pip install PyPDF2 mysql-connector-python
```

Requirement already satisfied: PyPDF2 in c:\users\asus\anaconda3\lib\site-packages (3.0.1)
Requirement already satisfied: mysql-connector-python in c:\users\asus\anaconda3\lib\site-packages (9.0.0)

Task1 :- Extraction of pdf file

```python
import PyPDF2

def extraction_of_pdf_to_text(path):

    # Extract text from a PDF file and return it as a list of lines.

    # Parameters:
    # path (str): The file path of the PDF document.

    # Returns:
    # list: A list containing each line of extracted text from the PDF.


    # Open the PDF file in binary read mode
    with open(path, 'rb') as pdf_file:
        # Create a PdfReader object to read the PDF
        reader = PyPDF2.PdfReader(pdf_file)

        # Initialize an empty string to accumulate the extracted text
        text = ""

        # Iterate through each page in the PDF
        for page_num in range(len(reader.pages)):
            # Get the specific page object using the page number
            page = reader.pages[page_num]

            # Extract text from the current page and append it to the text variable
            text = text + page.extract_text() or ""  # Use or "" to avoid NoneType errors

        # Split the accumulated text into lines and return as a list
        return text.split("\n")  # Split text into lines based on newline characters
```

```python
reader = PyPDF2.PdfReader("sample_bank_statement.pdf")
print(len(reader.pages))
```

2

```python
to_text = extraction_of_pdf_to_text("sample_bank_statement.pdf")
print(to_text)
```

['Zenith National Cooperative Bank', 'Reg. No. : ZNC/BLR/RSR/CR/159/03', 'Branch : Indiranagar', 'Statement Of Saving Account', 'Current Date : 30/06/2025', 'Transaction Details From Date : 01/01/2025 To Date : 30/06/2025', 'Name : MEHT A ARYAN D EVENDRA', 'JOINTHOLDER1 : MEHT A PRIY A ARYAN', 'A/c No. : 7024186', 'A/c Start Date : 15/09/2020', 'Interest Rate % : 4.7 5', 'IFSC Code : ZNCB0004567', 'Address : EVERGREEN RESIDENCY , FLA T 604, 12TH MAIN, INDIRANAGAR, BANGALORE', 'Mobile No.: 9876543210', 'Date Particular Debit Credit Balance', '01/01/2025 Opening Balance 0.00 0.00 52,415.75 Cr', '02/01/2025 TO TR F UPI/New Year Party Expenses 3,500.00 0.00 48,915.75 Cr', '05/01/2025 BY TRF Salary Credit 0.00 65,000.00 113,915.75 Cr', '07/01/2025 TO TRF UPI/Grocery Shopping 2,800.00 0.00 111,115.75 Cr', '10/01/2025 TO TRF Rent Payment 25,000.00 0.00 86,11 5.75 Cr', '12/01/2025 TO TRF UPI/Electricity Bill 1,850.00 0.00 84,265.75 Cr', '15/01/2025 TO TRF Credit Card Bill 12,500.0 0 0.00 71,765.75 Cr', '18/01/2025 TO TRF UPI/Dining Out 2,200.00 0.00 69,565.75 Cr', '20/01/2025 TO TRF UPI/Mobile Recharge 999.00 0.00 68,566.75 Cr', '25/01/2025 TO TRF UPI/Online Shopping 3,500.00 0.00 65,066.75 Cr', '28/01/2025 TO TRF UPI/Gym M embership 2,000.00 0.00 63,066.75 Cr', '31/01/2025 BY TRF Interest Credit 0.00 249.00 63,315.75 Cr', '02/02/2025 TO TRF UP I/W eekend Getaway 5,500.00 0.00 57,815.75 Cr', '05/02/2025 BY TRF Salary Credit 0.00 65,000.00 122,815.75 Cr', '07/02/2025 TO TRF UPI/Grocery Shopping 3,200.00 0.00 119,615.75 Cr', '10/02/2025 TO TRF Rent Payment 25,000.00 0.00 94,615.75 Cr', '1 2/02/2025 TO TRF Insurance Premium 8,500.00 0.00 86,115.75 Cr', '15/02/2025 TO TRF Credit Card Bill 15,000.00 0.00 71,115.7 5 Cr', "18/02/2025 TO TRF UPI/V alentine's Day Gift 4,000.00 0.00 67,115.75 Cr", '20/02/2025 TO TRF UPI/Car Service 3,500.0 0 0.00 63,615.75 Cr', '23/02/2025 BY TRF UPI/Freelance W ork 0.00 10,000.00 73,615.75 Cr', '25/02/2025 TO TRF UPI/Home Deco r Items 6,000.00 0.00 67,615.75 Cr', '28/02/2025 BY TRF Interest Credit 0.00 267.00 67,882.75 Cr', '02/03/2025 TO TRF UPI/M ovie Night 1,500.00 0.00 66,382.75 Cr', '05/03/2025 BY TRF Salary Credit 0.00 65,000.00 131,382.75 Cr', '07/03/2025 TO TRF UPI/Grocery Shopping 3,500.00 0.00 127,882.75 Cr', '10/03/2025 TO TRF Rent Payment 25,000.00 0.00 102,882.75 Cr', '12/03/20 25 TO TRF UPI/Electricity Bill 2,100.00 0.00 100,782.75 Cr', '15/03/2025 TO TRF Car Loan EMI 18,500.00 0.00 82,282.75 Cr', '18/03/2025 TO TRF UPI/Birthday Gift 2,500.00 0.00 79,782.75 Cr', '20/03/2025 TO TRF UPI/Online Shopping 7,250.00 0.00 72,5 32.75 CrDate Particular Debit Credit Balance', '23/03/2025 TO TRF UPI/Dental Appointment 3,000.00 0.00 69,532.75 Cr', '25/0 3/2025 BY TRF Tax Refund 0.00 8,500.00 78,032.75 Cr', '28/03/2025 TO TRF UPI/Home Repairs 4,500.00 0.00 73,532.75 Cr', '31/ 03/2025 BY TRF Interest Credit 0.00 290.00 73,822.75 Cr', '02/04/2025 TO TRF UPI/Charity Donation 2,000.00 0.00 71,822.75 C r', '05/04/2025 BY TRF Salary Credit 0.00 68,000.00 139,822.75 Cr', '07/04/2025 TO TRF UPI/Grocery Shopping 3,800.00 0.00 1 36,022.75 Cr', '10/04/2025 TO TRF Rent Payment 25,000.00 0.00 111,022.75 Cr', '12/04/2025 TO TRF UPI/Electricity Bill 2,30 0.00 0.00 108,722.75 Cr', '15/04/2025 TO TRF Credit Card Bill 20,000.00 0.00 88,722.75 Cr', '18/04/2025 TO TRF UPI/New Smar tphone 15,000.00 0.00 73,722.75 Cr', '20/04/2025 TO TRF UPI/W edding Gift 5,000.00 0.00 68,722.75 Cr', '23/04/2025 BY TRF U PI/Freelance W ork 0.00 12,000.00 80,722.75 Cr', '25/04/2025 TO TRF UPI/Home Appliance 8,000.00 0.00 72,722.75 Cr', '28/04/ 2025 TO TRF UPI/Car Insurance 6,500.00 0.00 66,222.75 Cr', '30/04/2025 BY TRF Interest Credit 0.00 261.00 66,483.75 Cr', '0 2/05/2025 TO TRF UPI/W eekend Trip 7,000.00 0.00 59,483.75 Cr', '05/05/2025 BY TRF Salary Credit 0.00 68,000.00 127,483.75 Cr', '07/05/2025 TO TRF UPI/Grocery Shopping 4,000.00 0.00 123,483.75 Cr', '10/05/2025 TO TRF Rent Payment 25,000.00 0.00 9 8,483.75 Cr', '12/05/2025 TO TRF Family V acation Expenses 75,000.00 0.00 23']

Removing initial rows that are not required

```python
def remove_first_rows(lines, num_rows):
    return lines[num_rows:]
pdf_lines_cleaned = remove_first_rows(to_text, 15)
pdf_lines_cleaned
```

```
['01/01/2025 Opening Balance 0.00 0.00 52,415.75 Cr',
 '02/01/2025 TO TRF UPI/New Year Party Expenses 3,500.00 0.00 48,915.75 Cr',
 '05/01/2025 BY TRF Salary Credit 0.00 65,000.00 113,915.75 Cr',
 '07/01/2025 TO TRF UPI/Grocery Shopping 2,800.00 0.00 111,115.75 Cr',
 '10/01/2025 TO TRF Rent Payment 25,000.00 0.00 86,115.75 Cr',
 '12/01/2025 TO TRF UPI/Electricity Bill 1,850.00 0.00 84,265.75 Cr',
 '15/01/2025 TO TRF Credit Card Bill 12,500.00 0.00 71,765.75 Cr',
 '18/01/2025 TO TRF UPI/Dining Out 2,200.00 0.00 69,565.75 Cr',
 '20/01/2025 TO TRF UPI/Mobile Recharge 999.00 0.00 68,566.75 Cr',
 '25/01/2025 TO TRF UPI/Online Shopping 3,500.00 0.00 65,066.75 Cr',
 '28/01/2025 TO TRF UPI/Gym Membership 2,000.00 0.00 63,066.75 Cr',
 '31/01/2025 BY TRF Interest Credit 0.00 249.00 63,315.75 Cr',
 '02/02/2025 TO TRF UPI/W eekend Getaway 5,500.00 0.00 57,815.75 Cr',
 '05/02/2025 BY TRF Salary Credit 0.00 65,000.00 122,815.75 Cr',
 '07/02/2025 TO TRF UPI/Grocery Shopping 3,200.00 0.00 119,615.75 Cr',
 '10/02/2025 TO TRF Rent Payment 25,000.00 0.00 94,615.75 Cr',
 '12/02/2025 TO TRF Insurance Premium 8,500.00 0.00 86,115.75 Cr',
 '15/02/2025 TO TRF Credit Card Bill 15,000.00 0.00 71,115.75 Cr',
 "18/02/2025 TO TRF UPI/V alentine's Day Gift 4,000.00 0.00 67,115.75 Cr",
 '20/02/2025 TO TRF UPI/Car Service 3,500.00 0.00 63,615.75 Cr',
 '23/02/2025 BY TRF UPI/Freelance W ork 0.00 10,000.00 73,615.75 Cr',
 '25/02/2025 TO TRF UPI/Home Decor Items 6,000.00 0.00 67,615.75 Cr',
 '28/02/2025 BY TRF Interest Credit 0.00 267.00 67,882.75 Cr',
 '02/03/2025 TO TRF UPI/Movie Night 1,500.00 0.00 66,382.75 Cr',
 '05/03/2025 BY TRF Salary Credit 0.00 65,000.00 131,382.75 Cr',
 '07/03/2025 TO TRF UPI/Grocery Shopping 3,500.00 0.00 127,882.75 Cr',
 '10/03/2025 TO TRF Rent Payment 25,000.00 0.00 102,882.75 Cr',
 '12/03/2025 TO TRF UPI/Electricity Bill 2,100.00 0.00 100,782.75 Cr',
 '15/03/2025 TO TRF Car Loan EMI 18,500.00 0.00 82,282.75 Cr',
 '18/03/2025 TO TRF UPI/Birthday Gift 2,500.00 0.00 79,782.75 Cr',
 '20/03/2025 TO TRF UPI/Online Shopping 7,250.00 0.00 72,532.75 CrDate Particular Debit Credit Balance',
 '23/03/2025 TO TRF UPI/Dental Appointment 3,000.00 0.00 69,532.75 Cr',
 '25/03/2025 BY TRF Tax Refund 0.00 8,500.00 78,032.75 Cr',
 '28/03/2025 TO TRF UPI/Home Repairs 4,500.00 0.00 73,532.75 Cr',
 '31/03/2025 BY TRF Interest Credit 0.00 290.00 73,822.75 Cr',
 '02/04/2025 TO TRF UPI/Charity Donation 2,000.00 0.00 71,822.75 Cr',
 '05/04/2025 BY TRF Salary Credit 0.00 68,000.00 139,822.75 Cr',
 '07/04/2025 TO TRF UPI/Grocery Shopping 3,800.00 0.00 136,022.75 Cr',
 '10/04/2025 TO TRF Rent Payment 25,000.00 0.00 111,022.75 Cr',
 '12/04/2025 TO TRF UPI/Electricity Bill 2,300.00 0.00 108,722.75 Cr',
 '15/04/2025 TO TRF Credit Card Bill 20,000.00 0.00 88,722.75 Cr',
 '18/04/2025 TO TRF UPI/New Smartphone 15,000.00 0.00 73,722.75 Cr',
 '20/04/2025 TO TRF UPI/W edding Gift 5,000.00 0.00 68,722.75 Cr',
 '23/04/2025 BY TRF UPI/Freelance W ork 0.00 12,000.00 80,722.75 Cr',
 '25/04/2025 TO TRF UPI/Home Appliance 8,000.00 0.00 72,722.75 Cr',
 '28/04/2025 TO TRF UPI/Car Insurance 6,500.00 0.00 66,222.75 Cr',
 '30/04/2025 BY TRF Interest Credit 0.00 261.00 66,483.75 Cr',
 '02/05/2025 TO TRF UPI/W eekend Trip 7,000.00 0.00 59,483.75 Cr',
 '05/05/2025 BY TRF Salary Credit 0.00 68,000.00 127,483.75 Cr',
 '07/05/2025 TO TRF UPI/Grocery Shopping 4,000.00 0.00 123,483.75 Cr',
 '10/05/2025 TO TRF Rent Payment 25,000.00 0.00 98,483.75 Cr',
 '12/05/2025 TO TRF Family V acation Expenses 75,000.00 0.00 23']
```

```python
# Importing Pandas library
import pandas as pd
```

Splitting the transactions

```python
def split_transaction(line):
    # Splits a transaction line into its individual components.

    # Split the line at spaces
    parts = line.split()

    # Date is the first part
    date = parts[0]

    # Amount is second to last, transaction type is the last part
    cr = parts[-1]
    balance = parts[-2]
    total_balance = " ".join(parts[-2:])
    credit = parts[-3]   # Credit amount
    debit = parts[-4]    # Debit amount

    # Everything in between is the description
    particular = " ".join(parts[1:-4])   # Join the description parts

    return [date, particular, debit, credit, total_balance]
```

```python
# Apply the split_transaction function to each line in data
split_data = [split_transaction(line) for line in pdf_lines_cleaned]
# This line uses a list comprehension to apply the split_transaction function to each line
# and store the results in the split_data list

# Create a DataFrame with the appropriate columns
df = pd.DataFrame(split_data, columns=['date', 'particular', 'debit', 'credit', 'total_balance'])
# Create a pandas DataFrame using the split_data list and specify the column names

df # printing the df
```

	date	particular	debit	credit	total_balance
0	01/01/2025	Opening Balance	0.00	0.00	52,415.75 Cr
1	02/01/2025	TO TRF UPI/New Year Party Expenses	3,500.00	0.00	48,915.75 Cr
2	05/01/2025	BY TRF Salary Credit	0.00	65,000.00	113,915.75 Cr
3	07/01/2025	TO TRF UPI/Grocery Shopping	2,800.00	0.00	111,115.75 Cr
4	10/01/2025	TO TRF Rent Payment	25,000.00	0.00	86,115.75 Cr
5	12/01/2025	TO TRF UPI/Electricity Bill	1,850.00	0.00	84,265.75 Cr
6	15/01/2025	TO TRF Credit Card Bill	12,500.00	0.00	71,765.75 Cr
7	18/01/2025	TO TRF UPI/Dining Out	2,200.00	0.00	69,565.75 Cr
8	20/01/2025	TO TRF UPI/Mobile Recharge	999.00	0.00	68,566.75 Cr
9	25/01/2025	TO TRF UPI/Online Shopping	3,500.00	0.00	65,066.75 Cr
10	28/01/2025	TO TRF UPI/Gym Membership	2,000.00	0.00	63,066.75 Cr
11	31/01/2025	BY TRF Interest Credit	0.00	249.00	63,315.75 Cr
12	02/02/2025	TO TRF UPI/W eekend Getaway	5,500.00	0.00	57,815.75 Cr
13	05/02/2025	BY TRF Salary Credit	0.00	65,000.00	122,815.75 Cr
14	07/02/2025	TO TRF UPI/Grocery Shopping	3,200.00	0.00	119,615.75 Cr
15	10/02/2025	TO TRF Rent Payment	25,000.00	0.00	94,615.75 Cr
16	12/02/2025	TO TRF Insurance Premium	8,500.00	0.00	86,115.75 Cr
17	15/02/2025	TO TRF Credit Card Bill	15,000.00	0.00	71,115.75 Cr
18	18/02/2025	TO TRF UPI/V alentine's Day Gift	4,000.00	0.00	67,115.75 Cr
19	20/02/2025	TO TRF UPI/Car Service	3,500.00	0.00	63,615.75 Cr
20	23/02/2025	BY TRF UPI/Freelance W ork	0.00	10,000.00	73,615.75 Cr
21	25/02/2025	TO TRF UPI/Home Decor Items	6,000.00	0.00	67,615.75 Cr
22	28/02/2025	BY TRF Interest Credit	0.00	267.00	67,882.75 Cr
23	02/03/2025	TO TRF UPI/Movie Night	1,500.00	0.00	66,382.75 Cr
24	05/03/2025	BY TRF Salary Credit	0.00	65,000.00	131,382.75 Cr
25	07/03/2025	TO TRF UPI/Grocery Shopping	3,500.00	0.00	127,882.75 Cr
26	10/03/2025	TO TRF Rent Payment	25,000.00	0.00	102,882.75 Cr
27	12/03/2025	TO TRF UPI/Electricity Bill	2,100.00	0.00	100,782.75 Cr
28	15/03/2025	TO TRF Car Loan EMI	18,500.00	0.00	82,282.75 Cr
29	18/03/2025	TO TRF UPI/Birthday Gift	2,500.00	0.00	79,782.75 Cr
30	20/03/2025	TO TRF UPI/Online Shopping 7,250.00 0.00 72,53...	Particular	Debit	Credit Balance
31	23/03/2025	TO TRF UPI/Dental Appointment	3,000.00	0.00	69,532.75 Cr
32	25/03/2025	BY TRF Tax Refund	0.00	8,500.00	78,032.75 Cr
33	28/03/2025	TO TRF UPI/Home Repairs	4,500.00	0.00	73,532.75 Cr
34	31/03/2025	BY TRF Interest Credit	0.00	290.00	73,822.75 Cr
35	02/04/2025	TO TRF UPI/Charity Donation	2,000.00	0.00	71,822.75 Cr
36	05/04/2025	BY TRF Salary Credit	0.00	68,000.00	139,822.75 Cr
37	07/04/2025	TO TRF UPI/Grocery Shopping	3,800.00	0.00	136,022.75 Cr
38	10/04/2025	TO TRF Rent Payment	25,000.00	0.00	111,022.75 Cr
39	12/04/2025	TO TRF UPI/Electricity Bill	2,300.00	0.00	108,722.75 Cr
40	15/04/2025	TO TRF Credit Card Bill	20,000.00	0.00	88,722.75 Cr
41	18/04/2025	TO TRF UPI/New Smartphone	15,000.00	0.00	73,722.75 Cr
42	20/04/2025	TO TRF UPI/W edding Gift	5,000.00	0.00	68,722.75 Cr
43	23/04/2025	BY TRF UPI/Freelance W ork	0.00	12,000.00	80,722.75 Cr
44	25/04/2025	TO TRF UPI/Home Appliance	8,000.00	0.00	72,722.75 Cr
45	28/04/2025	TO TRF UPI/Car Insurance	6,500.00	0.00	66,222.75 Cr
46	30/04/2025	BY TRF Interest Credit	0.00	261.00	66,483.75 Cr

	date	particular	debit	credit	total_balance
47	02/05/2025	TO TRF UPI/W eekend Trip	7,000.00	0.00	59,483.75 Cr
48	05/05/2025	BY TRF Salary Credit	0.00	68,000.00	127,483.75 Cr
49	07/05/2025	TO TRF UPI/Grocery Shopping	4,000.00	0.00	123,483.75 Cr
50	10/05/2025	TO TRF Rent Payment	25,000.00	0.00	98,483.75 Cr
51	12/05/2025	TO TRF Family V acation	Expenses	75,000.00	0.00 23

In the above df we can see that at the 30th index value is overlapping with the next page's title so in next we are changing that

In [168...
```python
#  Update the values at index 30
df['particular'][30] = 'TO TRF UPI/Online Shopping'
df['debit'][30] = '7,250.00'
df['credit'][30] = '0.00'
df['total_balance'][30] = '72,532.75 Cr'
```

In [169...
```python
#  Update the values at index 30
df['particular'][51] = 'TO TRF Family Vacation Expenses'
df['debit'][51] = '75,000.00'
df['credit'][51] = '0.00'
df['total_balance'][51] = '23 Cr'
```

In [170...
```python
# converting debit and credit columns with float data type
df['debit'] = df['debit'].str.replace(',', '').astype(float)
df['credit'] = df['credit'].str.replace(',', '').astype(float)

df['date'] = pd.to_datetime(df['date'], format='%d/%m/%Y').dt.strftime('%Y-%m-%d')
```

Generating unique id to each transaction

In [172...
```python
# As expected in assignment we don't have unique_id in pdf file so we are creating that

import uuid

# Function to generate unique IDs
def generate_unique_id():
    return str(uuid.uuid4())

# Apply the function to create a list of unique IDs
df['unique_id'] = [generate_unique_id() for _ in range(len(df))]

# Ensure DataFrame is processed from the first row
df.reset_index(drop=True, inplace=True)
```

In [173...
```python
# requirement from the pdf we are creating single column for the debit and credit

df['transaction_type'] = df.apply(lambda row: 'Credit' if row['debit'] == 0.00 else 'Debit', axis=1)

df['amount'] = df.apply(lambda row: row['debit'] if row['credit'] == 0.00 else row['credit'], axis=1)
df
df1 = pd.DataFrame(df, columns=['unique_id','date', 'particular', 'amount', 'transaction_type'])
df1
```

	unique_id	date	particular	amount	transaction_type
0	841253a4-dd36-4a6c-8117-f6068753dddf	2025-01-01	Opening Balance	0.0	Credit
1	30b67bac-348f-4a29-89d0-9b0074a6f3ff	2025-01-02	TO TRF UPI/New Year Party Expenses	3500.0	Debit
2	f607e090-0d46-4398-ab74-75b901501077	2025-01-05	BY TRF Salary Credit	65000.0	Credit
3	244f7399-f44c-4966-bff3-775ecff6f993	2025-01-07	TO TRF UPI/Grocery Shopping	2800.0	Debit
4	ef8e701b-8a94-49aa-82ea-d73156769766	2025-01-10	TO TRF Rent Payment	25000.0	Debit
5	ecc52839-6c55-4336-846c-da4ef7d0aef4	2025-01-12	TO TRF UPI/Electricity Bill	1850.0	Debit
6	0a0a64cf-e64f-4678-aaa0-768318c88798	2025-01-15	TO TRF Credit Card Bill	12500.0	Debit
7	98f24f85-ba0a-4a67-a8ad-cc618a7f8213	2025-01-18	TO TRF UPI/Dining Out	2200.0	Debit
8	8d013f1e-9b11-4aff-b73a-d66f92b43d46	2025-01-20	TO TRF UPI/Mobile Recharge	999.0	Debit
9	cc10537c-b469-4baf-af70-475499231230	2025-01-25	TO TRF UPI/Online Shopping	3500.0	Debit
10	29d37a82-9b26-4ead-8239-4572ad69d932	2025-01-28	TO TRF UPI/Gym Membership	2000.0	Debit
11	3c3e3209-3428-4f76-9f39-0eb0a8360ec6	2025-01-31	BY TRF Interest Credit	249.0	Credit
12	7f6e3720-460c-475c-8f40-16514e3f87e3	2025-02-02	TO TRF UPI/W eekend Getaway	5500.0	Debit
13	e0f410ec-0830-490a-bcc1-65acca502113	2025-02-05	BY TRF Salary Credit	65000.0	Credit
14	d88a86a2-d7ae-42d0-8d09-c33159676efb	2025-02-07	TO TRF UPI/Grocery Shopping	3200.0	Debit
15	2db0aef8-4c50-4859-882e-d7f29bdb65fe	2025-02-10	TO TRF Rent Payment	25000.0	Debit
16	63d74ecb-02c3-4c8f-9310-1c494347f5d5	2025-02-12	TO TRF Insurance Premium	8500.0	Debit
17	33c42244-afc5-433b-a601-9ad528287737	2025-02-15	TO TRF Credit Card Bill	15000.0	Debit
18	24c64049-4329-48e9-8f29-32b4cbbe7a77	2025-02-18	TO TRF UPI/V alentine's Day Gift	4000.0	Debit
19	83095003-b2ee-4fac-b0d0-634171917a41	2025-02-20	TO TRF UPI/Car Service	3500.0	Debit
20	16933ff3-ba4f-432e-8fab-b65f89f00157	2025-02-23	BY TRF UPI/Freelance W ork	10000.0	Credit
21	00d305d6-4e20-4eb9-b74f-1c6a810814dd	2025-02-25	TO TRF UPI/Home Decor Items	6000.0	Debit
22	afb33c51-7647-4f43-b5af-743725166ef6	2025-02-28	BY TRF Interest Credit	267.0	Credit
23	0a15388c-393c-457c-8305-467c12835765	2025-03-02	TO TRF UPI/Movie Night	1500.0	Debit
24	4529971b-68b9-48e0-a849-3d48d48c01bd	2025-03-05	BY TRF Salary Credit	65000.0	Credit
25	47bf9f85-919e-4397-ba24-722cb4b86c94	2025-03-07	TO TRF UPI/Grocery Shopping	3500.0	Debit
26	97c91bd0-4b29-4869-8b6b-f3532228c07e	2025-03-10	TO TRF Rent Payment	25000.0	Debit
27	4762f60d-9101-429b-a6eb-2e46847d6414	2025-03-12	TO TRF UPI/Electricity Bill	2100.0	Debit
28	48248e4d-1d4f-472f-8496-b6be07eb273a	2025-03-15	TO TRF Car Loan EMI	18500.0	Debit
29	efd061aa-93d7-4214-b651-522d62735a3a	2025-03-18	TO TRF UPI/Birthday Gift	2500.0	Debit
30	86928d79-28fc-4885-9c02-7ea3828cbce8	2025-03-20	TO TRF UPI/Online Shopping	7250.0	Debit
31	a70abe55-4fe1-408e-a7af-0c3e3b4f6c90	2025-03-23	TO TRF UPI/Dental Appointment	3000.0	Debit
32	898ceb04-0f97-4cee-b72c-4f0c2a6b835a	2025-03-25	BY TRF Tax Refund	8500.0	Credit
33	6bcd7e1b-7f06-498a-bd4a-c105ec2519e7	2025-03-28	TO TRF UPI/Home Repairs	4500.0	Debit
34	abb7a2dd-ff00-47d9-8603-7056a5b93732	2025-03-31	BY TRF Interest Credit	290.0	Credit
35	6735f3b9-e311-4621-818b-da2f89d4b3dd	2025-04-02	TO TRF UPI/Charity Donation	2000.0	Debit
36	e875486e-b86d-489e-b60e-81f964c713ac	2025-04-05	BY TRF Salary Credit	68000.0	Credit
37	7ea8c3fa-87db-45ec-adce-60b64fe11dbb	2025-04-07	TO TRF UPI/Grocery Shopping	3800.0	Debit
38	7399f619-649c-43c4-9dde-4048fcc7ae95	2025-04-10	TO TRF Rent Payment	25000.0	Debit
39	ea4e4c33-67df-45a6-a592-c9a4cec4a2b9	2025-04-12	TO TRF UPI/Electricity Bill	2300.0	Debit
40	bd349645-6713-4a87-95b5-0c6750a3e0bf	2025-04-15	TO TRF Credit Card Bill	20000.0	Debit
41	8583ea84-33f5-4e23-9fca-d61b9d7a4d15	2025-04-18	TO TRF UPI/New Smartphone	15000.0	Debit
42	60c92228-2fca-4723-8bd3-186af897de2f	2025-04-20	TO TRF UPI/W edding Gift	5000.0	Debit
43	035a3603-34d7-4951-842e-e09a577eda3a	2025-04-23	BY TRF UPI/Freelance W ork	12000.0	Credit
44	86f4c911-f86f-4157-bf4b-82d8d9990eda	2025-04-25	TO TRF UPI/Home Appliance	8000.0	Debit
45	ffbf25ef-5a7f-469e-b42b-f86c4eb285f9	2025-04-28	TO TRF UPI/Car Insurance	6500.0	Debit
46	17c2aa0d-6d31-40c5-b9df-03b890c31b6f	2025-04-30	BY TRF Interest Credit	261.0	Credit

	unique_id	date	particular	amount	transaction_type
47	960e6a11-2714-4551-ac93-6d27c1b7c8f9	2025-05-02	TO TRF UPI/W eekend Trip	7000.0	Debit
48	bd35abf4-a142-4c38-842b-5442daaded08	2025-05-05	BY TRF Salary Credit	68000.0	Credit
49	8aff9633-a970-4d90-80d3-5a6123746878	2025-05-07	TO TRF UPI/Grocery Shopping	4000.0	Debit
50	99d84344-2023-444b-bb4e-1f6071bb9098	2025-05-10	TO TRF Rent Payment	25000.0	Debit
51	52025227-400c-45a1-b079-331ef45c28ee	2025-05-12	TO TRF Family Vacation Expenses	75000.0	Debit

In [201...
```python
import json

# Define the JSON structure as a Python dictionary
json_data = {
    "start_row":1,
    "columns": {
        "unique_id":"unique_id",
        "date": "date",
        "description": "description",
        "amount": "amount",
        "transaction_type": "transaction_type"
    }
}
```

In [203...
```python
with open('data_config.json', 'r') as file:
    config = json.load(file)

start_row = config['start_row']
columns_mapping = config['columns']
```

In [176...
```python
df1.rename(columns=columns_mapping, inplace=True)

# Ensure the start row is properly accounted for
df1 = df1.iloc[start_row - 1:].reset_index(drop=True)

# Check the DataFrame
df1
```

	unique_id	date	particular	amount	transaction_type
0	841253a4-dd36-4a6c-8117-f6068753dddf	2025-01-01	Opening Balance	0.0	Credit
1	30b67bac-348f-4a29-89d0-9b0074a6f3ff	2025-01-02	TO TRF UPI/New Year Party Expenses	3500.0	Debit
2	f607e090-0d46-4398-ab74-75b901501077	2025-01-05	BY TRF Salary Credit	65000.0	Credit
3	244f7399-f44c-4966-bff3-775ecff6f993	2025-01-07	TO TRF UPI/Grocery Shopping	2800.0	Debit
4	ef8e701b-8a94-49aa-82ea-d73156769766	2025-01-10	TO TRF Rent Payment	25000.0	Debit
5	ecc52839-6c55-4336-846c-da4ef7d0aef4	2025-01-12	TO TRF UPI/Electricity Bill	1850.0	Debit
6	0a0a64cf-e64f-4678-aaa0-768318c88798	2025-01-15	TO TRF Credit Card Bill	12500.0	Debit
7	98f24f85-ba0a-4a67-a8ad-cc618a7f8213	2025-01-18	TO TRF UPI/Dining Out	2200.0	Debit
8	8d013f1e-9b11-4aff-b73a-d66f92b43d46	2025-01-20	TO TRF UPI/Mobile Recharge	999.0	Debit
9	cc10537c-b469-4baf-af70-475499231230	2025-01-25	TO TRF UPI/Online Shopping	3500.0	Debit
10	29d37a82-9b26-4ead-8239-4572ad69d932	2025-01-28	TO TRF UPI/Gym Membership	2000.0	Debit
11	3c3e3209-3428-4f76-9f39-0eb0a8360ec6	2025-01-31	BY TRF Interest Credit	249.0	Credit
12	7f6e3720-460c-475c-8f40-16514e3f87e3	2025-02-02	TO TRF UPI/W eekend Getaway	5500.0	Debit
13	e0f410ec-0830-490a-bcc1-65acca502113	2025-02-05	BY TRF Salary Credit	65000.0	Credit
14	d88a86a2-d7ae-42d0-8d09-c33159676efb	2025-02-07	TO TRF UPI/Grocery Shopping	3200.0	Debit
15	2db0aef8-4c50-4859-882e-d7f29bdb65fe	2025-02-10	TO TRF Rent Payment	25000.0	Debit
16	63d74ecb-02c3-4c8f-9310-1c494347f5d5	2025-02-12	TO TRF Insurance Premium	8500.0	Debit
17	33c42244-afc5-433b-a601-9ad528287737	2025-02-15	TO TRF Credit Card Bill	15000.0	Debit
18	24c64049-4329-48e9-8f29-32b4cbbe7a77	2025-02-18	TO TRF UPI/V alentine's Day Gift	4000.0	Debit
19	83095003-b2ee-4fac-b0d0-634171917a41	2025-02-20	TO TRF UPI/Car Service	3500.0	Debit
20	16933ff3-ba4f-432e-8fab-b65f89f00157	2025-02-23	BY TRF UPI/Freelance W ork	10000.0	Credit
21	00d305d6-4e20-4eb9-b74f-1c6a810814dd	2025-02-25	TO TRF UPI/Home Decor Items	6000.0	Debit
22	afb33c51-7647-4f43-b5af-743725166ef6	2025-02-28	BY TRF Interest Credit	267.0	Credit
23	0a15388c-393c-457c-8305-467c12835765	2025-03-02	TO TRF UPI/Movie Night	1500.0	Debit
24	4529971b-68b9-48e0-a849-3d48d48c01bd	2025-03-05	BY TRF Salary Credit	65000.0	Credit
25	47bf9f85-919e-4397-ba24-722cb4b86c94	2025-03-07	TO TRF UPI/Grocery Shopping	3500.0	Debit
26	97c91bd0-4b29-4869-8b6b-f3532228c07e	2025-03-10	TO TRF Rent Payment	25000.0	Debit
27	4762f60d-9101-429b-a6eb-2e46847d6414	2025-03-12	TO TRF UPI/Electricity Bill	2100.0	Debit
28	48248e4d-1d4f-472f-8496-b6be07eb273a	2025-03-15	TO TRF Car Loan EMI	18500.0	Debit
29	efd061aa-93d7-4214-b651-522d62735a3a	2025-03-18	TO TRF UPI/Birthday Gift	2500.0	Debit
30	86928d79-28fc-4885-9c02-7ea3828cbce8	2025-03-20	TO TRF UPI/Online Shopping	7250.0	Debit
31	a70abe55-4fe1-408e-a7af-0c3e3b4f6c90	2025-03-23	TO TRF UPI/Dental Appointment	3000.0	Debit
32	898ceb04-0f97-4cee-b72c-4f0c2a6b835a	2025-03-25	BY TRF Tax Refund	8500.0	Credit
33	6bcd7e1b-7f06-498a-bd4a-c105ec2519e7	2025-03-28	TO TRF UPI/Home Repairs	4500.0	Debit
34	abb7a2dd-ff00-47d9-8603-7056a5b93732	2025-03-31	BY TRF Interest Credit	290.0	Credit
35	6735f3b9-e311-4621-818b-da2f89d4b3dd	2025-04-02	TO TRF UPI/Charity Donation	2000.0	Debit
36	e875486e-b86d-489e-b60e-81f964c713ac	2025-04-05	BY TRF Salary Credit	68000.0	Credit
37	7ea8c3fa-87db-45ec-adce-60b64fe11dbb	2025-04-07	TO TRF UPI/Grocery Shopping	3800.0	Debit
38	7399f619-649c-43c4-9dde-4048fcc7ae95	2025-04-10	TO TRF Rent Payment	25000.0	Debit
39	ea4e4c33-67df-45a6-a592-c9a4cec4a2b9	2025-04-12	TO TRF UPI/Electricity Bill	2300.0	Debit
40	bd349645-6713-4a87-95b5-0c6750a3e0bf	2025-04-15	TO TRF Credit Card Bill	20000.0	Debit
41	8583ea84-33f5-4e23-9fca-d61b9d7a4d15	2025-04-18	TO TRF UPI/New Smartphone	15000.0	Debit
42	60c92228-2fca-4723-8bd3-186af897de2f	2025-04-20	TO TRF UPI/W edding Gift	5000.0	Debit
43	035a3603-34d7-4951-842e-e09a577eda3a	2025-04-23	BY TRF UPI/Freelance W ork	12000.0	Credit
44	86f4c911-f86f-4157-bf4b-82d8d9990eda	2025-04-25	TO TRF UPI/Home Appliance	8000.0	Debit
45	ffbf25ef-5a7f-469e-b42b-f86c4eb285f9	2025-04-28	TO TRF UPI/Car Insurance	6500.0	Debit
46	17c2aa0d-6d31-40c5-b9df-03b890c31b6f	2025-04-30	BY TRF Interest Credit	261.0	Credit

	unique_id	date	particular	amount	transaction_type
47	960e6a11-2714-4551-ac93-6d27c1b7c8f9	2025-05-02	TO TRF UPI/W eekend Trip	7000.0	Debit
48	bd35abf4-a142-4c38-842b-5442daaded08	2025-05-05	BY TRF Salary Credit	68000.0	Credit
49	8aff9633-a970-4d90-80d3-5a6123746878	2025-05-07	TO TRF UPI/Grocery Shopping	4000.0	Debit
50	99d84344-2023-444b-bb4e-1f6071bb9098	2025-05-10	TO TRF Rent Payment	25000.0	Debit
51	52025227-400c-45a1-b079-331ef45c28ee	2025-05-12	TO TRF Family Vacation Expenses	75000.0	Debit

As mentioned in requirements we converted dataframe into json file

```
In [178… df1.to_json('transaction_output.json', orient='records', lines=True)

print("DataFrame has been saved to 'transaction_output.json'")
```

DataFrame has been saved to 'transaction_output.json'

Connecting to the Mysql server

```
In [180… import mysql.connector

def create_database_and_table():
    connection = mysql.connector.connect(
        host='localhost',
        user='root',
        password='SQL123'
    )
    cursor = connection.cursor()

    # Create database
    cursor.execute("CREATE DATABASE IF NOT EXISTS bank")

    # Use the database
    cursor.execute("USE bank")

    # Create table
    cursor.execute('''
        CREATE TABLE IF NOT EXISTS transactions (
            unique_id CHAR(36) PRIMARY KEY,
            date DATE,
            description TEXT,
            amount DECIMAL(10, 2),
            transaction_type ENUM('CREDIT', 'DEBIT')
        )
    ''')

    cursor.close()
    connection.close()

create_database_and_table()
```

```
In [181… import pandas as pd

def connect_to_db():
    return mysql.connector.connect(
        host='localhost',
        user='root',
        password='SQL123',
        database='bank'
    )

def insert_dataframe(df):
    connection = connect_to_db()
    cursor = connection.cursor()

    # Insert each row into the database
    for _, row in df.iterrows():
        cursor.execute('''
            INSERT INTO transactions (unique_id, date, description, amount, transaction_type)
            VALUES (%s, %s, %s, %s, %s)
        ''', (row['unique_id'], row['date'], row['particular'], row['amount'], row['transaction_type']))

    connection.commit()
    cursor.close()
    connection.close()

insert_dataframe(df1)
```