



Q

ChatGPT ▾

Partilhar



image generator

Explorar GPTs

Hoje

Identifying Edible Plants

LeetCode Difficulty Insights

Data Inconsistencies Analysis ...

Check all the mistakes and inconsistencies here: "Transaction_ID
Customer_ID Transaction_Amount Date Account_Type
Loan_Amount Loan_Approval_Status Transaction_Type Branch_ID
Transaction_Location Customer_Age Customer_Income
Customer_Marital_Status Customer_Credit_Score
Transaction_Method Fee_Charged Transaction_Description
Transaction_Category Loan_Duration Interest_Rate
Previous_Loan_Status Account_Balance Total_Deposits
Total_Withdrawals Total_Transactions Pending_Transactions
Overdraft_Status

1 C001 200 2023-01-01 Savings 1000 Approved Deposit
B001 New York 35 50000 Single 720 Credit Card 5 Monthly
deposit Income 24 5.0% Approved 2000 5000 1500
10 No
2 C002 150 2023-01-02 Checking Rejected Withdrawal B002
Chicago 45 75000 Married 680 Debit Card 3 ATM
withdrawal Expense 36 4.5% Rejected 3500 7000 2000
8 Yes
3 C003 350 20...↓ J1-03 Savings 5000 Approved Deposit
B001 Boston 25 60000 Single 710 Debit Card 6 Monthly
deposit Income 48 4.8% Approved 4500 4000 1200
12 No
4 C001 50 2023-01-04 Savings Pending Withdrawal B003
New York 35 50000 Single 720 Credit Card 2 ATM withdrawal
Expense 24 5.0% Approved 1950 5100 1450 9 Yes
5 C004 1000 2023-01-05 Checking Approved Deposit
B004 Los Angeles 50 85000 Married 690 Debit Card 10
Direct deposit Income 60 5.5% Approved 5500 9000
5000 15 No
6 C005 75 2023-01-06 Savings 2000 Rejected Withdrawal
B005 Dallas 60 30000 Divorced 640 Debit Card 4 ATM
withdrawal Expense 36 6.0% Rejected 1500 6000 1200
10 Yes
7 C006 250 2023-01-07 Checking Pending Deposit B002
San Francisco 25 45000 Single 730 Credit Card 6 Monthly
deposit Income 48 4.3% Pending 4000 3500 1000 7
No
8 C003 500 20...↓ J1-08 Savings Approved Deposit B001
...ng... 20 60000 Single 710 Credit Card 7 Monthly
deposit Income 48 4.8% Approved 4300 4500 1300
11 No
9 C007 150 2023-01-09 Investment 10000 Approved
Withdrawal B004 Miami 40 95000 Married 740 Debit Card
12 Investment withdrawal Expense 72 3.9% Approved
7000 15000 7000 20 No
10 C002 500 2023-01-10 Savings 5000 Rejected Deposit
B002 Chicago 45 75000 Married 680 Debit Card 3
Monthly deposit Income 36 4.5% Rejected 3300 8000
2500 12 Yes
11 C008 400 2023-01-11 Checking Approved Deposit
B001 New York 30 60000 Single 700 Debit Card 5 ATM
deposit Income 24 4.2% Approved 3800 7200 1800
10 No
12 C009 200 2023-01-12 Savings 2000 Rejected
Withdrawal B003 Chicago 33 50000 Single 690 Credit Card
4 Online transfer Expense 30 4.0% Rejected 1600 5000
2200 9 Yes ↓
13 C010 1500 2023-01-13 Checking Approved Deposit
B004 San Francisco 26 72000 Single 750 Debit Card 15
Direct deposit Income 60 4.6% Approved 5500 10000
4500 14 No
14 C003 250 2023-01-14 Savings Approved Withdrawal
B001 New York 28 60000 Single 710 Debit Card 6 ATM
withdrawal Expense 36 4.8% Approved 4300 4500 1200
10 No
15 C006 120 2023-01-15 Savings Pending Deposit B005
Boston 25 45000 Single 725 Debit Card 3 Monthly deposit
Income 36 4.7% Pending 3700 3800 1500 8 No
16 C007 2000 2023-01-16 Investment Approved
Withdrawal B004 Miami 40 95000 Married 740 Debit Card
20 Investment withdrawal Expense 72 3.9% Approved
7000 12000 8000 18 Yes
17 C002 700 2023-01-17 Savings 7000 Approved Deposit
B003 Chicago 45 75000 Married 680 Credit Card 6 Online
transfer Income 48 5.2% Approved 4500 8500 2000
12 No ↓
18 C005 200 2023-01-18 Savings 2000 Rejected
Withdrawal B002 Los Angeles 60 30000 Divorced 640 Debit
Card 3 ATM withdrawal Expense 36 6.0% Rejected 1300
4200 2200 9 Yes

19 C004 500 2023-01-19 Checking Approved Deposit
B001 New York 50 85000 Married 690 Debit Card 7 Direct
deposit Income 60 5.3% Approved 6000 9500 5000
14 No
20 C008 300 2023-01-20 Savings 3000 Rejected
Withdrawal B004 Dallas 30 60000 Single 710 Credit Card 2
ATM withdrawal Expense 36 4.9% Rejected 2000 4900
1000 8 Yes
21 C001 1000 2023-01-21 Savings 4000 Approved
Deposit B002 New York 35 65000 Single 730 Credit Card 10
Monthly deposit Income 24 5.0% Approved 2200 5800
1700 12 No
22 C002 200 2023-01-22 Checking Rejected Withdrawal
B001 Chicago 45 72000 Married 675 Debit Card 5 ATM
withdrawal Expe ↓ 36 4.5% Rejected 3200 6700 1800
8 Yes
23 C003 1500 2023-01-23 Savings Approved Deposit
B003 Boston 28 60000 Single 710 Credit Card 8 Direct
deposit Income 48 4.8% Approved 4600 4700 1500
10 No
24 C004 500 2023-01-24 Investment Approved
Withdrawal B001 Los Angeles 50 85000 Married 690 Debit
Card 12 Investment withdrawal Expense 72 3.9%
Approved 5400 12000 6000 15 Yes
25 C005 300 2023-01-25 Savings 3000 Rejected
Withdrawal B005 Chicago 60 40000 Divorced 620 Debit
Card 6 Online transfer Expense 36 6.0% Rejected 1400
4300 1500 7 No
26 C006 400 2023-01-26 Savings Pending Deposit B003
San Francisco 30 50000 Single 710 Credit Card 7 Monthly
deposit Income 24 4.5% Pending 3900 3800 1300 9
No
27 C007 800 2023-01-27 Checking Approved Deposit
B004 Miami 40 ↓ 50000 Married 740 Debit Card 15 Direct
deposit Income 60 4.2% Approved 6500 9500 3500
14 No
28 C008 100 2023-01-28 Savings 2000 Rejected
Withdrawal B005 Dallas 32 60000 Single 700 Debit Card 4
ATM withdrawal Expense 36 4.7% Rejected 1700 4600
1200 10 Yes
29 C009 350 2023-01-29 Checking Approved Deposit
B003 New York 33 55000 Single 680 Credit Card 5 ATM
deposit Income 24 4.3% Approved 4200 5400 1800
12 No
30 C010 200 2023-01-30 Savings 5000 Approved
Withdrawal B002 Los Angeles 26 72000 Single 750 Debit
Card 8 Monthly withdrawal Expense 36 4.9% Approved
4600 4700 1300 10 No
31 C001 1500 2023-01-31 Checking Pending Deposit
B004 Boston 35 52000 Single 710 Debit Card 10 Direct
deposit Income 48 4.6% Pending 4000 6000 2200 13
No
32 C003 800 2023-02-01 Savings Approved Withdrawal
B001 Miami 28 ↓ 30000 Single 700 Debit Card 5 ATM
withdrawal expense 36 4.1% Approved 4600 4900 1500
11 Yes
33 C002 100 2023-02-02 Checking Rejected Deposit B003
Chicago 45 75000 Married 690 Debit Card 3 Online transfer
Income 36 4.2% Rejected 3200 6700 2000 8 No
34 C004 700 2023-02-03 Investment 5000 Approved
Deposit B001 San Francisco 50 85000 Married 690 Debit
Card 12 Investment deposit Income 60 5.3% Approved
5500 11000 3000 17 Yes
35 C005 250 2023-02-04 Savings Pending Withdrawal
B004 Dallas 60 30000 Divorced 640 Debit Card 6 ATM
withdrawal Expense 36 6.0% Pending 1400 4300 1500
9 No
36 C006 300 2023-02-05 Savings 2000 Approved Deposit
B005 Los Angeles 35 60000 Single 720 Debit Card 7 Online
transfer Income 36 5.0% Approved 4300 4800 1600
12 No
37 C007 500 2023-02-06 Checking Approved Deposit
B002 Miami 40 ↓ 50000 Married 700 Debit Card 8 Direct
deposit Income 45 5
38 C007 100 2023/02/21 Savings 2000 Approved Deposit
B003 Miami 95000 Married 740 Debit Card 10 Monthly
deposit Income 36 4.6 Approved 4500 8000 2000 15
No
39 C008 200 2023-02-22 Checking Rejected Withdrawal
B002 Los Angeles 50 82000 Single 700 Credit Card 5 ATM
withdrawal Expense 48 5.0% Rejected 4000 7500 2500
10 Yes
40 C009 300 2023/02/23 Savings 5000 Approved Deposit
B001 Chicago 38 71000 Married 720 Debit Card 7
Monthly deposit Income 24 4.7% Approved 3800 6700

Monthly deposit income	24	4.7%	Approved	5900	8700
3000	12	No			
41	C010	1500	2023/02/24	Investment	10000 Approved
Withdrawal	B004	Dallas	40	85000 Single	700 Debit Card 12
Investment withdrawal	Expense	72	4.1%	Approved	8000
13000	5000	20	No		
42	C001	1000	2023-02-25	Savings	Approved Deposit
B003	New York	↓	54000	Married	650 Debit Card 3
Monthly deposit Income	24	5.5%	Approved	2200	7000
1800	10	No			
43	C002	150	2023/02/26	Savings	Rejected Deposit
B002	Chicago	45	73000	Married	690 Credit Card 2 Online
transfer	Income	48	4.3%	Rejected	2500 6300 1200 8
Yes					
44	C003	350	2023-02-27	Checking	Approved Deposit
B004	Los Angeles	28	62000	Single	710 Debit Card 8 Direct
deposit	Income	36	4.8%	Approved	3700 8200 2000
12	No				
45	C004	400	2023-02-28	Savings	Rejected
Withdrawal	B005	Miami	45	75000	Married 700 Debit Card
6	ATM withdrawal	Expense	36	5.0%	Rejected 1500 6000
1800	15	Yes			
46	C005	250	2023-02-29	Savings	Approved Deposit
B001	Chicago	60	30000	Divorced	640 Credit Card 4 Online
transfer	Income	24	6.0%	Approved	3000 4000 1500
10	No				
47	C006	500	↓ 3/03/01	Investment	Approved
Withdrawal	B003	Dallas	55	85000	Married 730 Debit Card
15	Investment withdrawal	Expense	72	4.7%	Approved
6500	12000	4000	15	No	
48	C007	200	2023-03-02	Savings	Rejected Deposit
B004	Miami	35	60000	Single	690 Debit Card 7 ATM withdrawal
Income	36	4.9%	Rejected	2300 5200 1700 10	Yes
49	C008	800	2023-03-03	Checking	5000 Approved
Withdrawal	B002	Los Angeles	45	73000	Married 680 Credit
Card	12	ATM withdrawal	Expense	60	5.3% Approved
5200	10000	6000	20	No	
50	C009	1500	2023/03/04	Savings	Approved Deposit
B005	Chicago	50000	Single	710 Credit Card	10 Monthly
deposit	Income	36	4.5%	Approved	4200 6500 2000
13	No				
51	C010	500	2023-03-05	Investment	Rejected Deposit
B003	Miami	38	75000	Married	725 Debit Card 4 ATM
withdrawal	Income	36	5.4%	Rejected	3700 8500 3000
14	Yes				
52	C001	400	↓ 3/03/06	Savings	Approved Deposit
B002	New York	33	67000	Single	740 Credit Card 5 Online
transfer	Income	24	5.0%	Approved	3600 7400 1900
12	No	"(where does it need data cleaning?)			

There are several issues with inconsistencies and missing data in the dataset. Here are the observations:

1. Missing Values for Some Columns:

- **Loan_Amount**: Some transactions do not have a loan amount (e.g., transactions 2, 4, 7, 10, 13, etc.).
- **Loan_Approval_Status**: Some transactions lack a loan approval status (e.g., transactions 6, 7, 20, 21, etc.).
- **Transaction_Location**: In transactions like 37, the location for the transaction is missing.
- **Customer_Age**: In some cases, the customer age is missing (e.g., transaction 38).

2. Date Format Inconsistencies:

- Dates are inconsistent in format, e.g., transaction 38 and 40 have a different format (2023/02/21, 2023/02/23, etc.), while others follow YYYY-MM-DD. This inconsistency should be addressed by standardizing the date format.

3. Transaction_Type Inconsistencies:

- For example, transaction 37 has a missing transaction type, which can cause issues during data analysis.

4. Inconsistent Loan Data:

- Some transactions list loan-related data even when there is no loan (e.g., transaction 4, where **Loan_Amount** is blank but **Loan_Approval_Status** is "Pending").
- Inconsistent **Loan_Amount** values are also present, which may not correspond to a **Loan_Approval_Status**. For instance, transaction 9 shows a loan amount of 10,000 but does not provide loan approval.

5. Interest Rate Formatting:

- For some transactions, the **Interest_Rate** column does not have the percentage sign (e.g., transactions 37 and 38, where it's simply 4 instead of 4.5%).

6. Fee Charged:

- Some records are missing a fee (e.g., transaction 37, 38). It should be verified if the fee column is required for all transactions.

7. Pending Transaction Status:

- Transactions such as 6, 7, 26, etc., have pending statuses but no clear reasoning behind them, considering other columns such as transaction amount. Further clarification on pending statuses may be needed.

8. Customer Income and Other Customer Info:

- For a few records, the customer's income data seems inconsistent or missing (e.g., transaction 50 with missing income).

9. Column Name Inconsistencies:

- In some records, columns like `Total_Deposits`, `Total_Withdrawals`, and `Total_Transactions` have missing values. These should either be populated or marked clearly as "not applicable" if that's the case.



10. Data Values Inconsistencies:

- For example, in transaction 38, the `Customer_Age` is missing, which could affect demographic analysis.
- Similarly, transaction 47 is missing certain columns like `Transaction_Location`, `Transaction_Type`, and others.

11. Date Missing or Incorrect Formatting:

- Inconsistent date formats should be fixed (e.g., 2023/02/21 vs. 2023-02-21).

Recommended Data Cleaning Actions:

- Standardize the date format across the dataset (YYYY-MM-DD).
- Handle missing values in columns (e.g., `Loan_Amount`, `Loan_Approval_Status`, `Transaction_Location`, etc.).
- Ensure the `Interest_Rate` column includes the percentage sign where applicable.
- Fix the inconsistencies in the `Transaction_Type` and `Fee_Charged` columns.
- Resolve missing customer age and income data where applicable.
- Clarify and possibly remove records with contradictory or missing loan information.
- Double-check for any outliers or inconsistencies in the data entries that do not align with expected patterns (e.g., fees, income, and deposit/withdrawal amounts).



Pergunte qualquer coisa

+ Procurar Raciocina



O ChatGPT pode cometer erros. Considere verificar informações importantes.

Atualizar plano
Mais acesso aos melhores modelos

