Payment Gateway Reconciliation

Using Excel & SQL

By Preety Paramanick

Payment Gateway Reconciliation Analysis

Inspiration

Maintaining correct and uniform financial records is essential for all businesses, particularly those that depend significantly on online transactions. The practice of checking and aligning payment information from various platforms and sources, known as payment gateway reconciliation, is key to this task. This study seeks to spot and resolve any inconsistencies in the three given datasets related to payment gateway transactions.

Objectives:

1.Eliminate Duplicate Records.

2.Match Records Across Datasets.

3.Summarize Matching/Unmatching Records.

4.Investigate Amount Mismatches.

5.Document Matching Logic and Assumptions.

Steps

I got the data in excel format so, first I separate them and save them individual in CSV format.

Import Data into Excel

- Open Excel and create a new workbook.
- Import Data: Go to the "Data" tab, select "From Text/CSV", and choose Collection CSV file Repeat this for Siply.CSV and Cashfree.CSV as well.
- Load Data: Preview the data and click "Load" to import it into separate sheets within your workbook.

Import Data into SQL

- Open SQL Workbench and Create a new Database as “payment_gateway”.
- Import each CSV file by click in Table data import wizard. Or it also can be possible by write SQL queries manually but it would be a tedious job in this case.
- Make unique identifiers in each table.

Clean and Prepare Data in Excel

- It could be possible in two ways: 1.Using Power Query, 2.Using Excel Workbook’s tools.(I applied both)
- Make sure that every column should have its appropriate datatype.

Data Review in SQL

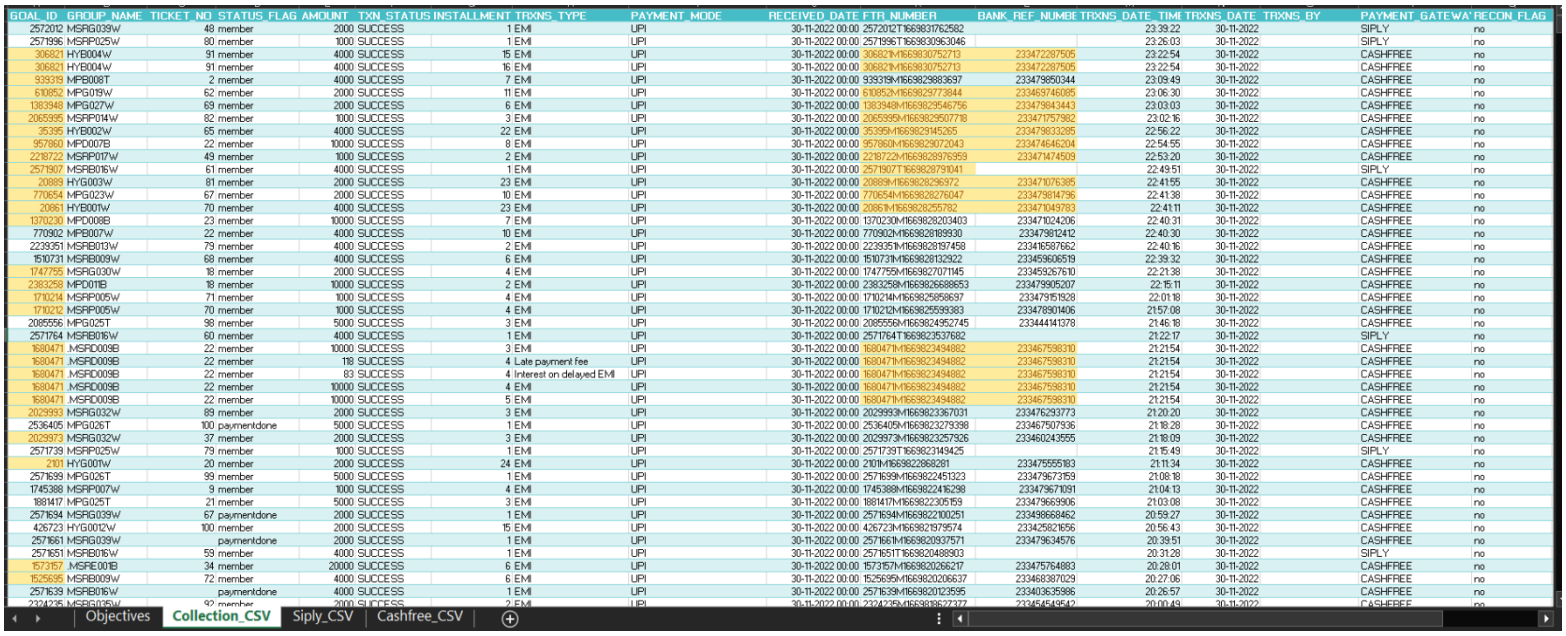
- By perform ‘Select * from Table_Name’ query I quickly review each table in SQL.

1. Find Duplicate records in three dataset

Find Duplicate in Excel

- Identifies duplicates by using conditional column formatting.
- Duplicate values are pointed as Yellow color.
- To find unique column I need to delete duplicates by using Remove duplicates under Data Ribbon.

Quick Snippet of the work



GOAL_ID	GROUP_NAME	TICKET_NO	STATUS_FLAG	AMOUNT	TXN_STATUS	INSTALLMENT	TRXNS_TYPE	PAYMENT_MODE	RECEIVED_DATE	FTR_NUMBER	BANK_REF_NUMBE	TRXNS_DATE_TIME	TRXNS_DATE	TRXNS_BY	PAYMENT_GATEWA	RECON_FLAG
2572012	MSRG039W	48	member	2000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2572012T1669831762582		23:39:22	30-11-2022		SIPLY	no
2571996	MSRP025W	80	member	1000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571996T1669830963046		23:26:03	30-11-2022		SIPLY	no
306821	HYB004W	91	member	4000	SUCCESS	15	EMI	UPI	30-11-2022 00:00	306821M1669830752713	233472287505	23:22:54	30-11-2022		CASHFREE	no
306821	HYB004W	91	member	4000	SUCCESS	16	EMI	UPI	30-11-2022 00:00	306821M1669830752713	233472287505	23:22:54	30-11-2022		CASHFREE	no
939319	MPB008T	2	member	4000	SUCCESS	7	EMI	UPI	30-11-2022 00:00	939319M1669829883697	233479850344	23:09:49	30-11-2022		CASHFREE	no
610852	MPG019W	62	member	2000	SUCCESS	11	EMI	UPI	30-11-2022 00:00	610852M1669829773844	233469746085	23:06:30	30-11-2022		CASHFREE	no
1383948	MPG027W	69	member	2000	SUCCESS	6	EMI	UPI	30-11-2022 00:00	1383948M1669829546756	233479843443	23:03:03	30-11-2022		CASHFREE	no
2065995	MSRP014W	82	member	1000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	2065995M1669829507718	233471757982	23:02:16	30-11-2022		CASHFREE	no
35395	HYB002W	65	member	4000	SUCCESS	22	EMI	UPI	30-11-2022 00:00	35395M1669829145265	233479833285	22:56:22	30-11-2022		CASHFREE	no
957860	MPD007B	22	member	10000	SUCCESS	8	EMI	UPI	30-11-2022 00:00	957860M1669829072043	233474646204	22:54:55	30-11-2022		CASHFREE	no
2218722	MSRP017W	49	member	1000	SUCCESS	2	EMI	UPI	30-11-2022 00:00	2218722M1669828976959	233471474509	22:53:20	30-11-2022		CASHFREE	no
2571907	MSRB016W	61	member	4000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571907T1669828791041		22:49:51	30-11-2022		SIPLY	no
20889	HYG003W	81	member	2000	SUCCESS	23	EMI	UPI	30-11-2022 00:00	20889M1669828296972	233471076385	22:41:55	30-11-2022		CASHFREE	no
770654	MPG023W	67	member	2000	SUCCESS	10	EMI	UPI	30-11-2022 00:00	770654M1669828276047	233479814796	22:41:38	30-11-2022		CASHFREE	no
20861	HYB001W	70	member	4000	SUCCESS	23	EMI	UPI	30-11-2022 00:00	20861M1669828255782	233471949793	22:41:11	30-11-2022		CASHFREE	no
1370230	MPD008B	23	member	10000	SUCCESS	7	EMI	UPI	30-11-2022 00:00	1370230M1669828203403	233471024206	22:40:31	30-11-2022		CASHFREE	no
770902	MPB007W	22	member	4000	SUCCESS	10	EMI	UPI	30-11-2022 00:00	770902M1669828189930	233479812412	22:40:30	30-11-2022		CASHFREE	no
2239351	MSRB013W	79	member	4000	SUCCESS	2	EMI	UPI	30-11-2022 00:00	2239351M1669828197458	233416587662	22:40:16	30-11-2022		CASHFREE	no
1510731	MSRB009W	68	member	4000	SUCCESS	6	EMI	UPI	30-11-2022 00:00	1510731M1669828132922	233459606519	22:39:32	30-11-2022		CASHFREE	no
1747755	MSRG030W	18	member	2000	SUCCESS	4	EMI	UPI	30-11-2022 00:00	1747755M1669827071145	233459267610	22:21:38	30-11-2022		CASHFREE	no
2383258	MPD011B	18	member	10000	SUCCESS	2	EMI	UPI	30-11-2022 00:00	2383258M1669826688653	233479905207	22:15:11	30-11-2022		CASHFREE	no
1710214	MSRP005W	71	member	1000	SUCCESS	4	EMI	UPI	30-11-2022 00:00	1710214M1669825858697	233479151928	22:01:18	30-11-2022		CASHFREE	no
1710212	MSRP005W	70	member	1000	SUCCESS	4	EMI	UPI	30-11-2022 00:00	1710212M1669825599383	233478901406	21:57:08	30-11-2022		CASHFREE	no
2085556	MPG025T	98	member	5000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	2085556M1669824952745	233444141378	21:46:18	30-11-2022		CASHFREE	no
2571764	MSRB016W	60	member	4000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571764T1669823537682		21:22:17	30-11-2022		SIPLY	no
1680471	MSRD009B	22	member	10000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	1680471M1669823494882	233467598310	21:21:54	30-11-2022		CASHFREE	no
1680471	MSRD009B	22	member	118	SUCCESS	4	Late payment fee	UPI	30-11-2022 00:00	1680471M1669823494882	233467598310	21:21:54	30-11-2022		CASHFREE	no
1680471	MSRD009B	22	member	83	SUCCESS	4	Interest on delayed EMI	UPI	30-11-2022 00:00	1680471M1669823494882	233467598310	21:21:54	30-11-2022		CASHFREE	no
1680471	MSRD009B	22	member	10000	SUCCESS	4	EMI	UPI	30-11-2022 00:00	1680471M1669823494882	233467598310	21:21:54	30-11-2022		CASHFREE	no
1680471	MSRD009B	22	member	10000	SUCCESS	5	EMI	UPI	30-11-2022 00:00	1680471M1669823494882	233467598310	21:21:54	30-11-2022		CASHFREE	no
2029993	MSRG032W	89	member	2000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	2029993M1669823367031	233476293773	21:20:20	30-11-2022		CASHFREE	no
2536405	MPG026T	100	paymentdone	5000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2536405M1669823279398	233467507936	21:18:28	30-11-2022		CASHFREE	no
2029973	MSRG032W	37	member	2000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	2029973M1669823257926	233460243555	21:18:09	30-11-2022		CASHFREE	no
2571739	MSRP025W	79	member	1000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571739T1669823149425		21:15:49	30-11-2022		SIPLY	no
2101	HYG001W	20	member	2000	SUCCESS	24	EMI	UPI	30-11-2022 00:00	2101M1669822868281	233475555183	21:11:34	30-11-2022		CASHFREE	no
2571699	MPG026T	99	member	5000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571699M1669822451323	233479673159	21:08:18	30-11-2022		CASHFREE	no
1745388	MSRP007W	9	member	1000	SUCCESS	4	EMI	UPI	30-11-2022 00:00	1745388M1669822416298	233479671091	21:04:13	30-11-2022		CASHFREE	no
1881417	MPG025T	21	member	5000	SUCCESS	3	EMI	UPI	30-11-2022 00:00	1881417M1669822305159	233479669906	21:03:08	30-11-2022		CASHFREE	no
2571694	MSRG039W	67	paymentdone	2000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571694M1669822100251	233498668462	20:59:27	30-11-2022		CASHFREE	no
426723	HYG001W	100	member	2000	SUCCESS	15	EMI	UPI	30-11-2022 00:00	426723M1669821979574	233425821656	20:56:43	30-11-2022		CASHFREE	no
2571661	MSRG039W		paymentdone	2000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571661M1669820937571	233479634576	20:39:51	30-11-2022		CASHFREE	no
2571651	MSRB016W	59	member	4000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571651T1669820488903		20:31:28	30-11-2022		SIPLY	no
1573157	MSRE001B	34	member	20000	SUCCESS	6	EMI	UPI	30-11-2022 00:00	1573157M1669820266217	233475764883	20:28:01	30-11-2022		CASHFREE	no
1525695	MSRB009W	72	member	4000	SUCCESS	6	EMI	UPI	30-11-2022 00:00	1525695M1669820206637	233468387029	20:27:06	30-11-2022		CASHFREE	no
2571639	MSRB016W		paymentdone	4000	SUCCESS	1	EMI	UPI	30-11-2022 00:00	2571639M1669820123595	233403635986	20:26:57	30-11-2022		CASHFREE	no
2374235	MSRB005W	92	member	2000	SUCCESS	2	EMI	UPI	30-11-2022 00:00	2374235M1669818672377	233454549542	20:00:43	30-11-2022		CASHFREE	no

Objectives | Collection_CSV | Siply_CSV | Cashfree_CSV

Find Duplicates in SQL

- We can find duplicate by this query for each table.

Quick Snippet of the work

```
-- Find Duplicate in Siply --
SELECT Customer_id,COUNT(DISTINCT user_goal_id)
FROM siply
GROUP BY Customer_id
HAVING COUNT(DISTINCT user_goal_id) > 1;
```

Customer_id	COUNT(DISTINCT goal_id)
643124	2
643342	3
645217	2
663962	5
672964	3
678314	2
681398	3
684234	2
694047	2
704162	2
715003	2
715219	2
718616	2
728328	2

2.Find Matching Records Between All Three Datasets by Tracking Payment

Find Matching Records in Excel

- As we know every column of the table is not required.
- By taking common column and using formula like VLOOKUP and INDEX-MATCH we can retrieve data that are common.

Quick Snippet of the work

```
=INDEX(Siply_CSV!C:C,MATCH(Objectives!L4,Siply_CSV!J:J,0))
```

```
=VLOOKUP(Objectives!$L4,Collection1,14,0)
```

Collection and Siply match

Total_UserGoalID = 1775

User_Goal_id	Transaction_Date(collection)	Transaction_Date(siply)	Amount(Collection)	Amount(siply)	Payment_Mode	Siply_CustomerID
1520164	01-11-2022	01-11-2022	1000	1000	UPI	1520084
2376791	01-11-2022	01-11-2022	1000	1000	UPI	2376741
2376795	01-11-2022	01-11-2022	1000	1000	UPI	2376786
1442418	01-11-2022	01-11-2022	4000	4000	UPI	1442416
1502705	01-11-2022	01-11-2022	1000	1000	UPI	1441890
2376851	01-11-2022	01-11-2022	4000	4000	UPI	2376841
2022118	26-11-2022	01-11-2022	4000	4000	UPI	2022110
2034935	01-11-2022	01-11-2022	2000	2000	UPI	1535624
2376900	01-11-2022	01-11-2022	1000	1000	UPI	2376890
702593	01-11-2022	01-11-2022	500	500	UPI	702589
2376983	01-11-2022	01-11-2022	1000	1000	UPI	2369503
1540490	01-11-2022	01-11-2022	1000	1000	UPI	1490713
2112439	01-11-2022	01-11-2022	1000	1000	UPI	2112321
954196	01-11-2022	01-11-2022	250	250	UPI	954189
1062693	01-11-2022	01-11-2022	2000	2000	UPI	1062688
1872779	01-11-2022	01-11-2022	2000	2000	UPI	1872756
2380809	01-11-2022	01-11-2022	2000	2000	UPI	2229973
739025	01-11-2022	01-11-2022	2000	2000	UPI	739022
966311	01-11-2022	01-11-2022	2000	2000	UPI	966035
1993667	01-11-2022	01-11-2022	1000	1000	UPI	1770742
2380912	01-11-2022	01-11-2022	1000	1000	UPI	2380908
1457596	01-11-2022	01-11-2022	719	1000	UPI	1457594
1924561	01-11-2022	01-11-2022	1000	1000	UPI	1901398
1001765	01-11-2022	01-11-2022	2000	2000	UPI	663962
663968	01-11-2022	01-11-2022	2000	2000	UPI	663962
663987	01-11-2022	01-11-2022	1500	1500	UPI	663962
664030	01-11-2022	01-11-2022	1500	1500	UPI	663962
951165	01-11-2022	01-11-2022	1500	1500	UPI	951162

Find Matching Data in SQL

- By the snippet query we can achieve that

Quick Snippet of the work

```
SELECT co.Goal_ID,c.bank_ref_no,co.Trxns_date, co.Amount, co.Payment_mode,co.Txn_status,Installment,Trxns_type
FROM Collection co
JOIN Siply s ON co.goal_id = s.goal_id
JOIN Cashfree c ON s.payment_mode = c.payment_mode
where co.payment_mode="UPI" ;
```

Goal_ID	bank_ref_no	Trxns_date	Amount	Payment_mode	Txn_status	Installment	Trxns_type
2572012	231000000000	2022-11-30	2000	UPI	SUCCESS	1	EMI
2571996	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI
2571907	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2571907	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2571764	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2571739	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI
2571651	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2571552	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI
2571513	231000000000	2022-11-30	2000	UPI	SUCCESS	1	EMI
2571455	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI
2571401	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2571088	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI
2570988	231000000000	2022-11-30	2000	UPI	SUCCESS	1	EMI
2570962	231000000000	2022-11-30	4000	UPI	SUCCESS	1	EMI
2570778	231000000000	2022-11-30	1000	UPI	SUCCESS	1	EMI

3.Prepare Summary of Matching/Unmatching Records and Amount.

Summary of Matching/Unmatching Records and Amount in EXCEL

- By applying functions like VLOOKUP and INDEX_MATCH we can fetch this records in excel. Also we can fetch the records by using Pivot table

Quick Snippet of the work

Dataset	Matched_Count	Unmatchted_count	Matched_Amount	Unmatched_Amount	Total_Amount
Siply	1568	207	₹ 24,12,500	₹ 4,14,709	₹ 28,27,209
Cashfree	3526	617	₹ 1,34,78,337	₹ 36,64,258	₹ 1,71,42,595
Collection	5094	4979	₹ 1,58,90,837	₹ 2,49,18,337	₹ 4,08,09,174
Total	5094	5803	₹ 1,58,90,837	₹ 2,89,97,304	₹ 4,48,88,141

Summery of Matching/Unmatching Records and Amount in SQL

- By using UNION, JOIN, WHERE clause We can summarize records and amount.

Quick Snippet of the work

```
SELECT
    'Matching Records' AS Summary,
    COUNT(*) AS NumberOfRecords,
    SUM(C.amount) AS TotalAmount
FROM Cashfree C
JOIN Collection Co ON C.bank_ref_no= Co.bank_ref_no
JOIN Siply S ON C.Amount = S.Amount
UNION
SELECT
    'Unmatching Records Cashfree1' AS Summary,
    COUNT(*) AS NumberOfRecords,
    SUM(C.amount) AS TotalAmount
FROM Cashfree C
LEFT JOIN Collection Co ON C.bank_ref_no = Co.bank_ref_no
LEFT JOIN Siply S ON C.payment_mode = S.payment_mode
WHERE Co.payment_mode IS NULL OR S.payment_mode IS NULL
UNION
SELECT
    'Unmatching Records Collection1' AS Summary,
    COUNT(*) AS NumberOfRecords,
    SUM(Co.amount) AS TotalAmount
FROM Collection Co
LEFT JOIN Cashfree C ON Co.bank_ref_no = C.bank_ref_no
LEFT JOIN Siply S ON Co.payment_mode = S.payment_mode
WHERE C.payment_mode IS NULL OR S.payment_mode IS NULL
UNION
SELECT
    'Unmatching Records Siply1' AS Summary,
    COUNT(*) AS NumberOfRecords,
    SUM(S.amount) AS TotalAmount
FROM Siply S
LEFT JOIN Cashfree C ON S.paymeny_mode= C.payment_mode
LEFT JOIN Collection Co ON S.goal_id = Co.goal_id
WHERE C.payment_mode IS NULL OR Co.payment_mode IS NULL
```

This query can throw error as it will fetch a huge amount of data, make sure to limit rows for quick result.

4.Identify Cases of Any Amount Mismatch

Identify cases of mismatch amount in EXCEL

- Mismatch records or data can be found in excel by Conditional Formatting by highlighting the mismatch data.

Quick Snippet of the work

Collection and Siply match

Total_UserGoalID = 1775							Matched_TrxnDate = 1600	Matched_Amount = 1568
User_Goal_id	**Transaction_Date(collection)**	**Transaction_Date(siply)**	**Amount(Collection)**	**Amount(siply)**	**Payment_Mode**	**Siply_CustomerID**	**Trxn_date_Match**	**Amount_Match**
1520164	01-11-2022	01-11-2022	1000	1000	UPI	1520084	Match	Match
2376791	01-11-2022	01-11-2022	1000	1000	UPI	2376741	Match	Match
2376795	01-11-2022	01-11-2022	1000	1000	UPI	2376786	Match	Match
1442418	01-11-2022	01-11-2022	4000	4000	UPI	1442416	Match	Match
1502705	01-11-2022	01-11-2022	1000	1000	UPI	1441890	Match	Match
2376851	01-11-2022	01-11-2022	4000	4000	UPI	2376841	Match	Match
2022118	26-11-2022	01-11-2022	4000	4000	UPI	2022110	Not Match	Match
2034935	01-11-2022	01-11-2022	2000	2000	UPI	1535624	Match	Match
2376900	01-11-2022	01-11-2022	1000	1000	UPI	2376890	Match	Match
702593	01-11-2022	01-11-2022	500	500	UPI	702589	Match	Match
2376983	01-11-2022	01-11-2022	1000	1000	UPI	2369503	Match	Match
1540490	01-11-2022	01-11-2022	1000	1000	UPI	1490713	Match	Match
2112439	01-11-2022	01-11-2022	1000	1000	UPI	2112321	Match	Match
954196	01-11-2022	01-11-2022	250	250	UPI	954189	Match	Match
1062693	01-11-2022	01-11-2022	2000	2000	UPI	1062688	Match	Match
1872779	01-11-2022	01-11-2022	2000	2000	UPI	1872756	Match	Match
2380809	01-11-2022	01-11-2022	2000	2000	UPI	2229973	Match	Match
739025	01-11-2022	01-11-2022	2000	2000	UPI	739022	Match	Match
966311	01-11-2022	01-11-2022	2000	2000	UPI	966035	Match	Match
1993667	01-11-2022	01-11-2022	1000	1000	UPI	1770742	Match	Match
2380912	01-11-2022	01-11-2022	1000	1000	UPI	2380908	Match	Match
1457596	01-11-2022	01-11-2022	719	1000	UPI	1457594	Match	Not Match
1924561	01-11-2022	01-11-2022	1000	1000	UPI	1901398	Match	Match
1001765	01-11-2022	01-11-2022	2000	2000	UPI	663962	Match	Match

Identify cases of mismatch amount in SQL

- After JOINING the table we can fine mismatch records in SQL.

Quick Snippet of the work

```
-- Identify Cases of Any Amount Mismatch --
-- Cashfree and Collection amount mismatch --
SELECT *
FROM Cashfree c
JOIN Collection co ON c.bank_ref_no = co.bank_ref_no
WHERE c.AMOUNT <> co.AMOUNT;
```

- We can perform this query with each dataset.

Conclusion and Assumption

During Data-cleaning and Processing

- The datasets were in Excel format with consistent column headers and data types.During the Cleaning and processing I need to rename some Common column and modify their datatypes as well.
- Data quality issues (missing values, inconsistencies) were addressed through cleaning and standardization by Power Query Editor present in EXCEL.

Duplicate Identification

- Duplicates were identified within each dataset based on a chosen criteria :

  1: Exact matches on all fields(Goal_ID,Bank_Ref_No.).

  2: Partial matches on key identifiers like Order_ID, customer ID, etc.

  - Assumptions: Uniqueness of chosen identifiers within each dataset.All duplicate records represent the same transaction.

Matching Logic

Records were matched across datasets using common identifiers such as:

Goal ID, Order ID, Bank_Ref_No, Customer ID

- Assumptions: Uniqueness of chosen identifiers across all datasets. Consistent representation and data quality of identifiers across datasets. Cases with missing or inconsistent identifiers were handled individually (explained further).

Matching/Unmatching Summary

- Matching counts were identified using COUNTIF or similar functions based on matching criteria.
- Unmatching counts were calculated by subtracting matching counts from total records.
- Transaction amounts were summed for matched and unmatching records using SUMIF or SUMIFS functions.
  - Assumptions: Accurate calculations based on chosen criteria.Completness of datasets captured.

Amount Mismatch Analysis

- Transactions with different amounts despite matching identifiers were identified using conditional formatting or comparisons.

- Potential causes for mismatches were considered, including:

  Rounding differences, Fee variations, Partial refunds, Data entry errors

  - Assumptions: Identified causes accurately explain most mismatches.

All assumptions made throughout the process are documented within this document.