Business Intelligence Analyst Technical Interview

Denick Garo

18/11/2025

SECTION A

1. **Combine the seven files into a single file that can be uploaded as a SQL table**

To meet the requirement of combining the seven Excel files into a single file suitable for SQL upload, I created a Python script (merge_calls.py) using Pandas. The goal was to automate the merging process, clean the resulting dataset, and generate a structured CSV file ready for import into a database.

Once the files were combined, I applied cleaning steps to ensure all fields were consistent and SQL-compatible. After cleaning, I exported the final Data Frame to a CSV. This produced a well-structured dataset with consistent column types, ready for direct SQL ingestion. After generating the CSV, I used DBeaver to complete the upload and get ready for the second phase of querying the data.

Attached here is the link for the merge.py file.

https://github.com/Denick12/BI_Assessment/blob/master/merge_data.py

SECTION B

For the next section, write SQL queries for each question that will return the requested results for the full time covered in Files 1-7.

2. **How many disconnect types are there?**
   7
3. **How many unique users are there?**
   673
4. **Who is the top 10 users who took calls? How many calls did each take?**

User_id	Calls taken
c4580193-23d1-4a28-aebb-136c27848142	8,402
58eea8b2-12ef-4e44-b229-af25799ac3bb	8,135
b8b5a757-7176-499d-b93d-8feebb36c09f	7,351
4ff996ef-9f41-4fc7-9894-333cd0364bcc	7,116

bed34990-b428-4868-b4b3-e18e0980dedf	6,836
8261bb15-26ea-48f4-b5e8-686de8c5ccea	6,601
57fe1954-e706-4ac9-a693-8f0d8b5faa5d	6,536
fefd5888-a8c0-4c46-8061-81b2e104efb0	6,383
5933760a-c198-489e-b828-0a95aaa66edb	6,217
96591502-f4f9-4357-b7ca-11fed55b1fc3	6,181

5. What is the average total handle time of a call?

avg_handle_time_seconds	avg_handle_time_hms
1180.8749134265634	03:19:40

6. Who are the top 10 users with the disconnect type of 'peer'?

UserId	peer_disconnects
e3169f6e-7561-44ed-9256-1e7b75c85990	2,691
58eea8b2-12ef-4e44-b229-af25799ac3bb	2,404
2af4b65f-e557-40b8-9876-f0a060c4ba68	1,663
b8b5a757-7176-499d-b93d-8feebb36c09f	1,635
57fe1954-e706-4ac9-a693-8f0d8b5faa5d	1,609
0bf95bf6-a789-4d91-8ad8-50d96a522c38	1,593
730bd0f7-b90b-46dd-b0c5-f310d4e9784a	1,503
4f33a4b0-1d58-402b-9046-16d2157e3b77	1,414
b4ad9889-08e7-4dd2-90c0-2fe3d2cb484a	1,409
bed34990-b428-4868-b4b3-e18e0980dedf	1,379

7. **Who are the top 10 users with the highest ratio of disconnect type 'peer' to total calls taken? What is the ratio for each of these users? How does this differ from disconnects 'client' and 'peer'? What can you infer from this?**

UserId	total_calls	peer_count	peer_ratio	client_count	client_ratio
dcfcf1bf-1435-4fcb-a677-7ce9b3368b36	5	5	1	0	0
53d692a1-d4da-4d9d-a256-33737d264b15	37	35	0.9459	0	0
3cee4f26-b17c-4a0e-8faa-22b2f5da37a9	33	30	0.9091	1	0.0303
c32de4f2-7feb-467f-b9ad-9ddfff4bdd39	10	9	0.9000	0	0
79fcd683-2d7d-4f33-b685-ea21c0bd72cb	38	34	0.8947	2	0.0526
9e320771-73b1-4592-819a-7f022b7fb43a	171	152	0.8889	8	0.0468
80676f08-9c92-4132-9233-30f00ebd5609	40	35	0.8750	4	0.1
248482c1-73b5-4e1c-bb47-8d6824ab1ca3	143	124	0.8671	9	0.0629
eb2417e8-cb69-4cd5-a7c3-a625c808bbc8	37	31	5	0.8378	0.1351
2e309d49-a3fa-40cc-bfb4-1d6dfda06a44	42	35	0.8333	2	0.0476

8. **How many inbound queue calls are received per month?**

Month	Inbound_calls
2023-01-01	57,871
2023-02-01	52,043
2023-03-01	54,822
2023-04-01	49,935
2023-05-01	47,810
2023-06-01	43,944
2023-07-01	41,282
2023-08-01	39,631
2023-09-01	33,387
2023-10-01	31,981
2023-11-01	26,157
2023-12-01	25,280
2024-01-01	38,177
2024-02-01	35,669
2024-03-01	38,793
2024-04-01	38,826
2024-05-01	36,221

2024-06-01	64,038
2024-07-01	35,151
2024-08-01	34,321
2024-09-01	31,029
2024-10-01	29,969
2024-11-01	27,771
2024-12-01	28,364

9. **What is the average number of calls taken during Q4 by month?**

Month & Year	avg_calls_q4
2023-10-01	28,255.333333333333
2023-11-01	28,255.333333333333
2023-12-01	28,255.333333333333
2024-10-01	28,255.333333333333
2024-11-01	28,255.333333333333
2024-12-01	28,255.333333333333

10. **What is the percentage difference of inbound queue calls taken in 2023 vs 2024?**

Calls_2024	Calls_2023	Pct_Difference
438,333	504,153	13.06

11. **What is the month and year with the most inbound queue calls?**

Month	Inbound calls
2024-06-01	64,038

SECTION C

For the next section, create a visualization to answer each question.

12. **How many calls were taken per month during 2023 and 2024?**
    - Attached is the snapshot of the Power BI
13. **What is the ratio of total talk time to total hold time per month for 2023 and 2024?**

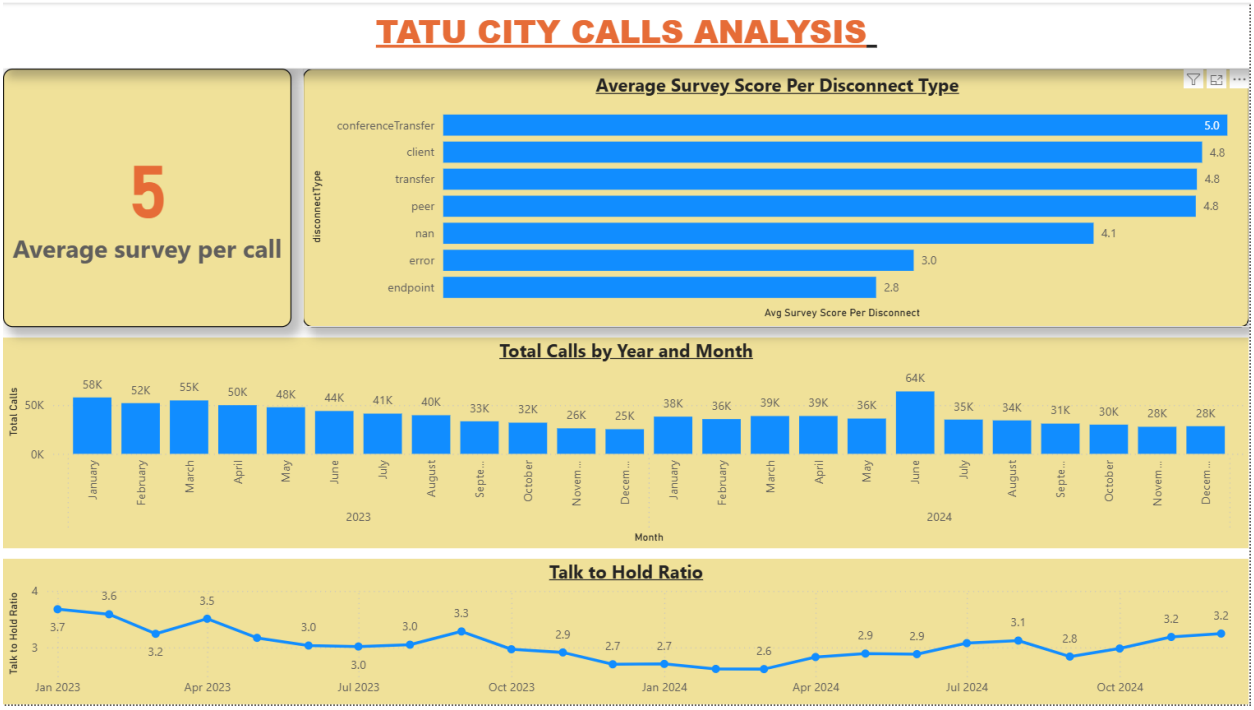
- Attached below is the snapshot of Power BI for your reference.

14. What is the average survey score per call?

- Attached below is the snapshot of Power BI for your reference.

15. What is the average survey score per disconnect type? What can you infer from this?

- Disconnect types strongly influence customer satisfaction. Normal call completions score highly, while technical or abnormal disconnections significantly reduce survey scores.
- Attached below is the snapshot of Power BI for your reference.



SECTION D

16. Is there a correlation between hold time and handle time?

Yes, there is typically a strong positive correlation between holding time and handle time. The correlation between hold time and handle time is approximately **0.75**, which indicates a strong positive relationship. This means that as hold time increases, handle time also tends to increase significantly. Calls with longer hold

times are likely to have longer overall handling times, suggesting that holding time is an important factor contributing to total handle time.

17. How is talk time affected by hold time?

The correlation between talk time and hold time is approximately **0.45**. This indicates a moderate positive relationship, meaning that calls with longer hold times tend to also have longer talk times, but the relationship is not very strong. It suggests that while holding time may influence talk time to some extent, other factors are also affecting talk time.

18. Is there a correlation between hold time and disconnect type? If so, please explain it.

Comparing the mean of hold times per disconnect type reveals:

- **More complex or error-prone disconnect types tend to have higher hold times**,
- While **simple or agent-ended disconnect types tend to have lower hold times**.

This provides useful operational insight even without a formal correlation coefficient.

19. Explain the steps you take to QA your data and results.

A. Data QA (Data Quality Assessment)

1. Schema validation
    - Check column types, date fields, numeric fields.
2. Missing values check
    - Identify nulls in key columns (StartDateTime, AgentID, CallType).
3. Outlier detection
    - Extremely high call durations, negative times, or impossible values.
4. Duplicate check
    - Ensure calls aren't double-logged.
5. Business rule consistency
    - HandleTime ≥ TalkTime + HoldTime
    - StartDateTime < EndDateTime

- SurveyScore within valid range.

6. Filtering invalid records
    - System tests, internal numbers, corrupted rows.

B. Analysis QA

1. Recalculate metrics manually for a random sample
    - Spot-check 5–10 rows.
2. Cross-validate totals
    - Compare sum of calculated fields vs. dataset raw fields.
3. Re-run calculations using a different method
    - SQL vs. Excel vs. Python.
4. Pivot table validation
    - Ensure grouping matches business logic.
5. Peer review
    - Another analyst or team member reviews formulas.

C. Visualization QA

1. Check axis values, labels, filters
2. Ensure no hidden filters misrepresent the data
3. Verify time granularity (month vs. week vs. day)

20. Group the shapes below based on their characteristics. List all possible groupings

By Shape Type

Circles / Curved Shapes

- Solid grey circle
- Blue ring (donut shape)
- Grey crescent

Polygons (5 sides)

- Purple striped pentagon
- Purple dotted pentagon

Quadrilaterals

- Hollow square (dashed outline)
- Purple striped parallelogram

By Fill Pattern / Texture

Solid Filled

- Grey circle
- Grey crescent

Striped Fill

- Purple striped pentagon
- Purple striped parallelogram

Dotted Fill

- Purple dotted pentagon

Hollow / No Fill

- Blue ring (only border)
- Hollow Square (dashed bord