Got it 👍 friend — since you have a **client call tomorrow**, I’ll prepare a crisp **summary + talking points** that you can use directly. This will cover:

1. **What we observed in the data (columns present, issues noticed)**
2. **Limitations & drawbacks**
3. **Extra data needed for deeper analysis**
4. **Smart questions to ask the client**

# **🔹 1. What’s in the Current Data**

* **Transcript-related fields**: Transcript text, Summary, Summary Title.
* **Call details**: Call ID, Timestamp, Caller Type, Phone Number.
* **Agent details**: Agent Name, Lan ID.
* **Outcome-related fields**: Call Outcome, Next Action, Disposition, Sentiment, Quality Score.
* **Dictionaries**: Complaint & Call Flags (detected, what happened, resolution, flags like Non-FCR, ID verification).

# **🔹 2. Limitations & Drawbacks**

* **Time Coverage**: Only **1 month of data** → cannot do monthly or seasonal trends. Only weekly insights possible.
* **Null / Missing Values**:  
  + **Summary / Summary Title** → many blanks → weakens topic analysis & word clouds.
  + **Disposition** → some rows missing → incomplete reason analysis.
  + **Next Action** → inconsistent → cannot fully map resolution journey.
  + **Complaint dictionary** → often incomplete (detected=true but what\_happened blank).
  + **Call Flags** → some truncated, not standardized.
* **Agent Metadata**: Only names given. Missing **team, supervisor, tenure, location** → can’t benchmark properly.
* **Customer Metadata**: Only caller type (Customer/Broker). Missing **segment, product type, tenure**.
* **Operational Metrics**: No **call duration, hold time, transfer counts, escalations, queue time**.
* **Feedback**: No **post-call CSAT/NPS surveys** to validate transcript analysis.

# **🔹 3. Extra Data You Should Request**

* **Extended timeframe**: At least **3–6 months of data** → for robust trend analysis.
* **Complete Complaint/Flag fields**: Ensure JSON is not truncated, include what\_happened, resolved\_on\_call, desired\_outcome.
* **Agent metadata**: Team, supervisor, location, shift, experience.
* **Customer metadata**: Segment, product type, tenure, churn risk indicators.
* **Operational data**: Call duration, hold time, transfers, escalations, queue wait.
* **Feedback data**: CSAT/NPS survey responses linked to calls.

# **🔹 4. Smart Questions to Ask the Client**

Here’s a list you can use in tomorrow’s call 👇

1. **Timeframe**
   * Can you provide more than one month of data (ideally 3–6 months) so we can show trends and seasonality?
2. **Complaint & Call Flags**
   * We see some complaints detected but missing details (“what happened”). Is that data available in another system?
   * Can we get fully populated Call Flag data (some rows are truncated)?
3. **Agent Metadata**
   * Do you have additional agent-level attributes (team, supervisor, tenure, shift, location)? This will help us show performance benchmarking and training needs.
4. **Customer Metadata**
   * Is there segmentation available (customer type, product, account tenure)? This will help us connect call reasons to customer value.
5. **Operational Metrics**
   * Do you track call duration, hold time, transfer counts, escalation flags, queue wait time? These will allow us to link operational pain points with customer sentiment.
6. **Feedback Validation**
   * Do you have CSAT or NPS survey data linked to these calls? This will validate transcript-based sentiment.

✅ **How you should frame it in the client call:**

* *“We did an initial analysis on the transcript dataset, and while it gives us a good start, we noticed some limitations: one month timeframe, nulls in summary/disposition/next action, incomplete complaints/flags. To provide stronger insights, especially for trends, root causes, and benchmarking, we’ll need some additional data.”*

👉 This way, you show **you worked on the data** but also push them to give **richer context**.