Absolutely—asking the **right questions** during exploratory data analysis (EDA) is key to uncovering **insights** that are **clear, relevant, and impactful** to your customer.

Here's a comprehensive list organized by theme. Think of it as your **EDA checklist**—your personal detective toolkit for getting to know any dataset and telling a compelling story about it.

---

1. General Data Understanding

- What is the shape of the dataset (rows, columns)?
- What does each column represent? What is the unit, if any?
- What are the data types (numerical, categorical, text, date)?
- Is this structured (tabular), time series, or unstructured (text/image) data?
- What is the primary entity or observation unit (a person, transaction, machine, etc.)?
- Is there a target variable? What are we trying to predict, understand, or optimize?

---

2. Data Quality

- Are there missing values? How many per column?
  - Are they random, patterned, or systemic?
- Are there duplicate rows?
- Are there constant or near-constant columns?
- Are there unexpected values (e.g., negative ages, future dates, special characters)?

---

3. Descriptive Statistics & Distributions

- What are the min, max, mean, median, std, skewness, and kurtosis for each numerical variable?
- Are distributions normal, skewed, or multimodal?
  - Are there extreme outliers? Are they errors or legitimate?
- Do any variables have large variance or a wide range?
- Which values appear most frequently (mode)?
- Do any numerical columns show heavy-tailed distributions?

---

4. Categorical Variables

- How many unique values are there per categorical column?
- What is the frequency of each category? Are any dominant?
- Are any categories rare or have too few occurrences?
- Are categories cleanly formatted (e.g., consistent spelling/casing)?

---

5. Time-Based Questions (if applicable)

- What is the time range of the data?
- Are there gaps or bursts in time?
- Are there temporal trends or seasonality?
- Are key metrics stable or changing over time?
- Is there a timestamp granularity issue (e.g., all dates the same, only daily resolution)?

---

6. Relationships Between Variables

- Which variables are strongly correlated (positively or negatively)?
- Are there multicollinear features (i.e., redundant information)?
- Are there interesting interactions between features (e.g., only when A and B happen together)?
- Are there group-level differences (e.g., average metric by category)?

---

7. Feature Engineering Opportunities

- Are there columns that can be combined (e.g., year + month)?
- Can you create ratios, flags, or aggregates?
- Are there temporal features to extract (e.g., day of week, hour)?
- Can you bucket or bin continuous variables (e.g., age groups, price bands)?
- Can you encode hierarchical or ordinal relationships?

---

8. Target Variable Exploration (if supervised task)

- Is the target balanced or imbalanced?
- How is the target variable distributed across key features?
- Are there data leakage risks (features that "cheat" by knowing the future)?
- How does the target evolve over time or by group?

---

9. Anomalies & Red Flags

- Are there clusters of missing data for certain rows/groups?
- Are there inconsistent relationships (e.g., younger person with high pension)?
- Are there suspicious timestamp patterns (e.g., all events logged at the same second)?
- Are there duplicated IDs, names, or transaction numbers?
- Do repeated exact values suggest automation, human error, or sensor defaults?

---

10. Communicating to the Customer

- What trends or insights matter most to the customer's domain?
- How does each insight tie back to the business problem or goal?
- Can you offer a possible explanation or hypothesis for each insight?
- What decisions could be influenced by these insights?
- Where should further data collection or deeper analysis happen?

Bonus: Your Narrative Template for Each Variable

For each important column, answer: 1. **What does this variable measure, and why does it matter?** 2. **How is it distributed?** 3. **What is typical (mean/median), and what is unusual (outliers)?** 4. **How does it relate to the target or other variables?** 5. **What should the customer know or do about it?**

Would you like me to turn this checklist into a printable reference card or Markdown template for reuse in your reports or projects?