**1. What are missing values and how do you handle them?**  
Missing values are blank spaces in the data. Like if someone didn’t write their age, it's missing.  
We can remove those rows or fill the blank with an average, zero, or "unknown".

**2. How do you treat duplicate records?**  
Duplicates mean the same row is repeated more than once. We remove extra copies to keep only one clean entry.

**3. What is the difference between dropna() and fillna()?**  
dropna means remove rows that have blanks.  
fillna means fill the blanks with a value like 0, average, or any text like "Not Available".

**4. What is outlier treatment and why is it important?**  
Outliers are values that are very different from others. Like age 150 when most are 20-40.  
We remove or fix them because they can give wrong results in analysis.

**5. What is standardizing data?**  
Standardizing means making data look the same.  
Like changing M to male, or writing all dates in the same format like day-month-year.

**6. How do you handle different formats in data (like date)?**  
If dates are written differently like 01-01-2024 or Jan 1, 2024, we convert all of them to one format using Excel or Python.

**7. What are some common data cleaning problems?**

* Missing values
* Duplicate rows
* Wrong data types
* Spelling mistakes
* Different formats
* Inconsistent text like "F", "female", "f"

**8. How do you check data quality?**  
I check if there are any blanks, duplicate entries, or wrong formats.  
Good data should be complete, correct, and clean before doing analysis.