```python
# 1. LOADING AND INSPECTING DATA
import pandas as pd
df = pd.read_csv("retail_sales_dataset.csv")
print(df)
```

```
     Transaction ID        Date Customer ID  Gender  Age Product
Category  \
0                 1  2023-11-24     CUST001    Male   34
Beauty
1                 2  2023-02-27     CUST002  Female   26
Clothing
2                 3  2023-01-13     CUST003    Male   50
Electronics
3                 4  2023-05-21     CUST004    Male   37
Clothing
4                 5  2023-05-06     CUST005    Male   30
Beauty
..              ...         ...         ...     ...  ...
...
995             996  2023-05-16     CUST996    Male   62
Clothing
996             997  2023-11-17     CUST997    Male   52
Beauty
997             998  2023-10-29     CUST998  Female   23
Beauty
998             999  2023-12-05     CUST999  Female   36
Electronics
999            1000  2023-04-12    CUST1000    Male   47
Electronics

     Quantity  Price per Unit  Total Amount
0           3              50           150
1           2             500          1000
2           1              30            30
3           1             500           500
4           2              50           100
..        ...             ...           ...
995         1              50            50
996         3              30            90
997         4              25           100
998         3              50           150
999         4              30           120

[1000 rows x 9 columns]
```

```python
# QUESTION 1:WHAT ARE  THE COILUMNS IN THE DATASET
# ANSWER:
# VERTICAL ATRIBUTES IN THE DATASET THAT SPECIFI THE CHARACTERISTICS
OF THE ENTITIES
# FOLLOWING ARE THE COLUMNS IN THE DATASET THAT I USED.
```

```
# # # there are total 9 columns:
# Transaction ID: A unique identifier for each transaction, allowing
tracking and reference.

# Date: The date when the transaction occurred, providing insights
into sales trends over time.

# Customer ID: A unique identifier for each customer, enabling
customer-centric analysis.

# Gender: The gender of the customer (Male/Female), offering insights
into gender-based purchasing patterns.

# Age: The age of the customer, facilitating segmentation and
exploration of age-related influences.

# Product Category: The category of the purchased product (e.g.,
Electronics, Clothing, Beauty), helping understand product
preferences.

# Quantity: The number of units of the product purchased, contributing
to insights on purchase volumes.

# Price per Unit: The price of one unit of the product, aiding in
calculations related to total spending.

# Total Amount: The total monetary value of the transaction,
showcasing the financial impact of each purchase.



# QUESTION2:
# DATATYPE OF EACH COILUMN
# finding datatype of columns in a dataset using pandas
import pandas as pd
df = pd.read_csv("retail_sales_dataset.csv")
print(df.dtypes)
```

```
Transaction ID       int64
Date                object
Customer ID         object
Gender              object
Age                  int64
Product Category    object
Quantity             int64
Price per Unit       int64
Total Amount         int64
dtype: object
```