```sql
7     -- 1st STEP IS EDA
8     -- Total transactions
9  •  SELECT COUNT(DISTINCT BillNo) FROM mytable;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

COUNT(DISTINCT BillNo)
21663

```sql
14        -- Average basket size
15  •    SELECT AVG(item_count)
16   ⊖   FROM (
17          SELECT BillNo, COUNT(DISTINCT Itemname) AS item_count
18          FROM mytable
19          GROUP BY BillNo
20        ) t;
21
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

AVG(item_count)
23.6015

```sql
11       -- Total products
12  •     SELECT COUNT(DISTINCT Itemname) FROM mytable;
13
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Con

COUNT(DISTINCT Itemname)
4169

```sql
    -- Data Quality Validation And Cleaning

    -- Handle Missing Values (NULL CHECK)
SELECT
    SUM(BillNo IS NULL) AS null_billno,
    SUM(Itemname IS NULL) AS null_itemname,
    SUM(Quantity IS NULL) AS null_quantity,
    SUM(Price IS NULL) AS null_price,
    SUM(CustomerID IS NULL) AS null_customer
FROM mytable;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content

null_billno	null_itemname	null_quantity	null_price	null_customer
0	0	0	0	0

```sql
46    -- Remove Cancelled Transactions (Retail Rule)
47  ● SELECT *
48    FROM mytable
49    WHERE BillNo NOT LIKE 'C%';
50
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

BillNo	Itemname	Quantity	Present_Date	Price	CustomerID	Country
536365	WHITE HANGING HEART T-LIGHT HOLDER	6	01-12-2010 08:26	2.550	17850	United Kingdom
536365	WHITE METAL LANTERN	6	01-12-2010 08:26	3.390	17850	United Kingdom
536365	CREAM CUPID HEARTS COAT HANGER	8	01-12-2010 08:26	2.750	17850	United Kingdom
536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	01-12-2010 08:26	3.390	17850	United Kingdom
536365	RED WOOLLY HOTTIE WHITE HEART	6	01-12-2010 08:26	3.390	17850	United Kingdom

```
51
52      -- Date Cleaning (Text to Date)
53 ●    SELECT
54        STR_TO_DATE(Present_Date, '%d-%m-%Y %H:%i') AS clean_date
55      FROM mytable;
56
```

Result Grid | ▦ ↔ Filter Rows: [   ] | Export: ▦ | Wrap Cell Content: 𝚰A | Fetch rows:

clean_date
▶ 2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:26:00
2010-12-01 08:28:00
2010-12-01 08:28:00
2010-12-01 08:34:00
2010-12-01 08:34:00
2010-12-01 08:34:00
2010-12-01 08:34:00
2010-12-01 08:34:00

```sql
57      -- FINAL CLEAN DATASET
58  •   SELECT DISTINCT
59        BillNo,
60        Itemname
61      FROM mytable
62      WHERE Quantity > 0
63        AND Price > 0
64        AND Itemname IS NOT NULL
65        AND BillNo NOT LIKE 'C%';
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

BillNo	Itemname
536365	CREAM CUPID HEARTS COAT HANGER
536365	GLASS STAR FROSTED T-LIGHT HOLDER
536365	KNITTED UNION FLAG HOT WATER BOTTLE
536365	RED WOOLLY HOTTIE WHITE HEART.
536365	SET 7 BABUSHKA NESTING BOXES
536365	WHITE HANGING HEART T-LIGHT HOLDER
536365	WHITE METAL LANTERN
536366	HAND WARMER RED POLKA DOT
536366	HAND WARMER UNION JACK
536367	ASSORTED COLOUR BIRD ORNAMENT
536367	BOX OF 6 ASSORTED COLOUR TEASPOONS
536367	BOX OF VINTAGE ALPHABET BLOCKS
536367	BOX OF VINTAGE JIGSAW BLOCKS
536367	DOORMAT NEW ENGLAND
536367	FELTCRAFT PRINCESS CHARLOTTE DOLL