TASK 5: RETAIL SALES DATA ANALYSIS

-- importing dataset

create database retail\_sales;

use retail\_sales;

show tables;

select \* from retails\_data limit 10;

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Transaction Id | Date | Customer Id | Gender | Age | Product Category | Quantity | Price per Unit | Total Amount |
| 1 | 2023-11-24 | CUST001 | Male | 34 | Beauty | 3 | 50 | 150 |
| 2 | 2023-02-27 | CUST002 | Female | 26 | Clothing | 2 | 500 | 1000 |
| 3 | 2023-01-13 | CUST003 | Male | 50 | Electronics | 1 | 30 | 30 |
| 4 | 2023-05-21 | CUST004 | Male | 37 | Clothing | 1 | 500 | 500 |
| 5 | 2023-05-06 | CUST005 | Male | 30 | Beauty | 2 | 50 | 100 |
| 6 | 2023-04-25 | CUST006 | Female | 45 | Beauty | 1 | 30 | 30 |
| 7 | 2023-03-13 | CUST007 | Male | 46 | Clothing | 2 | 25 | 50 |
| 8 | 2023-02-22 | CUST008 | Male | 30 | Electronics | 4 | 25 | 100 |
| 9 | 2023-12-13 | CUST009 | Male | 63 | Electronics | 2 | 300 | 600 |
| 10 | 2023-10-07 | CUST010 | Female | 52 | Clothing | 4 | 50 | 200 |

# -- Data cleaning -- check null values

select \* from retails\_data

where 'Transaction Id' is null

or 'Date' is null

or 'Customer Id' is null

or 'Gender' is null

or 'Age' is null

or 'Product Category' is null

or 'Quantity' is null

or 'Price per Unit' is null

or 'Total Amount' is null;

## output:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Transaction Id | Date | Customer Id | Gender | Age | Product Category | Quantity | Price per Unit | Total Amount |

# -- exploration and business insights

select COUNT(\*) as total\_transactions from retails\_data;

## output:

|  |
| --- |
| Total transaction |
| 1000 |

select COUNT(distinct `Customer ID`) as unique\_customers from retails\_data;

|  |
| --- |
| Unique\_customer |
| 1000 |

select distinct `Product Category` from retails\_data;

## output:

|  |
| --- |
| Product Category |
| Beauty |
| Clothing |
| Electronics |

select `Product Category`, SUM(`Total Amount`) as total\_sales from retails\_data group by `Product Category` order by total\_sales desc;

## output:

|  |  |
| --- | --- |
| Product Category | Total sales |
| Beauty | 156905 |
| Clothing | 155580 |
| Electronics | 143515 |

select `Gender`, COUNT(\*) as transactions from retails\_data group by `Gender`;

## output:

|  |  |
| --- | --- |
| Gender | Transaction |
| male | 490 |
| female | 510 |

select `customer id`,sum(`total amount`) as total\_spent from retails\_data group by `customer id`order by total\_spent desc;

## output:

|  |  |
| --- | --- |
| Customer Id | Total spent |
| CUST416 | 2000 |
| CUST124 | 2000 |
| CUST743 | 2000 |
| CUST561 | 2000 |
| CUST447 | 2000 |
| CUST946 | 2000 |
| CUST109 | 2000 |
| CUST139 | 2000 |
| CUST074 | 2000 |
| CUST970 | 2000 |
| CUST416 | 2000 |

select round(avg(`age`),2) as avg\_age

from retails\_data

where `product category` = 'Beauty';l

|  |  |
| --- | --- |
| Customer Id | Total spent |
| CUST416 | 2000 |
| CUST124 | 2000 |
| CUST743 | 2000 |
| CUST561 | 2000 |
| CUST447 | 2000 |
| CUST946 | 2000 |
| CUST109 | 2000 |
| CUST139 | 2000 |
| CUST074 | 2000 |
| CUST970 | 2000 |
| CUST416 | 2000 |

# -- view

create view v\_sales\_by\_category as select `product category`, sum(`total amount`) as total\_sales from retails\_data group by `product category`;

create view v\_top\_customers as select `customer id`, sum(`total amount`) as total\_spent from retails\_data group by `customer id` order by total\_spent desc limit 5;

create view v\_transactions\_by\_gender as select `gender`, count(\*) as transactions from retails\_data group by `gender`;

create view v\_high\_value\_txn as select \* from retails\_data where `total amount` > 1000;

select \* from v\_sales\_by\_category;

## output:

|  |  |
| --- | --- |
| Product Category | Total sales |
| Beauty | 143515 |
| Clothing | 155580 |
| Electronics | 156905 |

select \* from v\_top\_customers;

## output:

|  |  |
| --- | --- |
| Customer Id | Total spent |
| CUST416 | 2000 |
| CUST124 | 2000 |
| CUST743 | 2000 |
| CUST561 | 2000 |
| CUST447 | 2000 |

select \* from v\_transactions\_by\_gender;

## output:

|  |  |
| --- | --- |
| Gender | Transaction |
| male | 490 |
| female | 510 |

select \* from v\_high\_value\_txn;

## output:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Transaction Id | Date | Customer Id | Gender | Age | Product Category | Quantity | Price per Unit | Total Amount |
| 13 | 2023-08-05 | CUST013 | Male | 22 | Electronics | 3 | 500 | 1500 |
| 15 | 2023-01-16 | CUST015 | Female | 42 | Electronics | 4 | 500 | 2000 |
| 16 | 2023-02-17 | CUST016 | Male | 19 | Clothing | 3 | 500 | 1500 |
| 31 | 2023-05-23 | CUST031 | Male | 44 | Electronics | 4 | 300 | 1200 |
| 46 | 2023-06-26 | CUST046 | Female | 20 | Electronics | 4 | 300 | 1200 |
| 47 | 2023-11-06 | CUST047 | Female | 40 | Beauty | 3 | 500 | 1500 |
| 54 | 2023-02-10 | CUST054 | Female | 38 | Electronics | 3 | 500 | 1500 |
| 58 | 2023-11-13 | CUST058 | Male | 18 | Clothing | 4 | 300 | 1200 |
| 65 | 2023-12-05 | CUST065 | Male | 51 | Electronics | 4 | 500 | 2000 |
| 67 | 2023-05-29 | CUST067 | Female | 48 | Beauty | 4 | 300 | 1200 |

select column\_name, data\_type from information\_schema.columns where table\_name = 'retails\_data';

## output:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Transaction Id | Date | Customer Id | Gender | Age | Product Category | Quantity | Price per Unit | Total Amount |
| 13 | 2023-08-05 | CUST013 | Male | 22 | Electronics | 3 | 500 | 1500 |
| 15 | 2023-01-16 | CUST015 | Female | 42 | Electronics | 4 | 500 | 2000 |
| 16 | 2023-02-17 | CUST016 | Male | 19 | Clothing | 3 | 500 | 1500 |
| 31 | 2023-05-23 | CUST031 | Male | 44 | Electronics | 4 | 300 | 1200 |
| 46 | 2023-06-26 | CUST046 | Female | 20 | Electronics | 4 | 300 | 1200 |
| 47 | 2023-11-06 | CUST047 | Female | 40 | Beauty | 3 | 500 | 1500 |
| 54 | 2023-02-10 | CUST054 | Female | 38 | Electronics | 3 | 500 | 1500 |
| 58 | 2023-11-13 | CUST058 | Male | 18 | Clothing | 4 | 300 | 1200 |
| 65 | 2023-12-05 | CUST065 | Male | 51 | Electronics | 4 | 500 | 2000 |
| 67 | 2023-05-29 | CUST067 | Female | 48 | Beauty | 4 | 300 | 1200 |