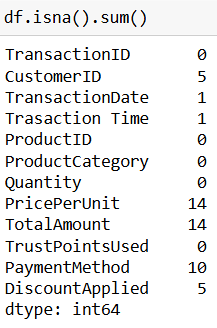
**Transaction Sales Data.**

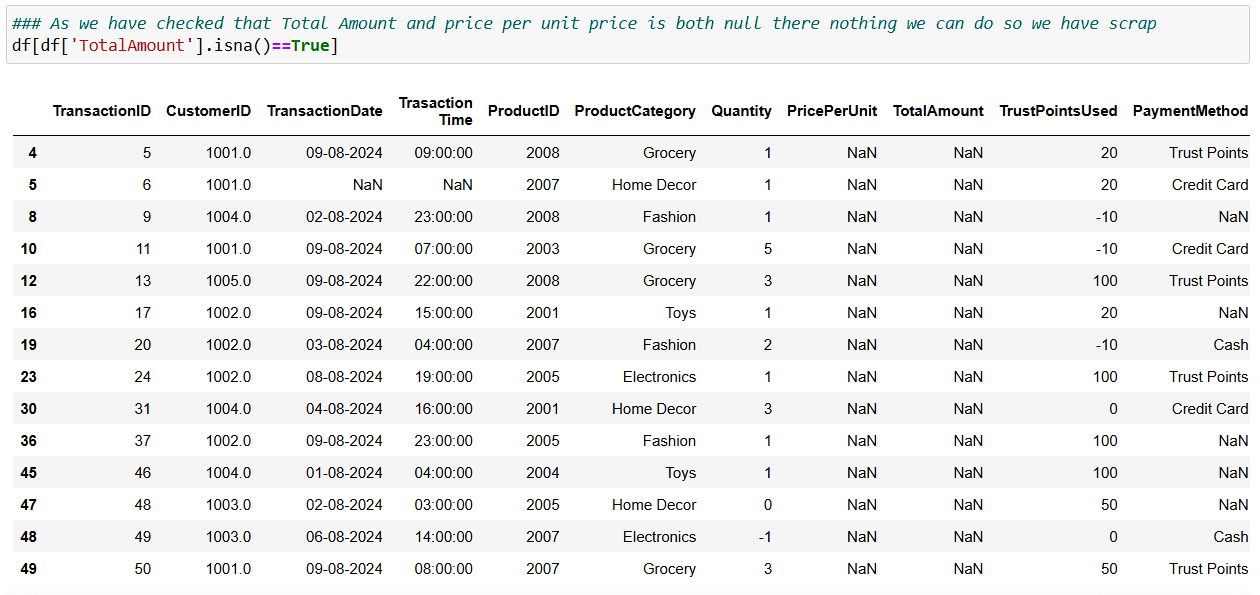
1. **Data preprocessing steps which I followed.**

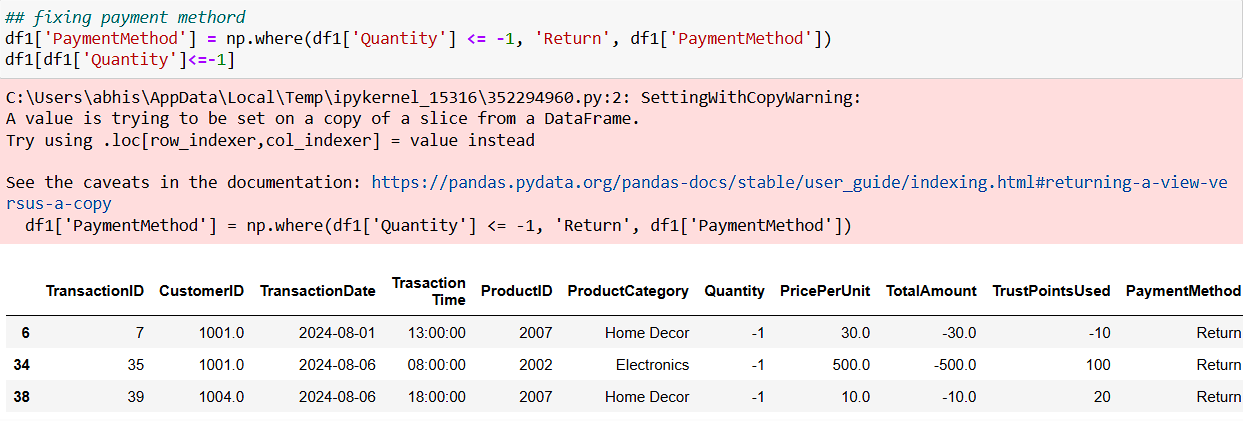
**First of all find the null values which is present in the datasets.**

****

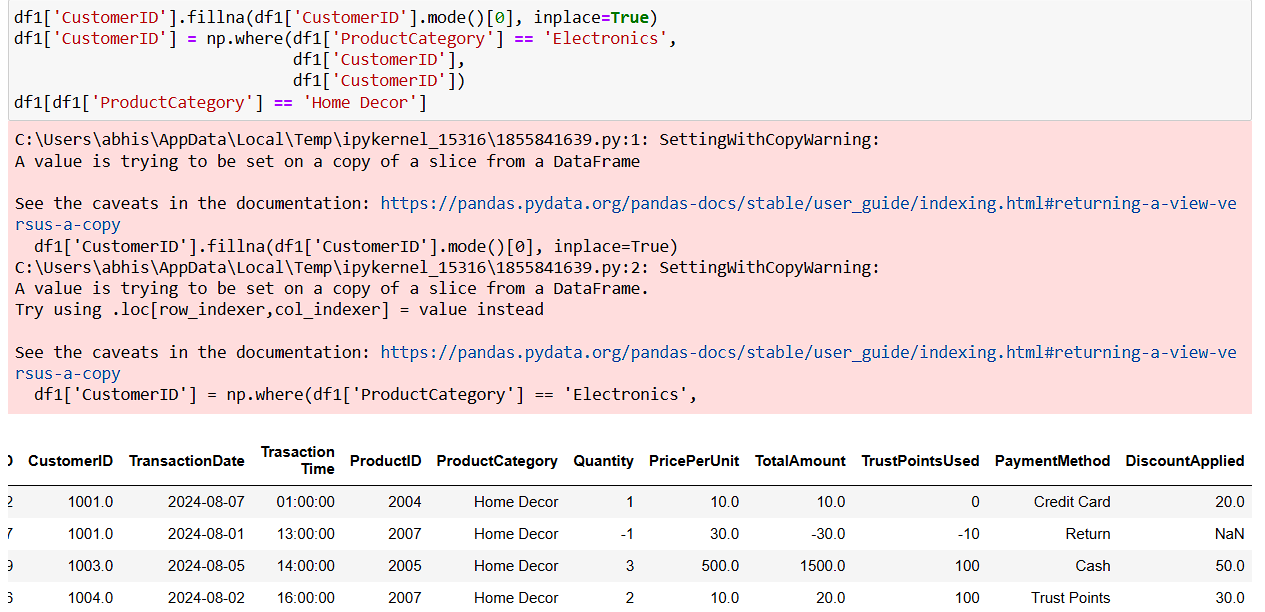
**Remove the all the rows which has maximum number of the null values as we have found that**

**price Per Unit and Total Amount.**

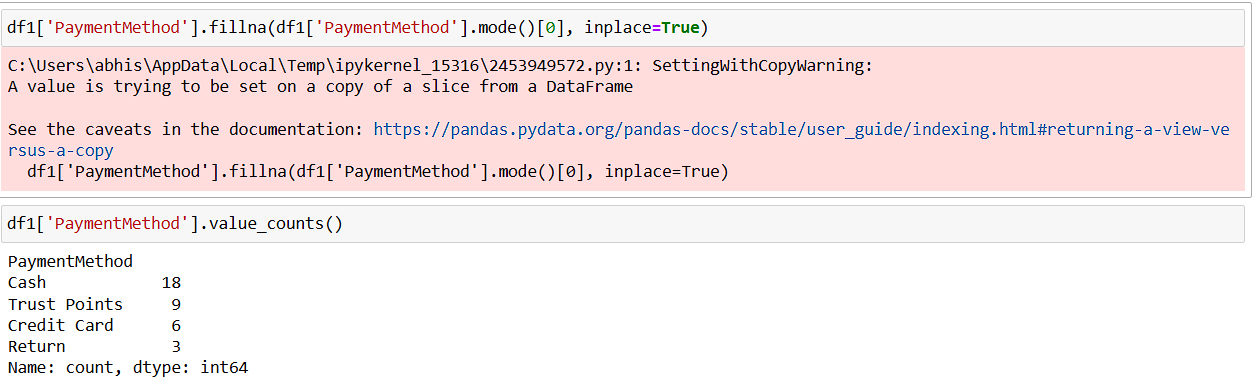
****

**As we see some quantity is -ve so we assume that -ve value shows that return the product** 

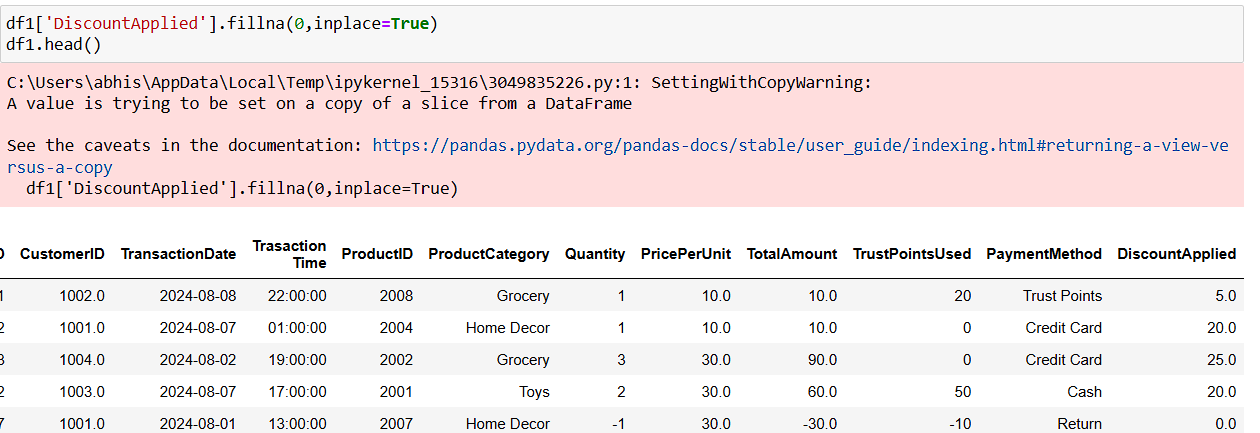
**We also found that CustomerID has some null values to clear those null values. So we look who is most frequent buyer for product from each category.**

****

**There is some transaction the pay method is not define to figure out this we replaces null values with most frequent payment method is used.**

****

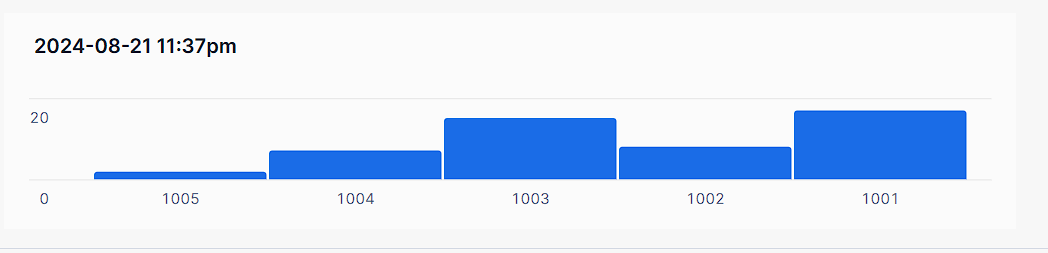
**We also found that discount value is also null so as we assume that in such cases no discount is provided.**

****

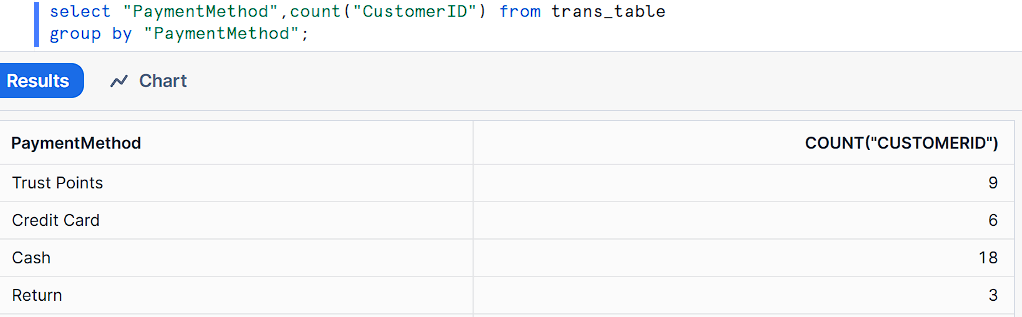
1. **Aggregation we used like sum and count is most used full in Quantity and Total amount**

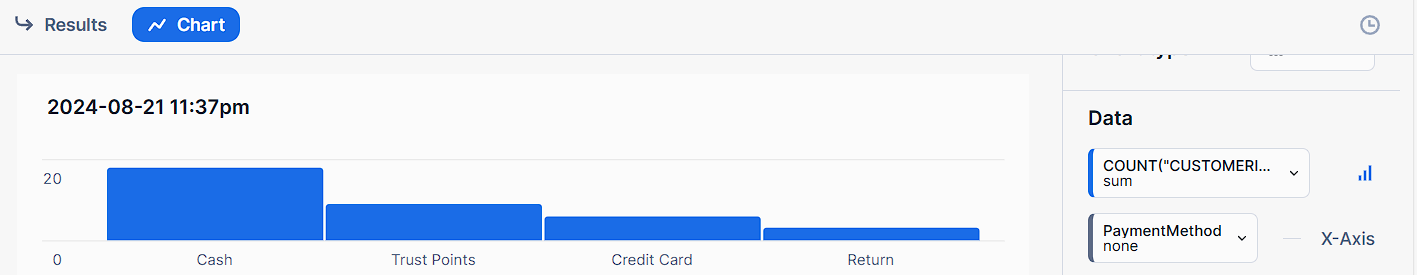
* **Which customer order most quantity product**

****

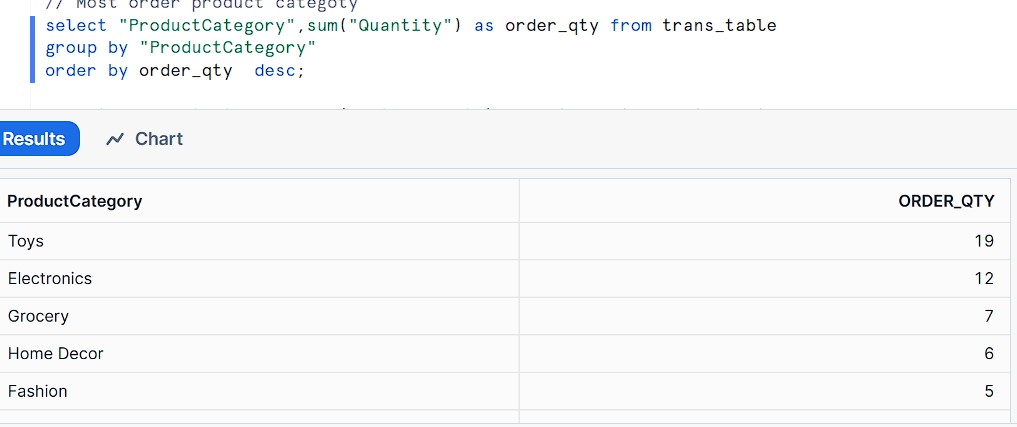
****

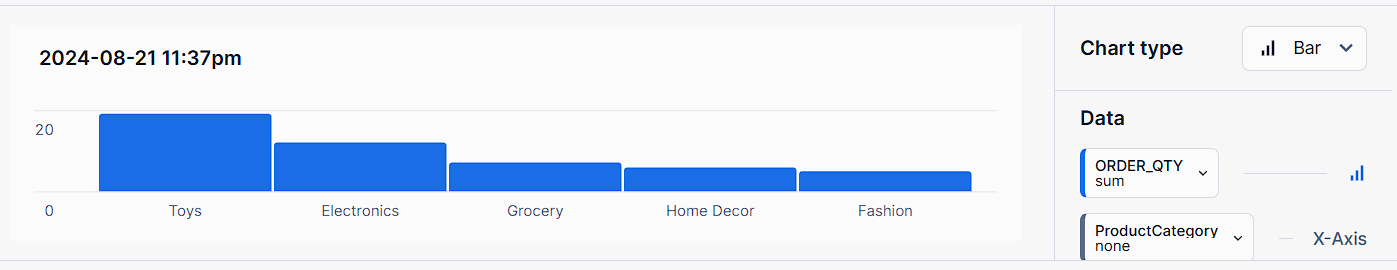
* **Which Payment Method is most frequently used**

****

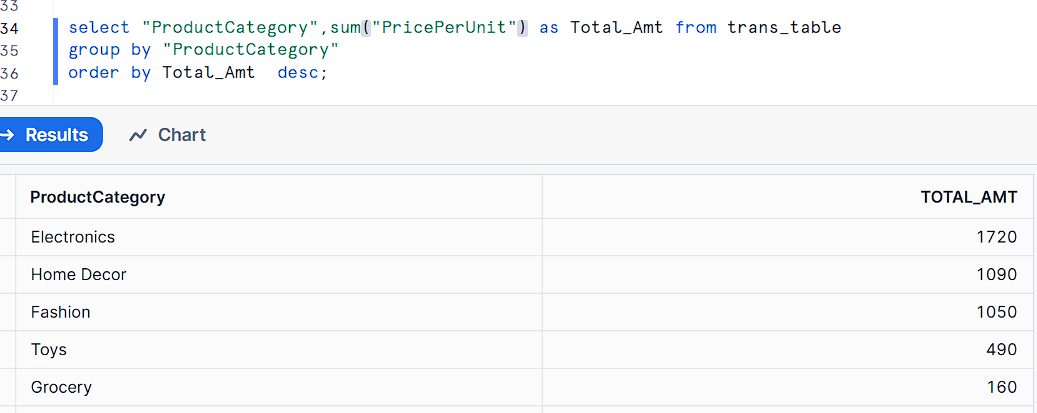
****

* **Which product category is most order by the customer.**

****

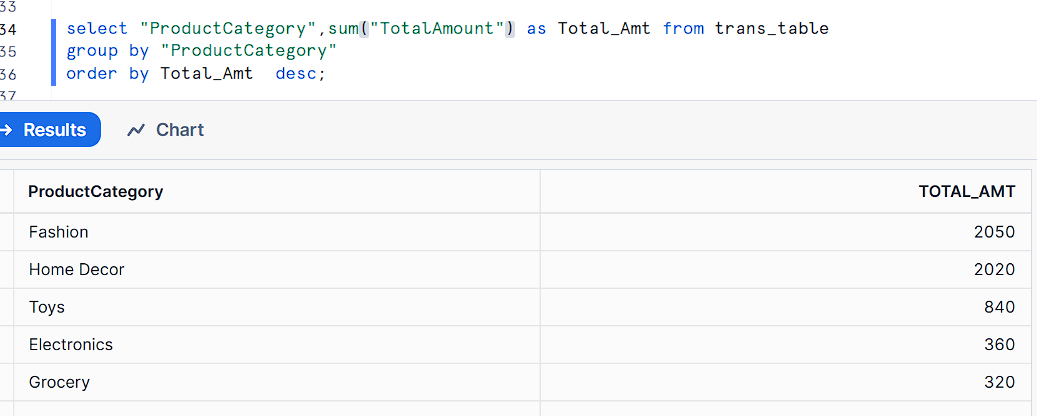
****

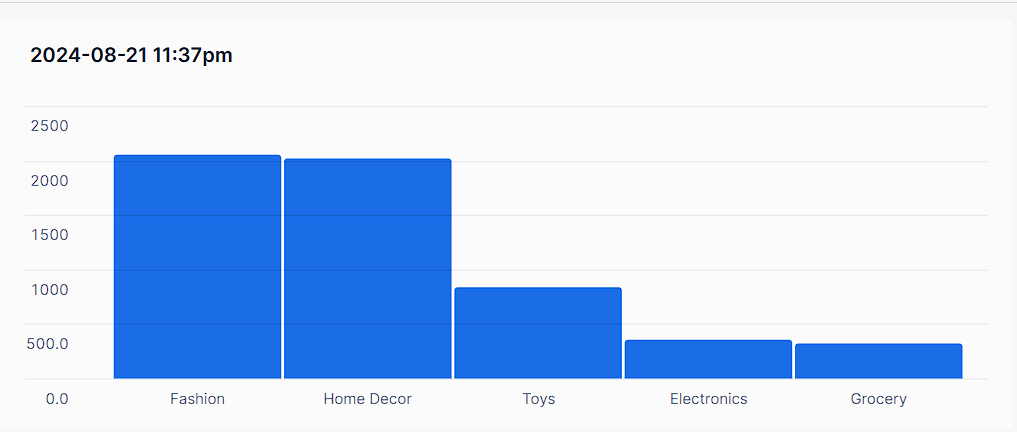
* **Which most is expensive product category.**

****

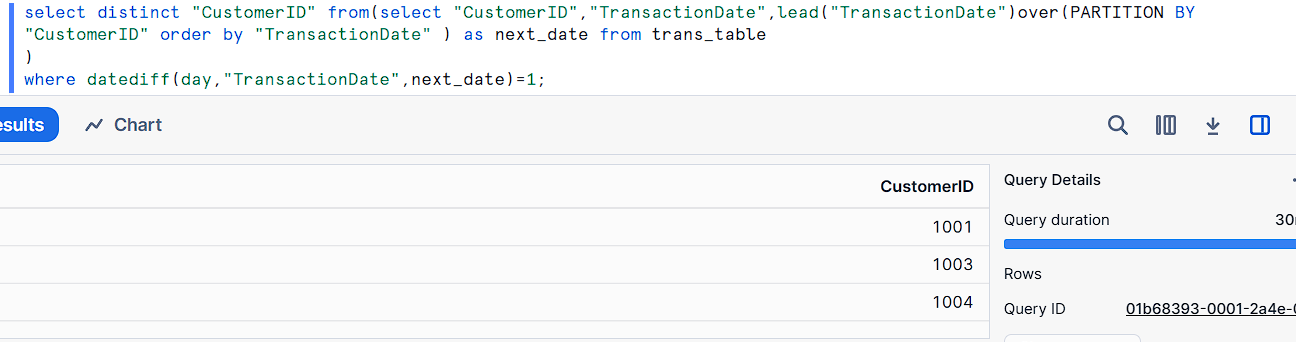
****

* **Which is most product category generates maximum revenue**

****

****

**3 Using Window function to found that who is most consecutive buyer 2 days.**

****

**All SQL CODE:**

**create database data\_yuma\_Assignment;**

**use data\_yuma\_Assignment;**

**CREATE or replace TABLE Trans\_table (**

**"Row\_Number" FLOAT NOT NULL,**

**"TransactionID" FLOAT NOT NULL,**

**"CustomerID" FLOAT NOT NULL,**

**"TransactionDate" DATE NOT NULL,**

**"Trasaction Time" varchar NOT NULL,**

**"ProductID" FLOAT NOT NULL,**

**"ProductCategory" VARCHAR NOT NULL,**

**"Quantity" FLOAT NOT NULL,**

**"PricePerUnit" FLOAT NOT NULL,**

**"TotalAmount" FLOAT NOT NULL,**

**"TrustPointsUsed" FLOAT NOT NULL,**

**"PaymentMethod" VARCHAR NOT NULL,**

**"DiscountApplied" FLOAT**

**);**

**//The customer which order most of thing**

**select "CustomerID",sum("Quantity")as orde\_qty from trans\_table**

**group by "CustomerID"**

**order by orde\_qty desc;**

**//The PaymentMethod used to most orders**

**select "PaymentMethod",count("CustomerID") from trans\_table**

**group by "PaymentMethod";**

**// Most order product categoty**

**select "ProductCategory",sum("Quantity") as order\_qty from trans\_table**

**group by "ProductCategory"**

**order by order\_qty desc;**

**select "ProductCategory",sum("TotalAmount") as Total\_Amt from trans\_table**

**group by "ProductCategory"**

**order by Total\_Amt desc;**

**select dayname("TransactionDate") as Day\_name,sum("TotalAmount")as Total\_Amt from trans\_table**

**group by Day\_name**

**order by Total\_Amt desc;**

**// The user who purchase in 2 consecutive days**

**select distinct "CustomerID" from(select "CustomerID","TransactionDate",lead("TransactionDate")over(PARTITION BY "CustomerID" order by "TransactionDate" ) as next\_date from trans\_table**

**)**

**where datediff(day,"TransactionDate",next\_date)=1;**