**Walmart Sales Data Analysis**

1. find total no of rows & duplicates rows.

select count(\*) from walmartsalesdata;

SELECT \*

FROM walmartsalesdata

WHERE `Invoice ID` IN (

SELECT `Invoice ID`

FROM walmartsalesdata

GROUP BY `Invoice ID`

HAVING COUNT(\*) > 1

);

1. find null values.

select \* from walmartsalesdata

where 'invoice id' is null;

1. if multiple null values found then replace it (currently no null value).

UPDATE walmartsalesdata

SET column1 = 'N/A', column2 = 0, column3 = 'Unknown'

WHERE column1 IS NULL OR column2 IS NULL OR column3 IS NULL;

1. how to check/identify the column datatype.

SELECT 'Quantity', DATA\_TYPE

FROM INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'walmartsalesdata'

and column\_name = 'Quantity';

1. how to change the column datatype.

ALTER TABLE walmartsalesdata

MODIFY `branch` VARCHAR(20) NOT NULL,

MODIFY `city` VARCHAR(20) NOT NULL,

MODIFY `Customer type` VARCHAR(30) NOT NULL,

MODIFY `Gender` VARCHAR(10) NOT NULL,

MODIFY `Product line` VARCHAR(100) NOT NULL,

MODIFY `Unit price` decimal(10,2) NOT NULL,

MODIFY `Quantity` int NOT NULL,

MODIFY `Tax 5%` decimal(6,4) NOT NULL,

MODIFY `Total` decimal(12,4) NOT NULL,

MODIFY `date` datetime NOT NULL,

MODIFY `time` time NOT NULL,

MODIFY `payment` VARCHAR(20) NOT NULL,

MODIFY `cogs` decimal(10,2) NOT NULL,

MODIFY `gross margin percentage` decimal(11,9) not null,

MODIFY `gross income` decimal(12,4) NOT NULL,

MODIFY `Rating` decimal(6,4) NOT NULL;

1. Add new column i,e. time\_of\_day.

alter table walmartsalesdata

modify time\_of\_day varchar(30);

set sql\_safe\_updates = 0;

update walmartsalesdata

set time\_of\_day =

(case

when time between '00:00:00' and '12:00:00' then 'Morning'

when time between '12:01:00' and '16:00:00' then 'Afternoon'

else 'Evening'

end);

select \* from walmartsalesdata;

1. create new column i.e, find day\_name.

alter table walmartsalesdata

add column day\_name varchar(20);

update walmartsalesdata

set day\_name = dayname(date)

where day\_name is null;

1. Add new column i,e. find month\_name

alter table walmartsalesdata

add column month\_name varchar(20);

update walmartsalesdata

set month\_name = monthname(date)

where month\_name is null;

1. How many unique cities does the data have ?

select distinct(city) as unique\_city from walmartsalesdata;

1. in which city is each branch ?

select distinct city , branch as unique\_city\_branch from walmartsalesdata;

1. How many unique product lines does the data have ?

select count(distinct `product line`) as unique\_product\_line from walmartsalesdata;

1. what is most common payment method

select payment , count(payment) from walmartsalesdata

group by payment;

1. what is most selling product line

select `Product line` , count(`Product line`) as most\_selling\_product from walmartsalesdata

group by `Product line`

order by most\_selling\_product desc;

1. what is total revenue by month

select month\_name , sum(total) as revenue from walmartsalesdata

group by month\_name

order by revenue desc;

1. what months had largest cogs

select month\_name , sum(cogs) as largest\_cogs from walmartsalesdata

group by month\_name

order by largest\_cogs desc;

1. what product line had the largest revenue ?

select `Product line` , sum(total) as revenue from walmartsalesdata

group by `Product line`

order by revenue desc;

1. what is the city with the largest revenue ?

select city , sum(total) as revenue from walmartsalesdata

group by city

order by revenue desc;

1. which branch sold more products than avg products sold ?

select branch , sum(quantity) as more\_sold from walmartsalesdata

group by branch

having more\_sold > (select avg(Quantity) from walmartsalesdata);

1. what is the most common poduct line by gender

select `Product line` , gender , count(\*) as total\_qty from walmartsalesdata

group by `Product line` , Gender

order by total\_qty desc;

1. what is average rating of each product line

select `Product line` , avg(rating) as avg\_rate from walmartsalesdata

group by `Product line`

order by avg\_rate desc;