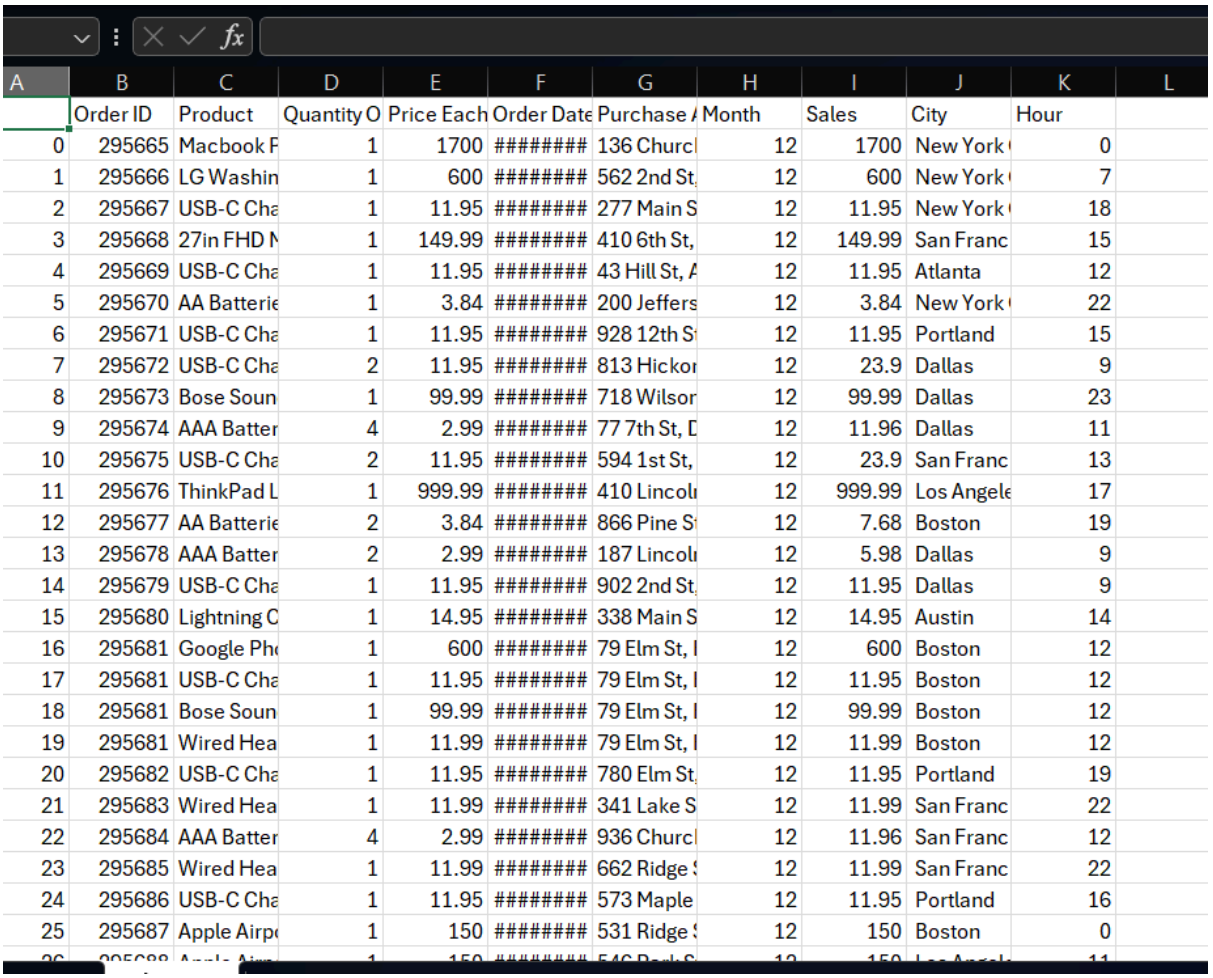
PROJECT - SALES DATA OF PRODUCTS USING EXCEL,SQL AND POWER BI

Our goal-

identify the best-selling products, calculate revenue metrics such as total sales and , and create visualisations to present our findings effectively from the sample sales data for various products sold (kaggle)

EXCEL-



	A	B	C	D	E	F	G	H	I	J	K	L
		Order ID	Product	Quantity O	Price Each	Order Date	Purchase A	Month	Sales	City	Hour	
0		295665	Macbook P	1	1700	########	136 Churcl	12	1700	New York (	0	
1		295666	LG Washin	1	600	########	562 2nd St,	12	600	New York (	7	
2		295667	USB-C Cha	1	11.95	########	277 Main S	12	11.95	New York (	18	
3		295668	27in FHD M	1	149.99	########	410 6th St,	12	149.99	San Franc	15	
4		295669	USB-C Cha	1	11.95	########	43 Hill St, A	12	11.95	Atlanta	12	
5		295670	AA Batterie	1	3.84	########	200 Jeffers	12	3.84	New York (	22	
6		295671	USB-C Cha	1	11.95	########	928 12th St	12	11.95	Portland	15	
7		295672	USB-C Cha	2	11.95	########	813 Hickor	12	23.9	Dallas	9	
8		295673	Bose Soun	1	99.99	########	718 Wilsor	12	99.99	Dallas	23	
9		295674	AAA Batter	4	2.99	########	77 7th St, D	12	11.96	Dallas	11	
10		295675	USB-C Cha	2	11.95	########	594 1st St,	12	23.9	San Franc	13	
11		295676	ThinkPad L	1	999.99	########	410 Lincolı	12	999.99	Los Angele	17	
12		295677	AA Batterie	2	3.84	########	866 Pine St	12	7.68	Boston	19	
13		295678	AAA Batter	2	2.99	########	187 Lincolı	12	5.98	Dallas	9	
14		295679	USB-C Cha	1	11.95	########	902 2nd St,	12	11.95	Dallas	9	
15		295680	Lightning C	1	14.95	########	338 Main S	12	14.95	Austin	14	
16		295681	Google Phc	1	600	########	79 Elm St, I	12	600	Boston	12	
17		295681	USB-C Cha	1	11.95	########	79 Elm St, I	12	11.95	Boston	12	
18		295681	Bose Soun	1	99.99	########	79 Elm St, I	12	99.99	Boston	12	
19		295681	Wired Hea	1	11.99	########	79 Elm St, I	12	11.99	Boston	12	
20		295682	USB-C Cha	1	11.95	########	780 Elm St,	12	11.95	Portland	19	
21		295683	Wired Hea	1	11.99	########	341 Lake S	12	11.99	San Franc	22	
22		295684	AAA Batter	4	2.99	########	936 Churcl	12	11.96	San Franc	12	
23		295685	Wired Hea	1	11.99	########	662 Ridge S	12	11.99	San Franc	22	
24		295686	USB-C Cha	1	11.95	########	573 Maple	12	11.95	Portland	16	
25		295687	Apple Airp	1	150	########	531 Ridge S	12	150	Boston	0	

This is the sample data with 185951 rows and 11 columns

As we can see that this dataset table has various inconsistencies like the column order_date is in the wrong format category and the product and purchase address has incorrect cell width and price is not in correct format and the integer row number doesn't have a column name at all.
So first we will clean the data using excel -

step 1

a) cleaning the data and changing basic column width and proper formatting category
b) fixed columns width for purchase address and used proper date format for (order_date) column and
checked for missing data and inconsistencies
c) gave first column the name (id) because it was empty and had attributes of a unique identifier

d) then converted the excel file to csv file and imported it to sql compiler (sqlite online)

id	Order ID	Product	Quantity O	Price Each	Order Date	Purchase Address	Month	Sales	City	Hour
0	295665	Macbook Pro Laptop	1	₹1,700.00	30-12-2019	136 Church St, New York City, NY 10001	12	1700	New York	0
1	295666	LG Washing Machine	1	₹600.00	29-12-2019	562 2nd St, New York City, NY 10001	12	600	New York	7
2	295667	USB-C Charging Cable	1	₹11.95	12-12-2019	277 Main St, New York City, NY 10001	12	11.95	New York	18
3	295668	27in FHD Monitor	1	₹149.99	22-12-2019	410 6th St, San Francisco, CA 94016	12	149.99	San Franc	15
4	295669	USB-C Charging Cable	1	₹11.95	18-12-2019	43 Hill St, Atlanta, GA 30301	12	11.95	Atlanta	12
5	295670	AA Batteries (4-pack)	1	₹3.84	31-12-2019	200 Jefferson St, New York City, NY 10001	12	3.84	New York	22
6	295671	USB-C Charging Cable	1	₹11.95	16-12-2019	928 12th St, Portland, OR 97035	12	11.95	Portland	15
7	295672	USB-C Charging Cable	2	₹11.95	13-12-2019	813 Hickory St, Dallas, TX 75001	12	23.9	Dallas	9
8	295673	Bose SoundSport Headphones	1	₹99.99	15-12-2019	718 Wilson St, Dallas, TX 75001	12	99.99	Dallas	23
9	295674	AAA Batteries (4-pack)	4	₹2.99	28-12-2019	77 7th St, Dallas, TX 75001	12	11.96	Dallas	11
10	295675	USB-C Charging Cable	2	₹11.95	13-12-2019	594 1st St, San Francisco, CA 94016	12	23.9	San Franc	13
11	295676	ThinkPad Laptop	1	₹999.99	28-12-2019	410 Lincoln St, Los Angeles, CA 90001	12	999.99	Los Angele	17
12	295677	AA Batteries (4-pack)	2	₹3.84	20-12-2019	866 Pine St, Boston, MA 02215	12	7.68	Boston	19
13	295678	AAA Batteries (4-pack)	2	₹2.99	06-12-2019	187 Lincoln St, Dallas, TX 75001	12	5.98	Dallas	9
14	295679	USB-C Charging Cable	1	₹11.95	25-12-2019	902 2nd St, Dallas, TX 75001	12	11.95	Dallas	9
15	295680	Lightning Charging Cable	1	₹14.95	01-12-2019	338 Main St, Austin, TX 73301	12	14.95	Austin	14
16	295681	Google Phone	1	₹600.00	25-12-2019	79 Elm St, Boston, MA 02215	12	600	Boston	12
17	295681	USB-C Charging Cable	1	₹11.95	25-12-2019	79 Elm St, Boston, MA 02215	12	11.95	Boston	12
18	295681	Bose SoundSport Headphones	1	₹99.99	25-12-2019	79 Elm St, Boston, MA 02215	12	99.99	Boston	12
19	295681	Wired Headphones	1	₹11.99	25-12-2019	79 Elm St, Boston, MA 02215	12	11.99	Boston	12
20	295682	USB-C Charging Cable	1	₹11.95	23-12-2019	780 Elm St, Portland, OR 97035	12	11.95	Portland	19
21	295683	Wired Headphones	1	₹11.99	23-12-2019	341 Lake St, San Francisco, CA 94016	12	11.99	San Franc	22
22	295684	AAA Batteries (4-pack)	4	₹2.99	04-12-2019	936 Church St, San Francisco, CA 94016	12	11.96	San Franc	12
23	295685	Wired Headphones	1	₹11.99	11-12-2019	662 Ridge St, San Francisco, CA 94016	12	11.99	San Franc	22
24	295686	USB-C Charging Cable	1	₹11.95	17-12-2019	573 Maple St, Portland, ME 04101	12	11.95	Portland	16
25	295687	Apple Airpods Headphones	1	₹150.00	10-12-2019	531 Ridge St, Boston, MA 02215	12	150	Boston	0

**********8

Now for the second step we are going to use sql to answer various business questions

step 2 -

Question 1

in sql compiler

we used this code below to analyse the top 5 best selling product in the table

****

with cte as
(SELECT * FROM sales)

select product,round(sum(Sales),0) su from cte

group by 1
order by su desc
Limit 5
****

SQLite

```sql
WITH cte AS
(SELECT * FROM sales)

SELECT product,round(sum(Sales),0) su  FROM cte
GROUP BY 1
ORDER BY su DESC
LIMIT 5
```

Product	su
Macbook Pro Laptop	8037600
iPhone	4794300
ThinkPad Laptop	4129959
Google Phone	3319200
27in 4K Gaming Monitor	2435098

this is the output for the question 1
****

Question 2

with this code we will find sales trend over month time along with its cumulative sum

****

```
with cte as
(SELECT * FROM sales)
,
cte2 as
(select product ,month,round(sum(Sales),0) as sum from cte
group by 1,2
order by 1,2)
select *,sum(sum) over(partition by product order by month) as cumul from cte2
*****
```

product	month	sum	cumul
20in Monitor	1	23978	23978
20in Monitor	2	27058	51036
20in Monitor	3	35857	86893
20in Monitor	4	43226	130119
20in Monitor	5	37507	167626
20in Monitor	6	35417	203043
20in Monitor	7	35967	239010

This is the output for question 2

****

Question 3

with this code when can see the month in which the product had maximum sales and its total sale amount till now.

****

```
with cte as
(SELECT * FROM sales)
,
cte2 as
(select product ,month,round(sum(Sales),0) as sum from cte
group by 1,2
order by 1,2),
cte3 as
(select *,sum(sum) over(partition by product order by month) as cumul from cte2 ),
cte4 as
(select product,month,max(sum) as maximum_sale,max(cumul) as total_sale
```

```
from cte3
group by 1
order by 1,2)
select * from cte4
```

****

product	month	maximum_sale	total_sale
20in Monitor	12	62584	454150
27in 4K Gaming Monitor	12	336951	2435097
27in FHD Monitor	12	144290	1132424
34in Ultrawide Monitor	12	322612	2355557
AA Batteries (4-pack)	12	14300	106120
AAA Batteries (4-pack)	12	12681	92741
Apple Airpods Headphones	12	311400	2349150

This is the output for question 3

****

Question 4

with this we can see best selling products,their sum of sales per month and total cumulative sum till now

****

```
with cte as
(SELECT * FROM sales)
,
cte2 as
(select product ,month,round(sum(Sales),0) as sum from cte
group by 1,2
order by 1,2),
cte3 as
(select *,sum(sum) over(partition by product order by month) as cumul from cte2 ),
cte4 as
(select product,month,max(sum) as maximum_sale,max(cumul) as total_sale
from cte3
group by 1
order by 1,2),
cte5 as
(select product,round(sum(sales),0) as sum from cte
group by 1
order by sum desc)
select * from cte3
order by sum desc
```

****

product	month	sum	cumul
Macbook Pro Laptop	12	1093100	8037600
Macbook Pro Laptop	10	892500	6196500
Macbook Pro Laptop	5	790500	3075300
Macbook Pro Laptop	4	771800	2284800
Macbook Pro Laptop	11	748000	6944500
Macbook Pro Laptop	3	644300	1513000
iPhone	12	634200	4794300

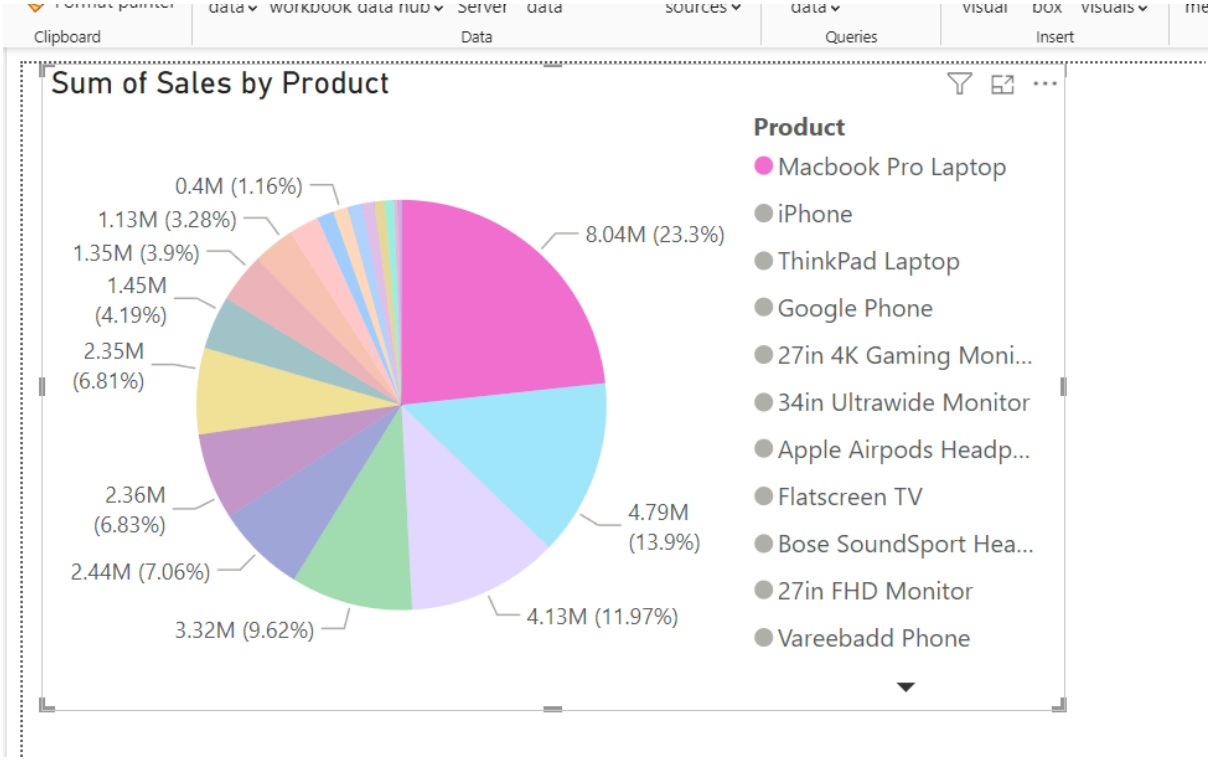
This is the output for question 4

****

Now for step 3 we are going to do some visualisation in power bi

Step 3

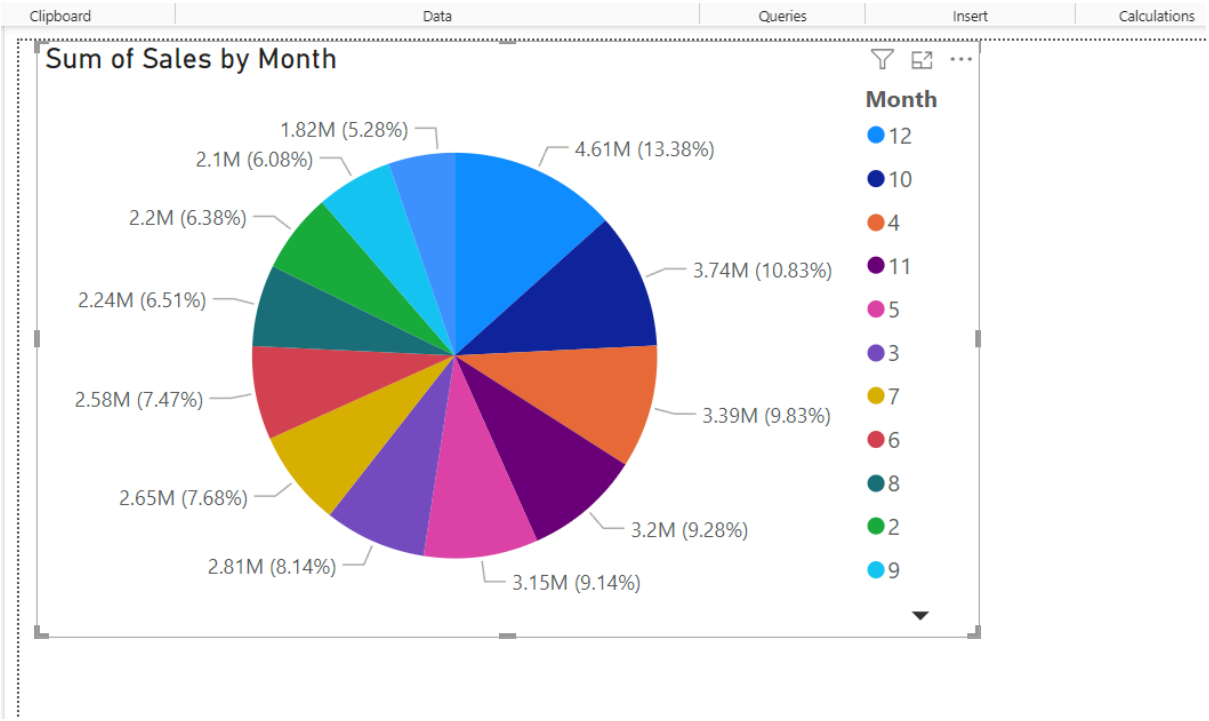
VISUAL 1 -



With this image we can easily figure out total sales per product And their percentage

****

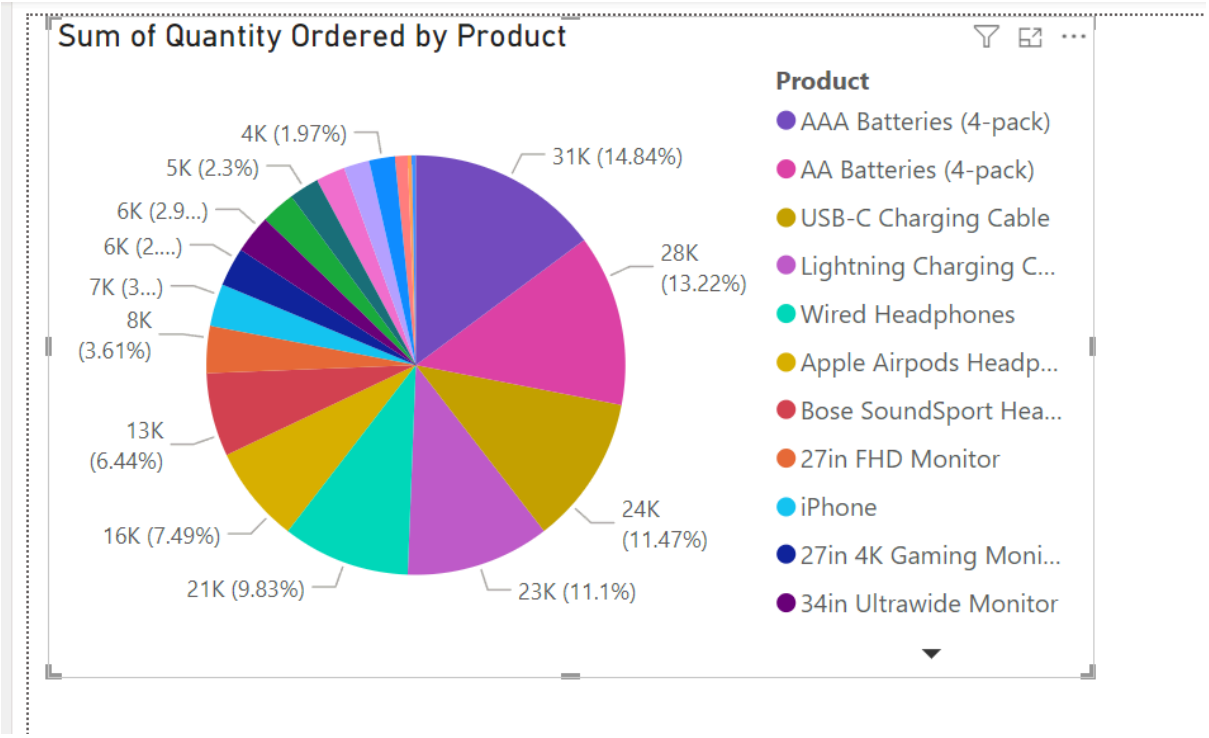
VISUAL 2 -



With this we can see total sales of all products per month

****

VISUAL 3 -

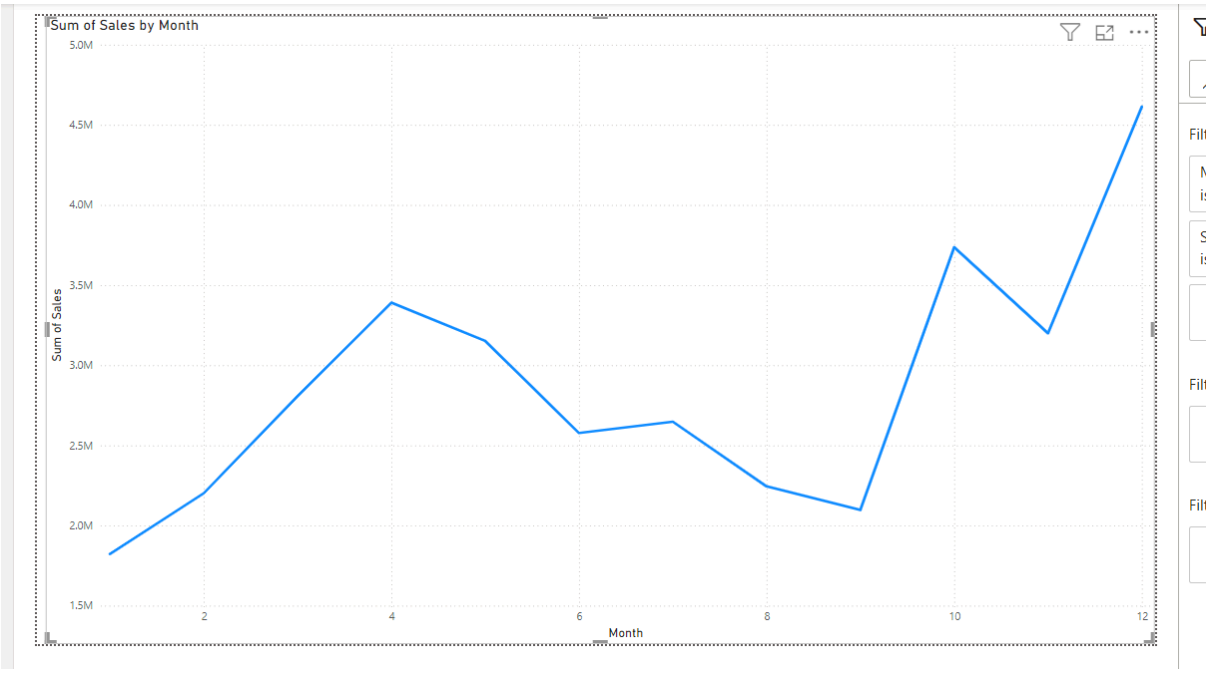


With the above visual we can see which of the product was sold mostly (total quantity per product)

AAA batteries were sold in huge quantity

****

Visual 4 -



With this we can see trends of sales over time
(amount of total sales of all products changed per month)

****