1. select month(Date) as months, sum(Sales_Amount) as 每月銷售總額,
count(distinct Customer_ID) as 每月消費人數,
sum(Quantity) as 每月消費數量,
count(id) as 每月消費次數
from 2020scanner_data
where year(Date) ='2020'
group by month(Date);

months	每月銷售總額	每月消費人數	每月消費數量	每月消費次數
2	122114.6099999993	3527	15253.71	10412
3	127924.54000000008	4145	15686.1	10865
4	138172.11000000095	4161	16633.485999999997	11342
5	142719.85999999967	4418	18295.257	12103
6	131305.34999999977	3720	16120.596000000001	10659
7	120591.96999999981	3353	14296.960000000001	9804
8	116908.98	3303	13901.519999999999	9583
9	141239.8899999999	4089	17399.23	11685
10	140853.51000000013	3980	17801.98	11668
11	132883.32999999975	3901	16683.920000000002	11053
12	152124.1899999999	4186	18677.3	12535

2.select month(Date) as 月 , avg(Sales_Amount) as 平均消費額
from 2020scanner_data
where year(Date) = '2020'
group by month(Date);

月	平均消費額
1	11.123365009502825
2	11.728256819054904
3	11.774002761159702
4	12.182340856991805
5	11.792106089399262
6	12.318730650154802
7	12.30028253773971
8	12.199622247730366
9	12.087281985451407
10	12.071778368186418
11	12.02237673029947
12	12.135954527323351

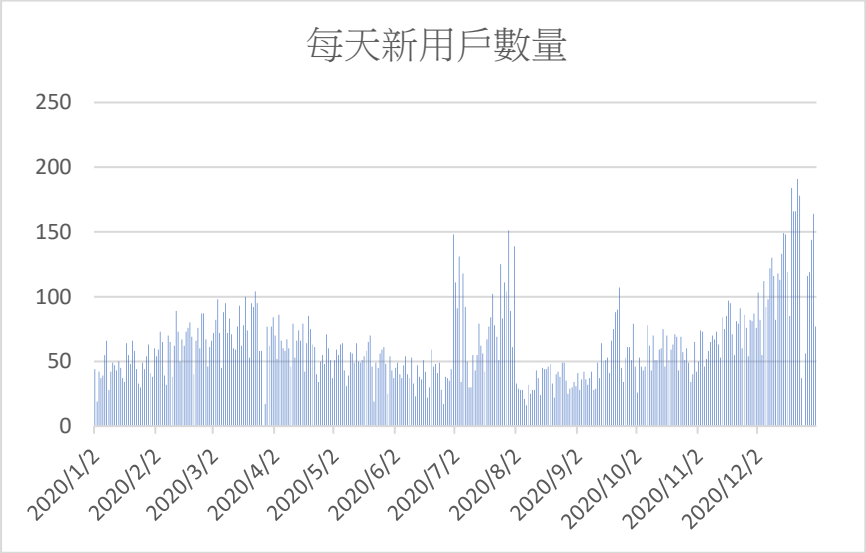
3. select sum(Sales_Amount) as 銷售總額,
count(distinct Customer_ID) as 消費人數,
sum(Quantity) as 消費數量,
count(id) as 消費次數
from 2020scanner_data
where year(Date) ='2020'
group by quarter(Date);

銷售總額	消費人數	消費數量	消費次數
361239.4299999871	8259	45814.18000000001	31274
412197.31999998743	9076	51049.33900000001	34104
378740.83999998716	7775	45597.71000000002	31072
425861.02999997843	8486	53163.2	35256

4. select avg(Sales_Amount) as 平均消費額
from 2020scanner_data
where year(Date) = '2020'
group by quarter(Date);

每季	平均消費額
1	11.550790752701616
2	12.086480178277924
3	12.189136199793634
4	12.079107953255477

5.SELECT DATE(Date) AS date, COUNT(DISTINCT Customer_ID) AS 每天新用戶數量
FROM 2020scanner_data AS t1
WHERE NOT EXISTS (
SELECT 1
FROM 2020scanner_data AS t2
WHERE t2.Customer_ID = t1.Customer_ID AND t2.id > t1.id
)
GROUP BY DATE(`Date`);



6.

```
SELECT MONTH(Date) AS Month, COUNT(DISTINCT Customer_ID) AS 每月新用戶數量
FROM 2020scanner_data AS t1
WHERE NOT EXISTS (
        SELECT 1
        FROM 2020scanner_data AS t2
        WHERE t2.Customer_ID = t1.Customer_ID AND t2.id > t1.id
    )
GROUP BY MONTH(`Date`);
```

Month	每月新用戶數量
1	1364
2	1846
3	2235
4	1865
5	1570
6	1203
7	2479
8	1164
9	1537
10	1706
11	2125
12	3531

7.

```
SELECT COUNT(DISTINCT Customer_ID) / (SELECT COUNT(DISTINCT Customer_ID) FROM 2020scanner_data) AS 僅消費一次客戶占比
FROM (
  SELECT Customer_ID, COUNT(DISTINCT Transaction_ID) AS num_transactions
  FROM 2020scanner_data
  GROUP BY Customer_ID
  HAVING num_transactions = 1
) subquery;
```

僅消費一次客戶占比
0.5098

8.

```
SELECT
  DATE_FORMAT(t1.Date, '%Y-%m') AS Month,
  COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) / COUNT(DISTINCT t1.Customer_ID) AS 當月回購率
FROM 2020scanner_data t1
```

```
JOIN (
    SELECT
        Customer_ID,
        DATE_FORMAT(Date, '%Y-%m') AS Month,
        COUNT(DISTINCT Transaction_ID) AS cnt
    FROM 2020scanner_data
    GROUP BY Customer_ID, DATE_FORMAT(Date, '%Y-%m')
) t2
ON t1.Customer_ID = t2.Customer_ID AND DATE_FORMAT(t1.Date, '%Y-%m') =
t2.Month
GROUP BY DATE_FORMAT(t1.Date, '%Y-%m');

SELECT AVG(Repurchase_Rate) AS 整體平均值
FROM (
    SELECT
        DATE_FORMAT(t1.`Date`, '%Y-%m') AS Month,
        COUNT(DISTINCT t1.Customer_ID) AS Total_Customers,
        COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) AS
Repeat_Customers,
        COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) /
COUNT(DISTINCT t1.Customer_ID) AS Repurchase_Rate
    FROM 2020scanner_data t1
    JOIN (
        SELECT
            Customer_ID,
            DATE_FORMAT(`Date`, '%Y-%m') AS Month,
            COUNT(DISTINCT Transaction_ID) AS cnt
        FROM 2020scanner_data
        GROUP BY Customer_ID, DATE_FORMAT(`Date`, '%Y-%m')
    ) t2
    ON t1.Customer_ID = t2.Customer_ID AND DATE_FORMAT(t1.Date, '%Y-%m') =
t2.Month
    GROUP BY DATE_FORMAT(t1.Date, '%Y-%m')
) t;
```

Month	當月回購率
2020-01	0.2691
2020-02	0.2628
2020-03	0.2268
2020-04	0.2410
2020-05	0.2456
2020-06	0.2586
2020-07	0.2469
2020-08	0.2637
2020-09	0.2570
2020-10	0.2611
2020-11	0.2640
2020-12	0.2797

整體平均值
0.25635833

9.

```
SELECT
    DATE_FORMAT(t1.Date, '%Y-%m') AS Month,
    COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) /
COUNT(DISTINCT t1.Customer_ID) AS 當月回購率
FROM 2020scanner_data t1
JOIN (
    SELECT
        Customer_ID,
        DATE_FORMAT(Date, '%Y-%m') AS Month,
        COUNT(DISTINCT Transaction_ID) AS cnt
    FROM 2020scanner_data
    GROUP BY Customer_ID, DATE_FORMAT(Date, '%Y-%m')
) t2
ON t1.Customer_ID = t2.Customer_ID AND DATE_FORMAT(t1.Date, '%Y-%m') =
t2.Month
GROUP BY DATE_FORMAT(t1.Date, '%Y-%m') limit 11;

SELECT AVG(Repurchase_Rate) AS 整體平均
FROM (
    SELECT
        DATE_FORMAT(t1.`Date`, '%Y-%m') AS Month,
        COUNT(DISTINCT t1.Customer_ID) AS Total_Customers,
        COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) AS
Repeat_Customers,
        COUNT(DISTINCT CASE WHEN t2.cnt > 1 THEN t2.Customer_ID END) /
COUNT(DISTINCT t1.Customer_ID) AS Repurchase_Rate
```

```
    FROM 2020scanner_data t1
    JOIN (
        SELECT
            Customer_ID,
            DATE_FORMAT(`Date`, '%Y-%m') AS Month,
            COUNT(DISTINCT Transaction_ID) AS cnt
        FROM 2020scanner_data
        GROUP BY Customer_ID, DATE_FORMAT(`Date`, '%Y-%m')
    ) t2
    ON t1.Customer_ID = t2.Customer_ID AND DATE_FORMAT(t1.Date, '%Y-%m') =
t2.Month
    GROUP BY DATE_FORMAT(t1.Date, '%Y-%m')
    limit 11
) t;
```

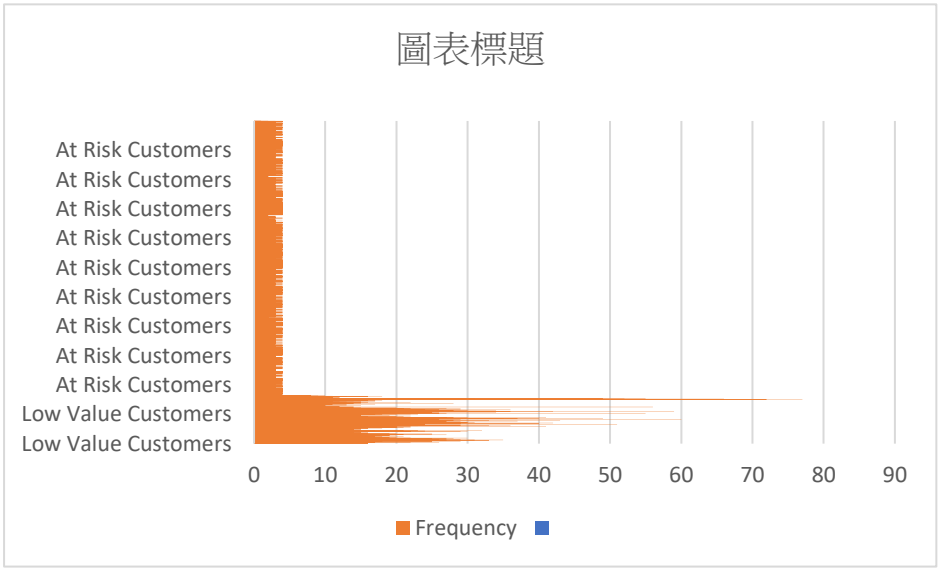
Month	當月回購率
2020-01	0.2691
2020-02	0.2628
2020-03	0.2268
2020-04	0.2410
2020-05	0.2456
2020-06	0.2586
2020-07	0.2469
2020-08	0.2637
2020-09	0.2570
2020-10	0.2611
2020-11	0.2640

整體平均
0.25423636

10.

```
SELECT
    Customer_ID,
    MAX(Date) AS Recency,
    COUNT(DISTINCT Date) AS Frequency,
    SUM(Sales_Amount) AS Monetary,
    CASE
        WHEN MAX(Date) >= DATE_SUB(NOW(), INTERVAL 30 DAY) AND
COUNT(DISTINCT Date) >= 10 AND SUM(Sales_Amount) >= 10000 THEN 'High Value
Customers'
        WHEN MAX(Date) >= DATE_SUB(NOW(), INTERVAL 60 DAY) AND
COUNT(DISTINCT Date) >= 5 AND SUM(Sales_Amount) >= 5000 THEN 'Recent Active
Customers'
```

```
        WHEN COUNT(DISTINCT Date) >= 10 AND SUM(Sales_Amount) >= 10000 THEN
'Frequent Customers'
        WHEN SUM(Sales_Amount) >= 10000 THEN 'Big Spenders'
        WHEN MAX(Date) >= DATE_SUB(NOW(), INTERVAL 30 DAY) THEN 'New
Customers'
        WHEN MAX(Date) < DATE_SUB(NOW(), INTERVAL 90 DAY) AND
COUNT(DISTINCT Date) < 5 AND SUM(Sales_Amount) < 5000 THEN 'At Risk
Customers'
        WHEN COUNT(DISTINCT Date) < 5 AND SUM(Sales_Amount) < 5000 THEN
'Infrequent Customers'
        ELSE 'Low Value Customers'
    END AS RFM_Category
FROM 2020scanner_data
GROUP BY Customer_ID
ORDER BY RFM_Category DESC;
```



https://drive.google.com/file/d/18VlTjjNdqYekrRft9Cfrv-znm1qeebl0/view?usp=sharing

11.

```
SELECT
        user_type,
        COUNT(DISTINCT Customer_ID) AS user_count
FROM (
        SELECT
                Customer_ID,
                CASE
```

```
            WHEN DATEDIFF(MAX(Date), min(Date)) > 180 AND COUNT(DISTINCT
DATE(Date)) > 1 THEN '忠誠用戶'
            WHEN DATEDIFF(MAX(Date), min(Date)) > 180 AND COUNT(DISTINCT
DATE(Date)) = 1 THEN '一次性用戶'
            WHEN DATEDIFF(MAX(Date), min(Date)) <= 180 AND COUNT(DISTINCT
DATE(Date)) > 1 THEN '新用戶'
            WHEN DATEDIFF(MAX(Date), min(Date)) <= 180 AND COUNT(DISTINCT
DATE(Date)) = 1 THEN '流失老用戶'
        END AS user_type
    FROM
        2020scanner_data
    GROUP BY
        Customer_ID
) AS user_type_table
GROUP BY
    user_type;
```

user_type	user_count
忠誠用戶	3082
新用戶	7684
流失老用戶	11859

12.

```
select SKU as 熱銷商品前十名 , SUM(Quantity) as 賣出數量
from 2020scanner_data
group by SKU
order by 賣出數量 DESC limit 10;
```

熱銷商品前十名	賣出數量
CKDW0	5769.2
TD3DD	3786
UNJKW	2179
EEI1Q	1568
MXKDP	1252
H8P2L	1118
2SVKS	1113
CYRX4	1071
W1ZMG	1033
C6TXL	1027

13.

```
select SKU as 商品購買人數前十名 , count(distinct Customer_ID) as 商品購買人
```

數
from 2020scanner_data
group by SKU
order by 商品購買人數 DESC limit 10;

商品購買人數前十名	商品購買人數
UNJKW	1648
COWU2	681
OV1P9	613
M6J9W	584
CZUZX	550
E3PAN	478
A59HR	473
C6TXL	471
7GQRJ	437
W1ZMG	426

14.

```
select
    month(`Date`) as 月份,
    s.SKU,
    sum(s.Quantity) as 銷售量,
    count(distinct s.Customer_ID) AS 購買人數
from 2020scanner_data s
inner join(
    select SKU
    from 2020scanner_data
    group by SKU
    order by sum(Quantity) DESC ,count(distinct Customer_ID) DESC
    limit 10
    )p
    on s.SKU = p.SKU
    group by 月份, s.SKU
    order by 銷售量 DESC, 購買人數 DESC
    limit 10;
```

月份	SKU	銷售量	購買人數
12	CKDW0	970	10
9	CKDW0	661	6
1	CKDW0	620	7
10	CKDW0	610.2	7
2	CKDW0	580	8
5	CKDW0	517	9
6	TD3DD	517	8
11	CKDW0	510	4
8	TD3DD	400	10
5	TD3DD	391	9

15.

select SKU_Category as 商品品類銷售數量前十名 , count(distinct Quantity) as 商品購買人數

from 2020scanner_data

group by SKU_Category

order by 商品購買人數 DESC limit 10;

商品品類銷售數量前十名	商品購買人數
JI5	42
XG4	30
C3Y	22
H8O	18
LPF	15
JPI	15
HXR	14
W41	12
U5F	12
SJS	12

16.

select SKU_Category as 商品品類購買人數前十名 , count(distinct Customer_ID) as 商品購買人數

from 2020scanner_data

group by SKU_Category

order by 商品購買人數 DESC limit 10;

商品品類購買人數前十名	商品購買人數
N8U	6201
R6E	3400
P42	3248
0H2	3176
LPF	2734
IEV	2416
H15	2299
29A	2209
FEW	2157
U5F	2014

17.

```
select
    month(`Date`) as 月份,
    s.SKU_Category,
    sum(s.Quantity) as 銷售量,
    count(distinct s.Customer_ID) AS 購買人數
from 2020scanner_data s
inner join(
    select SKU_Category
    from 2020scanner_data
    group by SKU_Category
    order by sum(Quantity) DESC ,count(distinct Customer_ID) DESC
    limit 10
    )p
    on s.SKU_Category = p.SKU_Category
    group by 月份, s.SKU_Category
    order by 銷售量 DESC, 購買人數 DESC
    limit 10;
```

月份	SKU_Category	銷售量	購買人數
6	JI5	1626	31
10	JI5	1443	41
12	JI5	1393	35
5	JI5	1380	35
12	N8U	1223	883
1	JI5	1184	23
9	N8U	1126	841
10	N8U	1103	814
11	JI5	1057	27
3	N8U	1039	801

18:
根據以上分析，我們可以得出結論
月度銷售表現穩定，季度銷售額和消費人數基本相近，但季度平均消費金額略有波動。
老用戶流失佔用戶總數的比例較高，一次性用戶的比例不容忽視。
最暢銷和最暢銷的商品並不完全重疊，但有些商品在兩個圖表中都表現良好。
各品類的銷售額和買家數量排名並不完全相同，但部分品類在兩個排名中均進入前十。
從以上分析可以看出，我們的客戶群基本穩定，但還需要提高客戶忠誠度，促進復購