

# Consumer Goods Ad\_Hoc Insights



**BY MANISH KUMAR CHAUDHARY**

1. Provide the list of markets in which customer "Atliq Exclusive" operates its business in the APAC region.

```
SELECT market
FROM dim_customer
WHERE customer = "Atliq Exclusive"
AND region = "APAC"
GROUP BY market;
```

**QUERY EXPLANATION:**

1. We are using WHERE condition along with AND condition to get records of Atliq Exclusive which are in APAC region only.
2. GROUP BY is used so that we can get distinct markets.  
**DISTINCT** is slower as compare to GROUP BY so I have used GROUP BY.

**OUTPUT:**

market
India
Indonesia
Japan
Philippines
South Korea
Australia
Newzealand
Bangladesh

**INSIGHTS:**

1. Atliq Exclusive deals in 8 countries in APAC region.

2. What is the percentage of unique product increase in 2021 vs. 2020? The final output contains these fields,  
 unique\_products\_2020  
 unique\_products\_2021  
 percentage\_chg

```

WITH unique_product_cte AS (
  SELECT (SELECT COUNT(DISTINCT product_code)
          FROM fact_sales_monthly
         WHERE fiscal_year = "2020")
        unique_product_2020
     ,(SELECT COUNT(DISTINCT product_code)
          FROM fact_sales_monthly
         WHERE fiscal_year = "2021")
        unique_product_2021
   )
  SELECT unique_product_2020
        ,unique_product_2021
        , ROUND((unique_product_2021-
unique_product_2020)*100/unique_product_2020,1)
  percentage_chg
 FROM unique_product_cte

```

#### QUERY EXPLANATION:

1. In CTE we are simply finding the count of unique products in year 2020 and 2021 by using SELECT statement in a subquery.  
Here, subquery was used because both year 2020 and 2021 is on same column.
2. After this we are now simply using SELECT statement to get the desired output.

#### OUTPUT:

unique_product_2020	unique_product_2021	percentage_chg
245	334	36.3

#### INSIGHTS:

1. There is change of approximately 36 % in unique products of year 2020 and 2021. Here, unique products in 2021 was increased by approximately 36%.
3. Provide a report with all the unique product counts for each segment and sort them in descending order of

product counts. The final output contains 2 fields,

segment

product\_count



```
SELECT segment
      ,COUNT(DISTINCT product_code) product_count
  FROM dim_product
 GROUP BY segment
 ORDER BY product_count DESC;
```

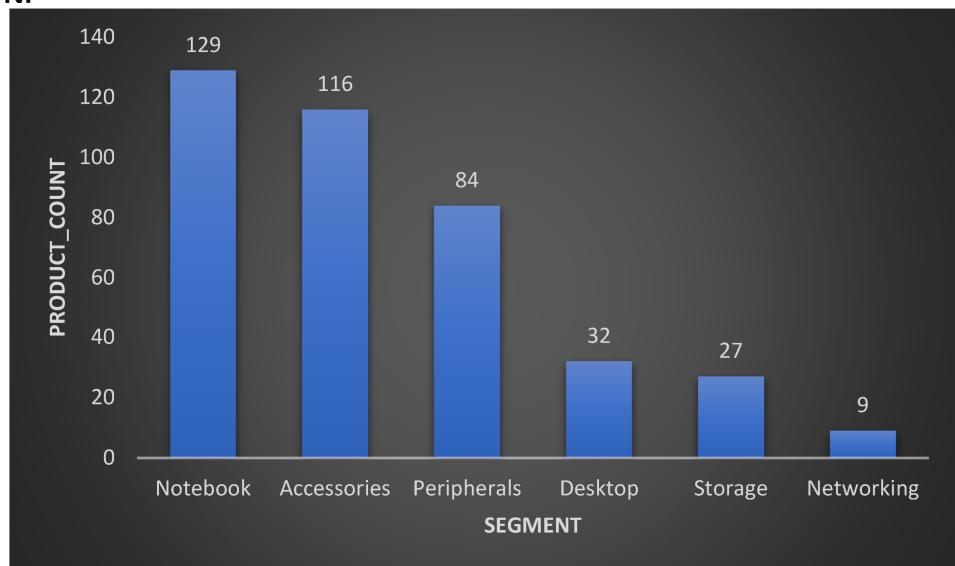
#### QUERY EXPLANATION:

1. We simply counting the distinct products for every segment that is present.
2. At last ordering them on the basis of descending order of count of product.

OUTPUT:

segment	product_count
Notebook	129
Accessories	116
Peripherals	84
Desktop	32
Storage	27
Networking	9

#### VISUALISATION:



## INSIGHTS:

1. Notebook segment is having the highest product count and is approximately 32% of overall unique products in all segments.
2. From the above visual we can see that for the Networking segment, there are only a few unique products.
3. Notebook, Accessories, and peripherals account for almost 82 % of overall unique products in all the segments.
  
4. Follow-up: Which segment had the most increase in unique products in 2021 vs 2020? The final output contains these fields,

segment

product\_count\_2020

product\_count\_2021

difference

```
WITH product_count_2020_cte AS (
    SELECT p.segment
        ,COUNT(DISTINCT m.product_code)
    product_count_2020
    FROM fact_sales_monthly m
    LEFT JOIN dim_product p
    ON m.product_code = p.product_code
    WHERE m.fiscal_year = "2020"
    GROUP BY segment
),
product_count_2021_cte AS (
    SELECT p.segment
        ,COUNT(DISTINCT m.product_code)
    product_count_2021
    FROM fact_sales_monthly m
    LEFT JOIN dim_product p
    ON m.product_code = p.product_code
    WHERE m.fiscal_year = "2021"
    GROUP BY segment
)
SELECT c20.segment
    ,c20.product_count_2020
    ,c21.product_count_2021
    ,MAX(ABS(c21.product_count_2021-
c20.product_count_2020)) difference
FROM product_count_2020_cte c20
INNER JOIN product_count_2021_cte c21
ON c20.segment = c21.segment
```

## QUERY EXPLANATION:

1. Inside CTE we are counting product for each segment in year 2020 and year 2021.
2. After this we are joining two CTE and finding the required output.  
Here, we have

#### OUTPUT:

segment	product_count_2020	product_count_2021	difference
Accessories	69	103	34

#### INSIGHTS:

1. From the output we can see that for Accessories segment there is maximum change in unique product count in year 2021 as compare to year 2020.
5. Get the products that have the highest and lowest manufacturing costs. The final output should contain these fields,  
product\_code  
product  
manufacturing\_cost

```

SELECT m.product_code
      ,p.product
      ,MAX(m.manufacturing_cost) manufacturing_cost
FROM dim_product p
INNER JOIN fact_manufacturing_cost m
ON p.product_code = m.product_code
UNION ALL
SELECT m.product_code
      ,p.product
      ,MIN(m.manufacturing_cost) manufacturing_cost
FROM dim_product p
INNER JOIN fact_manufacturing_cost m
ON p.product_code = m.product_code

```

#### QUERY EXPLANATION:

1. I have used separate queries for finding the MAX and MIN of manufacturing cost.
2. To get them in one column I have used UNION ALL to combine the result of both query in single column.

**OUTPUT:**

product_code	product	manufacturing_cost
A0118150101	AQ Dracula HDD – 3.5 Inch SATA 6 Gb/s 5400 RPM 256 MB Cache	240.5364
A0118150101	AQ Dracula HDD – 3.5 Inch SATA 6 Gb/s 5400 RPM 256 MB Cache	0.892

**INSIGHTS:**

1. We can see that for the same product we are having maximum and minimum manufacturing costs which means they are of a different variant.
6. Generate a report which contains the top 5 customers who received an average high pre\_invoice\_discount\_pct for the fiscal year 2021 and in the Indian market. The final output contains these fields,

customer\_code

customer

average\_discount\_percentage

```
SELECT d.customer_code
      ,c.customer
      ,AVG(d.pre_invoice_discount_pct) average_discount_percentage
  FROM dim_customer c
 INNER JOIN fact_pre_invoice_deductions d
    ON c.customer_code = d.customer_code
 WHERE d.fiscal_year = "2021"
   AND c.market = "India"
 GROUP BY d.customer_code, c.customer
 ORDER BY average_discount_percentage DESC
 LIMIT 5
```

**QUERY EXPLANATION:**

1. We are using AVG () function to calculate the average discount percentage and order them in descending order.
2. After that we are using LIMIT 5 so that we get the top 5 highest average discount percentages and corresponding records for them.

Here, for the general approach we could have used the WINDOW function like ROW\_NUMBER() or RANK() to get the top 5 records.

**OUTPUT:**

customer_code	customer	average_discount_percentage
90002009	Flipkart	0.3083
90002006	Viveks	0.3038
90002003	Ezone	0.3028
90002002	Croma	0.3025
90002016	Amazon	0.2933

7. Get the complete report of the Gross sales amount for the customer “Atliq Exclusive” for each month. This analysis helps to get an idea of low and high-performing months and take strategic decisions. The final report contains these columns:

Month

Year

Gross sales Amount

```
SELECT MONTHNAME(s.date) month
      ,YEAR(s.date) year
      ,SUM(s.sold_quantity*p.gross_price) gross_sales_amount
  FROM fact_sales_monthly s
  LEFT JOIN fact_gross_price p
  ON s.product_code = p.product_code
  LEFT JOIN dim_customer c
  ON s.customer_code = c.customer_code
 WHERE c.customer = "Atliq Exclusive"
 GROUP BY month, year
 ORDER BY year, MONTH(date)
```

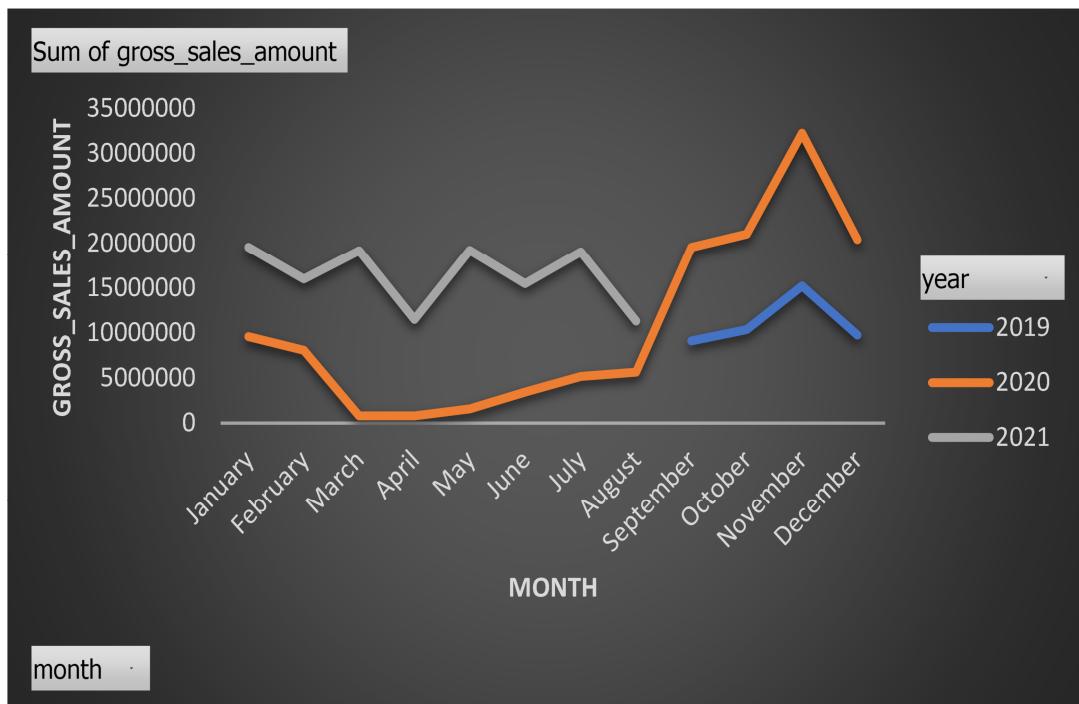
**QUERY EXPLANATION:**

1. As we want month and year-wise gross sales amount so we are using MONTHNAME() and YEAR() to extract the month name and year from the date column.
2. We want the total gross sales amount thus we are using SUM().
3. We are using GROUP BY to get records month and year-wise separated.
4. At last order them on the basis of year first and then on the basis of MONTH(date) so that we get MONTHNAME ordered in actual month-wise that is starting from January to December.

**OUTPUT:**

month	year	gross_sales_amount
September	2019	9092670.339
October	2019	10378637.6
November	2019	15231894.97
December	2019	9755795.058
January	2020	9584951.939
February	2020	8083995.548
March	2020	766976.4531
April	2020	800071.9543
May	2020	1586964.477
June	2020	3429736.571
July	2020	5151815.402
August	2020	5638281.829
September	2020	19530271.3
October	2020	21016218.21
November	2020	32247289.79
December	2020	20409063.18
January	2021	19570701.71
February	2021	15986603.89
March	2021	19149624.92
April	2021	11483530.3
May	2021	19204309.41
June	2021	15457579.66
July	2021	19044968.82
August	2021	11324548.34

## VISUALISATION:



## INSIGHTS:

1. From the line chart we can see that in November 2020 there is the highest gross\_sales\_amount.
  2. We can also see that for the year 2020 gross\_sales amount is increasing gradually and after November it started decreasing.
  3. For 2019 and 2021 it is fluctuating up and down.
8. In which quarter of 2020, got the maximum total\_sold\_quantity? The final output contains these fields sorted by the total\_sold\_quantity,

Quarter

```
SELECT QUARTER(date) quarter
      ,SUM(sold_quantity) total_sold_quantity
FROM fact_sales_monthly
WHERE fiscal_year = "2020"
GROUP BY quarter
ORDER BY total_sold_quantity DESC
LIMIT 1
```

### QUERY EXPLANATION:

1. In this we are using QUARTER to extract quarter out of date column.
2. We want only for year 2020 so we are using where condition to get data of 2020 only.
3. As we want quarter with maximum total sold quantity we are using LIMIT 1.

### OUTPUT:

quarter	total_sold_quantity
4	8425822

9. Which channel helped to bring more gross sales in the fiscal year 2021 and the percentage of contribution?

The final output contains these fields,

channel

gross\_sales\_mln

Percentage

```
● ● ●

WITH sales_cte AS(
    SELECT c.channel channel
        ,SUM(s.sold_quantity*p.gross_price) gross_sales_mln
        ,RANK() OVER(ORDER BY SUM(s.sold_quantity*p.gross_price) DESC) rnk
    FROM fact_sales_monthly s
    LEFT JOIN fact_gross_price p
    ON s.product_code = p.product_code
    LEFT JOIN dim_customer c
    ON s.customer_code = c.customer_code
    WHERE s.fiscal_year = "2021"
    GROUP BY c.channel
)
SELECT channel
    ,gross_sales_mln
    ,ROUND(gross_sales_mln*100/(SELECT SUM(gross_sales_mln) FROM sales_cte),1) percentage
FROM sales_cte
WHERE rnk = 1
GROUP BY gross_sales_mln
```

## QUERY EXPLANATION:

1. Here, in CTE we are querying for channel, gross sales by particular channel and then ranking them on the basis of gross sales.
2. Within CTE we are using the WHERE condition so that we get details of year 2021 only.
3. In SELECT statement after CTE we are querying for channel, gross sales and percentage of gross sales by particular channel.
4. As we want which one has highest sales we are using WHERE condition for rank =1 so that we get highest gross sales channel.

OUTPUT:

channel	gross_sales_mln	percentage
Retailer	1.92E+09	73.2

## INSIGHT:

1. From the output we can see Retailer channel is having highest gross sales.
2. The gross sales by Retailer accounts to 73.2 % of total gross sales by all channels.

10. Get the Top 3 products in each division that have a high total\_sold\_quantity in the fiscal\_year 2021?  
The final output contains these

fields,  
division  
product\_code  
product  
total\_sold\_quantity  
rank\_order

```
● ● ●

WITH rank_cte AS (
  SELECT p.division
    ,p.product_code
    ,p.product
    ,SUM(s.sold_quantity) total_sold_quantity
    ,ROW_NUMBER() OVER (PARTITION BY p.division ORDER BY SUM(s.sold_quantity) DESC) rank_order
  FROM fact_sales_monthly s
  LEFT JOIN dim_product p
  ON s.product_code = p.product_code
  WHERE s.fiscal_year = "2021"
  GROUP BY p.division, p.product_code, p.product
)
SELECT division
    ,product_code
    ,product
    ,total_sold_quantity
    ,rank_order
  FROM rank_cte
 WHERE rank_order <= 3;
```

**QUERY EXPLANATION:**

1. In CTE we are querying mainly total sold quantity and giving them row\_number/rank on the basis of sold quantity.  
Here, we are giving rank on the basis of total sold quantity in descending order for each division.
2. We are using WHERE condition so that we can get records of year 2021 only.
3. Now in SELECT statement we are querying for all required columns.
4. And giving condition for rank so that we can get top 3 highest selling products for each division.

**OUTPUT:**

division	product_code	product	total_sold_quantity	rank_order
N & S	A6720160103	AQ Pen Drive 2 IN 1	701373	1
N & S	A6818160202	AQ Pen Drive DRC	688003	2
N & S	A6819160203	AQ Pen Drive DRC	676245	3
P & A	A2319150302	AQ Gamers Ms	428498	1
P & A	A2520150501	AQ Maxima Ms	419865	2
P & A	A2520150504	AQ Maxima Ms	419471	3
PC	A4218110202	AQ Digit	17434	1
PC	A4319110306	AQ Velocity	17280	2
PC	A4218110208	AQ Digit	17275	3

**INSIGHTS:**

1. For N&S division product AQ Pen Drive 2 IN 1 is having highest total sold quantity.
2. For P&A division it is product AQ Gamers Ms.
3. For PC division it is product AQ Digit.
4. From the output we can also see that N&S division is having highest total sold quantity as compared to other divisions.