Consumer Goods : Ad_Hoc Insights – SQL Project

- *Domain: Consumer Goods*

- *Sub-Domain: Electronics and Durables*

- *Function: Executive Management*

Presented by : Vinayak Hiremath

Consumer Goods : Ad_Hoc Insights – SQL Project

Contents

- *What is Consumer Goods Domain*

- *About Company and Problem Statement*

- *Ad Hoc Requests from the Management*

- *Data Modelling – ERD Diagram*

- *SQL Query and Result for Each Ad Hoc request*

- *Key Takeaways*

Consumer Goods : Ad_Hoc Insights – SQL Project

What is Consumer Goods Domain ?

The consumer goods domain is a sector of the economy that includes companies that make and sell products for individual and household use. Consumer goods are also known as final goods or retail goods

About Company and Problem Statement

- **Domain:** Consumer Goods | *Function: Executive Management*

- *AtliQ Hardware (imaginary company) is one of the leading computer hardware producers in India and well expanded in other countries too.*

- *The management noticed the need to get insights to make quick and smart data-informed decisions.*

- *Then the management has given the 10 ad hoc requests to the Data analytics team.*

Consumer Goods : Ad_Hoc Insights – SQL Project

Ad_Hoc Requests

Codebasics SQL Challenge

Requests:

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

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

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

5. Get the products that have the highest and lowest manufacturing costs. The final output should contain these fields,
   product_code
   product
   manufacturing_cost

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

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:
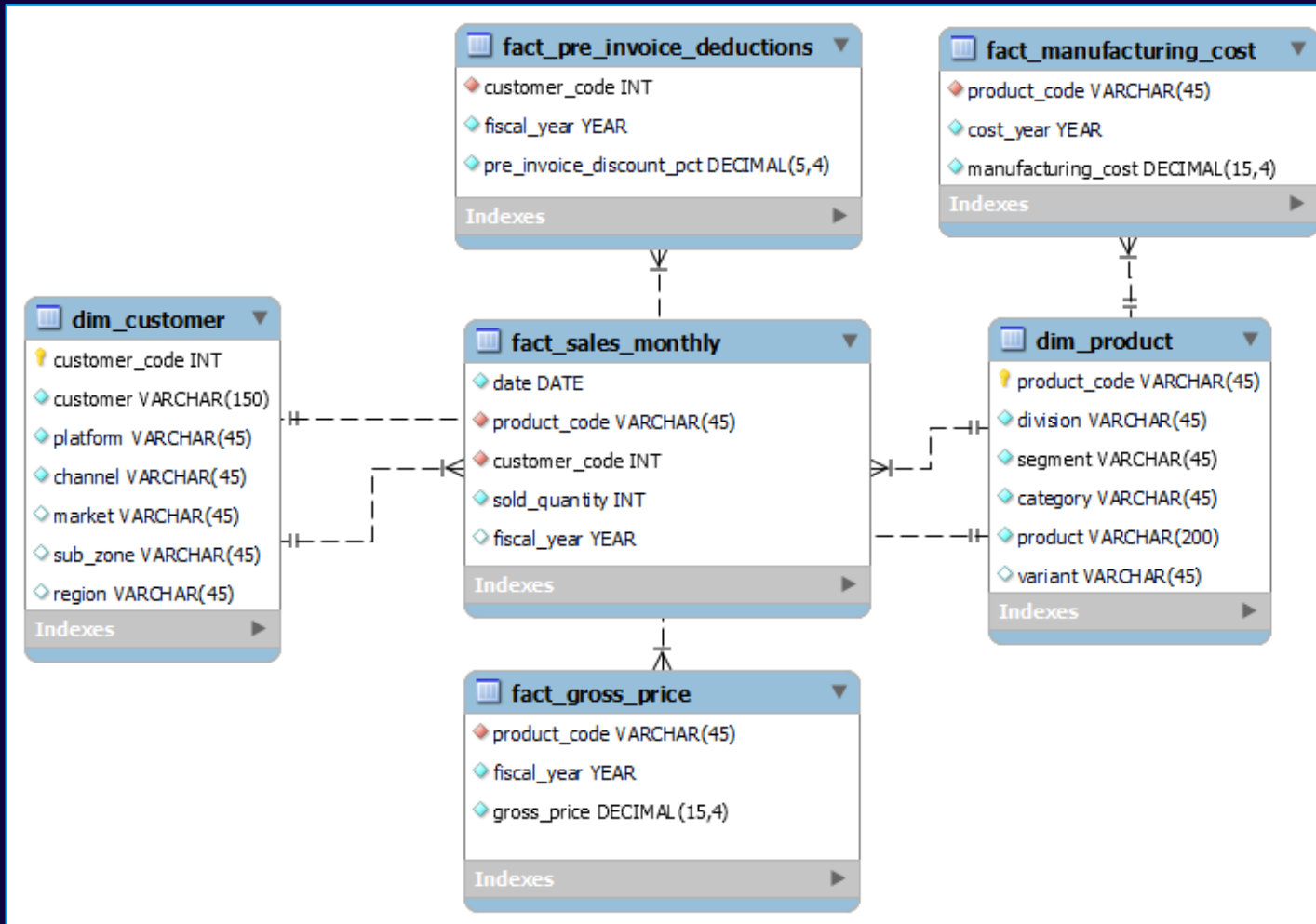   Month
   Year
   Gross sales Amount

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
   total_sold_quantity

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

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

Consumer Goods : Ad_Hoc Insights – SQL Project

❖ Data Modelling – ERD Diagram

fact_pre_invoice_deductions
- customer_code INT
- fiscal_year YEAR
- pre_invoice_discount_pct DECIMAL(5,4)
- Indexes

fact_manufacturing_cost
- product_code VARCHAR(45)
- cost_year YEAR
- manufacturing_cost DECIMAL(15,4)
- Indexes

dim_customer
- customer_code INT
- customer VARCHAR(150)
- platform VARCHAR(45)
- channel VARCHAR(45)
- market VARCHAR(45)
- sub_zone VARCHAR(45)
- region VARCHAR(45)
- Indexes

fact_sales_monthly
- date DATE
- product_code VARCHAR(45)
- customer_code INT
- sold_quantity INT
- fiscal_year YEAR
- Indexes

dim_product
- product_code VARCHAR(45)
- division VARCHAR(45)
- segment VARCHAR(45)
- category VARCHAR(45)
- product VARCHAR(200)
- variant VARCHAR(45)
- Indexes

fact_gross_price
- product_code VARCHAR(45)
- fiscal_year YEAR
- gross_price DECIMAL(15,4)
- Indexes

❑ Steps:

▪ Used Existing Database and Different tables.

▪ Applied Reverse Engineering in MySQL Database to get the ERD diagram.

▪ Defined the Proper Primary and Foreign Keys for Dimension and Fact tables.

▪ Finally the Data Model is Ready.

Q1. Provide the list of markets in which customer "AtliQ Exclusive" operates its business in the APAC region.

SQL Query

```
------------------------------------------------------------

SELECT market as List_of_Markets FROM dim_customer
WHERE customer='Atliq Exclusive' AND region = 'APAC';
```

Result

List_of_Markets
India
Indonesia
Japan
Philiphines
South Korea
Australia
Newzealand
Bangladesh
India

Q2. What is the percentage of unique product increase in 2021 vs. 2020?

```
----------------------------------------------------------

 WITH Total_Products AS
(select
 CASE when f.fiscal_year = 2020 then  p.product END as products_2020,
 CASE when f.fiscal_year = 2021 then  p.product END as products_2021
from dim_product p JOIN fact_sales_monthly f
ON p.product_code = f.product_code)

,Unique_products as
(SELECT count(distinct products_2020) as unique_products_2020,
 count(distinct products_2021) as unique_products_2021
FROM Total_Products)


SELECT unique_products_2020,unique_products_2021,
ROUND((unique_products_2021-unique_products_2020)*100/unique_products_2020,2)
as percentage_chg
FROM Unique_products;
```

SQL Query

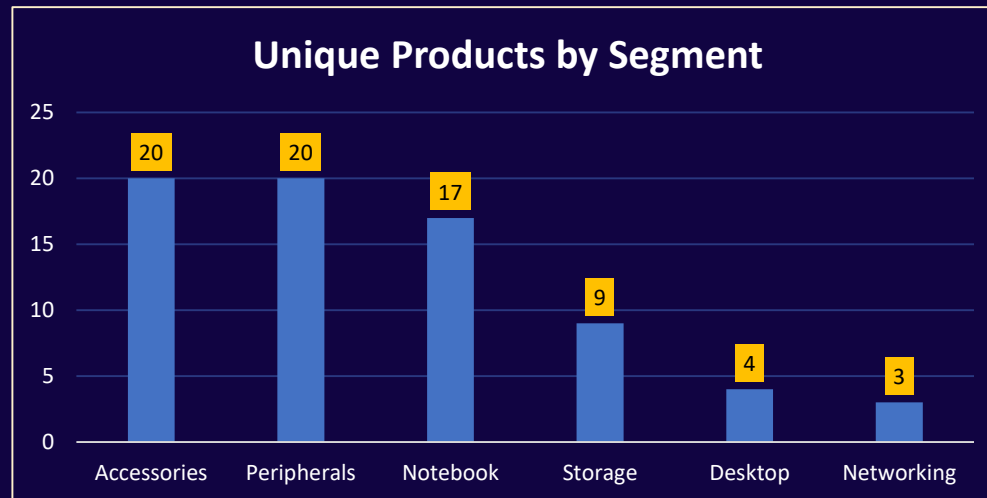| unique_products_2020	unique_products_2021	percentage_chg
51	68	33%

Result

Unique_Products 2020 Vs 2021
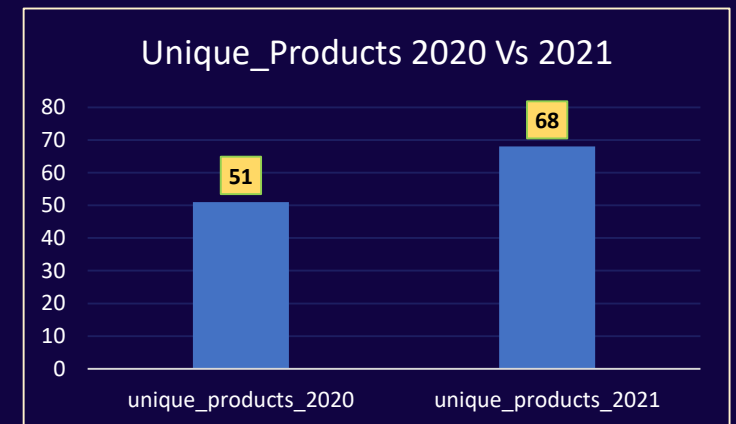
Q3. Provide a report with all the unique product counts for each segment and sort them in descending order of product counts.

SQL Query

```
-----------------------------------------------------------

SELECT segment,count(distinct product) as product_count from dim_product
GROUP BY segment
ORDER BY product_count desc;
```

Result

Unique Products by Segment

Q4. Follow-up: Which segment had the most increase in unique products in 2021 vs 2020?

SQL Query

```
--------------------------------------------------------

WITH cte as
(SELECT segment,
CASE when f.fiscal_year = 2020 then p.product END as products_2020,
CASE when f.fiscal_year = 2021 then p.product END as products_2021
from dim_product p JOIN fact_sales_monthly f
ON p.product_code = f.product_code)

,cte2 as
(SELECT segment, count(distinct products_2020) as unique_products_2020,
count(distinct products_2021) as unique_products_2021 from cte
GROUP BY segment)
SELECT segment,unique_products_2020,unique_products_2021,
(unique_products_2021-unique_products_2020) as diffrence
FROM cte2;
```

Result

Segment Wise Unique Products
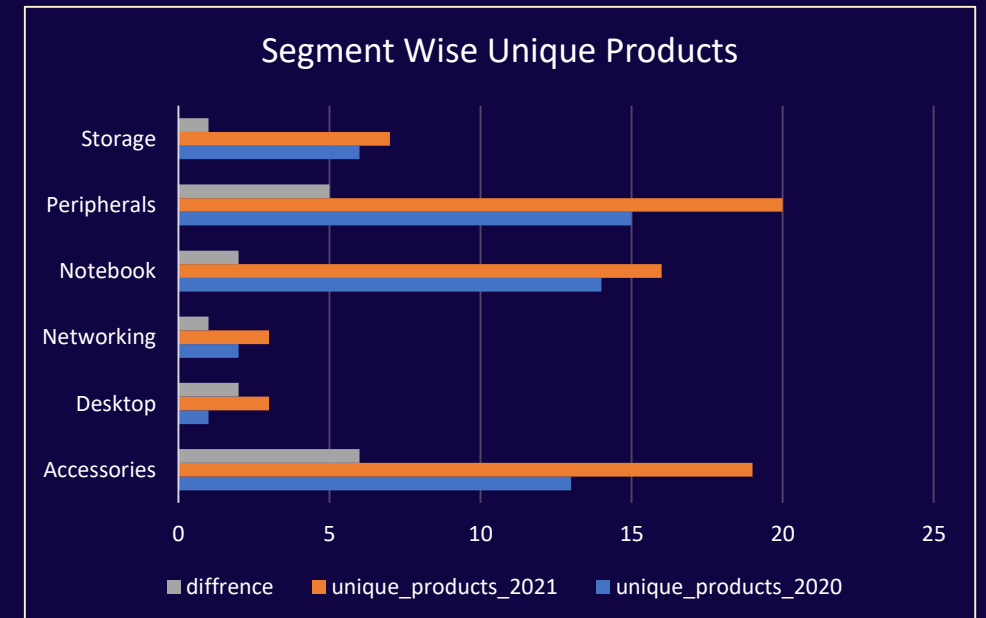
Q5. Get the products that have the highest and lowest manufacturing costs.

SQL Query

```
------------------------------------------------------------

SELECT p.product_code,p.product,min(f.manufacturing_cost) as lowest_mcost,
max(f.manufacturing_cost) as highest_mcost
FROM
dim_product p JOIN fact_manufacturing_cost f
ON p.product_code = f.product_code
GROUP BY p.product_code,p.product
ORDER BY highest_mcost desc;
```

Result

| product_code	cost_year	manufacturing_cost
A0118150101	2020	5.0207
A0118150101	2021	5.5172
A0118150102	2020	5.718
A0118150102	2021	6.2835
A0118150103	2020	6.3264
A0118150103	2021	6.59
A0118150104	2020	6.4789
A0118150104	2021	6.8199
A0219150201	2020	6.4858
A0219150201	2021	7.0498
A0219150202	2020	7.059
A0219150202	2021	7.2031
A0220150203	2020	7.0621
A0220150203	2021	7.3563
A0320150301	2020	6.8414
A0320150301	2021	7.3563
A0321150302	2021	7.8161

Q6. 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.

SQL Query

```
-----------------------------------------------------

SELECT c.customer_code,c.customer,
ROUND(avg(f.pre_invoice_discount_pct)*100,2) as average_discount_percentage
FROM dim_customer c
JOIN
fact_pre_invoice_deductions f
ON c.customer_code = f.customer_code
WHERE f.fiscal_year = 2021 AND market = 'india'
GROUP BY c.customer_code,c.customer
ORDER BY average_discount_percentage desc
LIMIT 5;
```

Result

Average Discount Percentage (%)

- Flipkart
- Viveks
- Ezone
- Croma
- Amazon

20% 20% 20% 20% 20%
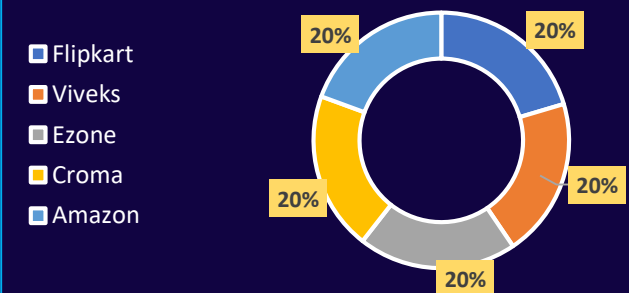
Q7. Get the complete report of the Gross sales for the customer "AtliQ Exclusive" for each month .
-- This analysis helps to get an idea of low and high-performing months and take strategic decisions.

SQL Query

```
--------------------------------------------------------

SELECT monthname(date_add(f.date,interval 4 month)) as Month_Name ,
year(date_add(f.date,interval 4 month)) as Fiscal_Year,
ROUND(SUM(fg.gross_price*f.sold_quantity),2) as Gross_sales_amount from dim_customer c
JOIN
fact_sales_monthly f ON c.customer_code = f.customer_code
JOIN
fact_gross_price fg ON fg.product_code = f.product_code AND fg.fiscal_year = f.fiscal_year
WHERE c.customer = 'Atliq Exclusive'
GROUP BY monthname(date_add(f.date,interval 4 month)),
year(date_add(f.date,interval 4 month))
ORDER BY year(date_add(f.date,interval 4 month)) desc;
```
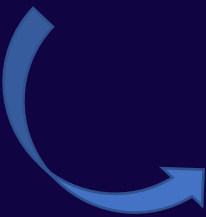
Result

Month Name	Fiscal Year	Gross sales amount
January	2021	₹ 12353.5M
February	2021	₹ 13218.6M
March	2021	₹ 20465.0M
April	2021	₹ 12944.7M
May	2021	₹ 12399.4M
June	2021	₹ 10129.7M
July	2021	₹ 12144.1M
August	2021	₹ 7312.0M
September	2021	₹ 12150.2M
October	2021	₹ 9824.5M
November	2021	₹ 12092.3M
December	2021	₹ 7178.7M
January	2020	₹ 4496.3M
February	2020	₹ 5135.9M
March	2020	₹ 7522.9M
April	2020	₹ 4830.4M
May	2020	₹ 4740.6M
June	2020	₹ 3996.2M
July	2020	₹ 378.8M
August	2020	₹ 395.0M
September	2020	₹ 783.8M
October	2020	₹ 1695.2M
November	2020	₹ 2551.2M
December	2020	₹ 2786.6M

Q8. In which quarter of 2020, got the maximum total_sold_quantity?

SQL Query

```
-------------------------------------------------------

with CTE as
  (SELECT quarter(date_add(date,interval 4 month)) as Quarter_No,
  sum(sold_quantity) as total_sold_quantity from fact_sales_monthly
  WHERE fiscal_year=2020
  GROUP BY quarter(date_add(date,interval 4 month)))
  SELECT Quarter_No,total_sold_quantity
  FROM CTE
  ORDER BY total_sold_quantity DESC limit 1;
```

Result

Quarter_No	total_sold_quantity
1	7005619

Q9. Which channel helped to bring more gross sales in the fiscal year 2021 and the percentage of contribution?
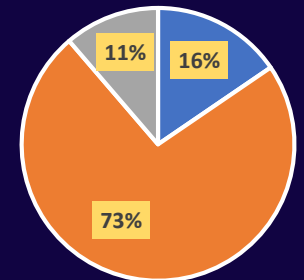
SQL Query

```sql
------------------------------------------------------
WITH CTE as
(SELECT c.channel,
ROUND(SUM(fs.sold_quantity*fp.gross_price)/1000000,2) as Total_Gross_Sales_mln
FROM
fact_sales_monthly fs JOIN dim_customer c
ON fs.customer_code = c.customer_code
JOIN
fact_gross_price fp
ON
fp.fiscal_year = fs.fiscal_year AND fp.product_code = fs.product_code
WHERE fs.fiscal_year = 2021
GROUP BY c.channel)
SELECT channel,Total_Gross_Sales_mln ,
ROUND(Total_Gross_Sales_mln*100/(SELECT sum(Total_Gross_Sales_mln) from CTE),2) as percentage
FROM CTE
GROUP BY channel;
```

Result

Total Gross Sales (mln) by Channel
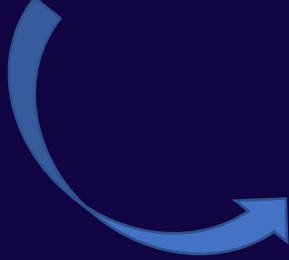

- Direct
- Retailer
- Distributor

11% 16% 73%

Q10. Get the Top 3 products in each division that have a high total sold quantity in the fiscal year 2021?

SQL Query

```sql
-----------------------------------------------------

with CTE as
(
SELECT fs.product_code,p.product,p.division, SUM(sold_quantity) as Total_sold_quantity
FROM fact_sales_monthly fs
JOIN dim_product p ON p.product_code = fs.product_code
WHERE fs.fiscal_year = 2021
GROUP BY fs.product_code,p.product,p.division
)
,CTE2 AS
(
SELECT division,product,product_code,Total_sold_quantity,
rank() over(partition by division order by Total_sold_quantity desc) as rnk
FROM CTE
)
SELECT division,product_code,product,Total_sold_quantity,rnk FROM CTE2 WHERE rnk≤3;
```

Result

division	product_code	product	Total_sold_quantity	rnk
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

Consumer Goods : Ad_Hoc Insights – *SQL Project*

Key Takeaways:

1. Accessories and Peripherals led with the most unique products (20).

2. Unique product offerings increased by 33% in FY 2021.

3. The top 5 customers received an average 20% pre-invoice discount.

4. The Retailer channel contributed 73% to total gross sales in FY 2021.

5. All segments saw an increase in unique products in FY 2021.