Ad-Hoc Insights

Presented by : Pratik Yadav

Content

1. Company Overview
2. Problem statement
3. Dataset and model
4. Ad-hoc requests , output and insights

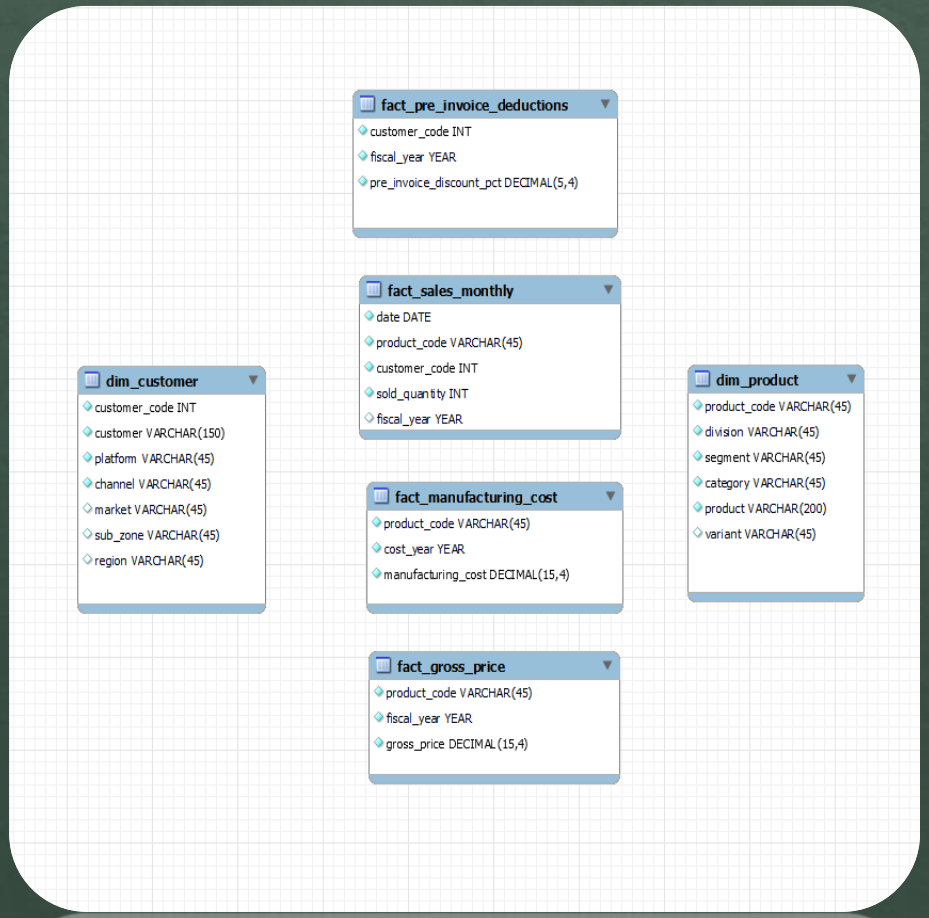
Company Overview

AtliQ Hardwares (fictitious company) is one of the leading computer hardware manufacturers in India, with strong presence in other nations.

Problem Statement

- Management noticed that they do not have sufficient insights to make data-driven decisions.
- Decision made to expand the data analytics team with several junior data analysts.
- Tony Sharma, the data analytics director, seeks candidates with strong tech and soft skills.
- To evaluate these skills, Tony decided to conduct a SQL challenge.
- The company seeks insights for 10 ad hoc requests.

Dataset and 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

codebasics.io

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

codebasics.io

Ad-hoc Requests, SQL Query, Output and Insights

Request-1

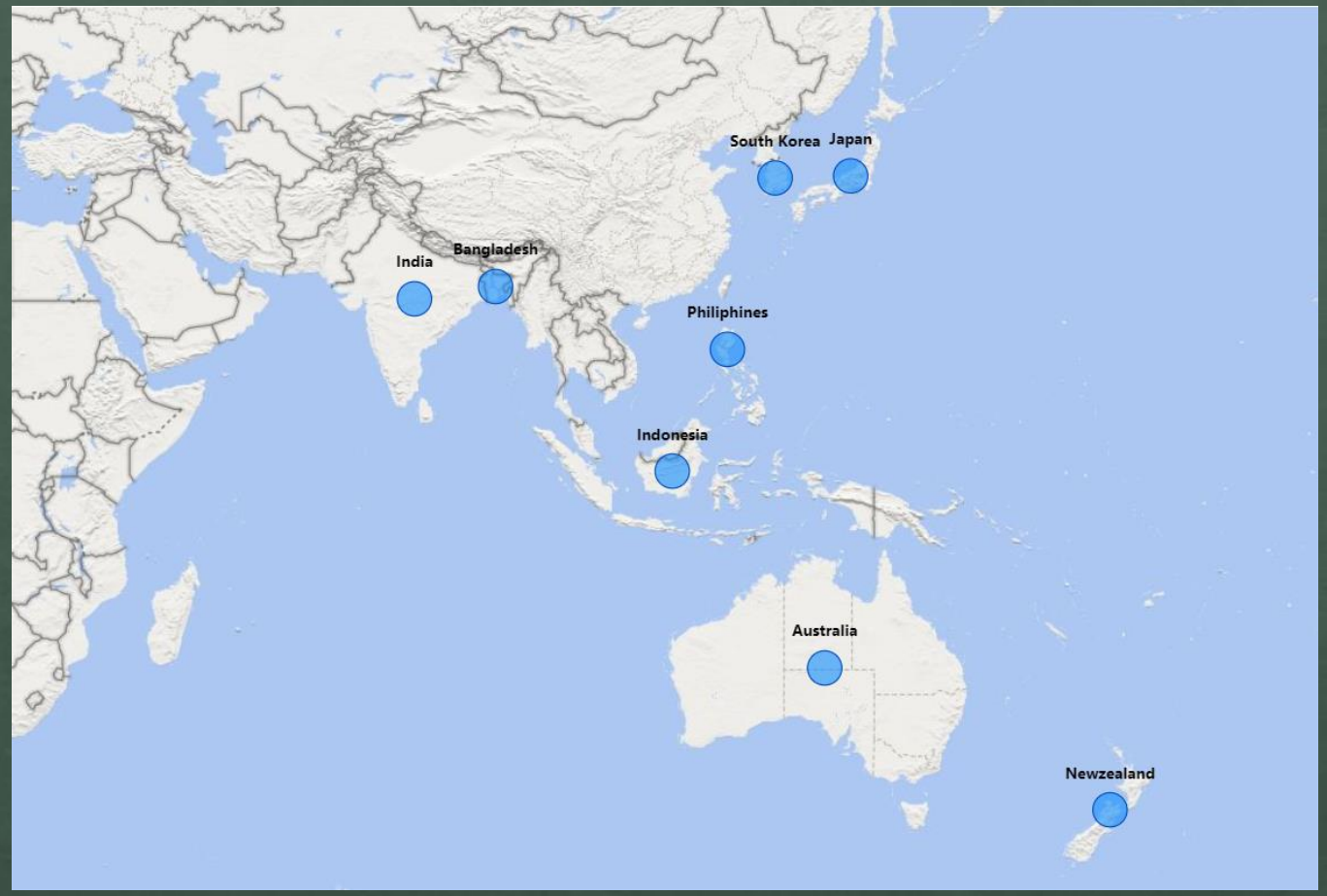
Provide the list of markets in which customer "AtliQ Exclusive" operates its business in the APAC region.

SQL Query:

```
Select
distinct market
from dim_customer
where
customer = "AtliQ Exclusive" and region
= "APAC”
;
```

Output

	market
▶	India
	Indonesia
	Japan
	Philiphines
	South Korea
	Australia
	Newzealand
	Bangladesh



Insights

AtliQ Exclusive is active in eight countries across the APAC region, showcasing its considerable market presence.

Request-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.

SQL Query:

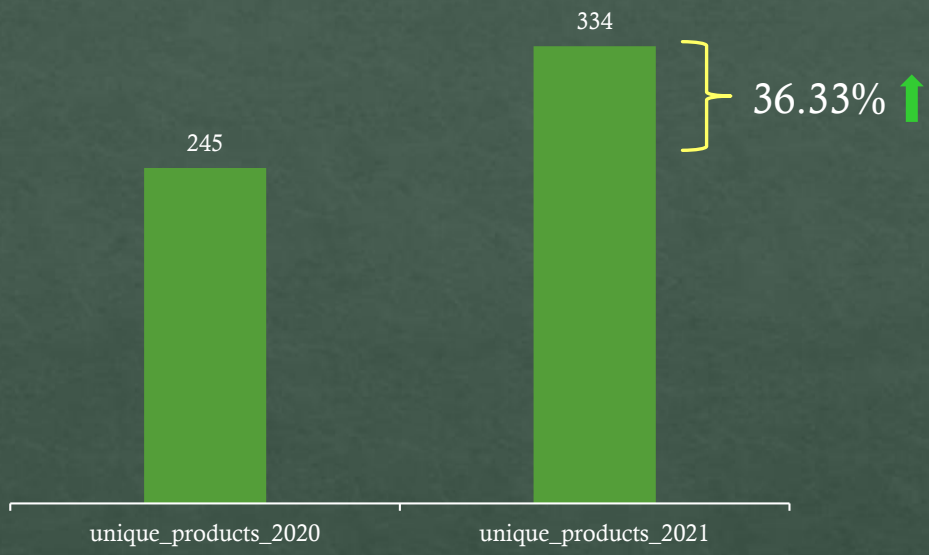
```
with cte1 as
(Select count(distinct product_code) as unique_products_2020
from fact_sales_monthly
where fiscal_year = 2020),

cte2 as
(Select count(distinct product_code) as unique_products_2021
from fact_sales_monthly
where fiscal_year = 2021)

Select unique_products_2020, unique_products_2021,
Round((unique_products_2021  unique_products_2020)*100/unique_products_2020,2) as percentage_chg
from cte1 cross join cte2
;
```

Output

	unique_products_2020	unique_products_2021	percentage_chg
▶	245	334	36.33



Insights

1. The number of unique products grew from 245 in 2020 to 334 in 2021, marking a 36.33% increase.

2. This significant increase highlights the company's commitment to innovation and meeting diverse customer needs.

Request-3

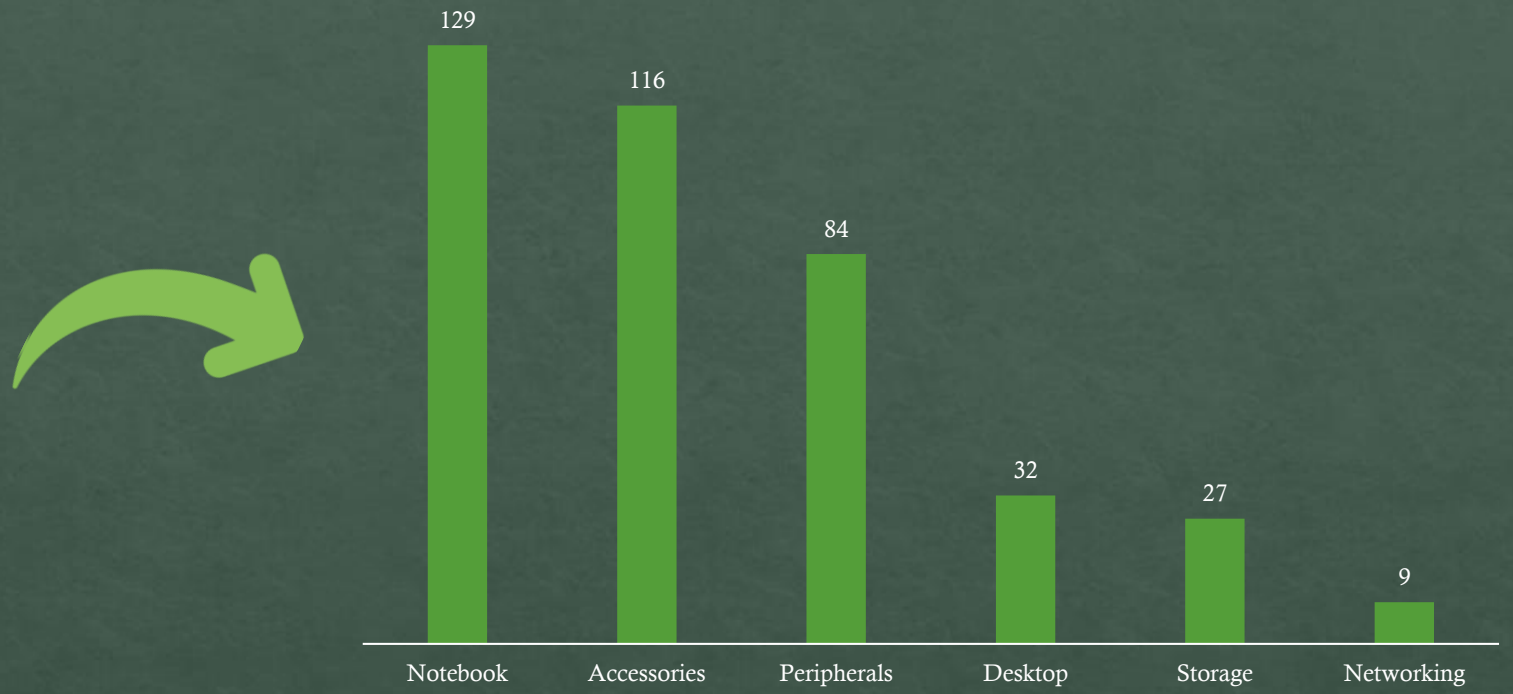
Provide a report with all the unique product counts for each segment and sort them in descending order of product counts. The final output contains 2 fields, segment, product_count.

SQL Query:

```
Select
segment,
count(distinct product_code) as product_count
from dim_product
group by segment
order by product_count desc
;
```

Output

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



Insights

1. AtliQ shines in Notebooks, Accessories, and Peripherals segment, accounting for a whopping 82.87% of product variety.

2. Desktops, Storage, and Networking currently hold a minor share (17.13%) of AtliQ's product portfolio.

3. In Desktops, Storage, and Networking segment AtliQ should focus on products that align with current customer trends and industry demands.

Request-4

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.

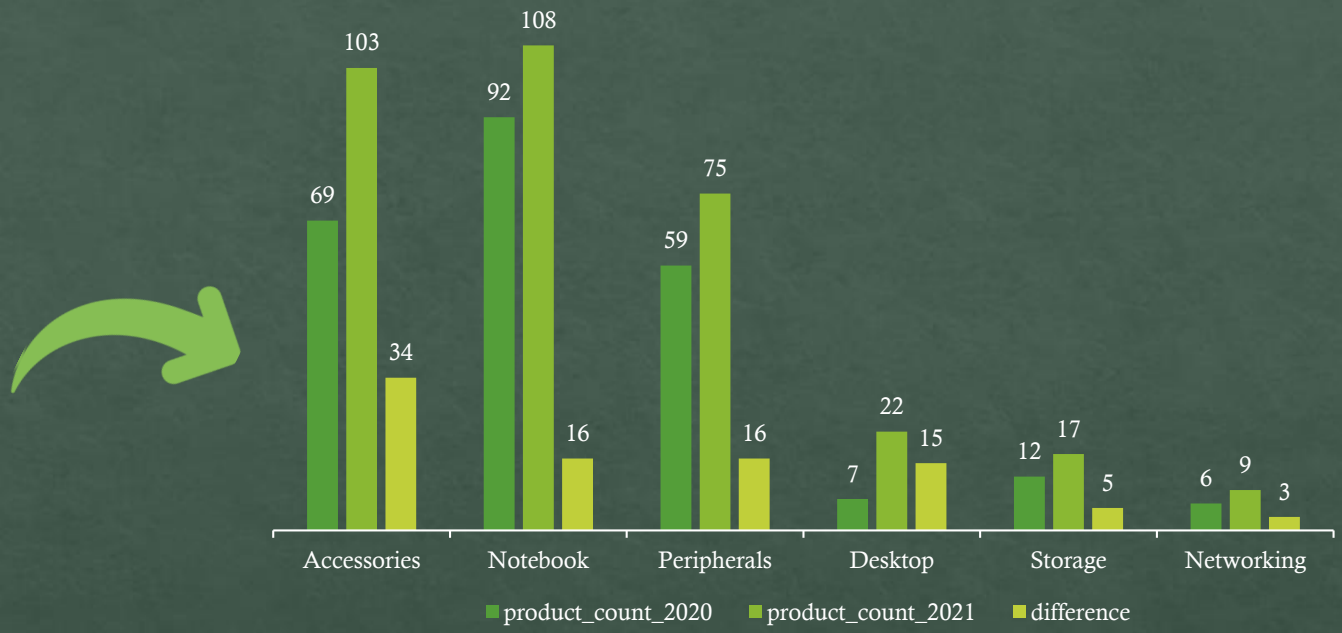
SQL Query:

```
with cte1 as
(Select
 p.segment,
count(distinct(case when s.fiscal_year=2020 then p.product_code end)) as product_count_2020,
count(distinct(case when s.fiscal_year=2021 then p.product_code end)) as product_count_2021
from dim_product p
join fact_sales_monthly s
on p.product_code = s.product_code
group by p.segment)

Select *,
product_count_2021-product_count_2020 as difference
from cte1
order by difference desc;
```

Output

segment	product_count_2020	product_count_2021	difference
Accessories	69	103	34
Notebook	92	108	16
Peripherals	59	75	16
Desktop	7	22	15
Storage	12	17	5
Networking	6	9	3



Insights

1. Accessories have the most increased products from 2020 to 2021, with an increase of 34 new products.

2. Desktop products production increased drastically by approximately 214% from 2020 to 2021.

3. Storage and Networking segments are producing the least new products from 2020 to 2021.

Request-5

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

SQL Query:

```
(Select
p.product_code, p.product, m.manufacturing_cost
from dim_product p
join fact_manufacturing_cost m
on p.product_code = m.product_code
order by m.manufacturing_cost desc
limit 1)
Union
(Select
p.product_code, p.product, m.manufacturing_cost
from dim_product p
join fact_manufacturing_cost m
on p.product_code = m.product_code
order by m.manufacturing_cost asc
limit 1)
;
```

Output

	product_code	product	manufacturing_cost
▶	A6121110208	AQ HOME Allin1 Gen 2	263.4207
	A2118150101	AQ Master wired x1 Ms	0.8654

Insights

1. The AQ HOME Allin 1 Gen 2 has the highest manufacturing cost of 263.42 among all the products.

2. The AQ Master Wired x1 Ms has the lowest manufacturing cost of 0.87 among all the products.

Request-6

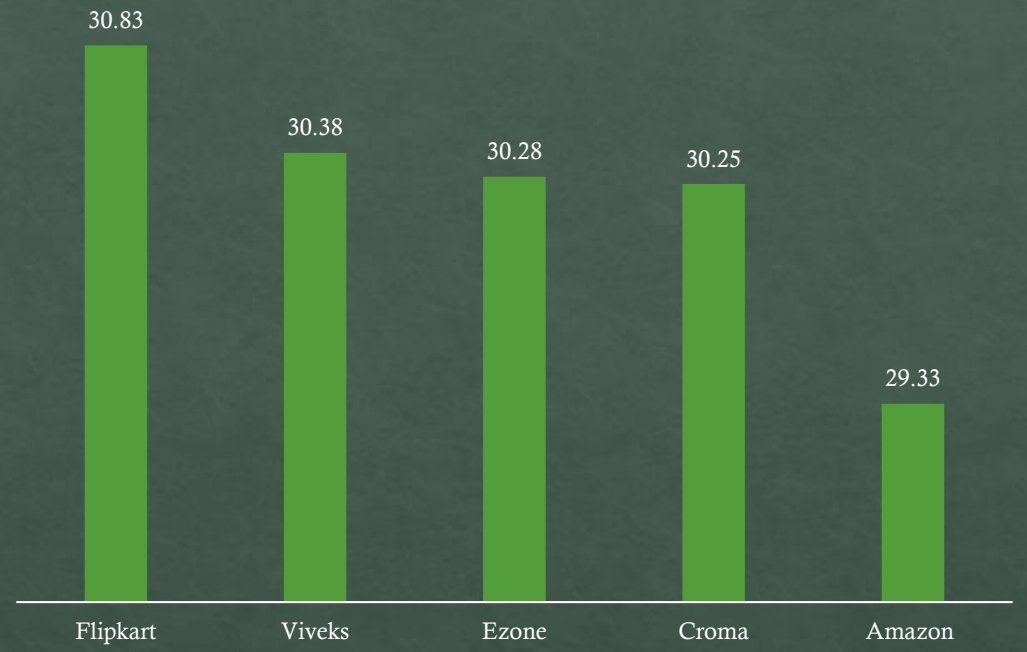
Generate a report which contains the top 5 customers who received an average high pre_invoice_discount_pct for the fiscal year 2021 and in the Indian market. The final output contains these fields, customer_code, customer, average_discount_percentage .

SQL Query:

```
Select
c.customer_code, c.customer, round(avg(p.pre_invoice_discount_pct*100),2) as average_discount_percentage
from dim_customer c
join fact_pre_invoice_deductions as p
on c.customer_code = p.customer_code
where p.fiscal_year = 2021 and c.market = "India“
group by c.customer_code
order by p.pre_invoice_discount_pct desc
limit 5
;
```

Output

customer_code	customer	avg_discount_pct
90002009	Flipkart	30.83%
90002006	Viveks	30.38%
90002003	Ezone	30.28%
90002002	Croma	30.25%
90002016	Amazon	29.33%



Insights

1. Flipkart leads the list with the highest average discount percentage at 30.83%.

2. AtliQ has provided relatively uniform discount percentages to its top 5 customers, with a narrow range of discounts from 29.33% to 30.83%.

Request-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.

SQL Query:

```
Select monthname(s.date) as Month, Year(s.date) as Year,
Round(sum(s.sold_quantity*g.gross_price),2) as Gross_Sales_amount
from fact_sales_monthly s
join fact_gross_price g
on s.product_code = g.product_code
join dim_customer c on s.customer_code = c.customer_code
where c.customer = "Atliq Exclusive“
group by Month, Year
;
```

Output

	September	2019	17040562.24
	October	2019	19475069.34
	November	2019	28512004.15
	December	2019	18322529.10
	January	2020	18009123.49
	February	2020	15171931.84
	March	2020	1422525.16
	April	2020	1492369.18
	May	2020	2971173.85
	June	2020	6451963.41
	July	2020	9685828.63
	August	2020	10599401.49
	September	2020	37752848.19
	October	2020	40444450.56
	November	2020	62302295.57
	December	2020	39306619.91
	January	2021	37704996.30
	February	2021	30852326.32
	March	2021	36972600.02
	April	2021	22147394.99
	May	2021	37037201.24
	June	2021	29887193.88
	July	2021	36677913.91
	August	2021	21839400.72

Insights

1. March, April, May recorded the lowest sales in 2020, likely due to the COVID-19 pandemic. Sales began to recover after June 2020.

2. November generated the highest sales for 2020, with around 28.51 million.

3. Sales figures in 2021 show significant growth compared to 2020.

4. In 2021, the lowest sales were in August, accounting for 21.83 million, while the highest sales were in November, reaching 62.3 million.

Request-8

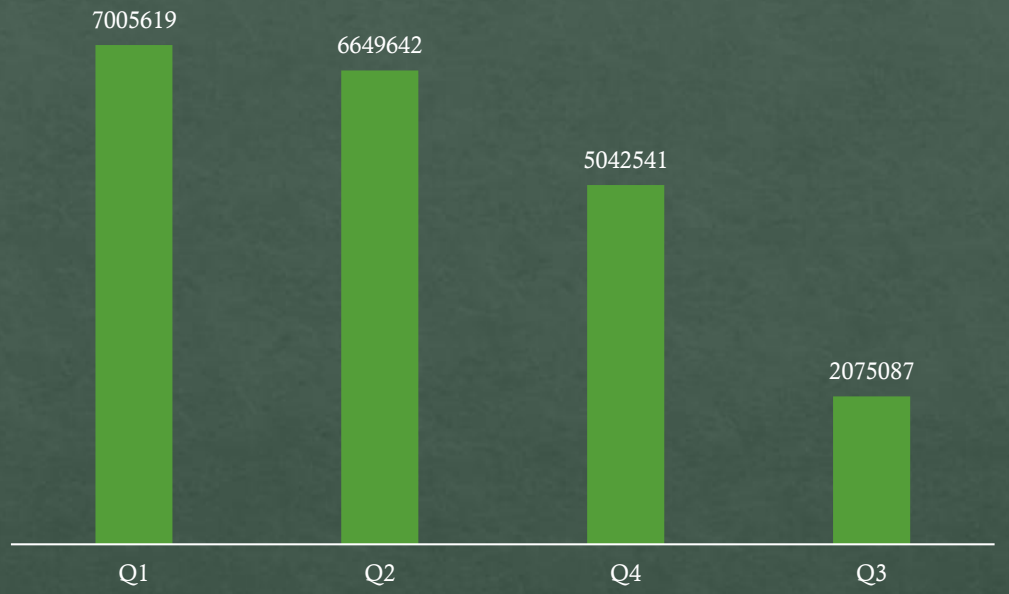
In which quarter of 2020, got the maximum total_sold_quantity? The final output contains these fields sorted by the total_sold_quantity, Quarter, total_sold_quantity.

SQL Query:

```
Select
        (case
        when month(date) in (9,10,11) then "Q1"
        when month(date) in (12,1,2) then "Q2"
        when month(date) in (3,4,5) then "Q3”
        else "Q4" end) as Quarter,
        sum(sold_quantity) as total_sold_quantity
from fact_sales_monthly
where fiscal_year = 2020
group by Quarter
order by total_sold_quantity desc
;
```

Output

Quarter	total_sold_qty
Q1	7005619
Q2	6649642
Q4	5042541
Q3	2075087



Insights

1. Q1 (2020) saw the highest sales volume, reaching around 7 million units, indicating strong performance early in the year.

2. AtliQ experienced a significant decline in sales during Q3 (March, April, May), possibly due to the COVID-19 pandemic, reflecting challenging market conditions and shifting consumer behavior.

Request-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.

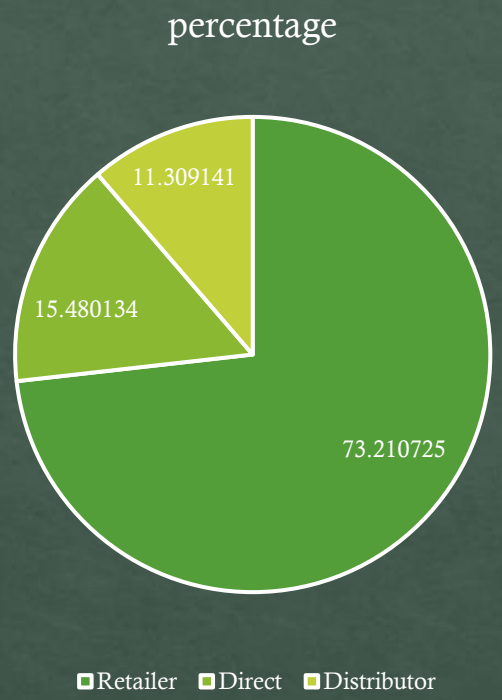
SQL Query:

```
 with cte1 as
(Select c.channel, round(sum(g.gross_price*s.sold_quantity/1000000),2) as gross_sales_mln
from fact_sales_monthly s
join fact_gross_price g
on s.product_code = g.product_code
join dim_customer c on s.customer_code = c.customer_code
where s.fiscal_year = 2021
group by c.channel)

Select *, gross_sales_mln*100/sum(gross_sales_mln) over() as percentage
from cte1
order by gross_sales_mln desc
;
```

Output

	channel	gross_sales_mln	percentage
▶	Retailer	3708.46	73.210725
	Direct	784.14	15.480134
	Distributor	572.86	11.309141



Insights

1. The majority of AtliQ's sales come from retailers, accounting for a substantial 73.21% of the total gross sales.

2. Direct and distributor channel together contributes only 26.79% of the total sales.

Request-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.

SQL Query:

```
with cte1 as
(Select p.division, s.product_code, p.product, sum(s.sold_quantity)  as total_sold_quantity
from fact_sales_monthly s
join dim_product p
on s.product_code = p.product_code
where s.fiscal_year = 2021
group by p.product_code, p.division),

cte2 as
(Select division, product_code, product, total_sold_quantity,rank() over(partition by division order by
total_sold_quantity desc) as 'rank_order’
from cte1)

Select cte1.division, cte1.product_code, cte1.product, cte2.total_sold_quantity, cte2.rank_order
from cte1 join cte2
on cte1.product_code = cte2.product_code
where cte2.rank_order in (1,2,3)
;
```

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. In the **N & S** division, **pen drives** dominate the top three spots, showcasing their popularity and high sales performance.

2. The **P & A** division's top three products are exclusively **mouse**, reflecting a clear demand for these devices.

3. The **PC** division's bestsellers are all **personal laptops**.

Thankyou