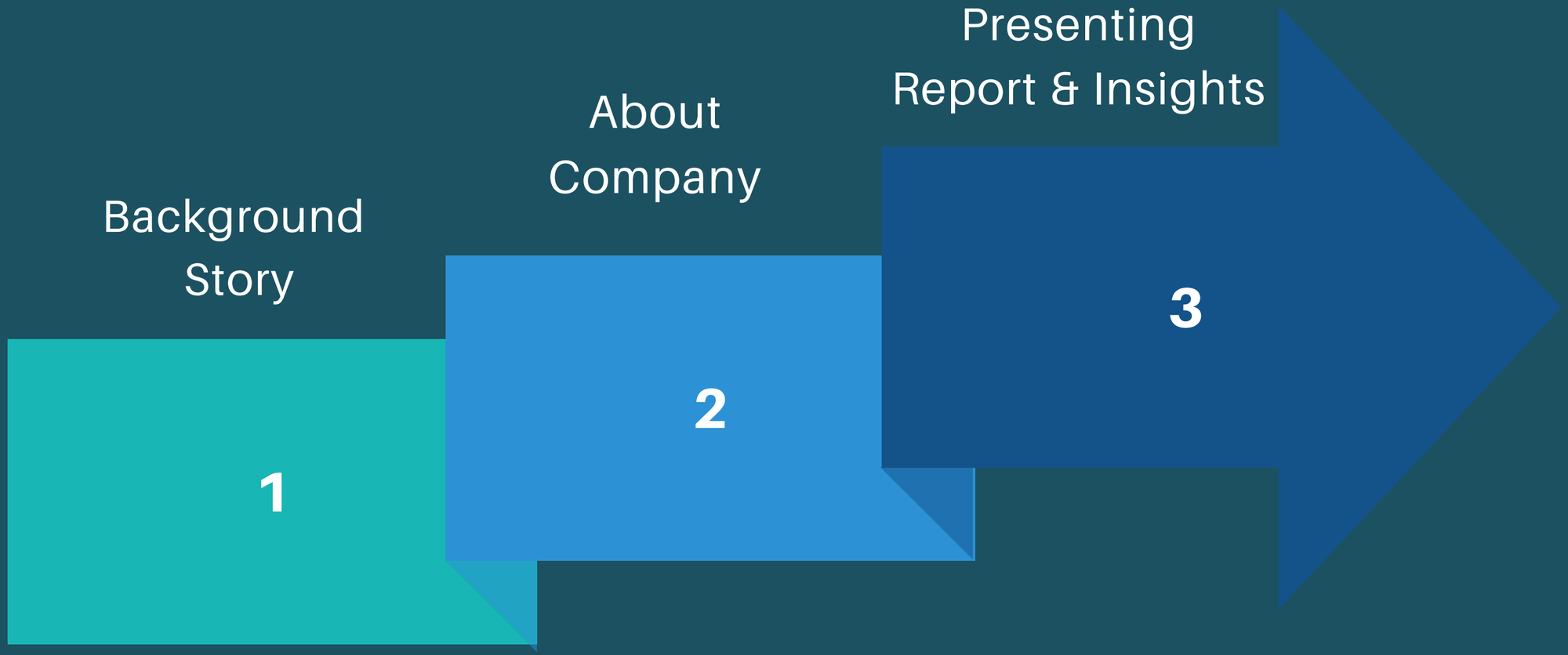
Codebasics
4th Resume Project Challenge

AtliQ Hardware

Consumer Goods Ad-hoc Insights

-Presented by Devesh Singh

Today's Agenda



Background Story

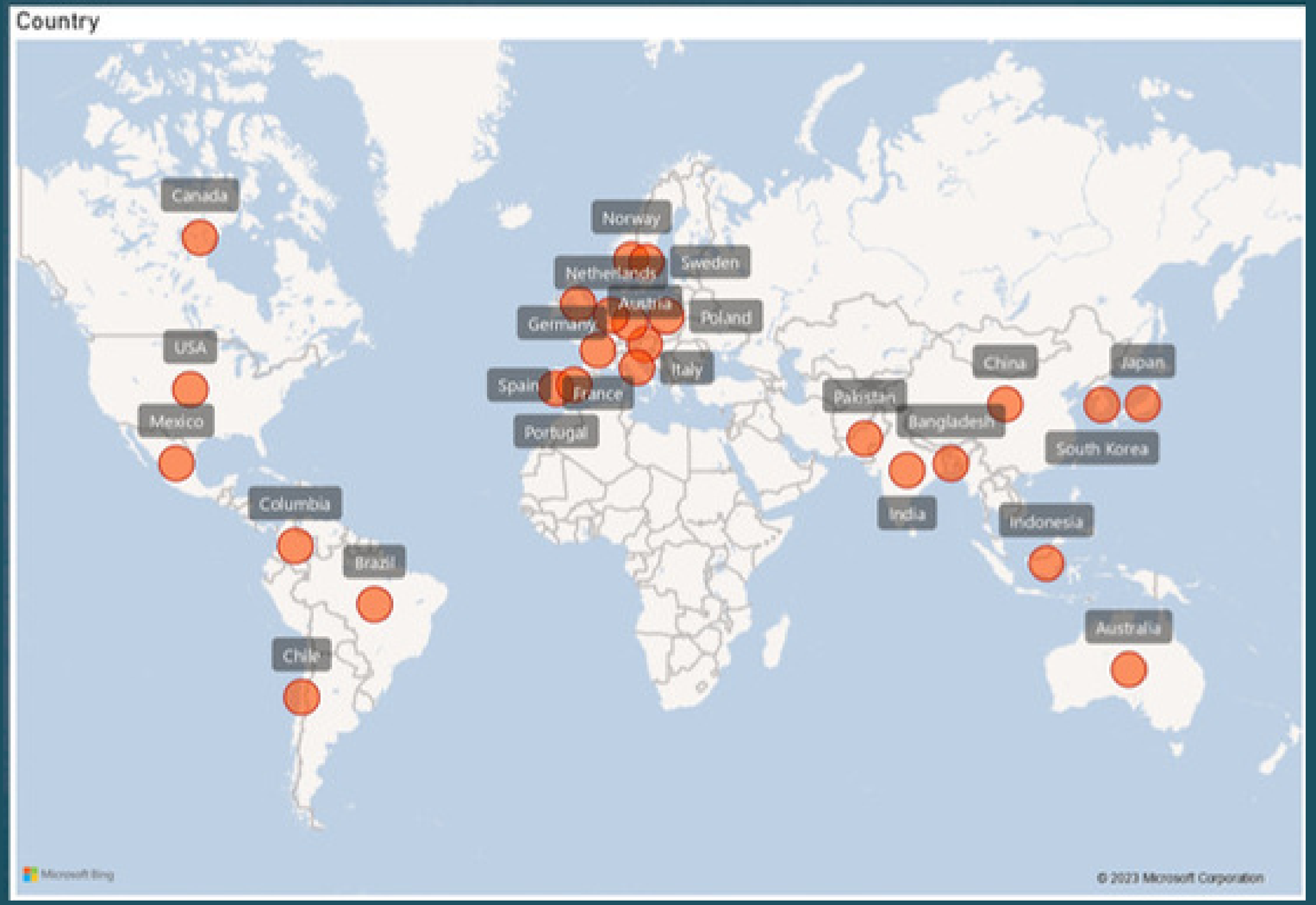
- Atliq Hardwares (imaginary company) is one of the leading computer hardware & accessory producers in India and well expanded in other countries too.

- However, the management noticed that they do not get enough insights to make quick and smart data-informed decisions. They want to expand their data analytics team by adding several junior data analysts.

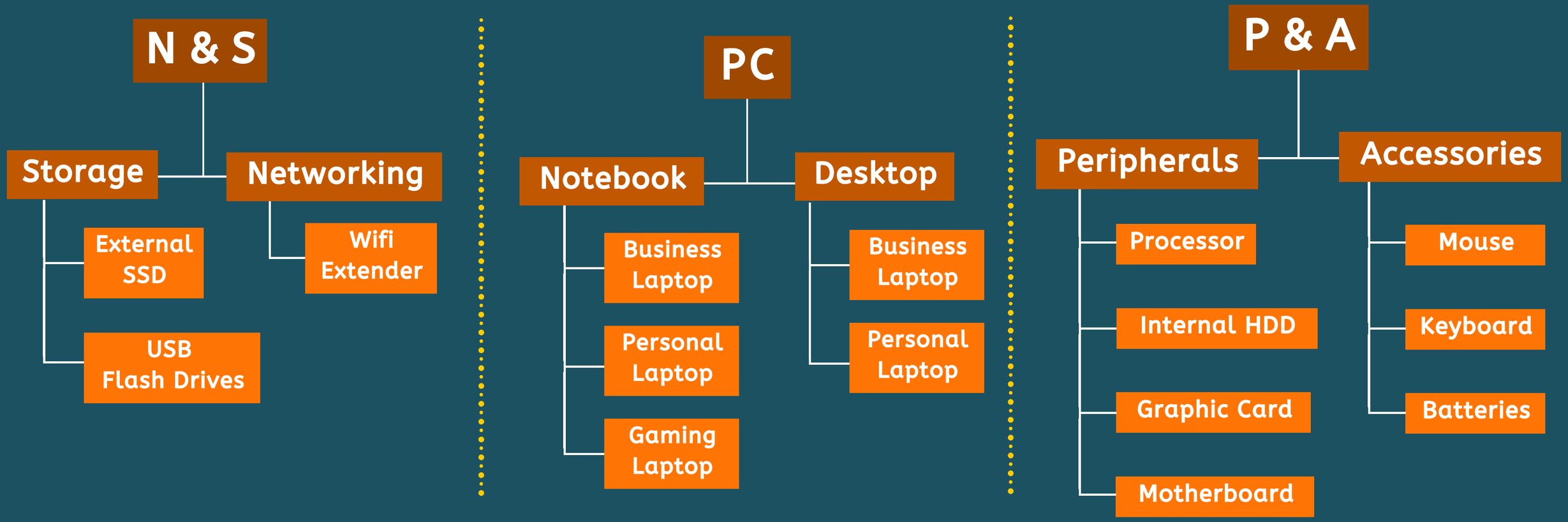
- To identify the right talent for their team, they are conducting a SQL challenge that will evaluate candidates' technical as well as interpersonal skills necessary to collaborate effectively in a dynamic environment.

About Company
(Global Presence)

Region	Countries (Market)
APAC	Australia
APAC	Bangladesh
APAC	China
APAC	India
APAC	Indonesia
APAC	Japan
APAC	Newzealand
APAC	Pakistan
APAC	Philiphines
APAC	South Korea
EU	Austria
EU	France
EU	Germany
EU	Italy
EU	Netherlands
EU	Norway
EU	Poland
EU	Portugal
EU	Spain
EU	Sweden
EU	United Kingdom
LATAM	Brazil
LATAM	Chile
LATAM	Columbia
LATAM	Mexico
NA	Canada
NA	USA



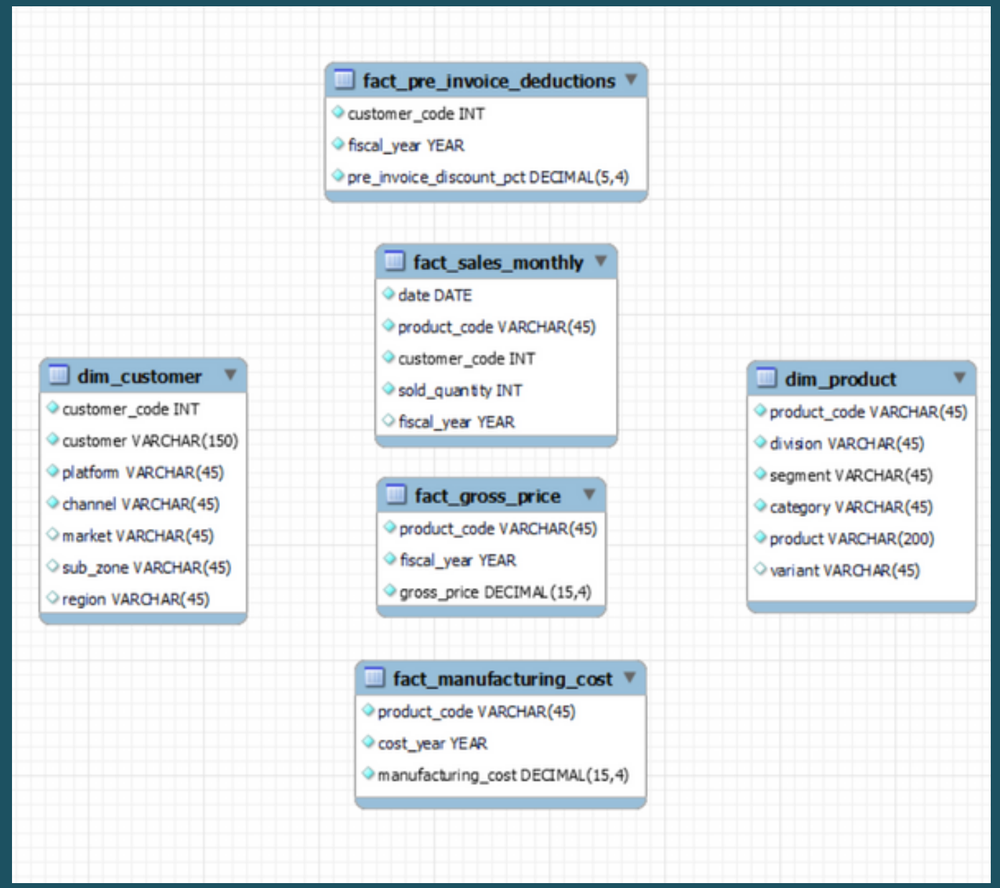
(AtliQ's Product Lines)



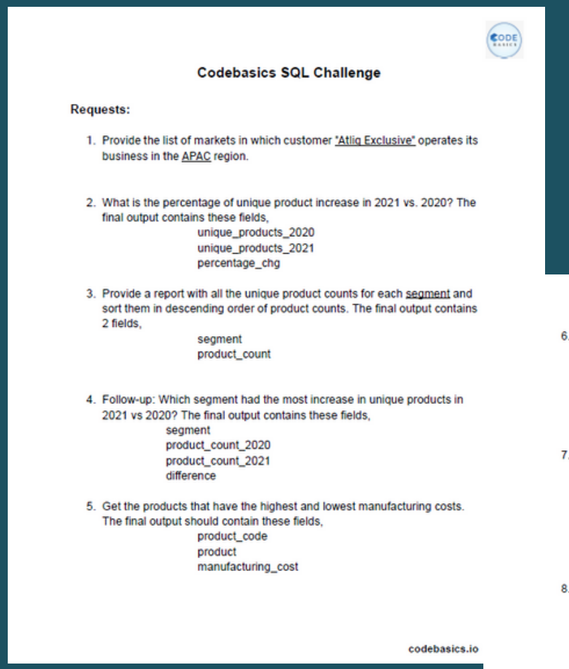
(Data, Requests, Tools Used)

Fiscal Years: 2020 & 2021

September - August



MySQL



Codebasics SQL Challenge

Requests:

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

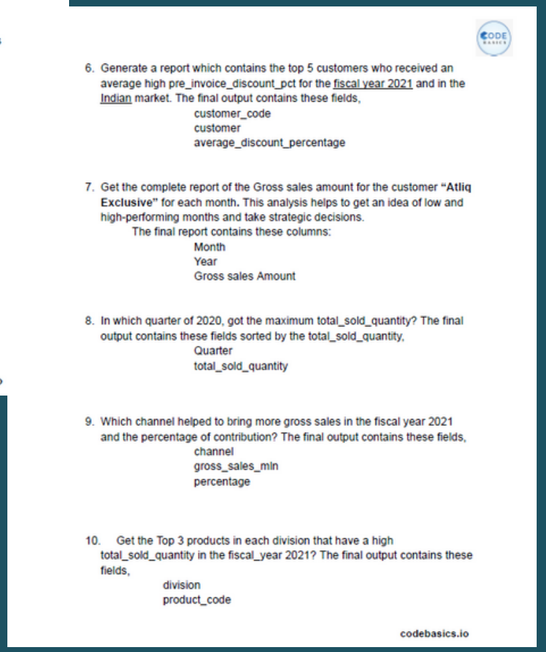
2. What is the percentage of unique product increase in 2021 vs. 2020? The final output contains these fields,
   unique_products_2020
   unique_products_2021
   percentage_chg

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

Reports & Insights

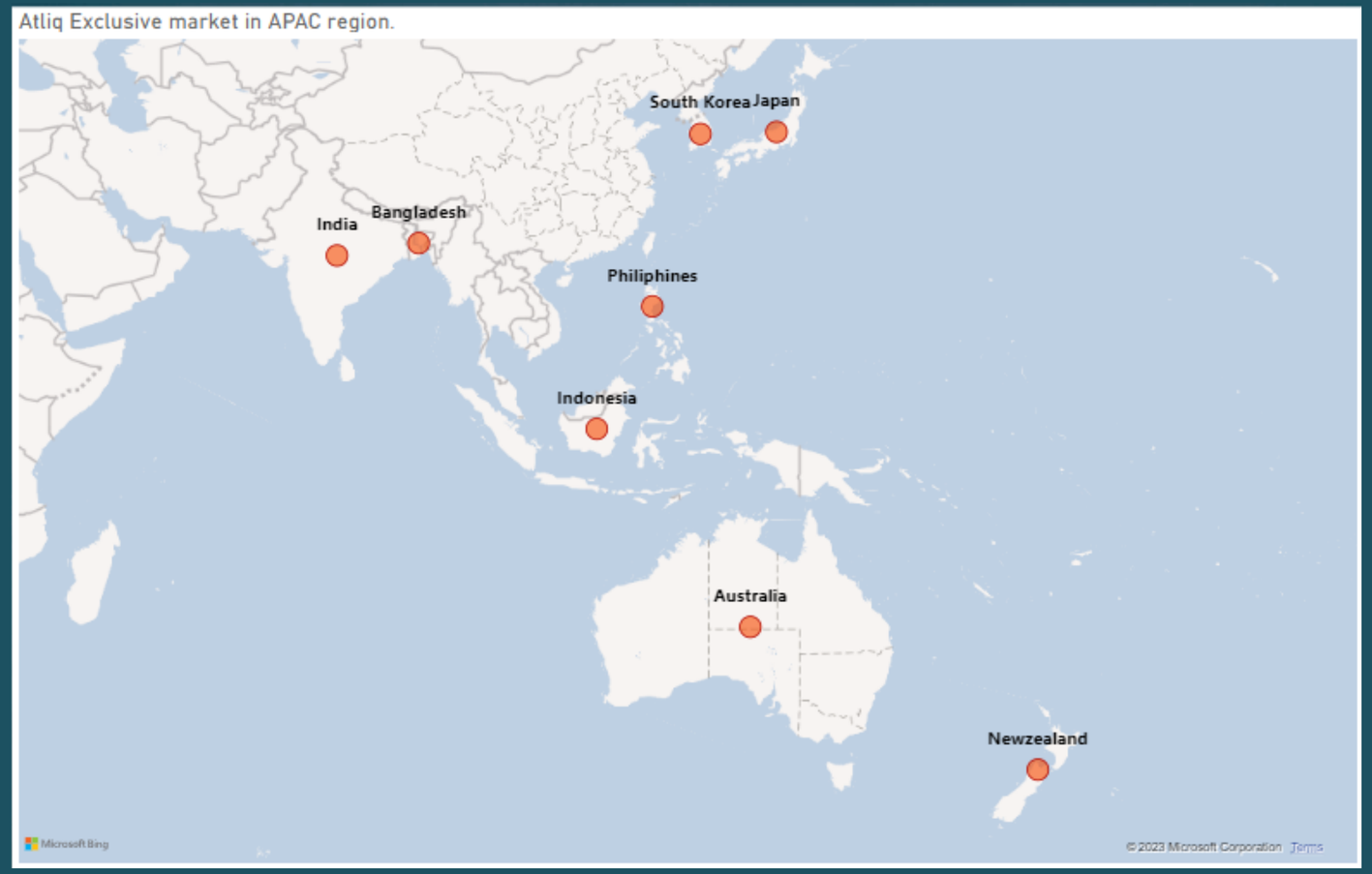
- Ad-hoc request
- SQL queries result
- Visualization and Insights

Request -1

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

```
3
4              ####################    Request -1    ####################
5
6 •  SELECT distinct market
7    FROM dim_customer
8    WHERE customer = 'Atliq Exclusive' AND region = 'APAC';
9
```

	market
▶	India
	Indonesia
	Japan
	Philiphines
	South Korea
	Australia
	Newzealand
	Bangladesh



Request -2

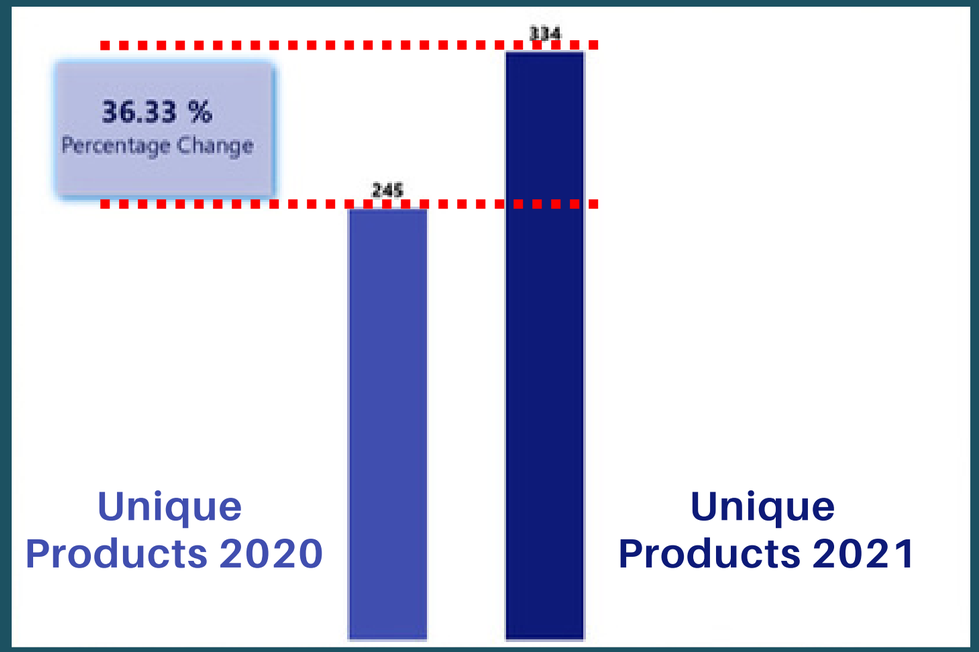
--> What is the percentage of unique product increase in 2021 vs. 2020? The final output contains these fields:

unique_products_2020 , unique_products_2021, percentage_chg

```sql
                    ####################    Request -2    ####################

-- By using CTE and CASE statement
/* Here we are providing the condition for fiscal year from inside by using CASE statement */
WITH cte as (
        SELECT
            COUNT(DISTINCT CASE WHEN fiscal_year = 2020 THEN product_code END) AS unique_product_2020,
            COUNT(DISTINCT CASE WHEN fiscal_year = 2021 THEN product_code END) AS unique_product_2021
        FROM fact_sales_monthly
        WHERE fiscal_year IN (2020, 2021)
       )
SELECT unique_product_2020, unique_product_2021,
       ROUND((unique_product_2021 - unique_product_2020)*100/unique_product_2020, 2) as percentage_chg
       FROM cte;
```

	unique_product_2020	unique_product_2021	percentage_chg
▶	245	334	36.33



Insight:

- New 89 product were introduced in 2021, indicating that company is bringing new products in the market to increase its customer base.

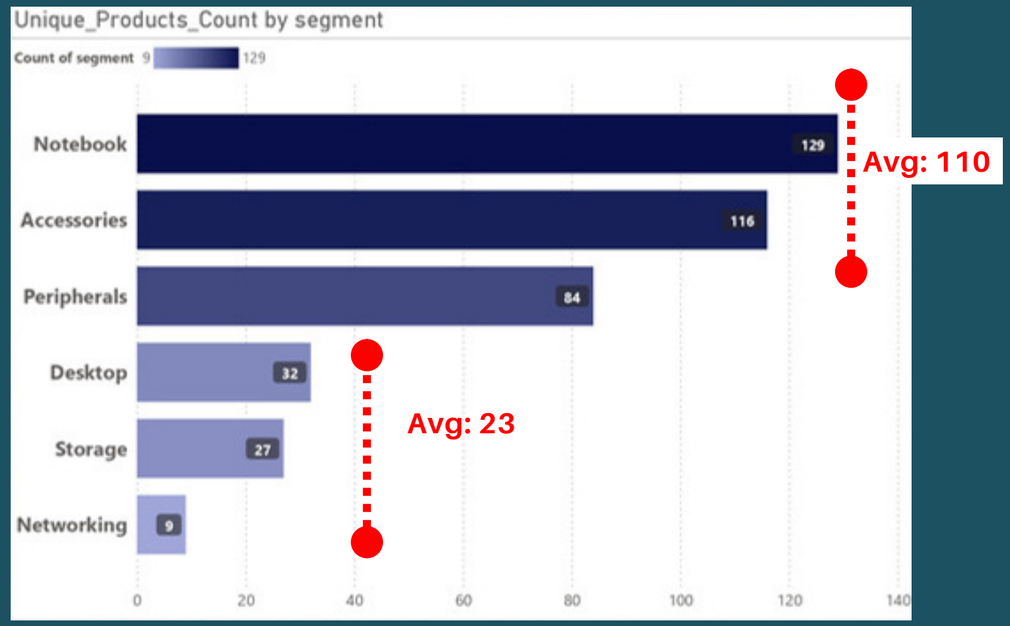
Request -3

-->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

```
#####################    Request -3

SELECT
      segment, count(product_code) as product_count
      FROM dim_product
      GROUP BY segment
      ORDER BY product_count DESC;
```

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



Insights:

- The average of top 3 segments is almost 5 times the average of lower 3 segments, means around 83% of the different products lies in the Notebook, Accessories and Peripherals segments.
- Both segments Notebook and Desktop belongs to same Division (PC) but the gap is of **97 units** i.e. very high even though desktop is a demanding category which indicates that we need to introduce more product in Desktop segment.

Request -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

```sql
53
54                    ####################    Request -4    ####################
55
56  WITH cte as(
57          SELECT
58                  dp.segment,
59                  count(distinct CASE WHEN fiscal_year=2020 THEN fs.product_code END) as product_count_2020,
60                  count(distinct CASE WHEN fiscal_year=2021 THEN fs.product_code END) as product_count_2021
61                FROM fact_sales_monthly fs
62                JOIN dim_product dp
63                USING(product_code)
64                WHERE fiscal_year IN (2020, 2021)
65                GROUP BY dp.segment
66      )
67
68    SELECT segment, product_count_2020,  product_count_2021,
69          (product_count_2021 -  product_count_2020) as difference
70          FROM cte;
71
```

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

segment	Unique Products 2020	Unique Products 2021	Difference
Networking	6	9	3
Storage	12	17	5
Desktop	7	22	15
Notebook	92	108	16
Peripherals	59	75	16
Accessories	69	103	34
Total	245	334	89

Insights:

- Accessories has highest increase in the number of total unique products from 2020 to 2021 i.e. 34.
- Desktop has around 200% increase in the list which is record breaking among all segment. It shows that we are bringing different categories for our customers but still there is lot more to explore.

Request -5

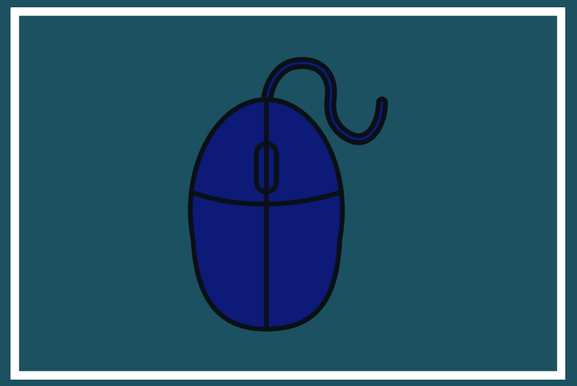
-->Get the products that have the highest and lowest manufacturing costs. The final output should contain these fields: product_code, product, manufacturing_cost

```
78
79                    ####################    Request -5    ####################
80
81 •  SELECT
82        fc.product_code, dp.product, manufacturing_cost
83        FROM fact_manufacturing_cost fc
84        JOIN dim_product dp
85        USING (product_code)
86        WHERE manufacturing_cost IN
87              (SELECT max(manufacturing_cost) FROM fact_manufacturing_cost
88                 UNION
89               SELECT min(manufacturing_cost) FROM fact_manufacturing_cost);
90
91
```

product_code	product	manufacturing_cost
A2118150101	AQ Master wired x1 Ms	0.8920
A6120110206	AQ HOME Allin1 Gen 2	240.5364

240.54 Dollar

AQ HOME Allin1 Gen 2
(Plus 3)
Personal Desktop



0.89 Dollar

AQ Master wired x1 Ms
(Standard 1)
Mouse

Request -6

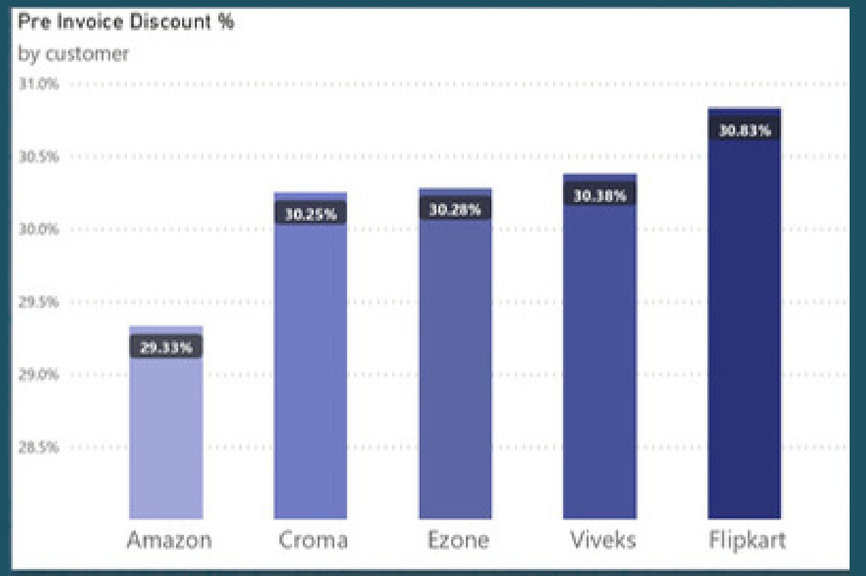
-->Generate a report which contains the top 5 customers who received an average high pre_invoice_discount_pct for the fiscal year 2021 and in the Indian market. The final output contains these fields: customer_code, customer, average_discount_percentage

```
#####################    Request -6    #####################

SELECT
    pid.customer_code, dc.customer,
    ROUND(AVG(pid.pre_invoice_discount_pct), 4) as average_discount_percentage
    FROM fact_pre_invoice_deductions pid
    JOIN dim_customer dc
    USING (customer_code)
    WHERE dc.market="India" AND pid.fiscal_year= 2021
    GROUP BY dc.customer,pid.customer_code
    ORDER BY average_discount_percentage DESC
    LIMIT 5;
```

FY: 2021

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



Insights:

- In 2021, all top 5 customer got almost equal pre-invoice discount which is around 30 %.
- Gven that Flipkart was the most discounted customer and whereas Amazon was least.

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:

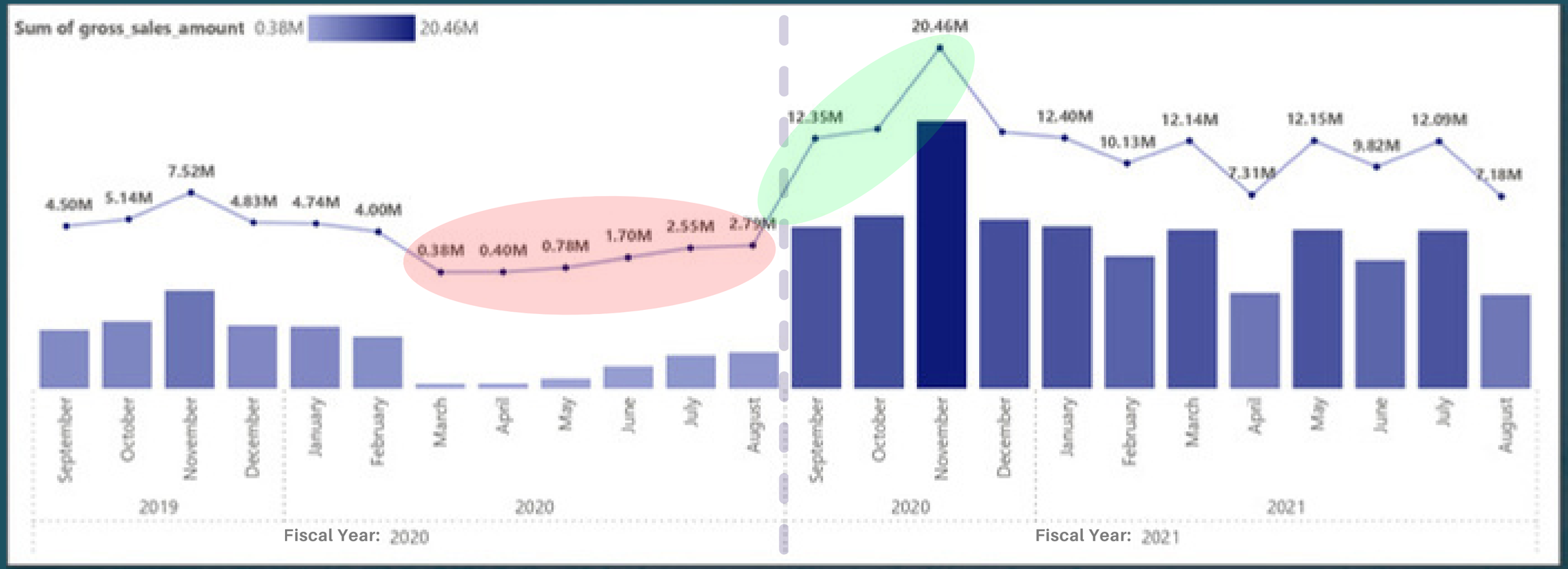
Month, Year, Gross sales Amount

```
9
0                  ####################    Request -7    ####################
1
2 •  SELECT
3        CONCAT(MONTHNAME(fs.date),' ',YEAR(fs.date) ) as month,
4        fs.fiscal_year,
5        SUM(fg.gross_price*fs.sold_quantity) as gross_sales_amount
6       FROM fact_sales_monthly fs
7       JOIN dim_customer dc USING(customer_code)
8       JOIN fact_gross_price fg USING(product_code, fiscal_year)
9       WHERE dc.customer= 'Atliq Exclusive'
0       GROUP BY month, fs.fiscal_year
1       ORDER BY fs.fiscal_year;
2
```

Customer: AtliQ Exclusive
Market: India

month	fiscal_year	gross_sales_amount
September 2019	2020	4496259.6724
October 2019	2020	5135902.3467
November 2019	2020	7522892.5608
December 2019	2020	4830404.7285
January 2020	2020	4740600.1605
February 2020	2020	3996227.7661
March 2020	2020	378770.9700
April 2020	2020	395035.3535
May 2020	2020	783813.4238
June 2020	2020	1695216.6008
July 2020	2020	2551159.1584
August 2020	2020	2786648.2601
September 2020	2021	12353509.7938
October 2020	2021	13218636.1966
November 2020	2021	20464999.0997
December 2020	2021	12944659.6509
January 2021	2021	12399392.9788
February 2021	2021	10129735.5675
March 2021	2021	12144061.2501
April 2021	2021	7311999.9547
May 2021	2021	12150225.0139
June 2021	2021	9824521.0110
July 2021	2021	12092346.3245
August 2021	2021	7178707.5902

Customer: AtliQ Exclusive | Market: India



Insights:

- The highest Gross Sales for both the fiscal years 2020 & 2021 was in the month of November, one of the possible reason is the Diwali season.
- Lowest Gross Sales for 2020 & 2021 was in March and August respectively.
- Company had very low sales from March - to - August then it increased quickly in fiscal year 2020, this could be due the COVID-19 pandemic.

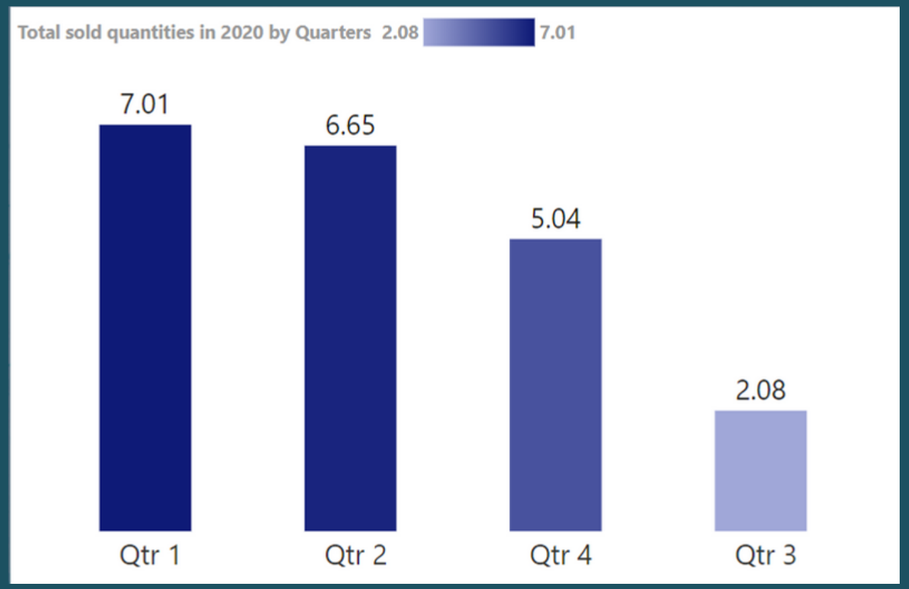
Request -8

-->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

```
125
126                    ####################    Request -8    ####################
127
128 •   SELECT
129         CASE
130             WHEN MONTH(date) IN (9,10,11) THEN 'Q1'
131             WHEN MONTH(date) IN (12,1,2) THEN 'Q2'
132             WHEN MONTH(date) IN (3,4,5) THEN 'Q3'
133             ELSE 'Q4'
134             END as quaters,
135           ROUND(SUM(sold_quantity)/1000000, 2) as total_sold_quantity_mln
136         FROM fact_sales_monthly
137         WHERE fiscal_year = 2020
138       GROUP BY quaters;
139
```

FY: 2020

	quaters	total_sold_quantity_mln
▶	Q1	7.01
	Q2	6.65
	Q3	2.08
	Q4	5.04



Insights:

- Quarter 1 of fiscal year 2020 saw the most units sold overall, while Quarter 3 had the fewest, COVID-19 was one the reason as discussed above.
- Quarter 1 accounts for approximately 34% of the total sold quantity for FY2020.

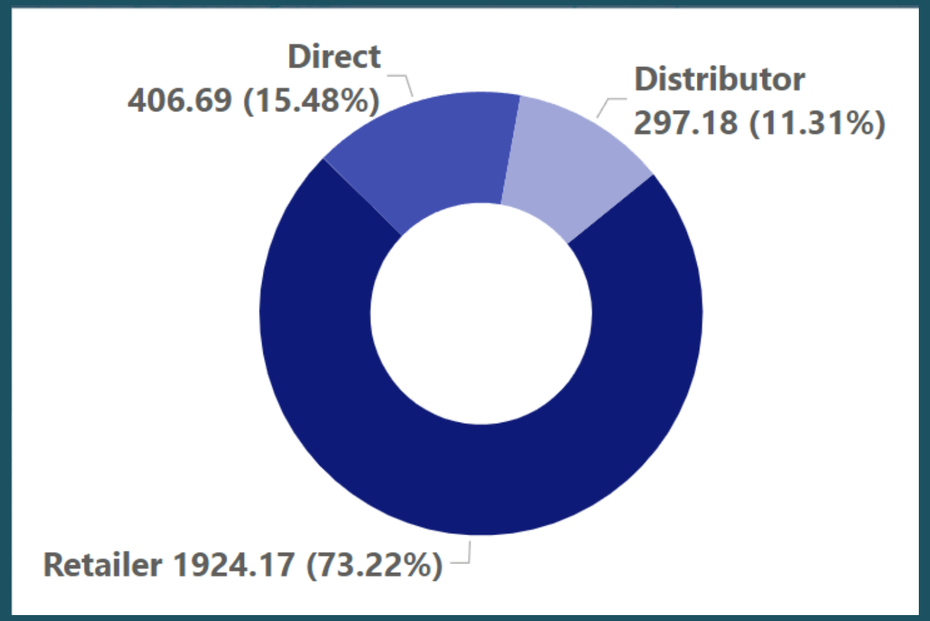
Request -9

-->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
142
143                    ####################    Request -9    ####################
144
145 • WITH cte as(
146         SELECT
147             dc.channel,
148             ROUND(SUM(fg.gross_price*fs.sold_quantity)/1000000, 2) as gross_sales_mln
149         FROM fact_sales_monthly fs
150         JOIN dim_customer dc USING(customer_code)
151         JOIN fact_gross_price fg USING(product_code)
152         WHERE fs.fiscal_year = 2021
153         GROUP BY dc.channel
154         )
155 SELECT *,
156         CONCAT(ROUND(gross_sales_mln*100/(SUM(gross_sales_mln) OVER()), 2), ' %') as percentage
157         FROM cte;
158
```

FY: 2021

channel	gross_sales_mln	percentage
Direct	406.69	15.48 %
Distributor	297.18	11.31 %
Retailer	1924.17	73.22 %



Insights:

- Retailers helped bring maximum sales to the company with 73.22% as the contribution percentage.
- Distributor makes the least contribution at a percentage of 11.31%.

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

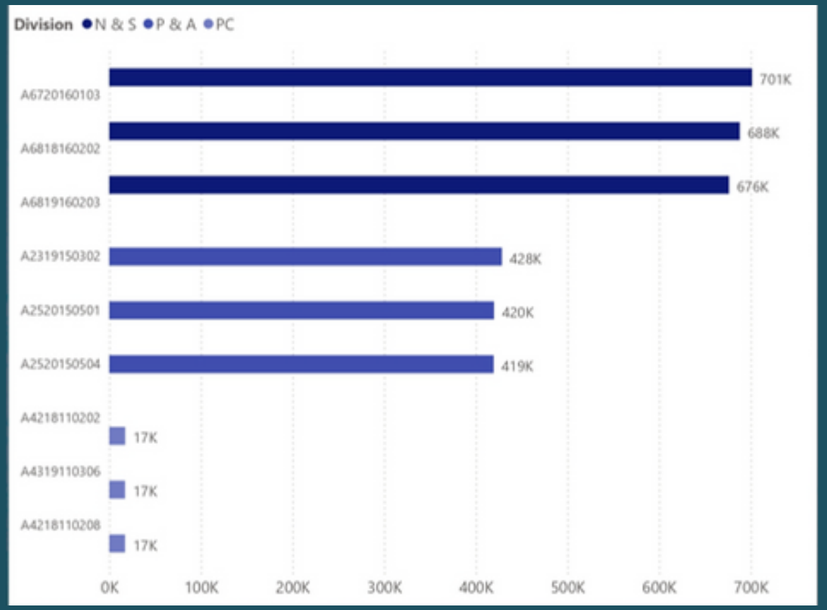
```sql
                    ####################    Request -10    ####################

WITH cte1 as(
        SELECT
             dp.division, dp.product_code, dp.product,
             SUM(fs.sold_quantity) as total_sold_quantity
        FROM fact_sales_monthly fs
        JOIN dim_product dp USING(product_code)
        WHERE fs.fiscal_year = 2021
        GROUP BY dp.product_code, dp.product, dp.division
        ),
cte2 as(
         SELECT *,
            ROW_NUMBER()
            OVER(partition by division ORDER BY total_sold_quantity DESC) as rank_order
         FROM cte1
        )

SELECT * from cte2 where rank_order <= 3;
```

FY: 2021

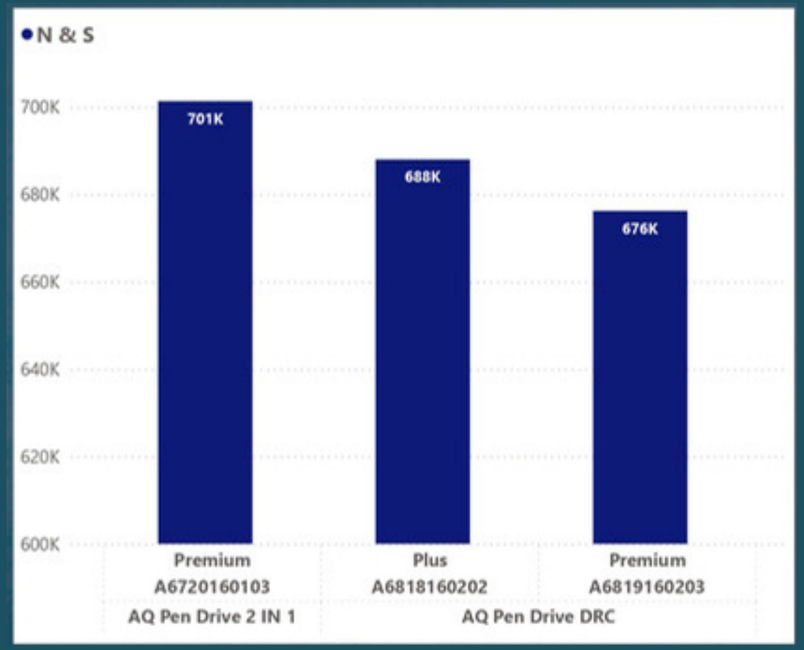
division	product_code	product	total_sold_quantity	rank_order
N & S	A6720160103	AQ Pen Drive 2 IN 1	701373	1
N & S	A6818160202	AQ Pen Drive DRC	688003	2
N & S	A6819160203	AQ Pen Drive DRC	676245	3
P & A	A2319150302	AQ Gamers Ms	428498	1
P & A	A2520150501	AQ Maxima Ms	419865	2
P & A	A2520150504	AQ Maxima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



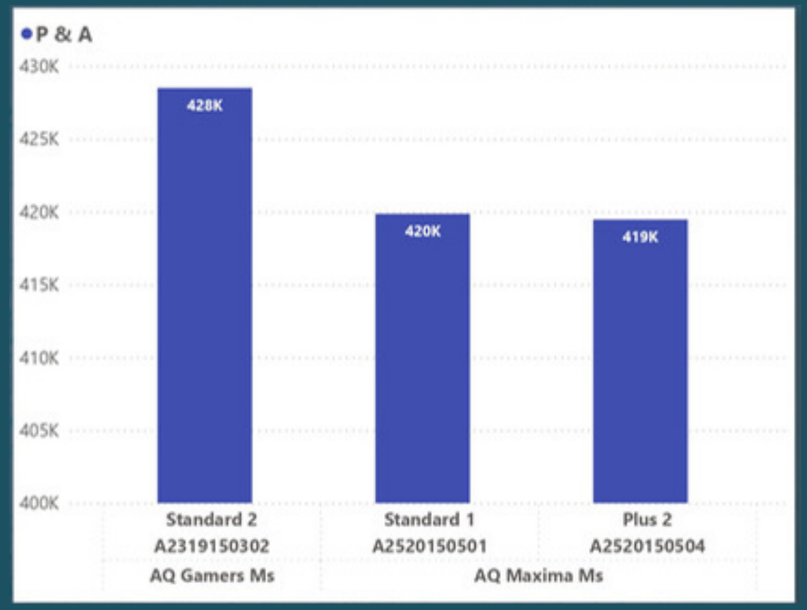
Insights:

- Total sold quantity of products in the 3 divisions is decreasing from N&S, P&A to PC.
- Even though Networking and Storage has least unique products as seen above, but their sale is far greater than other segments. It indicates that N&S has good contribution in overall revenue of company.

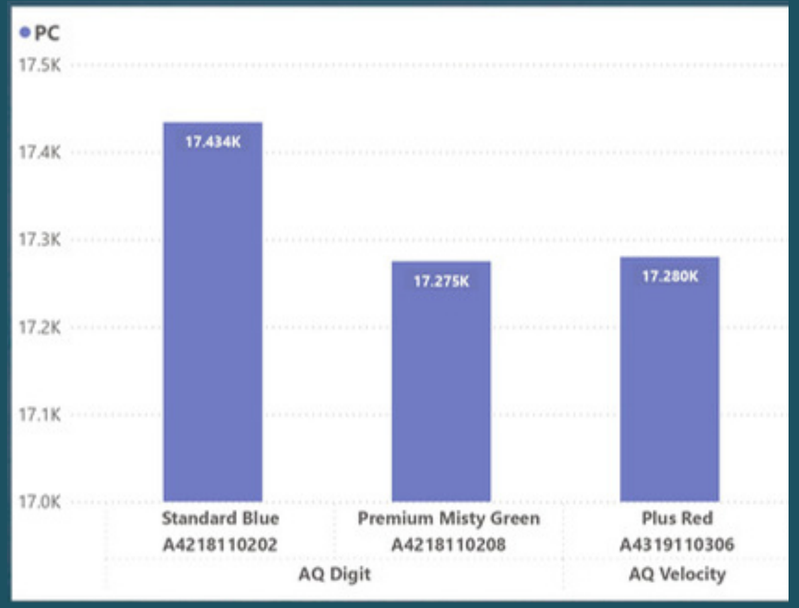
Decrease in the no. of quantities of product sold in different divisions.



N & S



P & A



PC

Thank You
for your valuable time