# Low Level Design (LLD)

## Consumer Goods Ad-Hoc Insights

Revision Number: 1.6 Last date of revision: 23/03/2023

**Pranit Patil** 

## **Document Version Control**

Date Issued	Version	Description	Author
06/02/2023	1.0	Abstract, Introduction	Pranit Patil
08/02/2023	1.1	Problem Statement	Pranit Patil
12/02/2023	1.2	Dataset Information	Pranit Patil
16/02/2023	1.3	Architecture	Pranit Patil
18/02/2023	1.4	Architecture Description	Pranit Patil
28/02/2023	1.5	Deployment	Pranit Patil
13/03/2023	1.6	Final Revision	Pranit Patil

## **Contents**

D	ocı	ımer	nt Version Control	2
Al	osti	ract.		4
1	ļ	Intro	duction	5
	1.1		Why this Low-Level Design Document?	
	1.2		Scope	
	1.3		Problem Statement and Task	
2	-	Tech	nnical specifications	6
	2.1	1	Dataset	6
3	ļ	Prop	oosed Solution	9
	3.1	1 .	Architecture	9
4	ç	SQL	Queries	9
5	5 Reporting			.17
6	6 Deployment			
7	7 Key performance indicators (KPI)			

## **Abstract**

Atliq Hardware's is one of the leading computer hardware producers in India and well expanded in other countries too. However, the management wanted to get enough insights such as understanding changing consumer preferences, focusing on sustainability, building brand loyalty, customer behavior, preferences, and trends. By using data to inform decision-making across functions such as product development, marketing, and supply chain management, companies can better understand their customers and create more effective strategies to make quick and smart data-informed decisions.

#### 1 Introduction

#### 1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the heart disease diagnostic analysis. LLD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

The LLD will be focusing on the below objectives:

- Problem Understanding.
- Data Acquisition.
- Data Pre-Processing and Exploratory Analysis.
- Dashboard report for important activities.

#### 1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step by step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

#### 1.3 Problem Statement and Task

Atliq Hardware's is one of the leading computer hardware producers in India and well expanded in other countries too. However, the management wanted to get enough insights such as understanding changing consumer preferences, focusing on sustainability, building brand loyalty, customer behavior, preferences, and trends. By using data to inform decision-making across functions such as product development, marketing, and supply chain management, companies can better understand their customers and create more effective strategies to make quick and smart data-informed decisions.

#### Task:

- 1. Check 'ad-hoc-requests.pdf' there are 10 ad hoc requests for which the business needs insights.
- 2. You need to run a SQL query to answer these requests.
- 3. The target audience of this dashboard is top-level management hence you need to create a presentation to show the insights.

## 2 Technical specifications

#### 2.1 Dataset

Database Link: <u>Consumer-Goods-Ad\_Hoc-Project/atliq\_hardware\_db.zip\_at\_main\_-</u> Pp11112000/Consumer-Goods-Ad\_Hoc-Project (github.com)

Below provided a comprehensive overview of the tables found in the 'gdb023' (atliq\_hardware\_db) database (Link is Provided). It includes information for six main tables:

- 1. dim customer: contains customer-related data.
- 2. dim\_product: contains product-related data.
- 3. *fact\_gross\_price:* contains gross price information for each product.
- 4. *fact\_manufacturing\_cost:* contains the cost incurred in the production of each product.
- 5. *fact\_pre\_invoice\_deductions*: contains pre-invoice deductions information for each product.
- 6. fact\_sales\_monthly: contains monthly sales data for each product.

#### Column Description for dim\_customer table:

- 1. *customer\_code:* The 'customer\_code' column features unique identification codes for every customer in the dataset. These codes can be used to track a customer's sales history, demographic information, and other relevant details. For example, the codes could look like '70002017', '90005160', and '80007195' respectively.
- 2. *customer:* The 'customer' column lists the names of customers, for example 'Atliq Exclusive', 'Flipkart', and 'Surface Stores' etc.
- 3. *platform:* The 'platform' column identifies the means by which a company's products or services are sold. "Brick & Mortar" represents the physical store/location, and "E-Commerce" represents online platforms.
- 4. *channel:* The 'channel' column reflects the distribution methods used to sell a product. These methods include "Retailers", "Direct", and "Distributors". Retailers refer to physical or online stores that sell products to consumers.

Direct sales refer to sales made directly to consumers through a company's website or other direct means, and distributors refer to intermediaries or middlemen between the manufacturer and retailer or end consumers.

- 5. *market:* The 'market' column lists the countries in which the customer is located.
- 6. **region:** The 'region' column categorizes countries according to their geographic location, including "APAC" (Asia Pacific), "EU" (Europe), "NA" (North America), and "LATAM" (Latin America).
- 7. **sub\_zone:** "The 'sub\_zone' column further breaks down the regions into sub-regions, such as "India", "ROA" (Rest of Asia), "ANZ" (Australia and New Zealand), "SE" (Southeast Asia), "NE" (Northeast Asia), "NA" (North America), and "LATAM" (Latin America)."

#### • Column Description for dim\_product table:

- 1. **product\_code:** The 'product\_code' column features unique identification codes for each product, serving as a means to track and distinguish individual products within a database or system.
- 2. **division:** The 'division' column categorizes products into groups such as "P & A" (Peripherals and Accessories), "N & S" (Network and Storage) and "PC" (Personal Computer).
- 3. **segment:** The 'segment' column categorizes products further within the division, such as "Peripherals" (keyboard, mouse, monitor, etc.), "Accessories" (cases, cooling solutions, power supplies), "Notebook" (laptops), "Desktop" (desktops, all-in-one PCs, etc), "Storage" (hard disks, SSDs, external storage), and "Networking" (routers, switches, modems, etc.).
- 4. *category:* The 'category' column classifies products into specific subcategories within the segment.
- 5. *product:* The 'product' column lists the names of individual products, corresponding to the unique identification codes found in the 'product\_code' column.
- 6. *variant:* The "variant" column classifies products according to their features, prices, and other characteristics. The column includes variants such as "Standard", "Plus", "Premium" that represent different versions of the same product.

#### Column Description for fact\_gross\_price table:

- 1. **product\_code:** The 'product\_code' column features unique identification codes for each product.
- 2. *fiscal\_year:* The 'fiscal\_year' column contains the fiscal period in which the product sale was recorded. A fiscal year is a 12-month period that is used for accounting purposes and often differs from the calendar year. For Atliq Hardware, the fiscal year starts in September. The data available in this column covers the fiscal years 2020 and 2021.
- 3. *gross\_price:* The 'gross\_price' column holds the initial price of a product, prior to any reductions or taxes. It is the original selling price of the product.

#### Column Description for fact\_manufacturing\_cost:

- 1. **product\_code:** The 'product\_code' column features unique identification codes for each product
- 2. **cost\_year:** The "cost\_year" column contains the fiscal year in which the product was manufactured.
- 3. *manufacturing\_cost:* The "manufacturing\_cost" column contains the total cost incurred for the production of a product. This cost includes direct costs like raw materials, labor, and overhead expenses that are directly associated with the production process.

#### • Column Description for fact\_pre\_invoice\_deductions:

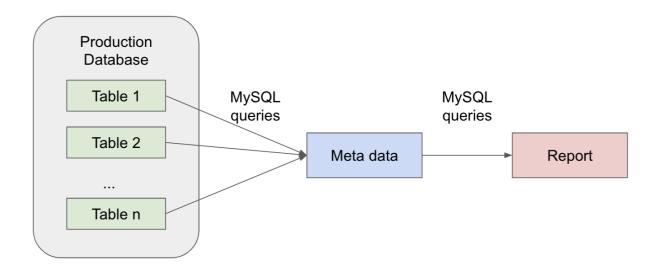
- 1. *customer\_code:* The 'customer\_code' column features unique identification codes for every customer in the dataset. These codes can be used to track a customer's sales history, demographic information, and other relevant details. For example, the codes could look like '70002017', '90005160', and '80007195' respectively.
- 2. *fiscal\_year:* The "fiscal\_year" column holds the fiscal period when the sale of a product occurred.
- 3. *pre\_invoice\_discount\_pct:* The "pre\_invoice\_discount\_pct" column contains the percentage of pre-invoice deductions for each product. Pre-invoice deductions are discounts that are applied to the gross price of a product before the invoice is generated, and typically applied to large orders or long-term contracts.

#### Column Description for fact sales monthly:

- 1. **date:** The "date" column contains the date when the sale of a product was made, in a monthly format for 2020 and 2021 fiscal years. This information can be used to understand the sales performance of products over time.
- 2. **product\_code:** The "product\_code" column contains a unique identification code for each product. This code is used to track and differentiate individual products within a database or system.
- 3. *customer\_code:* The 'customer\_code' column features unique identification codes for every customer in the dataset. These codes can be used to track a customer's sales history, demographic information, and other relevant details. For example, the codes could look like '70002017', '90005160', and '80007195' respectively.
- 4. **sold\_quantity:** The "sold\_quantity" column contains the number of units of a product that were sold. This information can be used to understand the sales volume of products and to compare the sales volume of different products or variants.
- 5. *fiscal\_year:* The "fiscal\_year" column holds the fiscal period when the sale of a product occurred.

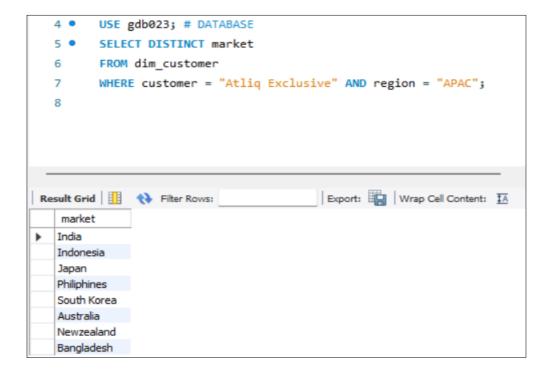
## **3 Proposed Solution**

#### 3.1 Architecture



### **4 SQL Queries**

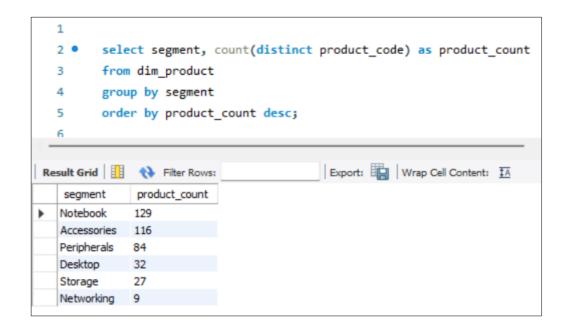
1. Provide the list of markets in which customer "Atliq Exclusive" operates its business in the APAC region.



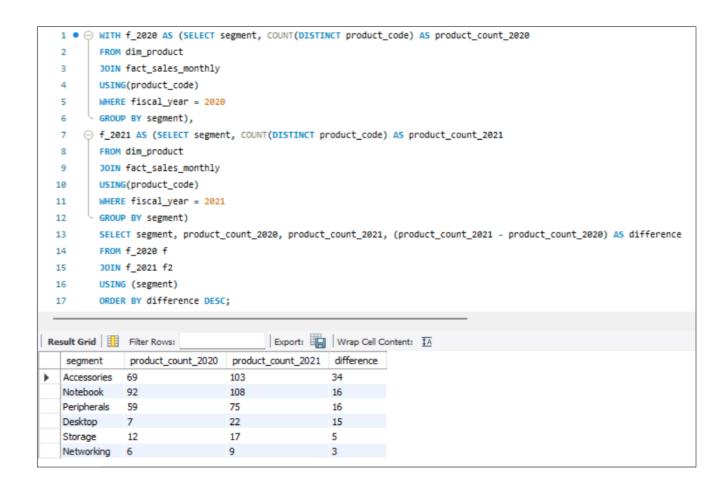
2. What is the percentage of unique product increase in 2021 vs. 2020? The final output contains below fields: unique\_products\_2020, unique\_products\_2021, percentage\_chg.

```
USE gdb023
  2 □ with up_2020 as (select count(distinct product_code) as unique_products_2020
                         from fact_sales_monthly
  4
                         where fiscal_year=2020 ),
  5
              up_2021 as (select count(distinct product_code) as unique_products_2021
                         from fact sales monthly
  6
  7
                         where fiscal year=2021),
              percent_chng as (select (((unique_products_2021 - unique_products_2020) / unique_products_2020)*100) as percentage_chng
  8
  9
                         from up_2020,up_2021)
 10
        select *
 11
         from up_2020, up_2021, percent_chng;
Result Grid Filter Rows:
                                     Export: Wrap Cell Content: TA
   unique_products_2020 unique_products_2021 percentage_chng
245
                                        36.3265
```

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



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.

```
SELECT product code, product, manufacturing cost
 2
      FROM dim product
      JOIN fact manufacturing cost
 3
 4
      USING (product code)
 5
   6
      (SELECT MAX(manufacturing_cost) FROM fact_manufacturing_cost),
      (SELECT Min(manufacturing cost) FROM fact manufacturing cost)
 7
 8
Export: Wrap Cell Content: IA
  product_code product
                       manufacturing_cost
 A2118150101 AQ Master wired x1 Ms
                       0.8920
 A6120110206 AQ HOME Allin1 Gen 2 240.5364
```

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 below fields: customer code, customer, average discount percentage.

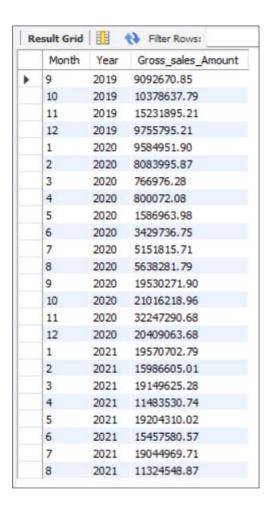
```
1 \bullet \ominus \mathsf{WITH} \mathsf{\ cte1\ AS\ (SELECT\ customer\_code,\ customer,\ \mathsf{ROUND}(\mathsf{AVG}(\mathsf{pre\_invoice\_discount\_pct}) * 100)} \mathsf{\ AS\ average\_discount\_percentage}
       FROM dim_customer c
 2
 3
       JOIN fact_pre_invoice_deductions d
      USING (customer_code)
 5
      WHERE d.fiscal_year = 2021 AND c.market = "India"
      GROUP BY (customer_code)
 7 ORDER BY average discount percentage DESC LIMIT 5)
 8
     SELECT customer_code, customer, CONCAT(average_discount_percentage, "%") AS average_discount_percentage
 9
10
     FROM cte1;
Result Grid Filter Rows:
                                    Export: Wrap Cell Content: IA
  customer_code customer average_discount_percentage
  90002009
               Flipkart
                       31%
  90002006 Viveks 30%
  90002002
               Croma
                       30%
  90002003 Ezone 30%
              Amazon 29%
  90002016
```

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 MONTH(date) AS Month,
      YEAR(date) AS Year,
 2
      SUM(ROUND((g.gross_price * s.sold_quantity),2)) AS Gross_sales_Amount
 3
4
      FROM fact_sales_monthly s
 5
      JOIN dim customer c
6
      ON s.customer_code = c.customer_code
7
      JOIN fact gross price g
      ON g.product_code = s.product_code
9
      WHERE customer = 'Atliq Exclusive'
      GROUP BY Month, Year
10
11
      ORDER BY Year, Month;
```

Result of question 7 is at next page.

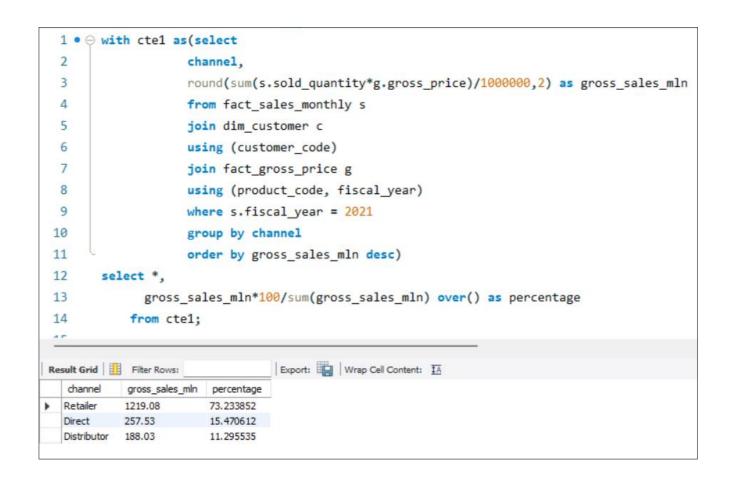


8. In which quarter of 2020, got the maximum total\_sold\_quantity? The final output contains these fields sorted by the total\_sold\_quantity:

Quarter, total\_sold\_quantity.

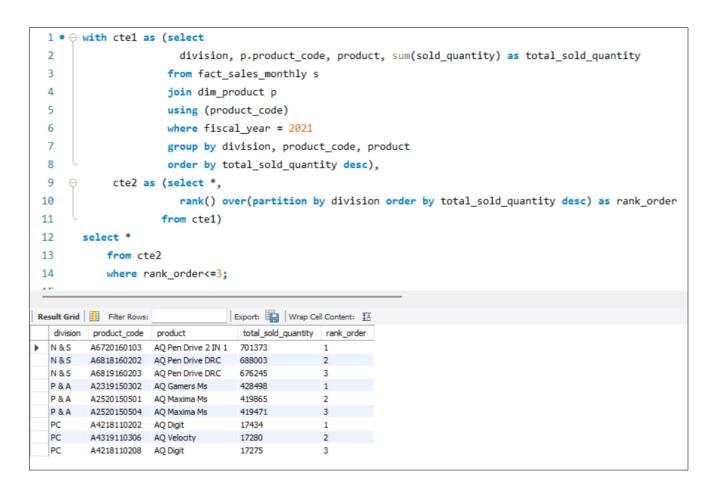
```
1 • ⊖ with quarters as (select *,
  2
                         case
                         when month(date) in (9,10,11) then 'Q1'
  3
  4
                         when month(date) in (12,1,2) then 'Q2'
  5
                         when month(date) in (3,4,5) then 'Q3'
                         when month(date) in (6,7,8) then 'Q4'
  6
  7
                         end as Quarter
                         from fact_sales_monthly
  8
  9
                         where fiscal year = 2020)
10
        select Quarter, sum(sold_quantity) as total_sold_quantity
11
        from quarters
12
        group by Quarter
13
        order by total_sold_quantity desc;
Result Grid Filter Rows:
                                 Export: Wrap Cell Content: IA
  Quarter total_sold_quantity
          7005619
 Q1
  Q2
         6649642
          5042541
  Q4
         2075087
  Q3
```

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, product, total\_sold\_quantity, rank\_order.



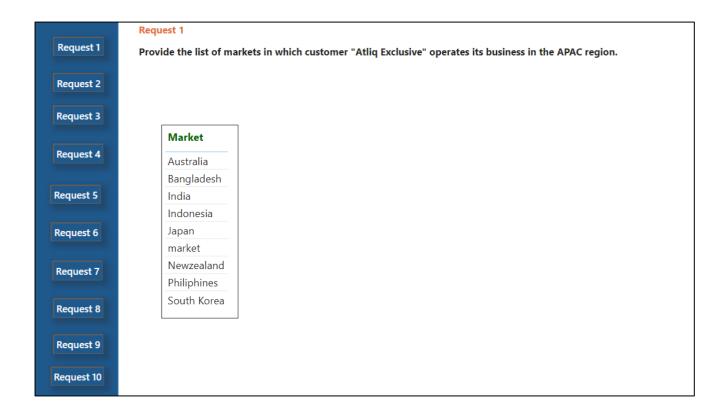
## 5 Reporting

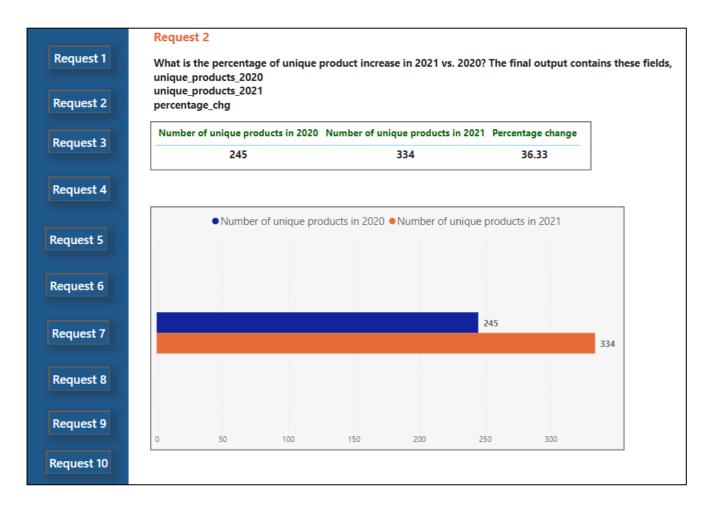
Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self explanatory report because your model will be used by many stakeholders who are not from technical background.

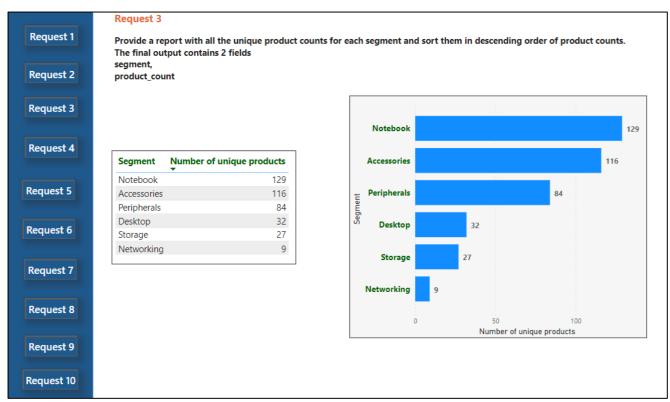
- a) High Level Design Document (HLD)
- b) Low Level Design Document (LLD)
- c) Architecture
- d) Wireframe
- e) Detailed Project Report
- f) Power Point Presentation

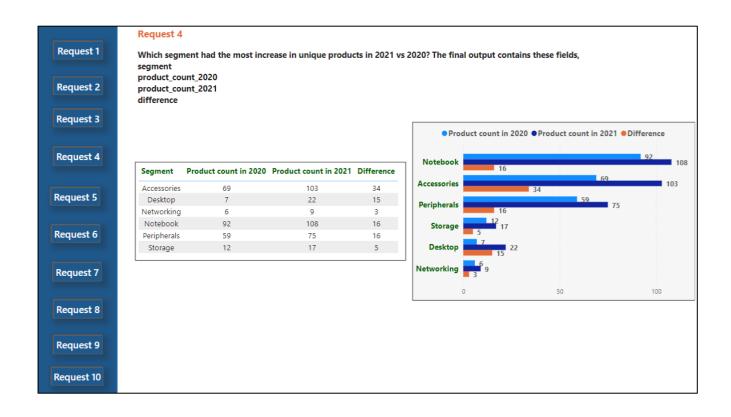
## 6 Deployment

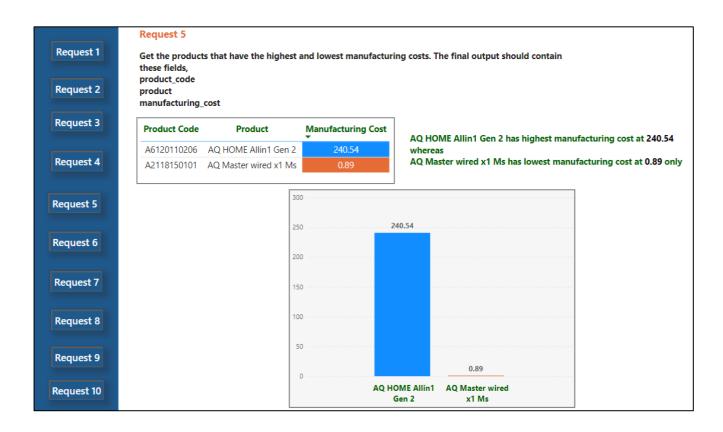
We created a Power BI Dashboard and published it on Power BI Service.

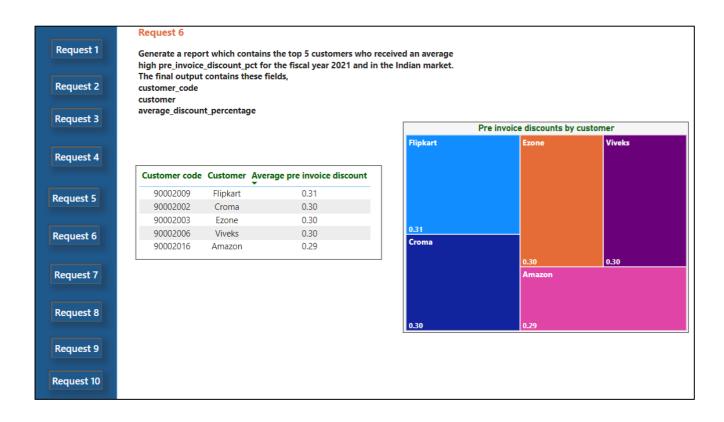


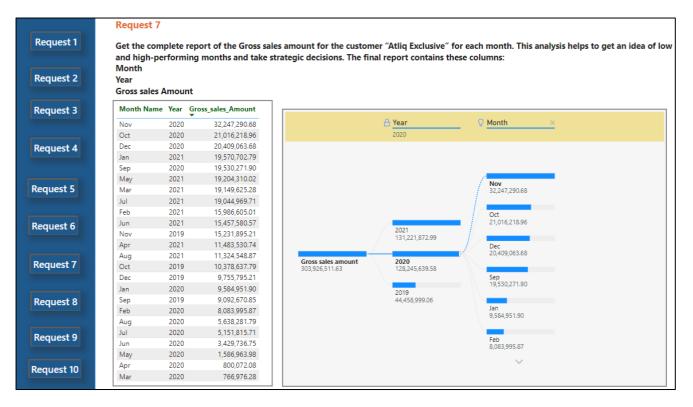


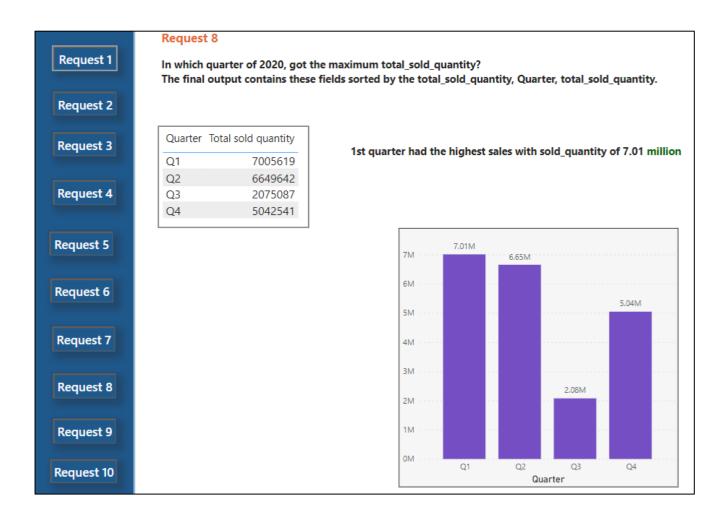


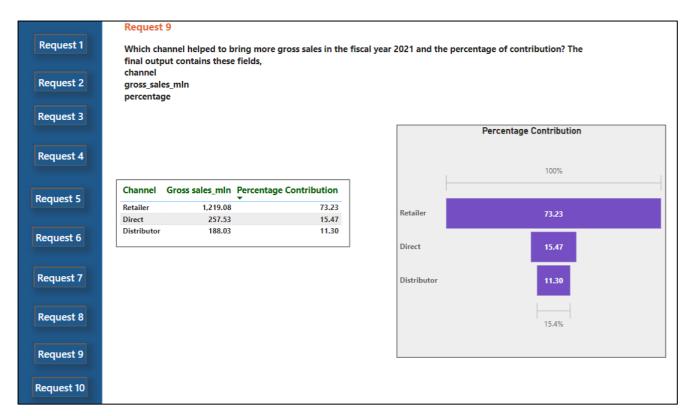


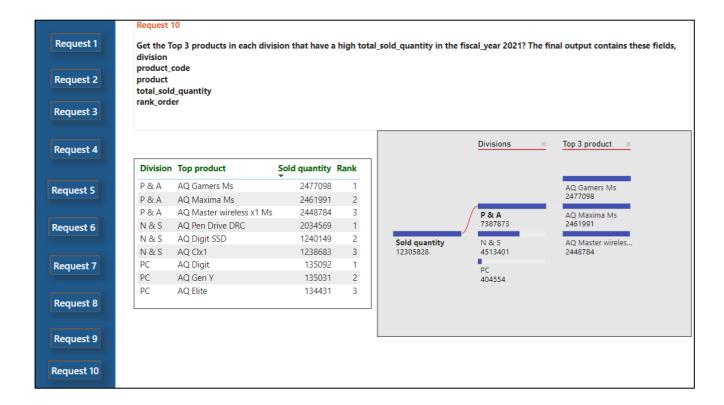












## 7 Key performance indicators (KPI)

Dashboard are implemented to display and indicate certain KPIs and relevant indicators. As and when, the system starts to capture the historical/periodic data for a user, the dashboard will be included to display charts over time with progress on various indicators or factors.

Key indicators displaying a summary of the Consumer Goods Ad-Hoc Project and its relationship with different metrics:

- 1. Market Distribution of Atliq Exclusive in the APAC Region.
- 2. Percentage increase in Unique Products (2020 VS 2021).
- 3. Segment Wise Product Count.
- 4. Product Count 2020 & 2021 by Segment & Difference by Segment.
- 5. Highest and Lowest Costing Products.
- 6. Top 5 Customers with High Average Discount.
- 7. Monthly Gross Sales Amount for Fiscal Year 2020 and 2021.
- 8. Total Sold Quantity Per Quarter of 2020.
- 9. Gross Sales and Percentage Contribution through each Channel.
- 10. Total Sold Quantity by Division, Product, Rank Order.