Capstone Project Report
-
Analysis and Visualization for Sourcing Request and Quotation Data for Sourcy

Group 1
Group Leader: April Liu
Members: Leah Yam, Morris Chan,
Venus Mok, Joyce Lo

Project Objectives

Sourcy aims to revolutionize global B2B sourcing and international trade by offering customers data-driven insights to enhance their procurement strategies. This involves detailed analysis and visualization of sourcing requests and quotations from Sourcy's Postgres database to:

- Identify trends and performance among customers, products, and suppliers.
    - GMV performance, including the Potential & Gained GMV, total cost and Gross Profit;
    - Product Categories and SKUs performance, including the Quoted & Purchased Categories;
    - Sourcing Request (SR) including the quantities of SR, GMV per SR, average initial quotation per SR, and No. of Product, Categories & SKUs per SR.

- Generate actionable insights for identifying cross-sell and upsell opportunities.
    - Sourcy has created the Data analytics (DA) report for supplier's insights usage.
        - generate the dashboard for providing the quantities of DA requests, DA reports per DA, and Category types per report

Dataset Description

The organization under consideration maintains a comprehensive database that serves as the foundation for various business operations. This report provides a detailed description of the dataset contained within the database.

The database under consideration comprises a total of 45 tables, which store a diverse range of information crucial to the organization's operations. The organization utilizes DBeaver as the primary platform for database management and development activities.

The data within the database spans from May 2024 onwards, providing a comprehensive record of the organization's activities and transactions during this period.

The tables within the database contain a wide variety of information, including:

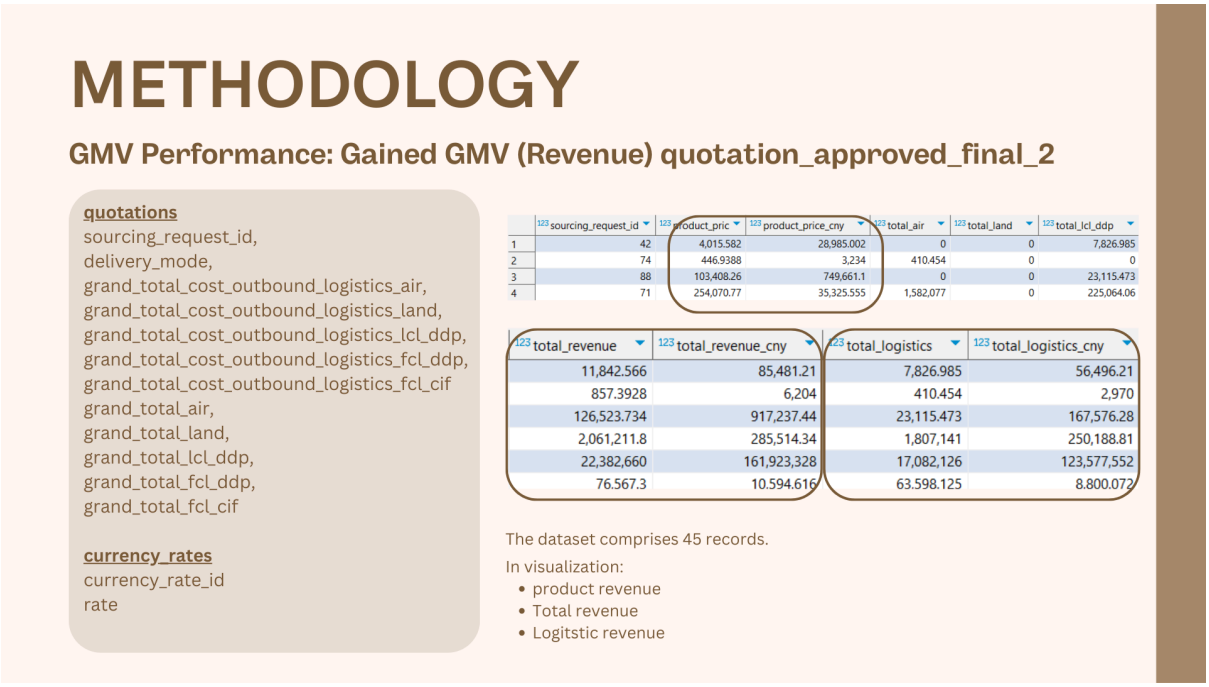
- Customer basic information with foreign key
- Countries, regions, and cities
- Sourcing requests, items, and events
- Quotation including items, and events
- Product specifications
- Shipping items, types, and costs
- Supplier basic information
- Customer-requested data analysis reports

To prevent the accidental modification of the original data, the organization has employed SQL scripts within the 'test_dwh' schema. This approach ensures the integrity and

preservation of the primary data sources while enabling the team to safely conduct testing and development activities.

Dataset characteristics - Methodology& Findings

- **GMV Performance**



- Methodology
  First, we use SQL syntax to extract all "approved" data from our existing Sourcing quotation data. These "approved" data represent orders where the goods have been delivered to the customer and the customer has completed payment. For these "approved" quotes, we calculate product revenue, total revenue, and logistics revenue based on the client's chosen delivery mode. We then convert the total amount from USD to CNY (Chinese Yuan) and create a new dataset called "quotation_approved_final_2" to store this clean data.
  The "quotation_approved_final_2" contains 45 records and includes the following field: sourcing_request_id, product_price, product_price_cny, total_air, total_land, total_lcl_ddp, total_revenue, total_revenue_cny, total_logistics, total_logistics_cny.
  The visualization will show the following metrics: Gain GMV (revenue).

METHODOLOGY

GMV Performance: Gained GMV (Revenue)

quotations
sourcing_request_id,
delivery_mode,
grand_total_cost_outbound_logistics_air,
grand_total_cost_outbound_logistics_land,
grand_total_cost_outbound_logistics_lcl_ddp,
grand_total_cost_outbound_logistics_fcl_ddp,
grand_total_cost_outbound_logistics_fcl_cif
grand_total_air,
grand_total_land,
grand_total_lcl_ddp,
grand_total_fcl_ddp,
grand_total_fcl_cif

currency_rates
currency_rate_id
rate

sourcing_request_id	SUM(total_revenue_cny)
51	162M
49	10.4M
66	4.77M
28	3.62M
77	2.05M
Totals	195M

SUM(product_price_cny)	SUM(total_logistics_cny)
38.3M	124M
4.55M	5.85M
2.17M	2.59M
502k	3.12M
1.73M	326k
54.3M	141M

- This dashboard reveals that the total revenue from scoring is 195 million CNY. Breaking down this data further, we see the total product revenue is 54.3 million CNY and the total logistics revenue is 14 million CNY.

METHODOLOGY

GMV Performance: aggregated_sourcing_request_costs_l

quotations
sourcing_request_id,
delivery_mode,
status

sourcing_request_items
sourcing_request_id,
total_price_exfactory,
total_cost_outbound_logistics_air,
total_cost_outbound_logistics_lcl_ddp,
total_cost_outbound_logistics_fcl_ddp

sourcing_request_id	total_price_exfactory	total_cost_outbound_logistics_air	total_cost_outbound_logistics_lcl_ddp	total_cost_outbound_logistics_fcl_ddp
56	110,768.9375	891,769.8125	135,132.859375	[NULL]
11	961,875	442,680	74,373.609375	[NULL]
70	28,700	1,669,150	273,477.5	[NULL]
44	98,518	46,884.58203125	5,813.78222656	[NULL]
49	3,047,271	3,485,448.25	551,205.875	35,973,448
28	328,761.5	511,400.15625	81,406.8046875	[NULL]
42	111,650	570,000	3,951,360.25	[NULL]
45	8,343.75	5,032	2,850	[NULL]

total_cost_air	total_cost_lcl_ddp	total_cost_fcl_ddp	delivery_mode	total_costs
982,597.0625	225,960.140625	[NULL]	{AIR,LCL_DDP}	1,137,671.625
1,088,130	719,823.8125	[NULL]	{AIR,LCL_DDP}	1,478,928.625
1,698,205	302,532.5	[NULL]	{AIR,LCL_DDP}	1,971,327.5
105,234.5859375	73,549.78125	[NULL]	{AIR,LCL_DDP}	151,216.359375
6,540,010	3,605,767.75	39,028,000	{AIR,LCL_DDP}	7,083,925

The dataset comprises 45 records.
In visualization:

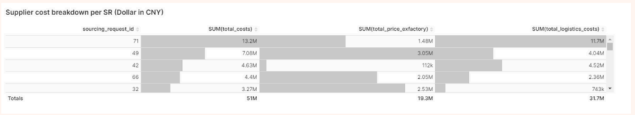
- Total logistics cost: $141M
- Item costs: $54.3M
- Overall total costs: $195M

METHODOLOGY

GMV Performance: aggregated_sourcing_request_costs_l

quotations
sourcing_request_id,
delivery_mode,
status

sourcing_request_items
sourcing_request_id,
total_price_exfactory,
total_cost_outbound_logistics_air,
total_cost_outbound_logistics_lcl_ddp,
total_cost_outbound_logistics_fcl_ddp



The dataset comprises 45 records.
In visualization:

- Total logistics cost: $141M
- Item costs: $54.3M
- Overall total costs: $195M

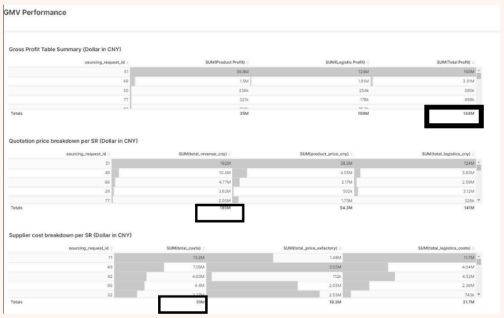
GMV performance (cost): We inner join “quotations” & “sourcing_request_items” table, taking sourcingn request as the key to find out the exfactory and logistic cost per each sourcing request, and then we have found out the breakdown of cost (total logistic & item costs)

METHODOLOGY

GMV Performance: Gross Profit (Revenue-Cost)

quotations
ID, date, type, status

sourcing_request_items
sourcing_request_id,
total_price_exfactory,
total_cost_outbound_logistics_air,
total_cost_outbound_logistics_lcl_ddp,
total_cost_outbound_logistics_fcl_ddp



The dataset comprises 45 records.
In visualization:

- Items profit: $35M
- Logistics Profit: $100M
- Total Profit: $144M

GMV performance (cost): To calculate the profit, so we simply just deduct the cost from revenue to get the results, and regarding to our calculator, the results has reveal as above.

-

- **Product Categories & SKUs:**

METHODOLOGY

Product Categories & SKUs: test_skusl

quotations
ID, date, type, status

quotation_items
ID

sourcing_request_items
quantity, customerID, productID

products
scrape product ID

customers_readonly
countries, subcategory

taxonomies
category level 1,2,3

quotation_id	quotation_year	quotation_month	quotation_date	quotation_type	status	sourcing_request_item_id	quantity
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	363	1
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	364	2
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	365	250
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	366	1
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	367	1
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	368	20

customer_id	country	subcategory	product_id	scrape_product_id	taxonomy_id	category_level_1	category_level_2	category_level_3
116	PH	Hospitality	124	1,356	1,687	Hardware	Plumbing	Plumbing Fixtures
116	PH	Hospitality	125	1,357	1,687	Hardware	Plumbing	Plumbing Fixtures
116	PH	Hospitality	125	1,357	1,687	Hardware	Plumbing	Plumbing Fixtures
116	PH	Hospitality	126	1,358	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holde
116	PH	Hospitality	127	1,359	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holde
116	PH	Hospitality	128	1,360	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holde
116	PH	Hospitality	129	1,361	595	Home & Garden	Decor	Mirrors

The dataset comprises 3504 records.

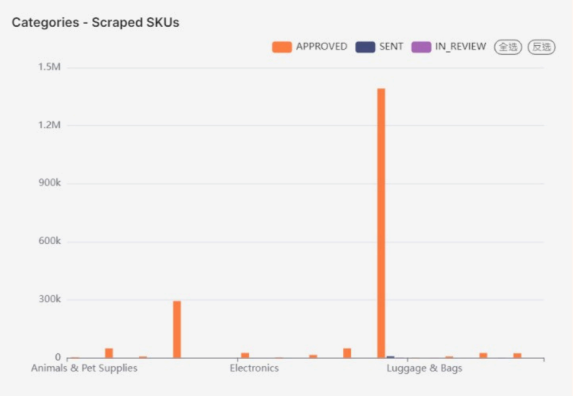
In visualization:

- auto scraping/manual scraping
- customer's business category
- Pinpoint the categories that demonstrate the highest popularity

- For the second objective, to count the value of product and category SKUs, a new table called 'test_sku1' should be created. To achieve this result, six tables have been combined using left and right join functions. The table and column structure should explain the requested items per quotations, quantity, customer details, and categories of the products. This will allow the task to be completed at the required SKUs.

PRODUCT CATEGORIES & SKUS

Quoted & Purchased Scraped Categories SKUs



Quoted Categories (SENT + IN_REVIEW):

1. Home & Garden - 7K
2. Furniture - 50
3. Business & Industrial - 7

Purchased Categories (APPROVED):

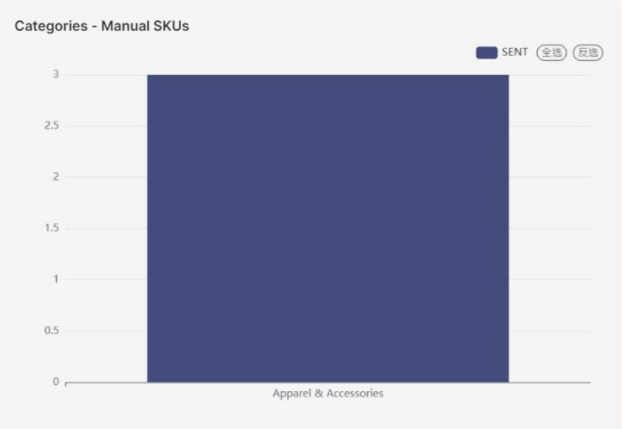
1. Home & Garden - 1.39M
2. Business & Industrial - 0.29M
3. Apparel & Accessories, Health & Beauty - 48K

- In the scraped product data, we have combined the information for both 'quoted' and 'purchased' products. For the quoted items, we looked at the 'sent' and 'in_review' statuses. For the purchased items, we focused on the 'approved' status.

The top three quoted categories are 'Home & Garden' with 7K, 'Furniture' with 50, and 'Business & Industrial' with 7. The top three purchased categories are 'Home & Garden' with 1.39M, 'Business & Industrial' with 0.29M, and 'Apparel & Accessories' and 'Health & Beauty' both with 48K.
Based on these results, the 'Home & Garden' category appears to be the highest requested and purchased. This suggests the company should continue prioritizing and sourcing more high-performing products in this popular segment.

PRODUCT CATEGORIES & SKUS

Quoted & Purchased Manual Categories SKUs



- Only **3 Apparel & Accessories** SKUs were quoted in Manual Categories

- For the manual product data, where the scrape_product_id column has a null value, there are only 3 items that were quoted in the 'Apparel & Accessories' category. This contrasts with the scraped product data, which showed the 'Home & Garden', 'Business & Industrial', and 'Apparel & Accessories' / 'Health & Beauty' categories as the top quoted and purchased segments. The limited manual product data in the 'Apparel & Accessories' category suggests this could be an area with untapped potential that the company may want to further explore and expand upon.

PRODUCT CATEGORIES & SKUS

Total Product Categories SKUs per Month in 2024

April - 25K SKUs (25,508)
May - 8.2M SKUs (8,293,778)
June - 108K SKUs (108,851)

- May has the highest number of SKUs at 8.2M, compared to 108K SKUs in June and 25.5K SKUs in April. This indicates that May is the peak month for product categories and SKUs in 2024 for this business. Understanding this peak in May will be crucial for the business to optimize its inventory management, product planning, and sales strategies to capitalize on the heightened product availability and potential customer demand during that time.

Map - Quantity and Customers

SKU - Group by tables: Categories vs SKUs, customer, region

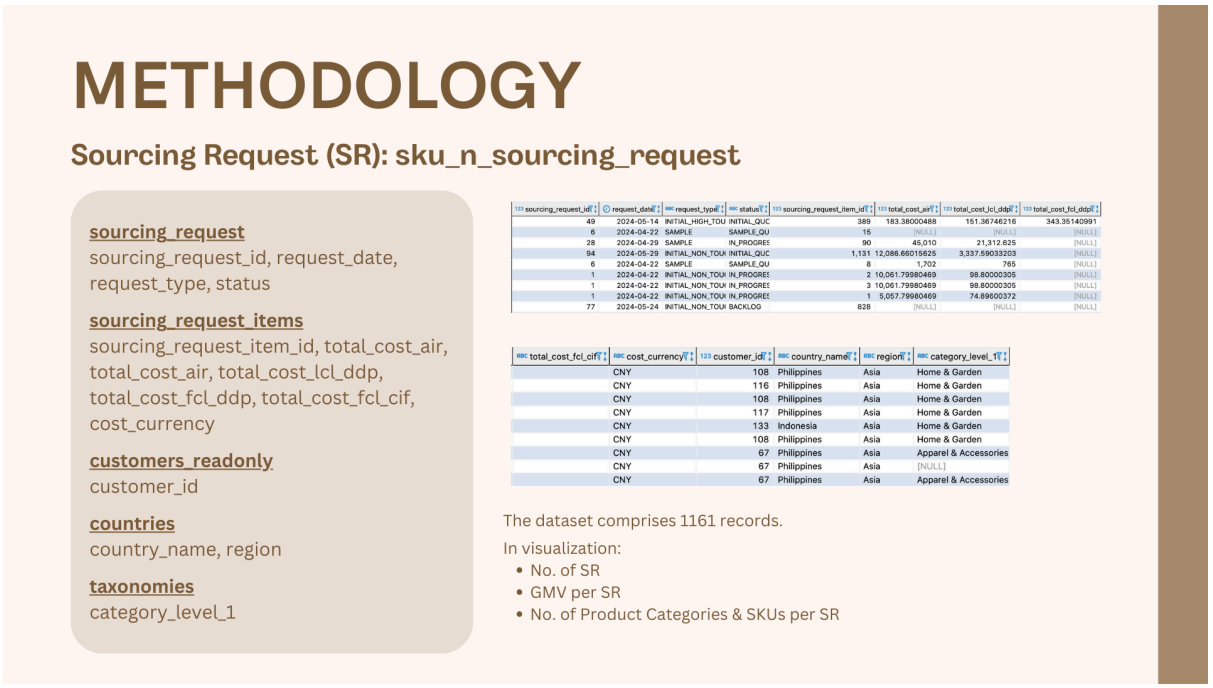
customer_id	category_level_1	region	SUM(quantity)
117	Home & Garden	Asia	6.19M
116	Business & Industrial	Asia	460k
125	Business & Industrial	Asia	300k
128	Home & Garden	Asia	136k
128	Sporting Goods	Asia	114k
116	Home & Garden	Asia	88.3k
122	Health & Beauty	Asia	76.3k
66	Home & Garden	Asia	70k
128	N/A	Asia	67k
128	Electronics	Asia	66.6k
128	Toys & Games	Asia	65k
135	Apparel & Accessories	Asia	60.8k
118	Home & Garden	Europe	51k
111	Business & Industrial	Asia	50k
65	N/A	Asia	43k
Totals			8.43M

- The region with the highest total quantity is Asia, which accounts for the majority of the products across various categories such as ‘Home & Garden’, ‘Business & Industrial’, ‘Health & Beauty’ and ‘Electronics’.

- The top three product categories by total quantity are:
    - Home & Garden: 460k
    - Business & Industrial: 300k
    - Sporting Goods: 114k

- Customer 117 (restaurant business) has the highest total quantity of 619k across multiple product categories, including ‘Home & Garden’, ‘Business & Industrial’, and ‘Sporting Goods’.

- **Souring Request (Methodology)**



Creating a new dataset is called “sku_n_sourcing_request”, and tying up the relative tables, using the SQL syntax to join the data columns. The dataset contains 1161 records and includes the following fields: sourcing_request_id, request_date, request_type, status, sourcing_request_items, total_cost_air, total_cost_lcl_ddp, total_cost_fcl_ddp, total_cost_fcl_cif, cost_currency, customers_readonly, customer_id, country_name, region, category_level_1. The visualization will show the following metrics: No. of SR, GMV per SR, and No. of Product Categories & SKUs per SR.

Sourcing Request (SR) – Total no. of SR

99

Total No of Sourcing Request

According to the table “sourcing request”, there are ninety-nine no. of sourcing requests.

Sourcing Request (SR) - No. of SR per status in Summary

status	COUNT(sourcing_request_id)
INITIAL_QUOTATION_DONE	29
SAMPLE_QUOTATION_DONE	11
BULK_QUOTATION_DONE	2
Totals	42

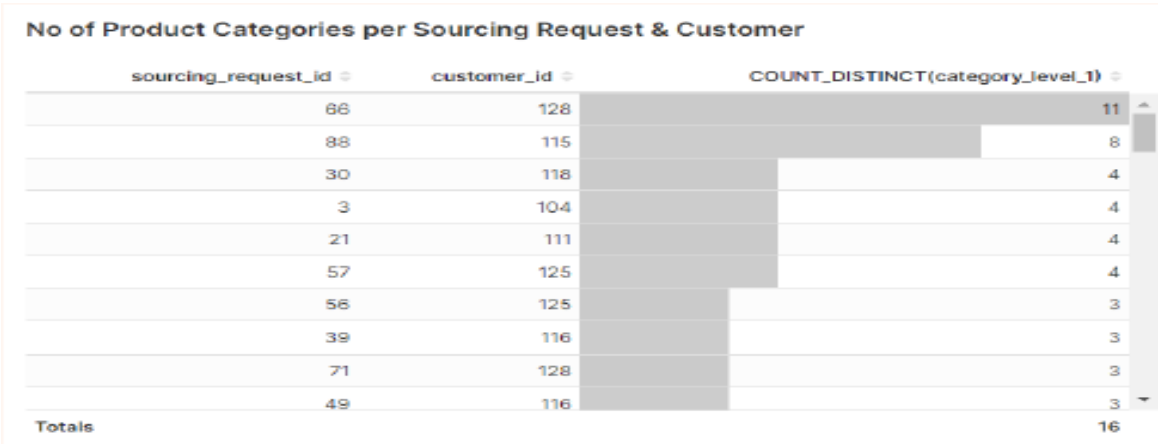
This graph summarizes the number of Sourcing Requests (SR) per status. The dataset contains a total of 42 Sourcing Requests. The breakdown by status is:
INITIAL_QUOTATION_DONE: 29 SRs
SAMPLE_QUOTATION_DONE: 11 SRs
BULK_QUOTATION_DONE: 2 SRs

SOURCING REQUEST

ɔurcing Request (SR) - latest SR with status, timestamp

	Metric	COUNT(sourcing_request_id)			
	status	BULK_QUOTATION_DONE	INITIAL_QUOTATION_DONE	SAMPLE_QUOTATION_DONE	Total (Sum)
request_date					
2024-04-22			1	1	2
2024-04-24			1		1
2024-04-25			1		1
2024-04-26		1			1
2024-04-29				1	1
2024-05-02			2		2
2024-05-03			2		2
2024-05-07			1		1
2024-05-08			1		1
2024-05-13			3		3
2024-05-14			1	2	3
2024-05-15			2	1	3
2024-05-16			3	1	4
2024-05-17			1		1
2024-05-20			2	2	4
2024-05-22		1		1	2
2024-05-24			2		2
2024-05-27			1		1
2024-05-28			2		2
2024-05-29			3	1	4
Total (Sum)		2	29	11	42

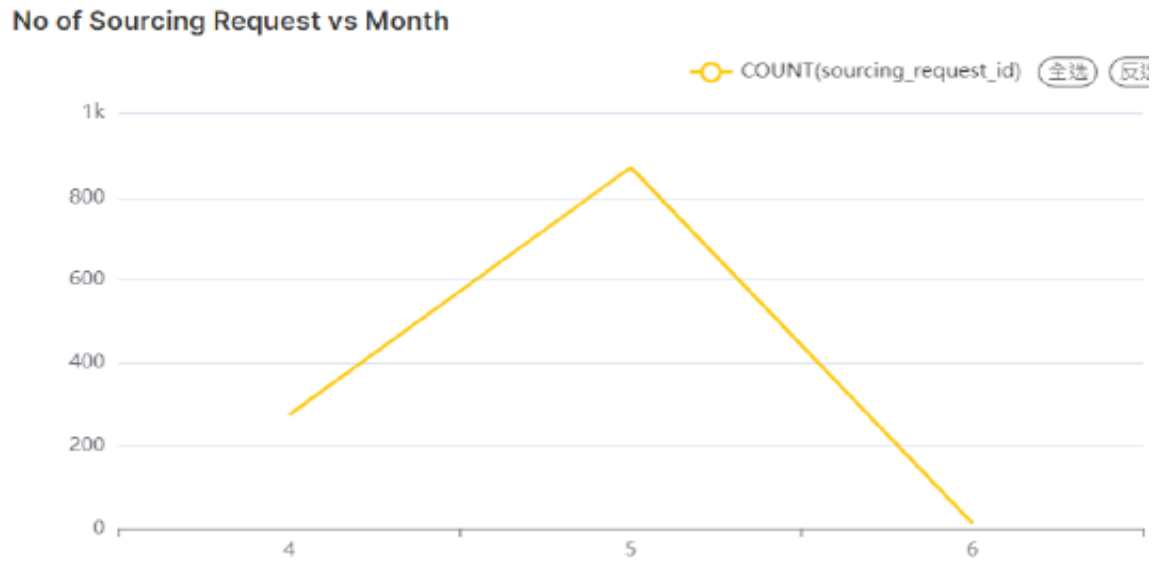
The graph shows a breakdown of Sourcing Requests (SR) by their status and the timestamp (request_date) for the latest SR with each status from April to May. Same as in the last graph, There are a total of 42 Sourcing Requests in the dataset.
The breakdown by status is:
BULK_QUOTATION_DONE: 2 SRs
INITIAL_QUOTATION_DONE: 29 SRs
SAMPLE_QUOTATION_DONE: 11 SRs



No of Product Categories per Sourcing Request & Customer

sourcing_request_id	customer_id	COUNT_DISTINCT(category_level_1)
66	128	11
88	115	8
30	118	4
3	104	4
21	111	4
57	125	4
56	125	3
39	116	3
71	128	3
49	116	3
Totals		16

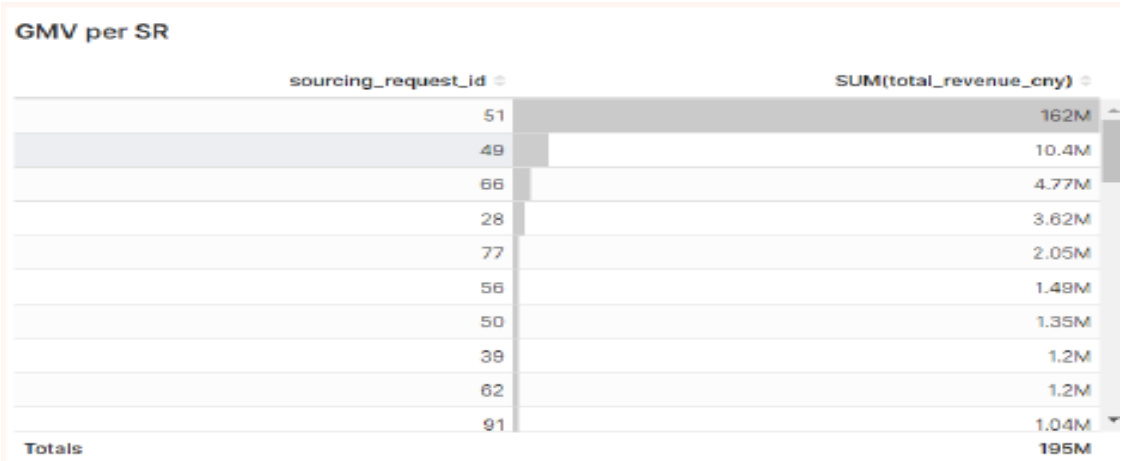
The graph shows a visualization of the number of distinct product categories per Sourcing Request and Customer. For this sample, the total number of distinct product categories across all Sourcing Requests is 16. The customer with ID 128 has Sourcing Requests that cover the highest number of distinct product categories at 11. The customer with ID 115 has Sourcing Requests that cover 8 distinct product categories. The customers with IDs 111, 104, 125, and 118 all have Sourcing Requests that cover 4 distinct product categories each. The customer with ID 57 has Sourcing Requests that cover 3 distinct product categories.



The graph shows a line chart of the number of Sourcing Requests (SR) over time, specifically the number of SRs per month(from April to June). The number of SRs starts at a low level of around 100 in April. It then rises sharply to a peak of around 800 SRs in May of the period shown. After the peak, the number of SRs declines rapidly back down to around 100 by the end of the period.

The graph shows a world map visualization of the geographic distribution of Sourcing Requests.

The majority of Sourcing Requests seem to be coming from the Asia-Pacific region, specifically China and surrounding countries. There are also some Sourcing Requests originating from Southeast Asia, particularly Indonesia and Malaysia.

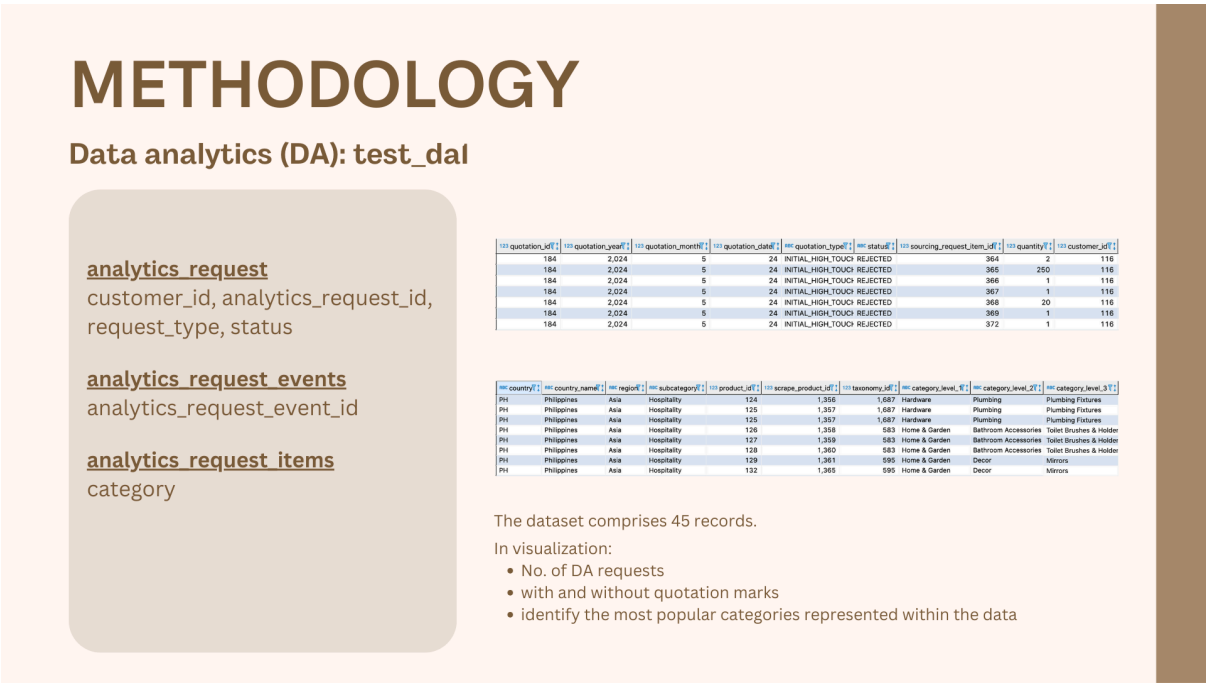


GMV per SR

sourcing_request_id	SUM(total_revenue_cny)
51	162M
49	10.4M
66	4.77M
28	3.62M
77	2.05M
56	1.49M
50	1.35M
39	1.2M
62	1.2M
91	1.04M
Totals	195M

The graph shows a bar chart displaying the Gross Merchandise Value (GMV) per Sourcing Request (SR) for different Sourcing Requests. For this sample, The Sourcing Request with ID 51 has the highest GMV per SR at around $162M. The Sourcing Request with ID 49 has the second-highest GMV per SR at around $10.4M. The Sourcing Request with ID 66 has a GMV per SR of around $4.77M. The Sourcing Request with ID 28 has a GMV per SR of around $3.62M. The GMV per SR values decrease as you go down the chart, with the lowest being around $1.04M.

- **Data Analytics**



METHODOLOGY

Data analytics (DA): test_dal

analytics_request
customer_id, analytics_request_id, request_type, status

analytics_request_events
analytics_request_event_id

analytics_request_items
category

quotation_id	quotation_year	quotation_month	quotation_date	quotation_type	status	sourcing_request_item_id	quantity	customer_id
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	364	2	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	365	250	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	366	1	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	367	1	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	368	20	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	369	1	116
184	2,024	5	24	INITIAL_HIGH_TOUCH	REJECTED	372	1	116

country	country_name	region	subcategory	product_id	scrape_product_id	taxonomy_id	category_level_1	category_level_2	category_level_3
PH	Philippines	Asia	Hospitality	124	1,356	1,687	Hardware	Plumbing	Plumbing Fixtures
PH	Philippines	Asia	Hospitality	125	1,357	1,687	Hardware	Plumbing	Plumbing Fixtures
PH	Philippines	Asia	Hospitality	125	1,357	1,687	Hardware	Plumbing	Plumbing Fixtures
PH	Philippines	Asia	Hospitality	126	1,358	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holder
PH	Philippines	Asia	Hospitality	127	1,359	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holder
PH	Philippines	Asia	Hospitality	128	1,360	583	Home & Garden	Bathroom Accessories	Toilet Brushes & Holder
PH	Philippines	Asia	Hospitality	129	1,361	595	Home & Garden	Decor	Mirrors
PH	Philippines	Asia	Hospitality	132	1,365	595	Home & Garden	Decor	Mirrors

The dataset comprises 45 records.

In visualization:

- No. of DA requests
- with and without quotation marks
- identify the most popular categories represented within the data

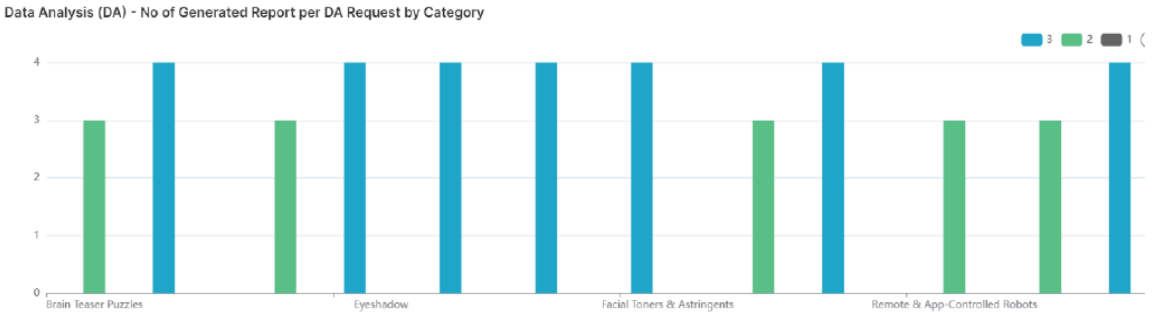
-

Data Analytics (DA)

DA - no of report per customer

customer_id	COUNT(analytics_request_event_id)
130	28
115	15
35	2
Totals	45

- The data shows that there were a total of 3 requests for data analytics during this period. These requests resulted in the generation of 45 reports overall. Further analysis reveals that Customer 130 made the most requests out of all customers. Understanding the specific needs and data requirements of the customers making the highest number of requests can help the company better allocate resources and optimize its data analytics services to meet their needs effectively.



- There were a total of 3 data analytics requests made, which resulted in the generation of 45 reports. The first data analytics request had 2 reports canceled. The second and third data analytics requests covered the same categories: ‘Brain Teaser Puzzle’, ‘Eyeshadow’, ‘Facial Toners & Astringents’, and ‘Remote & App-Controlled Robots’.

  The data analytics requests and reporting appear to be focused on product categories related to beauty, technology, and brain teasers, suggesting these are areas of interest or focus for the business.

  Overall, the graph provides a high-level overview of the data analytics activities, highlighting the number of requests, reports generated, and the specific product categories being analyzed. This information can help the business understand its data analytics needs and optimize its processes accordingly.

DA - no of report per cat

category	COUNT(analytics_request_event_id)
Lip Balms & Moisturizers	4
Facial Day Care	4
Electric Eyelash Curlers	4
Facial Masks	4
Facial Toners & Astringents	4
Eyeshadow	4
Wrinkle & Anti-Aging Devices	4
Remote & App-Controlled Robots	3
Jigsaw Puzzles	3
Toy Building Sets	3
Totals	45

- The data shows a total of 45 reports were generated across 3 data analytics requests. Further analysis reveals that the product categories with the highest number of reports per item were ‘Lip Balms & Moisturizers’, ‘Facial Day Care’, ‘Electric Eyelash Curlers’, ‘Facial Masks’, ‘Facial Toners & Astringents’, ‘Eyeshadow’, and ‘Wrinkle & Anti-Aging Devices’, each with 4 reports per item. This suggests these may be priority categories for the business, warranting deeper data analysis to optimize performance and drive growth in these key product segments.

Challenges and Limitations

- **Time Scarcity**: Limited time to thoroughly study the database and understand its structure and data relationships; Narrow scope for fully utilizing the diverse technical capabilities available for data analysis and visualization.
- **Platform Adoption**: The analysis was conducted in the early phases of adopting and acclimating to DBeaver and Superset, which limited the team's familiarity and proficiency with these tools.
- **Dashboard Capabilities**: The current capabilities of the dashboards are limited in terms of robustness and comprehensiveness, constraining the depth and breadth of the insights that can be derived.
- **Large-scale Data Import**: Importing large datasets into the analysis tools proved to be time-consuming, which impacted the overall efficiency of the exploration and reporting process.
- **Missing Values**: While moving the data from the database to the analysis tools, some values were found to need to be included, requiring additional time and effort to cross-check and validate the data.

Future Work

- **Deeper Analysis of GMV Preference**: Conduct a more in-depth analysis of the Gross Merchandise Value (GMV) performance, exploring factors that drive customer preference and purchasing behavior; Investigate the relationship between product categories, pricing, and customer segmentation to uncover insights into GMV drivers;

Perform statistical analysis and modeling to identify significant variables influencing GMV and understand the underlying dynamics; Leverage advanced analytical techniques, such as regression analysis, clustering, or predictive modeling, to extract more comprehensive and actionable insights.

- **ETL Automation**: Develop reusable ETL (Extract, Transform, Load) scripts to automate the data extraction, transformation, and loading processes; Implement a robust and scalable data pipeline that can handle the growing volume and complexity of the dataset; Ensure the data pipeline is adaptable to changes in the source database schema or reporting requirements; Incorporate data quality checks and validation routines within the ETL process to maintain data integrity and reliability; Explore the use of data orchestration tools or platforms to streamline the end-to-end data processing workflow.

Conclusion

- **Understand the Business Model:** To ensure the success of the project, it is crucial to have a thorough understanding of the client's business model, their goals, and the underlying data and processes that drive their operations.
- **Fulfill All Stakeholder's Expectations:** It is essential to engage with all relevant stakeholders, including the client, end-users, and any other key stakeholders, to clearly understand their expectations and requirements. This will help to align the project deliverables with their needs.
- **Provide Suitable Tools:** Based on the project requirements, it is recommended to utilize tools such as Superset, a powerful and flexible data visualization platform, and PostgreSQL, a robust and reliable relational database management system, to support the data processing and dashboard development needs.
- **Setup a Possible Project Timetable:** Establish a comprehensive project timetable that outlines the key milestones, tasks, and timelines for the project. This will help to ensure the timely delivery of the project and manage stakeholder expectations effectively.
- **Communication is Important:** Regular and effective communication with the client and all stakeholders is crucial throughout the project lifecycle. This will help to address any concerns or issues promptly and maintain a collaborative working relationship.
- **Regular Review and Modification:** Implement a process for regularly reviewing the dashboard and making necessary modifications based on user feedback and changing business requirements. This will ensure that the dashboard remains relevant and meets the evolving needs of the client and end-users.

Reference (e.g. dataset)

- superset
- DBeaver

Distribution of Work

- Dataset movement, cleaning, query scripts - All
- Powerpoint - All
- Final report - All