SCIENCE AND INFORMATION TECHNOLOGY
SINTECH
JOURNAL

SINTECH JOURNAL | ISSN 2598-7305 | E-ISSN 2598-9642
Vol. 7 No 2 – Agustus 2024| https://s.id/sintechjournal
DOI: https://doi.org/10.31598
Publishing: Prahasta Publisher

Business Intelligence Implementation for Groceries Commodity Price Data Analytics

Fathe Hibatulwafi[1], Taufik Asmiyanto[2]

[1,2]Department of Library and Information Science,
Faculty of Humanities, Universitas Indonesia

e-mail: fathe.hibatulwafi@ui.ac.id, tasmiy@ui.ac.id.

Received : July, 2024	Accepted : August, 2024	Published : August, 2024

Abstract

As a startup company specializing in agricultural technology, currently PT XYZ does not have any effective and integrated tools to analyze groceries commodity for purchase and sale data-driven decision making. This research will integrating multiple data sources for accurate analysis, tailoring the business intelligence (BI) system to PT XYZ's specific needs, and developing a robust solution that supports comprehensive decision-making. BI implementation solution will be developed using Pentaho Data Integration & Apache Superset. Applied research used as its method, that contains processes to gather user requirements to gather user needs from each information, design the solution, and dashboard development. The study focused on analyzing data from both internal sources, such as purchase and sales transactions, and external sources, including market price data, using these BI tools to provide comprehensive insights. The developed dashboard contains several sections that allow users to see price recommendation price and get the insight for the last seven days trend, monitor the market position, see team performance, and find the summary about estimated and actual margin comparison. Pentaho successfully handled ETL and data modeling, while Apache Superset enabled straightforward dashboard setup, though chart customization is limited.

Keywords: *Business Intelligence, Decision Making, Data Analysis, Data Visualization, Apache Superset*

Abstrak

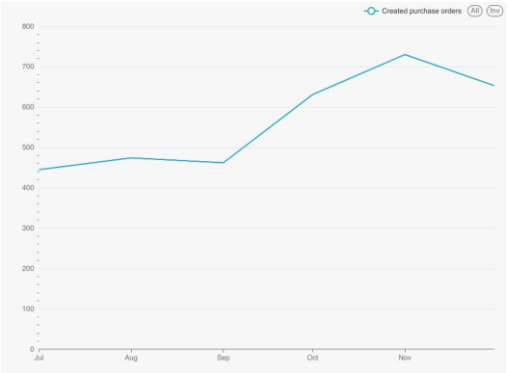
Sebagai perusahaan rintisan yang mengkhususkan diri dalam teknologi pertanian, saat ini PT XYZ tidak memiliki alat yang efektif dan terintegrasi untuk membantu mereka menganalisis harga komoditas bahan pangan, guna mendukung pengambilan keputusan pembelian dan penjualan berbasis data. Penelitian ini menggabungkan berbagai sumber data untuk analisis yang akurat, merancang solusi business intelligence (BI) yang spesifik dengan kebutuhan PT XYZ, da megnembangkan solusi yang komprehensif untuk mendukung pengambilan keputusan. Solusi implementnasi BI akan dikembagkan dengan menggunakan Penntaho Data Integration & Apache Superset. Penelitian terapan akan digunakan sebagai metode penelitia, dengan melibatka proses pengumpulan kebutuhan pengguna, desain solusi, dan pengembangan dashboard. Penelitian berfokus pada analisis data internal seperti transaksi jual beli, dan data eksternal seperti data pasar, untuk memperoleh wawasan yang komprehensif. Dashboard yang dikembangkan mencakup fitur yang memungkinkan pengguna untuk melihat harga rekomendasi, trend seminggu terakhir, memonitor posisi pasar, melihat performa tim, dan melihat rangkuman estimasi dan aktual dari keuntungan. Pentaho berhasil mengelola proses ETL dan data modellingn, sedangkan Apache Superset berhasil secara mudah membantu perancangan dashboard, walaupun terdapat limtiasi pada chart customization.

Kata Kunci: *Intelejensi Bisnis, Pengambilan Keputusan, Analisis Data, Visualisasi Data, Apache Superset*

1. INTRODUCTION

PT XYZ, a startup company specializing in agricultural technology with a vision to digitize the entire farming process comprehensively, has several departments in their main business process. One of those are the Sourcing Department which is responsible for procuring various types of groceries commodity from the producer or suppliers and the Sales Marketing Department that is responsible for selling those commodities to the retail or business customer. These two departments are critical for PT XYZ's business due to their direct impact on business performance, particularly in the context of price volatility and market fluctuations common in agricultural commodities. Procurement process, including economic order quantity calculations, spending decisions, and supplier selection, has the key role in meeting business goals by securing the essential resources needed for effective operations (Madarac, 2020)[1].

Based on the data data collected from PT XYZ regarding their purchasing transaction, currently they has 156,6k purchase transactions every month on average. As in Picture 1 shows, the procurement intensity can be seen from high purchase orders data in July-December 2023. It shows that even with the numerous data they owned and produced, PT XYZ has not optimized their data yet to support their business decision. This lack of data optimization, particularly in analyzing price trends and forecasting future market movements, underscores the need for a more specialized approach to data analytics in the agricultural sector. In other perspective, the urgency of having the tools to facilitate better data-drive decision making to support their high frequency transaction can be considered important.



Picture 1. Purchase orders PT XYZ

Considering the large volumes of data, Elgendy et al (2022) [2] explained that big data analytics and decision-making are crucial aspects of project success. There is a significant impact of big data analytics on project success, as well as an improvement in relationships through decision-making interactions between big data analytics and the three dimensions of project success.

In the context of PT XYZ's existing operations, the lack of effective and integrated tools or methods for defining purchase or sale prices is a significant gap. In other words, even with all the existing data resources they have, there is no integrated data to help them analyze the groceries commodity price, to support their purchase and sale decision making. Since every organization routinely makes decisions, from minor to crucial decisions, the quality of the decision must be considered as exceptional for the organization to function in an efficient manner. In the agricultural sector, where price fluctuations are frequent and unpredictable, the needs of accurate and timely decision-making is even more pronounced. Those decisions need to be taken with caution, because a single wrong action could lead to more damage than envisioned (Akshaya et al, 2023)[3].

Generally, business intelligence is the concepts that facilitate the process of company data processing which will later be used as the basis for the decision-making in their business process, such as purchasing or sales process. process of the sales process, which also allows company to manage and analyze data that can help them get visibility of profit prediction in the future (Nurdin, 2023)[4].

Based on the initial problem, this research will propose a business intelligence implementation solution for groceries commodity price data analysis in PT XYZ. Business intelligence is a discipline to understand the historical data that will be integrated and combined to obtain information and insights that leads to better business decision-making (van Gils, 2023)[5]. Implementing business intelligence and using it to get and explore the insights, such as identifying the business status quo and predict the trend, could drive good business decisions with competitive value (Maaitah, 2023)[6].

Although there have been previous studies on commodity price analysis, they often lack a focus on the specific challenges associated with groceries or food commodities and its implementation using the business intelligence concept. Previously, there were several and similar research for business intelligence implementation, such as the first research "Perancangan Data Warehouse Harga Pangan Di Wilayah Perumda Pasar Jaya" by Nisa (2021)[7]. The objective of this research is to develop a data warehouse for price food commodity in Perumda Pasar Jaya, so that it can be used as decision making consideration. In this research, they build the data warehouse using Kimball method and using Petaho Data Integrationn as its Extract, Transform, and Load (ETL) tool. The research completed the data warehouse development as the decided tools. The limitation in this research is the data warehouse could not quite gives the overall view for the decision making, due lack of data sources and visualization for further price analysis.

The second research, *"Aplikasi Monitorinng dan Prediksi Harga Komoditas Pasar Daerah Indramayu Berbasis Fuzzy Time Series"* by Samarudin (2022)[8], initially decided to develop a web-based app to monitor commodity price that can predict the price fluctuation. It developed in PHP using Laravel 5.8 framework, using Fuzzy Time Series as its prediction method. As the result, prediction can be acquired for a month ahead so that it can be anticipated. This research also has the limitation in the data sources that using to analyze the food commodity price that should be able to be optimized by combined multiple data source to get the further wider analysis for the prediction. The other limitation might be related in how the design and solution delivered without analyzing the requirements first, such as the needs of the data and what question should be answered by this developed solution.

The third one, *"Analisis Fluktuasi Harga Pangan di Kota Bogor"* by Bahtiar (2022)[9]. This research gives the overview how the pattern of food commodity price fluctuation based on the month-over-month trend and average comparison. It shows that there are several factors for the fluctuation such as commodity demand, weather, commodity production, etc. The insights from this commodity price analysis allows the users to get wide visibility regarding price fluctuation. However, in order to use it decision-making consideration it may not be a right fit, since the data is only present to be analyzed further. The other limitation is the source of analyzed data is only limited to single data source.

Lastly, "Business Intelligence Untuk Memantau Perkembangan Harga Pangan Provinsi DKI Jakarta" by Hidayat (2024)[10]. The objective of this research is to provide a solution to monitor food and groceries commodity in DKI Jakarta using business intelligence implementation. In order to delivering the business intelligence solution, this research used Pentaho and PowerBI as its business intelligence tool. The results shows the business intelligence solution can answer the potential price fluctuation, average, max, and min prices based on multiple dimensio such as commodity, location, and period of time. Similar like other previous research, this research is only combining one single of data source to analyze the commodity price. The business requirement is not deeply captured yet, so that the answer that can be answered in business intelligence solutions seems quite limited.

Based on those previous research, it shows that there are found several research gap, such as most of the research only analyzing data using one single data source, which restricted the depth and accuracy of its business intelligence analysis. It also found that the prior research did not thoroughly and explicitly analyze or incorporate specific business requirements, leading to solutions that were not fully aligned with user needs, and the limited question that can be answered in their analysis or business intelligence output. Lastly, the existing studies provided limited support for decision-making, often focusing only on data monitoring or visualization without offering actionable insights.

In essence, several novelties that can be delivered in this research are:

1. Provide multiple data sources (externally and internally) for the business intelligence implementation, expecting more comprehensive and accurate analysis of groceries commodity prices.
2. Empathize and address the specific business requirements of PT XYZ, as trader, ensuring that the business

intelligence solution is tailored to their unique needs.
3. Develop a robust business intelligence system that designed to support comprehensive decision-making, including price trend analysis and market forecasting, directly influencing purchasing and sales strategies.

Based on the problem and considered research novelty, the objective of this research are:
1. Gather the needs of data and information in order to support the groceries commodity price in PT XYZ.
2. Design the business intelligence solution, including architecture and the data modelling.
3. Implement business intelligence to analyze groceries commodity price data.

It is expected that the business intelligence implementation research for PT XYZ can drive the business growth and performance with better decision making in groceries commodity price data analysis. Also it could be a reference for any other business or institution who wants to implement the business intelligence from any field or industry.

1.1 LITERATURE REVIEW

Groceries Commodity Price Analysis

In general definition, groceries or agricultural commodities are products like food, fruits, vegetables that traded in supply chain. It can be purchased as inputs for the company's goods or as components of purchased items from a company's suppliers as part of a company's operations and overhead costs (Hamidu, 2022)[11]. When analyzing the groceries commodity it involves utilizing time series data to track the production ad price fluctuation, aiding in economic stability, forecasting, and also decision-making for wide and multiple stakeholders (Marina, 2024)[12].

As one of the approach, data analytics can be used to optimize the supply chain. It allows to drive lower costs and lower prices for consumers. Data analytics can help to identify the most efficient way to transport food products from farms to processing plants to retail stores (Irabor et al, 2023)[13]. Real-time data warehouse implementation in groceries and agricultural commodity allow to developing the agricultural sector to support decision-making quickly and accurately (Yudhistira et al, 2023)[14].

Data Warehouse

Data warehouse (DWH) is a large repository that consolidates data from various sources (usually from the transactional databases, internally & externally) to provide consistent data & information for end users, in order to enhance data-driven knowledge generation within the organizations (Aljuwaiber, 2022)[15].

Manikandan & Selvakumar (2023)[16] explained the several benefits to implement data warehousing in a organization, those are:
1. Improve forecasting abilities by recognizing the market trend using the end-to-end captured data.
2. Allows to handle massive volumes of data.
3. Enables the end user to navigate, comprehend, query the data due the curated and structured data in DWH.
4. Simpler design and manage queries than normalized databases, especially for analytics purposes.
5. Effectively handle the large demand from users through DWH.
6. Allows to evaluate a sizable volume of historical data that is provided by data warehousing.

Paliwal & Saraswat (2022)[17] also explain that as the vast repository of corporated, consolidated, and curated data from various sources that is used to provide insights for decision makers, there are several tools to gather all the data in DWH for reporting and analytics, such as:
1. DWH Appliances and Big Data technologies like Hadoop and Map Reduce
2. CDC (Change Data Capture)
3. Compression
4. Data Reduplication
5. Data Replication
6. EAI (Enterprise Application Integration)
7. ETL (Extract, Transform, Load)

ETL (Extract, Transform, Load)

In business intelligence perspective, an ETL tool helps to extract the data from one or multiple sources, cleanse and load it into DW. As integration techniques, ETL has an important role in integrating data and displaying it to users

in a single and functional format known as data integration (Sreemathy et al, 2021)[18].

Business Intelligence

DWH and business intelligence (BI) both are complementing disciplines, which by the contrast BI is discipline of making sense of collected historic, integrated, aggregated data in an approach to gain insights that lead to better business decisions, so that DWH is one of the tools that can be useful for implementing effective BI (van Gils, 2023)[5].

Orlovskyi & Kopp (2020)[19] explained that BI is an umbrella term, which includes reports, scorecards, dashboards, notification alerts, user pre-defined queries, ad hoc queries, multidimensional analyses, statistical analyses, prediction, and simulation models, and those BI capabilities helps business users to make vital business decisions in order to reduce costs, increase revenues, etc. Core of the BI capabilities are data extraction, storage, and visualization.

These capabilities are supported by the respective information technologies:

1. Extract-Transform-Load (ETL) pipelines.
2. Data Warehouses (DWH) and Data Marts (DM).
3. Reports and dashboards.

Business Intelligence Tools

There are several common tools and technologies that will support the BI implementation such as ETL pipeline, DWH, reports and dashboards. Those are:

1. **Pentaho Data Integration (PDI)**
   A software from Pentaho (also called Kettle) that can be used for ETL processes such as data migration, cleanse, and load to the transactional database or OLTP (vice versa). PDI has two primary components, those are the transformation component, which is the sets of instructions to change the input and output as expected or designed, and job component, which is the sets of instructions to run or execute the transformation (Pramata & Widhiasih, 2020)[20].
2. **Power BI**
   Cloud-based reporting service developed by Microsoft that allows the report creation and turns it into meaningful information and insights. It intended for users to independently create their own dashboards and reports without any technical skills or help from IT persons. Power BI allows their users to get and analyze any data from different sources that can be easily retrieved, visualized and shared with other users (Ozdemir et al, 2023)[21].
3. **Tableau**
   Data visualization tool used for Businesses Intelligence that allows to consume the raw data and represents it into understandable format. The reports & dashboard generated using Tableau can be easily understood by experts in the business, including the non-technical. One of Tableau's best features is Data merging, streaming data analysis, association of data (Kanchan & Khedikar, 2021) [22]
4. **Metabase**
   A tool that used for business intelligence based on open source. Metabase allows users to process their own data, and display it in an easily understood format, such as bar graphs or detailed tables (Anwar, 2020)[23].
5. **Apache Superset**
   An open source BI tool used for visualization and data exploration. It provides an advance of charts to build interactive dashboards. In Superset, the dataset can be brought from any source like defect through API calls. The data has to be processed and stored into a database (or DWH), so that the same dataset can be populated in the dashboard using superset. Superset also supports a wide range of SQL databases by using python ORM (Halundi & Moharir, 2021)[24].

2. RESEARCH METHOD

This research uses applied research as its method using the case study approach. Applied research is the research that is not limited only to understand and define the problems, but also the development of problem-solving with action for functional purpose (Firdaus et al, 2021)[25]. Overall, this research several phases. It initially started with gathering user requirement and analyzing the captured requirements to scope the solution features and capabilities. Based on captured requirements, the BI solution will be designed. It includes the blueprint design, architecture and the data modelling, how the data should stream from the extraction to load and presented as ready to consume visualization. Lastly, all the required data will be acquired and captured, the based on the gathered and available data, dashboard will be designed in BI development process.

Picture 2. Research stages

In the gathering user requirements process, the data will be collected throught structured interview and observation in order to capture the details such as how they currently analyze the price data before deciding the purchase or sell price, what is the data and information they need to support their purchase or sell price decision, and how they measure or define the ideal price for purchase or sell.

Population for the interview is the team members from the Sourcing & Sales Marketing Department, considering those are the departments that are facing the groceries commodity price analysis problem. Purposive sampling will be used to select the most relevant informants who can effectively answer the research questions. The criteria for inclusion are team members responsible for analyzing and setting the purchase and sale prices for groceries commodities. In this case, 5 individuals meet these criteria. One of each of informants are Head of Sourcing, Head of Sales Marketing, Sourcing Supervisor, Sales Supervisor, and Management. Those users highly considered due their day-to-day responsibility to analyzing the commodity price that will be purchased or sold.

After the requirements captured from the users, this research will convert the needs into a proposed solution that contains the architecture and blueprint. Which leads to the development process that will deliver the blueprint solution into the final output, business intelligence dashboard.

This research will use two primary tools for data warehouse integration and business intelligence tools. Pentaho will be used as a data integration tool, it provides the ETL solution that contains the main solution for data integration, transformation and job (Saraswati & Martarini, 2020) [23]. Apache Superset as the business intelligence tool. Superset is an open source business intelligence tool that is used for data exploration and visualization for a wide variety of charts and helps to build interactive dashboards (Halundi & Moharir, 2021)[21]. The implementation using Superset as business intelligence tools can be explored more due its current limited research that implement the tools. Furthermore, the effectiveness of these tools in supporting decision-making solutions will be reexamined to ensure they fully meet the research goals.

3. RESULT AND DISCUSSION

3.1 User Requirement Analysis

Based on the data that is captured from the gathering requirement session, here are several business intelligence requirements for commodity price analysis. Unlike the other studies, this research initially gathered and analyzed all of user needs in context of analyzing commodity price to decide purchase and sales price decision making.

Business Needs

In order to conduct the price analysis, user expects that they can gain the following insights:

1. The user wants to see every commodity recommendation price in current date daily and retrospectively (back-date), so that they can use the today recommendation price as the purchase/sell baseline.
2. The user wants to see the trend line of each commodity price fluctuation, so

that they can see and analyze the trend from each commodity.

3. The user wants to see the market trend position based on the highest and lowest market price, so that they can find the right market to source the commodity with fair and good price.
4. The user wants to see the trend between the market, competitors, and PT XYZ price, so that they can gather insight about PT XYZ market positioning.
5. The user wants to see the comparison and gap between actual PT XYZ purchase/sell price and the recommendation price for each commodity, so that they can use it to measure team performance.
6. The user wants to see the comparison between expected margin and the actual margin for each commodity, so that they can see if there is a higher margin potential for each commodity.

Data Sources

In order to get the price reference data, these are the internal & external data sources with multi-dimensional data, that has to be collected in commodity price analysis. This decision also considered based on the user requirements and the approach to optimize accuracy and wide range of price commodity analysis. It also contrasts previous studies that only consider one single data source to analyze the

Particularly for external data, it will be scraped from each source such as from the website. So in order to get the data, this process only develop a micro engine to scrap all of external data and extract and store to structured format.

Table 1: Data Sources

Source	Type	URL
Purchase orders	Internal	Transactional db
Sales orders	Internal	Transactional db
Informasi Pangan Jakarta	External	infopangan.jakarta.go.id
Paskomnas	External	trading.paskomnas.id
Bonrojo	External	t.me/BonRojoTelurBlitar

Market

Based on Business Needs #4, here are the whitelisted market in Jakarta that will be monitored in this commodity price analysis:

Table 2: Whitelisted Market

Province	Market Name
Jakarta Barat	Pasar Jembatan Lima
Jakarta Barat	Pasar Palmerah
Jakarta Pusat	Pasar Senen
Jakarta Pusat	Pasar Cempaka Putih
Jakarta Selatan	Pasar Kebayoran Lama
Jakarta Selatan	Pasar Minggu
Jakarta Timur	Pasar Klender
Jakarta Timur	Pasar Induk Kramat Jati
Jakarta Utara	Pasar Sunter
Jakarta Utara	Pasar Pluit

Each market represents the provinces and districts in DKI Jakarta. Other considerations, the whitelisted market is also considered as large traditional market and wholesale market.

Commodity

Regarding there are numerous groceries commodities in PT XYZ, this initiative will be only focusing on whitelisted commodities based on its impact for PT XYZ (fast moving and high transaction commodities). Those selected commodities are:

Table 3: Selected Commodity

Commodity	Source
Bawang Merah Kupas 1 Kg	IPJ, Paskomnas
Bawang Putih Kupas 1 Kg	IPJ, Paskomnas
Beras Ramos 1 Kg	IPJ, Paskomnas
Cabe Merah Besar 1 Kg	IPJ, Paskomnas
Cabe Merah Keriting 1 Kg	IPJ, Paskomnas
Cabe Rawit Merah 1 Kg	IPJ, Paskomnas
Kentang 1 Kg	IPJ, Paskomnas
Telur Ayam Negeri Curah	IPJ, Paskomnas, Bonrojo
Tomat Merah 1 Kg	IPJ, Paskomnas

3.2 Architecture & Solution Design

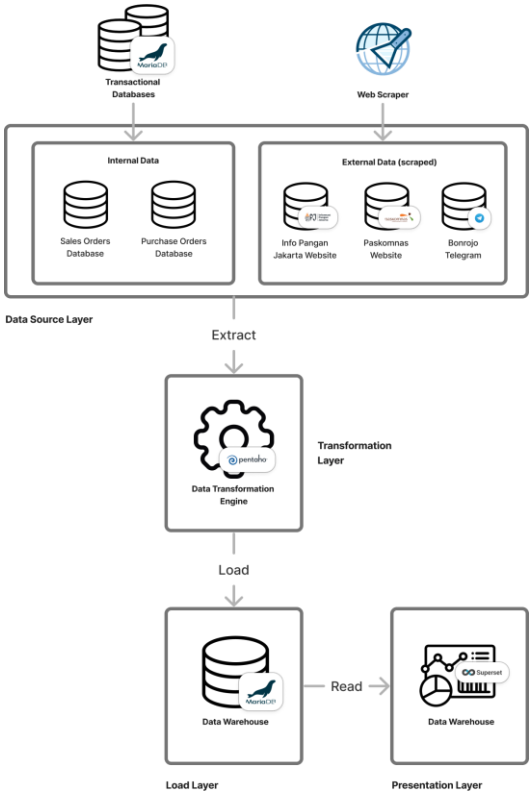
Based on the user requirement *analysis*, there are several blueprints that must be defined

before starting to develop the dashboard. Those are the data pipeline, data model, and data dictionary.

Data Pipeline

As it shows in Picture 3. Data Pipelinen, the designed data pipeline contains three layers:

1. Data sources layer that includes the external and internal data sources.
2. Transformation layer that uses Pentaho to transform all the extracted data into designed data cubes.
3. Load layer is where the data cubes are stored in the data warehouse (using MariaDB).
   Presentation layer that shows the data visualization in Superset.



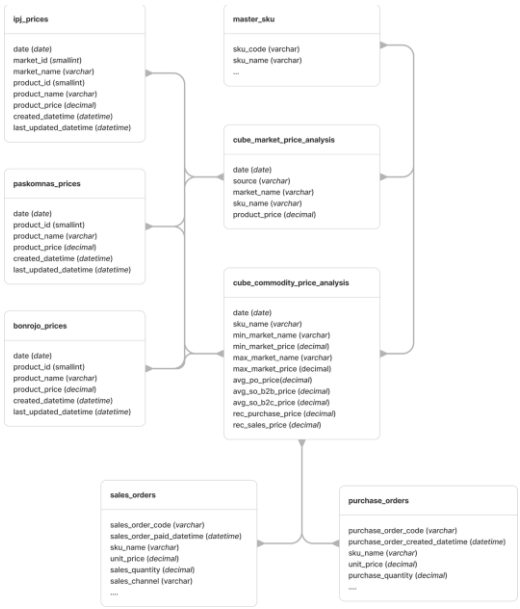
Picture 3. Data Pipeline

Data Model & Dictionary

Based on the requirements, there are two data cubes that has to be provided in order to show and analyze the data as the business needs. Those two data cubes are market price analysis cube that will show the market price from each source and commodity price analysis cube that will show price recommendation for purchase/sales, market recommendation, and purchase/sales performance.

In order to create those two cubes, there are several data table from external and internal that has to be extracted and transformed to the cube, such as ipj_price, paskomnas_price, and bonrojo_price for the external data source, and sku_master, sales_orders, and purchase orders for internal data source.

Refer to Picture 4. Data Model, those two cubes related with multiple external and internal sources. Primarily, those two cubes are the ones that will later on be analyzed and visualized in the dashboard. Each column definition will be descriptively explain in data dictionary.



Picture 4. Data Model

Based on Table 4 and 5, here are the data dictionary for the required two cubes:

Table 4: Market price analysis cube

Column	Type	Desc
date	date	Created date
source	varchar	Price data source
market_name	varchar	Market name for IPJ
sku_name	varchar	Commodity name
product_price	decimal	Commodity price

Table 5: Commodity price analysis cube

Column	Type	Desc

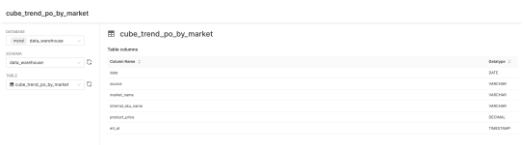
date	date	Created date
sku_name	varchar	Commodity name
min_market_name	varchar	Source & market name that has lowest price
min_market_price	decimal	Price of the lowest market
max_market_name	varchar	Source & market name that has highest price
max_market_price	decimal	Price of the highest market
avg_po_price	decimal	Average purchase orders
avg_so_b2b_price	decimal	Average B2B sales orders
avg_so_b2c_price	decimal	Average B2C sales orders
rec_purchase_pr	decimal	Purchase price recommendation
rec_sales_price	decimal	Sales price recommendation

Due to data confidentiality PT XYZ, the exact formulation for recommendation purchase and sales price is not explicitly mentioned in the data dictionary.

Using Pentaho, the data model successfully implemented and covered without any re-adjustment due Its incapability. In this research, Pentaho completely stream all the data as it architected in the data pipeline and ETL process.

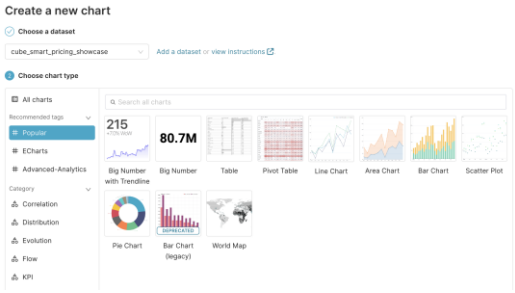
3.2 Commodity Price Analysis Dashboard Development

Initially, all created cubes in data warehouse (as designed in data model) must be added first in Superset. Once the cube added, Superset will link to the data warehouse to get the data. It has to ensure that Superset should be already connected to the data warehouse and database origin (including the scheme and table) can be accessed by Superset.



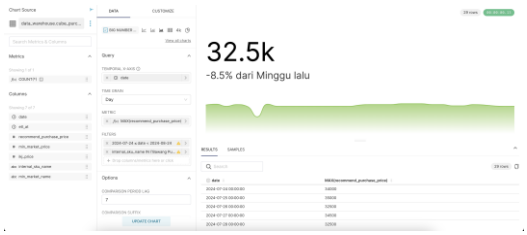
Picture 5. Adding Dataset in Superset

After the cubes is already available as dataset in Superset, the next step is designing the chart. In general concept, Superset can visualize a chart based on a particular dataset. So, in other words, one chart always has one dataset that linked to it. By default, Superset has widely rich chart design option that can easily be configured.



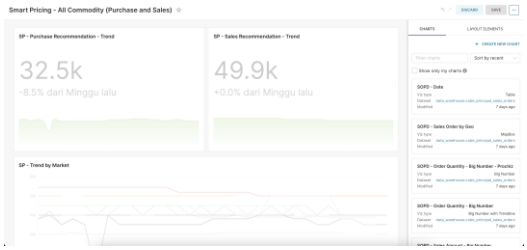
Picture 6. Adding Chart in Superset

When configuring a chart, Superset allows the parameter selection such as metrics, dimensions, filters, and style customization. As an example, here's are the configuration of one of Superset chart template called Big Number.



Picture 7. Configuring Chart in Superset

Lastly, once the all chart already designed and created, the next step is creating the dashboard. This process as simple as drag and drop the already created charts and designing the chart size and its layout. In Dashboard configuration, it also allows to setup filter for the dashboard.



Picture 8. Adding Dashboard

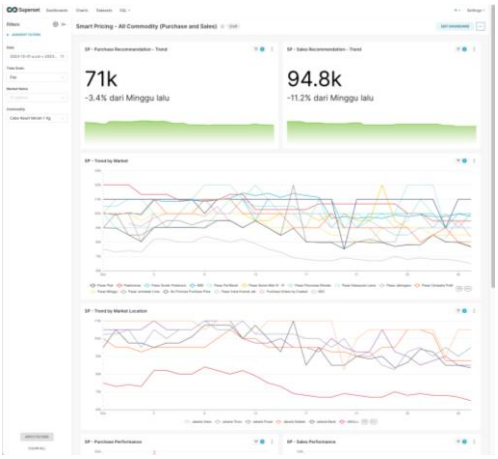
In dashboard development process, Superset can easily whole requirement as expected. The configuration process also can be finished seamlessly. The simple hierarchy between the datasets, charts, and dashboard make it easier to configure. Other than wide range of chart design library, Superset has limitation in

advance customization. It might be a hassle when there is a requirement to design a nearly brand-new type of chart that is not available in Superset. In essence, Superset could be a considered as easy to use business intelligence tool.

Dashboard Result

Based on the business requirements and designed solution, the Commodity Price Analysis Dashboard is successfully developed in Apache Superset. The dashboard contains several elements, such as datasets, charts, and filters. The datasets are connected to the designed market & commodity price analysis cube, the charts and filters are designed based on the business needs, so that the dashboard users will be able to interact and analyze the dashboard easily.

Due to data confidentiality from PT XYZ, the showcase that will be presented will be only the data that has date range limited to 1-31 December 2023 and filtered only for *"Cabe Rawit Merah 1 Kg"* commodity price.

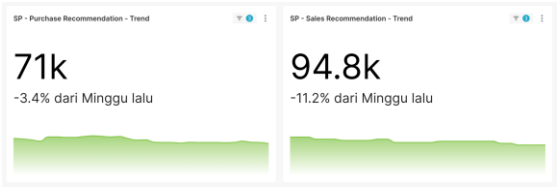


Picture 9. Overview Dashboard

The dashboard contains several charts and each of them designed for particular use cases that are already considered based on the business needs.

Purchase & Price Recommendation

This dashboard section shows the recommendation price from purchase and sales perspective and also insight trends from the last 7 days, using two identical charts. It designed using Big Number template. The number metric (eg. 71k and 94.8k) mentioned about the current price, so that when the user wants to purchase this recommendation price will help them to negotiate to the suppliers in the secondary market. For the sales perspective, the recommendation price will help them to set the minimum price to sell the product. Based on this business intelligence insight, the Sourcing and Sales Department can easily get the information about current date price and its historical fluctuation as negotiation or quote the sell price to the customer.



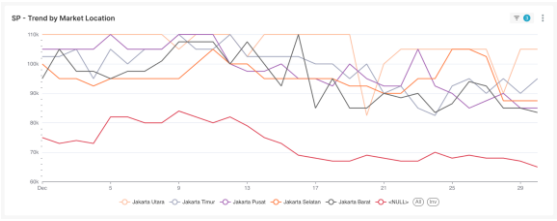
Picture 10. Price Recommendation

Trend by Market & Location

This dashboard section shows the market position from each selected market (refer to Table 2: Whitelisted Market). This dashboard designed using Time-series chart template. It gives users the visibility to understand which market has the lowest and the highest price. The insight from this chart can be used by the Sourcing Department to recommend the market to source and buy their product. For example, based on this chart, Pasar Cempaka Putih has the lowest price for Cabe Rawit Merah 1 Kg, so the Sourcing Department should be considered to source from Pasar Cempaka Putih market. So that based on this business intelligence insight, the Sourcing Department can consider to source from Pasar Cempaka Putih to get lowest price, compared to other market in Jakarta.

Picture 11. Market Trend

This section also came-up with the market location perspective price on average, based on each whitelisted market for a particular commodity.



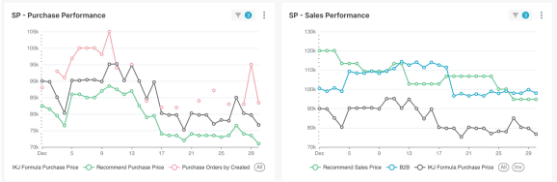
Picture 12. Market Trend by Location

Purchase & Sales Performance

This dashboard section shows two identical perspectives between purchase and sales performance, and each perspective has their own comparison metrics. For purchase performance, it compares the price recommendation price, average purchase order price, and primary market price (Pasar Induk Kramat Jati) to represent the market trend. For sales performance, it compares the recommendation price, average sales order price, and also Pasar Induk Kramat Jati price.

Each perspective allows the user to gain insight to see their current team performance. For example, in the purchase scenario, it seems they were sourcing Cabe Rawit Merah 1 Kg commodity too high than the purchase recommendation price, since the purchase recommendation price is the highest price threshold to source a particular commodity. In summary they should be allowed to source cheaper in order to optimize the margin.

In the sales scenario, it might be slightly different. The sales recommendation price acts as the lowest threshold to sell a particular commodity. Based on the data, it seems like the Sales Department is having a quite performance, since their average sales price is always catching-up and overcome the sales recommendation price.

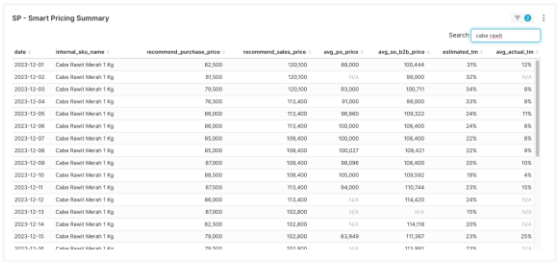


Picture 13. Performance

Estimated vs Actual Margin

This dashboard section contains the overall summary in tabular perspective. It allows the users to see the estimated vs actual margin. The actual margin is measured by subtracting the average sales orders price to average purchase order price. The estimated margin measured by subtracting the sales recommendation price to purchase recommendation price. If the team performs to follow the recommendation price from each purchase and sales, they can get the actual margin.

Based on this data for Cabe Rawit Merah 1 Kg commodity, the team was not always catching-up with the estimated margin (up-to -19% difference), but there was a time when the actual margin overcome the estimated margin (up-to +2%).



Picture 14. Estimated vs Actual Margin

4. CONCLUSION

This research successfully implemented business intelligence for groceries commodity price data analysis. As its objective, this research gathered the data and information as the fundamental business needs for commodity price analysis.

Based on the captured business needs, this research designed some blueprints regarding the business intelligence dashboard development, including data pipeline, data model, and data dictionary. Those blueprints implemented successfully into functional and ready to use business intelligence dashboard for commodity price analysis. The dashboard contains several sections that allows users to see the price recommendation price and get the insight for the last seven days trend, monitor the market position in order to select the right market to source, have visibility for the team performance in order to gain insight when the

team is not sourcing and selling based on the recommendation price, and lastly find the summary about estimated and actual margin comparison.

As the tools that are using in this research, Pentaho and Apache Superset successfully facilitate the objective to design and develop business intelligence solution as it should be. The implementation of ETL and data model completely covered by Pentaho to stream all data as it architected in the pipeline and ETL process. Superset as the visualization tool also provide easy to setup and configure dashboard. Starting from the data warehouse connection, widely rich chart templates, and straight forward dashboard design. However, if the dashboard development requires complex and advance chart design, Superset might be a hassle due its limited customization in their existing chart template. But overall, Pentaho and Superset could be a highly recommended tools for another business intelligence solution research and implementation.

For further research, the usability and business impact aspect would be a potential to continue this research, to ensure how usable this dashboard and the impact from the business-side quantitatively.

REFERENCES

[1] S. M. Mađarac, Z. Filipović, and M. Eljuga, "Purchasing Business in the Conditions of the Pandemic Crisis," in *ITEMA*, 2020, pp. 219–224. doi: 10.31410/itema.2020.219.

[2] N. Elgendy, A. Elragal, and T. Päivärinta, "DECAS: a modern data-driven decision theory for big data and analytics," *J Decis Syst*, vol. 31, no. 4, pp. 337–373, 2022, doi: 10.1080/12460125.2021.1894674.

[3] R. Akshaya, V. Ravi, S. R. Lakshmi, and R. Aparna, "Big Data to Big Impact: Effect of Big Data in Modern Decision Making," 2023, pp. 363–373. doi: 10.2991/978-94-6463-162-3_32.

[4] A. A. Nurdin, G. N. Salmi, K. Sentosa, A. R. Wijayanti, and A. Prasetya, "Utilization of Business Intelligence in Sales Information Systems," *Journal of Information System Exploration and Research*, vol. 1, no. 1, pp. 39–48, 2022, doi: 10.52465/joiser.v1i1.101.

[5] B. van Gils, *Data in Context*. Cham: Springer Nature Switzerland, 2023. doi: 10.1007/978-3-031-35539-4.

[6] T. Maaitah, "The Role of Business Intelligence Tools in the Decision Making Process and Performance," 2023.

[7] K. Nisa, D. Sugiarto, and T. Siswanto, "Perancangan Data Warehouse Harga Pangan Di Wilayah Perumda Pasar Jaya," *Explore:Jurnal Sistem informasi dan telematika*, vol. 12, no. 1, p. 47, 2021, doi: 10.36448/jsit.v12i1.1744.

[8] A Sumarudin, Adi Suheryadi, Bahrainsyah Oksareinaldi, and Lia Nurfadilah, "Aplikasi Monitoring dan Prediksi Harga Komoditas Pasar Daerah Indramayu Berbasis Fuzzy Time Series," *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 1, no. 1, pp. 15–24, 2020, doi: 10.30630/jitsi.1.1.4.

[9] R. Bahtiar and F. D. Raswatie, "Analisis Fluktuasi Harga Pangan di Kota Bogor," *Indonesian Journal of Agriculture Resource and Environmental Economics*, vol. 1, no. 2, pp. 70–81, 2023, doi: 10.29244/ijaree.v1i2.42020.

[10] M. K. Hidayat and D. Putri, "Business Intelligence Untuk Memantau Perkembangan Harga Pangan Provinsi DKI Jakarta," *Jurnal Infortech*, vol. 6, no. 1, pp. 30–36, 2024, doi: 10.31294/infortech.v6i1.21832.

[11] Z. Hamidu, P. B. Oppong, E. Asafo-Adjei, and A. M. Adam, "On the Agricultural Commodities Supply Chain Resilience to Disruption: Insights from Financial Analysis," *Math Probl Eng*, vol. 2022, 2022, doi: 10.1155/2022/9897765.

[12] I. Marina, D. Sukmawati, E. Juliana, and Z. N. Safa, "Dinamika Pasar Komoditas Pangan Strategis: Analisis Fluktuasi Harga Dan Produksi," *Paspalum: Jurnal Ilmiah Pertanian*, vol. 12, no. 1, p. 160, 2024, doi: 10.35138/paspalum.v12i1.700.

[13] Blessing Otohan Irabor, Adekunle Abiola Abdul, Bankole Ibrahim Ashiwaju, and Gbolahan Olaoluwa Oladayo, "a Comprehensive Review of the Role of Data Analytics in Shaping Food Pricing Strategies in the United States: Historical Perspectives, Current Trends,

and Future Projections," *Computer Science & IT Research Journal*, vol. 4, no. 1, pp. 36–53, 2023, doi: 10.51594/csitrj.v4i1.603.

[14] A. Yudhistira, I. S. Sitanggang, and H. A. Adrianto, "Development ETL (Extract, Transform and Load) Module in Indonesian Agricultural Commodities OLAP System," *ILKOM Jurnal Ilmiah*, vol. 15, no. 2, pp. 335–343, 2023, [Online]. Available: https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/1758

[15] A. Aljuwaiber, "Data Warehousing as Knowledge Pool : A Vital Component of Business Intelligence," *International Journal of Computer Science, Engineering and Information Technology*, vol. 12, no. 4, pp. 21–26, Aug. 2022, doi: 10.5121/ijcseit.2022.12402.

[16] Dr. L. C. Manikandan and Dr. R. K. Selvakumar, "A Review on Data Warehousing Concepts, Challenges and Applications," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 25–31, Jan. 2023, doi: 10.32628/cseit239015.

[17] M. Paliwal and P. Saraswat, "APPROACHES OF DATA WAREHOUSING AND THEIR APPLICATIONS: A REVIEW," *International Journal of Innovative Research in Computer Science & Technology*, pp. 117–121, Jan. 2022, doi: 10.55524/ijircst.2022.10.1.21.

[18] J. Sreemathy, R. Brindha, M. Selva Nagalakshmi, N. Suvekha, N. Karthick Ragul, and M. Praveennandha, "Overview of ETL Tools and Talend-Data Integration," in *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 1650–1654. doi: 10.1109/ICACCS51430.2021.9441984.

[19] D. Orlovskyi and A. Kopp, "A Business Intelligence Dashboard Design Approach to Improve Data Analytics and Decision Making," 2020.

[20] I. Putu, A. Eka Pratama, N. P. Nirmala, and D. Widhiasih, "PERANCANGAN DATA WAREHOUSE UNTUK PREDIKSI PENJUALAN PRODUK PADA ORBA EXPRESS MENGGUNAKAN PENTAHO," 2020.

[21] M. OZDEMİR, E. E. ÜLKÜ, and K. YILDIZ, "Analysis and Comparison of Business Intelligence Tools Most Preferred by Companies in Turkey," *Balkan Journal of Electrical and Computer Engineering*, vol. 11, no. 2, pp. 144–155, Jun. 2023, doi: 10.17694/bajece.1123171.

[22] K. A. Khedikar, "Data Analytics for Business Using Tableau," *SSRN Electronic Journal*, 2021, doi: 10.2139/ssrn.3835030.

[23] F. Anwar, I. Wahyuddin, and D. Hidayatullah, "Data Visualization Application 'JIRA' at PT. Jati Piranti Solusindo Using the Metabase," *Jurnal Mantik*, vol. 4, no. 1, pp. 289–293, 2020.

[24] A. Halundi, M. Moharir, and A. Professor, "CREATING DASHBOARD USING SUPERSET BY API CALL TO DIFFERENT SOURCE," 2021. [Online]. Available: www.jetir.org

[25] Firdaus, Zulfadilla, and F. Caniago, "RESEARCH METHODOLOGY : TYPES IN THE NEW PERSPECTIVE," *Jurnal Manajemen dan Ilmu Pendidkan*, vol. 3, no. 1, pp. 1–16, 2021, [Online]. Available: https://ejournal.stitpn.ac.id/index.php/manazhim