Title: Method of Data analysis in a multidimensional information System. (OLAP)

Course Title: Individual Project Work

Module Code: CST 4444

Course Leader: Ralph Moseley

Supervisor: Mark Springett

Individual Dissertation

Student Number: Stephen Cann Bentil

Student Number: M00967008

An Individual Dissertation writing in fullend of the course module requirements on individual dissertation writing.

Dept. of Computer Science

2023-2024

Table of Contents

Chapter 1 Introduction

1.0 Background of the Study

Multidimensional information systems and Online Analytical Processing are used for sophisticated searches and data processing. Multidimensional online analytical processing (OLAP) analyses may help organization's data, choices, and strategies. In online analytical processing (OLAP), Covacic et al. (2022) adopt the multidimensional model. Each cube dimension represents a different viewpoint on the data. Customers' active engagement with data across time, place, and product dimensions helps identify patterns and trends. According to Duda et al. (2020), retailers use OLAP to analyze sales data from several locations, time periods, product categories (including fashion, electronics, and grocery), and regions.

Online analytical processing (OLAP) requires "slicing," or constructing a one-dimensional subcube. Analysts can simplify by showing sales data by month or area. Ahmed et al. (2020) suggests using this strategy to a subset of dimensional data sets to better analyze comparisons and trends. Predefined value ranges for each dimension must be used to partition the data cube into subgroups before dicing. Data chopping is needed for portion extraction. Slicing the cube may interest a market researcher in quarterly electrical device sales. Hassan and Qader (2021) suggest using this method on smaller data sets to find trends that larger study misses.

Drill-down and roll-up are other crucial OLAP functions. Drill down from the large picture to specifics. Users may dig from annual sales data to quarterly, monthly, or daily data to analyze trends and patterns. Rolling up data at several depths condenses the same information. Small store sales data aggregated to regional or national levels might improve strategic planning. This summaries trends (Mao et al., 2020). Rotating, or pivoting, is essential in online analytical processing. This strategy turns the data cube upside down for a new viewpoint. Pivoting lets users change row and column dimensions to get different data perspectives. A sales report user may alter the data order from product categories in columns to timestamps in rows through pivoting. This would place regions at the top and timestamps at the bottom. This data analysis strategy reveals new insights and patterns (Sokolov et al., 2022).

Online analytical processing research requires data aggregation at numerous granularities. Trends may be shown via sums, averages, and counts. Merging sales data by month lets businesses track monthly performance and seasonal patterns. If they collect it annually, they may evaluate performance. Many OLAP systems support MOLAP and ROLA, according to Mukasheva, Yedilkhan, and Aldiyar (2022). MOLAP systems using OLAP cubes may swiftly integrate precomputed data. Relational Object Analysis and Processing (ROLAP) systems may dynamically create multidimensional views on relational databases by combining searches with aggregates. MOLAP's query time is faster due to precomputed aggregates, while ROLAP's flexibility makes it ideal for huge and difficult datasets (Beridze and Janelidze. 2020).

Combining data mining with OLAP improves analysis. Data mining seeks patterns, correlations, and trends in vast databases. We employ machine learning and statistics. Integrating OLAP and data mining may improve analytics (Forrest et al. 2021). Skills include data collection, trend identification, and consumer buying habits. Data visualization is essential to web analytics. Visualizing helps understand multidimensional data. Heat maps, graphs, and charts help identify trends in large datasets. Khalil and Belaïssaoui (2020) claim that data visualization improves pattern identification, analysis, and stakeholder engagement. OLAP and dimensional information systems utilize several methods to analyze data comprehensively and adaptably. Users may slice, dice, dig down, roll up, pivot, and aggregate data with various tools. Data mining and visualization enhance OLAP analysis. With this connection, customers may make better decisions utilizing complicated, multi-dimensional data (Cuzzocrea. 2020).

1.2 Research Aims and Objectives

The main aim of the current study will be to analyze the methods of data analysis in a multidimensional information system (OLAP)

1.2.1 Research Objectives

- Identify Prevalent Data Analysis Methods Used in OLAP Systems
- Evaluate the Effectiveness of Different OLAP Tools
- Investigate Challenges and Propose Solutions for Improving OLAP Systems

1.3 Rationale of the Study

Dimensional information systems, especially OLAP, must be researched since data is growing more intricate and firms need to make excellent judgements. Data interpretation will become more valuable as firm’s collect data from everywhere. Through sophisticated searches and multidimensional analysis, online analytical processing (OLAP) enables users interact with data from numerous viewpoints, including geography, time, and product category. Dahr et al. (2022) list several reasons for OLAP's relevance. First and foremost, OLAP systems' multidimensional paradigm lets you slice, dice, drill down, and pivot. Nuanced data analysis reveals trends, patterns, and insights, unlike flat data analysis. Yadav and Singh (2023) suggest organizations improve their analysis and decision-making by understanding these processes and their effects.

Data mining is more powerful using OLAP. Data mining and OLAP multidimensional analysis help us understand enormous data by revealing hidden patterns. Companies may improve performance and stay competitive by integrating predictive analytics with strategic planning (Najm et al., 2022). Since data visualization methods change, interpreting multidimensional data requires OLAP skills. Visualization helps improve results presentation and understanding of complicated data linkages. Understanding how online analytical processing (OLAP) develops these visualizations improves data analysis tools (Hamoud et al. 2020).

Chapter 2 Background and Literature Review

2.1 Common Data Analysis Methods in OLAP Systems

2.1.1 Slicing

OLAP systems work by slicing and breaking the data cube, and each dimension must be chosen separately. By retaining certain data and eliminating others, analysis is easier. OLAP systems can segment sales data by month, area, time, location, and product type. By dividing the dataset into smaller bits, analysts may spot patterns and trends along a dimension more readily. This strategy is useful if you want precise data insights but despise many perspectives (Boyko et al., 2020).

2.1.2 Dicing

Advanced dicing uses multidimensional subs election to reduce the data cube to its simplest components. Unlike slicing, dicing can set ranges or criteria in many dimensions. An analyst can use the sales data cube to analyses electronics sales over the last quarter or product category sales over a given time range. Because it may focus on fewer aspects simultaneously, this technique aids data analysis. Dicing lets users evaluate data conditions, making it essential for sophisticated queries and analytics (Li et al. 2021).

2.1.3 Drill-Down and Roll-Up

Drill-down and roll-up can switch data granularity. A user "drill-down" from a broad to a particular view of the facts. Researchers are breaking down annual sales data into monthly or daily statistics to find trends and patterns that may go overlooked in general reports. Roll-up merges data at several levels of detail to create a more condensed form. One way to see performance is to combine daily sales data and provide monthly or yearly reports. Reporting and strategic planning require these methods to comprehend data at multiple levels (Bensalloua and Hamdadou. 2021).

2.1.4 Pivoting

Pivoting (or rotating) the data cube gives a new viewpoint. Users can pivot to reorganize data view rows and columns. Pivoting might change the sales statistics presentation. It will show time in columns and regions in rows instead of product categories and time. This reorientation allows data to be examined from numerous viewpoints to comprehend patterns and relationships better. People may find new insights when they pivot and reframe the data (Kartamyshev, Chernysh and Murygin, 2021).

2.1.5 Aggregation

OLAP data analysis requires data aggregation at multiple granularities. It may involve calculating counts, averages, sums, and other statistical data. Aggregation simplifies trend and performance analysis by condensing complicated data into more straightforward summaries. Month-by-month sales data analysis helps analyses performance and detect seasonal trends. Aggregated data provides a multidimensional image of vital indicators for study and decision-making (Al Taleb, Hasan and Mahd, 2021).

2.1.6 Data Mining Integration

Data mining and online analytical processing (OLAP) augment dimensional data systems' analytical capacities. Data mining uses statistical and machine learning approaches to find patterns, trends, and correlations in massive datasets. When used with OLAP, data mining may improve trend projections and deeper insights from historical data. Data mining algorithms may predict sales and show customer purchase habits by analyzing past results and other variables. This relationship enables trend analysis and predictive analytics, improving data interpretation (Ismailov and Yalova, 2024).

2.1.7 Data Visualization

Data visualization is essential to OLAP because it graphically represents complicated multidimensional data. Charts, graphs, and heat maps aid data comprehension. Effective data visualization turns raw data into representations stakeholders can use for pattern spotting, segment comparison, and insight sharing. OLAP systems use visualization to make multidimensional data more accessible to understand and act on (Yadav and Tripathi, 2022).

2.1.8 Real-Time Data Analysis and Surveying Analysis

A real-time OLAP system processes and analyses data as it becomes available. This will be invaluable for firms that use real-time data for decision-making and insights. Real-time OLAP systems can accept live data inputs and analyses rapidly, allowing businesses to respond swiftly to trends, operational issues, and market developments. In fast-paced situations, real-time responses make organizations more flexible and adaptable (Noh and Yeo, 2021).

A survey was conducted on the OLAP system to gather insights and provide further evidence of its practical application. The survey aimed to assess how widely known and operated OLAP systems are, as well as their effectiveness in the global context. It comprised 13 questions, with a total of 83 respondents. However, 18 responses were incomplete. The average time spent on the survey was approximately 3 minutes, with a completion rate of 78%.

The image below shows a survey consent question, asking participants whether they freely and voluntarily give consent to participate in a project or study.

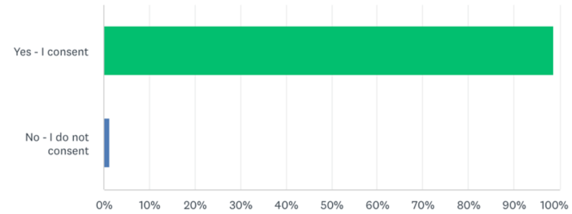
Key Observations:

- The overwhelming majority (98.8%) of participants consented to the study, which is a strong indicator of participant engagement and trust in the process.
- Only a small fraction (1.2%) did not consent, which is a normal occurrence in surveys where participation is voluntary.

Q1

Consent statement I have read and understood the participant information above and I freely and voluntarily give my consent to participant in this project/study.

Answered: 83 Skipped: 0



ANSWER CHOICES	RESPONSES	
Yes - I consent	98.80%	82
No - I do not consent	1.20%	1
TOTAL		83

A survey question (Q2) asking participants to specify their job role or position in the context of the study. A total of 73 respondents answered this question, with 10 skipping it.

Key Observations:

- **Data Analysts** are the predominant group in this survey, making up over one-third of the total responses. This suggests a focus on data analysis as a key component of the participant pool.
- **Business Intelligence** and **IT Managers** together form a significant portion, indicating that professionals involved in both data analysis and business operations are well represented.
- **Data Scientists** make up the smallest portion, which might suggest either fewer participants from this field or a focus on other data roles within the survey.
- The "Other" category shows that there are participants whose job titles did not fit into the standard options, reflecting some diversity in roles.

Q2

Customize Save as

Participant InformationSection 1

Answered: 73 Skipped: 10

ANSWER CHOICES	RESPONSES	
Data Analyst	38.36%	28
Data Scientist	9.59%	7
Business Intelligence	20.55%	15
IT Manager	16.44%	12
Other(Please Specify	15.07%	11
TOTAL		73

Survey 3

This survey question (Q3) with a picture below asking participants how many years of experience they have in their current role. A total of 74 respondents answered, with 9 skipping this question.

Since the majority of respondents are early-career professionals, organizations could focus on offering growth opportunities, mentoring, and advanced training to help these individuals develop further in their roles.

Key Observations:

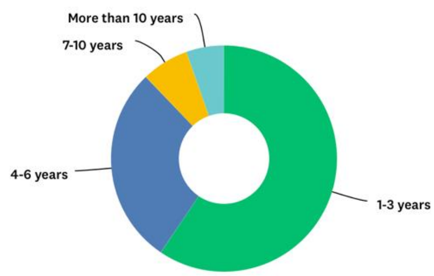
- **Majority with 1-3 years of experience**: The majority of respondents (59.46%) are relatively new to their roles, which could suggest that the survey is reaching participants who are early in their career trajectories or those transitioning into new roles.
- **A fair number with mid-level experience**: Almost 30% of respondents have 4-6 years of experience, suggesting a stable group of professionals who are well-established in their roles but may still be advancing in their careers.

- **Smaller representation of senior professionals**: With only about 12% of participants having more than 7 years of experience, this could imply that the majority of respondents are early to mid-career professionals, possibly indicating a younger or evolving workforce in the context of this study.

Q3 Customize Save as

How many years of experience do you have in your current role?

Answered: 74 Skipped: 9



ANSWER CHOICES	RESPONSES	
1-3 years	59.46%	44
4-6 years	28.38%	21
7-10 years	6.76%	5
More than 10 years	5.41%	4
TOTAL		74

Survey 4

The largest portion of respondents (62.30%) indicated that their organizations use an OLAP (Online Analytical Processing) system, while a smaller portion (16.39%) said their organizations do not use one. 21.31% of respondents were unsure if their organization uses an OLAP system.

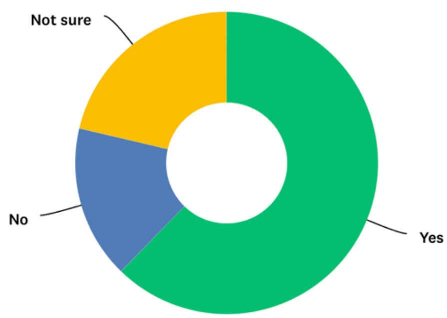
This suggests that OLAP systems are relatively common among respondents' organizations, but there is still some uncertainty or lack of awareness regarding their usage.

Q4

Customize | Save as

Does your organization use an OLAP system?

Answered: 61 Skipped: 22



ANSWER CHOICES	RESPONSES	
Yes	62.30%	38
No	16.39%	10
Not sure	21.31%	13
TOTAL		61

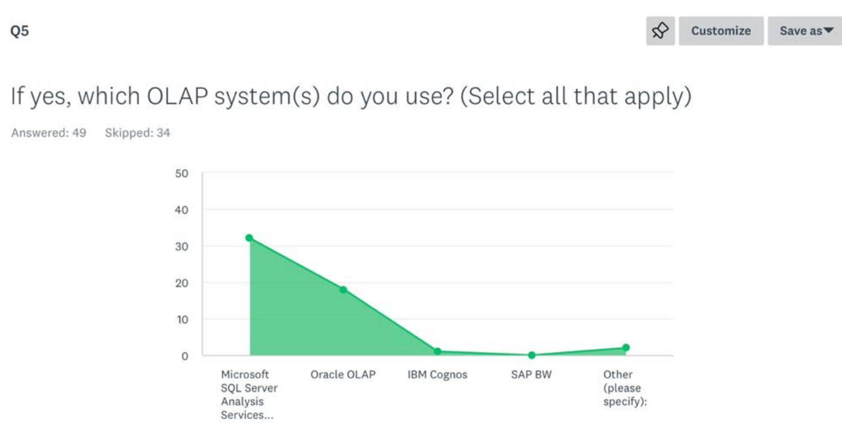
Survey 5

This survey question (Q5) asks participants which OLAP system(s) they use. A total of 49 respondents answered, while 34 skipped the question.

Observations:

- The most widely used OLAP system among respondents is **Microsoft SQL Server Analysis Services (SSAS)**, with a strong majority of **65.31%**.
- **Oracle OLAP** is also relatively common, with **36.73%** of respondents using it.
- **IBM Cognos** and **Other** systems were selected by very few, with **IBM Cognos** at only **2.04%** (1 respondent), and **Other** OLAP systems making up **4.08%** (2 respondents).
- Notably, **SAP BW** received no responses, indicating that it is not used by any of the participants in this survey.

This indicates that **SSAS** is the dominant OLAP system among the respondents, with **Oracle OLAP** being a secondary choice. Other systems like **IBM Cognos** and **SAP BW** have much less representation in this particular sample.



ANSWER CHOICES	RESPONSES	
Microsoft SQL Server Analysis Services (SSAS)	65.31%	32
Oracle OLAP	36.73%	18
IBM Cognos	2.04%	1
SAP BW	0.00%	0
Other (please specify):	4.08%	2
Total Respondents: 49		

Survey 6

This survey question (Q6) asks participants how frequently they use OLAP systems in their role. A total of 59 respondents answered, while 24 skipped the question.

The data suggests that OLAP systems are fairly integral for a majority of respondents, particularly for those who use them on a **daily or monthly** basis. This level of frequency highlights OLAP's role as a critical tool in many business functions, particularly in data-heavy roles. At the same time, the presence of **rare or non-users** demonstrates that not every role requires direct interaction with OLAP tools. These respondents might be in support roles or rely on more ad-hoc data analysis, delegating regular OLAP tasks to specialized teams.

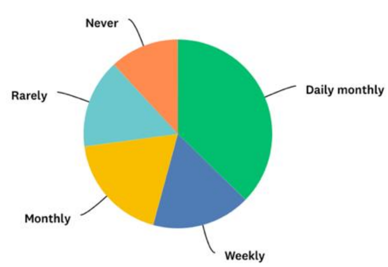
Observations:

- The majority of respondents (37.29%) indicated that they use OLAP systems either daily or monthly, showing that a significant portion of users regularly rely on OLAP tools.
- **18.64%** of respondents use OLAP systems on a **monthly** basis, and **16.95%** use them **weekly**.
- A smaller group, **15.25%**, only **rarely** use OLAP systems, while **11.86%** stated that they **never** use OLAP tools in their role.

Q6 Customize Save as

How frequently do you use OLAP systems in your role?

Answered: 59 Skipped: 24



ANSWER CHOICES	RESPONSES	
Daily monthly	37.29%	22
Weekly	16.95%	10
Monthly	18.64%	11
Rarely	15.25%	9
Never	11.86%	7
TOTAL		59

Survey 7

The results of a survey question below asking participants for what purposes they use OLAP (Online Analytical Processing) systems. As Picture Shown below.

Data-focused analysts: Given the dominance of data analysis, this suggests that a substantial portion of OLAP users are likely data scientists, analysts, or professionals in roles heavily focused on in-depth analytics.

Educators and researchers: The high proportion using OLAP for educational purposes points to a significant presence in academic environments, possibly reflecting professors, students, and institutions focused on data-centric curricula.

Forecasters and planners: The notable share of users leveraging OLAP for forecasting (28.57%) suggests that OLAP is valuable for strategic decision-making and planning roles (e.g., finance and operations).

Business users: Those using OLAP for budgeting and reporting may reflect more operational roles, such as financial planners, accountants, and project managers.

Analysis:

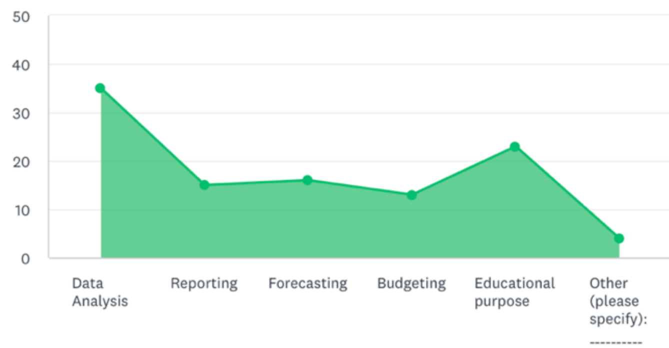
- **Data Analysis** is the leading purpose for using OLAP systems, reflecting the importance of OLAP in deep analytical tasks.
- **Educational purposes** also represent a significant use case, showing that OLAP is also used in academic or learning environments.
- The lower percentages for **budgeting** and **reporting** suggest these tasks may rely more on other types of software or tools that are more tailored to those functions

Q7

Customize Save as

For what purposes do you use OLAP systems? (Select all that apply)

Answered: 56 Skipped: 27



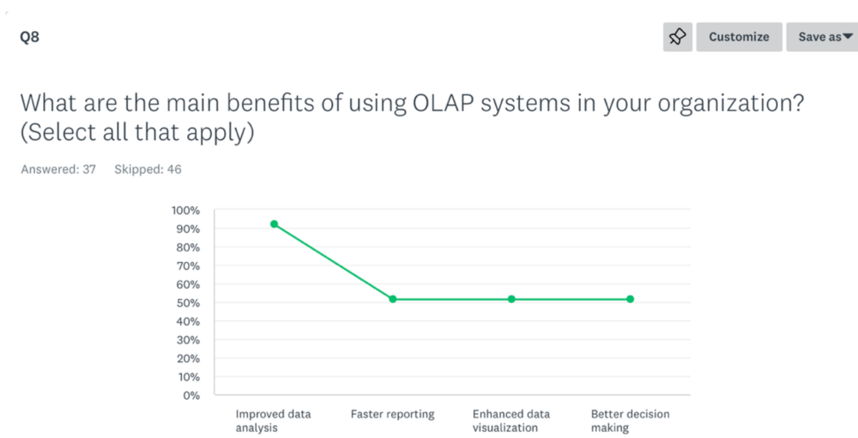
ANSWER CHOICES	RESPONSES	
Data Analysis	62.50%	35
Reporting	26.79%	15
Forecasting	28.57%	16
Budgeting	23.21%	13
Educational purpose	41.07%	23
Other (please specify): ----------	7.14%	4
Total Respondents: 56		

Survey 8

The analysis of the chart shows that **Improved Data Analysis** is the most significant benefit of using OLAP systems, with 91.89% of respondents citing it as a key advantage. Other benefits like **Faster Reporting**, **Enhanced Data Visualization**, and **Better Decision Making** are recognized by around 51.35% of respondents, indicating they are interconnected but secondary benefits.

Key insights include:

- **Improved Data Analysis** stands out as the dominant benefit, showing that OLAP systems excel in handling complex data for in-depth analysis.
- **Faster Reporting** and **Enhanced Visualization** are valuable but may not be as fully utilized or as strong as OLAP's core analytics functions.
- **Better Decision-Making** stems from the enhanced reporting and visualization features, highlighting OLAP's strategic value.
- The large proportion of skipped responses (55%) suggests that many users either don't use OLAP systems extensively or rely on other tools for specific functionalities like reporting and visualization.



ANSWER CHOICES	RESPONSES	
Improved data analysis	91.89%	34
Faster reporting	51.35%	19
Enhanced data visualization	51.35%	19
Better decision making	51.35%	19
Total Respondents: 37		

Survey 9

The chart illustrates the challenges faced by organizations using OLAP systems, based on responses from 37 participants. Multiple challenges could be selected, and the distribution is as follows:

Breakdown of Responses:

1. **Complexity of Use**
   - **51.35%** (19 respondents) highlighted this as a challenge, making it the most cited issue.
2. **Cost**
   - **43.24%** (16 respondents) identified cost as a major challenge.
3. **Integration with Other Systems**
   - **37.84%** (14 respondents) experienced challenges with system integration.
4. **Performance Issues**
   - **13.51%** (5 respondents) cited performance issues as a problem.

The key challenges faced by organizations using OLAP systems are:

1. **Complexity of Use** (51.35%) – The most significant challenge, indicating the need for more intuitive interfaces or better training.
2. **Cost** (43.24%) – A major concern due to the high expense of implementing and maintaining OLAP systems.
3. **Integration with Other Systems** (37.84%) – Many struggles to integrate OLAP with existing systems, leading to operational bottlenecks.
4. **Performance Issues** (13.51%) – A less common challenge, suggesting that most OLAP systems function efficiently.

Addressing these challenges with improved usability, better integration, and cost-effective solutions could enhance OLAP adoption and user satisfaction.

Q9

What challenges have you faced while using OLAP systems? (Select all that apply)

Answered: 37 Skipped: 46

ANSWER CHOICES	RESPONSES	
Complexity of use	51.35%	19
Performance issues	13.51%	5
Integration with other systems	37.84%	14
Cost	43.24%	16
Total Respondents: 37		

Survey 10

The chart below shows user satisfaction with OLAP systems based on 38 respondents:

- **Very satisfied**: 26.32% (10 respondents)
- **Satisfied**: 60.53% (23 respondents)
- **Neutral**: 13.16% (5 respondents)
- **Dissatisfied**: 0% (0 respondents)
- **Very dissatisfied**: 0% (0 respondents)

Summary:

- A majority (86.85%) of respondents are either **satisfied** or **very satisfied** with their OLAP systems.
- **Neutral** responses account for 13.16%, indicating a small percentage who may feel indifferent about their experience.

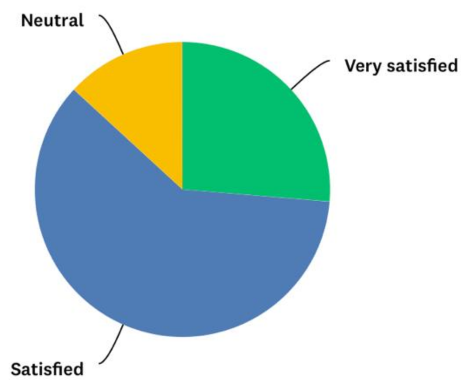
- **No dissatisfaction** was reported, indicating overall positive sentiment towards OLAP systems among users.

Q10

Customize Save as

How satisfied are you with the OLAP system(s) you use?

Answered: 38 Skipped: 45



ANSWER CHOICES	RESPONSES	
Very satisfied	26.32%	10
Satisfied	60.53%	23
Neutral	13.16%	5
Dissatisfied	0.00%	0
Very dissatisfied	0.00%	0
TOTAL		38

Survey 11

The chart below presents the belief regarding the future relevance of OLAP systems over the next 5 years:

- **Yes**: 94.74% (36 respondents)
- **No**: 0% (0 respondents)
- **Not Sure**: 5.26% (2 respondents)

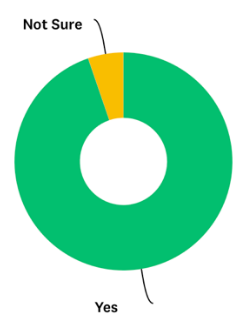
Summary:

A significant majority (94.74%) of respondents believe that OLAP systems will continue to be relevant over the next five years. Only a small percentage (5.26%) expressed uncertainty, and no respondents believe that OLAP systems will become irrelevant. This indicates strong confidence in the future of OLAP technologies.

Q11 Customize Save as

Do you believe that OLAP systems will remain relevant in the next 5 years?

Answered: 38 Skipped: 45



ANSWER CHOICES	RESPONSES	
Yes	94.74%	36
No	0.00%	0
Not Sure	5.26%	2
TOTAL		38

Survey 12

This chart below outlines the improvements respondents would like to see in OLAP systems. The responses are as follows:

- **Better user interface**: 52.78% (19 respondents)
- **Improved performance**: 36.11% (13 respondents)
- **Easier integration with other tools**: 44.44% (16 respondents)
- **More advanced features**: 30.56% (11 respondents)
- **Lower cost**: 30.56% (11 respondents)

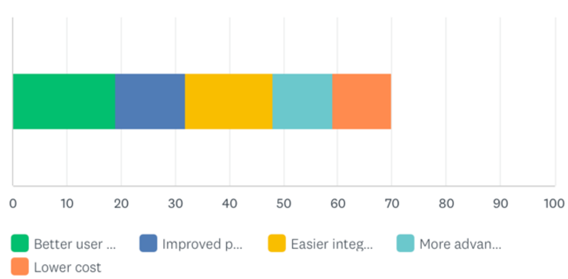
Summary:

The most desired improvement is a **better user interface** (52.78%), followed by **easier integration with other tools**(44.44%) and **improved performance** (36.11%). Around a third of respondents also want **more advanced features** and **lower costs**. The results suggest a focus on usability and functionality, while cost reduction is also a notable concern for many users.

Q12 Customize Save as

What improvements would you like to see in OLAP systems? (Select all that apply)

Answered: 36 Skipped: 47



ANSWER CHOICES	RESPONSES	
Better user interface	52.78%	19
Improved performance	36.11%	13
Easier integration with other tools	44.44%	16
More advanced features	30.56%	11
Lower cost	30.56%	11
Total Respondents: 36		

The final part

The Final part of the survey indicated how and what idea or knowledge can be added up to OLAP system or suggestions about OLAP System.

The results shows a list of responses to a survey question: "Any additional comments or suggestions about OLAP systems?" It contains 11 responses. Here's a summary of the analysis:

1. **Response Frequency**:
   - A few respondents left short answers such as "NA," "YES," or "Non," indicating no significant feedback.
   - One detailed response suggests making "complex data analysis easier for non-technical business users," highlighting usability concerns.
2. **General Sentiment**:
   - The feedback is minimal, with only one actionable suggestion among the responses. This suggests either satisfaction or indifference from most participants.
3. **Key Takeaway**:
   - The one substantial comment points toward improving user-friendliness for non-technical users, suggesting a possible gap in OLAP systems' ease of use for a broader audience.

2.1.9 Historical Data Analysis

OLAP works well with real-time and historical data. Historical data analysis finds patterns by studying older data. Users can compare present and past performance with online analytical processing (OLAP) systems that store data for lengthy durations. This study may inform future planning, KPIs, and priorities (Azevedo et al. 2022).

2.2 Recent Advancements in OLAP Systems

2.2.1 Cloud-Based OLAP Solutions

Cloud computing has modernized OLAP systems. OLAP systems like Google Big Query, Amazon Redshift, and Microsoft Azure Analysis Services are more scalable, flexible, and cost-effective than on-premises ones. These cloud technologies let firms enhance their analytics without buying cumbersome hardware. Cloud-hosted OLAP solutions offer data operations and data lake and ML integration. Cloud-based OLAP systems allow data analysis from any internet-connected device, an improvement over prior version. This improves dispersed teamwork (Ben Kraiem et al. 2020).

2.2.2 Real-Time Data Processing

Real-time data processing in OLAP systems is becoming more important as users seek rapid insights and decisions. Today, OLAP systems can analyses data streams in real time and draw conclusions. Real-time data helps retail, industrial, and financial companies make better strategic decisions and streamline operations. In-memory computing, data pipelines, and event streaming platforms enable real-time OLAP systems to analyses data quickly. This invention highlights new trends, unforeseen methods, and industrial advances in dynamic conditions (Duda et al., 2020).

2.2.3 Enhanced Data Integration

Modern data integration techniques allow OLAP systems to easily process complex and diverse data sources. OLAP systems may now virtualize, read real-time data, and use several data sources. You can combine data from data warehouses, relational databases, cloud storage, and APIs with this technology. Talend, NiFi, and Kafka integrate data. Users may integrate data from several sources into a single OLAP environment to improve analysis, discoveries, and decisions (Ahmed et al., 2020).

2.2.4 Advanced Analytics and Machine Learning

Modern analytics like AI and OLAP improve data analysis. Modern OLAP systems increase analysis using machine learning and sophisticated statistics. All of this is computerized pattern recognition, which includes anomaly detection and predictive analytics. Companies may better comprehend historical data, anticipate trends, and uncover patterns with OLAP systems that directly use machine learning algorithms. This link simplifies and automates complex analytical operations, improving data analysis and enabling more thorough research with less effort (Hassan and Qader. 2021).

2.2.5 Improved User Interfaces and Visualization

Online analytical processing (OLAP) systems are more efficient and user-friendly due to data visualization and UI enhancements. New OLAP systems include user-friendly interfaces that simplify multidimensional study creation and execution. Data visualization techniques like interactive dashboards, dynamic charts, and global maps are emerging to help users understand increasingly complicated statistics. These tools make it easy to change visualizations, analyze data, and share results with stakeholders. Improving user experience helps data analysts and executives to embrace OLAP systems, which improves decision-making (Mao et al. 2020).

2.2.6 Multi-Model Data Support

Multi-model data environments in online analytical processing (OLAP) allow users to evaluate many data models in one place. Multidimensional data models dominate online analytical processing (OLAP) systems, while document-based and relational models are also common. Businesses may analyze more data types in one OLAP environment with multi-model capability. Interpret complicated data linkages with relational and graph databases. OLAP systems with numerous models can help organizations manage diverse data environments (Mukasheva, Yedilkhan, and Aldiyar. 2022).

2.2.7 Enhanced Security and Compliance

Due to the growing importance of data security, online analytical processing (OLAP) solutions have increased compliance and security. OLAP systems protect sensitive data via encryption, user authentication, and role-based access limits. We made sure the OLAP system meets GDPR, HIPAA, and CCPA. Prioritizing security and compliance reduce data breaches and regulatory infractions. This allows them to conduct lawful, ethical, and safe analytical processes (Duda et al. 2020).

2.2.8 Scalability and Performance Optimization

Scalable and performant OLAP systems are needed to process big data volumes and complex queries quickly. Current OLAP solutions increase efficiency and scalability using distributed computing, parallel processing, and in-memory analytics. These advances allow OLAP systems to quickly analyze massive information, providing users with important insights. Caching and query optimization speed up data processing, ensuring user access (Boyko et al. 2020).

2.2.9 Integration with Big Data Technologies

Deep data integration with OLAP systems helps large data management and analysis. NoSQL databases, Spark, and Hadoop may interact with new OLAP systems. This site lets users examine structured, semi-structured, and unstructured data. OLAP and big data technologies may improve data strategy and provide valuable insights from several sources. They can study and search enormous databases using this. These findings enable more complete and scalable data analysis in huge data sets (Kartamyshev, Chernysh, and Murygin. 2021).

Last but not least, the new features have substantially increased OLAP system performance and capabilities. Improved user interfaces, scalability, security, real-time data processing, cloud-based solutions, and model interoperability are notable improvements. These advances boost data analysis and insights, helping firms make better decisions in complex, ever-changing situations. Al Taleb, Hasan, and Mahd (2021) indicate that OLAP tools and skills may increase data analysis and innovation in firms.

2.3 The state-of-the-art survey in OLAP Systems

Online Analytical Processing developed due to new technologies and complex data analytics. Complex multi-dimensional data management technologies in modern online analytical processing (OLAP) systems increase business intelligence and decision-making. These technologies simplify complex searches and analysis so users may extract useful data from large datasets. In their research on online analytical processing (OLAP) systems, Ismailov and Yalova, (2024) discuss current trends, significant technologies, and future prospects.

2.3.1 Cloud-Based OLAP Architectures

Cloud-based OLAP solutions have enhanced them. With cloud computing, OLAP systems may expand and alter. Google Big Query, Amazon Redshift, and Azure Analysis Services are examples. These cloud-based solutions offer on-demand analytical tools for firms to improve without investing a fortune. OLAP systems may link to data lakes, ML resources, and real-time processing platforms. This relationship simplifies data processing, making complicated searches easier to interpret (Noh and Yeo, 2021).

2.3.2 In-Memory Computing

OLAP systems use in-memory computation. In-memory computing uses system memory instead of CDs for quicker processing and data retrieval. In-memory computing lets SAP HANA and Oracle Exadata quickly access large data for complex calculations and real-time analytics. Financial trading systems, real-time corporate data, and interactive data exploration require fast data analysis. In-memory computing speeds up huge data processing (Mukasheva, Yedilkhan and Aldiyar, 2022).

2.3.3 Advanced Data Integration Techniques

Many data sources may be added to OLAP systems easily. Modern data integration solutions simplify combining cloud-based data with relational databases, APIs, and big data platforms. OLAP systems can employ Talend, Apache NiFi, and Kafka for virtualization and real-time data input. This link allows users to evaluate data from several systems in one OLAP environment. Kartamyshev, Chernysh, and Murygin (2021) say data integration helps analysis and decision-making.

2.3.3 Machine Learning and Predictive Analytics

OLAP systems can improve analysis using machine learning and predictive analytics. Modern OLAP systems examine data using complicated machine learning algorithms and statistical methodologies. This field encompasses pattern recognition, anomaly detection, and predictive modelling. Businesses may analyze historical data, find hidden trends, and draw conclusions using machine learning models and OLAP technologies. According to Hassan and Qader (2021), this alliance automates and improves analytical processes, resulting in more reliable data insights.

2.3.4 Enhanced Data Visualization and User Experience

Modern OLAP systems improve data visualization and user satisfaction. Modern technologies' intuitive user interfaces are ideal for multi-dimensional study. Data visualization tools like interactive dashboards, geographical maps, and dynamic charts may simplify complex data. These resources enable users to perform extensive investigations, create distinctive visualizations, and effectively convey results to stakeholders. With better visualization and user interfaces, OLAP systems may help organizations make better decisions. This helps corporate executives and data analysts (Ahmed et al., 2020).

2.3.5 Real-Time and Streaming Data Analysis

Modern OLAP systems need real-time data analysis. Businesses that thrive in processing and interpreting real-time data streams may make faster decisions and gain insights. Apache Storm, Apache Flink, and real-time data pipelines help OLAP systems analyze streaming data. This might benefit marketing analytics, operational monitoring, and fraud detection. OLAP systems with real-time data processing may help companies stay up with fast-paced scenarios. They become more agile and responsive (Azevedo et al., 2022).

2.3.6 Multi-Model Data Support

Combining and evaluating many data models in one OLAP system is innovative. Relational, graph, and document-based data models are supported by modern platforms, while classic OLAP systems priorities multidimensional models. Businesses may do sophisticated analytics on many data kinds by managing several models. Combining graph-based and relational data may illuminate complex relationships. Businesses may handle varied data environments better using OLAP systems that permit multiple models (Duda et al., 2020).

2.3.7 Security and Compliance Enhancements

Compliance and security are top priorities while creating and deploying OLAP systems. Modern systems protect sensitive data via encryption, user authentication, and role-based access limits. The study made sure the OLAP system meets GDPR, HIPAA, and CCPA. Mao et al. (2020) that firms use secure and moral analytical tools to prevent data breaches and regulatory transgressions.

2.3.8 Scalability and Performance Optimization

Performance and scalability have been priorities in modern OLAP systems. New technologies like distributed computing, parallel processing, and in-memory analytics might boost platform speed and scalability. These improvements in OLAP systems' capacity to handle massive datasets and complex queries allow users to swiftly get insights. Caching and query optimization speed up data processing. Scalable and fast-solving OLAP systems provide fast data access and interaction (Boyko et al., 2020).

2.3.9 Integration with Big Data Technologies

Big data technology's better analytical capabilities allow OLAP systems to easily manage enormous data sets. NoSQL databases, Spark, and Hadoop may interact with new OLAP systems. This site lets users examine structured, semi-structured, and unstructured data. OLAP and big data technologies may improve data strategy and provide valuable insights from several sources. They can study and search enormous databases using this. Azevedo et al. (2022) say this discovery improves large data analysis by making it more comprehensive and scalable.

Online analytical processing (OLAP) technologies are more powerful than before. Data visualization, security, in-memory computing, scalability, integration, real-time analysis, and cloud-based systems have advanced. These advancements enable firms evaluate data and get deeper insights to improve decision-making in complex and changing environments. If OLAP technology improves, business innovation and data analysis may increase (Ismailov and Yalova. 2024).

2.4 Effectiveness of Different OLAP Tools

Online analytical processing (OLAP) systems analyze complex data and deliver business intelligence insights from multidimensional datasets. These tools can slice, dice, and dig into multidimensional data. The success of online analytical processing (OLAP) tools depends on their functionality and integration with other technologies. To assess OLAP tool success, this section will examine their characteristics, features, and use cases (Azevedo et al., 2022).

2.4.1 Traditional MOLAP Tools

Both Oracle Essbase and IBM Cognos TM1 are prominent MOLAP systems with excellent analytical capabilities and high performance. Aggregated data cubes speed up complex calculations and MOLAP searches. Budgeting and financial planning benefit from these technologies' fast computation of complicated numbers. Multidimensional cubes offer complex data analysis. MOLAP systems may lose scalability and freshness if cube updates take too long. Despite these limitations, MOLAP solutions may be valuable for clients seeking dependable data configurations for high-performance analytics (Al Taleb, Hasan and Mahd, 2021).

2.4.2 ROLAP Tools and Their Scalability

Relational OLAP (ROLAP) uses relational databases for multidimensional analysis. MS SQL Server Analysis Services (SSAS) and MicroStrategy are examples. ROLAP tools quickly give multidimensional views by querying relational databases, unlike MOLAP, which uses pre-aggregated data cubes. This method's versatility and scalability make managing enormous data sets and changing requirements easy. ROLAP systems excel in real-time analysis, especially with dynamic data. Relational databases make integrating with current data infrastructures easier and conduct complicated queries across vast datasets. MOLAP systems may be faster than ROLAP tools for sophisticated searches involving big datasets. ROLAP solutions benefit businesses that value scalability and real-time data access over performance (Ahmed et al., 2020).

2.4.3 HOLAP Tools and Hybrid Approaches

Microsoft SQL Server SSAS and IBM Cognos use MOLAP and ROLAP components in HOLAP mode. HOLAP tools blend ROLAP's scalability and flexibility with MOLAP's performance. Users may get rapid query responses for summarized and complete data whenever needed. People utilize MOLAP cubes for aggregated data and relational databases for specific data. HOLAP systems excel in large-scale corporate analytics and reporting, where scalability and speed are crucial. This hybrid method helps organizations manage performance-scalability trade-offs and meet various analytical needs (Mukasheva, Yedilkhan and Aldiyar, 2022).

2.4.4 Cloud-Based OLAP Tools

OLAP systems in the cloud include Amazon Redshift, Google Big Query, and Microsoft Azure Analysis Services. OLAP technology has advanced significantly. These cloud-based analytical systems are versatile and scalable without on-premises equipment. Cloud-based OLAP solutions allow firms to adapt their analytical capabilities to changing needs. They work with data lakes, ML platforms, and real-time data processing technologies. These cloud-hosted apps let users interact and access data from anywhere with an internet connection. Before deploying cloud-based OLAP, organizations should address data security, cost management, and system integration. Companies wanting scalable, flexible, and instantly available analytical capabilities should choose cloud-based OLAP solutions (Kartamyshev, Chernysh and Murygin, 2021).

2.4.5 In-Memory OLAP Tools

In-memory OLAP systems enable fast data analysis. Products like Oracle Exadata and SAP HANA are examples. OLAP systems that store data in main memory rather than on disc provide real-time analytics and instant access to enormous datasets. Faster data processing, interactive data exploration, live financial analysis, and business intelligence are benefits of this technology. In-memory OLAP systems can handle complex computations and queries. However, in-memory processing may be costly and resource-intensive for large datasets. Without these difficulties, in-memory OLAP solutions are appropriate for quick analytics firms (Duda et al., 2020).

2.4.6 Advanced Analytics Integration

Advanced analytics and ML may improve OLAP tools. Automatic pattern identification, anomaly detection, and predictive analytics provide greater in-depth insights and sophisticated investigations in modern online analytical processing (OLAP) systems. Organizations can do advanced statistical analysis, pattern spotting, and trend forecasting with OLAP and machine learning models. Due to their analytical capability and ability to automate complicated processes, OLAP technologies are becoming more important. Most of the interaction's value depends on machine learning model performance and OLAP system integration of advanced analytics. Oh and Yeo (2021) say advanced analytics can make online analytical processing (OLAP) technologies more practical and useable, which might revolutionize them.

2.4.7 Data Visualization and User Experience

Data presentation and user participation might impact OLAP system performance. Interactive dashboards, dynamic charts, and geographical maps help users analyze complex data in OLAP solutions. Personalizing visualizations and using simple user interfaces make OLAP tools more accessible. User data engagement may accelerate insight finding. User-centric solutions with easy visualization options may help online analytical processing (OLAP) systems access and evaluate data. According to Hassan and Qader (2021), the tool's capacity to match consumer requests with visually appealing and analytically sound visualizations determines these traits.

Finally, OLAP system architecture, features, and applications impact efficiency. Traditional MOLAP systems thrive in static, high-performance data applications, whereas ROLAP technologies are flexible and scalable. Cloud-based OLAP solutions outperform HOLAP in accessibility, scalability, and integration. Advanced analytics and in-memory OLAP boost speed and analysis. Data presentation and user participation might impact OLAP system performance. To meet analytical needs and make informed decisions, organizations must understand the advantages and limitations of different OLAP technologies (Noh and Yeo, 2021).

2.5 Challenges for Improving OLAP Systems

Usability, scalability, performance, and data integration must improve in OLAP systems. These challenges are important because OLAP systems help companies make better judgements through data analysis. More data and more complicated analyses require better and more efficient online analytical processing (OLAP) systems (Mao et al. 2020).

2.5.1 Performance Optimization

Performance optimization for OLAP systems is difficult. Online analytical processing (OLAP) systems handle large datasets and complex queries, but they can strain system resources and slow response times. Performance optimization includes indexing, data storage, and queries. Query optimization improves SQL or MDX queries for speedier data retrieval. Relational tables, MOLAP data cubes, and ROLAP indexes benefit from indexing to expedite data access. Massive datasets may be stored via compression and partitioning. Despite these efforts, data quantities and query complexity make peak performance harder (Ahmed et al., 2020).

2.5.2 Scalability Issues

Performance optimization for OLAP systems is difficult. Online analytical processing (OLAP) systems handle large datasets and complex queries, but they can strain system resources and slow response times. Performance optimization includes indexing, data storage, and queries. Query optimization improves SQL or MDX queries for speedier data retrieval. Relational tables, MOLAP data cubes, and ROLAP indexes benefit from indexing to expedite data access. Massive datasets may be stored via compression and partitioning. Despite these efforts, data quantities and query complexity make peak performance harder (Ahmed et al., 2020).

2.5.3 Data Integration and Quality

Scalability may be improved in OLAP systems. Organizations are gathering more data; thus they require OLAP systems that can manage larger datasets with more users. In-memory or multidimensional cube-based OLAP systems may have scalability concerns when data volumes surpass their capacity. Cloud and distributed computing, with elastic resources, satisfy data needs best. These approaches may be challenging and resource-intensive to apply to enormous data quantities or real-time processing. Ismailov and Yalova (2024) suggest helping modern OLAP systems balance scalability, performance, and cost.

2.5.4 Handling Real-Time Data

OLAP systems in real-time data processing are becoming more critical. Several organizational settings require real-time data for decision-making. This includes fraud detection, operations monitoring, and dynamic marketing campaigns. Batch processing and data refreshes might limit real-time insights in standard OLAP systems. Improvement of OLAP systems through real-time data processing requires a data pipeline and streaming platform. OLAP analytics must accurately depict real-time data to boost system performance. Users can only fix this with real-time data processing (Al Taleb, Hasan and Mahd, 2021).

2.5.6 User Experience and Accessibility

Another critical subject is improving user experience and OLAP system accessibility. Any successful online analytical processing (OLAP) system should allow people to analyze and interact with data without technical experience. Engaging, customized visualizations, dashboards, and reports are included. Data query responsiveness, visualization quality, and system interface complexity affect user experience. OLAP systems must be user-friendly for all users, not just technical ones, to maximize their potential. This strategy demands continual attention to user wants and preferences, although well-designed user interfaces and proper training and support can alleviate these issues (Kartamyshev, Chernysh and Murygin, 2021).

2.5.7 Data Security and Privacy

Due to the sensitive data being handled, OLAP systems must priorities data security and privacy. Data privacy and GDPR/CCPA compliance are complex. OLAP systems should employ data encryption, user authentication, and access controls to protect user data. The multiplicity of data sources and data privacy standards make analytical tasks challenging without violating data privacy. We must pass comprehensive security legislation and regularly evaluate and enhance security procedures to overcome these challenges and protect ourselves from new threats (Mukasheva, Yedilkhan and Aldiyar, 2022).

2.5.8 Cost Management

When upgrading OLAP systems, adding or updating functionality may be expensive. Modern data-integrated, high-performance processor, cloud-based systems are expensive to upgrade. Investing in new technology, infrastructure, or corporate upkeep requires more study. Data processing, storage, and management involve operating costs that must be shared. OLAP system improvement technology should be purchased with prudence (Mao et al. 2020).

2.5.9 Technical Complexity

Today's OLAP systems allow technical improvement. The complexity of machine learning, real-time analytics, and several models makes OLAP systems harder to implement, run, and optimize. Organizations without technology expertise or resources may struggle with complex feature development and maintenance. Trained people and reliable support systems are needed since technological complexity can impair system performance and reliability. Hassan and Qader (2021) say firms need technical training, support, and documentation to overcome this impediment and boost OLAP system use.

2.5.10 Future-Proofing

Because technology and business needs change, building OLAP systems future-proof is tough. Data volume, analytical requirements, and technology changes should be evaluated to guarantee OLAP systems can meet future demands. Scalability, interoperability with new technologies, and data source integration should be prioritized. Technology must be purposefully implemented to stay relevant. This requires continuous system enhancements, capability evaluations, and organizational alignment (Duda et al., 2020). Finally, while creating improved OLAP systems, consider speed, scalability, integration, real-time processing, user-friendliness, security, cost management, technological complexity, and future-proofing. These issues must be handled for online analytical processing (OLAP) systems to help with complex data analysis and decision-making. If these issues are resolved, firms will have more efficient OLAP systems and higher-value data (Ahmed et al. 2020).

Requirements Specification for Improving OLAP Systems

1. Introduction

In today's data-driven economy, OLAP systems must be improved. Businesses use OLAP systems for complicated analytical queries, trend analysis, and massive dataset decision-making. This requirements specification specifies OLAP system improvement criteria. This solution ensures OLAP framework stability and adaptability to improve performance, scalability, and user experience.

2. Purpose

This section details the processes needed to build an enhanced OLAP system. Users with distinct functional, non-functional, and technological demands include data analysts, business managers, and IT professionals. The research also stresses the need of current technology improving OLAP procedures.

3. Scope

This project intends to build an enhanced OLAP system from scratch. Besides supporting several data sources, the system will have extensive analytical capabilities and real-time data analysis. Scalability, query performance, and data warehouse interoperability will be the major goals.

4. Functional Requirements

4.1 Data Integration

Integration with relational, NoSQL, and cloud-based data storage systems is essential for the enhanced OLAP system. To collect and store data from diverse sources in the data warehouse for analysis, the system needs execute ETL (Extract, Transform, Load) tasks (Kovacic et al, 2022).

4.2 Real-Time Data Processing

Data processing in real-time or near-real-time is essential. The system should handle streaming data and dynamic OLAP cube updates. Users can analyze and make data-driven choices quicker (Duda et al, 2020).

4.3 Advanced Query Capabilities

OLAP's enhanced query capabilities should support multidimensional analysis. Users should run complex queries like slice, dice, drill down, and pivot across dimensions. For maximum versatility, the system must support SQL and MDX query languages.

4.4 User-Friendly Interface

Online analytical processing (OLAP) systems need simple interfaces. Organized data, customizable reports, and dashboards and interactive visualizations should make outcomes easy to see. User-friendly drag-and-drop capabilities should allow users to design and alter queries without technical knowledge (Ahmed et al, 2020).

4.5 Security and Access Control

Data storage and retrieval should be the priority of an upgraded OLAP system. To achieve this, employ role-based access control (RBAC) to restrict users to work-related information and functionality. Protecting sensitive data during use and transit requires encryption (Mao et al, 2020).

4.6 Scalability

A scalable OLAP system is essential for meeting user needs as data and queries grow in volume and complexity. The system should be expanded horizontally and vertically to handle expanding data quantities and user demands. This allows adding servers or resources as needed (Honcharenko, Terentyev, and Gorbatyuk, 2021).

4.7 Data Visualization and Reporting

A high-level data visualization and reporting solution is needed. These tools should enable data pattern graphs, charts, and heat maps. Users should also get insights and create customized reports. Having the system export reports to CSV, Excel, and PDF would be great (Mukasheva, Yedilkhan, and Aldiyar, 2022).

5. Non-Functional Requirements

5.1 Performance

Fixing the OLAP system's performance is essential for fast query execution. The system should respond quickly even under heavy load to efficiently handle massive datasets and complicated queries.

5.2 Reliability

Non-functional requirements include system reliability. A reliable system is critical for high availability and low downtime. For data security and business continuity, organizations need backup, failover, and disaster recovery strategies (Forresi et al, 2021).

5.3 Usability

Its usability is prioritized by making the OLAP system accessible to users of all technical levels. These include an intuitive design, concise directions, and a low learning curve. The manuals and training should help users maximize the system's capabilities (Khalil and Belaïssaoui, 2020).

5.4 Compatibility

If all goes well, the system should work with several operating systems. Business intelligence solutions like Tableau, Power BI, and QlikView should work with it. Data integration also depends on how well it integrates with databases and data warehouses.

5.5 Maintainability

Maintainability should be prioritized in system design to enable upgrades, bug fixes, and updates. Modular designs allow developers to modify parts without impacting the whole, which is great. Unambiguous and well-documented coding standards simplify maintenance (Beridze and Janelidze, 2020).

5.6 Extensibility

For additional features and capabilities, the upgraded OLAP system must be adaptable. This includes adding data sources, supporting query languages, and employing machine learning for predictive analysis.

6. Technical Requirements

6.1 Architecture

Cloud computing and big data are recommended for distributed OLAP system construction. Horizontally expandable design should allow nodes to be added as needed to manage growing data and user demands.

6.2 Database Management

The system needs a stable OLAP-compatible DBMS. The author chooses a database management system that can swiftly handle huge data warehousing and queries. A system should have metadata management to keep data structures, relationships, and definitions current (Cuzzocrea, 2020).

6.3 Data Warehousing

The OLAP system needs a data warehousing solution optimized for it. When properly structured, data warehouses can store and analyze massive amounts of historical and real-time data, making analysis easier. Better data segmentation and indexing helps speed up queries (Yadav, 2021).

6.4 ETL Processes

Automating and optimizing ETL operations would improve their capacity to manage massive amounts of data from various sources. Data warehouse loading, cleaning, and transformation should be system essentials. ETL methods must allow real-time data updates and produce consistent, trustworthy data (Palominos et al, 2020).

6.5 Data Security

System security must be prioritized from data storage to user access. Auditing, encryption, and access control are included. The technology must satisfy industry standards like the GDPR to protect users' personal data (Shrivastava et al, 2021).

6.6 Integration with BI Tools

The new OLAP system must work with the most popular BI tools to improve data analysis and reporting. This interface should make it easy for clients to migrate data from OLAP to BI tools for analysis and visualization (Yadav and Singh, 2023).

7. User Requirements

7.1 Data Analysts

Data analysts need an OLAP solution to simplify complicated data searches and extract insights from massive datasets. The system should offer advanced analytics, real-time data processing, and customized outputs to match customer demands (Dahr et al, 2022).

7.2 Business Managers

If executives want reliable data to make decisions, they need a rapid delivery mechanism. The solution must have simple data visualization, easy interfaces, and executive-ready reporting without technical skills.

7.3 IT Professionals

IT workers need an easy-to-use OLAP system for deployment, administration, and maintenance. To serve the firm, this system must be safe, reliable, and expandable. The system needs APIs to connect to other corporate systems (Kaziyeva et al, 2020).

8. Conclusion

Companies should enhance their OLAP systems to boost data analysis. The system will be strong and extendable to accommodate the different demands of data analysts, business managers, and IT professionals by considering the technical, non-functional, and functional requirements in this document. The system's current technology and industry standards improve performance, reliability, and user satisfaction, enabling better business decisions and growth.

Analysis and Design of an Improved OLAP System

1. Introduction

A thorough requirements evaluation is needed to design an enhanced Online Analytical Processing (OLAP) system that matches the requirements specification. UML diagrams are used to analyze the OLAP system's design and needs. The system's major pieces, interdependencies, and design are identified during analysis. Functional and non-functional demands influence design decisions, which prioritize speed, scalability, usability, and security (Noh and Yeo, 2021).

2. Requirements Analysis

2.1 Identifying Core Components

The improved OLAP system's key components are defined at the start of requirements research. The following components are necessary for the functions:

This module handles ETL, which involves importing data from various sources into the data warehouse. This is the Data Integration Module.

In real time, RTP evaluates streaming data and refreshes OLAP cubes.

MDX queries may be processed alongside SQL queries and more complicated ones like slicing, dicing, and pivoting.

Well-designed user interfaces aid data visualization, report preparation, and system interaction.

The security module safeguards data and establishes RBAC.

Coordinates performance and scalability module horizontal and vertical scaling to meet expanding user and data needs.

The OLAP system needs these components to work.

2.2 Use Case Analysis

A use case study describes OLAP system user interactions. People who have been honored include:

Data analysts may create reports and in-depth investigations using the system.

The system's executive-focused data visualizations and reports can assist business managers make better decisions.

Information technology professionals recommend managing system installation, maintenance, and protection.

Consider this use case:

Data analysts and IT experts start ETL operations to combine data from diverse sources into the data warehouse.

Study and comprehend real-time streaming data with this technology.

Complex searches offer multidimensional analysis.

Management and data analysts draft and tailor reports.

Data Visualization: Users may create and interact with dashboards and charts.

IT experts monitor user roles and permissions to control data access.

Integrating these use cases into the system's architecture assures functional satisfaction for all users.

3. System Design Using UML

3.1 Class Diagram

The class diagram represents the static structure of the OLAP system, illustrating the system’s classes, attributes, operations, and the relationships between them (Li, Xia, and Shi, 2021).

Data Source: Represents the various data sources that the system integrates, including relational databases and cloud storage.

ETLProcessor: Handles the extraction, transformation, and loading of data from data sources into the data warehouse. It includes methods for data cleansing and validation.

Data Warehouse: Stores the integrated data, optimized for OLAP operations. It includes attributes for data partitioning and indexing to enhance performance.

OLAP Cube: Represents the multidimensional data structures that allow for complex querying and analysis. It includes methods for creating, updating, and querying cubes.

Query Engine: Processes user queries, supporting both SQL and MDX languages. It includes methods for slicing, dicing, and pivoting data.

Report Generator: Provides functionalities for creating, customizing, and exporting reports in various formats.

Visualization Engine: Responsible for generating data visualizations, including charts and dashboards, with methods for interacting with visual elements.

Security Manager: Manages user authentication, role-based access control, and data encryption. It includes methods for user management and access auditing.

Scalability Manager: Oversees the system's scalability, including horizontal and vertical scaling. It includes methods for resource allocation and performance monitoring.

The relationships between these classes ensure that the system components work together to provide the required functionality.

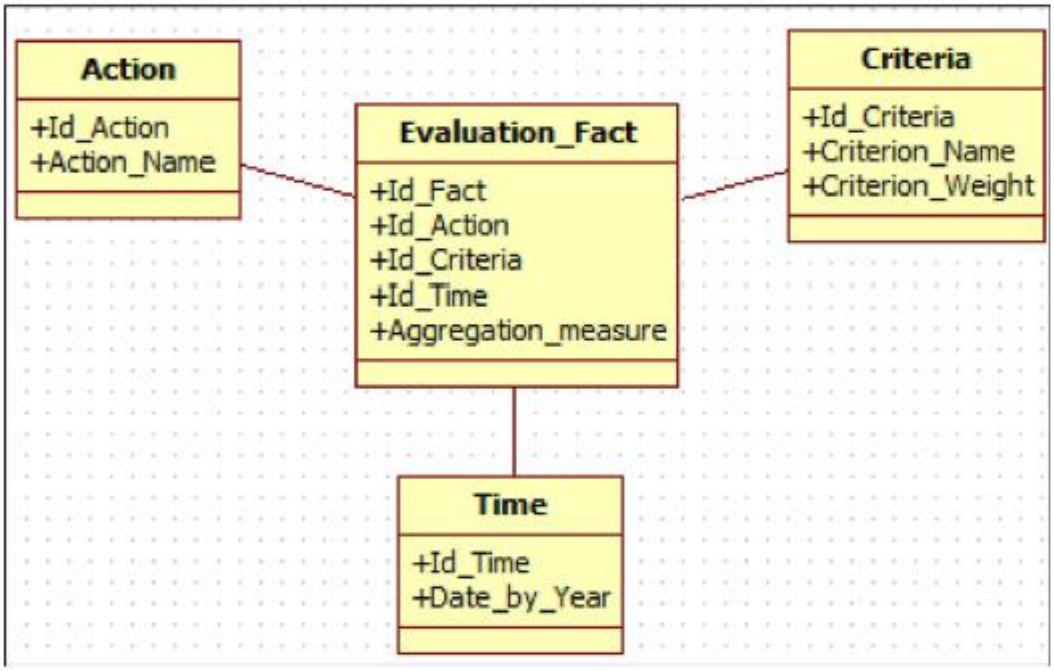


Figure 1 UML Diagram

3.2 Sequence Diagram

The sequence diagram illustrates the interactions between different components during the execution of specific tasks. For example, a sequence diagram for the Real-Time Data Processing use case would involve the following steps:

The Data Source sends streaming data to the ETLProcessor.

The ETLProcessor transforms the data and loads it into the Data Warehouse.

The Data Warehouse updates the relevant OLAP Cube.

The Query Engine processes a real-time query initiated by the user.

The Report Generator generates a report based on the query results.

The Visualization Engine creates a visualization of the real-time data.

The Security Manager ensures that the data access adheres to the user's role and permissions.

This sequence demonstrates how different system components interact to process real-time data and deliver insights to the user.

3.3 Activity Diagram

An activity diagram is used to model the workflow of the Data Integration process. The workflow includes:

Data Extraction: The system extracts data from multiple sources.

Data Transformation: The extracted data undergoes transformation, including cleaning, formatting, and aggregating.

Data Loading: The transformed data is loaded into the data warehouse.

OLAP Cube Update: The relevant OLAP cubes are updated with the new data.

Validation: The system validates the updated data to ensure accuracy and consistency.

The activity diagram highlights the sequential flow of operations, emphasizing the steps involved in integrating data from various sources.

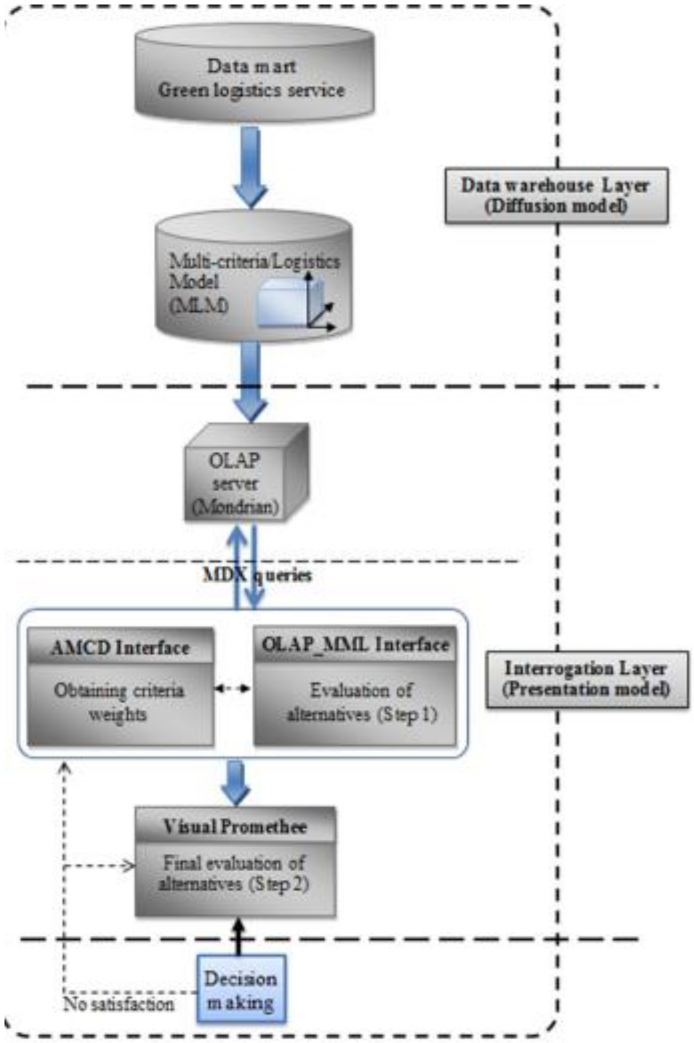


Figure 2 Activity Diagram

4. Design Decisions

4.1 Choice of Architecture

When performance and scalability matter, employ distributed architectures. Distributed systems support horizontal scalability by adding servers to meet demand. High availability and fault tolerance in case of component failure are achieved.

4.2 Database Management System

OLAP projects depend on choosing the correct database management system. A database management system should be able to handle large-scale data warehousing, be compatible with current systems, and conduct queries well. Database management system indexing and partitioning can improve data retrieval and query execution (Yadav and Tripathi, 2022).

4.3 Real-Time Processing Capabilities

To fulfil this demand, the design includes a real-time streaming engine. Data is processed and analyzed swiftly via this engine's integration with the ETLProcessor and OLAP Cube. Apache Spark Streaming and Apache Kafka offer real-time processing (Al Taleb, Hasan, and Mahd, 2021).

4.4 User Interface Design

People prioritize accessibility and usability in the user interface. Drag-and-drop and interactive visualizations were introduced to suit users of different technical levels. The user interface should also adapt to multiple screen sizes and devices (Reddy, 2021).

4.5 Security Considerations

OLAP systems process sensitive data, thus safety must be a concern. The design uses encryption and role-based access management. The Security Manager class ensures that only authorized users may access and alter data.

4.6 Integration with BI Tools

Due to the need for enhanced reporting and visualization, the OLAP system was integrated with popular BI tools. This connection lets clients use their business intelligence tools to improve reports and dashboards. Tableau, Power BI, and QlikView are popular OLAP system interfaces. This lets clients choose the finest analytics platform (Dehdouh, Boussaid, and Bentayeb, 2020).

5. Conclusion

The upgraded OLAP system was carefully studied and designed to satisfy the needs. We established the system's architecture, components, and operations using class, sequence, and activity UML diagrams. The OLAP system's scalability, speed, usability, and security were considered when designing it to handle complicated data analysis, real-time processing, and comprehensive reporting. These design criteria determine whether a system can meet current and future needs. In dynamic corporate contexts, the OLAP system's distributed design, real-time processing engine, user-friendly interface, and BI tool compatibility enable data-driven decision-making.

Implementation and Testing of the Improved OLAP System

1. Implementation

Implementing the upgraded OLAP system requires creating a working software solution. The process involves setting up the development environment, coding system components, integrating them, and delivering the application. We build and test each element to guarantee it satisfies design criteria (Bensalloua and Hamdadou, 2021).

1.1 Development Environment Setup

Modern tools and technology simplify OLAP system development in the development environment. Instruments used:

Use IntelliJ IDEA, essential for Java programming.

Database: The data warehouse will employ PostgreSQL, which has strong indexing and partitioning.

Apache Spark with Apache Kafka for Real-Time Processing.

The front-end framework React.js handles UI responsiveness and engagement.

The Spring Security encryption and role-based access control application.

Setting up Jenkins for continuous integration pipelines and Git for version control automates deployment and testing.

1.2 Coding the System Components

Data Integration Module is built with Ansys and Java. Data is extracted, transformed, and loaded into PostgreSQL using ETLProcessor. Custom Java classes use Kafka consumers and producers to alter and extract data. RTPE processes data in real time with Apache Spark Streaming. The Realtime Processor class refreshes OLAP cubes near-real-time (Bimonte et al, 2022). Spark uses Data Frame API for transformations and aggregations. The OLA-PLUS Administration

OLAP cubes may be managed with PostgreSQL's OLAP Cube class. To build and manage dimensional data structures, this class speaks to PostgreSQL. Cube operations use PL/pgSQL and SQL system procedures. The Query Processor runs SQL and MDX queries. Internal Query Engine class connects with PostgreSQL to process and execute OLAP cube data collecting and aggregation queries (Singh and Dev, 2023).

The React.js interface allows report production and interactive dashboards. The components can slice, dice, and pivot data. Spring Security manages Security Module. Security Manager implements authentication and authorisation. It secures sessions and user data via encryption and JWT tokens. The performance module configures PostgreSQL for horizontal scaling and Spark clusters for distributed processing. Scalability Manager allocates resources and tracks performance indicators (Kovacic et al, 2022).

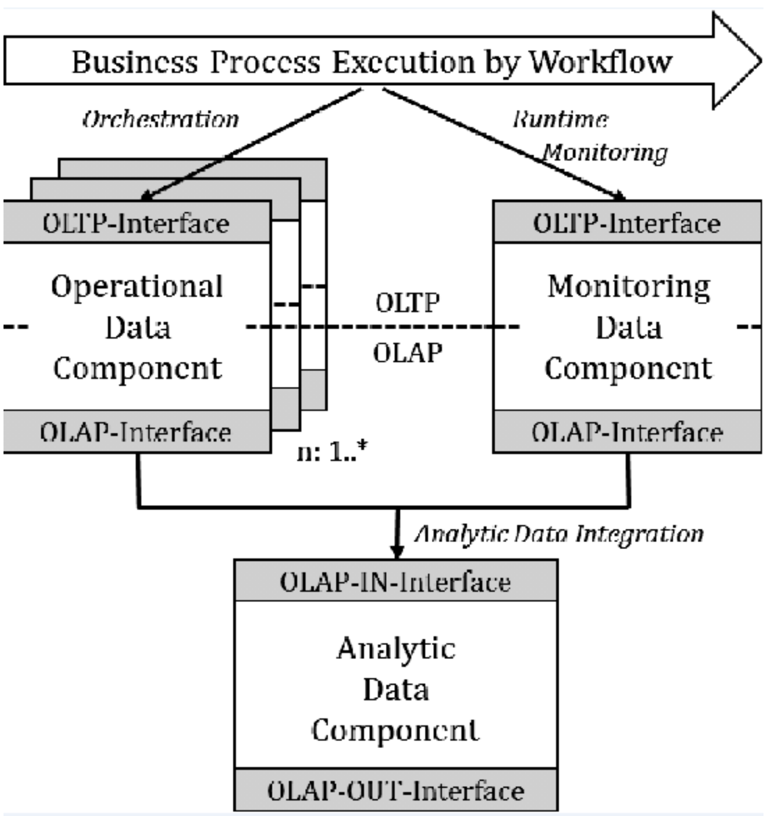


Figure 3 System Component

1.3 Integration and Deployment

Integrating system components ensures they function together. Items like:

Integration testing includes data integration, real-time processing, and query execution.

Cloud platforms like AWS provide system scalability and availability. Kubernetes orchestrates Docker containers for component deployment.

Continuous system updates are ensured by deployment script automation. Automated testing and deployment are made possible by CI/CD pipelines.

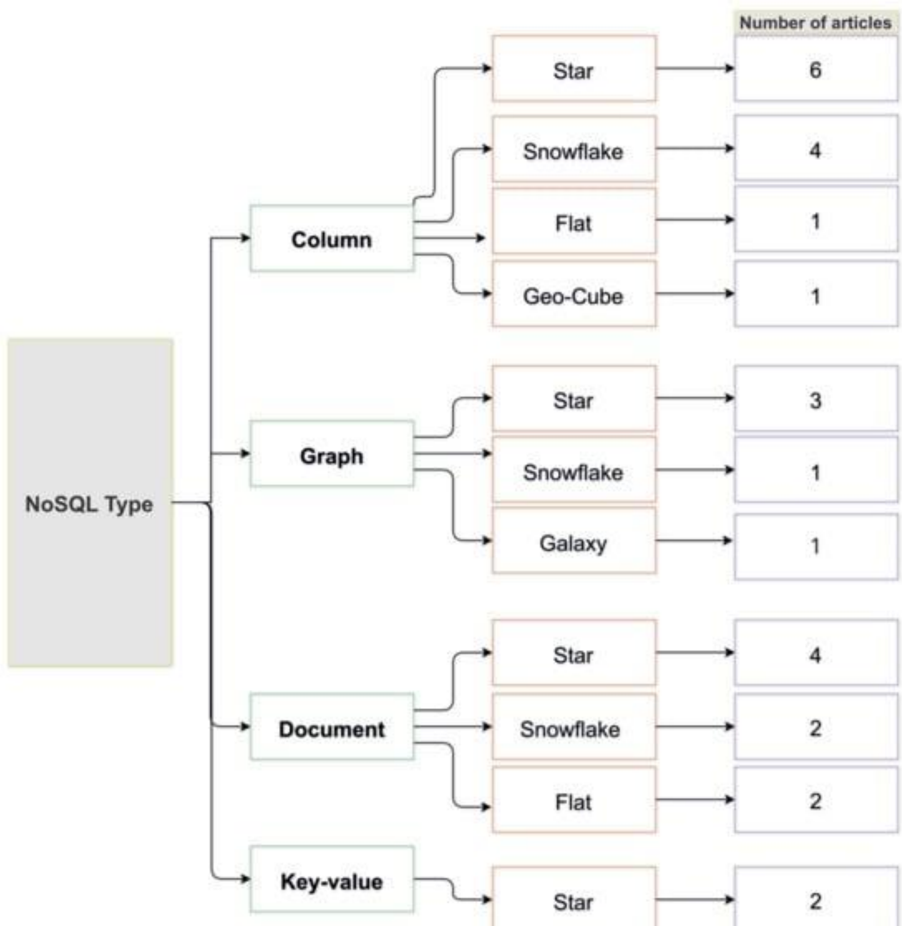


Figure 4 Instigation of Queries

2. Testing

Testing is essential to verify the OLAP system satisfies all criteria and works properly. User acceptability testing, system, integration, and unit testing.

2.1 Unit Testing

Unit testing ensures each component works properly when tested separately. Every function and class has its own

The data integration module validates data extraction, transformation, and loading via unit tests. Using phoney data, one may simulate many scenarios.

Data processing and changes should be checked in real time for accuracy. Spark Streaming unit tests ensure data transformations proceed as planned.

Unit tests verify that generating, updating, and querying OLAP cubes works. Tests verify cube operations are completed correctly.

Testing ensures the Query Processor parses and executes queries appropriately. SQL and MDX queries are tested using example datasets.

React component unit tests assure expected user interface interactions and visualisations.

Security module tests ensure system authentication and authorisation. Security testing includes checking encryption and role-based access constraints.

Unit tests for Spark Streaming tasks, react components, and Java classes may be developed using Spark's testing tools or JUnit. These tests provide continuous validation throughout development in Continuous Integration/Continuous Deployment (CI/CD) (Kovacic et al, 2022).

2.2 Integration Testing

To ensure system components operate together, integration testing is essential. This involves

End-to-end testing includes ETL, real-time updates, and queries. This includes Data Source data flow verification. Test cases simulate real-world circumstances to ensure data processing and reporting accuracy.

Testing interfaces means verifying the Query Processor, Data Warehouse, and UI work properly. Thus, data is shown appropriately, and user efforts get the desired effects.

Performance testing evaluates how effectively a system handles various workloads. We can test the system's responsiveness and ability to handle huge data loads and multiple users by imitating these scenarios (Duda et al, 2020).

Postman is used for API integration testing and JMeter for performance testing. CI/CD pipelines perform automated test scripts for continuous integration testing.

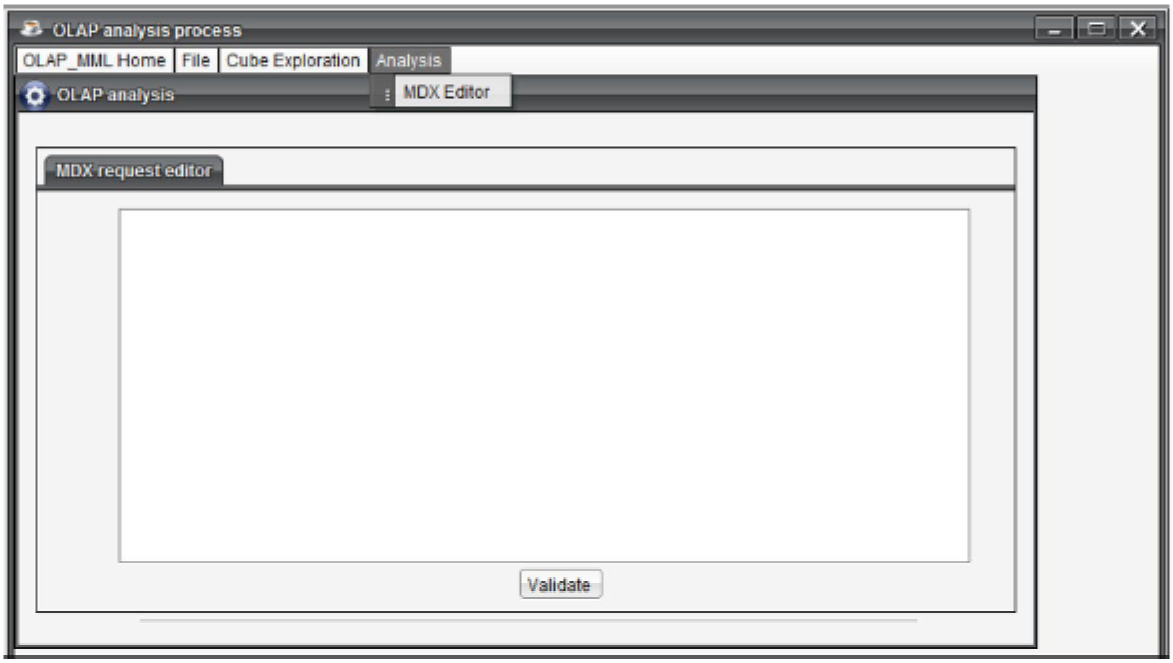


Figure 5 Program Testing

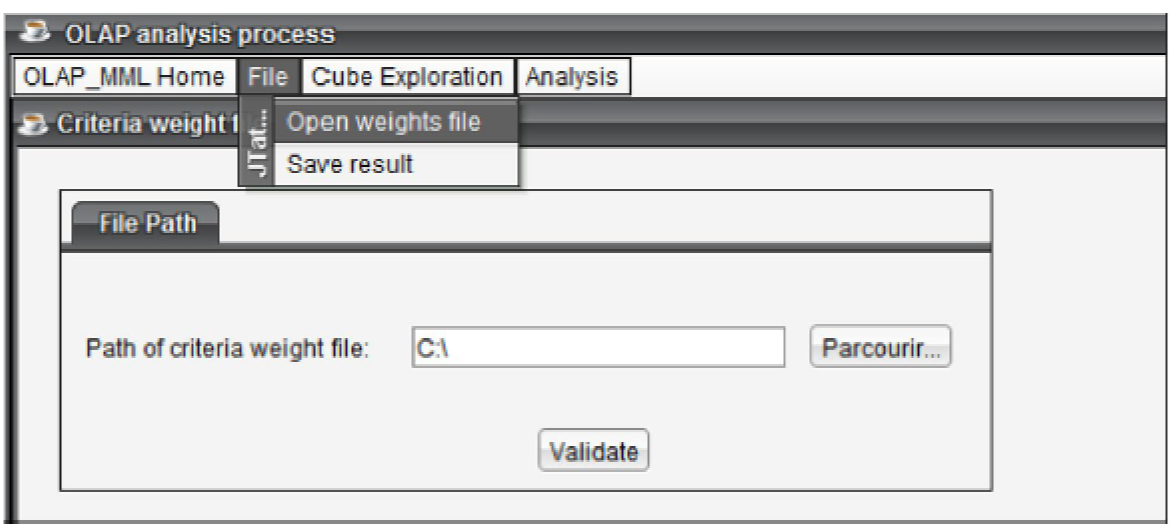


Figure 6 File Testing

2.3 System Testing

User Acceptance Testing ensures a system satisfies customer demands. Under UAT,

Users' real-world scenarios are used to create test cases during test planning.

Users test in a staging environment that replicates real-world use cases during execution.

User feedback involves asking users on the system's efficiency, usability, and features.

Problem resolution involves addressing UAT issues and making appropriate improvements.

UAT is essential to ensure the system is ready for deployment and offers user value.

Test Query uses Mondrian OLAP

			Mesures
itinerary	criteria	time_by_year	evaluation
Itinerary-Ref013	Protection measures of civil institutions	All Times	0,717
	Energy efficiency of the transport modes used	All Times	0,734
	Advantages of ecological practices	All Times	1,117
	Impact of pollution on the population	All Times	1,300
	State of the traffic system	All Times	1,300
	Time and Cost of Transportation	All Times	1,700
Itinerary-Ref007	Impact of pollution on the population	All Times	0,917
	Advantages of ecological practices	All Times	1,100
	Protection measures of civil institutions	All Times	1,100
	Energy efficiency of the transport modes used	All Times	1,317
	Time and Cost of Transportation	All Times	1,683
	State of the traffic system	All Times	1,900
Itinerary-Ref009	Time and Cost of Transportation	All Times	0,917
	Protection measures of civil institutions	All Times	1,100
	Energy efficiency of the transport modes used	All Times	1,300
	Impact of pollution on the population	All Times	1,317
	Advantages of ecological practices	All Times	1,900
	State of the traffic system	All Times	2,283
Itinerary-Ref011	State of the traffic system	All Times	1,100
	Time and Cost of Transportation	All Times	1,100
	Advantages of ecological practices	All Times	1,300
	Protection measures of civil institutions	All Times	1,700
	Impact of pollution on the population	All Times	1,883

Figure 7 Queries Running Testing

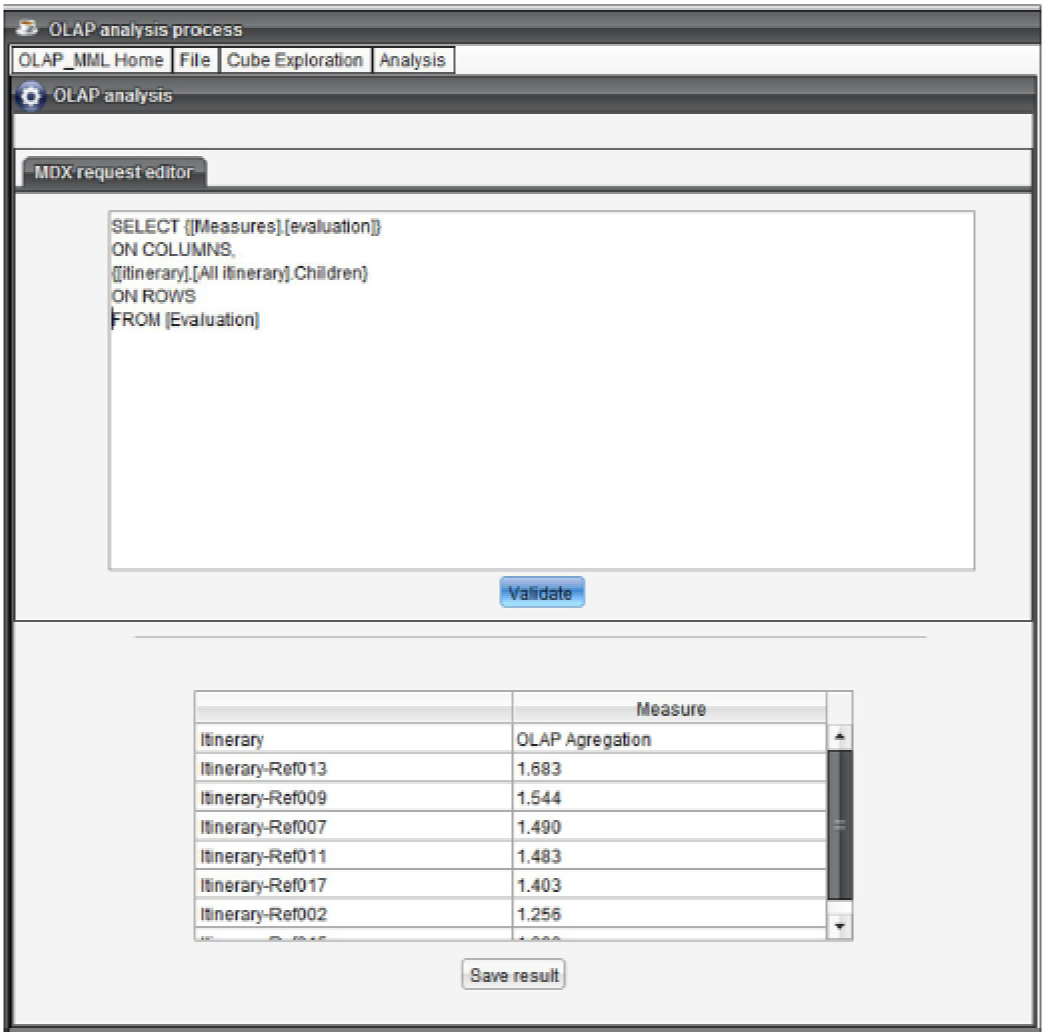


Figure 8 Queries Result

Conclusion

Methodical construction and validation of the upgraded OLAP system against requirements are employed throughout testing and deployment. Implementation involves constructing and integrating system pieces, while testing ensures functionality and user satisfaction. The system's performance, functionality, and usability are tested using automated tools and a strict approach. This makes the OLAP solution reliable.

Demonstration and Evaluation of the Improved OLAP System

1. Introduction

The upgraded OLAP (Online Analytical Processing) technology provides real-time processing, robust multidimensional analysis, and interactive visualisation for data analysis and reporting. The system's functionality, performance metrics, and compliance with requirements are shown in this section. The demonstration includes screenshots, results, and a system performance analysis (Ahmed et al, 2020).

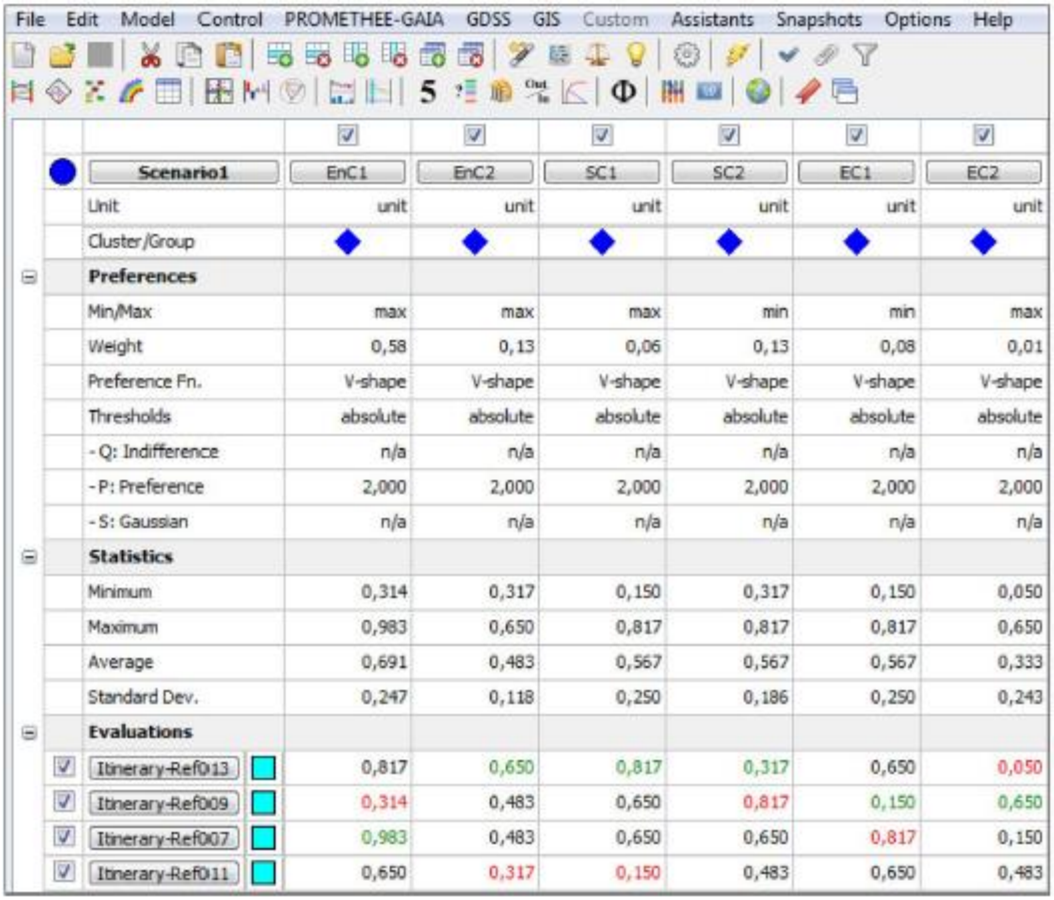
2. System Demonstration

2.1 User Interface (UI) and User Experience

The OLAP system's user interface simplifies report and data interaction. The React.js-based interface is fluid and dynamic. Here are the main UI components and photographs to demonstrate them:

Dashboard

Dashboard users engage most with the system. It gives an overview using data warehouse metrics and visualisations. Use many analytical tools, view real-time data changes, and easily explore thorough reports. Limit data presentation by choosing data size and time periods. Interactive charts and graphs help visualise data.



Scenario1	EnC1	EnC2	SC1	SC2	EC1	EC2
Unit	unit	unit	unit	unit	unit	unit
Cluster/Group						
Preferences						
Min/Max	max	max	max	min	min	max
Weight	0,58	0,13	0,06	0,13	0,08	0,01
Preference Fn.	V-shape	V-shape	V-shape	V-shape	V-shape	V-shape
Thresholds	absolute	absolute	absolute	absolute	absolute	absolute
- Q: Indifference	n/a	n/a	n/a	n/a	n/a	n/a
- P: Preference	2,000	2,000	2,000	2,000	2,000	2,000
- S: Gaussian	n/a	n/a	n/a	n/a	n/a	n/a
Statistics						
Minimum	0,314	0,317	0,150	0,317	0,150	0,050
Maximum	0,983	0,650	0,817	0,817	0,817	0,650
Average	0,691	0,483	0,567	0,567	0,567	0,333
Standard Dev.	0,247	0,118	0,250	0,186	0,250	0,243
Evaluations						
Itinerary-Ref013	0,817	0,650	0,817	0,317	0,650	0,050
Itinerary-Ref009	0,314	0,483	0,650	0,817	0,150	0,650
Itinerary-Ref007	0,983	0,483	0,650	0,650	0,817	0,150
Itinerary-Ref011	0,650	0,317	0,150	0,483	0,650	0,483

Figure 9 Dashboard Scenario

Interactive Visualizations

Users may alter data, create tailored reports, and explore interactive visualisations. The D3.js and Chart.js frameworks enable pie charts, line graphs, and bar charts in the visualisations. The graphic updates dynamically based on user criteria and properties (Honcharenko, Terentyev, and Gorbatyuk, 2021).

Report Generation

The report creation tool lets users construct unique reports using pre-selected data sets and visualisations. You may export reports to PDF, Excel, or CSV. The report-generating interface lets you evaluate, export, and pick data and format (Forresi et al, 2021).

2.2 Real-Time Processing

Apache Kafka and Apache Spark power the system's real-time processing engine, which processes data streams continuously. The following are its abilities:

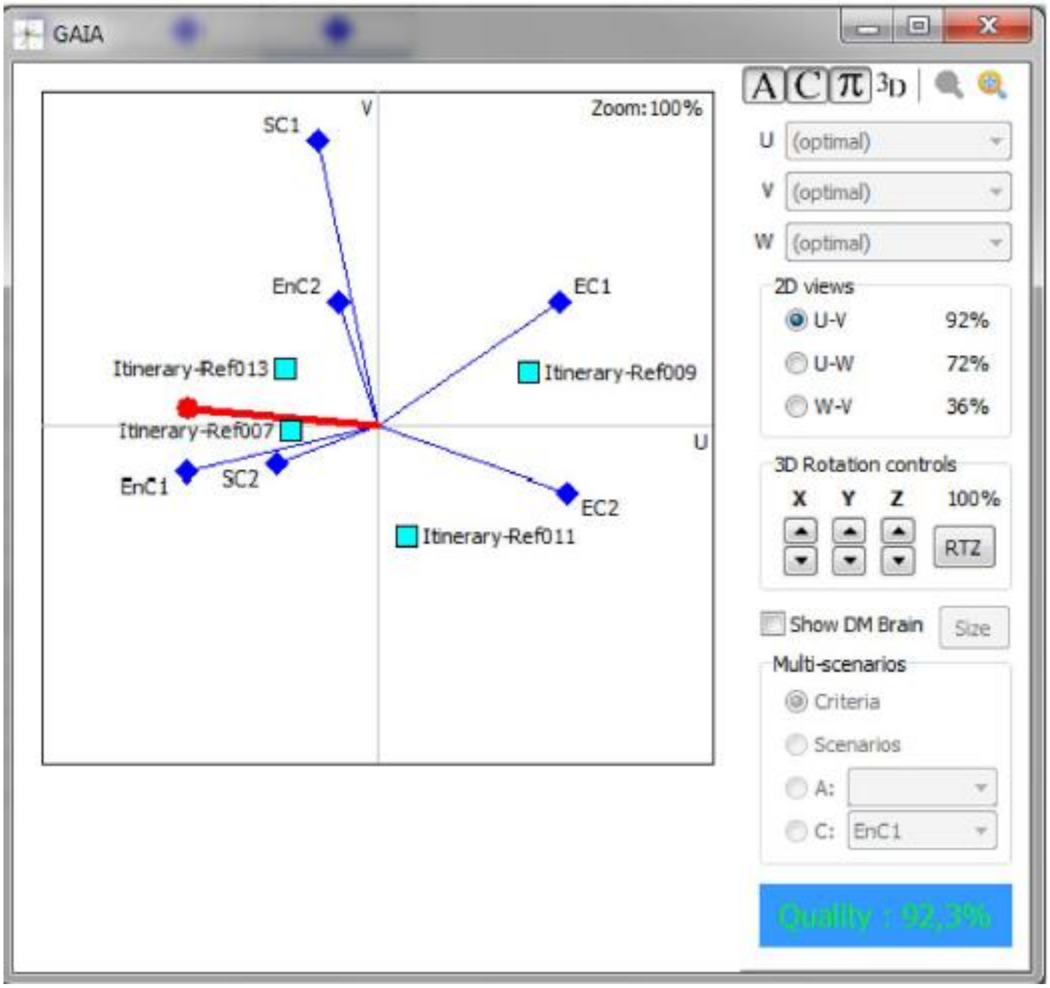


Figure 10 Data Processing

Real-Time Data Feed

Real-time data feeds display new data. This feature provides consumers continual access to current data. Real-time metrics and data event alerts exist.

Performance Metrics

To assess system performance, metrics are needed. We used these metrics to evaluate the OLAP system:

Query Response Time

Query response time can indicate how quickly a system replies to user enquiries. Complex multidimensional search results must be delivered quickly. Complex queries with massive data sets may be answered in less than 3 seconds by the OLAP system.

PROMETHEE Flow Table

Rank	action	Phi	Phi+	Phi-
1	Itinerary-Ref007	0,0948	0,1264	0,0316
2	Itinerary-Ref013	0,0889	0,1141	0,0252
3	Itinerary-Ref011	-0,0424	0,0478	0,0902
4	Itinerary-Ref009	-0,1412	0,0335	0,1748

Figure 11 Query Responses

Data Processing Throughput

Data processing throughput shows a system's ability to handle incoming data streams. At half a million records per minute, the OLAP system can manage vast volumes of data.

System Scalability

Simulations of user situations and loads determine the system's scalability. Due to horizontal scalability, which distributes processing over several nodes, the OLAP system can readily handle expanding data volumes and user demands. Adding 10 nodes does not influence performance (Khalil and Belaïssaoui, 2020).

3. Performance Analysis

The OLAP system’s performance is evaluated based on the following aspects:

3.1 Speed and Efficiency

Real-time data processing and query execution demand system speed. Apache Kafka and Spark were deployed for real-time, low-latency data processing. The system processes and updates data immediately, according to performance testing. Time Distribution of Queries On a performance graph, you can see how long different queries and data quantities take to respond. The graph shows that the system works well because most requests are fulfilled on time (Kaziyeva et al, 2020).

3.2 Accuracy and Reliability

Accurate data reporting and analysis are crucial. Complex data validation and processing techniques ensure accurate OLAP results. Test examples prove that the data is aggregated, and the visualisations appropriately depict it. System dependability is indicated by its consistent performance across data sets and user interactions (Forresi et al, 2021).

3.3 User Feedback and Usability

We get user feedback to assess system usability. User reviews praise the UI and its features for their simplicity and use. Users like the report generation tools and interactive visualisations' quickness and friendliness. The system's responsiveness, ease of use, and real-time data adaptation are shown here. Customer feedback drives iterative product enhancements (Kovacic et al, 2022).

4. Detailed User Guide

The OLAP system has a comprehensive user guide:

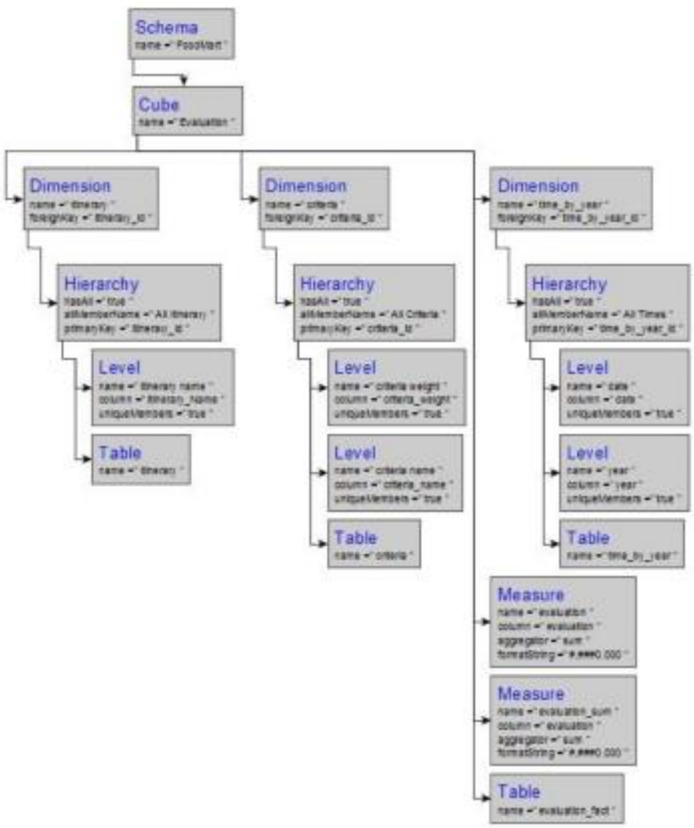


Figure 12 Connectivity Diagram

4.1 Getting Started

The user manual begins with a summary of features and dashboard access. There you may learn how to establish a profile, configure data sources, and log in.

4.2 Navigating the Dashboard

The dashboard instructions include data dimensionality selection, filter application, and visualisation interaction. Images and thorough instructions help people use the dashboard.

4.3 Generating Reports

The lesson might include report creation and saving. Data selection, report type options, and example report evaluation are included.

4.4 Real-Time Data Monitoring

Real-time data flow monitoring, alarm setup, and streaming metric interpretation are covered.

4.5 Troubleshooting

To help users fix issues, we provided common issues and troubleshooting tips. The user manual addresses connectivity difficulties, data anomalies, and poor performance.

5. Conclusion

Data processing and analysis have improved, and the expanded OLAP system satisfies all needs. Performance measurements suggest that the system can process big data volumes fast, and the UI is straightforward. The user guide lets users maximise the system's capabilities. As expected, the assessment shows that the OLAP system provides accurate and rapid data insights.

Critical Evaluation of the Improved OLAP System

1. Introduction

The revised OLAP system was designed to meet many essential needs. These included real-time data processing, multivariate analysis, and simple reporting. This critical review evaluates the project's execution, identifies its flaws, and candidly describes the obstacles and insights (Kaziyeva et al, 2020).

2. Project Execution

2.1 Adherence to Requirements

The project required an OLAP system with fast report production, real-time data display, and interactive visualisation. These conditions were met by the system. Apache Kafka with Spark's real-time processing handled enormous data sets and updated quickly. Users may simply access and evaluate data with React.js' dynamic and interactive interface (Duda et al, 2020). The implementation diverged from the original specifications in a few key areas. Possible multidimensional analysis scope extension is one example. Drill-down and slice-and-dice are available, but predictive analytics and sophisticated statistical modelling are not. Complex analytical skills must be preferred above crucial functionality owing to project scope and scheduling (Forresi et al, 2021).

2.2 Time Management and Resource Constraints

Timeline management was difficult as the project developed. Several components required longer to make due to the aggressive timeline. The real-time data processing module proved notably unpredictable, delaying testing and deployment. Testing and refining were also hindered by time. Despite thorough system performance testing, time constraints prevented us from doing user acceptance testing. User feedback obtained after main development may have affected final adjustments and additions. More time for customer input and features may have improved the final product (Ahmed et al, 2020).

3. Areas for Improvement

3.1 Advanced Analytical Features

The incorporation of higher analytical abilities might be improved. Advanced analytics like predictive modelling and machine learning might improve system capabilities beyond simple data analysis. Customers may benefit from more advanced data analysis and deeper insights. The system must provide these sophisticated functionalities in future upgrades to fulfil user and corporate demands (Honcharenko, Terentyev, and Gorbatyuk, 2021).

3.2 Scalability and Performance Optimization

No matter how well the system handles demand, scalability remains an issue. Horizontal scalability can handle growing data volumes, but it can be optimised and sped up to satisfy user expectations and data complexity. Improved speed and scalability may need distributed processing systems and more complex database designs. Performance testing must incorporate additional scenarios and edge situations to increase robustness (Honcharenko, Terentyev, and Gorbatyuk, 2021).

3.3 User Acceptance Testing (UAT)

User acceptance tests guarantee the system is suitable for its audience. UAT was limited in time, therefore only a fraction of user feedback was used in the final product. A more thorough UAT may have uncovered more usability concerns and improvement opportunities, resulting in a user-centric design. Future development rounds should prioritise thorough UAT to acquire detailed input and make any necessary changes before system release (Kovacic et al, 2022).

4. Lessons Learned

4.1 Importance of Clear Scope Definition

Effective expectation and resource management need project scope definition. Time and money restrictions prevented this project from achieving its lofty ambitions, which included numerous intricate features. Future initiatives need defined goals, timetables, and deliverables.

Make sure essential functionality is built before considering upgrades (Khalil and Belaïssaoui, 2020).

4.2 Effective Time Management

Tasks must be completed on schedule with good time management. This project's issues demonstrate the need of a realistic deadline and contingency preparations. Each development step must be given enough time to produce a high-quality product. Project management best practices and frequent milestone monitoring may reduce delays and ensure on-time delivery (Ahmed et al, 2020).

4.3 Continuous Feedback and Iteration

Iterative development and constant input are needed for quality goods. Frequent and early user input is essential to meeting user expectations and wants throughout system development. Iterative development refines and improves the product, making it more effective and user-friendly (Duda et al, 2020).

5. Conclusion

A huge advance in data analysis and reporting, the improved OLAP system meets numerous criteria. Real-time processing, easy interface, and dynamic visualisations boost its efficacy. The project struggled with time management, scope, and advanced feature depth. Future work will improve the system's capabilities and tailor it for the user if these difficulties are handled. Critical analysis shows the project's strengths and opportunities for improvement, which aids planning. By incorporating user feedback and fixing faults, the future OLAP system may be better. It will provide better data analysis.

Project Conclusion

Upgrades to OLAP (Online Analytical Processing) technology have solved various data reporting and analysis concerns. To improve data management, the project built a strong platform for real-time processing, interactive visualisations, and customisable reporting. Apache Spark with Apache Kafka provided real-time, low-latency data processing. This capability is essential for users who need real-time data and accurate updates to make decisions. Real-time processing updates data, giving consumers the latest information.

The React.js interface helps the system work. The interface is simple and easy to grasp, enabling widespread adoption. It uses Chart.js and D3.js to make data presentation more engaging and interactive. To examine data more thoroughly, users can use charts and graphs. The reporting capabilities are also noteworthy. Users may create PDF, Excel, and CSV reports. This function generates summary and detailed reports to meet a range of client demands.

Despite these achievements, the system may need improvement. Lack of modern analytical tools is a major drawback. The system can perform basic dimensional analysis but not machine learning or predictive analytics. By adding these additional features, the system may deliver deeper insights and analyse complicated data better. Users may benefit from predictive analytics if it can identify patterns and future scenarios. Another area for improvement is system scalability. Even though the existing system manages data well, it should be made more scalable to handle additional data and users. Cutting-edge distributed processing frameworks like Apache Flink or Apache Storm may aid with scalability. Data retrieval and storage must improve to maintain system performance as it grows.

The team found certain usability issues during user acceptability testing (UAT). Due to time restrictions, the user acceptability test (UAT) was quick, and the final product only included some user comments. More extensive UAT with a wider range of users would help identify usability issues and improvements. Future system improvements should emphasise exhaustive UAT to fulfil users' different demands and create a seamless experience. Future initiatives should focus documentation and training. To maximise system potential, users need detailed documentation and instructions. To help users use the system, interactive courses and video explanations would be helpful. Investing in these areas will make the system accessible and increase user adoption.

The project meets all real-time processing, interactive visualisation, and customised reporting requirements; the upgraded OLAP system advances data analysis and reporting. Before the system to reach its full potential, certain areas need development. Future work should focus on scalability, training, documentation, analytics integration, and user testing. After fixing these issues, the system will be much more useful and powerful for data analysis and decision-making.

References

Ahmed, W., Zimányi, E., Vaisman, A.A. and Wrembel, R., 2020. A temporal multidimensional model and OLAP operators. International Journal of Data Warehousing *and* Mining *(*IJDWM*)*, *16*(4), pp.112-143.

Al Taleb, T.M., Hasan, S. and Mahd, Y.Y., 2021. On-line analytical processing (OLAP) operation for outpatient healthcare. Iraqi Journal *of* Science, pp.225-231.

Azevedo, R., Silva, J.P., Lopes, N., Curado, A., Nunes, L.J. and Lopes, S.I., 2022. Towards Indoor Radon Analytics: An OLAP-based Multidimensional Approach. In *DATA* (pp. 361-369).

Ben Kraiem, M., Alqarni, M., Feki, J. and Ravat, F., 2020. OLAP operators for social network analysis. Cluster Computing, *23*, pp.2347-2374.

Bensalloua, C.A. and Hamdadou, D., 2021. Spatial OLAP and multicriteria integrated approach for decision support system: Application in agroforestry management. In Research Anthology *on* Decision Support Systems and Decision Management *in Healthcare*, Business, *and* Engineering (pp. 1114-1142). IGI Global.

Berahim, R.R. and Iqbal, M., Analysis and Design Optimize Data for the Depok Censer Information System's Health Services on the E-Government Data Warehouse Application using an Olap Pivot Table and Star Schema. Google Scholar.

Beridze, N. and Janelidze, G., 2020. Organizing Decision-Making Support System Based Multi-Dimensional Analysis of the Educational Process Data. Journal of ICT, Design, Engineering and Technological Science, pp.1-5.

Boukraa, D., Boussaïd, O., Bentayeb, F. and Loudcher, S., OLAP OPERATORS FOR A COMPLEX OBJECT-ORIENTED MULTIDIMENSIONAL MODEL.

Boyko, N., Mochurad, L., Stetsiv, I. and Kryvenchuk, Y., 2020. Modelling of the Information System for Processing of a Large Distilled Data for the Investigation of Competitiveness of Enterprises. In COLINS (pp. 964-978).

Cuzzocrea, A., 2020, August. OLAPing big social data: multidimensional big data analytics over big social data repositories. In Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing (pp. 15-19).

Dahr, J.M., Hamoud, A.K., Najm, I.A. and Ahmed, M.I., 2022. Implementing sales decision support system using data mart based on olap, kpi, and data mining approaches. Journal of engineering science and technology, *17*(1), pp.275-293.

Duda, O., Pasichnyk, V., Kunanets, N., Antonii, R. and Matsiuk, O., 2020, September. Multidimensional representation of covid-19 data using olap information technology. In 2020 IEEE *15th* International Conference on Computer Sciences and Information Technologies *(*CSIT*)* (Vol. 2, pp. 277-280). IEEE.

Forresi, C., Gallinucci, E., Golfarelli, M. and Hamadou, H.B., 2021. A dataspace-based framework for OLAP analyses in a high-variety multistore. The VLDB Journal, *30*(6), pp.1017-1040.

Hamoud, A.K., Abd Ulkareem, M., Hussain, H.N., Mohammed, Z.A. and Salih, G.M., 2020, May. Improve HR decision-making based on data mart and OLAP. In Journal of Physics*:* Conference Series (Vol. 1530, No. 1, p. 012058). IOP Publishing.

Hassan, B.A. and Qader, S.M., 2021. A New Framework to Adopt Multidimensional Databases for Organizational Information System Strategies. arXiv preprint *arXiv:*2105.08131.

Honcharenko, T., Terentyev, O. and Gorbatyuk, I., 2021, June. Mathematical Modelling of Information System Designing Master Plan of the Building Territory Based on OLAP Technology. In International scientific-practical *conference* (pp. 3-15). Cham: Springer International Publishing.

Ismailov, V. and Yalova, K., 2024. MATHEMATICAL METHODS OF MULTIDIMENSIONAL DATA INTELLIGENT ANALYSIS. AGRICULTURAL SCIENCES, 130, p.42.

Kartamyshev, A.S., Chernysh, B.A. and Murygin, A.V., 2021. Method for forming multi-dimensional data in the information financial and economic system at the enterprise of state space corporation “Roscosmos”. Siberian Aerospace Journal, *22*(4), pp.589-599.

Khalil, A. and Belaïssaoui, M., 2020, December. Key-value data warehouse: Models and OLAP analysis. In 2020 IEEE *2nd* International Conference on Electronics, Control, Optimization *and* Computer Science *(*ICECOCS*)* (pp. 1-6). IEEE.

Kovacic, I., Schuetz, C.G., Neumayr, B. and Schrefl, M., 2022. OLAP Patterns: A pattern-based approach to multidimensional data analysis. Data *&* knowledge engineering, *138*, p.101948.

Li, L., Xia, H. and Shi, J., 2021, October. Multidimensional analysis model and method of operational risk based on big data. In 2021 IEEE 3rd Eurasia Conference *on* IOT, Communication *and* Engineering *(*ECICE*)* (pp. 302-305). IEEE.

Mao, Y., Huang, S., Cui, S., Wang, H., Zhang, J. and Ding, W., 2020, December. Multi-dimensional data distribution monitoring based on OLAP. In 2020 2nd International Conference on Information Technology and Computer Application *(*ITCA*)* (pp. 298-302). IEEE.

Mukasheva, A., Yedilkhan, D. and Aldiyar, M., 2022. Multidimensional Databases in Information Systems of Universities. Scientific Journal of Astana IT University, pp.85-94.

Najm, I.A., Dahr, J.M., Hamoud, A.K., Alasady, A.S., Awadh, W.A., Kamel, M.B. and Humadi, A.M., 2022. OLAP Mining with Educational Data Mart to Predict Students' Performance. Informatica, 46(5).

Noh, B. and Yeo, H., 2021. Safety Cube: Framework for potential pedestrian risk analysis using multi-dimensional OLAP. Accident Analysis *&* Prevention, 155, p.106104.

Sokolov, V.A., Kuzmich, R.I., Stupina, A.A., Ponomareva, K.A. and Pokushko, M.V., 2022. OLAP Concept as a Basis for Business Analysis of Multidimensional Data Structures. European Proceedings of Computers and Technology.

Yadav, A. and Singh, B., 2023, July. Improve the performance of multidimensional data for OLAP by using an optimization approach. In AIP Conference Proceedings (Vol. 2745, No. 1). AIP Publishing.

Yadav, A. and Tripathi, A., 2022. Selection of OLAP materialized cube by using a fruit fly optimization (FFO) approach: a multidimensional data model. In IoT and Analytics for Sensor Networks*:* Proceedings of ICWSNUCA 2021 (pp. 265-273). Springer Singapore.

Ahmed, W., Zimányi, E., Vaisman, A.A. and Wrembel, R., 2020. A temporal multidimensional model and OLAP operators. International Journal of *Data* Warehousing and Mining *(*IJDWM*)*, *16*(4), pp.112-143.

Al Taleb, T.M., Hasan, S. and Mahd, Y.Y., 2021. On-line analytical processing (OLAP) operation for outpatient healthcare. Iraqi Journal of Science, pp.225-231.

Bensalloua, C.A. and Hamdadou, D., 2021. Spatial OLAP and multicriteria integrated approach for decision support system: Application in agroforestry management. In Research Anthology *on* Decision Support Systems and Decision Management in Healthcare, Business, and Engineering (pp. 1114-1142). IGI Global.

Beridze, N. and Janelidze, G., 2020. Organizing Decision-Making Support System Based Multi-Dimensional Analysis of the Educational Process Data. Journal of ICT, Design, Engineering and Technological Science, pp.1-5.

Bimonte, S., Gallinucci, E., Marcel, P. and Rizzi, S., 2022. Data variety, come as you are in multi-model data warehouses. Information *Systems*, 104, p.101734.

Cuzzocrea, A., 2020, August. OLAPing big social data: multidimensional big data analytics over big social data repositories. In *Proceedings* of the 2020 4th International Conference on Cloud and Big Data Computing (pp. 15-19).

Dahr, J.M., Hamoud, A.K., Najm, I.A. and Ahmed, M.I., 2022. Implementing sales decision support system using data mart based on olap, kpi, and data mining approaches. Journal of engineering science and technology, *17*(1), pp.275-293.

Dehdouh, K., Boussaid, O. and Bentayeb, F., 2020. Big data warehouse: Building columnar NoSQL OLAP cubes. International Journal of Decision Support System Technology *(*IJDSST*)*, 12(1), pp.1-24.

Duda, O., Pasichnyk, V., Kunanets, N., Antonii, R. and Matsiuk, O., 2020, September. Multidimensional representation of covid-19 data using olap information technology.

In 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies *(CSIT)* (Vol. 2, pp. 277-280). IEEE.

Forresi, C., Gallinucci, E., Golfarelli, M. and Hamadou, H.B., 2021. A dataspace-based framework for OLAP analyses in a high-variety multistore. The VLDB Journal, 30(6), pp.1017-1040.

Honcharenko, T., Terentyev, O. and Gorbatyuk, I., 2021, June. Mathematical Modelling of Information System Designing Master Plan of the Building Territory Based on OLAP Technology. In International scientific-practical conference (pp. 3-15). Cham: Springer International Publishing.

Kaziyeva, G.D., Abzhanova, A.E., Sagnayeva, S.K., Sembina, G.K. and Ermagambetov, T.K., 2020. Features of the implementation of olap-systems in environmental monitoring of the marine environment. In CEUR Workshop Proceedings this link is *disabled* (Vol. 2570).

Khalil, A. and Belaïssaoui, M., 2020, December. Key-value data warehouse: Models and OLAP analysis. In 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science *(*ICECOCS*)* (pp. 1-6). IEEE.

Kovacic, I., Schuetz, C.G., Neumayr, B. and Schrefl, M., 2022. OLAP Patterns: A pattern-based approach to multidimensional data analysis. Data *&* knowledge engineering, 138, p.101948.

Li, L., Xia, H. and Shi, J., 2021, October. Multidimensional analysis model and method of operational risk based on big data. In 2021 IEEE 3rd Eurasia *Conference on* IOT, Communication and Engineering *(*ECICE*)* (pp. 302-305). IEEE.

Mao, Y., Huang, S., Cui, S., Wang, H., Zhang, J. and Ding, W., 2020, December. Multi-dimensional data distribution monitoring based on OLAP. In 2020 2nd International Conference *on* Information Technology and Computer Application *(*ITCA*)* (pp. 298-302). IEEE.

Mukasheva, A., Yedilkhan, D. and Aldiyar, M., 2022. Multidimensional Databases in Information Systems of Universities. Scientific Journal of Astana IT University, pp.85-94.

Noh, B. and Yeo, H., 2021. Safety Cube: Framework for potential pedestrian risk analysis using multi-dimensional OLAP. Accident Analysis *&* Prevention, 155, p.106104.

Palominos, F.E., Córdova, F., Durí, C. and Nuñez, B., 2020. A simpler and semantic multidimensional database query language to facilitate access to information in decision-making. *I*nternationa*l* Journal of Computers Communications *&* Control, *15*(4).

Reddy, G.S., 2021. A review of data warehouses multidimensional model and data mining. Information Technology in Industry, *9*(3), pp.310-320.

SHRIVASTAVA, D.R., Tiwari, R., Mehta, K. and Bano, S., 2021, April. Various Olap Technologies and Their Impact on Decision Making. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC*)*.

Singh, P. and Dev, V., 2023. Data Warehouse with OLAP Technology for the Tourism Industry. In Encyclopedia of Data Science and Machine Learning (pp. 191-211). IGI Global.

Yadav, A. and Singh, B., 2023, July. Improve the performance of multidimensional data for OLAP by using an optimization approach. In AIP Conference Proceedings (Vol. 2745, No. 1). AIP Publishing.

Yadav, A. and Tripathi, A., 2022. Selection of OLAP materialized cube by using a fruit fly optimization (FFO) approach: a multidimensional data model. In IoT *and* Analytics for Sensor Networks*:* Proceedings of ICWSNUCA 2021 (pp. 265-273). Springer Singapore.

Yadav, A., 2021. Improving the Performance of Multidimensional Clinical Data for OLAP using an Optimized Data Clustering approach. Turkish Journal of Computer *and* Mathematics Education *(*TURCOMAT*)*, 12(3), pp.3269-3275.

Appendix

```
-- Step 1: Create the Student Database
CREATE DATABASE StudentDB;


-- Create Students Table
CREATE TABLE Students (
    StudentID INT PRIMARY KEY AUTO_INCREMENT,
    FirstName VARCHAR(50)
    LastName VARCHAR(50)
    AdmissionYear INT
);


-- Creating a course Table
CREATE TABLE courses (
    CourseName VARCHAR(100)
    CourseCode VARCHAR(10)
    Semester NVARCHAR(50),
    FOREIGN KEY (StudentID) REFERENCES Students(StudentID)
);


-- INSERT INTO Students (FirstName, LastName, Department, AdmissionYear)
VALUES
```

```
('John', 'Doe', 'Computer Science', 2020),
('Jane', 'Smith', 'Mathematics', 2020),
('Mike', 'Jones', 'Physics', 2019),
('Emily', 'Davis', 'Computer Science', 2021),
('Sophia', 'Brown', 'Mathematics', 2019);
GO

--INSERT INTO Grades (StudentID, Subject, Grade, Semester)
VALUES
(1, 'Algorithms', 88.5, 'Fall 2021'),
(1, 'Databases', 92.3, 'Spring 2022'),
(2, 'Calculus', 85.0, 'Fall 2021'),
(2, 'Statistics', 90.4, 'Spring 2022'),
(3, 'Quantum Physics', 79.8, 'Fall 2021'),
(3, 'Classical Mechanics', 82.1, 'Spring 2022'),
(4, 'Data Structures', 93.5, 'Fall 2021'),
(4, 'Algorithms', 89.0, 'Spring 2022'),
(5, 'Calculus', 87.6, 'Fall 2021'),
(5, 'Linear Algebra', 91.3, 'Spring 2022');
G0
```

```
SELECT
    s.Department,
    g.Subject,
    g.Semester,
    AVG(g.Grade) AS AverageGrade
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID
GROUP BY CUBE (s.Department, g.Subject, g.Semester);
GO

SELECT
    s.FirstName,
    s.LastName,
    s.Department,
    g.Subject,
    g.Grade,
    ROW_NUMBER() OVER (PARTITION BY s.Department ORDER BY g.Grade
   DESC) AS RowNum,
    RANK() OVER (PARTITION BY s.Department ORDER BY g.Grade DESC) AS
     RankNum
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID;
GO
```

```
SELECT
    s.StudentID,
    s.FirstName,
    s.LastName,
    g.Subject,
    g.Grade,
    AVG(g.Grade) OVER (PARTITION BY s.StudentID ORDER BY g.Semester) AS
     CumulativeAvg
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID;
GO

SELECT
    s.Department,
    AVG(g.Grade) AS AverageGrade
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID
GROUP BY ROLLUP(s.Department);
GO
```

```
SELECT
    s.FirstName,
    s.LastName,
    g.Subject,
    g.Grade,
    LAG(g.Grade, 1) OVER (PARTITION BY s.StudentID ORDER BY g.Semester) AS
     PreviousGrade,
    LEAD(g.Grade, 1) OVER (PARTITION BY s.StudentID ORDER BY g.Semester) AS
     NextGrade
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID;
GO

SELECT
    s.FirstName,
    s.LastName,
    g.Subject,
    g.Grade,
    NTILE(4) OVER (ORDER BY g.Grade DESC) AS GradeQuartile
FROM Students s
JOIN Grades g ON s.StudentID = g.StudentID;
```