**Data Driven Decision Making:**

**Distinctions between online transaction processing and online analytical processing**

Alejandro Garabatos Tormo, Eli Maurer, Elizabeth Kim, Prachi Shivam Parashar

Muma College of Business, University of South Florida

ISM 6021: Management Information Systems

**Abstract**

Data-driven decision-making is evolving as a crucial strategy in all organizations. We explored the significance of data-driven decision-making and how it leverages both Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) data environments to drive informed decisions. We broke down the architecture and infrastructure employed, skill sets required, along with examples to support the value that data-driven decision-making can provide. OLTP systems serve as a foundational layer of data capture and management, designed for real-time transactional operations. OLAP analyzes bigger datasets, resulting in multidimensional analysis. The symbiotic relationship between these two makes data-driven decision-making effective. While OLTP systems capture real-time data, OLAP is leveraged by decision-makers to identify trends, patterns, and a holistic understanding of the data received. A strong architectural base and infrastructure are essential for successful data-driven decision-making. Composed of several layers, architecture begins with its basic data collection and storage, usually through databases, data lakes, or clouds. These warehouses or lakes store structured and unstructured data and are critical in meeting the demands of the volume of data received. As data-driven decision-making consists of several layers, this report dissects the complex exterior into sections to understand each dimension.

*Keywords:* Data-driven decision-making, online transaction processing (OLTP), online analytical processing (OLAP), architecture, infrastructure, data capture, trends, patterns, data collection, data storage, databases, data lakes, clouds, structured data, unstructured data, semi-structured data

**Introduction**

Data-driven decision-making (DDDM) is a strategic approach for businesses today, allowing them to base their choices on factual data, metrics, and key performance indicators (KPIs). By meticulously analyzing data from sources such as market research and customer feedback, organizations can uncover valuable insights that align with their long-term vision. It involves a structured process comprising of five key steps: understanding the company's vision, identifying reliable data sources, efficiently organizing the data, conducting in-depth data analysis using tools and drawing conclusions with the help of visualization tools. DDDM is important in today's highly competitive business environment because it enhances operational excellence, pinpointing, and resolving inefficiencies and bottlenecks. It empowers informed choices in marketing, sales, product development, and customer service, boosting overall performance. Furthermore, it enables cost-effective planning through market insights and efficient resource allocation. Companies utilizing DDDM gain a competitive edge identifying opportunities, innovating, and expanding into new markets for sustained growth in a data-centric world.

One challenge in DDDM is the complexity of data. This is where OLAP and OLTP come into play. OLAP (Online Analytical Processing) optimizes analytical queries, while OLTP (Online Transaction Processing) handles transactional processing. OLAP and OLTP collaborate in DDDM by collecting transactional data with OLTP and then using OLAP to analyze, detect trends, and generate reports and dashboards for decision-makers. (Reddy et al., 2010)

**Literature review**

While Data-Driven Decision-Making is becoming widely popular in various fields, one literature review by Tanmay Sinha shares the distinct differences between the foundation of DDDM, OLTP and OLAP. Tanmay iterates the relational database incorporated within OLTP while expanding on the multidimensional database of that in the OLAP. (Sinha, 2021) The fundamentals of data warehousing and the architecture of the end-to-end process is elaborately explored in the literature review provided by Saagari et al. Further dissection of the similarities and differences are identified in their writing when comparing the pillars of OLTP and OLAP to that of Sinha’s interpretation. Saagari, and company, analyze the data content, database design, and use orientation within both OLTP and OLAP systems. (Saagari, 2013) Another interesting study provided by Dr. Samuel Conn provided in-depth analysis of OLTP and OLAP integration while examining potential usages of components of OLTP and OLAP. In this literature review, it was concluded that one data model cannot support both OLTP and OLAP environments, the feasibility of integrating OLTP and OLAP data through materialized views is increasing, and it is possible to use the same database engine for both OLTP and OLAP applications. (Conn, 2021) Clearly this topic is of great discussion through several scholarly articles and its understanding and continued learning is critical for incorporation into DDDM.

**Discussion and Findings**

Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) are the basis behind Data-Driven Decision-Making. An OLTP system typically processes many transactions like those found in daily operations. OLTP data is stored within databases and OLTP systems are commonly used for simple queries such as updating, deleting, or adding data. OLTP systems are designed to handle immense amounts of transactions by people in real time. A common example of an OLTP system at work is in purchasing an airline ticket. Within an airline’s website, one round-trip ticket from Tampa, FL to Houston, TX is selected from a list of departure flights and return flights. Passenger details and credit card information are shared and when “confirm payment” is selected, several processes occur instantaneously. The seats chosen for the flight immediately became unavailable, and the amount paid is withdrawn from the payment source. The same instant processing applies to any retail store Point of Sale (POS) system. Upon purchase of an item, that item is immediately removed from the store’s inventory and the money is withdrawn from its source to purchase the product. While the data from OLTP systems is quick and effective, the architecture of OLTP systems is row-based storage. This limits the companies’ ability to use OLTP data for business intelligence. Through an ETL (Extract, Transform, Load) process, the data within OLTP databases can be inserted into OLAP warehouses where they can more effectively be used to make data-driven decisions.

OLAP, unlike OLTP, is the storage of historical data and is meant to handle complex analytical queries in hopes of identifying trends and patterns. Additionally, as OLTP systems utilize row-based storage, OLAP systems use column-based storage. It allows data to be queried and acquired based on one dimension, rather than all the information on one transaction. Dimensions can vary from location (county, zip code, state), time (quarter, month, year, day, season), or product characteristics (toys, brand, color, adult/children). Considering the vast possibilities of dimensions, the core of most OLAP systems is the OLAP cube which is a multidimensional database that enables quicker processing than traditional relational databases.

**Types of Data Structures**

Three types of data structures currently exist: structured data, unstructured data, and semi-structured data (Vishwakarma, 2019). Structured data is typically defined as data that is organized and formatted. It is data that can exist in tables, relational databases, and has clearly defined data attributes. Examples are point-of-sale details and financial information (Alexsoft), and this type of data can be analyzed relatively quickly. Unstructured data is not organized and does not exist in any database structure because it has no rules or formats. Some examples of unstructured data are videos, images, surveys, and Word documents. The nature of this data is that collection, processing, and analysis is more difficult and time consuming (Microsoft Azure) than structured data. Semi-structured data have some elements of structure but do not exist in tables or relational databases. It is data that can contain metadata, which is data about other data–like information and data of an image, and tags that help group and describe the data. Some common examples of semi-structured data are emails and web files.

**Data warehouses vs Data Lakes**

The purpose of data warehouses and data lakes is to store and process data, and enterprise-level organizations will have both (Microsoft Azure). Data warehouses contain structured and relational data and have a very specific purpose whereas data lakes will contain structured, unstructured, and semi-structured data, both relational and non-relational. Data lakes will therefore contain big data, IoT (Internet of Things), social media, and streaming data that is meant to be accessible by data scientists and data engineers (Alex The Analyst, 2022). Data warehouses will contain transactional data, batch reporting, application, and business data that is accessed and used by business analysts.

**OLTP/OLAP in the Job Market**

There are multiple distinct roles within a company that would be critical in the data driven decision making, more specifically on the data collection and transforming that data into information so it is useful to others within the company. Some of those roles are Business Analysis, Data Analyst, Data Scientist and Data Modeler. If we look closer at the Data Modeler Job market in the US in Figure 1, we can see that the average salary is around $121,000 per year for someone with an average experience of about 5-10 years. Also, we can see that this pay varies by state within the US having NY as the top average salary state amongst Data Modelers. On the other hand, states such as FL and TX come close to the bottom of the rank, which is very curious given the number of large, well-known corporations headquartered in both states.

**Figure 1:** Data Modeler US average Salary

A map of the united states with a graph

Description automatically generated

**Summary**

The goal of this project was to define concepts of online transaction processing (OLTP) verses online analytical processing (OLAP), how the concepts applied to enterprise data, and applied to real-world scenarios. The goal was achieved through a comprehensive analysis of academic sources and industry literature to define OLTP and OLAP. The project group engaged in detailed discussions to outline the study’s scope and sequence of the study, methodology to meet project objectives, and the presentation format of project outcomes. The project was methodically executed through a well-structured division of tasks, efficient project management, and a cohesive approach to presenting the finds in the final report. This methodical approach provided a clear framework for the project’s execution. The presentation prompted additional questions regarding the practical applications of OLTP versus OLAP in enterprise systems, which are addressed in the final report. The final project report encompasses a statement of goals and objectives, a comprehensive review of academic and industry-specific literature, the practical application of these concepts in a real-world context, and conclusions related to the study’s objectives.

**Conclusion**

In conclusion, the project studied critical concepts of data driven decision making (DDDM) and the foundational technologies behind it, shedding light on strategic significance for businesses and organizations. The team’s meticulous analysis and review of academic and industry-related literature resulted in several key conclusions. First, understanding DDDM is crucial when data, metrics, and key performance indicators are used to guide strategic decisions. Thorough analysis of data from various sources yields invaluable insights. Second, the project outlined a structured framework for making data-driven decisions, beginning with a clear understanding of the organization’s vision and reliable data sources. The exploration of data sources included structured, unstructured, and semi-structured data, as well as the distinctions between data warehouses and data lakes that house these data. Third, this project studied and explored the varying roles within organizations critical to DDDM, bringing to light the significance of these roles. Understanding the job markets for these data-related roles is essential for professionals seeking employment and organizations seeking those talented in this ever-evolving field. Finally, this project's success was achieved through a methodical approach that included a structured division of tasks, efficient project management, and a cohesive presentation. The framework facilitated the project’s goals that led to valuable questions during the presentation. A comprehensive exploration and evaluation of DDDM, OLTP, and OLAP offers valuable insights into the importance of data and how it shapes business practices and strategic decision-making processes.

**References**

Alexsoft.(n.a.). (16 Apr. 2021). W*hat is OLAP: A Complete Guide to Online Analytical Processing.* Altexsoft.com. Retrieved from <https://www.altexsoft.com/blog/olap-online-analytical-processing/>

Alex The Analyst. (2022, April 26). *Database vs Data Warehouse vs Data Lake | What is the Difference?* [Video]. YouTube. <https://www.youtube.com/watch?v=-bSkREem8dM>

Conn, S. (2021). *OLTP and OLAP data integration: a review of feasible implementation methods and architectures for real time data analysis.* Usfsm. <https://www.academia.edu/479> 43403/OLTP\_and\_OLAP\_data\_integration\_a\_review\_of\_feasible\_implementation\_methods\_and\_architectures\_for\_real\_time\_data\_analysis

Kour, A. (May 2015). Data warehousing, data mining, OLAP and OLTP technologies are indispensable elements to support Decision-Making process in Industrial World. *International Journal of Scientific and Research Publications, 5*(5). <https://www.ijsrp.org/research-paper-0515.php?rp=P414064>

Microsoft Azure. (n.d.). *What is a Data Lake? Data Lake vs. Warehouse*. Azure.microsoft.com. <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-a-data-lake/#data-lake-vs-data-warehouse>

Pearlson, K. E., Saunders, C. S., & Galletta, D. F. (2019). *Managing & Using Information Systems: A Strategic Approach* (7th ed.). Hoboken, New Jersey: Wiley. (ISBN-13: 978-1-119-56056-2)

Reddy, G.Satyanarayana, Srinivasu, & Rikkula, S. R. (2010). Data Warehousing, Data Mining, OLAP and OLTP Technologies Are Essential Elements to Support. *International Journal on Computer Science and Engineering, 2*(9). https://www.researchgate.net/publication/50235148\_DATA\_WAREHOUSING\_DATA\_MINING\_OLAP\_AND\_OLTP\_TECHNOLOGIES\_ARE\_ESSENTIAL\_ELEMENTS\_TO\_SUPPORT\_DECISION-MAKING\_PROCESS\_IN\_INDUSTRIES

Saagari, S., Anusha, P. D., Priyanka, C. L., Sailaja, V. S. S. N. (2013). Data Warehousing, Data Mining, OLAP and OLTP Technologies Are Essential Elements to Support. *International Journal of Innovative Technology and Exploring Engineering*, 2(6). https://www.ijitee.org/wp-content/uploads/papers/v2i6/F0801052613.pdf

Sinha, T. (2021, March 16). *OLAP vs. OLTP: What’s the Difference?*. IBM Blog. https://www.ibm.com/blog/olap-vs-oltp/

Vishwakarma, A. (2019, April 8). *Difference between Structured, Semi-structured and Unstructured data*. GeeksforGeeks. <https://www.geeksforgeeks.org/difference-between-structured-semi-structured-and-unstructured-data/>