

Assignment 1

by Ashleigh Simons

Submission date: 22-Mar-2024 08:54AM (UTC+0000)

Submission ID: 227530384

File name: 23605_Ashleigh_Simons_Assignment_1_565908_539203235.docx (758.08K)

Word count: 4659

Character count: 28801

Module: Data Warehousing and Data Mining

Module Code: ACCA6026

Student Name: Ashleigh Simons

Student Number: 2003948

Task 1

Data Warehouse System compared to OLTP.

OLTP (Online-Transaction Processing) is a “class of software programs capable of supporting transaction-oriented applications” [1]. Companies like Cambrian Real Estate use this for transactions like listings, property management data, leases, and financial transactions. These transactions are used to store data in a “transactional database that includes insertion, deletion, and update while also supporting query requirements” [2]. They provide the necessary infrastructure to handle high volumes of transactions while ensuring data integrity, consistency and reliability.

The main difference between the OLTP system and data warehousing is that data warehousing is geared towards “business intelligence, historical data analysis, reporting, and decision making” [3]. This is beneficial as it provides a platform for strategic insights and decision making which can be used by the company to drive change and growth within the company and possibly lead to competitive advantages. Other benefits of these new systems would provide are “collaboration and global insights” [4]. This would be important as the company grows to have more estate agent locations and headquarters. Meaning, that the data will be stored together and accessible for each department within the company, providing an environment for effective collaboration whenever they should need it.

Additionally, storage will also be a benefit with data warehouses being able to be “100’s of gigabytes to petabytes” [5]. Taking advantage of the access to this additional data the company can use historical analysis and as stated before to make more “informed strategic decisions and identify opportunities and mitigate risks” [6]. Furthermore, making use of a data warehouse would also suggest the incorporation of “data quality and cleansing processes” [7]. This would ensure that the data being stored in the warehouse is accurate and reliable, leading to more meaningful insights.

Finally, implementing a data warehouse into the company might seem like a large initial investment, however, “the returns that you’ll reap in terms of business intelligence will be invaluable” [8]. This suggests that as you will be hopefully profiting more from the implemented system with smarter business decisions and lower risks being taken by the company, the initial cost and risk will be worth taking to benefit the company and its growth.

Data Warehouse Architecture

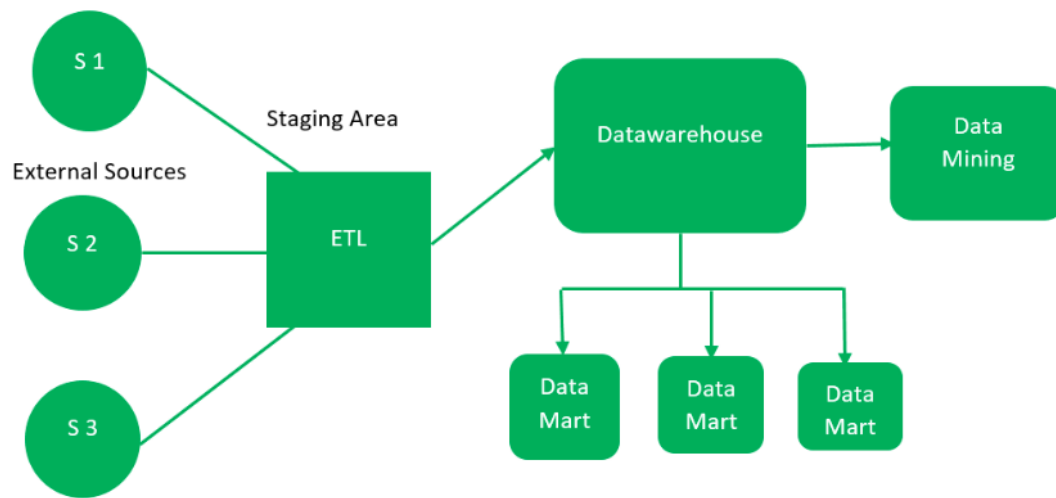


Figure 1 - The architecture of a standard data warehouse.

The figure above shows the standard architecture of a data warehouse. This data warehouse will be a centralized repository where all the data that is collected can be used for data analysis and reporting. The first stage on the far left of this diagram is the source systems. These are the internal and external data sources used to populate the data warehouse. These can be anything that provides raw data to the warehouse. Which includes, “Sales transactions, inventory transactions, customer interactions” [10] and any data within “Files, Servers, and devices” [11].

When these sources are outlined, the data is then put through an ETL process. ETL stands for “Extract, Transform, Load” [12]. This process is used to collect the data from the sources, transform it into data usable within the data warehouse and load it into it. This process can be relatively quick and sometime quite long depending on the volume of data and the complexity of the transformations that are being conducted. However, the more thorough this is, and more thought put into this process can result in data being more useful and result in the analysis being more “accurate and reliable” [13]. When this process is complete, the newly structured data can be loaded onto the data warehouse to be available to be used.

Next, data marts can be created. A data mart is “a structured data repository purpose-built to support the analytical needs of a particular department” [14]. This means you can have a data mart created for specific operations within the company and provide faster performance using data tailored to the tasks and objectives required from those positions. Finally, data mining can be conducted from the data warehouse, and this is the final section (on the right) of the

diagram. This will allow the company to identify business problems or targets and understand what is being looked for. Then data can be selected from the data warehouse required to perform analysis and prepared by data cleaning. Which is the process of “fixing or removing incorrect, corrupted, incorrectly formatted, duplicated, or incomplete data within a dataset.” [15]. This is required as if there is data which is not complete or usable, then the analysis might not be able to be performed as effectively causing errors. This can produce inaccurate results, creating unreliable analysis. Meaning that the analysis will not be used to solve the highlighted issues.

Data Marts

As mentioned in the previous section, the company could use data marts tailored to the specific needs of roles within the company. The managing director, finance director, and sales and marketing director will all be able to have data marts created to see the relevant data that they require to make well-rounded management decisions.

Finance Director -

The financial director would require data marts to include detailed financial data and analytics to monitor the company's finances and be able to track the revenue and expenses of the company to make sure that they are performing as required or if there are issues, they need to change to become more successful. There would also be marts required to track “financial transactions, budgets, forecasts, cash flow, balance sheets” [16] and other financial data that may be relevant. Being able to access all this data and information, they would be able to perform analysis and have reports on all the financial data that the company produces. This can then be used to justify decisions being made and act on them.

Sales and Marketing Director -

The sales and marketing director would also be able to make use of data marts. Accessing sales data, customers, market trends, and customer acquisition through a data mart could then provide the data required to create reports, dashboards, and conduct analysis. This is leading to the creation of sales and marketing strategies that make the company more successful within those areas. This could be done by developing targeted marketing strategies, that will then “optimize sales processes and improve the customer experience” [17].

Managing Director -

The managing director would oversee the day-to-day operations of the company. This means that they would need access to the same data marts, reports created, and analysis that the Finance and Sales and Marketing directors use as well as their own for their specific requirements. These data marts will need to provide targeted and relevant information to

support their decision-making process. Having “aggregated data, customized reporting, and real time updates, and ensuring data governance and security” [18]. These data marts will be able to support the director in making management decisions that will drive the company forward and achieve their goals.

Normalized and Dimensional Structures

Commonly in OLTP systems the structure is normalized. The difference between this and a dimensional structure of a data warehouse is that the dimensional model consists of “fact tables surrounded by dimension tables” [19]. This allows for “simpler querying for analytical purposes” [20]. This is done by denormalization and then sorting the data into groups based on their purposes, like finance, marketing, or other purposes which they may need. This is in comparison to the OLTP systems which is typically normalized and includes tables which data is organized into, with “each table attempting to reduce as much redundancy” [21]. This attempts to minimize data duplication and ensure data integrity.

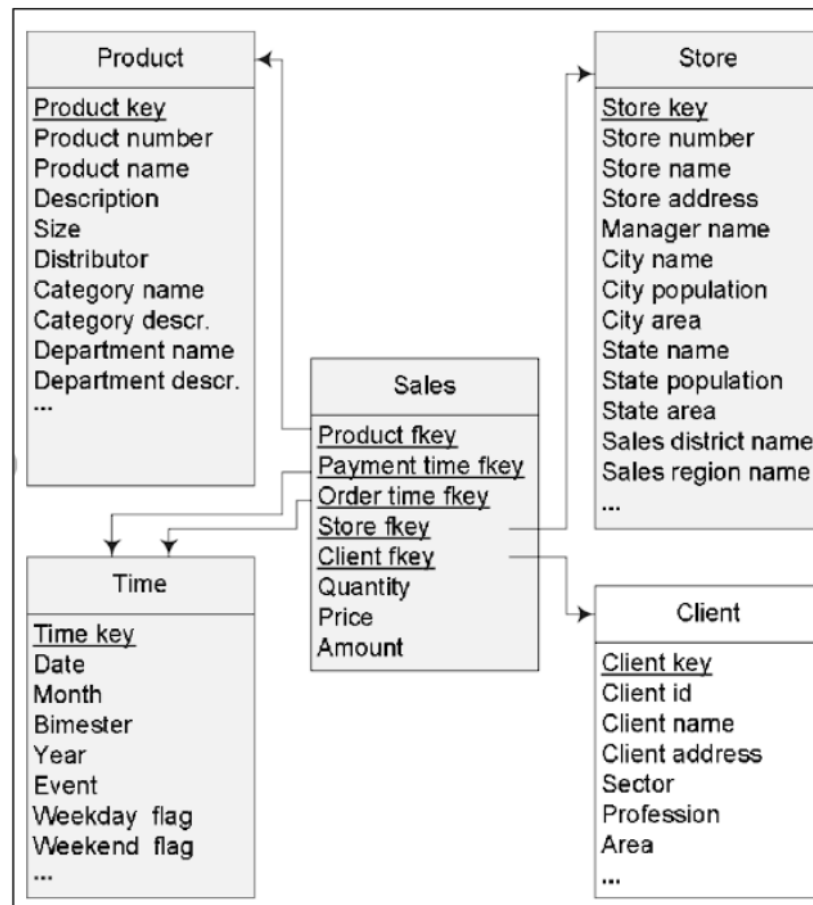


Figure 2 – Example of a Star Schema.

The diagram above shows an example of a star schema with fact and dimension tables for analyzing sales. You can see that the 'sales' table is in the middle and this is the fact table, it calls information from the dimension tables which are 'Product', 'Store', 'Time', and 'Client'. These tables will store all the data corresponding to each of the fields within the table. The 'Sales' fact then uses the fields of the dimension tables to create a fact table that can be used to analyze data within those fields. This current example would be for the sales of products; however, the same scheme can be used and implemented for the purposes required within the company. This would differentiate us from an OLTP system as the tables do not need to be connected or have relationships to each other. The fact tables only use what they are required to for the analysis, reports, and findings to be conducted. An example of how the company could make use of this is to identify market trends. Looking at the types of properties being sold and bought the prices, locations, and types of sellers to understand if there is a competitive advantage to be made by the company.

Task 2

Internal and External Source Systems

We can see that Cebrian Real Estate Company has many internal and external source systems. Some of these internal sources include databases for details on customers, property listings, lease agreements, and finances. The external sources this company has access to are market trend data, property valuation data, demographic data from many different locations, and competitor data. The company makes use of all of these currently, some being stored online, and others being stored on-site. This is where all the data warehouse data will come from, and these sources will be based on the intended structure to be implemented.

Star Schema

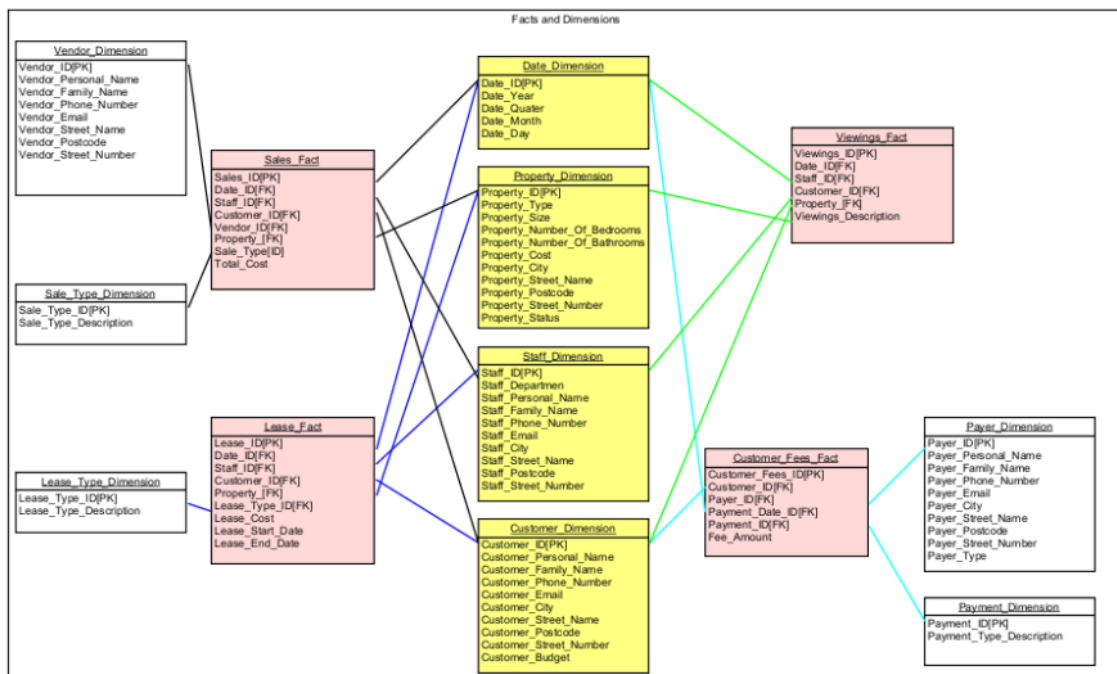


Figure 3 – Star Schema created for the Cebrian Real Estate Company

As you can see above is the star schema that has been created for the Cebrian Real Estate Company. These tables are split into four stars for each of the fact tables created through the 9-dimension tables.

Description of Facts and Dimensions Attributes

To find a complete list of attributes, their attribute types, source, and description please see appendix [1].

Meta Data

Metadata			
m_Sales	m_Lease	m_Viewings	m_CustomerFees
m_Sales_ID[PK] Table_Modified_ID Row_Modified AttributeModified_ID Old_Value New_Value Data_Source_ID User_ID Timestamp Comment	m_Sales_ID[PK] Table_Modified_ID Row_Modified AttributeModified_ID Old_Value New_Value Data_Source_ID User_ID Timestamp Comment	m_Sales_ID[PK] Table_Modified_ID Row_Modified AttributeModified_ID Old_Value New_Value Data_Source_ID User_ID Timestamp Comment	m_Sales_ID[PK] Table_Modified_ID Row_Modified AttributeModified_ID Old_Value New_Value Data_Source_ID User_ID Timestamp Comment

Figure 4 - Cebrian Real Estate Company Star Schema Metadata.

There are a number of reasons for the following meta data, some of these include “Auditing Changes, Versioning, Data Integrity, and Compliance” [23]. This allows modifications of data to be recorded down to the tables and attributes to be changed. It also allows you to see the differences in the data before and after the changes have been made so that if there are any circumstances you need to know it is stated. Furthermore, having data integrity provides recognition of changes being made and by who, this can also be useful for troubleshooting anomalies. Additionally, as they are in the estate agent's industry maintaining a record of modifications is important, especially since customer and staff data is involved. Finally, as all of these will be in place the tables will aim to improve data governance, making sure that the company is transparent and honest about and changes that are being made, when they are occurring, and who is making the changes.

Master Data

Master data	
Users	Data_Sources
User_ID[PK] User_Name User_Type	Source_ID[PK] Source_Location
Tables_Attributes	
Tables_Attributes_ID[PK] Table_Name Attribute_Name	

Figure 5 - Cebrian Real Estate Company Star Schema Metadata.

There are several reasons for having master data, some of these reasons include “Consistency, Integration, Quality, Analysis, and Business Understanding” [24]. Firstly, consistency is important as data will be integrated from a lot of different sources. So, having master tables makes sure that there is consistency across these different sources. Next master data tables provide a unified view of the businesses entities, this will all the data warehouse to integrate data more easily from these different sources and systems. Furthermore, data quality can be monitored through master data tables. Allowing this makes sure that the data being used and stored is accurate and reliable for future analysis to be conducted. Analysis can be done through these tables and the data within the other dimensions on the data warehouse as relationships can be understood and suggested, leading to discoveries in the data. Finally, business understanding can be found within these tables as their entities and attributes, describing the data with context about its operations and objectives within the company.

Aggregate Data

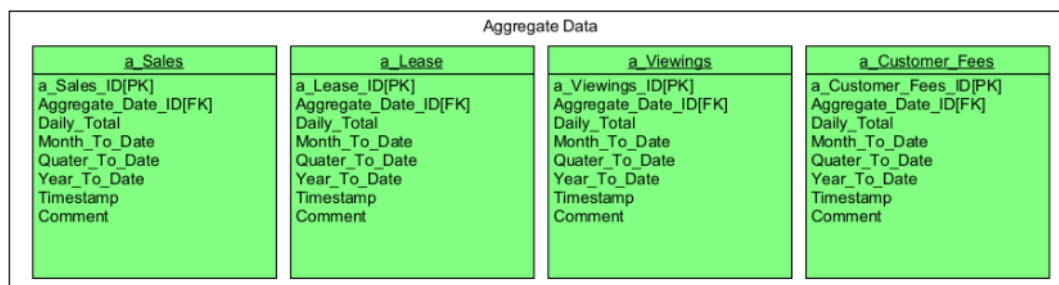


Figure 6 - Cebrian Real Estate Company Star Schema Metadata.

Within the aggregate tables are several fields and attributes that can be used to, and for, “Query Performance, Reduced Workload, Analytical Capabilities and Optimized Reporting” [25]. Firstly, query performance refers to stored data that has already been calculated by different levels, for instance in this case it is dated. But these are also months, quarters, and yearly so the data within these fields is already calculated. By doing this in advance it can lead to faster queries and be more efficient when performing complex queries as this data is already there to use. This in turn, reduces the workload of certain queries as this step has already been thought out and conducted, meaning that the queries that are run for analysis and research will be less strenuous on the system as it would not need to be calculated each time. Another reason for doing this is to enhance analytical capabilities. This suggests that potential patterns, trends, and insights might be discovered from this data, compared to just analyzing the raw data by itself. This will then lead to optimized reporting as these tables can be used for reporting and dashboards. Since these tables and data already exist, the reporting queries can be executed quickly.

Identification of Data Marts

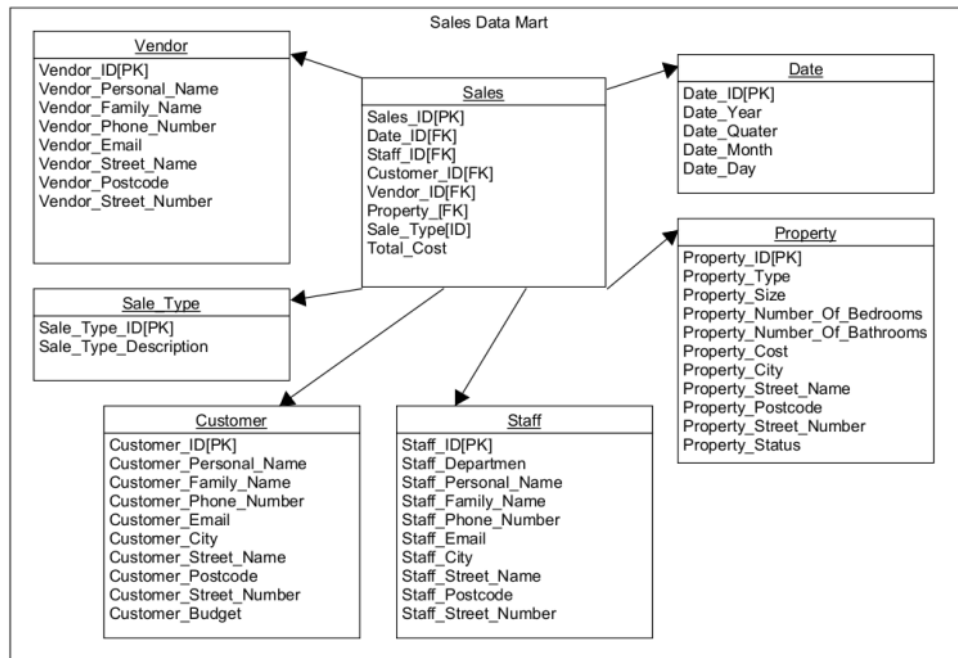


Figure 7 - The Sales Data Mart.

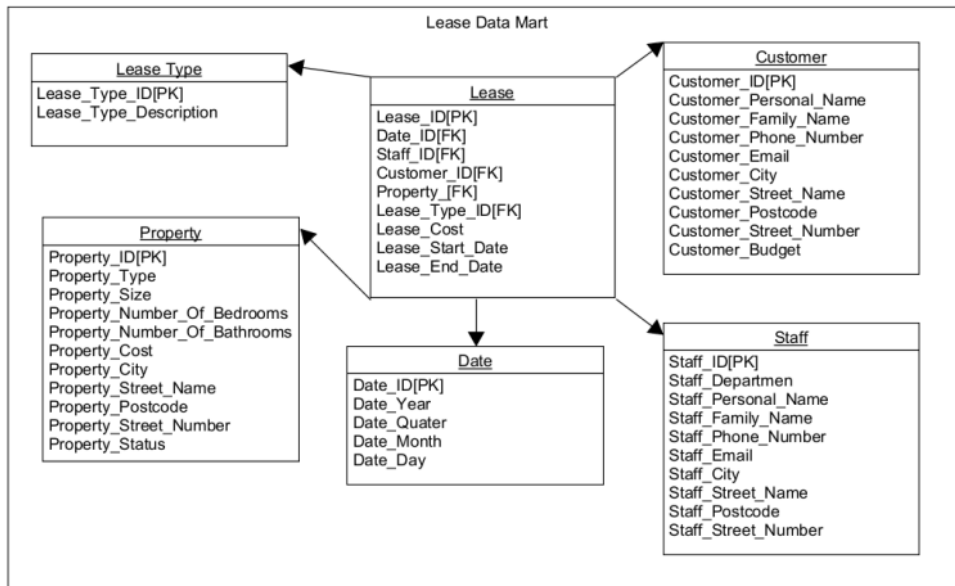


Figure 8 - The Lease Data Mart.

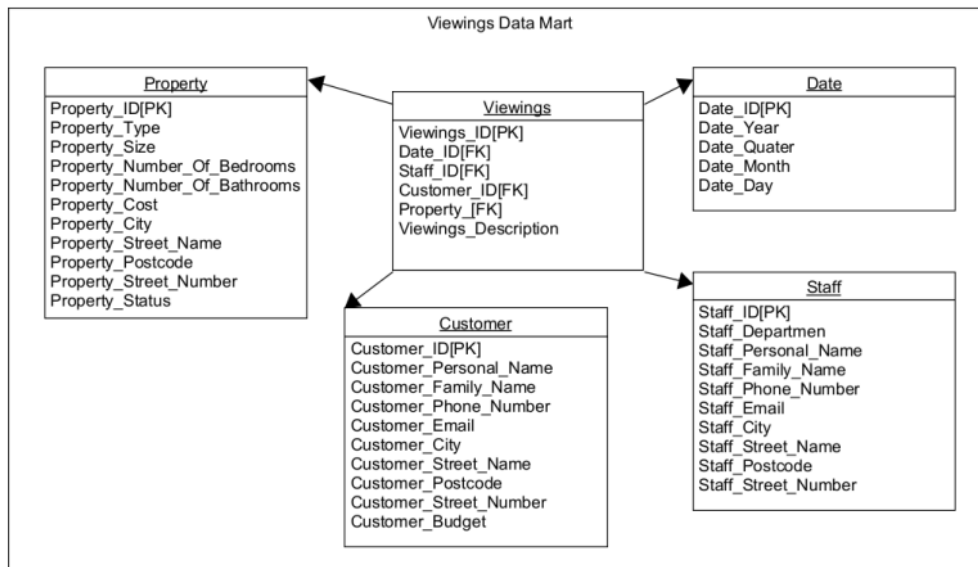


Figure 9 - The Viewings Data Mart.

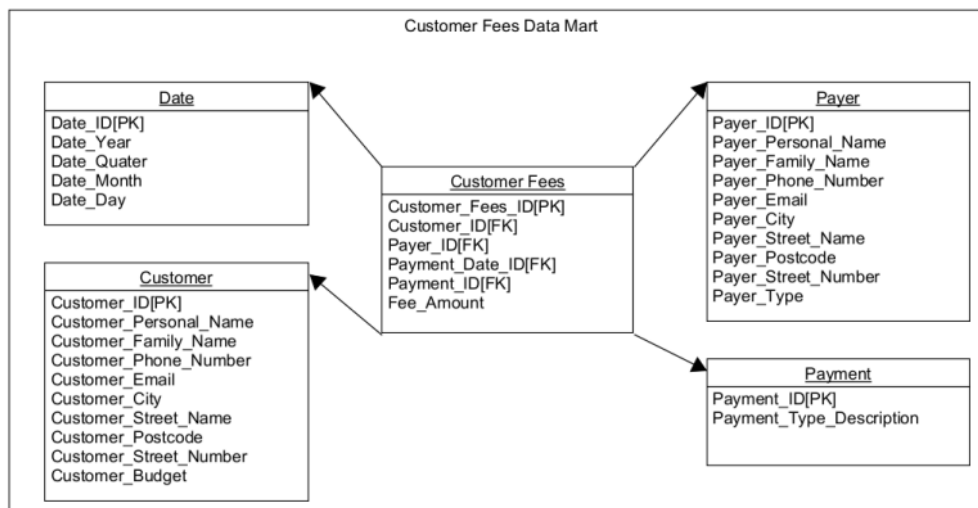


Figure 10 - The Customer Fees Data Mart.

The previous four data marts have been created for the Sales, Lease, Viewings, and Customer Fees. The goal of these data marts is to perform as a “subset of data that is designed to serve a specific function within the organization” [26]. In this case, the sales data mart would be beneficial for the sales and marketing director as they will be able to use this data mart for analysis and reporting. Some of the ways this can be used are, to show insight into the sales that are being made, highlight by which customers are making purchases and how often, how

often sales are being made generally throughout the company, and the revenue that is being made from these sales. Similarly, the sales and marketing director would also make use of the lease data mart as this would provide similar uses and outcomes. However, for the lease data instead of sales being made. Having insight into both these from their own data marts will prove a more efficient and effective way of producing reports and analysis.

The managing director would make more use of the viewings data mart, with this data surrounding customers wanting to book viewings on properties. Having this data about what customers are looking for and wanting to see could provide useful insight after analysis. Furthermore, the data could compare or differ to market trends and potentially highlight a gap or need within the market that the company can use to grow or boost revenue by focusing more time and effort on specific properties, locations, and house sizes.

The customer fees data mart would be especially useful with the financial director. Having this data mart will allow for business insights to be created as well as make sure that the company is contributing to “regularity compliance” [27]. This means that the company will be monitoring customer fees and transactions and be able to facilitate any audits or reporting processes that are required.

ETL Plan

ETL (Extract transform, and load), is the process of “combining data from multiple sources into a large central repository” [28]. This is done to clean and organize raw data and prepare it for potential use in the data warehouse. The first phase of this Plan is to “Extract” [29] data. The data extracted will be from the source systems identified in the Internal and External Source Systems section of Task 2. This will be performed from the sources they have on-site and external in the cloud. This step's main aim is to collect all the relevant data, for instance, in this case, it could be all the customer details within the customer database on site. The attributes and data held within this database will be extracted and prepared for transformation.

Once the data has been extracted and is ready, the transformation phase can start. Within this phase data from the database will be “Sorted, organized, and cleansed” [30]. This means that any duplicated entries will be deleted. For instance, if a customer is extracted twice with more than one ID for a single customer, the duplicate will be deleted. Furthermore, any missing values will be either eliminated from the database or attempted to be fixed. Finally, the data will be checked to make sure that the data being passed on to the next phase is reliable, consistent, and usable.

With the previous phase now complete, the final part of the plan can be carried out, which is the loading phase. Loading is the process of “placing data into the target system, where it is ready to be analyzed by BI tools or data analytics tools” [31]. This means that once the data is

then loaded into the data warehouse. Data mining can be conducted to produce reports, infographics, Viz, and Dashboards. These will provide insight into the data that the directors can then use to inform business decisions.

Data Quality and Information Governance Issues

The main data quality and information governance issues that need to be considered during the ETL process are “Data Accuracy, Consistency, and Reliability” [32]. This means the data should be checked properly through the ETL process as data can be inaccurate if not checked, providing less reliable data that will be used and create unreliable insights. This is also something that needs to be done consistently as if some data is reliable and others are not, the results that are found and created could cause harmful actions to be taken by the company that would damage the revenue and efficiency of day-to-day operations within the company.

For the normal operation of the data warehouse system, some things need to be considered like “monitoring, cleaning, access controls, auditing” [33]. This suggests data needs to be looked at and cleaned regularly if there are issues with new data being brought in and this will also be to avoid the previous issues within the ETL process. Restricting control of who can access the data will also benefit the company as only those with the correct purpose or need to use the data will reduce problems from occurring and prevent data from being used correctly and in a way that positive and beneficial for the company.

References

- [1] B. Lutkevich, "What is OLTP (online transaction processing)?", TechTarget Data Center, [Online], Available: <https://www.techtarget.com/searchdatacenter/definition/OLTP>
- [2] J. Ayush, "Difference between data warehousing and online transaction processing (OLTP)," GeeksforGeeks, [Online], Available: <https://www.geeksforgeeks.org/difference-between-data-warehousing-and-online-transaction-processing-oltp/>
- [3] Databricks, "What is a data warehouse?", Databricks, [Online], Available: <https://www.databricks.com/glossary/data-warehouse>
- [4] Fivetran, "Top 6 benefits of a Data Warehouse," Top 6 Benefits of a Data Warehouse, [Online], Available: <https://www.fivetran.com/learn/benefits-of-data-warehouse>
- [5] Amazon, "Data Lake vs Data Warehouse vs Data Mart - difference between cloud ...," Amazon.com, [Online], Available: <https://aws.amazon.com/compare/the-difference-between-a-data-warehouse-data-lake-and-data-mart/>
- [6] D. Tobin, "Data Warehouse vs. database: 7 key differences," Integrate.io, [Online], Available: <https://www.integrate.io/blog/data-warehouse-vs-database-what-are-the-key-differences/>
- [7] T. Naeem, "The 3 stages of Data Cleansing," Data Warehouse Information Center, [Online], Available: <https://datawarehouseinfo.com/the-3-stages-of-data-cleansing-to-ensure-data-accuracy-in-your-data-warehouse/>
- [8] A. Dearmer, "Data Warehouse Pricing & Cost: What you'll really spend," Integrate.io, [Online], Available: <https://www.integrate.io/blog/the-true-cost-of-a-data-warehouse/>
- [9] "Figure 1 - The architecture of a standard data warehouse.", GeeksforGeeks, Diagram, Geeksforgeeks.org, [Online], Available: <https://www.geeksforgeeks.org/data-warehouse-architecture/>
- [10] N. Halder, "A comprehensive guide to data warehouse architecture in 2023: Components, design, and best...", Medium, [Online], Available: <https://medium.com/@HalderNilimesh/a-comprehensive-guide-to-data-warehouse-architecture-in-2023-components-design-and-best-df7d5275ed18>
- [11] N. Halder, "A comprehensive guide to data warehouse architecture in 2023: Components, design, and best...", Medium, [Online], Available: <https://medium.com/@HalderNilimesh/a-comprehensive-guide-to-data-warehouse-architecture-in-2023-components-design-and-best-df7d5275ed18>
- [12] IBM, "What is ETL (extract, transform, load)?", IBM, [Online], Available: <https://www.ibm.com/topics/etl>
- [13] QLIK, "What is ETL? (extract, transform, load) The Complete Guide," Qlik, [Online], Available: <https://www.qlik.com/us/etl>
- [14] QLIK, "What is a Data Mart? definition, benefits, types," Qlik, [Online], Available: <https://www.qlik.com/us/data-warehouse/data-mart>

- [15] Tableau, "Guide to data cleaning: Definition, benefits, components, and how to clean your data," Tableau, [Online], Available: <https://www.tableau.com/learn/articles/what-is-data-cleaning#:~:text=tools%20and%20software-,What%20is%20data%20cleaning%3F,to%20be%20duplicated%20or%20mislabeled>
- [16] Mrseo, "What does a finance director do?", Practical CFO, [Online], Available: <https://pcfo.co.uk/post/what-does-a-finance-director-do/>
- [17] H. Higa, "The 9 goals to consider when creating a marketing strategy," HubSpot Blog, [Online], Available: <https://blog.hubspot.com/marketing/goals-of-marketing#:~:text=A%20marketing%20goal%20is%20a,essentially%20a%20waste%20of%20money>
- [18] J. Palmer, "What is a managing director? (overview, roles, and Responsibilities): Onboard board intelligence," OnBoard Board Management Software | Board Portal | Board Intelligence, [Online], Available: <https://www.onboardmeetings.com/blog/managing-director/>
- [19] Data Sleek, "Dimensional data modeling in data warehousing: Data Sleek," Data Sleek, [Online], Available: <https://data-sleek.com/blog/dimensional-data-modeling/#:~:text=Dimensional%20modeling%20is%20a%20data,used%20to%20quantify%20the%20data>
- [20] D. Taylor, "What is dimensional modeling in data warehouse? learn types," Guru99, [Online], Available: <https://www.guru99.com/dimensional-model-data-warehouse.html>
- [21] B. Scalzo, "Data modeling 101: OLTP Data Modeling, design, and normalization for the cloud," erwin Expert Blog, [Online], Available: <https://blog.erwin.com/blog/data-modeling-101-oltp-data-modeling-design-and-normalization-for-the-cloud/>
- [22] E. Malinowski, "Figure 2 – Example of a Star Schema." researchgate.net, [Online], Available: https://www.researchgate.net/figure/Example-of-a-star-schema-for-analyzing-sales_fig2_235934897
- [23] Atlan, "Metadata in Data Warehouse: 8 ways it is beneficial for you," Atlan, [Online], Available: <https://atlan.com/metadata-in-data-warehouse/#:~:text=Metadata%20in%20data%20warehouses%20is,governance%2C%20integration%2C%20and%20accessibility>
- [24] C. Stedman and J. Vaughan, "What is Master Data Management (MDM) and why is it important?", Data Management, [Online], Available: <https://www.techtarget.com/searchdatamanagement/definition/master-data-management#:~:text=Master%20data%20is%20often%20called,%2C%20products%2C%20suppliers%20and%20materials>
- [25] E. Crockett, "What is data aggregation? process, benefits, & tools," Datamation, [Online], Available: <https://www.datamation.com/big-data/data-aggregation/#mechanics>
- [26] Databricks, "What is a data mart," Databricks, [Online], Available: <https://www.databricks.com/glossary/data-mart>

[27] B. Cole, "What is regulatory compliance? - definition from techtarget.com," CIO, [Online], Available: <https://www.techtarget.com/searchcio/definition/regulatory-compliance>

[28] Amazon, "What is ETL? - extract transform load explained – aws", [Online], Available: <https://aws.amazon.com/what-is/etl/>

[29] Talend, "What is data extraction? definition and examples," Talend, [Online], Available: <https://www.talend.com/uk/resources/data-extraction-defined/>

[30] Talend, "What is data extraction? definition and examples," Talend, [Online], Available: <https://www.talend.com/uk/resources/data-extraction-defined/>

[31] Qlik, "What is ETL? (extract, transform, load) The Complete Guide," Qlik, [Online], Available: <https://www.qlik.com/us/etl>

[32] Dufrain, "Data quality in data warehousing: Challenges and Solutions," LinkedIn, [Online], Available: <https://www.linkedin.com/pulse/data-quality-warehousing-challenges-solutions-dufrain-consulting-iyode/>

[33] Dufrain, "Data quality in data warehousing: Challenges and Solutions," LinkedIn, [Online], Available: <https://www.linkedin.com/pulse/data-quality-warehousing-challenges-solutions-dufrain-consulting-iyode/>

Appendix

[1] Description of Dimensions and Facts

Dimension/ Fact Attributes	Data Type	Source	Description
Vendor_ID[PK]	Numeric	Customers Database	ID to uniquely Identify the vendor.
Vendor_Personal_Name	Text/String	Customers Database	The vendor's first name.
Vendor_Family_Name	Text/String	Customers Database	The vendor's surname.
Vendor_Phone_Number	Text/String	Customers Database	The vendor's phone number.
Vendor_Email	Text/String	Customers Database	The vendor's email address.
Vendor_Street_Name	Text/String	Customers Database	The street name where the vendor lives.
Vendor_Postcode	Text/String	Customers Database	The postcode where the vendor lives.
Vendor_Street_Number	Text/String	Customers Database	The street number of the property where the vendor lives.
Customer_ID[PK]	Numeric	Customers Database	ID to uniquely Identify the customer.
Customer_Personal_Name	Text/String	Customers Database	The customer's first name.
Customer_Family_Name	Text/String	Customers Database	The customer's surname.
Customer_Phone_Number	Text/String	Customers Database	The customer's phone number.
Customer_Email	Text/String	Customers Database	The customer's email address.
Customer_City	Text/String	Customers Database	The city where the customer lives.
Customer_Street_Name	Text/String	Customers Database	The street name where the customer lives.

Customer_Postcode	Text/String	Customers Database	The post code where the customer lives.
Customer_Street_Number	Text/String	Customers Database	The street number of the property where the customer lives.
Customer_Budget	Text/String	Customers Database	The budget that the customer has.
Property_ID[PK]	Numeric	Property Listings Database	ID to uniquely identify the property.
Property_Type	Text/String	Property Listings Database	The type of property that is being sold or leased.
Property_Size	Text/String	Property Listings Database	The size of the property.
Property_Number_Of_Bedrooms	Text/String	Property Listings Database	The number of bedrooms that the property has.
Property_Number_Of_Bathrooms	Text/String	Property Listings Database	The number of bathrooms that the property has.
Property_Cost	Text/String	Property Listings Database	The cost of the property.
Property_City	Text/String	Property Listings Database	The city where the property is.
Property_Street_Name	Text/String	Property Listings Database	The street name where the property is.
Property_Postcode	Text/String	Property Listings Database	The post code of the property.
Property_Street_Number	Text/String	Property Listings Database	The street number of the property.
Property_Status	Text/String	Property Listings Database	The status on if the property is sold, leased or neither.
Staff_ID[PK]	Numeric	Staff Database	ID to uniquely identify the staff member.
Staff_Personal_Name	Text/String	Staff Database	The staff member's first name.
Staff_Family_Name	Text/String	Staff Database	The staff member's surname.

Staff_Phone_Number	Text/String	Staff Database	The staff member's phone number
Staff_Email	Text/String	Staff Database	The staff member's email address.
Staff_City	Text/String	Staff Database	The city where the staff member lives.
Staff_Street_Name	Text/String	Staff Database	The staff member's street name.
Staff_Postcode	Text/String	Staff Database	The staff member's postcode.
Staff_Street_Number	Text/String	Staff Database	The staff member's home number.
Payer_ID[PK]	Numeric	Customers Database	ID to uniquely identify the paying customer.
Payer_Personal_Name	Text/String	Customers Database	The payer's first name.
Payer_Family_Name	Text/String	Customers Database	The payer's family name.
Payer_Phone_Number	Text/String	Customers Database	The payer's phone number.
Payer_Email	Text/String	Customers Database	The payer's email address.
Payer_City	Text/String	Customers Database	The city where the property is where the payer lives.
Payer_Street_Name	Text/String	Customers Database	The property street name where the payer lives.
Payer_Postcode	Text/String	Customers Database	The property postcode where the player lives.
Payer_Street_Number	Text/String	Customers Database	The property number where the payer lives.
Payer_Type	Text/String	Customers Database	Description of the type of payer that is being used to pay fees.
Payment_ID[PK]	Numeric	Finance Database	ID to uniquely identify the payment.

Payment_Type_Description	Text/String	Finance Database	A description of the payment type that has been used.
Sales_Type_ID[PK]	Numeric	Finance Database	ID to uniquely Identify the sale.
Sales_Type_Description	Text/String	Finance Database	A description of the type of sale that has been made.
Lease_Type_ID[PK]	Numeric	Finance Database	ID to uniquely Identify the lease.
Lease_Type_Description	Text/String	Finance Database	A description of the type of lease that has been chosen.
Date_ID[PK]	Text/String	Defined Generated Script/ Calendar	ID to uniquely Identify the current date.
Date_Year	Text/String	Defined Generated Script / Calendar	Use for handing and organizing data by year.
Date_Quarter	Text/String	Defined Generated Script / Calendar	Use for handing and organizing data by the quarter of the year.
Date_Month	Text/String	Defined Generated Script / Calendar	Use for handing and organizing data by the month of the year.
Date_Day	Text/String	Defined Generated Script / Calendar	Use for handing and organizing data by day of the week in the year.
Sales_ID[PK]	Numeric	Property Listings Database	ID to uniquely Identify the sale.
Date_ID[FK]	Numeric	Date Dimension	ID to uniquely Identify the current date.
Staff_ID[FK]	Numeric	Staff Dimension	ID to uniquely Identify the staff member.

Customers_ID[FK]	Numeric	Customers Dimension	ID to uniquely Identify the customer.
Vendor_ID[FK]	Numeric	Vendor Dimension	ID to uniquely Identify the vendor.
Property_ID[FK]	Numeric	Property Dimension	ID to uniquely Identify the property.
Sale_Type_ID[FK]	Numeric	Sale_Type Dimension	ID to uniquely Identify the sale.
Total_Cost	Text/String	Property Listings Database	The total cost of the sale that is being made.
Lease_ID[PK]	Numeric	Property Listings Database	ID to uniquely Identify the lease.
Date_ID[FK]	Numeric	Date Dimension	ID to uniquely Identify the current date.
Staff_ID[FK]	Numeric	Staff Dimension	ID to uniquely Identify the staff member.
Customers_ID[FK]	Numeric	Customers Dimension	ID to uniquely Identify the customer.
Property_ID[FK]	Numeric	Property Dimension	ID to uniquely Identify the property.
Lease_Type_ID[FK]	Numeric	Lease_Type Dimension	ID to uniquely Identify the lease.
Lease_Cost	Text/String	Property Listings Database	The cost of the lease that is being made.
Lease_Start_Date	Text/String	Lease Agreement Database	The date on which the lease starts.
Lease_End_Date	Text/String	Lease Agreement Database	The date on which the lease ends.
Viewings_ID[PK]	Numeric	Customer Database	ID to uniquely Identify the viewing.
Date_ID[FK]	Numeric	Date Dimension	ID to uniquely Identify the current date.
Staff_ID[FK]	Numeric	Staff Dimension	ID to uniquely Identify the staff member.

Customer_ID[FK]	Numeric	Customer Database	ID to uniquely Identify the customer.
Property_ID[FK]	Numeric	Property Listings Database	ID to uniquely Identify the property.
Viewings_Description	Text/String	Customer Database	A brief description of the property viewing being made and by who.
Customer_Fees_ID[PK]	Numeric	Customer Database	ID to uniquely Identify the customer fee.
Cutsomer_ID[FK]	Numeric	Customer Database	ID to uniquely Identify the customer.
Payer_ID[FK]	Numeric	Payer Dimension	ID to uniquely Identify the paying customer.
Payment_Date_ID[FK]	Numeric	Date Dimension	ID to uniquely Identify the current date.
Payment_ID[FK]	Numeric	Payment Dimension	ID to uniquely Identify the payment.
Fee_Amount	Text/String	Property Listings Database	Text to know the amount being paid.

Assignment 1

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

74/100

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

