

in collaboration with



## Data Management techniques and analytical approaches to improve sales in supermarkets.



**Submitted By:**

Suyesh Bhukaju Shrestha

Coventry:10175039

BSC (HONS) Computing

**Submitted To**

Mr. Manoj Shrestha

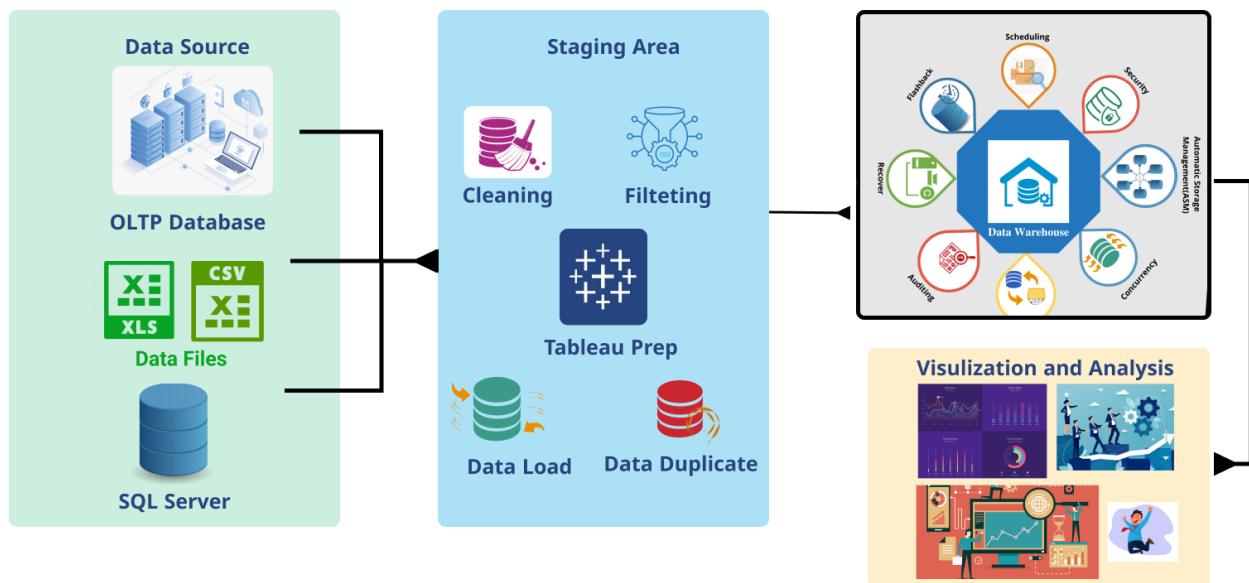
Project Supervisor

# Data Management technique and analytical approaches

To improve sales in supermarket

Suyesh Bhukaju Shrestha

10175039



## **Abstract**

Nowadays, information is a major source of competitive advantage in the commercial sector. In a similar vein, developments in technology and declining costs make it possible to capture enormous data sets. Large volumes of customer information are collected by the majority of supermarket. According to the findings, many companies have benefited from data management and analytical tools, which have helped them increase profits and obtain new insights. This dissertation focuses on data management and analytical strategies to improve sales of supermarket business. Oracle 11g is the most practical data management technology, and Tableau is the best data visualization solution for displaying charts in dashboards to apply relevant analytical tools and get business insights for informed decision making.

Further, an agile development approach was used for creating the product, while a desk-based research strategy was employed to write the literature review and a crisp-dm data strategy was used for data-mining. The various forms of analytics technique are outlined, and then the methods, platforms, and software packages used in this research are detailed. The Oracle data warehouse capabilities are also covered in this study. In the same vein, this study defines oracle security, backup, recover, and autonomous storage management approaches that may be used in the system to bolster data protection. This analysis demonstrates many types of visualization and dashboards that might be used in the decision-making process. The methods used to create this system and its effectiveness are outlined in this study.

# Table of Contents

Abstract.....	3
Keyword.....	9
Introduction.....	10
Aim .....	11
Objectives .....	12
Justification.....	12
Research question .....	14
Scope.....	15
Ethical consideration.....	16
Literature Review.....	17
Desk-based Research .....	17
Walmart .....	17
7-Eleven .....	19
Amazon .....	20
Development methodology .....	22
Agile.....	22
Why Agile? .....	23
Satisfied customers .....	24
Improve Quality.....	24
Predictability .....	24
Flexibility .....	24
Reduce risk.....	24
Data Analytics Methodology .....	24
Crisp –DM “Cross-Industry Standard Process for Data Mining Methodology” .....	24
Understanding the Business .....	25
Data Understanding .....	25
Data Preparation.....	26
Modeling.....	26
Evaluation .....	26
Deployment.....	26
Type of analytics.....	27
Descriptive data analytics .....	27

Diagnostic Data analytics .....	28
Predictive data analytics .....	28
Prescriptive data analytics.....	28
Tools .....	29
Technology .....	30
Techniques .....	31
Data warehouse.....	31
Extra, Transform, and Load (ETL) .....	32
Oracle Linux .....	33
Automatic Storage Management (ASM) .....	34
Database Security.....	36
User Security.....	36
System Security .....	39
Auditing .....	39
Backup .....	40
Recovery Manager (RMAN) .....	40
Whole Backup.....	40
Full Backup.....	40
Increment Backup .....	41
Physical backup .....	42
Logical Backup.....	42
Virtual Private Catalog (VPC) .....	42
Recovery .....	43
Flashback Technology .....	43
Data Concurrency .....	44
Database Maintenances.....	45
Automatic Workload Repository .....	45
Integration .....	46
Data Visualization.....	46
Executive Dashboard .....	47
Decision Support Dashboard .....	47
Operational Support Dashboard.....	47
Finding and Discussion.....	55

How data analytics is useful for supermarket? .....	55
How does data visualization help to take a better decision and improve business performance? .....	56
What are the ethical issue related to the application of Hardware Retail Information System in retail sector? .....	56
Project Plan .....	57
Issue Log.....	60
Conclusion .....	60
Future Work.....	61
Bibliography .....	63
Appendix.....	68

## Table of Figure

Figure 1: Real Time Customer Engagement.....	<b>Error! Bookmark not defined.</b>
Figure 2: Aim.....	11
Figure 3: Objective .....	12
Figure 4: Problem and Solution .....	13
Figure 5: Scope .....	15
Figure 6: Ethical Consideration.....	16
Figure 7: Walmart .....	18
Figure 8:7-Eleven.....	19
Figure 9: Customer Relation Management System .....	<b>Error! Bookmark not defined.</b>
Figure 10: Amazon .....	21
Figure 11: Agile .....	22
Figure 12: Benefit of Agile .....	23
Figure 13: Crisp-DM.....	25
Figure 14: Type of Data Analysis .....	27
Figure 15: Tools .....	29
Figure 16: Technology .....	30
Figure 17: Techniques.....	31
Figure 18: Data Warehouse .....	32
Figure 19: ETL Process .....	33
Figure 20: Oracle Linux.....	34
Figure 21: Automatic Storage Management .....	35
Figure 22: Data Security .....	36
Figure 23: User Security .....	37
Figure 24: Privilege.....	<b>Error! Bookmark not defined.</b>
Figure 25: System security.....	39
Figure 26: Full Backup .....	41
Figure 27: Increment Backup.....	41
Figure 28: Virtual Private Catalog.....	42
Figure 29: Flashback Data Archive.....	43
Figure 30:Database Maintenance.....	45
Figure 31: Dashboard 1.....	48
Figure 32: Dashboard 2.....	49
Figure 33: Dashboard 3.....	50
Figure 34:Dashboard 4.....	51
Figure 35: Dashboard 5.....	52
Figure 36: Dashboard 6.....	53
Figure 37: Dashboard 8.....	54
Figure 38: Dashboard 7.....	54
Figure 39: Dashboard 10.....	55
Figure 40: Project Plan 1.....	57
Figure 41: Project Plan 2.....	58
Figure 42: : Project Plan 3 .....	58
Figure 43: Project Plan 4.....	59

Figure 44: Project Plan 5.....	59
Figure 45: Issue Log .....	60
Figure 46: Future Works .....	62

## Keyword

REAL TIME DATA  
DASHBOARD DIAGNOSTIC ANALYSIS  
DATA WAREHOUSE DATA CLEANING  
AGILE METHODOLOGY ORACLE VISUALIZATION CRISP-DM  
**SUPERMARKET INTELLIGENCE SYSTEM**  
DECISION MAKING ETHIC TABLEAU  
ANALYTIC TABLEAU PREP  
PRESCRIPTIVE ANALYSIS RESEARCH  
DATA MANAGEMENT

## Introduction

One of the largest and most rapidly expanding industries is the supermarket business. Data and information are produced in large quantities by this sector. Data storage, data recovery, system security, user security, and analysis are all problems that plague many supermarkets. Data storage has grown more challenging in the supermarket sector due to the volume and variety of data created regarding customers, items, transactions, etc. Difficult or troublesome data storage may lead to system crashes and file corruption due to increased system complexity and demand. In the corporate world, that information is analyzed and visualized. The inability to reliably save data raises the possibility of issues in the analysis and display of marketing data, which in turn raises the risk of financial loss for businesses. Due to insufficient system and user protection, the supermarket's data has been compromised. Many supermarkets lacked high-security systems and their customers' safety because of the prohibitive expense of such measures. This is the main issue with data hacking and loss in grocery stores. There is a greater possibility of company failure, decreased revenue, and a tarnished image in the market as a whole. Although it's true that humans may make errors even when a recovery backup mechanism is in place, that's not always the case. When information is erased by accident, it is gone forever. It might cause issues in the framework, analysis, and presentation of data. Sometimes supermarkets make poor choices and use flawed logic. Supermarkets don't do it since it may be complicated and the study of costs might be expensive, so making the choice is tough.

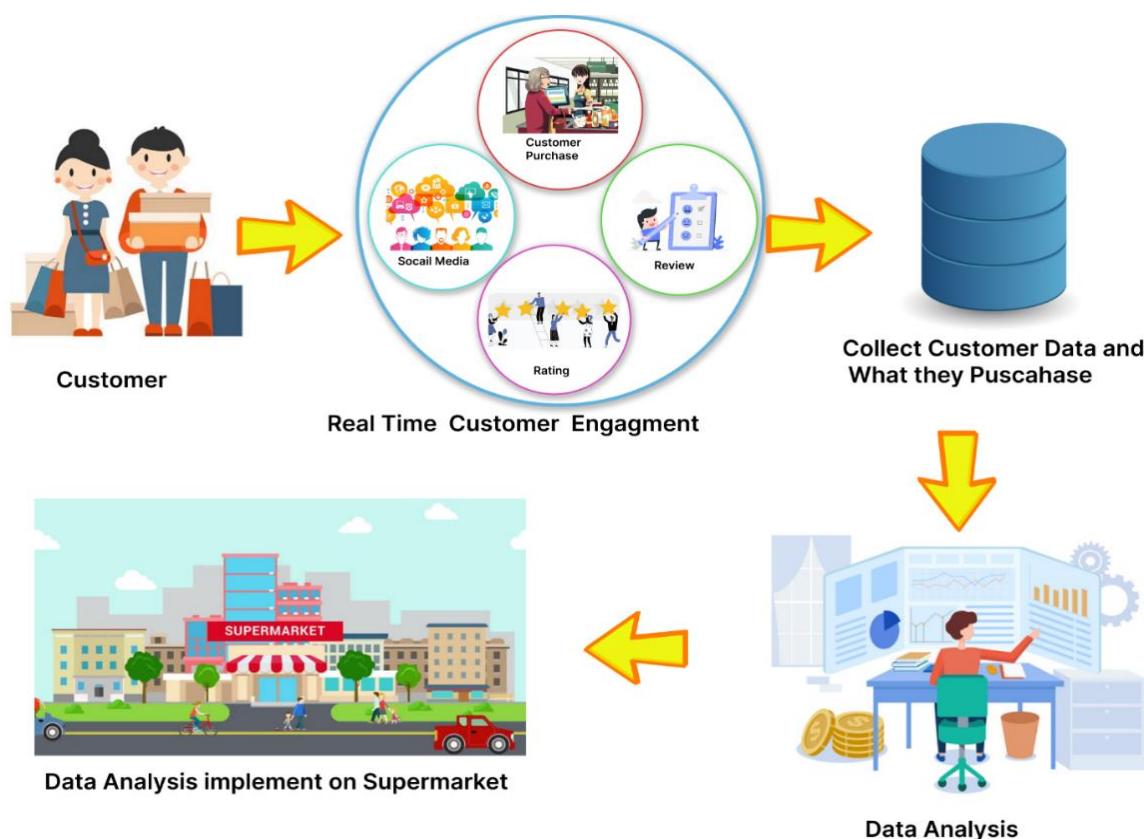


Figure 1: Real Time Customer Engagement

However, supermarkets are attempting to enhance their company and overcome these difficulties. The use of ASM (automatic storage management) is recommended. Data from databases is replicated over many hard drives, and a storage space is created and managed. It creates disk groups where they are required and removes disks mechanically. Logical administration of drives and file systems is unnecessary. System security and user security should be established for a secure system and data. System administrators have implemented security measures such as password files and operating system permissions and privileges to restrict access to authorized users. Users will have the option to restrict their own access to sensitive data and systems by limiting their system, object, and role privileges. Password complexity verification, password again and expiration, account lockout (wrong password more than 3 times), and password history will be established for increased user and profile safety. Data loss may be remedied with the use of the Undo administration system. Wherever this operation modifies database information, roll back to a prior backup. Deleted data can readily retrieved with the use of read-consistent and flashback queries, which undo enables.

Australian grocery chain Coles sells its wares online in addition to in-store. In order to maintain and improve their business across all of their platforms, these organizations depend significantly on data analytics-driven decisions. It analyzes massive amounts of data collected by the corporation from any and all sources to boost efficiency across the board. By examining more than 2500 data points, Coles is able to provide its suppliers unprecedented insight into their goods' performance. With over two million employees and 20,000 stores in 28 countries, Walmart is the largest business in the world based on sales. Walmart recognizes the importance of studying real-time data in shaping business outcomes.

## Aim



**Aim**  
**Data Management techniques and analytical approaches to improve sales in supermarkets.**



*Figure 2: Aim*

# Objectives

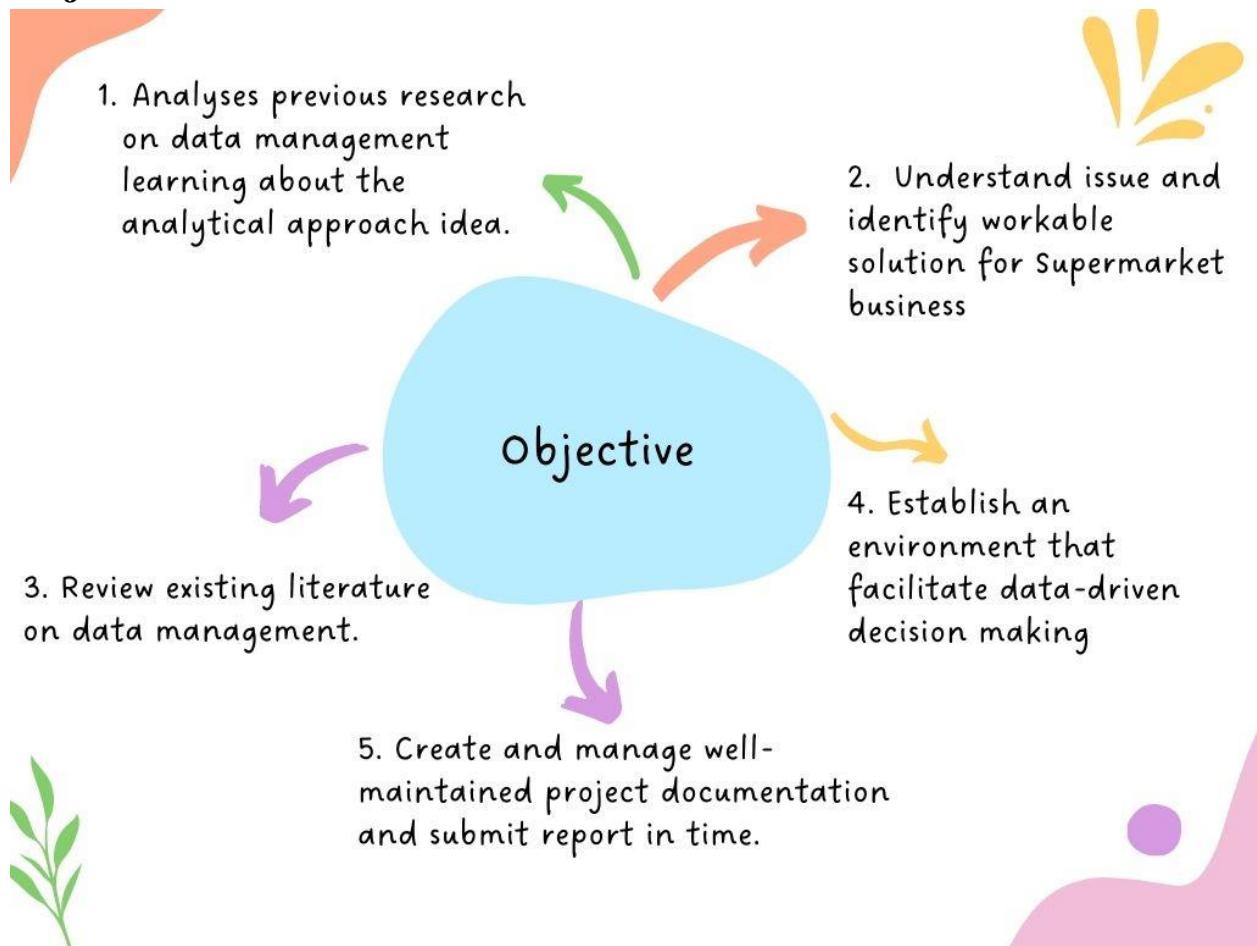


Figure 3: Objective

## Justification

Supermarkets are one of the world's fastest-growing industries. It generates huge amounts of data about consumers, items, users, and workers. Most supermarkets have cloud data and hardware difficulties. Many stores have historically relied on text-based or antiquated methods of data storing like Excel or spreadsheets. Many grocery stores have security holes in their systems or with their customers, making others vulnerable to attack and damage via compromised information. Despite producing copious amounts of data, these sectors have made zero measures toward implementing effective data management. Due to this, the information generated by these sectors cannot be relied upon. Inaccurate information leads to bad choices and poor results, which in turn causes problems for the business.

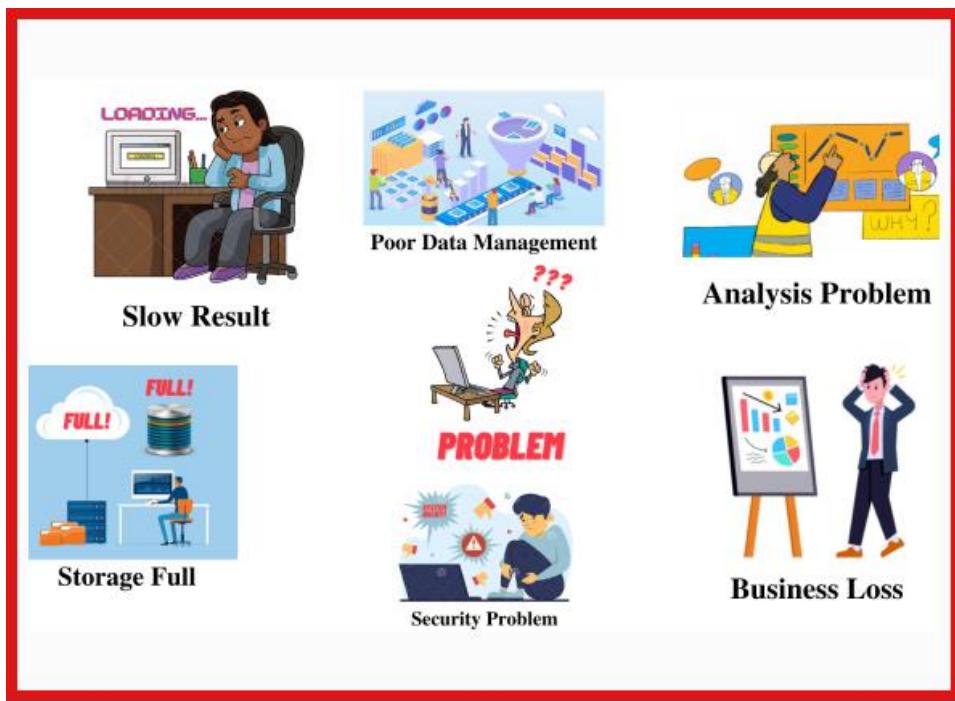


Figure 4: Problem and Solution

Any issue may be fixed with the help of Oracle's data warehouse technology. The massive data volumes produced by this sector need for ASM systems capable of handling these enormous files. Wherever it is required to do so, it builds and maintains a disk grid, deletes the disk, and redistributes the data automatically. Oracle Security provides features for both system and user security, including authentication of users, assignment of privileges and roles, password complexity verification, password aging and

expiration, account locking, and a password history. Data corruption or loss may be remedied by using an undo management system, which ensures that old versions of data are stored again in undo data whenever a method modifies data. With a flashback strategy, it's possible to retrieve lost information. Data may be appropriately evaluated and presented using Tableau. It does data analysis and displays the findings in graphical form. By analyzing and visualizing data, you may get insight into target market, product offerings, business performance, and the market's potential for growth.

## **Research question**

- How data analytics is useful for supermarket?
- How does data visualization help to take a better decision and improve business performance?
- What are the ethical issue related to the application of Hardware Retail Information System in retail sector?

# Scope

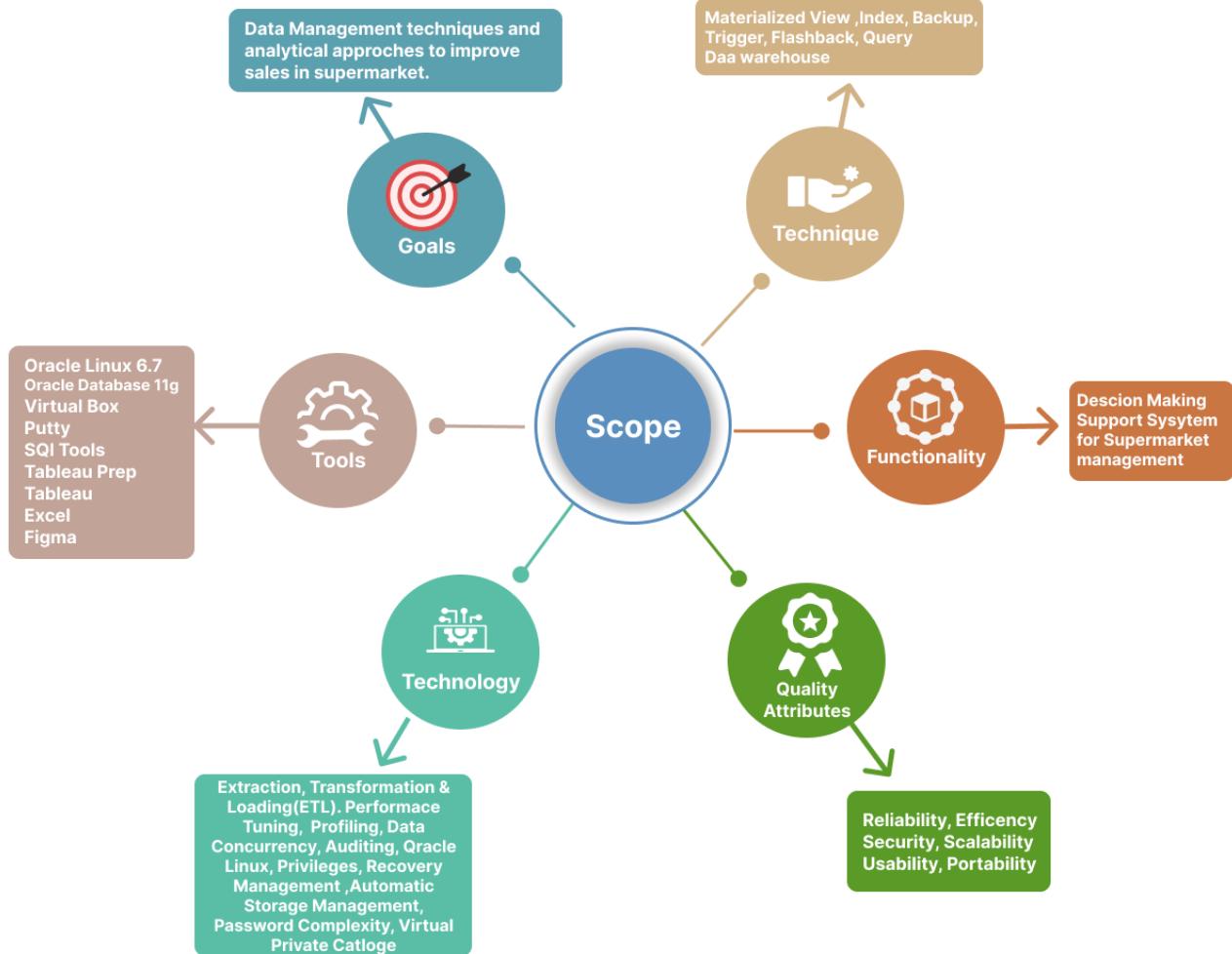


Figure 5: Scope

## Ethical consideration

Following ethical guidelines entails abiding by a set of rules and principles whenever it is practical to do so. Ethical considerations guarantee that no one will do anything to endanger others or the community as a whole. Individuals and organizations alike are discouraged from engaging in risky behavior as a consequence of this aspect. The narrative was primarily concerned with the question of what constituted good and poor conduct in interpersonal interactions.



Figure 6: Ethical Consideration

In any case, knowing who your customers, users, and sellers are is crucial. As a result, the authors have given their blessing to the publication of their names and the data they gathered. Access to sensitive information is restricted to a small group of trusted users to prevent it from falling into the wrong hands. The goal of this initiative was to create a system that would safeguard the privacy and anonymity of its users. Both the study's authoring and the system's construction were guided primarily by the General Data Protection Regulation (GDPR). Make sure to include only reputable sources, such as scholarly journals and well-known research papers, while writing your paper's section on data analysis or any other topic. All of the cited sources utilized in the study are included in the references section. None of the above is meant to be offensive or even close to anybody. No harm or modification is done to any of the values, customs, or beliefs that are inherent to the community and culture in issue.

## **Literature Review**

### **Desk-based Research**

Desk research is a method of gathering in-depth knowledge on a topic by analyzing a large number of research datasets that have already been collected and are publicly accessible. For legal and ethical considerations, this study uses a data sample from a major retailer. Information used in the study came from a wide range of publicly available resources such as the web, books, newspapers, and articles. Since all the necessary data can be collected rapidly and for nothing, this investigation is more productive in terms of both time and effort. Finding the right research subject is easier with the aid of desk research.

### **Walmart**

Walmart's sales volume exceed that of any other retail establishment in the planet put together. That employs over 2 million people in over 20,000 locations across 28 countries. It's to be expected that a corporation of this magnitude would acknowledge the significance of data analytics. Scientists studying the devastating storm that hit the United States in 2004 learned that data might provide unexpected insights when looked at in aggregate rather than as separate sets. CIO Linda Dillman stumbled into some surprising numbers while attempting to predict the need for emergency supplies in the aftermath of Hurricane Sandy. As consumers stockpiled supplies in anticipation of the upcoming storm, strawberry Pop Tarts were a popular item at many retailers. Walmart launched WalmartLabs and their Fast Big Data Team in 2011 as they saw the potential of data to help them better understand their customers' wants and needs and provide the items they were looking to purchase.

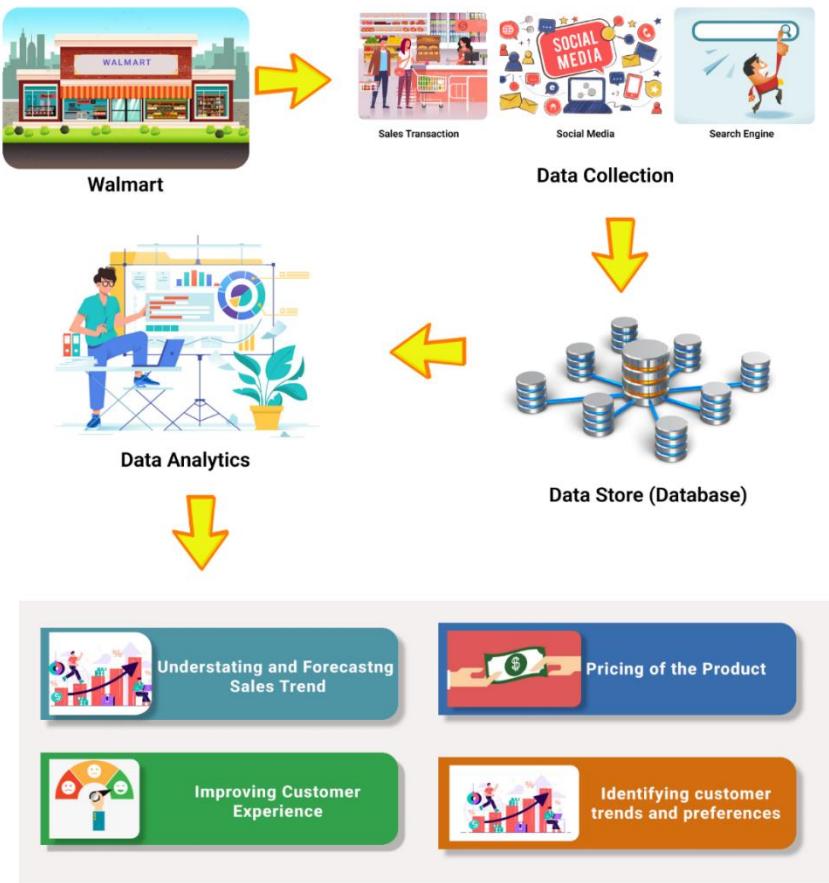


Figure 7: Walmart

The company's headquarters in Bentonville, Arkansas, is home to The Data Cafe, a cutting-edge analytics center. The Cafe's analytics team can monitor a 40-petabyte database of sales transactions in real time. Walmart's Senior Statistical Analyst Naveen Peddamail said, "If you can't obtain insights for a week or a month, you've lost revenue." WALMART's purpose is to supply business partners with data rapidly so they can respond and decrease turnaround time. It is a hybrid of predictive analytics and operational research. The Data Cafe is a place where teams from throughout the firm may come together to discuss data concerns and collaborate with analysts to find answers. The company also has a system in place that monitors key performance indicators and sends out notifications when those indicators increase over a certain level, at which time the relevant departments are expected to get in contact with the data team to determine what should be done. Peddamail cites a grocery store probe into a sudden drop in sales of a specific fruit as an example. By the time the figures reached the Cafe's specialists, it was obvious that a price error was to blame for the decline. Once a few days, everything had back to normal after the issue had been fixed. In addition, sales data from a variety of locations and regions may be monitored in real time. Peddamail claims that one Halloween, analysts monitoring sales data of novelty cookies found that the cookies were not selling at all across a number of retail locations. Thankfully, they were able to alert the store's merchandising management team, who promptly recognized they had overlooked displaying the products in question. It's

not a very complex algorithm, but it wouldn't have worked without access to real-time statistics. To anticipate what its consumers will purchase, Walmart's Social Genome Project, for instance, tracks their public social media activity. They have their own search engine called Polaris that delves deep into user queries, and they also provide a service called Shopycat that estimates how social influences effect consumers' decisions (again, utilizing social network data).

## 7-Eleven

7-Eleven is one of the most well-known convenience store chains in the world, and it is probably the biggest as well, with over 56,000 locations around the world. But until recently, the company didn't know much about its customers who paid cash. The 88-year-old retail company began setting up its first digital customer loyalty system for the U.S. in 2012. First, 7-Eleven's leadership (from operations, marketing, merchandising, and IT) included franchisees. In 2013, the firm collaborated with Teradata and Brierley+Partners to establish a CRM system that integrated transactional data with customer segmentation. 7Rewards, the customer-facing reward app, debuted last autumn. Each user gets a bar code and ID. The deployment began in 100 D.C. locations and was complete by March. Teradata's Real-Time Interaction Manager connected 8,500 POS systems in real time. Enterprise service bus was used.

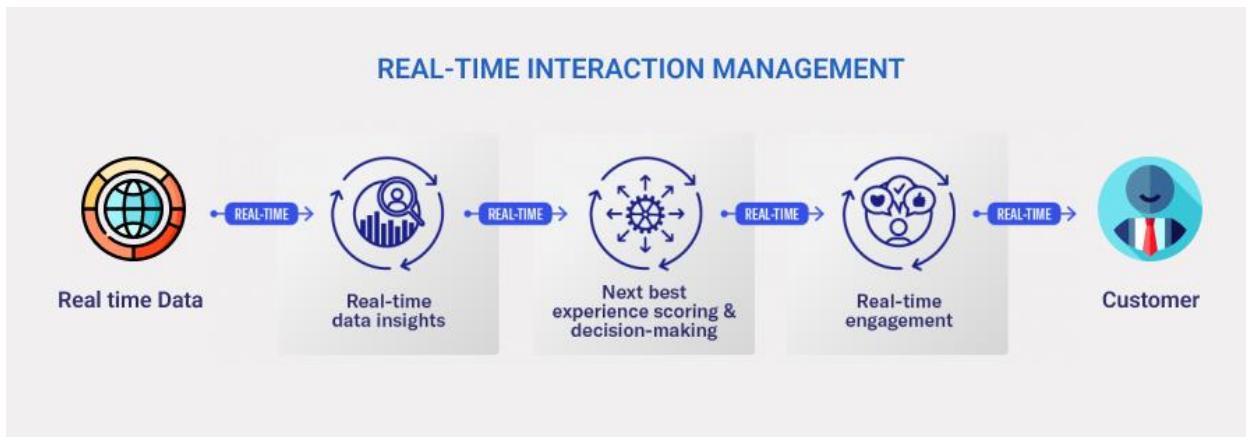
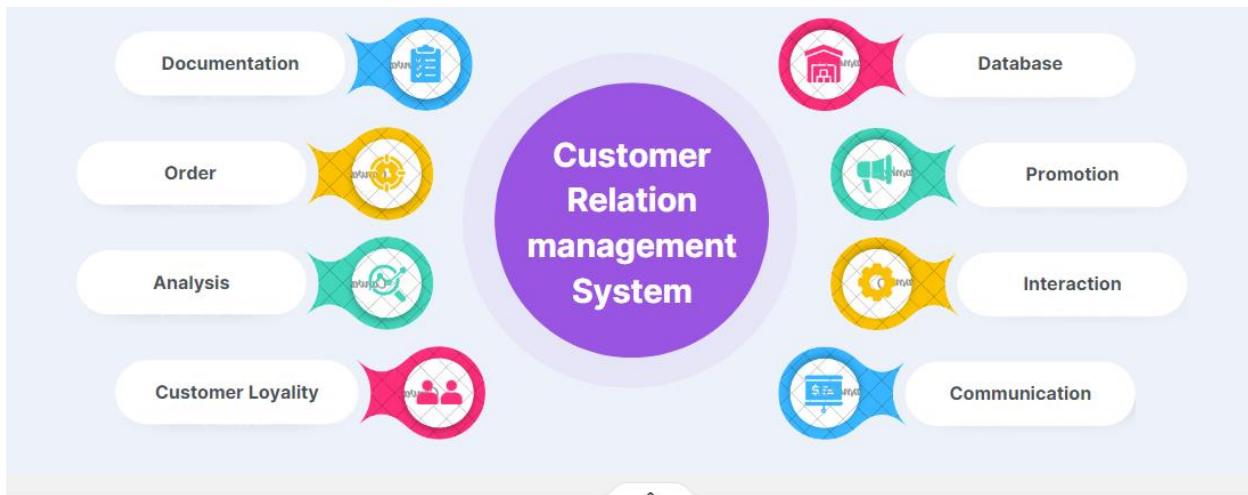


Figure 8:7-Eleven

Customer relationship management (CRM) is a method for managing interactions between a company and its existing and future customers. This is based on the straightforward premise of establishing connections in order to grow one's company. CRM software helps businesses stay in touch with their consumers, streamline their operations, and boost their earnings. When people talk about "CRM," they usually mean a Customer Relationship Management system, which is software designed to facilitate tasks like contact management, sales management, agent productivity, and so on. Customer relationship management (CRM) software may now be utilized throughout the customer's whole lifecycle, including but not limited to marketing, sales, e-commerce, and customer support. A customer relationship management (CRM) system helps a company in many ways, including attracting and retaining new customers, as well as providing existing customers with assistance and new services. These people might be clients, patients, colleagues, or vendors.



*Figure 9: Customer Relation Management System*

7-Eleven is the biggest chain retailer in the United States, with more than 5,500 franchised and company-operated outlets, and it faces the same pricing and product supply challenges as other companies. SAP claims that the software firm's demand management functions have helped 7-Eleven make more rational pricing selections. Scott McCombs, director of product intelligence at Seven-Eleven, recently discussed the need of rapid reactions in the present economic context. SAP's demand management solutions, founded on the SAP NetWeaver technology platform, aim to provide enterprise-wide decision support for pricing, marketing, and discounting. On a unified and scalable software platform, retailers can set pricing and stick to them with ease. The SAP retail solution portfolio includes products for retailers in the food, clothing, and hardware sectors. In addition to supply chain, finance, and human resources, the range of solutions also covers merchandising (including demand management) and retail operations (including a number of POS systems).

## Amazon

Jeff Bezos started Amazon in 1994 and is still its CEO today. Amazon.com, headquartered in Seattle, was once only an online bookstore, but now it is the biggest non-travel e-commerce firm in the world by a considerable margin. They've expanded from selling books and videos to selling eBooks, online videos, and now web services, making them a global retail giant. This is primarily because to their forerunner role in popularizing what is now often referred to as "recommendation engine" technology. This consists of algorithms that attempt to anticipate our desires and needs before presenting us with the option to pay for satisfying those wants and needs. Amazon has recently expanded its Amazon Now service, which offers fresh food with significantly quicker delivery than regular grocery shops, in an effort to compete with them. The stakes in the business world are greater than ever before as a direct result of the damaging impacts of information overload. It is crucial for online stores to provide their consumers as wide a selection as possible of their goods and services. Online retailers like Amazon and brick-and-mortar stores like Walmart have found success by trying to compete with supermarkets by providing "everything under one roof."



*Figure 10: Amazon*

Big Data collected from consumers' behaviors on Amazon was used to develop and improve the company's recommendation engine. While Amazon did not create the first recommendation engine, they did pioneer making them widely accessible to consumers. Collaborative filtering is crucial to Amazon's recommendation algorithm. This means it analyzes your profile data in order to make educated guesses about your preferences and then suggests goods that other people with similar profiles have purchased. When using Amazon's services, the business keeps tabs on every one of its approximately 250 million users. They monitor not only what you purchase but also what you look at, customer location (from which they might infer your income) and when and if customer provide feedback in the form of reviews and comments. It's used to build up a complete picture of who you are as a consumer. Amazon's main business is handled in its core data warehouse, which is comprised of Hewlett-Packard servers running Oracle on Linux, to handle the company's 187 million unique monthly website users and over two million third-party Amazon Marketplace sellers. The aggregated information is utilized to create a "360-degree vision" of their customers. Amazon then uses this information to identify additional users who they believe fall into the same consumer group, and makes product suggestions to them based on these users' previous purchases and preferences. The company's marketing approach would benefit from knowing the preferences of its clientele, which would be made possible by the information collected from clients. As a result, Amazon has begun to provide this information to marketers in order to facilitate the rollout of Big Data-driven advertising campaigns. Amazon's big data-driven customer service and shopping have made it a worldwide brand. Their customer-focused strategy and recommendation engine helped them succeed. Amazon's global revenues climbed 34.8% in the first half of 2021, from \$164.36 billion to \$221.60 billion. Net product sales rose 25.4% to 115.50 billion dollars from \$92.09 billion. Net service sales were \$106.10 billion, up 46.8% from \$72.28 billion previous year. Their use of data analytics has been a driving force in their company's rise to prominence. Sales of Amazon Web Services (AWS) increased by 34.7%, from \$21.03 billion to

\$28.31 billion. Big data recommendation engine simplifies anticipating client demands by profiling them based on purchases. These insights promote client happiness, boosting revenues. Amazon uses big data for price optimization by discounting best-selling goods and profiting from less popular things.

## Development methodology

### Agile

Agile is an iterative method of software and project management development that helps teams deliver early and often. Agile is one of the popular and productive strategy owing to flexibility, changeable and customer involve. The goal of agile was deliver value product with continuous delivery of product and cooperation and communication between team and product owner. The four values of the agile manifest are: It's based on a project, and it's always becoming better and more efficient during its life cycle. Kanban, Scrum, the Adaptive Project Framework (APF), and Extreme Programming are just a few of the numerous frameworks available (XP).

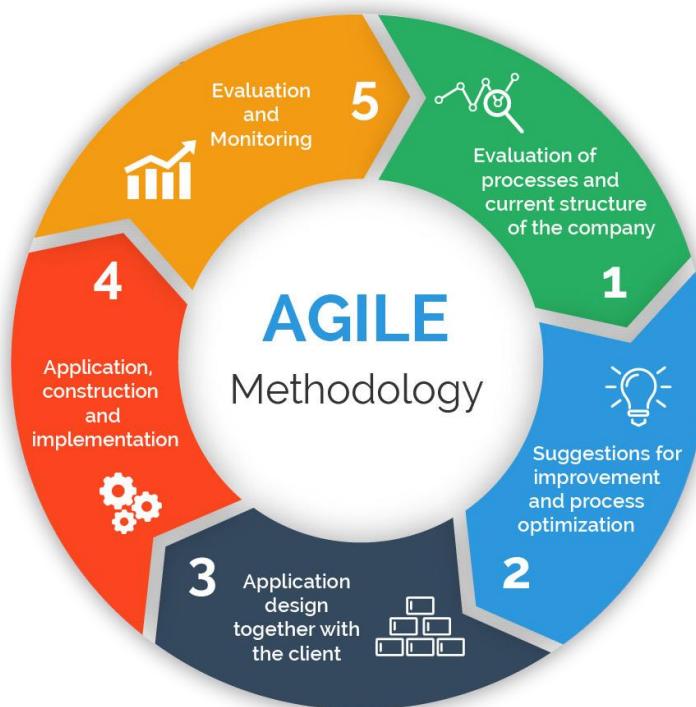


Figure 11: Agile

## Why Agile?

The agile technique has been shown to successfully construct software projects. The fact that it offers advantages over the waterfall technique has contributed to the widespread adoption of this methodology. My preference for the agile technique may be explained by a number of different factors. The teams are able to produce products swiftly and constantly, quickly reacting to changes and providing minimum viable products (MVPs) in subsequent iterations. Communication and frequent updates between the team and the stakeholders who are providing their comments are essential.

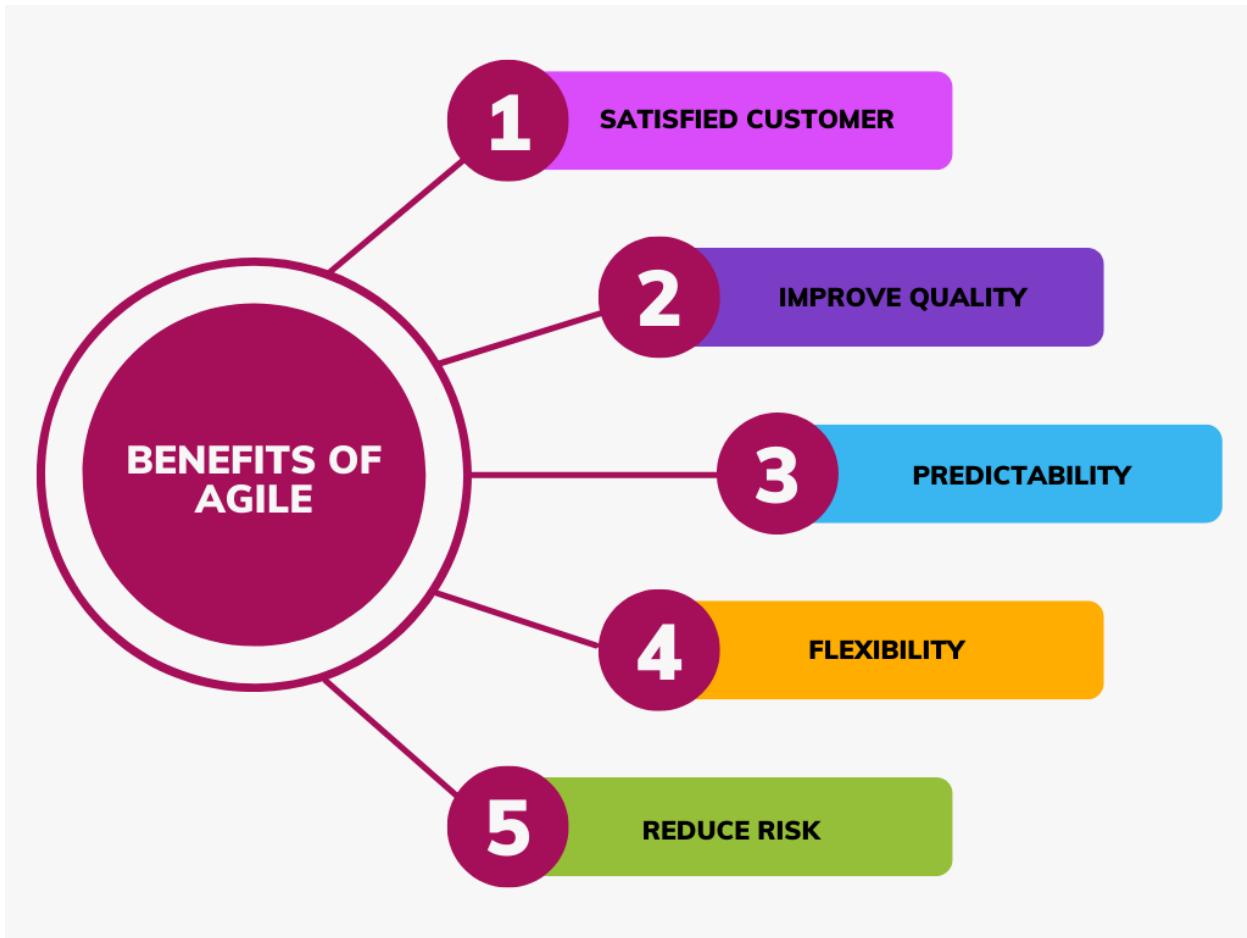


Figure 12: Benefit of Agile

### **Satisfied customers**

Participants in the agile life cycle are given opportunities to provide input and help shape the final deliverable. Enhanced client satisfaction and loyalty are two additional benefits of these procedures.

### **Improve Quality**

Agile is an iterative technique, which implies that with each cycle, the processes are refined and optimized more. This procedure is often used in places where the quality of the final product may be enhanced by doing so. Test-driven development (TDD) and behavior-driven development (BDD) are two types of testing used with this technique.

### **Predictability**

The agile technique operates in cycles called "sprints," during which the agile team produces significant results. The predictability of project outcomes and team outputs may be readily measured with the use of sprints. Time frames, budgets, and rough estimates of the whole working process are all easily calculable.

### **Flexibility**

The agile development method allows for more adaptability. Agile teams are often small, but they maintain continual contact with all team members and stakeholders in order to adapt the product's features and functions to the needs of its customers.

### **Reduce risk**

Developers using the agile technique provide working software to stakeholders on a continuous basis, while also identifying and resolving any roadblocks in a flash. By addressing them early on, we can prevent them from becoming major problems and strengthen our project's resilience.

## **Data Analytics Methodology**

### **Crisp –DM “Cross-Industry Standard Process for Data Mining Methodology”**

Crisp-DM helps firms with data mining, analytics, and data science. Life cycle management plans, organizes, and executes data mining projects. Using this method, data mining project planning began in 1996. Well-developed it processes boost flexibility, practicality, and strategic potential. Iteratively improving hypothesis testing and data analysis, and creating management-ready function templates.

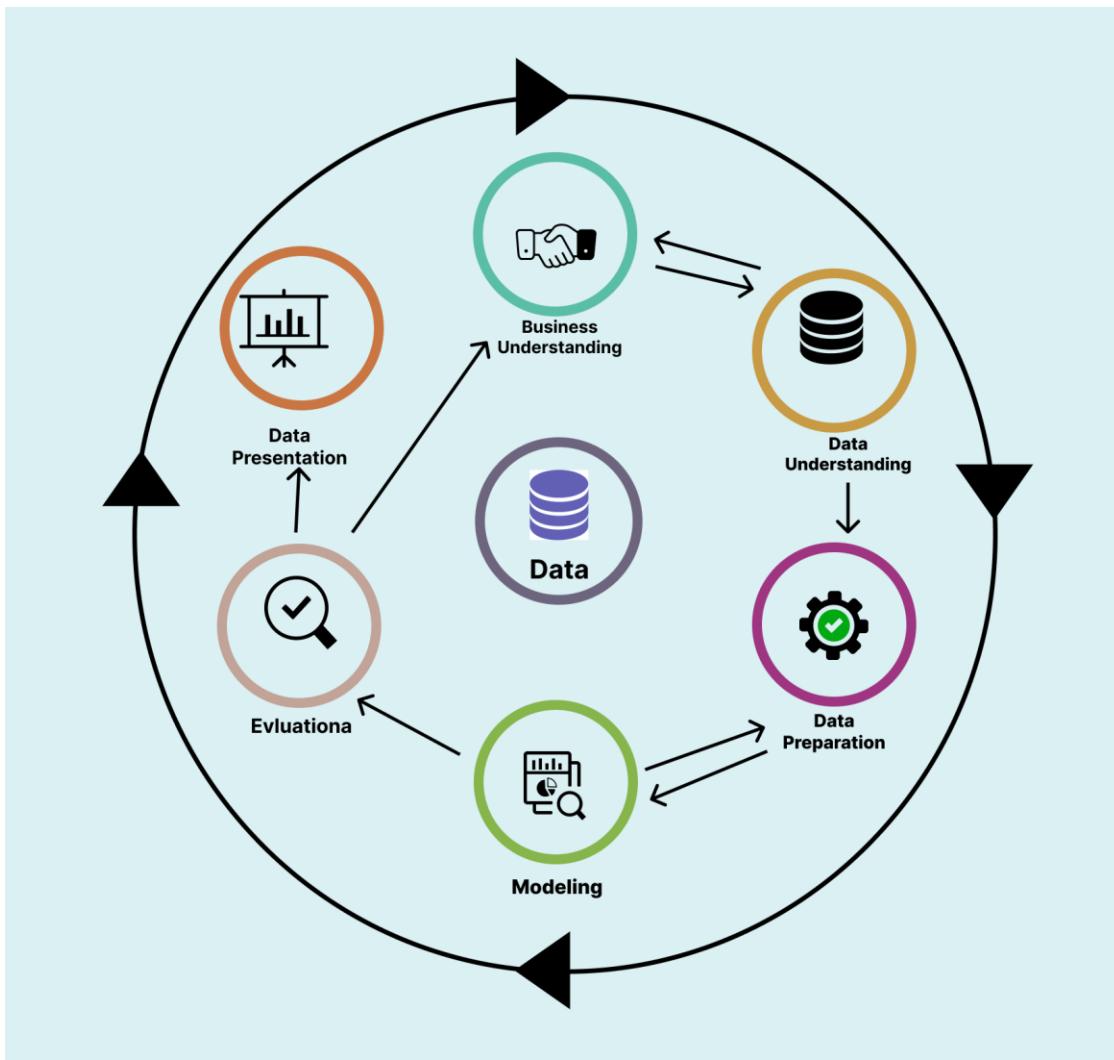


Figure 13: Crisp-DM

## Understanding the Business

Strategy development phases for businesses Consider the project's aims and requirements in the perspective of the business world. Following this realization, the analyst formulates a data mining challenge and plan of action. The viability of the project is evaluated through consideration of its costs, resources, dangers, and prospective rewards. Spend some time thinking about how the project will be organized and how the tasks will be accomplished.

## Data Understanding

This phase, sometimes known as "data munging," is used to get the final data set(s) ready for modeling. Once all of the data sources have been located, the next step is to choose, delete, build, and format the data. In this phase of data preparation, the data is checked for errors, built and integrated, and formatted.

## **Data Preparation**

During this process, which is also known as "data munging" in certain circles, the final data set (or sets) are cleaned up and made ready for modeling. After you have located all of the potential sources, you may then go on to selecting the data, removing it, building it, and formatting it. In order to properly prepare the data, they must first be cleaned, then created, then integrated, and last formatted.

## **Modeling**

The analyst is responsible for assessing, choosing, and putting into practice the relevant modeling methodologies. Because some methods, like as neural networks, have very precise criteria when it comes to the format of the data. A connection may be made here to the data preparation.

## **Evaluation**

Using the specified loss functions, the analyst develops and chooses models that have a good chance of performing well. Next, the analyst makes sure the models can adapt to new information by testing them. The analyst then verifies that all critical business concerns are covered by the models. Success is determined on which model is selected.

## **Deployment**

Most of the time, this involves inserting the model's underlying representational code into the system's underlying software. There are additional mechanisms here to rank and organize freshly uncovered content as it comes. The approach should apply the new information to the primary business problem. It is crucial that the code representation include all steps of data preparation prior to modeling. As a result, you may be certain that the model will handle new raw data in the same manner it did during training.

## Type of analytics

Data analytics focuses on drawing inferences from large amounts of unstructured data. Data analytics entails a wide variety of methods, many of which have been transformed into mechanical processes and algorithms that operate on raw data for human consumption. Statistics is the study of drawing inferences from numerical data. assist an organization in reaching its full potential in terms of performance, efficiency, profit, or strategic decision-making. Data analytics may be used to investigate what occurred (descriptive analytics), why an event occurred (diagnostic analytics), what will occur (predictive analytics), or what actions should be taken (prescriptive analytics) (prescriptive analytics).

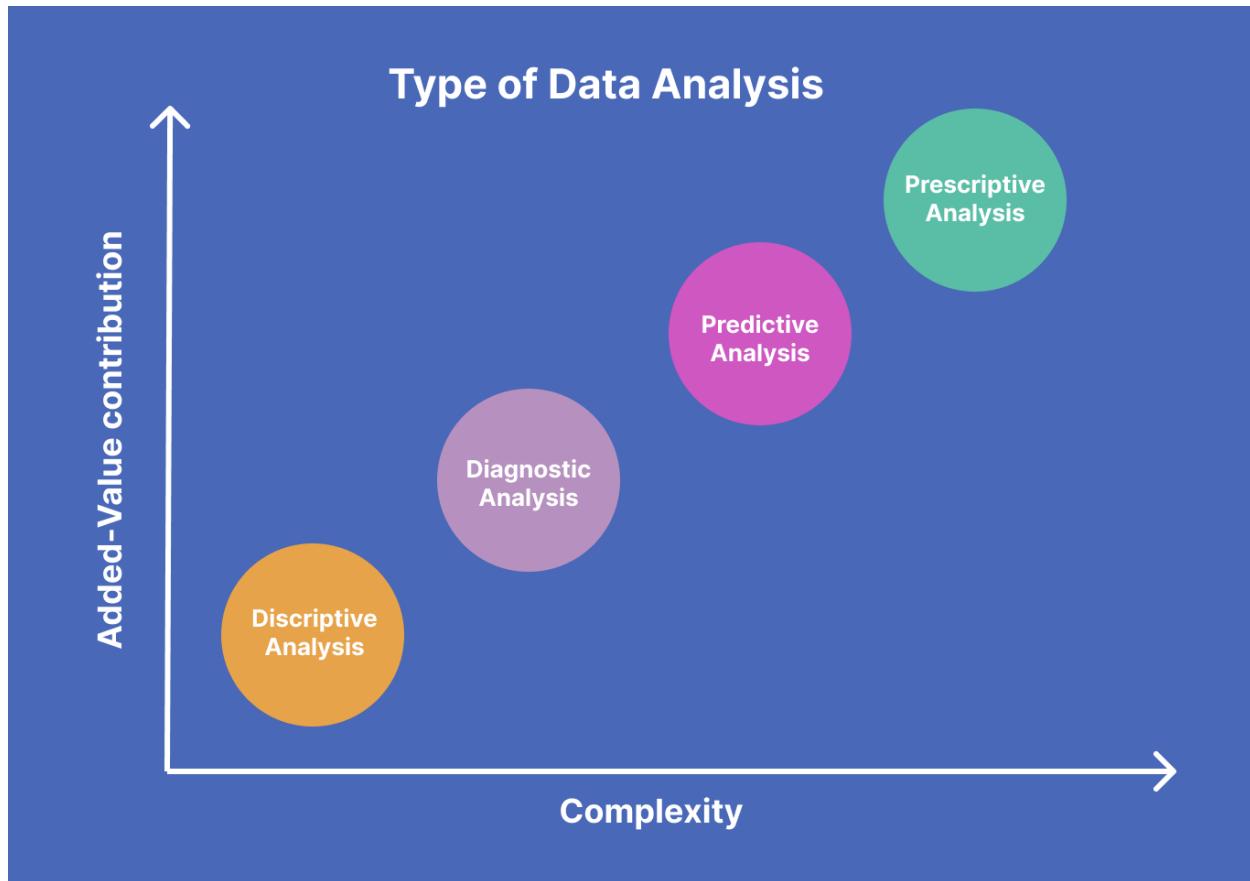


Figure 14: Type of Data Analysis

## Descriptive data analytics

Simply said, descriptive analytics describes what happened and adjusts raw data from various data sources to provide light on the past. But without explanation, these findings barely show whether something is right or incorrect. Most often, businesses utilize descriptive analysis for KPI tracking (KPIs). Key performance indicators reveal how well a firm is doing in comparison to predetermined standards.

## **Diagnostic Data analytics**

Diagnostic analytics allows for the comparison of historical data with other data in order to determine the root causes of occurrences. Diagnostic analytics provide in-depth knowledge about an issue. Diagnostic analysis digs deeper than descriptive analytics to determine the causes of observed outcomes. Organizations utilize this kind of analytics because it better draws conclusions from data and recognizes patterns of behavior.

## **Predictive data analytics**

Predictive data analytics has quickly become the industry standard. Analytics like this may be used in the business world for a variety of reasons, including the discovery of patterns, linkages, and causes. What will happen, for example, may be predicted with the use of this analytics. Complex analyses using deep learning or machine learning, as well as preemptive methods made feasible by projections, are two interesting facets of predictive analytics.

## **Prescriptive data analytics**

Big data and AI are used in prescriptive data analytics to make predictions and indicate next steps. Analytics like this are put to use in optimizations and random tests. Questions like "What if we try this?" and "What is the ideal action?" might benefit from prescriptive analytics, which makes use of ML advancements to provide solutions. The correct components, and maybe suggest some novel ones, with which to achieve the desired outcome.

## Tools

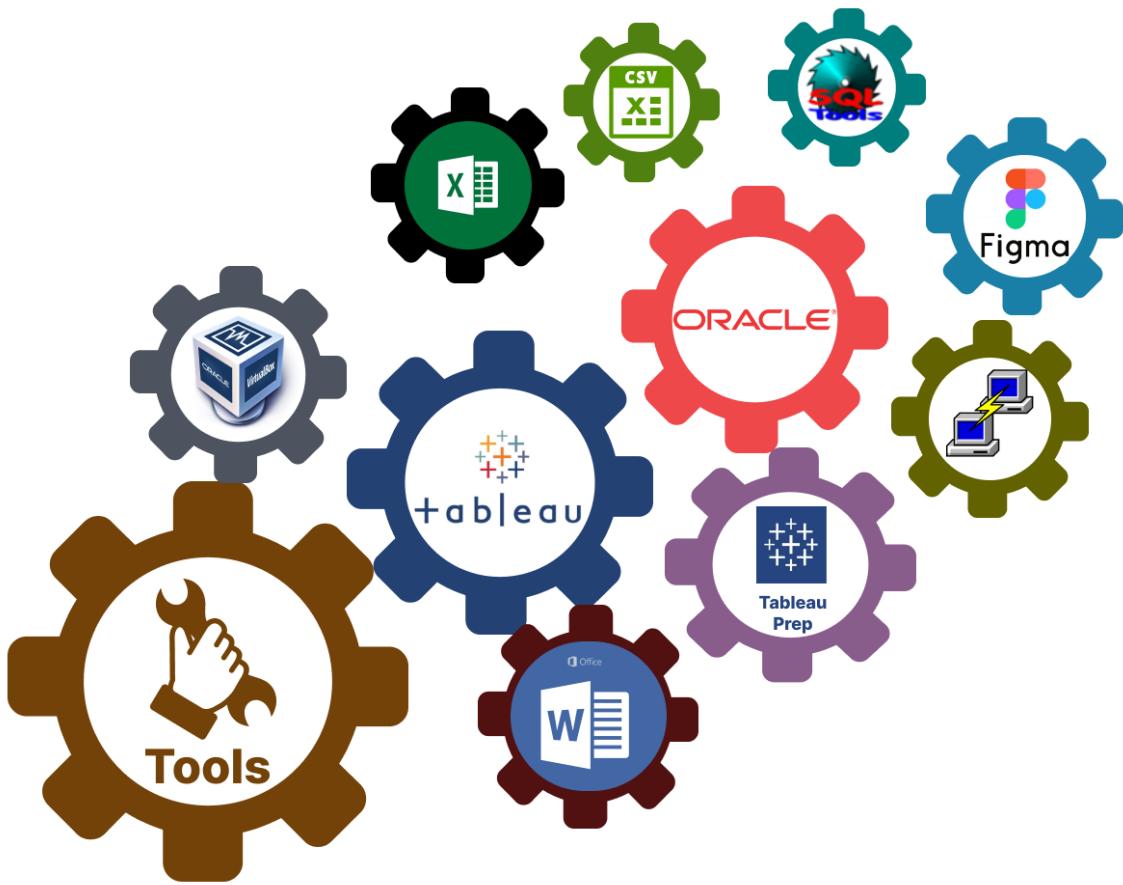


Figure 15: Tools

# Technology



Figure 16: Technology

# Techniques

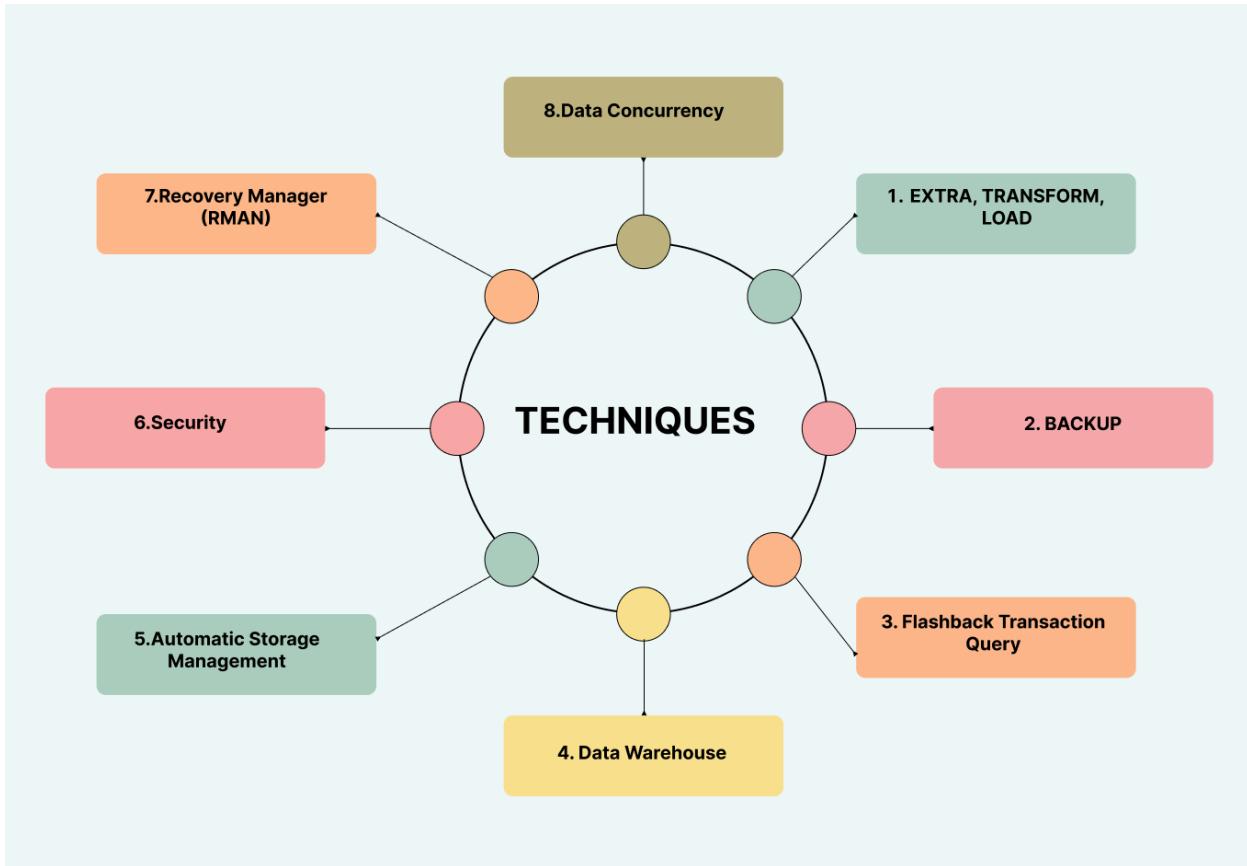


Figure 17: Techniques

## Data warehouse

Data warehouses were developed in the late 1980s with the goal of facilitating the transfer of information between operational and decision-support systems (DSSs). In the beginning, data warehouses needed a lot of backups. Typically, businesses would use a few distinct DSS settings to accommodate their many customer bases. While much of the same data was utilized across DSS settings, data collection, cleansing, and integration typically had to be performed several times. As data warehouses improved in performance, they grew from simple data storage to robust analytics infrastructures that serve a wide range of needs. These include operational analytics and performance management, among others. When it comes to analytics and other forms of business intelligence (BI), a "data warehouse" is a specific kind of data management system designed to help and speed things up. The primary purpose of a data warehouse is to serve as a research and analytical tool, hence they tend to store a lot of historical data.

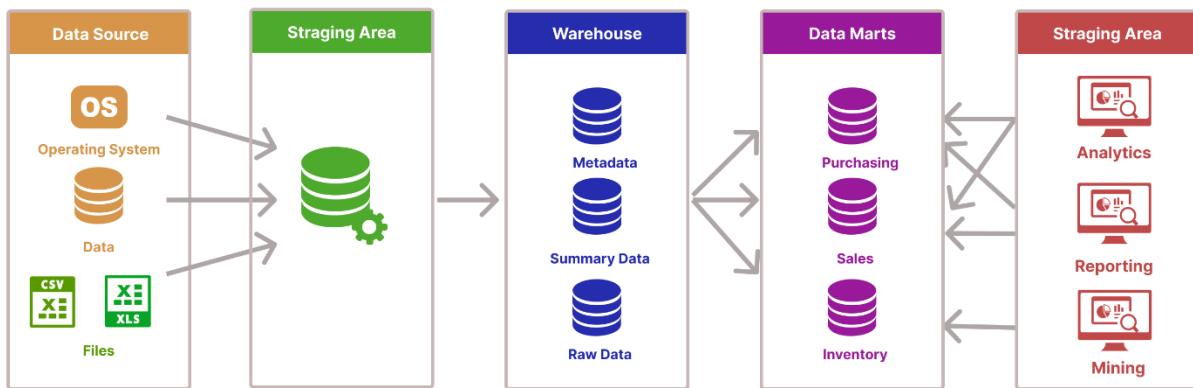


Figure 18: Data Warehouse

An information warehouse is a repository for massive amounts of data collected from a variety of sources. Massive volumes of data gathered from a wide range of sources are often stored in a facility known as an information warehouse. Businesses have the potential to gain significant business insights by putting their analytical capabilities to the challenge of improving decision-making. Data warehouses typically include, in addition to a relational database, an ELT solution, statistical analysis, reporting, and data mining capabilities, as well as client analysis tools for visualizing and presenting data to business users. This is done in order to maximize the value of the data warehouse. They are capable of providing data analysis according to a certain topic or job function (such as sales). A vast range of data coming from a number of different sources may be standardized inside a data warehouse. When information is stored in a data warehouse, such information does not change over the course of time. The study of data warehouses takes into account changes that have occurred over time. A well-constructed data warehouse should have a high data throughput, lightning-fast query results, and a wealth of different possibilities for customization.

## Extra, Transform, and Load (ETL)

Extract, transform, and load (ETL) is a data integration process that takes information from many sources and transforms it so that it can be put into a data warehouse or another target system with little loss of quality. The data analytics and machine learning processes depend on the information gathered and able to handle more complex analytics, which may enhance back-end operations or end-user experiences, as well as more particular business intelligence demands, such as monthly reporting. ETL extracts data from older systems. Improve data quality and consistency by cleaning it. Data-load a target database Check transformation, aggregation, and calculation rules.

# ETL Process

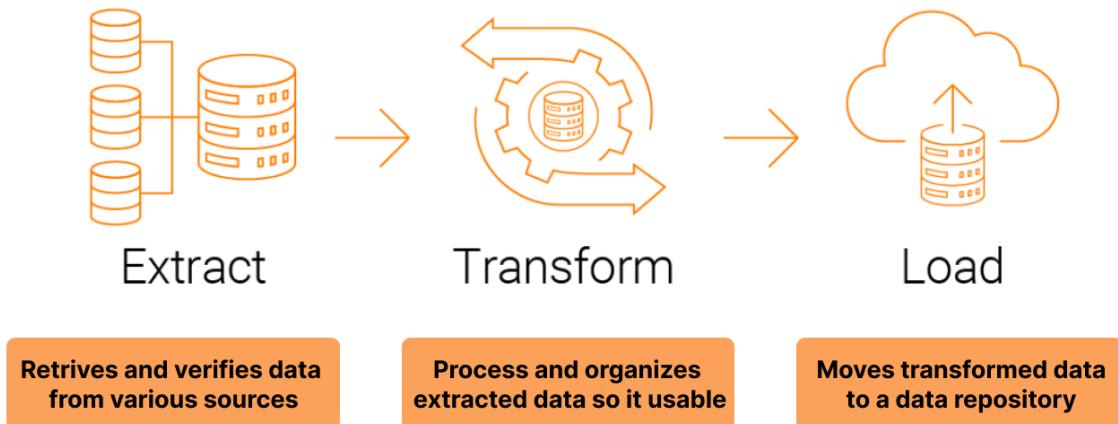


Figure 19: ETL Process

## Oracle Linux

Oracle Linux delivers a reliable, secure operating system to speed digital transformation. It has industry-leading performance and security for hybrid and multicolor systems. Oracle Linux and Red Hat Enterprise Linux are software-binary compatible. Oracle Autonomous Linux is a Linux distribution that has enhanced independence thanks to the strength of Oracle Linux. Because of its unique capacity to reduce complexity and human error, Oracle Autonomous Linux is the first and only autonomous operating system available today.

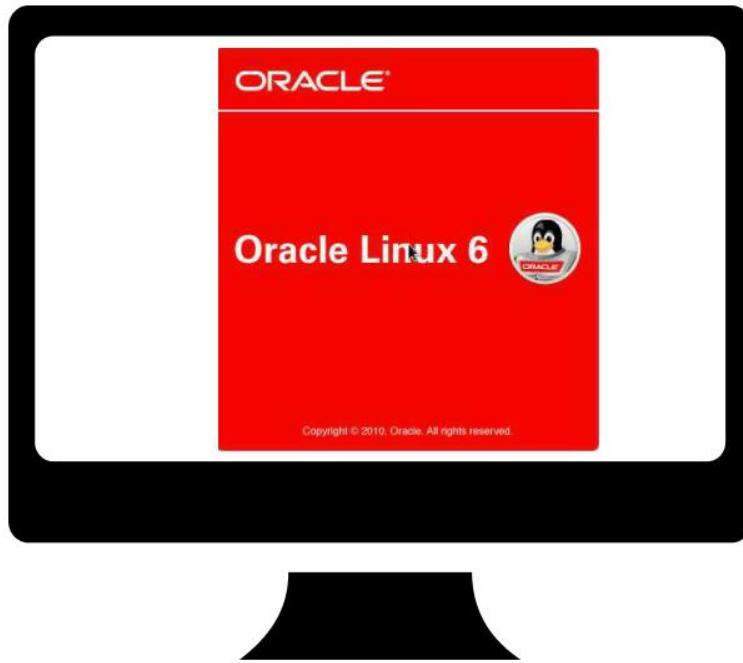


Figure 20: Oracle Linux

Oracle is one of a kind in that it offers the whole Linux-based solution stack, from applications and middleware to databases and management tools, operating systems, virtualization, hardware, engineered systems, and the cloud (operating system, virtualization, and hardware Premier Support customers get access to the AWX and Ansible-based Oracle Linux Automation Manager and Engine, which are used in the distribution of Oracle Linux 6.7. Put altogether, they offer a solid, scalable, and secure framework for automating business infrastructure at a fair price. Oracle Autonomous Linux is the first autonomous operating system that reduces complexity and human error to maximize cost savings, security, and availability. Gluster is a distributed file system that combines the disk space of several servers into a single, shared namespace, making it ideal for scaling data centers. Oracle Linux 6.7 features Common Criteria (CC) Certification and FIPS 140-2 certification of its cryptographic modules and STIG in Security Content Automation Protocol (SCAP). Oracle Linux 6.7 adds Unbreakable Enterprise Kernel, Containers, Persistent Memory, and Data Integrity.

## Automatic Storage Management (ASM)

Automatic storage management is cluster file system is both lightweight and speedy. Files and the Oracle database were both controlled by it. It creates a second hard drive to spread data evenly and creates a copy of the data in case the first one fails. Problems associated with storage are eliminated. Oracle's Automatic Storage Management unifies the database's file and volume administration (ASM). Oracle RAC features may be delivered using ASM either on a single SMP PC or across many cluster nodes. In order to manage files for applications other than the Oracle Database, such as executables, reports, BFILEs, video, audio, text, photos, and other general-purpose file data, ASM may be expanded with ACFS, a cross-platform, scalable file system and storage management solution.

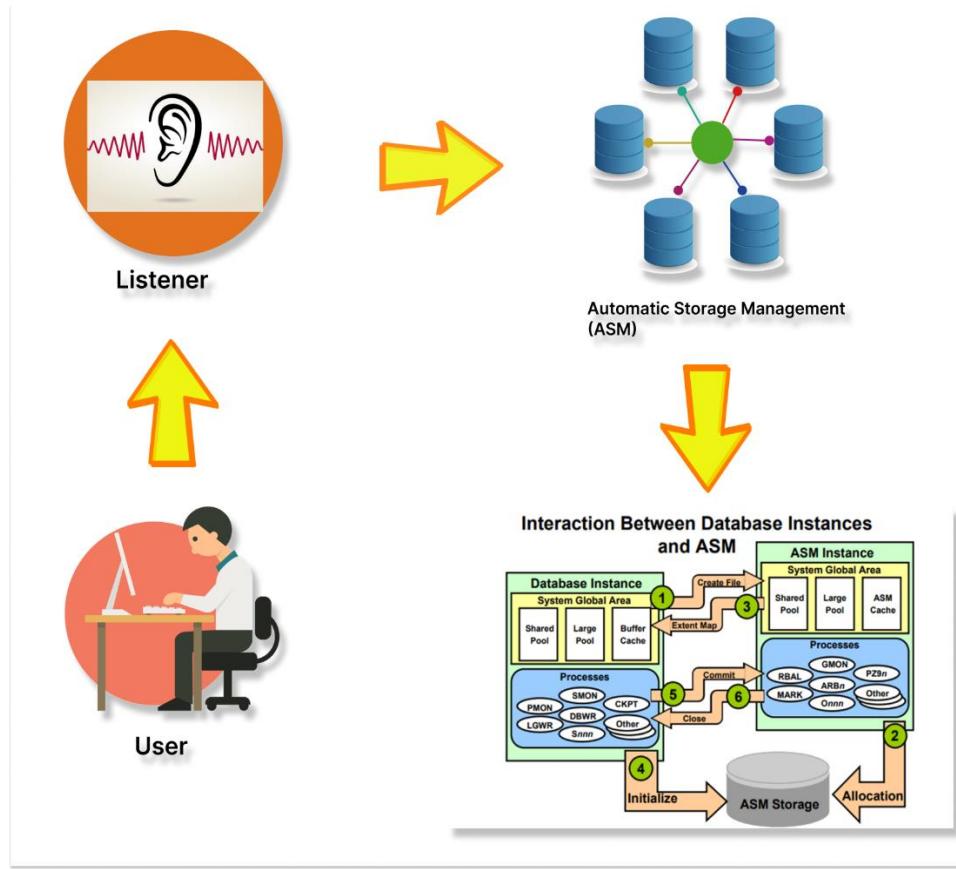


Figure 21: Automatic Storage Management

Furthermore, the database is requesting the ASM file. The ASM foreground process makes a COD entry and sets up space in the disk group for the new file. When a new file is added, its extent map is received by the ASMB database process and instantly initialized. A file commit is necessary to complete the database procedure once it has begun. In order to create the file and remove COD, the foreground ASM technique must be used. It is necessary to reopen the file for further I/O operations on the database instance. Style is synchronized between ASM and Database Instance. In order to map files to extents, database instances will need to communicate with ASM. Since Oracle ASM streamlines and automates file management, it is a useful addition to RAC. If the storage setup is altered, Oracle Automatic Storage Management will redistribute the database files across all of the drives (RAID 11). Redundancy is provided by ASM by copying database files to several hard drives. By removing the need to manually arrange data files on drives, ASM reduces Oracle's administrative load. The monotony of handling thousands of files is eliminated since the administrator only has to handle a small number of disk groups. When using ASM to set up a database, the disk groups are within the purview of the DBA. Once disk groups have been created, the DBA's role is limited to monitoring and making changes to disk allocations inside the groups. Each file is split into numerous smaller pieces called "extents" by ASM, and then those extents are spread over all of the disks in a disk group. Oracle databases create and delete files using ASM disk groups. Whereas earlier versions of ASM needed a manual command, ASM automatically deletes superfluous data files. When using drives that are prone to failure, ASM improves database consistency.

## Database Security

Database security tools, rules, and approaches preserve confidentiality, integrity, and availability. Data security includes all information security technologies and practices. It hurts database usage. More accessible and usable databases are more prone to security risks; more secure databases are harder to access and use.

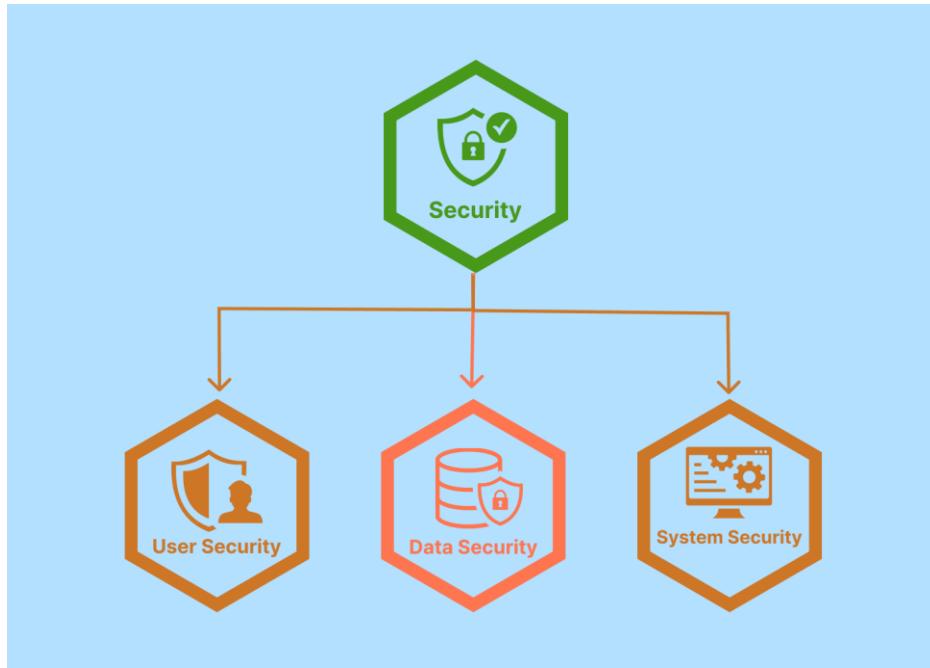
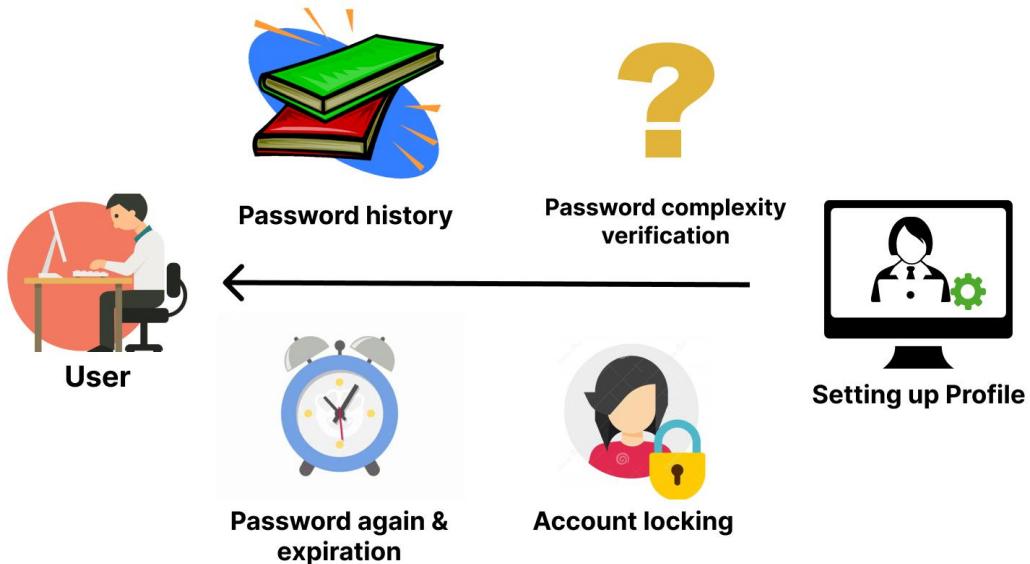


Figure 22: Data Security

## User Security

Password-protected user accounts give digital agency by limiting system access to authorized users. Secure systems need password authentication. Passwords secure user data. Pick a complex, secure password that's hard for insiders, outsiders, and invaders to guess. Oracle encrypts passwords using AES. Changing the password often may make it tougher to crack.



*Figure 23: User Security*

This project could utilize several Oracle capabilities, including the ability to allow many failed login attempts, specify a reuse time and expiry date, and set a minimum password length and lock duration. The company's needs decide who receives what. This project employs common password verification methods for users and profiles. This protects system users' credentials.

Oracle's inbuilt features include using a username as a password, allowing multiple unsuccessful login attempts, specifying a minimum password length and lock duration, and setting a reuse time and expiration date. This project implements widely-used techniques for guaranteeing password validity. This protects system passwords. Users may only have one profile. Changes to a user's profile won't be visible until they log out and back in. All succeeding profiles are based on the DEFAULT profile. Depending on the context, a profile's constraints may be expressed directly (CPU/Session) or left unsaid (CPU/Call) or referenced to from the DEFAULT profile's value (as in Connect Time).

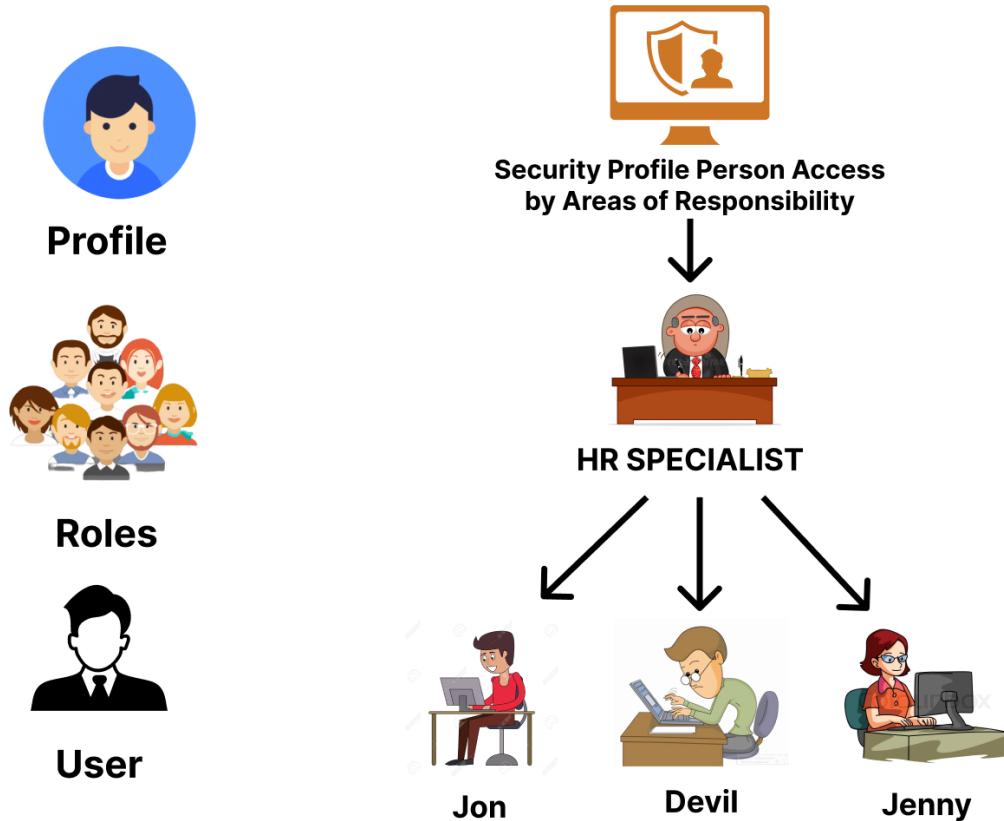


Figure 24: Privileges

Oracle characterizes privilege as the ability to connect, utilize resources, create, select, insert, edit, or delete data across schemas. Each job needs its own privileges. Object and system permissions are similar. Users with rights can alter databases. Users with the necessary credentials may create database tablespaces, sessions, etc. Object rights let users access a table, function, procedure, view, sequence, or package. The database administrator or someone with permissions may give access. Owner or authorized user may grant access. Owner consent is needed to use someone else's possessions. This project creates several roles, each with its own name, to provide users varied access capabilities. Comparable-level users connect, resource, insert, update, select, delete, etc. Role personnel may insert and choose on customer tables. All personnel may input and pick from the client table. Staff employees don't have access to data or information beyond their region, avoiding identity theft, phishing, and other unethical practices. It assigns roles and rights to protect user data.

## System Security

Maintaining data's usability for its intended purpose in the face of threats like loss or corruption is fundamental to system security. System security relies heavily on the development and implementation of preventative measures. There are a variety of tools at a system administrator's disposal for keeping data and infrastructure secure, including firewalls, data encryption, passwords, and biometrics. The Hardware Retail Information System (HRIS) must be kept secure to prevent unauthorized access to the sensitive financial, consumer, and other data it stores. Computer threats, assaults, and hackers may wipe out a business and all of its client data in an instant if the company's network isn't secure.

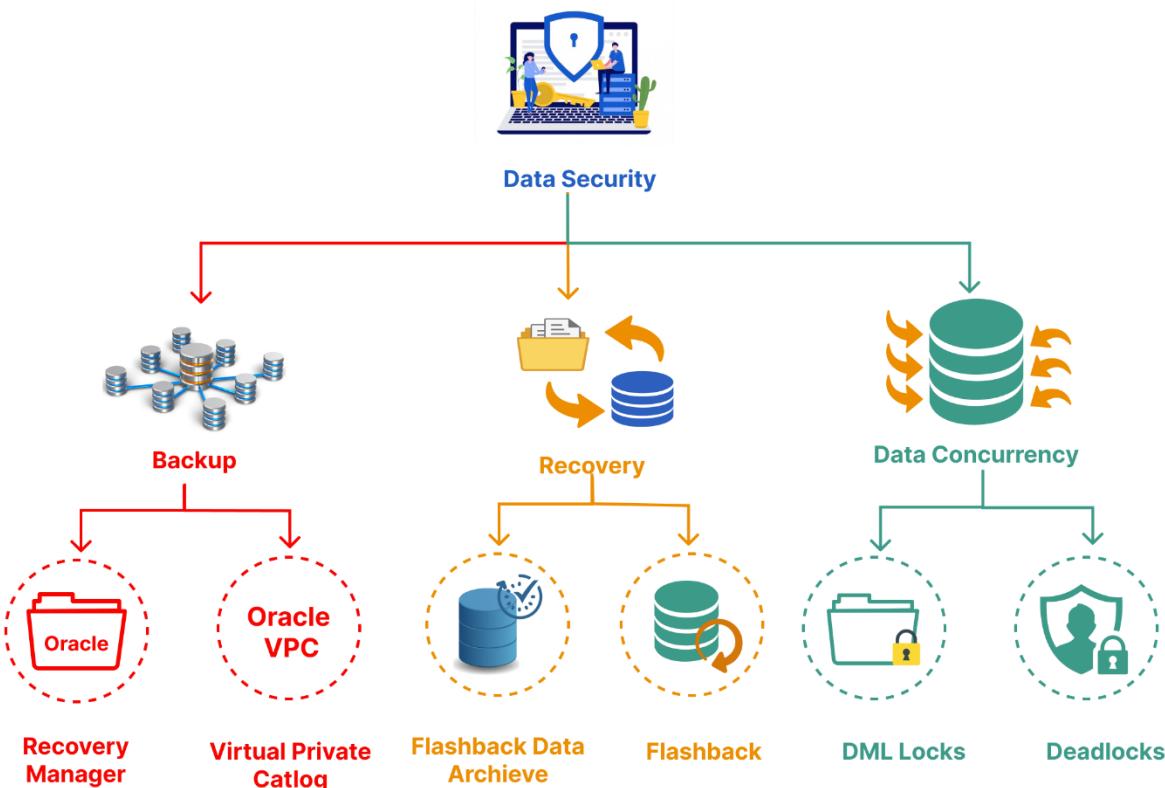


Figure 25: System security

## Auditing

Auditing creates more work for the system since it requires gathering and storing data. During an audit, only the most relevant occurrences should be recorded. When auditing is performed selectively, it has little to no effect on system speed. Auditing in Oracle entails keeping tabs on and recording database operations by authorized users. Users' actions in an organization may be tracked and assessed with the help of Oracle Auditing.

Secure systems prohibit unauthorized data access. Oracle users shouldn't have complete system data access. Database access and management may be limited by business needs, consumer preferences, and regulatory

constraints. The data shows this. Health and social data on ID cards, credit cards, and other cards must be secured. Oracle's database auditing covers obligatory, standard, value-based, fine-grain, and SYSDBA. These are auditing methods. Auditing may assist monitor staff system behavior. Helps firms find frauds, bad habits, and adjustments to avert damage. If an audit exposes employee misconduct, that individual should be investigated and reprimanded. If a company's audit mechanism is well-known, it may dissuade workers from cheating.

## **Backup**

### **Recovery Manager (RMAN)**

Oracle Recovery Manager (RMAN) offers native backup and recovery architecture for Oracle Database and Oracle Database Appliance deployments. RMAN commands conduct backup, restore, and recovery. RMAN parallelizes backups across RAC nodes. All storage, network connections, and CPUs may contribute to backups. RMAN block change tracking speeds up incremental backups. Block change tracking reads only updated database regions since the latest incremental or full backup. RMAN uses Image Copy or Backup Set to store data. A single data file, archived redo log file, or control file is copied as an Image Copy. RMAN-specific image copies aren't kept. They're the same as operating system file copying. RMAN can employ Image Copies with non-RMAN restore and recovery approaches. A backup set comprises one or more data, archived redo log, control, or server parameter files.

This project uses Oracle recovery management, backup, and FRA. This project method secures data. RMAN uses backup sets to move databases, data files, and tablespaces. Cross-platform data transmission reduces backup sizes via block compression. This project utilizes Oracle backup and FRA. It secures data. RMAN moves databases, files, and tablespaces using backup sets. Block compression minimizes cross-platform backup sizes Flashback Database needs a quick recovery area to store flashback logs. Not by default. Flashback log space is automatically maintained by the database and balanced with other rapid recovery files. RMAN uses a recovery catalog to hold Oracle database information. Recovery catalogs hold metadata longer than control files. This is important for recovering data before the control file's history. Extended backup history may mitigate the burden of keeping a recovery catalog database. Oracle database backups include:

#### **Whole Backup**

Each and every data file in the database, as well as the control file, should be backed up in full for a complete database backup. Most backups are copies of the whole database. Either the backup of the whole database is consistent, or it is not. Whether or not redo logs need to be applied after a backup restoration depends on the consistency of the backup.

#### **Full Backup**

The database's control file and all of its data files must be copied in their entirety for a full backup. The majority of databases' backups are complete copies. The database backup is either consistent across the board or it is not. The consistency of the backup determines whether or not redo logs must be applied after a restoration.

## Full Backup

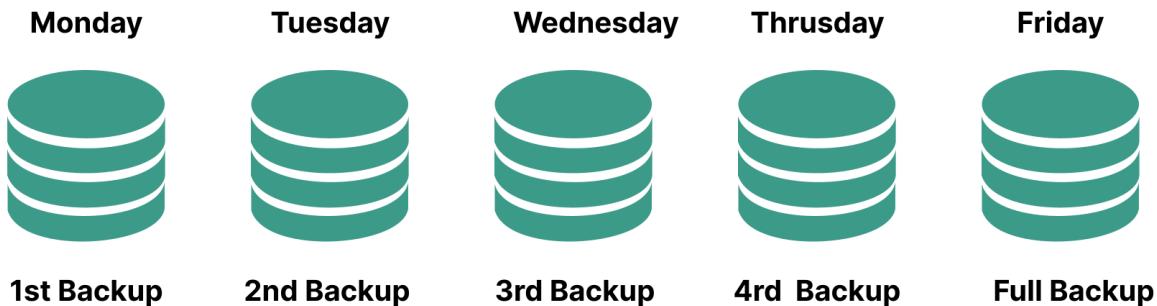


Figure 26: Full Backup

## Increment Backup

Incremental backups only copy changed data blocks. An incremental backup starts with a level 0 incremental backup, which replicates all data file blocks. A level 1 incremental backup copies only modified blocks from the previous backup. Level 1 backups may be cumulative, including all blocks modified since the last level 0 backup, or differential, including just blocks changed since the last level 0 or level 1 incremental backup. An incremental method creates daily level 1 backups. Oracle Recovery Manager (RMAN) automatically applies incremental backups and redo logs during recovery to restore the database to the specified time.

## Increment Backup

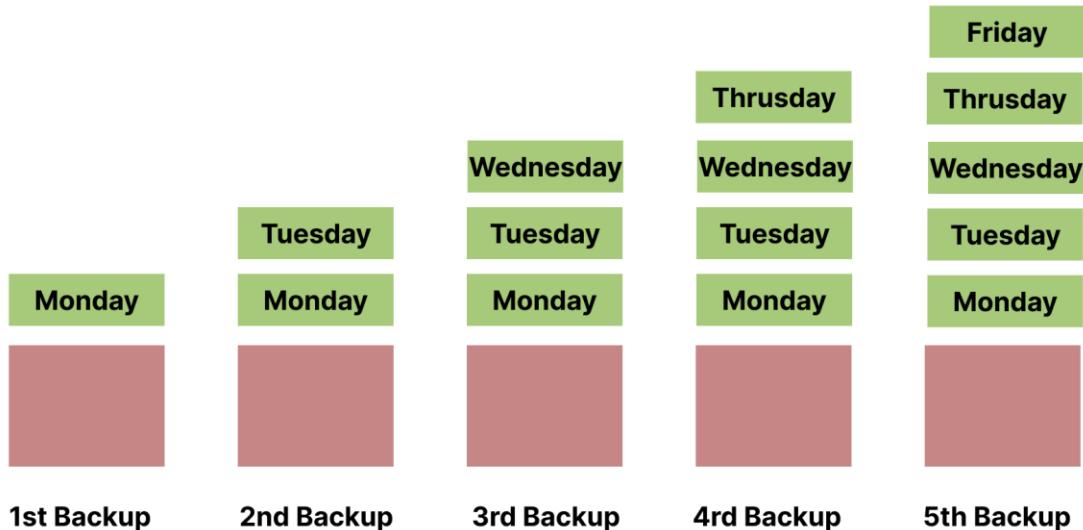


Figure 27: Increment Backup

## Physical backup

Data files, control files, and archived redo logs are all examples of physical files used in database storage and recovery. All backups, whether they're stored on disk or an offline medium like tape, are, at their core, copies of files containing database information.

## Logical Backup

Oracle Data Pump (export/import) tools are used to extract logical data from a database (such tables or stored procedures) and store it in a logical backup. Information is included in a binary file that may be read by Oracle Database.

## Virtual Private Catalog (VPC)

Oracle introduced Virtual Private Database (VPD), now a standard feature in Oracle Database 11g Enterprise Edition. When the object privileges and database roles provided by the application are inadequate, VPD may be utilized to provide the extra layer of security that is missing. Depending on the level of safety you need, a VPD policy might be quite straightforward or very complex. To ensure privacy and regulatory compliance, VPD may be used in tandem with the "application context" function to impose complex row and/or column level security requirements. One recovery catalog may have an unlimited number of virtual private catalogs built below it. There is a user in the database schema who owns each virtual private catalog, but that user is not the same as the user who owns the recovery catalog. If you are a virtual private catalog user, the recovery catalog administrator will provide you access to the databases you have registered in the recovery catalog. The recovery catalog administrator may also provide virtual private catalog access for registering new databases.

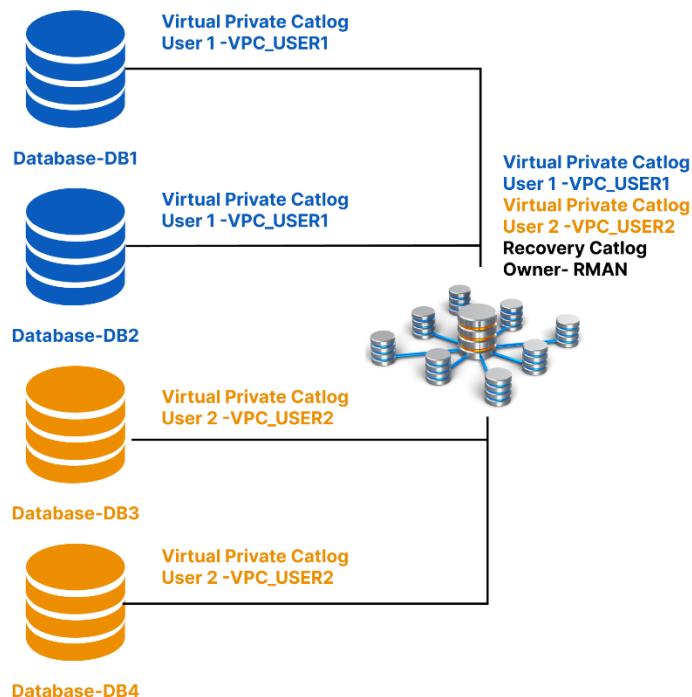


Figure 28: Virtual Private Catalog

## Recovery

### Flashback Technology

Oracle Flashback complements physical backup and recovery. Oracle Flashback Technology protects data. Oracle Flashback can access prior data states and rewind your database without restoring backups or conducting point-in-time recovery. Flashback is generally more efficient and less disruptive than media recovery. A flashback data archive uses logic to retrieve old data. Flashback data archives consist of tablespaces or tablespace sections. Specify the name, retention period, and tablespace when creating a flashback data archive. Specifies a default flashback data archive. After the retention term ends, the database automatically deletes old historical data.

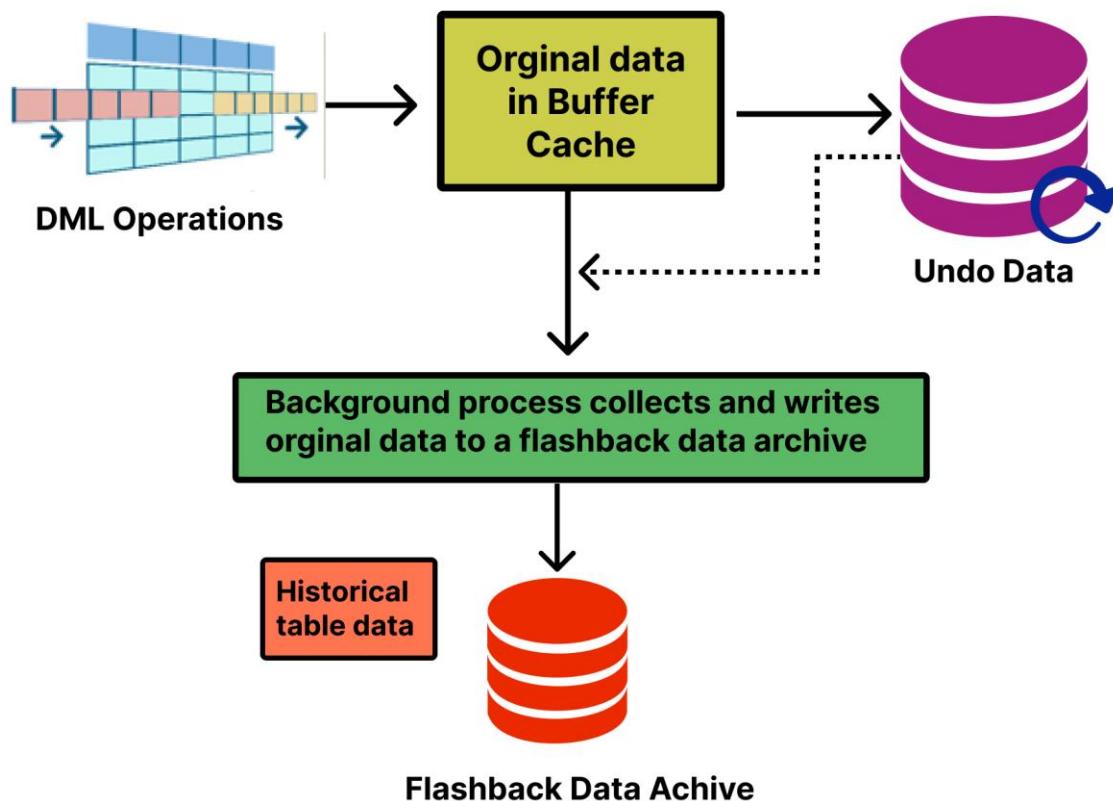


Figure 29: Flashback Data Archive

Oracle's flashback capabilities are mostly logical, allowing users to examine and edit database objects. Except for Oracle Flashback Drop, logical flashback capabilities use undo data, which captures database change impacts and overwritten values. It runs database queries at a specific time and displays results. To recover from an undesirable modification like an update to a database, perform a query to retrieve the lost rows. Oracle Flashback Database is a physical alternative to database point-in-time recovery (DBPITR). If

the present data files contain undesirable modifications, use the RMAN command FLASHBACK DATABASE to reverse them. The final result is similar to a DBPITR, but it's quicker since it doesn't involve recovering data files from backup or media recovery.

## Data Concurrency

In a single-user database, the user may alter data without worrying about other users. Multiple simultaneous transactions may update the same data in a multiuser database. Concurrent transactions must yield meaningful outcomes. Multiuser databases must govern data concurrency and consistency. Many users may access data at once. Each user sees the same data, including their own and other users' transactions. Database experts have defined serializability to explain consistent transaction behavior when transactions execute simultaneously. The serializable mode of transaction behavior ensures that transactions process serially, not simultaneously. While transaction isolation is typically preferable, running numerous programs in this manner might reduce performance. One transaction cannot insert into a table being searched by another.

Oracle query read consistency means all data comes at once (statement-level read consistency). Oracle's TWRC (transaction-level read consistency). Consistent Oracle rollback pieces. Transaction-altered data is in rollback chunks. Consistently read statements. One query returns data from the start. Query execution never commits unclean data. Displays just committed data. After execution starts, the query ignores changes. Oracle transactional read consistency. Serializable mode represents the database's state at transaction start. All transactions view consistent data at the same time, excluding serializable ones. Phantoms are eliminated through transaction-level consistency. Both read committed and serializable transactions lock uncommitted rows. First row change commits or undoes, then releases lock. If another transaction rolls back, the waiting transaction may update the locked row. Oracle locks prohibit data access and user harm. Oracle prevents other transactions from using a transaction's resource. When a transaction no longer needs a resource, locks are released.

A DML lock (data lock) ensures the integrity of simultaneously accessible data. DML locks prevent damaging DML or DDL meddling. DML statements obtain table- and row-level locks. Row-level locks prohibit transactions from altering the same row. When modifying a row, a transaction acquires a row lock. A statement or transaction may hold any number of row locks, and Oracle does not escalate them. Row locking gives the greatest concurrency and performance. Table-level locks prevent a table from being dropped in the midst of a DML process. A table lock is obtained using DDL or DML statements. Locking tables doesn't effect DML concurrency. Partitioned tables may have table and sub partition locks. A row exclusive table lock (SX) indicates that the transaction holding the lock has updated table rows. A transaction's row exclusive table lock enables other transactions to query, insert, update, remove, or lock rows simultaneously. Multiple transactions may gain row exclusive and row share table locks for the same table. A data dictionary lock (DDL) secures a schema object's definition during a DDL process. DDL commits implicitly. Say a user constructs a process. Oracle gains DDL locks on all procedure-referenced schema objects for the user's single-statement transaction. DDL locks prevent procedure objects from being changed or discarded before compilation.

## Database Maintenances

Oracle's powerful architecture simplifies proactive database maintenance. AWR is a built-in Oracle database repository. The database makes frequent backups, refreshes optimizer statistics, and health checks using automated activities. Oracle database server delivers server-generated notifications for situations that can't be fixed automatically (such running out of space). Oracle database server monitors itself and delivers notifications about issues. Alerts tell you of problems and frequently provide solutions. Each subsystem's adviser makes recommendations. Each subsystem's adviser makes recommendations.

## Automatic Workload Repository

The Automatic Workload Repository (AWR) collects, processes, maintains, and uses performance information for issue identification and self-tuning. ASH stores recent session activity in the AWR. Statistics describe the database and its items. Query optimizer uses optimizer statistics to find the optimum SQL execution plan. Database statistics monitor performance. AWR snapshots comprise database, application, operating system, and other metrics. AWR baselines are time-stamped AWR pictures. The baseline is used to compare performance to the baseline or to another benchmark. Oracle Database 11g automatically collects SMW baseline. The System Moving Window baseline includes the past eight snapshots by default. After enough data is gathered and analyzed, this baseline becomes legitimate. Statistics are calculated every Saturday at midnight.

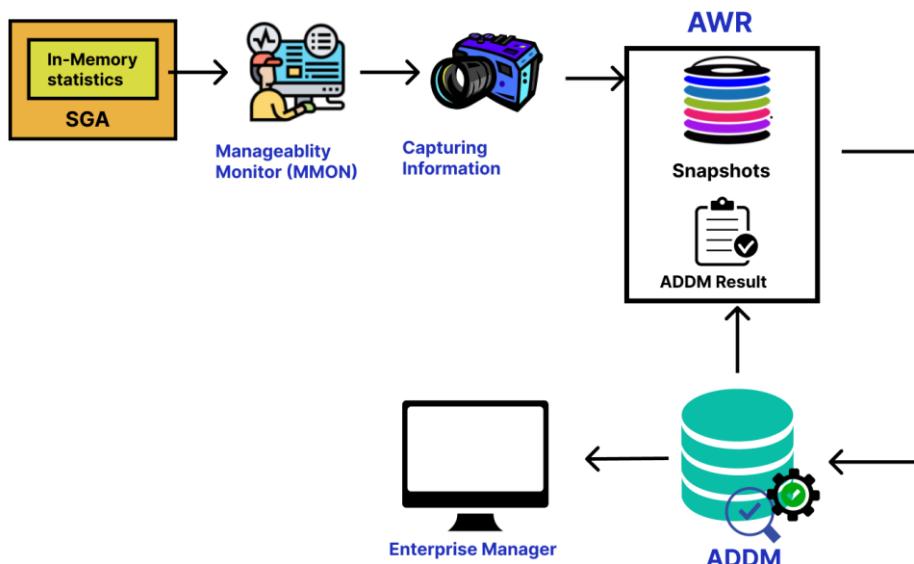


Figure 30: Database Maintenance

## **Integration**

With months of research behind it, a data warehouse, system security, backup, and other features were integrated into a supermarket intelligence system. This technique was made for increase the sales development of supermarket. Several new oracle features are used by the oracle database system to store, manage, visualize, and analyze the data. The information on grocery stores comes from Tableau's Data sample. When it comes to ensuring data integrity, backups, and recovery, nothing beats loading up the data in problem into an oracle feature data warehouse. A data warehouse is a storage management tool that enables scaling. Through the use of profile, roles, tablespace, and enabled password complexity, automatic storage management technology ensures the safety of each user's account and their data. System security was maintained via the use of auditing to track and record user actions. This prevents misuse of the system. In addition, apply backup and recovery techniques by enabling several backup kinds such as full backup, incremental backup, logical backup, and physical backup, all of which protect the data. Agile methodology, which uses an iterative approach to both software and project management, was used to create the project. The last step involves importing the grocery store data into tableau prep, where it will be cleaned before similar charts are compiled into a dashboard and shown. Tools from Tableau are used to create the visualization. Various metrics that show a company's health—including customer retention rate, profit, sales, and more—may be depicted graphically. The color schemes used in the creation of executive dashboards, decision support dashboards, and operational dashboards are accessible to anyone with color-blindness. These dashboards allow the retail system to track its performance and make informed choices.

## **Data Visualization**

Charts, graphs, infographics, and even animations may all be used for data visualization. These visual representations of data make complex relationships obvious and provide data-driven insights in a natural way. For this task, we use the data visualization capabilities of the Tableau software to study the trends, patterns, and larger data set provided by the Outliner. Pictures are readily recognized by the human brain, making them ideal for use in dashboard visualizations where they may provide much-needed context and information without the need for lengthy documentation or repeated explanations. The usefulness of data might be enhanced by visualization. Data visualization has many uses and isn't only for data teams. Data analysts and data scientists utilize it to uncover and explain patterns and trends. Data visualization stimulates team ideation. During brainstorming or Design Thinking meetings at the outset of a project, they help gather multiple viewpoints and identify common challenges. Visual discovery helps data analysts, data scientists, and other data professionals detect patterns and trends in a dataset. Everyday data viz enables storytelling once a new insight is discovered.

## **Executive Dashboard**

An executive dashboard compiles information on several KPIs in one place, allowing for quick, informed decision-making by company leaders. Modern executive dashboards aggregate information from every department and provide users the freedom to go through the data at their own pace. Executive dashboards are crucial for the contemporary CEO. It is much easier for CEOs to monitor and assess the most important indicators for their company's success when they are all shown in one convenient location. The CEO may get insights previously buried in data silos by using an executive dashboard that consolidates important data sources from throughout the enterprise to get a consolidated perspective, or by diving into individual data sets from any line of business.

## **Decision Support Dashboard**

The decision support system that is being developed for this project will assist analytical levels in the collecting and analyzing of vast data sets for the objectives of problem-solving and making choices that are based on accurate information. At this point in the process, the analytical models are utilized to extrapolate patterns from historical data. This dashboard is helpful for sales forecasting as well as delivering information to customers in a manner that is easy to understand. The market basket analysis chart displays the categories of products that work well together. It displays a connection between a numbers of different things.

## **Operational Support Dashboard**

The development of an operational dashboard enables high-level monitoring of the efficiency of operational processes. In most instances, the information is up to date and relevant to the actions that are typically carried out. This dashboard is designed to monitor and evaluate operational processes, provide insight into potential operational impediments, and communicate with management on whether or not their firm is on track with its operations. These dashboards are made accessible to the whole team in order to encourage a culture of leadership and shared responsibility for any modifications to the operational procedures. As a consequence of this, this dashboard offers information in real time to aid businesses in determining whether or not their processes are functioning well.

## Visulization on Supermarket



### YTD Sales Comparision with Previous Year



Figure 31: Dashboard 1

## Visulization on Supermarket

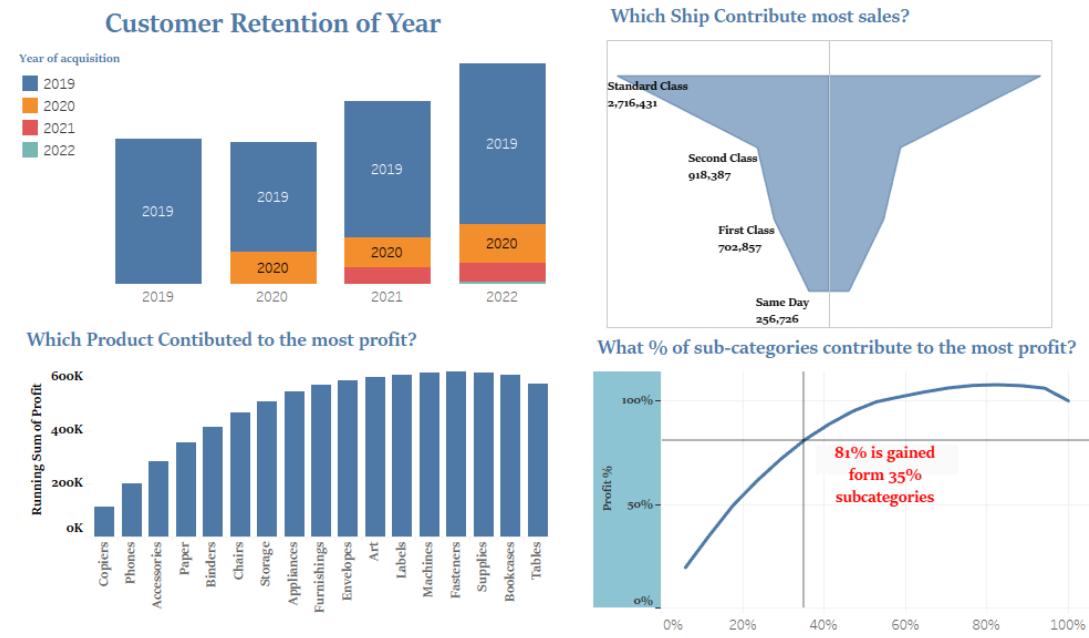


Figure 32: Dashboard 2

## Visulization on Supermarket

Executive Dashboard	Executive Dashboard	Decision Dashboard	Sunburst	Top 10 & Bottom 10	Operational Dashboard	Bump Ch
---------------------	---------------------	--------------------	----------	--------------------	-----------------------	---------

### Market Basket What goes together

Sub-Category	Accesso..	Applian..	Art	Binders	Bookcas..	Chairs	Copiers	Envelop..	Fasteners	Furnish..	Labels	Machines	Paper	Phones	Storage	Supplies	Tables	
Accessories		46.8%	24.0%		7.8%	30.6%	42.9%	12.4%	22.6%	34.3%	32.1%	3.4%	-0.1%	31.1%	33.0%	29.1%	-7.1%	36.7%
Appliances	46.8%		24.0%		7.4%	9.5%	45.7%	-5.3%	-0.4%	15.7%	29.9%	8.9%	14.9%	27.8%	16.3%	32.2%	-1.3%	14.6%
Art	24.0%	24.0%		12.1%	8.0%	52.8%	50.1%	59.7%	62.4%	31.7%	28.8%	16.6%	47.2%	22.0%	24.8%	26.9%	40.2%	
Binders	7.8%	7.4%	12.1%		22.3%	14.6%	6.1%	10.0%	28.8%	27.7%	2.4%	-0.2%	24.2%	23.5%	14.7%	9.7%	26.5%	
Bookcases	30.6%	9.5%	8.0%	22.3%		32.0%	36.0%	9.0%	34.6%	26.3%	32.6%	2.4%	32.0%	19.8%	4.6%	-5.2%	18.2%	
Chairs	42.9%	45.7%	52.8%	14.6%	32.0%		21.2%	45.3%	32.7%	36.0%	14.6%	48.3%	51.5%	56.7%	42.3%	4.9%	78.4%	
Copiers	12.4%	-5.3%	50.1%	6.1%	36.0%	21.2%		27.5%	17.0%	-0.6%	8.4%	18.8%	21.5%	26.2%	9.1%	76.9%	55.6%	
Envelopes	22.6%	-0.4%	59.7%	10.0%	9.0%	45.3%	27.5%		22.9%	13.2%	62.9%	40.0%	38.6%	18.2%	29.2%	62.7%	10.4%	
Fasteners	34.3%	15.7%	62.4%	28.8%	34.6%	32.7%	17.0%	22.9%		22.3%	18.9%	24.7%	53.4%	44.3%	21.9%	-10.4%	30.7%	
Furnishings	32.1%	29.9%	31.7%	27.7%	26.3%	36.0%	-0.6%	13.2%	22.3%		9.5%	45.6%	35.3%	53.1%	22.5%	26.8%	26.4%	
Labels	3.4%	8.9%	28.8%	2.4%	32.6%	14.6%	8.4%	62.9%	18.9%	9.5%		13.7%	27.5%	26.0%	23.6%	-5.2%	8.8%	
Machines	-0.1%	14.9%	16.6%	-0.2%	2.4%	48.3%	18.8%	40.0%	24.7%	45.6%	13.7%		9.8%	6.2%	33.7%	-5.3%	25.7%	
Paper	31.1%	27.8%	47.2%	24.2%	32.0%	51.5%	21.5%	38.6%	53.4%	35.3%	27.5%	9.8%		44.3%	46.2%	19.3%	48.9%	
Phones	33.0%	16.3%	22.0%	23.5%	19.8%	56.7%	26.2%	18.2%	44.3%	53.1%	26.0%	6.2%	44.3%		35.5%	23.1%	40.3%	
Storage	29.1%	32.2%	24.8%	14.7%	4.6%	42.3%	9.1%	29.2%	21.9%	22.5%	23.6%	33.7%	46.2%	35.5%		20.4%	36.7%	
Supplies	-7.1%	-1.3%	26.9%	9.7%	-5.2%	4.9%	76.9%	62.7%	-10.4%	26.8%	-5.2%	-5.3%	19.3%	23.1%	20.4%		-6.8%	
Tables	36.7%	14.6%	40.2%	26.5%	18.2%	78.4%	55.6%	10.4%	30.7%	26.4%	8.8%	25.7%	48.9%	40.3%	36.7%	-6.8%		

Figure 33: Dashboard 3

## Visulization on Supermarket

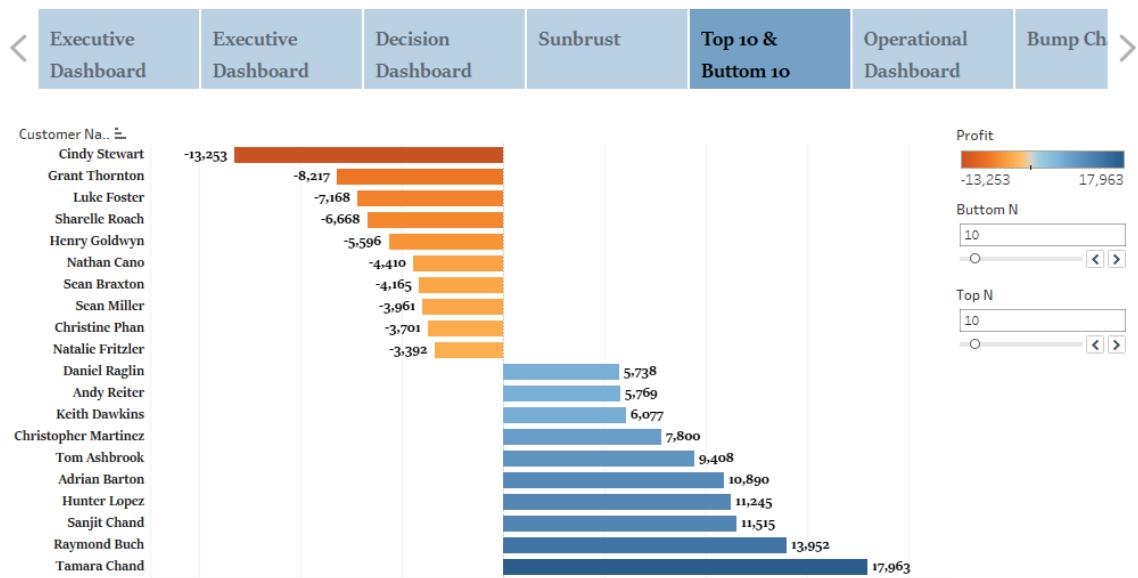


Figure 34: Dashboard 4

## Visulization on Supermarket

<a href="#">Executive Dashboard</a>	<a href="#">Executive Dashboard</a>	<a href="#">Decision Dashboard</a>	<b>Sunburst</b>	<a href="#">Top 10 &amp; Bottom 10</a>	<a href="#">Operational Dashboard</a>	<a href="#">Bump Ch</a>
-------------------------------------	-------------------------------------	------------------------------------	-----------------	--	---------------------------------------	-------------------------

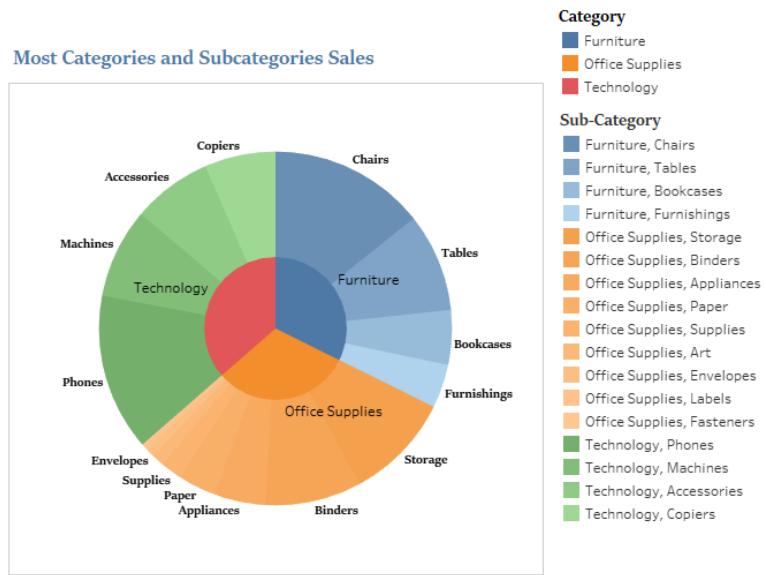
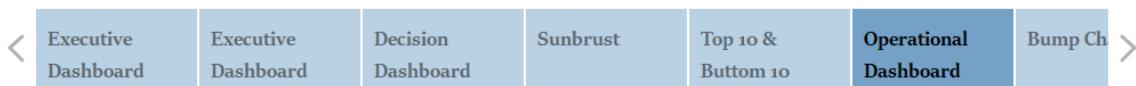


Figure 35: Dashboard 5

## Visulization on Supermarket



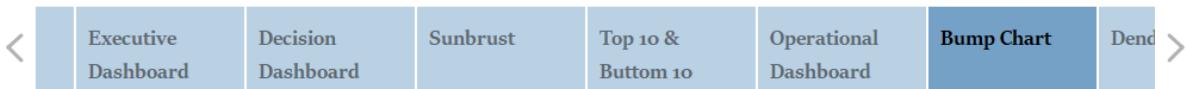
Actual/Forecast	Start Date	End Date	Select Country A
<input type="checkbox"/> Actual <input checked="" type="checkbox"/> Forecast	2000	2027	China
	<input type="button" value="&lt;"/>	<input type="button" value="&gt;"/>	<input type="button" value="&lt;"/>

### Sales Forecast by Year

Year of Year	Country A Population along Table (Down)	Country A YOY Change along Table (Down)	Country B Population along Table (Down)	Country B YOY
2000	1,262,645,000		1,042,261,758	
2001	1,271,850,000	0.73%	1,050,500,888	
2002	1,280,400,000	0.67%	1,076,705,723	
2003	1,288,400,000	0.63%	1,093,786,762	
2004	1,296,075,000	0.60%	1,110,626,108	
2005	1,303,720,000	0.57%	1,127,143,548	
2006	1,311,020,000	0.55%	1,143,289,350	
2007	1,317,885,000	0.53%	1,150,095,250	
2008	1,324,655,000	0.51%	1,174,662,334	
2009	1,331,260,000	0.50%	1,190,138,069	
2010	1,337,705,000	0.48%	1,205,624,648	
2011	1,344,130,000	0.48%	1,221,156,319	
2012	1,350,695,000	0.49%	1,236,686,732	
2013	1,317,602,973	-2.45%	1,221,846,491	
2014	1,285,321,700	-2.45%	1,207,184,333	
2015	1,253,831,318	-2.45%	1,192,609,121	
2016	1,223,112,451	-2.45%	1,178,385,744	
2017	1,193,146,196	-2.45%	1,164,245,115	
2018	1,163,914,114	-2.45%	1,150,274,174	
2019	1,135,398,218	-2.45%	1,136,470,883	

Figure 36: Dashboard 6

## Visulization on Supermarket



Ranking Analysis For Sub-category

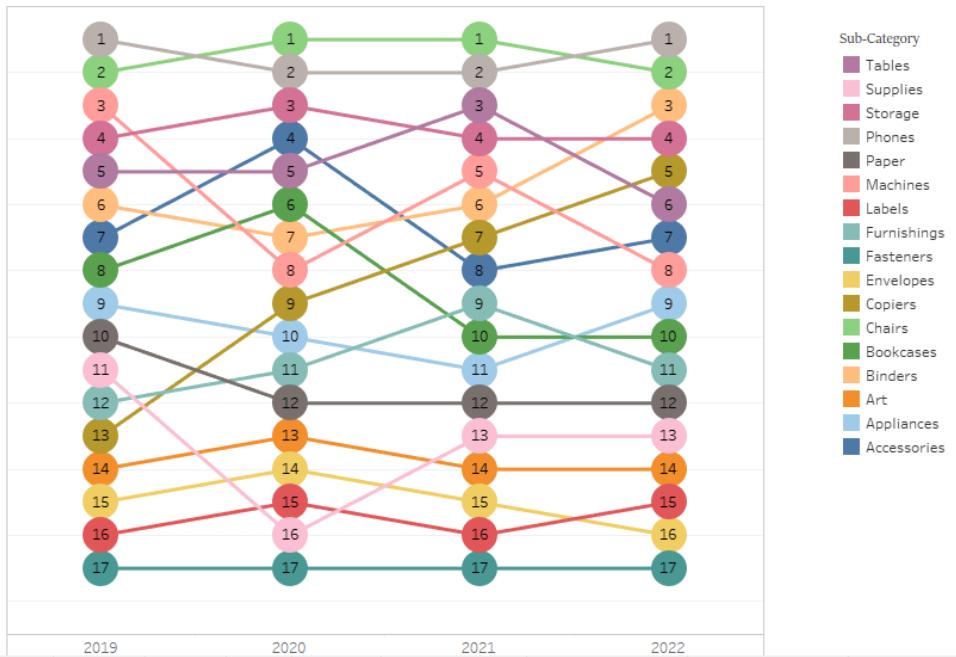


Figure 38: Dashboard 7

## Visulization on Supermarket



Year Profit

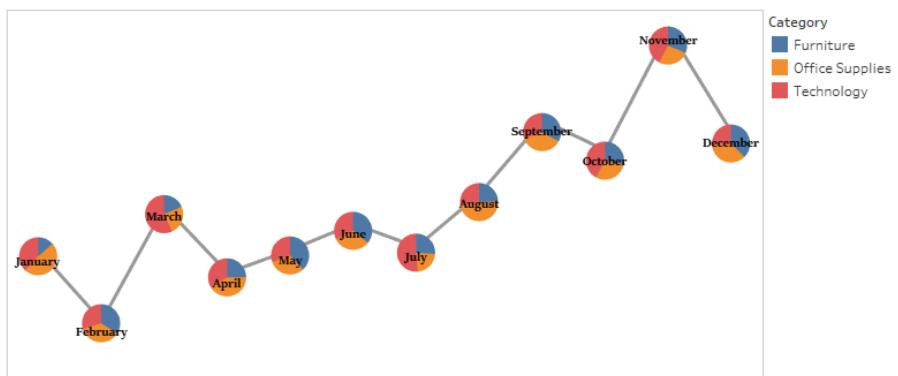


Figure 37: Dashboard 8

## Visulization on Supermarket

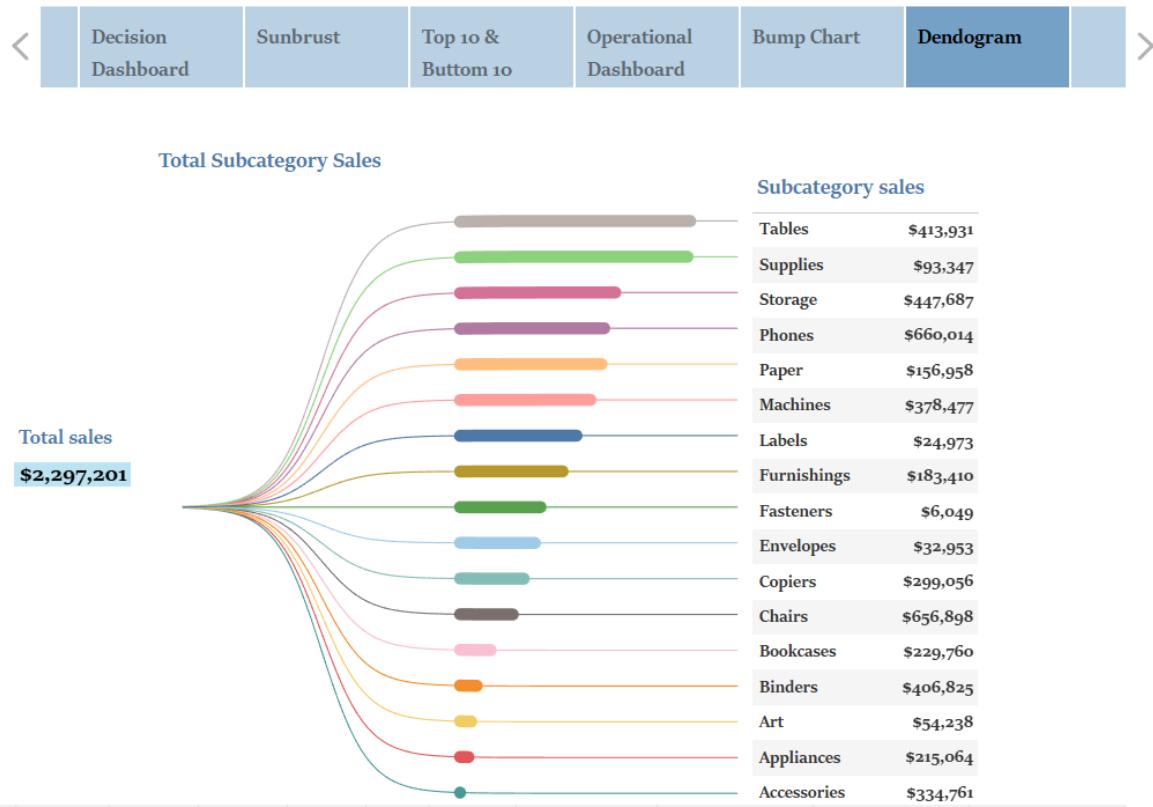


Figure 39: Dashboard 10

## Finding and Discussion

### How data analytics is useful for supermarket?

Data analysis is defined as the scientific study of evaluating rows of data to draw conclusions about information and visualization. Data analytics entails a wide variety of methodologies, many of which have been transformed into mechanical processes and algorithms that operate on raw data for human consumption. The supermarket industry is one of the fastest-growing in the commercial world, and with that comes a greater appreciation for the importance of data analysis, which may facilitate better decision-making and provide a clearer picture of the issue at hand. When used correctly, data analytics may aid a company in achieving its performance goals, whether that be increasing efficiency, increasing profits, or making better strategic choices. Data analytics may be used to investigate what occurred (descriptive analytics), what caused an event to occur (diagnostic analytics), what will occur in the future (predictive analytics), and what actions should be taken (prescriptive analytics) (prescriptive analytics). Spreadsheets, data visualization and reporting tools, data mining applications, and open-source languages are only some of the software tools used in data analytics. Business performance may be greatly improved with the use of data analytics. By integrating technology into their operations, businesses will be able to save money

by discovering new, more cost-effective methods of conducting business and by storing vast volumes of information. Data analytics may also aid in the development of new and improved goods and services by providing insight into client behavior and opinions.

## **How does data visualization help to take a better decision and improve business performance?**

Our brains swiftly analyze images while investigating and locating. This project created a tableau visualization and dashboard. Data visualization includes graphically representing data. When human look at statistics, human must read, absorb, and understand the data in context. When humans look at statistics, humans must read, absorb, and understand the data in context. Data visualizations simplify the steps our brains must take to understand data by using our innate capacity to perceive our environment. Humans grasp things better whether they're closer, further away, smaller, larger, taller, or shorter. Visualizations may transport information swiftly, helping our brains build connections quicker and improving decision-making. Data visualization saves time by simplifying enormous amounts of data so they can be understood quickly. Graphs provide real-time data analysis. This lets you monitor business performance and respond as required. It helps commercial decision-makers absorb information quickly. This helps them see problems and patterns, shortening response times. Politicians and fast-paced companies like retail and healthcare need speedier decisions. These companies need quick choices from their decision-makers. Once a choice is made, it must be communicated to carry out future processes. Data visualization may also ease communication. Data visualization may be accurate and easy to understand, increasing accuracy. To draw accurate conclusions from the research, use the right data visualization. Visual and contextual data help decision-makers focus on the essentials and define priorities. Even though all data is shown simultaneously, the most significant information is highlighted.

## **What are the ethical issue related to the application of Hardware Retail Information System in retail sector?**

All superstore data is in HRIS, says a study. This sector's HRIS raises several ethical issues. Restricting data exchange, protecting data, correcting data. Companies have an ethical duty to gather just the information they need, keep it safe, limit its spread, and correct any misuse. Retail Personally Identifiable Information (PII) is similar to other firms' PII. This is harmful for the retail environment and the company's image and finances. Ethical management of consumer data may help organization security and the retail business. Businesses must pay and treat workers decently. Randomly cutting hours and expecting workers to do more is common. Retailers may dismiss concerns or terminate workers for speaking out. Workers should minimize the collection of unnecessary client data, such as browser histories and transaction records, and should not share this information with other parties (even associated businesses). When providing personal information to a business, customers have an ethical obligation. Among these include providing accurate and complete information and refraining from exploiting or disclosing proprietary corporate information.

# Project Plan

The screenshot shows a project management interface with the following details:

**Project Title:** Data Management techniques and analytical improv...

**Navigation:** Overview, List, Board, Timeline, Calendar, Workflow, Dashboard, Messages, Files

**Search and Filter:** Share, Search, +, su

**Task List:**

Task name	Assignee	Due date
✓ Research on the Project	suyeshshres...	May 20 – 23
✓ Topic Chosen	suyeshshres...	May 2 – 7
✓ Project Initialization	suyeshshres...	May 1 – 5

Add task...

**Topic Selection Process**

✓ Research on the Project	suyeshshres...	May 20 – 23
✓ Topic Chosen	suyeshshres...	May 2 – 7
✓ Project Initialization	suyeshshres...	May 1 – 5

Add task...

**Proposal Process**

✓ Proposal Draft	suyeshshres...	May 9 – 13
✓ Final Proposal	suyeshshres...	Aug 14 – 18

Add task...

Figure 40: Project Plan 1

**Data Management techniques and analytical improv...**

Task name	Assignee	Due date	
<b>▼ Dissertation Writing Process</b>			
✓ Research on the Project and Resources	su suyeshshres...	Aug 1 – 27	
✓ Introduction	su suyeshshres...	Jul 13 – 17	
✓ Aim and Objectives	su suyeshshres...	Jul 15 – 17	
✓ Justification	su suyeshshres...	May 22 – 28	
✓ Scope	su suyeshshres...	May 25 – 27	
✓ Case Studies	su suyeshshres...	May 30 – Jun 22	
✓ Ethical Consideration	su suyeshshres...	May 26 – 31	
✓ Research Methodology	su suyeshshres...	Aug 20 – 23	
✓ Development Methodology	su suyeshshres...	Aug 18 – 21	
✓ Tools, Technology, and Techniques	su suyeshshres...	Aug 2 – 6	

Figure 41: Project Plan 2

**Data Management techniques and analytical improv...**

Task name	Assignee	Due date	
<b>▼ Development Process</b>			
✓ Finding data from Kaggle	su suyeshshres...	Aug 23 – 28	
✓ Data Wrangling - ETL	su suyeshshres...	Jul 27 – 31	
✓ Build and Installation of Oracle Linux 6.7	su suyeshshres...	Aug 1 – 6	

Figure 42: : Project Plan 3

**Data Management techniques and analytical improv...**

Overview List Board Timeline Calendar Workflow Dashboard Messages Files

+ Add task

All tasks Filter Sort Customize

Task name	Assignee	Due date
Conclusion	su suyeshshres...	Aug 21 – Today
Abstract	su suyeshshres...	Aug 18 – 21

Add task...

**Development Process**

Finding data from Kaggle	su suyeshshres...	Aug 23 – 28
Data Wrangling - ETL	su suyeshshres...	Jul 27 – 31
Build and Installation of Oracle Linux 6.7	su suyeshshres...	Aug 1 – 6
Creating Oracle Database and add user security and system security	su suyeshshres...	Aug 3 – 15
Create dashboard and Story in Tableau Desktop	su suyeshshres...	Aug 11 – 21

Add task...

+ Add section

Figure 43: Project Plan 4

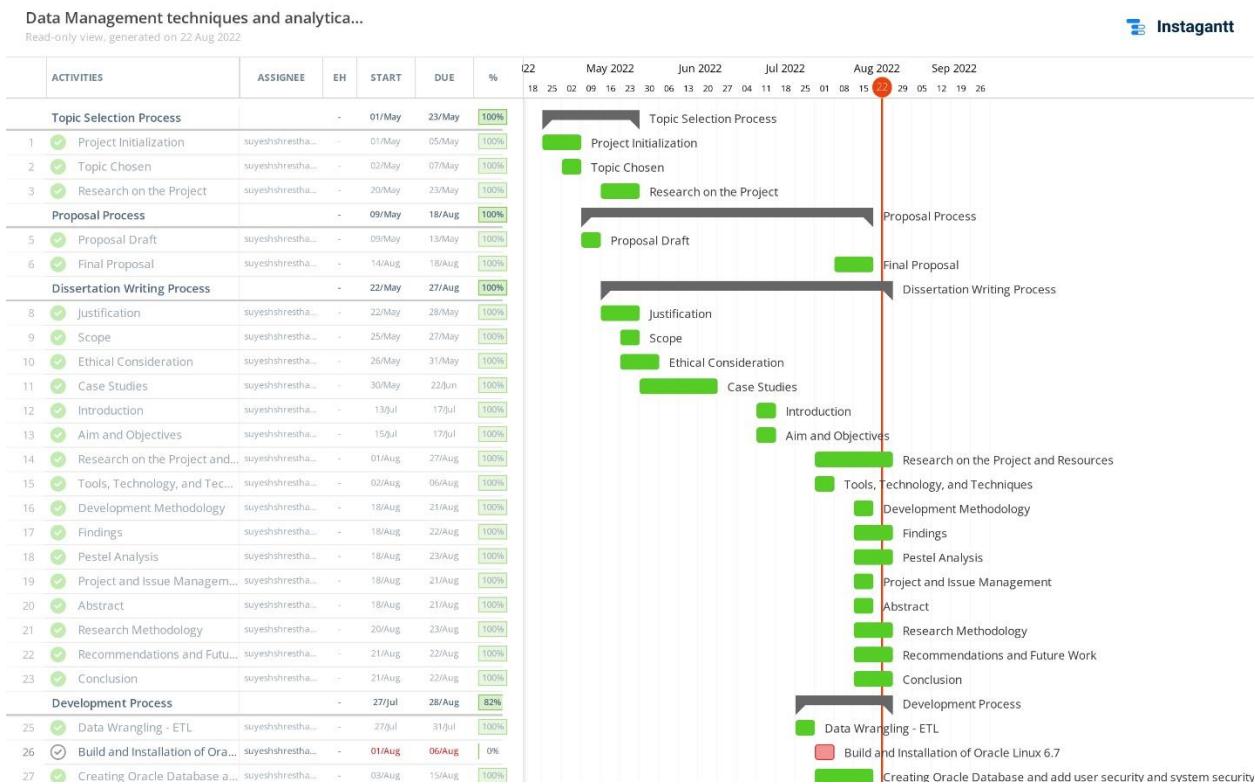


Figure 44: Project Plan 5

## Issue Log

S.N	Risk Description	Source of Occurrence	Impact	Priority	Solution
1.	Inappropriate research topic chosen	Lack of background information in deciding on a research project	High	High	Find a subject of interest that focuses on the problem you are experiencing.
2.	Insufficient understanding of the research subject	Improper the research done on the chosen topic	High	Medium	Different websites, books, scholarly articles, and related projects were researched and evaluated.
3.	Insufficient storage	Less amount of storage allocated	Medium	Low	External hard drive was used.
4.	Failure to follow procedures	Lack of dissertation knowledge.	High	High	Accept feedback from supervisor
5.	Resources Limitation	Lack of research resources, tools and other resources	High	Medium	Try to adapt to the resources provided.
6.	Database Corruption	Human error and improper scheduling	Medium	Medium	Backup data and maintained database.
7.	Difficult to design Dashboards	Lack of experience in tools and experiences.	High	High	Multiple charts and dashboards were practice.
8.	Inaccurate Estimates	Research requires more time than expected.	Medium	Low	Estimating time manage and finishing the project in proper time.

Figure 45: Issue Log

## Conclusion

One of the most competitive and quickly developing industries in the modern world is undoubtedly supermarket. The operation of this company results in the collection of a large amount of data, which may include information on products, customers, transactions, and other forms of data. As a result, they retain vast quantities of the data that they have acquired, use analytical methods helped by data visualization tools, and extract vital insights in order to make educated judgments on their company's operations. Data-driven initiatives are becoming an increasingly popular option for growth-oriented firms. The use of data visualization may help firms achieve greater success and development by improving their judgment, which in turn facilitates better decision making. Research involving the display of data for supermarkets.

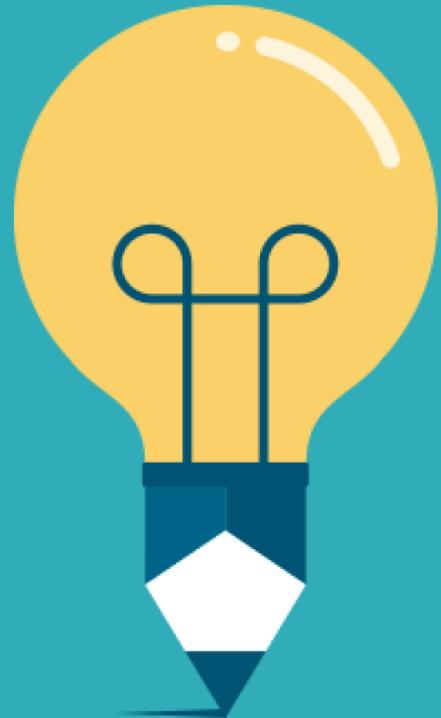
The purpose of this study is to improve the business via data analytics and visualization. Multiple sources were consulted, including scholarly articles, books, and online databases, to determine the research's focus, gather relevant data, and explain the study's findings. This investigation makes use of a wide range of methods and apparatus. Oracle databases are used for managing and protecting sensitive data. A variety of methods, including a data warehouse, automated storage management, data security, system security, data recovery and backup, and so on, are used to ensure the integrity of the stored data and the integrity of the

data management itself. The Tableau software is used for data visualization, whereas Tableau Prep is used for data cleaning, filtering and preparation. The diagrams, charts, and graphs that make up a dashboard might vary in form and style based on the specific use case being analyzed. In summary, this study shows that the inclusion of data visualization, analytical methods, and data management may boost supermarket profits.

## **Future Work**

This research study was an academic level project; yet, its technology, findings, and research were restricted due to a lack of supervision, expertise, and resources. This project has the potential to be improved in the future by doing thorough research, making further efforts, and receiving feedback from the supervisor. Small data are utilized for this research; but, in the future, huge datasets that are updated in real time will be used for improved decision making and visualization. In the not-too-distant future, an Oracle server will be built inside of a virtual box in order to provide enhanced data protection and make advantage of the innovative Oracle data Guard technology. The dashboards are exported to the public server of Tableau, where they are available for anybody to view, download, and make changes to. This might result in data breaches, which is a concern in a high-profile private firm where ensuring the confidentiality of customer information is essential. It is recommended that a Tableau server be purchased in order to enhance and save dashboards, graphs, and charts that are privately stored and cannot be seen. It is planned to construct a responsive dashboard, graph, and chart for mobile applications as well as various sizes of screens, which will then be linked to direct tableau server.

## Future Works



- 1 Implementing on Lagge data set
- 2 Implementing Oracle Data Guard
- 3 Seprate Dedicated Server for Oracle Database
- 4 Hosting Tableau Chard on Private Server
- 5 Creating more enhanced charts and dashboard
- 6 Developing a web and mobile application

Figure 46: Future Works

## Bibliography

- Management, S., Technology, S., States, U. and Kingdom, U., 2022. *Data Analytics Case Study Of Coles Supermarket / Total Assignment Help*. [online] Totalassignmenthelp.com. Available at: <<https://www.totalassignmenthelp.com/free-sample/data-analytics-case-study-of-coles-supermarket>> [Accessed 14 February 2022].
- Wrike.com. 2022. *What Is Agile Methodology in Project Management?*. [online] Available at: <<https://www.wrike.com/project-management-guide/faq/what-is-agile-methodology-in-project-management>> [Accessed 20 February 2022].
- PGBS. 2022. *Six Steps in CRISP DM - The Standard Data Mining Process*. [online] Available at: <<https://www.proglobalbusinesssolutions.com/six-steps-in-crisp-dm-the-standard-data-mining-process>> [Accessed 23 February 2022].
- Medium. 2022. *CRISP-DM Methodology For Your First Data Science Project*. [online] Available at: <<https://towardsdatascience.com/crisp-dm-methodology-for-your-first-data-science-project-769f35e0346c>> [Accessed 29 February 2022].
- Sridharan, M. and Sridharan, M., 2022. *CRISP-DM - A Framework For Data Mining & Analysis*. [online] Think Insights. Available at: <<https://thinkinsights.net/data-literacy/crisp-dm/>> [Accessed 2 March 2022].
- Oracle.com. 2022. *Can Data Analytics Improve Business Decisions?*. [online] Available at: <<https://www.oracle.com/business-analytics/data-analytics/>> [Accessed 7 March 2022].
- Kumari, R., 2022. *What is Data Analytics and its types / Analytics Steps*. [online] Analyticssteps.com. Available at: <<https://www.analyticssteps.com/blogs/what-data-analytics-and-its-types>> [Accessed 9 March 2022].
- Chartio. 2022. *Types of Data Analysis*. [online] Available at: <<https://chartio.com/learn/data-analytics/types-of-data-analysis/>> [Accessed 12 March 2022].
- Oracle Help Center. 2022. *Enterprise User Security Administrator's Guide*. [online] Available at: <<https://docs.oracle.com/en/database/oracle/oracle-database/19/dbimi/introducing-enterprise-user-security.html#GUID-080DE6BD-8D54-49EC-BAC4-4EC5CC6A25BF>> [Accessed 18 March 2022].
- Oracle.com. 2022. *What is a Data Warehouse?*. [online] Available at: <<https://www.oracle.com/database/what-is-a-data-warehouse/>> [Accessed 25 March 2022].
- Education, I., 2022. *What is ETL (Extract, Transform, Load)?*. [online] Ibm.com. Available at: <<https://www.ibm.com/cloud/learn/etl>> [Accessed 8 April 2022].

- Docs.oracle.com. 2022. *Oracle Help Center*. [online] Available at: <<https://docs.oracle.com/en/operating->>> [Accessed 12 March 2022].
- Docs.oracle.com. 2022. *Backup and Recovery*. [online] Available at: <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14220/backrec.htm](https://docs.oracle.com/cd/B19306_01/server.102/b14220/backrec.htm)> [Accessed 25 March 2022].
- Investopedia. 2022. *What Is Data Analytics?*. [online] Available at: <<https://www.investopedia.com/terms/d/data-analytics.asp>> [Accessed 29 March 2022].
- Bhasin, H., 2022. *What are Ethical Considerations in Research?*. [online] Marketing91. Available at: <[https://www.marketing91.com/ethical-considerations/#:~:text=Ethical%20consideration%20is%20a%20collection,from%20indulging%20in%20vicious%20conduct.](https://www.marketing91.com/ethical-considerations/#:~:text=Ethical%20consideration%20is%20a%20collection,from%20indulging%20in%20vicious%20conduct.>)> [Accessed 7 April 2022].
- Docs.oracle.com. 2022. *Introduction to Oracle Automatic Storage Management*. [online] Available at: <[https://docs.oracle.com/cd/E11882\\_01/server.112/e18951/asmcon.htm#OSTMG036](https://docs.oracle.com/cd/E11882_01/server.112/e18951/asmcon.htm#OSTMG036)> [Accessed 25 May 2022].
- Oracle.com. 2022. *Recovery Manager (RMAN)*. [online] Available at: <<https://www.oracle.com/database/technologies/high-availability/rman.html>> [Accessed 5 June 2022].
- Docs.oracle.com. 2022. *Using Oracle Flashback Technology*. [online] Available at: <[https://docs.oracle.com/cd/E11882\\_01/appdev.112/e41502/adfns\\_flashback.htm#ADFNS1008](https://docs.oracle.com/cd/E11882_01/appdev.112/e41502/adfns_flashback.htm#ADFNS1008)> [Accessed 11 June 2022].
- Hunter.cuny.edu. 2022. [online] Available at: <<http://www.hunter.cuny.edu/ccpd/repository/files/walmart.pdf>> [Accessed 19 June 2022].
- BI Blog | Data Visualization & Analytics Blog | datapine. 2022. *Types of Dashboards: Strategic, Operational & Analytical*. [online] Available at: <<https://www.datapine.com/blog/strategic-operational-analytical-tactical-dashboards/>> [Accessed 16 June 2022].
- Docs.oracle.com. 2022. *Automatic Workload Repository*. [online] Available at: <<https://docs.oracle.com/database/121/RACAD/GUID-C3CD2DCE-38BD-46BA-BC32-7A28CAC9A7FD.htm#RACAD951>> [Accessed 15 July 2022].
- Shichor, S. and Shichor, S., 2022. *Walmart Retail Link System Talks to Suppliers Through Data*. [online] E-Commerce Strategies | Selling Online. Available at: <<https://18knowledge.com/blog/walmart-retail-link-talks-to-suppliers-through-data/>> [Accessed 6 July 2022].

Linkedin.com. 2022. *When Data Met Retail: A #lovedata story*. [online] Available at: <<https://www.linkedin.com/pulse/when-data-met-retail-lovedata-story-david-becker?articleId=6016964218585178112>> [Accessed 10 May 2022].

E-research.siam.edu. 2022. [online] Available at: <[http://e-research.siam.edu/wp-content/uploads/2017/02/6.2Chaper\\_2.pdf](http://e-research.siam.edu/wp-content/uploads/2017/02/6.2Chaper_2.pdf)> [Accessed 15 August 2022].

*CREATE PROFILE* (2021) available from <[https://docs.oracle.com/cd/E11882\\_01/server.112/e41084/statements\\_6010.htm#SQLRF01310](https://docs.oracle.com/cd/E11882_01/server.112/e41084/statements_6010.htm#SQLRF01310)> [16 April 2021]

*Privileges* (2021) available from <<https://docs.oracle.com/database/121/TTSQL/privileges.htm#TTSQL338>> [20 April 2022]

*Oracle Database 11G: The Top New Features For Dbas And Developers / Security* (2021) available from <<https://www.oracle.com/technical-resources/articles/database/sql-11g-security.html>> [25 April 2022]

*Managing Security For Oracle Database Users - 11G Release 2 (11.2)* (2021) available from <[https://docs.oracle.com/cd/E28271\\_01/network.1111/e16543/users.htm#i1006230](https://docs.oracle.com/cd/E28271_01/network.1111/e16543/users.htm#i1006230)> [26 April 2022]

*Database Security* (2021) available from <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14220/security.htm#i16445](https://docs.oracle.com/cd/B19306_01/server.102/b14220/security.htm#i16445)> [28 April 2022]

*Performing Backup And Recovery* (2021) available from <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14196/backrest.htm#ADMQS009](https://docs.oracle.com/cd/B19306_01/server.102/b14196/backrest.htm#ADMQS009)> [29 April 2022]

*RMAN Backup Types* (2021) available from <<https://web.stanford.edu/dept/itss/docs/oracle/10gR2/backup.102/b14191/rcmconc1005.htm>> [5 May 2022]

*Backup And Recovery* (2021) available from <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14220/backrec.htm](https://docs.oracle.com/cd/B19306_01/server.102/b14220/backrec.htm)> [7 May 2022]

*Managing A Recovery Catalog* (2021) available from <<https://docs.oracle.com/database/121/BRADV/rcmcatdb.htm#BRADV89657>> [8 May 2022]

*CREATE FLASHBACK ARCHIVE* (2021) available from <[https://docs.oracle.com/database/121/SQLRF/statements\\_5011.htm#SQLRF20008](https://docs.oracle.com/database/121/SQLRF/statements_5011.htm#SQLRF20008)> [9 May 2022]

*Managing The Undo Tablespace* (2021) available from <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14231/undo.htm](https://docs.oracle.com/cd/B19306_01/server.102/b14231/undo.htm)> [10 May 2022]

DPA (2021) [online] available from <<https://www.itpro.co.uk/data-protection/28085/what-is-the-data-protection-act-1998>> [15 May 2022]

*Using Oracle Flashback Technology* (2021) available from <[https://docs.oracle.com/database/121/ADFNS/adfns\\_flashback.htm#ADFNS598](https://docs.oracle.com/database/121/ADFNS/adfns_flashback.htm#ADFNS598)> [17 May 2022]

*Data Concurrency And Consistency* (2021) available from <[https://docs.oracle.com/cd/B19306\\_01/server.102/b14220/consist.htm](https://docs.oracle.com/cd/B19306_01/server.102/b14220/consist.htm)> [18 May 2021]

*DBMS Deadlock In DBMS - Javatpoint* (2021) available from <<https://www.javatpoint.com/deadlock-in-dbms>> [19 May 2022]

*Automatic Workload Repository* (2021) available from <<https://docs.oracle.com/database/121/RACAD/GUID-C3CD2DCE-38BD-46BA-BC32-7A28CAC9A7FD.htm#RACAD951>> [19 May 2022]

*Database Performance Tuning Guide* (2021) available from <[http://ora-srv.wlv.ac.uk/oracle19c\\_doc/tgdba/automatic-performance-diagnostics.html#GUID-F1FBA929-4FCF-4872-AEE6-2DF716B5BD43](http://ora-srv.wlv.ac.uk/oracle19c_doc/tgdba/automatic-performance-diagnostics.html#GUID-F1FBA929-4FCF-4872-AEE6-2DF716B5BD43)> [20 May 2022]

*Creating Operations Dashboards / Smartsheet* (2021) available from <<https://www.smartsheet.com/creating-operations-dashboards>> [27 September 2022]

*The 3 Types Of Dashboards: Operational, Analytical & Strategic* (2021) available from <<https://www.idashboards.com/blog/2018/08/15/operational-analytical-and-strategic-the-three-types-of-dashboards/>> [28 September 2022]

*Challenges In A Data-Driven World* (2021) available from <<https://www.knowit.eu/services/experience/digital-strategy-and-analysis/strategy-and-digitalization/knowits-take-on-tomorrow/challenges-in-a-data-driven-world/>> [29 September 2022]

*The 3 Types Of Dashboards: Operational, Analytical & Strategic* (2021) available from <<https://www.idashboards.com/blog/2018/08/15/operational-analytical-and-strategic-the-three-types-of-dashboards/>> [29 September 2022]

*3 Challenges Of Data Adoption* (2021) available from <<https://towardsdatascience.com/3-challenges-of-data-adoption-790a87ae3472>> [29 September 2022]

*Overcoming The Top 3 Analytics Adoption Challenges* (2021) available from <<https://searchbusinessanalytics.techtarget.com/feature/Overcoming-the-top-3-analytics-adoption-challenges>> [26 September 2022]

*What Is Retail Analytics? / HCL Technologies* (2021) available from <<https://www.hcltech.com/technology-qa/what-is-retail-analytics>> [29 September 2021]

*What Is Retail Analytics? / Sisense* (2021) available from <<https://www.sisense.com/glossary/retail-analytics/>> [30 September 2022]

*Analytics: Retail Analytics* (2021) available from <<https://www.spcommerce.com/what-is-retail-analytics/>> [1 October 2021]

23 Sep Ultimate Guide to Retail Analytics (Definition, T. (2021) *Ultimate Guide To Retail Analytics (Definition, Types, Examples)* - *Retalon* [online] available from <<https://retalon.com/blog/retail-data-analytics>> [1 October 2022]

(2021) available from <<https://www.vendhq.com/blog/how-retailers-can-use-data-to-boost-productivity-customer-service-sales/>> [1 October 2022]

(2021) [online] available from <<https://www.yodlee.com/data-analytics/big-data-retail-analytics>> [ 1 October 2022]

Marr, B. (2017) *Big Data In Practice*. Chichester, West Sussex: Wiley

Pathak, R. (2021) *How Amazon Uses Big Data? / Analytics Steps* [online] available from <<https://www.analyticssteps.com/blogs/how-amazon-uses-big-data>> [1 October 2022]

Alter, A. (2021) Irresistible - The Rise Of Addictive Technology And The Business Of Keeping Us Hooked. Penguin

Helbing, D. (2021) Thinking Ahead - Essays On Big Data, Digital Revolution, And Participatory Market Society. Springer

Shron, M. (2021) Thinking With Data - How To Turn Information Into Insights. O'REILLY

*Data Management: What It Is And Why It Matters* (2021) available from <[https://www.sas.com/en\\_us/insights/data-management/data-management.html](https://www.sas.com/en_us/insights/data-management/data-management.html)> [2 October 2022]

Small Business - Chron.com. 2021. The Importance of an Audit System to Companies. [online] Available at: [Accessed 18 June 2022].

Oracle.com. 2021. Database Security Best Practices. [online] Available at: [Accessed 25 June 2021]. Docs.oracle.com. 2021. Introduction to Oracle Database Security. [online] Available at: [Accessed 30 June 2022].

Docs.oracle.com. 2021. Backup and Recovery. [online] Available at: [Accessed 5 July 2021]. Kuhn, D., Alapayi, S. and Nanda, A., n.d. RMAN recipes for Oracle database 11g. > [Accessed 8 July 2022].

Dba-oracle.com. 2021. RMAN backup types. [online] Available at: [Accessed 12 July 2021]. Easeus.com. 2021. Differential Backup VS Incremental Backup - EaseUS Todo Backup. [online] Available at: [Accessed 15 July 2022].

Gunukula, S. and Gunukula, S., 2022. Why Virtual Private Catalog? - Simple Talk. [online] Simple Talk. Available at: [Accessed 19 July 2022]. Oracle.com. 2022. Flashback Technologies | Oracle. [online] Available at: [Accessed 22 July 2022]

. Docs.oracle.com. 2022. Using Oracle Flashback Technology. [online] Available at: [Accessed 22 July 2022]. SearchDataBackup. 2021. What is Oracle Flashback Technology and is it a form of backup?. [online] Available at: [Accessed 27 July 2022].

Dba-oracle.com. 2022. Flashback Data Archive (FDA) tips. [online] Available at: [Accessed 30 July 2022]. Docs.oracle.com. 2022. Using Oracle Flashback Technology. [online] Available at: [Accessed 4 August 2022].

Docs.oracle.com. 2022. Data Concurrency and Consistency. [online] Available at: [Accessed 8 August 2022].

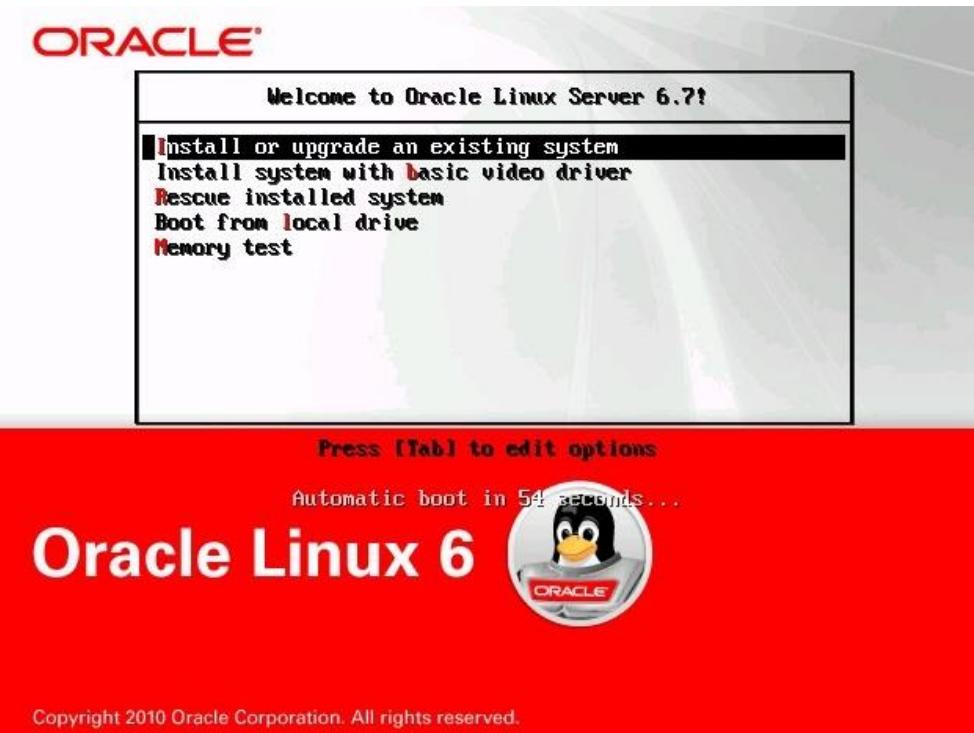
Tableau. 2022. Data visualization beginner's guide: a definition, examples, and learning resources. [online] Available at: [Accessed 14 February 2022]. Sas.com. 2022. Data Visualization: What it is and why it matters. [online] Available at: [Accessed 14 February 2022].

## Appendix

### Tableau Link:

[https://public.tableau.com/views/SupermarketDashboardforThesis/Finalstory  
?:language=en-US&publish=yes&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/SupermarketDashboardforThesis/Finalstory?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link)

### Oracle Installation 6.7



Screen displayed by the installation program depends on whether you selected the Legacy BIOS Boot Mode or the UEFI Boot Mode.

What type of devices will your installation involve?

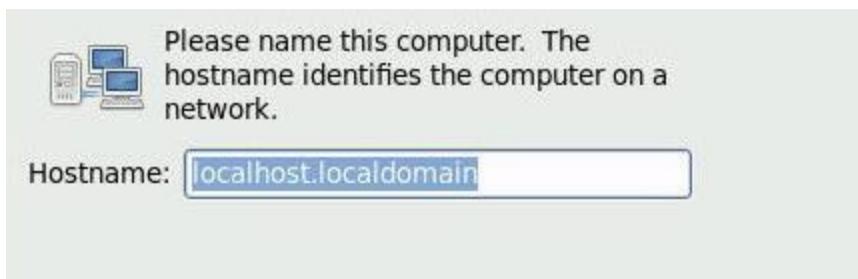
**Basic Storage Devices**

- Installs or upgrades to typical types of storage devices. If you're not sure which option is right for you, this is probably it.

**Specialized Storage Devices**

- Installs or upgrades to enterprise devices such as Storage Area Networks (SANs). This option will allow you to add FCoE / iSCSI / zFCP disks and to filter out devices the installer should ignore.

Above screen, select Specialized Storage Devices



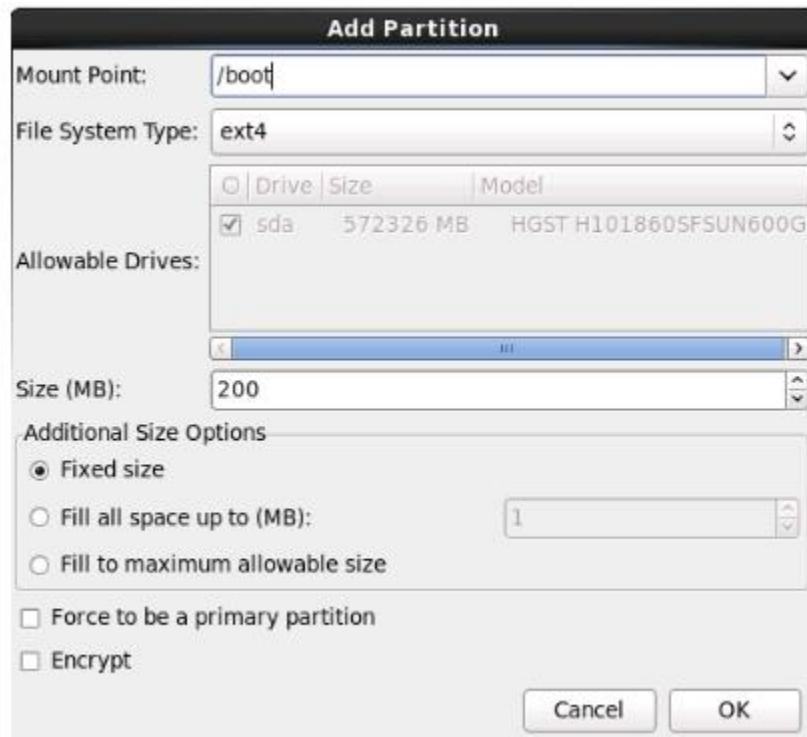
Entering Host name



Select the appropriate region and city



Select Standard Partition, and click



Set the Mount point to /boot and leave the File System Type set to ext4 and the Size (MB) set to 200.

Hard Drives				
sda (/dev/sda)				
sda1	200	/boot/efi	ext4	✓
sda2	200	/boot	ext4	✓
sda3	40000	/	ext4	✓
sda4	20000		swap	✓
Free	511924			

Click Next to apply the partitions

The screenshot shows the partitioning step of the Oracle Linux Server installation. It displays a table of partitions on disk sda. The table includes columns for the partition name, size, mount point, file system type, and a checkmark column indicating they are correctly configured. The 'sda3' partition is mounted at '/'.

**Partition Table:**

Partition	Size	Mount Point	Type	Status
sda1	200	/boot/efi	ext4	✓
sda2	200	/boot	ext4	✓
sda3	40000	/	ext4	✓
sda4	20000		swap	✓
Free	511924			

**Server Selection:**

Please select any additional repositories that you want to use for software installation.

Basic Server  
 Database Server  
 Web Server  
 Identity Management Server  
 Virtualization Host  
 Desktop  
 Software Development Workstation  
 Minimal

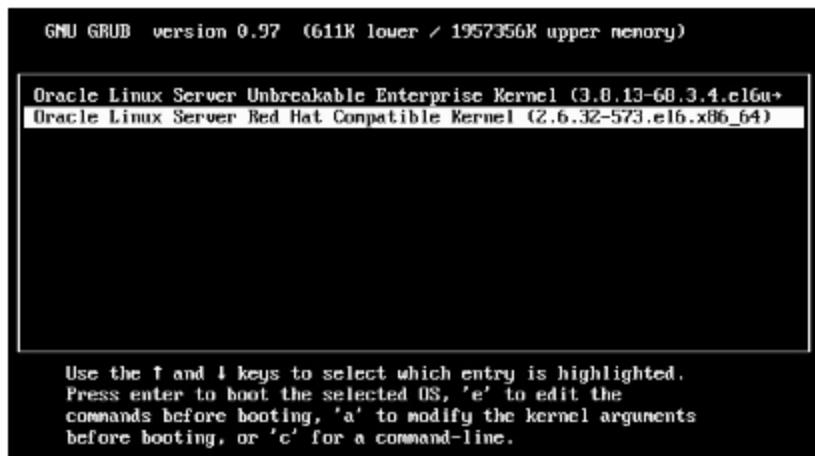
High Availability  
 Load Balancer  
 Oracle Linux Server

[+ Add additional software repositories](#) [Modify repository](#)

For the purposes of this sample installation, accept the Basic Server default



Wait until the Oracle Linux OS installation completes.



Choose Oracle Linux Server Red Hat Compatible Kernel (2.6.32-573.el6.x86\_64)

```

[UEK4]
name=UEK4
baseurl=http://public-yum.oracle.com/repo/oraclelinux/OL6/UEK4/x86_64
enabled=1
skipcheck=1
12. Run yum update kernel-uek or yum update
The update screen appears:

[root@engash-dhcp-05-47 ~]# yum update kernel-uek
Loaded plugins: security, ulmInfo
Setting up Update Process
UEK4
UEK4/primary
UEK4
Localrepo
Resolving Dependencies
-> Running transaction check
--> Package kernel-uek.x86_64 0:4.1.12-32.el6uek will be installed
--> Processing dependency: kernel-firmware = 4.1.12-32.el6uek for package: kernel-uek-4.1.12-32.el6uek.x86_64
--> Processing dependency: kernel = 4.1.12-32.el6uek for package: kernel-uek-4.1.12-32.el6uek.x86_64
--> Package kernel-firmware.noarch 0:2.6.32-573.el6 will be obsoleted
--> package kernel-uek-firmware.noarch 0:4.1.12-32.el6uek will be installed
--> Package linux-firmware.noarch 0:20140911-0.1.git365e80c.0.7.el6 will be obsoleting
--> package q12300-firmware.noarch 0:3.03.27.3.1.el6 will be obsoleted
--> Package q12300-firmware.noarch 0:3.03.27.3.1.el6 will be obsoleted
--> Package q12300-firmware.noarch 0:3.03.27.3.1.el6 will be obsoleted
--> Package q12300-firmware.noarch 0:7.05.00-1.el6 will be obsoleted
--> Package q12300-firmware.noarch 0:7.05.00-1.el6 will be obsoleted
--> package rt73uhf-firmware.noarch 0:1.8.7.el6 will be obsoleted
--> Finished Dependency Resolution
Dependencies Resolved

Available Packages
Package           Arch      Version            Repository      Size
installing:
kernel-uek        x86_64   4.1.12-32.el6uek          UEK4           50 M
linux-firmware    noarch   20140911-0.1.git365e80c.0.7.el6  UEK4           19 M
replacing kernel-firmware.noarch 2.6.32-573.el6
replacing q12300-firmware.noarch 1.15.2-1.15.3.1.el6
replacing q12300-firmware.noarch 1.15.2-1.15.3.1.el6
replacing q12300-firmware.noarch 3.03.27.3.1.el6
replacing q12300-firmware.noarch 7.04.00-1.el6
replacing q12300-firmware.noarch 7.05.00-1.el6
replacing rt73uhf-firmware.noarch 1.2-1.el6
replacing rt73uhf-firmware.noarch 1.8-7.el6
installing for dependencies:
kernel-uek-firmware    noarch   4.1.12-32.el6uek          UEK4           1.8 M

Transaction Summary
=====
Install   3 Package(s)

Total download size: 72 M
Is this ok [y/N]:
```

The update is complete, reboot the server with the reboot command.  
The server starts a new Oracle Unbreakable Enterprise Kernel.

## Configure Automatic Storage Management

Automatic Storage Management: +ASM\_dbhost.example.com

Home Performance Disk Groups Configuration Users ASM Cluster File System

Data Retrieved May 1, 2009 12:11:53 PM PDT Refresh

**General**

Current Status Up  
Up Since Apr 24, 2009 11:59:16 PM PDT  
Availability (%) 100  
(Last 24 hours)  
Instance Name +ASM  
Version 11.2.0.0.2  
Host dbhost.example.com  
Oracle Home /u01/app/oracle/product/11.2.0/grid

**Disk Group Usage (GB)**

Unallocated(0.29)	Internal(0.06)	orcl.example.com(3.65)
-------------------	----------------	------------------------

**Diagnostic Summary**

Alert Log No ORA- errors  
Active Incidents 0

**Serviced Databases**

Name	Disk Groups	Failure Groups	Allocated Space (GB)	Availability	Alerts
orcl.example.com	DATA	n/a	3.65	0	1

**Serviced ASM Cluster File Systems**

Mount Point	Availability	State	Used (%)	Used (GB)	Size (GB)	Allocated Space (GB)	Volume	Disk Group
(No ASM Cluster File Systems)								

**Other Volumes**

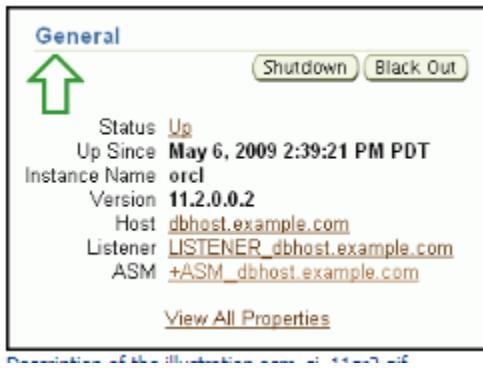
Volume	Volume Device	Usage	State	Disk Group	Size (GB)	Allocated Space (GB)	Redundancy
No object found							

**Alerts**

Severity	Category	Name	Impact	Message	Alert Triggered
×	Disk Group Usage	Disk Group Used % of Safely Usable		Disk group DATA has used 92.87% of safely usable free space.	May 1, 2009 12:08:01 PM

All Oracle Automatic Storage Management (Oracle ASM) administration tasks begin with the Oracle ASM Home page in Database Control.

- The status of the Oracle ASM instance
- A chart that shows the used and free space of each disk group
- A list of databases and Oracle ASM Cluster File Systems serviced by the Oracle ASM instance



## Accessing the Oracle ASM Home Page on a Single-Instance System

The screenshot shows the Oracle Enterprise Manager Disk Groups page. It displays the following information:

Select	Name	State	Redundancy	Size (GB)	Used (GB)	Used (%)	Usable Free (GB)	Member Disks
<input type="checkbox"/>	DATA	MOUNTED EXTERN		4.00	2.70	67.58	1.30	4
<input type="checkbox"/>	RECOVERY	MOUNTED EXTERN		4.00	1.72	43.12	2.28	4

**TIP** The usable free space specifies the amount of space that can be safely used for data. A value above zero means that redundancy can be properly restored after a disk failure.

**TIP** Mount All and Dismount All operation will only mount and dismount the disk groups specified in the Auto Mount Disk Groups parameter.

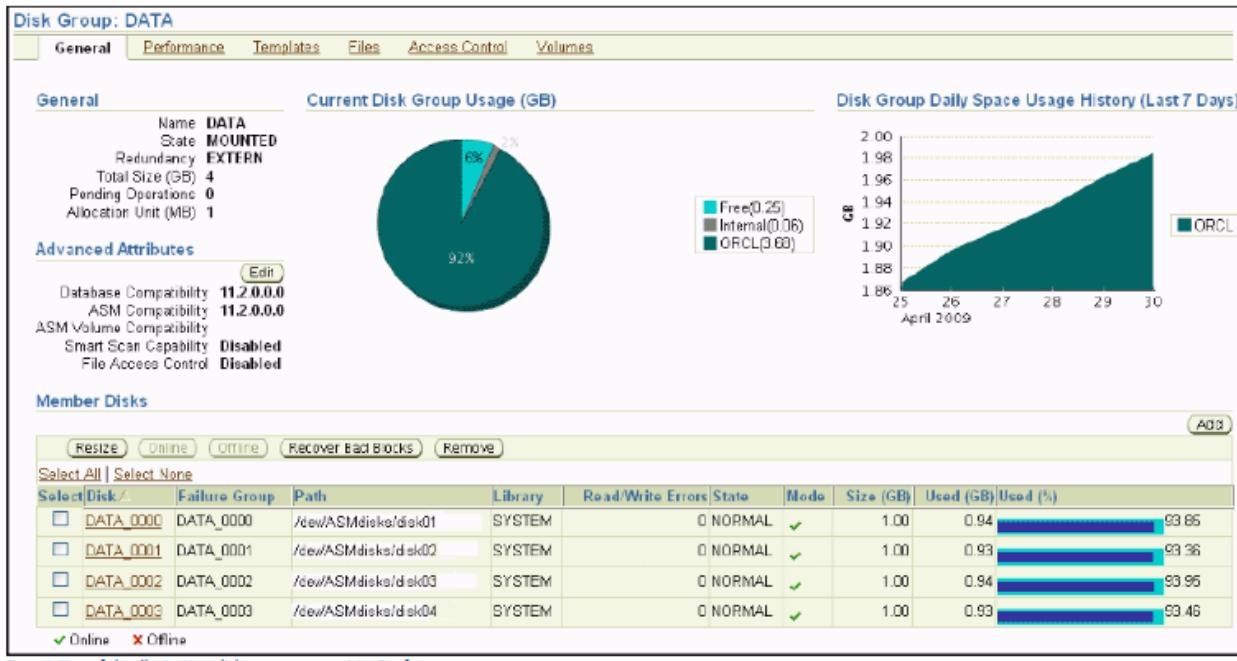
Using Oracle Enterprise Manager, you can monitor the total capacity of your disk groups, including the amount of unused space and the amount of unused space that can be safely utilized after taking mirroring into account.

The screenshot shows the Oracle Enterprise Manager Create Disk Group dialog box. It includes the following fields:

- \* Name:
- Redundancy:  HIGH  NORMAL  EXTERNAL
- Allocation Unit (MB):
- Member Disks: A table showing three available disks:
 

Select Path	Header Status	Library	Label	ASM Disk Name	Size	Unit	Failure Group	Force Reuse
/dev/ASMDisks/disk05	CANDIDATE	SYSTEM			256	MB		<input type="checkbox"/>
/dev/ASMDisks/disk06	CANDIDATE	SYSTEM			256	MB		<input type="checkbox"/>
/dev/ASMDisks/disk07	FORMER	SYSTEM			1024	MB		<input type="checkbox"/>

Create a disk group



**Add Disks**

Show SQL Cancel OK

**Rebalance Power**

Rebalance operations redistribute data evenly across all drives. ASM automatically rebalances a disk group whenever disks are added or dropped. To manually rebalance all disk groups, you must specify the Rebalance Power. Higher values use more I/O bandwidth and complete rebalance more quickly. Lower values cause rebalance to take longer, but use less I/O bandwidth. Values range from 0 to 11.

Rebalance Power: 1

**Member Disks**

Select Member Disks Only Candidate Disks

Select Path	Header Status	Library	Label	ASM Disk Name	Size	Unit	Failure Group	Force Reuse
<input type="checkbox"/> /dev/ASMDisks/disk05	CANDIDATE	SYSTEM			256	MB		<input type="checkbox"/>
<input type="checkbox"/> /dev/ASMDisks/disk06	CANDIDATE	SYSTEM			256	MB		<input type="checkbox"/>
<input type="checkbox"/> /dev/ASMDisks/disk07	FORMER	SYSTEM			1024	MB		<input type="checkbox"/>
<input type="checkbox"/> /dev/ASMDisks/disk08	FORMER	SYSTEM			1024	MB		<input type="checkbox"/>

## Adding Disks to a Disk Group

## Security

**Roles**

Object Type: Role

Search: Enter an object name to filter the data that is displayed in your results set.

Object Name: Go

By default, the search returns all uppercase matches beginning with the string you entered. To run an exact or case-sensitive match, double quote the search string. You can use the wildcard symbol (%) in a double quoted string.

Selection Mode: Single

Create

Edit View Delete Actions Create Like Go Previous 1-25 of 51 Next 25

Select	Role	Authentication
<input checked="" type="radio"/>	ADM_PARALLEL_EXECUTE_TASK	NO
<input type="radio"/>	AQ_ADMINISTRATOR_ROLE	NO
<input type="radio"/>	AQ_USER_ROLE	NO
<input type="radio"/>	AUTHENTICATEDUSER	NO
<input type="radio"/>	CONNECT	NO
<input type="radio"/>	CSW_USR_ROLE	NO
<input type="radio"/>	CTXAPP	NO
<input type="radio"/>	CWM_USER	NO
<input type="radio"/>	DATAPUMP_EXP_FULL_DATABASE	NO
<input type="radio"/>	DATAPUMP_IMP_FULL_DATABASE	NO

### Viewing Roles

Database Instance: orcl.example.com > Roles > Logged in As SYSTEM

Create Role

Show SQL Cancel OK

General Roles System Privileges Object Privileges Consumer Group Privileges

System Privilege Admin Option Edit List

No items found

General Roles System Privileges Object Privileges Consumer Group Privileges

### Creating a Role

Database Instance: orcl.us.oracle.com > Roles > Logged in As SYSTEM

Modify System Privileges

Cancel OK

Available System Privileges

ACCESS\_ANY\_WORKSPACE  
ADMINISTER ANY SQL TUNING SET  
ADMINISTER DATABASE TRIGGER  
ADMINISTER RESOURCE MANAGER  
ADMINISTER SQL MANAGEMENT OBJECT  
ADMINISTER SQL TUNING SET  
ADVISOR  
ALTER ANY ASSEMBLY  
ALTER ANY CLUSTER  
ALTER ANY CUBE

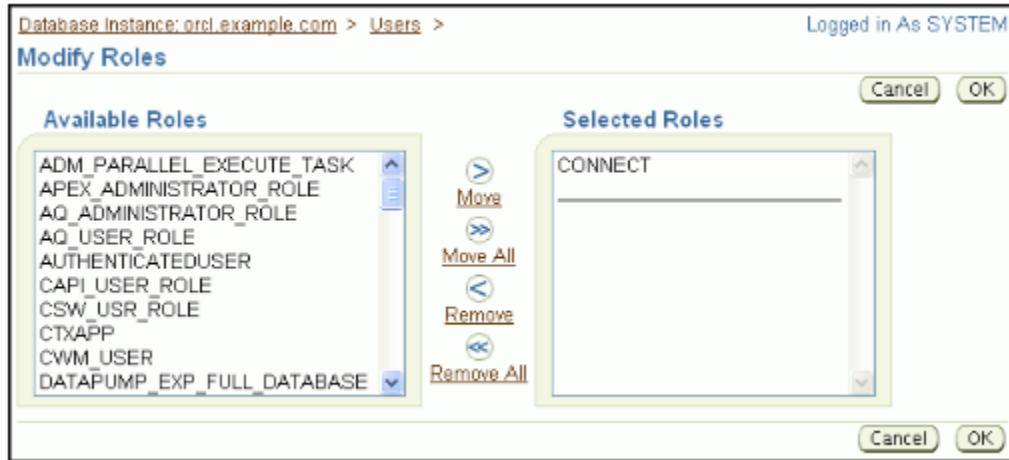
Selected System Privileges

CREATE PROCEDURE  
CREATE SEQUENCE  
CREATE SYNONYM  
CREATE TABLE  
CREATE TRIGGER  
CREATE VIEW

> Move  
Move All  
< Remove  
< Remove All

Cancel OK

Modify system privileges



To grant privileges and roles to the user Nick

Database Instance: orcl.example.com > Users >

Logged in As SYSTEM

Edit User: NICK

Actions Create Like Go Show SQL Revert Apply

General Roles System Privileges Object Privileges Quotas Consumer Group Privileges Proxy Users

Tablespace	Quota	Value	Unit
EXAMPLE	None	0	MBytes
SYSAUX	None	0	MBytes
SYSTEM	None	0	MBytes
TEMP	None	0	MBytes
UNDOTBS1	None	0	MBytes
USERS (Default)	None	0	MBytes

General Roles System Privileges Object Privileges Quotas Consumer Group Privileges Proxy Users

Actions Create Like Go Show SQL Revert Apply

## Assigning a Tablespace Quota to a User Account

Database Instance: orcl.example.com > Profiles > Logged in As SYSTEM

Edit Profile: DEFAULT

Actions Create Like Go Show SQL Revert Apply

General Password

**Password**

Expire in (days)   

Lock (days past expiration)   

**History**

Number of passwords to keep   

Number of days to keep for   

**Complexity**

Complexity function   

**Failed Login**

Number of failed login attempts to lock after   

Number of days to lock for   

## Modifying the Default Password Policy