



## **Project Assignment**

### **Phase 2: Physical Database Design and Tuning**

**By**

**Miss Pattanan Korkiattrakool      6388022**

**Mr. Dhammawat Siribunchawan 6388055**

**Miss Vipavee Ngamyingsurat      6388094**

**Submitted to**

**Asst.Prof. Dr. Charnyote Pluempitiwiriyawej**

**-A Report Submitted in Partial Fulfillment of**

**the Requirements for ITCS413\_Database Design**

**Faculty of Information and Communication Technology**

**Mahidol University**

**Table of Contents**

a. The overview of the database application project	1
Mission Statement and Objectives:	2
Scope and Boundaries:	3
Target Groups of Users:	3
b. The data requirements specifications	4
c. The transaction requirements specifications	4
d. The system requires specifications	6
e. The final Entity-Relationship diagram represents the conceptual database design	6
f. The ready-to-implement relational database schema and constraints representing the logical database design	7
g. The results of the analysis of at least two important queries or transactions for physical design	8
h. The results of the analysis of important indexes for physical design	11
i. The results of the analysis of important user views for physical design	12
j. References	13

## **a. The overview of the database application project**

### Mission Statement and Objectives:

Our mission is to create a robust and user-friendly database application that enables efficient storage, retrieval, and management of data, empowering businesses and organizations to streamline operations, enhance decision-making, and boost productivity. Our objectives include:

- **Enhanced Data Management:** Develop a reliable system for storing and organizing large volumes of data effectively.
- **Seamless Data Retrieval:** Optimize query performance to ensure quick and accurate retrieval of information.
- **User-Focused Design:** Create a user-friendly interface catering to diverse user needs and skill levels.
- **Performance Optimization:** Implement appropriate file organization techniques and indexes to enhance database performance.
- **Empowering Decision-Making:** Design user views and reports to facilitate data-driven decision-making.
- **Data Integrity and Security:** Enforce data integrity constraints and implement security measures.
- **Thorough Testing and Validation:** Conduct comprehensive testing to ensure functionality, performance, and reliability.
- **Continuous Improvement:** Explore denormalization possibilities to further optimize performance while considering trade-offs.

By accomplishing these objectives, we aim to deliver an efficient and user-centric database application that enables organizations to effectively manage their data, gain insights, and make informed decisions to drive success.

### Scope and Boundaries:

The project scope includes the physical design and tuning aspects of the database. It encompasses the selection of appropriate file organization techniques, index creation, view design, and the possibility of implementing triggers and assertions. The project does not involve the conceptual or logical design of the database or the development of the front-end application.

### Target Groups of Users:

The designed database application will cater to various target groups of users, including but not limited to:

- Data Analysts: Users who need to perform complex queries and generate reports based on the data stored in the database.
- Business Managers: Users who require access to aggregated and summarized data for decision-making and planning.
- System Administrators: Users responsible for managing the database, optimizing performance, and ensuring data integrity.
- End Users: Users who interact with the database through user views, providing a simplified interface for retrieving and updating data.

By successfully completing this project, the team aims to deliver a robust and optimized physical database design that meets the performance requirements of the target user groups, enhances data access efficiency, and ensures the integrity and reliability of the system.

**b. The data requirements specifications.**

- Customer

The data required on Customers including CustomerID(Unique identifier), CustomerName, Segment, Country, City, State, Postal Code, Region

- Product

The data required on Product including ProductID(Unique identifier), Category, SubCategory, Sales

- Payment

The data required on Payment including InvoiceID(Unique identifier), Payment Type, Discount, Profit, CustomerID, OrderID, PaymentMethod, AddressID ●

Payment method

The data required on Payment method including PaymentMethod, Name, Description

- Order

The data required on Order including OrderID(Unique identifier), OrderDate, GrossMarginPercentage, Product\_ID

- Address

The data required on Address including AddressID, Region, Postal Code, State, City, Country

**c. The transaction requirements specifications.**

- Data Entry

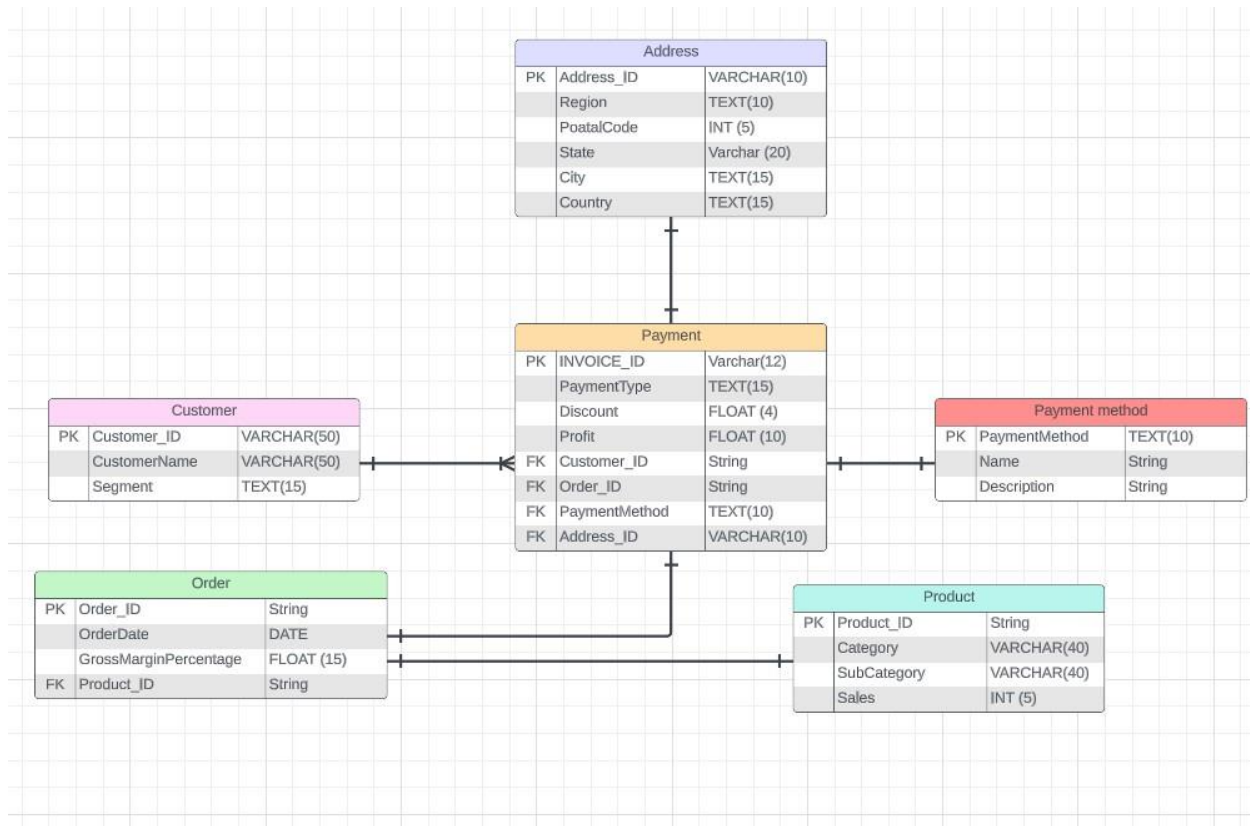
- Enter the details of the new Customers (111-22-3333, Senyai Yaiyai, ...)
- Enter the details of the new Product (111-22-3333, 74.69, 4.761904762, ...)
- Enter the details of the new Payment (111-22-3333, Ewallet, 26.1415, ...)
- Enter the details of the new Payment method (Ewallet, Senyai Yaiyai, ...)
- Enter the details of the new Order (111-22-3333, 74.69, 4.761904762, ...)
- Enter the details of the new Address ( North, 42420, Tamil Nadu, ...)

- Data Update/Delete
  - Update/Delete the details of the Customer (111-22-3333, Senyai, Yaiyai, ...)
  - Update/Delete the details of the Product (111-22-3333, 15.28, 4.761904762, ...)
  - Update/Delete the details of the Payment (111-22-3333, Credit card, 3.82, ...)
  - Update/Delete the details of the Payment method (Credit card, Senyai Yaiyai, ...)
  - Update/Delete the details of the Order (111-22-3333, 46.33, 4.761904762, ...)
  - Update/Delete the details of the new Address (South, 90036, Tamil Nadu, ...)
- Data Queries
  - Example of queries required by Branch views
    - View the Customers
    - View the Payment
    - View the Payment method
  - Example of queries required by Staff views
    - View the Product
    - View the Order
    - View the Address

**d. The system requires specifications.**

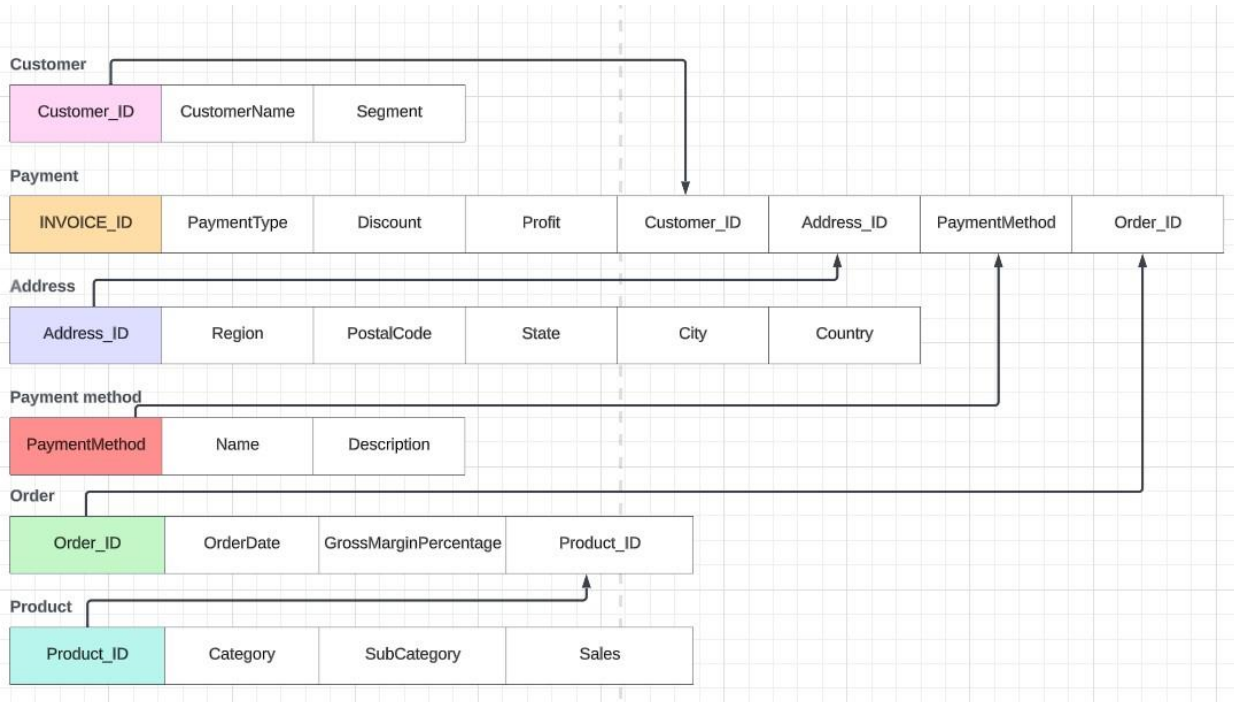
- Initial database size
  - There are approximately 793 Customer records in our database.
  - There are a total of 1861 Products in our database.
  - There are approximately 4922 Order records in our database.
  - There are approximately 9798 Invoice records in our database.
- Data rate of growth
  - We can summarize that in 5/9/2017 have the most sale in our database
  - We can summarize that Kanyakumari city have the most sale in India for our database
  - We can summarize that the product ID OFF-PA-10001970 have the most sale in our database

**e. The final Entity-Relationship diagram represents the conceptual database design**



**f. The ready-to-implement relational database schema and constraints representing the logical database design.**

- Relational schema





**g. The results of the analysis of at least two important queries or transactions for physical design**

(A) Insert values into the table

(B) List the total sales and products of products indicated in the specific categories.

(C) List the total sales product with total profit from 2017 - 2018

(D) Identify payment type

(E) Identify all items in the table

Transaction/ Relation	(A)				(B)				(C)				(D)				(E)			
	I	R	U	D	I	R	U	D	I	R	U	D	I	R	U	D	I	R	U	D
Customer	X																	X		
Product	X					X				X								X		
Payment	X												X					X		
Order	X					X				X								X		

I = Insert, R = Read, U = Update, D = Delete

### Transaction Analysis Form

**Transaction** List the total sales and products of products indicated in the specific categories.

**Transaction volume** Average: 381

Peak: 2,499

```
SELECT Category, AVG(Profit) AS Average_Profit, SUM(Sales) AS Total_Sales,
MAX(Sales) AS Peak_Sale
FROM Geocery
GROUP BY Category
ORDER BY Total_Sales DESC
LIMIT 1;
```

	Category	Average_Profit	Total_Sales	Peak_Sale
►	Eggs, Meat and Fish	381.834726727074	2229779	2499

Access	Entity	Type of Access	No. of References		
			Per Transaction	Avg. Per Transaction	Peak Per Transaction
1	OrderDate	R	2,229,779	381	2,499
Total References			2,229,779	381	2,499

## Transaction Analysis Form

**Transaction** List the total sales product with total profit from 2017 - 2018

**Transaction volume** Average: 387

Peak: 2,499 sale items

```
SELECT Category, AVG(Profit) AS Average_Profit, SUM(Sales) AS Total_Sales,
MAX(Sales) AS Peak_Sale
FROM Geocery
WHERE YEAR(OrderDate) BETWEEN 2017 AND 2018
GROUP BY Category
ORDER BY Total_Sales DESC
LIMIT 1;
```

	Category	Average_Profit	Total_Sales	Peak_Sale
►	Eggs, Meat and Fish	387.78844488059485	1342861	2499

Access	Entity	Type of Access	No. of References		
			Per Transaction	Avg. Per Transaction	Peak Per Transaction
1	OrderDate	R	1,342,861	387	2,499
Total References			1,342,861	387	2,499

**h. The results of the analysis of important indexes for physical design**

Table	Indexes
Customer	CustomerID
Product	ProductID
Payment	InvoiceID
Order	OrderID

## i. The results of the analysis of important user views for physical design

We create two of user view to analysis in business part the first user view is for the order that have OrderID, OrderDate, CustomerName, Category, and Sale for analysis only the part of payment process

	OrderID	OrderDate	CustomerName	Category	Sales
▶	CA-2017-152156	2017-11-08	Claire Gute	Oil and Masala	1254
	CA-2017-152156	2017-11-08	Claire Gute	Beverages	749
	CA-2017-138688	2017-06-12	Darrin Van Huff	Food Grains	2360
	US-2016-108966	2016-10-11	Sean O'Donnell	Fruits and Veggies	896
	US-2016-108966	2016-10-11	Sean O'Donnell	Food Grains	2355
	CA-2015-115812	2015-06-09	Brosina Hoffman	Food Grains	2305
	CA-2015-115812	2015-06-09	Brosina Hoffman	Fruits and Veggies	826
	CA-2015-115812	2015-06-09	Brosina Hoffman	Fruits and Veggies	1847
	CA-2015-115812	2015-06-09	Brosina Hoffman	Bakery	791
	CA-2015-115812	2015-06-09	Brosina Hoffman	Bakery	1795
	CA-2015-115812	2015-06-09	Brosina Hoffman	Snacks	1903
	CA-2015-115812	2015-06-09	Brosina Hoffman	Eggs, Meat and Fish	701
	CA-2018-114412	2018-04-15	Andrew Allen	Snacks	1659
	CA-2017-161389	2017-12-05	Irene Maddox	Fruits and Veggies	1277
	US-2016-118983	2016-11-22	Harold Pawlan	Eggs, Meat and Fish	831
	US-2016-118983	2016-11-22	Harold Pawlan	Oil and Masala	1440
	CA-2015-105893	2015-11-11	Pete Kriz	Bakery	1678
	CA-2015-167164	2015-05-13	Alejandro Grove	Beverages	1617
	CA-2015-143336	2015-08-27	Zuschuss Donatelli	Eggs, Meat and Fish	1757
	CA-2015-143336	2015-08-27	Zuschuss Donatelli	Beverages	692
	CA-2015-143336	2015-08-27	Zuschuss Donatelli	Beverages	522
	CA-2017-137330	2017-12-09	Ken Black	Food Grains	948

*Example of output for user view 1*

The second user view is for the order that have the Sales greater than 1,000 and have Profit greater than 500 to see what kind of product that have the most sales in our database

OrderID	OrderDate	GrossMarginPercentage	CustomerID	CustomerName	Segment	Country	City	State	PostalCode	Region	ProductID	Category	SubCategory	Sales	Discount	Profit	InvoiceID	PaymentType
▶ CA-2017-114601	2017-08-26	4.7619	AA-10480	Andrew Allen	Consumer	India	Ooty	Tamil Nadu	48234	West	TEC-PH-10002170	Bakery	Breads and Buns	1786	0.17	500.08	IN6-967-0308	Credit card
CA-2017-165169	2017-06-05	4.7619	JA-15215	Janet Lee	Consumer	India	Virudhunagar	Tamil Nadu	19711	East	OFF-ST-10001496	Oil and Masala	Spices	1220	0.29	500.2	IN5-073-2894	Cash
US-2017-111290	2017-07-22	4.7619	DK-13375	Dennis Kane	Consumer	India	Vellore	Tamil Nadu	48185	West	OFF-PA-10002262	Eggs, Meat and Fish	Mutton	1220	0.23	500.2	IN0-146-8125	Ewallet
CA-2018-152912	2018-11-09	4.7619	BM-11650	Brian Moss	Consumer	India	Nagercoil	Tamil Nadu	21044	Central	OFF-ST-10003208	Eggs, Meat and Fish	Fish	1137	0.18	500.28	IN6-116-5782	Credit card
US-2016-122784	2016-07-20	4.7619	RA-19915	Russell Applegate	Consumer	India	Coimbatore	Tamil Nadu	60035	West	TEC-PH-10001557	Fruits and Veggies	Fresh Fruits	1137	0.32	500.28	IN4-298-9793	Cash
CA-2018-147767	2018-12-02	4.7619	SV-20935	Susan Vittorini	Consumer	India	Trichy	Tamil Nadu	85345	South	OFF-BE-10001670	Snacks	Cookies	1390	0.32	500.4	IN9-120-3511	Cash
CA-2016-111199	2016-11-13	4.7619	JK-15730	Joe Kamberova	Consumer	India	Coimbatore	Tamil Nadu	28540	Central	OFF-PA-10001790	Snacks	Chocolates	2176	0.19	500.48	IN8-719-8675	Cash
US-2016-128090	2016-08-16	4.7619	JM-15985	John Murray	Consumer	India	Vellore	Tamil Nadu	85023	West	OFF-AR-10002255	Food Grains	Rice	1925	0.27	500.5	IN0-677-1560	Ewallet
US-2017-161844	2017-12-09	4.7619	DK-12835	Daniela Kotsonis	Consumer	India	Coimbatore	Tamil Nadu	37064	Central	OFF-BE-10003712	Beverages	Soft Drinks	1726	0.2	500.54	IN9-421-0807	Cash
CA-2015-145317	2015-03-18	4.7619	SM-20330	Sean Miller	Consumer	India	Cumbum	Tamil Nadu	32216	Central	TEC-MA-10003626	Oil and Masala	Masalas	1615	0.29	500.65	IN7-255-6137	Cash
CA-2018-121300	2018-09-29	4.7619	MG-17680	Maureen Gastreanu	Consumer	India	Virudhunagar	Tamil Nadu	44060	West	OFF-PA-10002787	Eggs, Meat and Fish	Fish	1669	0.16	500.7	IN7-278-4718	Ewallet
US-2015-115189	2015-12-30	4.7619	AR-10345	Alex Russell	Consumer	India	Vikupuram	Tamil Nadu	19143	East	TEC-PH-10002170	Fruits and Veggies	Fresh Fruits	1391	0.14	500.76	IN8-793-5847	Cash
CA-2015-138527	2015-09-12	4.7619	BN-11470	Brad Norvell	Consumer	India	Chennai	Tamil Nadu	27511	South	OFF-AP-10001469	Eggs, Meat and Fish	Chicken	1354	0.35	500.98	IN0-112-5577	Cash
CA-2017-134936	2017-12-19	4.7619	ES-14080	Erin Smith	Consumer	India	Cumbum	Tamil Nadu	85705	West	OFF-BE-10000174	Oil and Masala	Spices	1927	0.11	501.02	IN1-899-2557	Cash
CA-2016-155761	2016-12-11	4.7619	SC-20800	Stuart Calhoun	Consumer	India	Hamakal	Tamil Nadu	77041	Central	OFF-ST-10000943	Snacks	Noodles	1728	0.31	501.12	IN2-286-0573	Ewallet
US-2015-105137	2015-10-10	4.7619	RB-19435	Richard Berner	Consumer	India	Tirunelveli	Tamil Nadu	43229	South	OFF-BE-10002439	Fruits and Veggies	Organic Fruits	1966	0.24	501.12	IN7-082-6053	Cash
CA-2018-120019	2018-02-16	4.7619	NG-18355	Nat Gilpin	Consumer	India	Ooty	Tamil Nadu	4824	East	OFF-ST-10001932	Eggs, Meat and Fish	Chicken	2179	0.3	501.17	IN6-895-2553	Cash
CA-2017-109563	2017-07-14	4.7619	RB-19360	Raymond Buch	Consumer	India	Ooty	Tamil Nadu	94122	East	FUR-FU-10000073	Snacks	Cookies	1617	0.28	501.27	IN1-402-4928	Ewallet
CA-2016-150714	2016-10-26	4.7619	KH-16690	Kristen Hastings	Consumer	India	Bodi	Tamil Nadu	97477	Central	OFF-ST-10001809	Food Grains	Atta and Flour	1617	0.25	501.27	IN1-974-8567	Cash
CA-2017-144344	2017-10-28	4.7619	PG-18820	Patrick Gardner	Consumer	India	Pudukottai	Tamil Nadu	33437	Central	FUR-FU-10003553	Snacks	Cookies	1857	0.27	501.39	IN7-536-0742	Cash
CA-2018-138289	2018-01-16	4.7619	AR-10540	Andy Reiter	Consumer	India	Coimbatore	Tamil Nadu	49201	South	OFF-BE-10004995	Bakery	Cakes	1791	0.28	501.48	IN6-332-8517	Ewallet
CA-2018-102197	2018-12-20	4.7619	DK-13150	David Kendrick	Consumer	India	Chennai	Tamil Nadu	10009	East	OFF-PA-10001970	Oil and Masala	Spices	1286	0.12	501.54	IN4-410-8226	Credit card

*Example of output for user view 2*

## j. References

- [1] Mohamed Harris. Supermart Grocery Sales Retail Analytics Dataset. Available online: <https://www.kaggle.com/datasets/mohamedharris/supermart-grocery-sales-retail-analytics-dataset> (accessed on 11 May 2023).
- [2] Rohit Sahoo (2020). Sales Forecasting Dataset. Available online: <https://www.kaggle.com/datasets/rohithsahoo/sales-forecasting> (accessed on 11 May 2023).
- [3] Aung Pyae (2019). Supermarket Sales Dataset. Available online: <https://www.kaggle.com/datasets/aungpyacap/supermarket-sales> (accessed on 12 May 2023).