

zenius

Kampus
Merdeka
INDONESIA JAYA

Database: SQL Query I

Tuesday, 9 May 2023

Data Analytics

Program Zenius Studi Independen Bersertifikat
Zenius Bersama Kampus Merdeka



Quick Intro

Theo Jeremiah

Roles:

- CURRENTLY | Data Scientist at AirAsia
- 20 - 23 | Data Scientist at Allianz Indonesia
- 19 - 20 | Business Development at Mineski Indonesia
- 18 - 19 | Data Analyst at Excite Indonesia



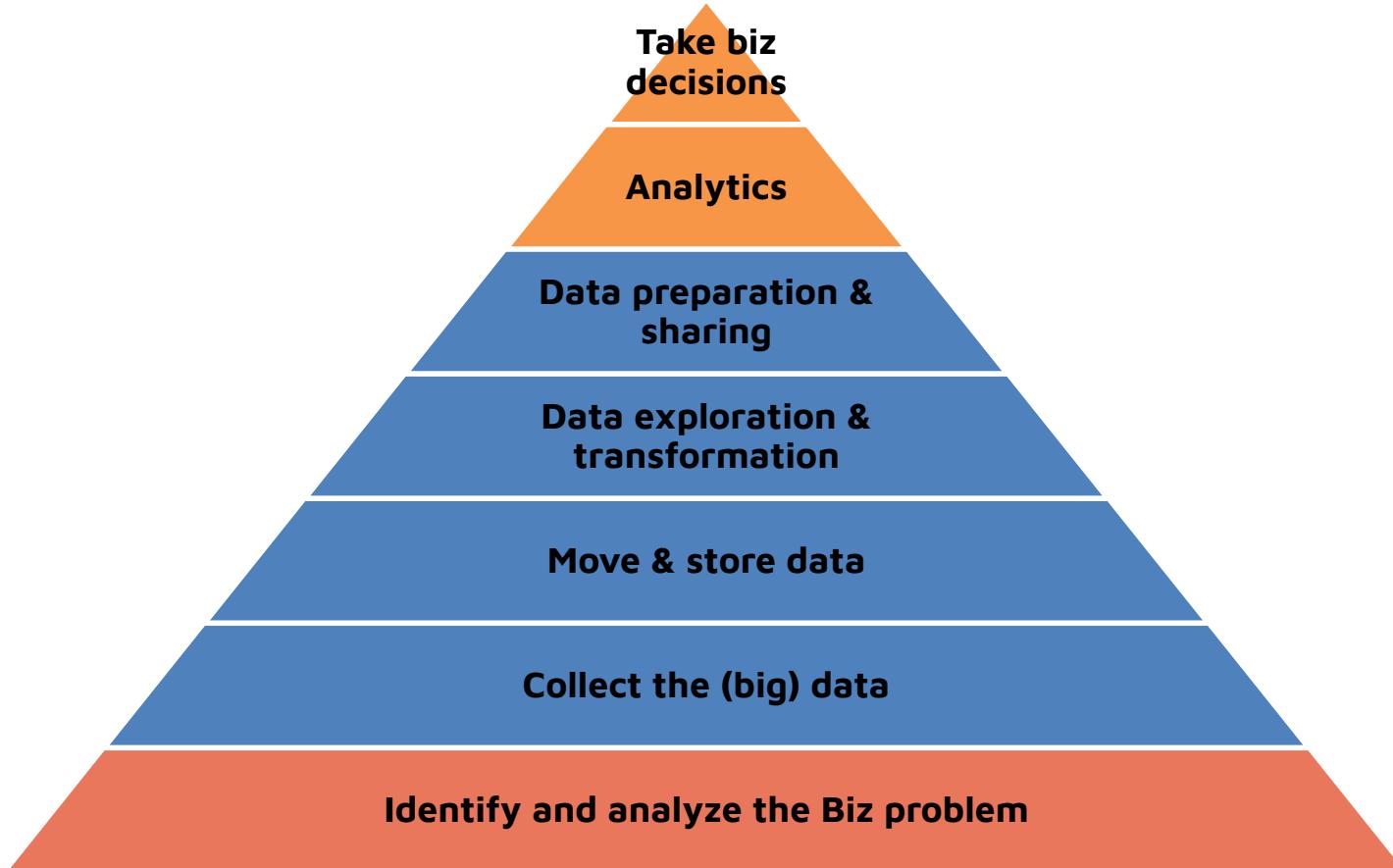
www.linkedin.com/in/theojeremiah/

- 1. OLTP vs OLAP**
- 2. RDBMS and SQL Basics**
- 3. Introduction to PostgreSQL**
- 4. SQL Commands and Operators**

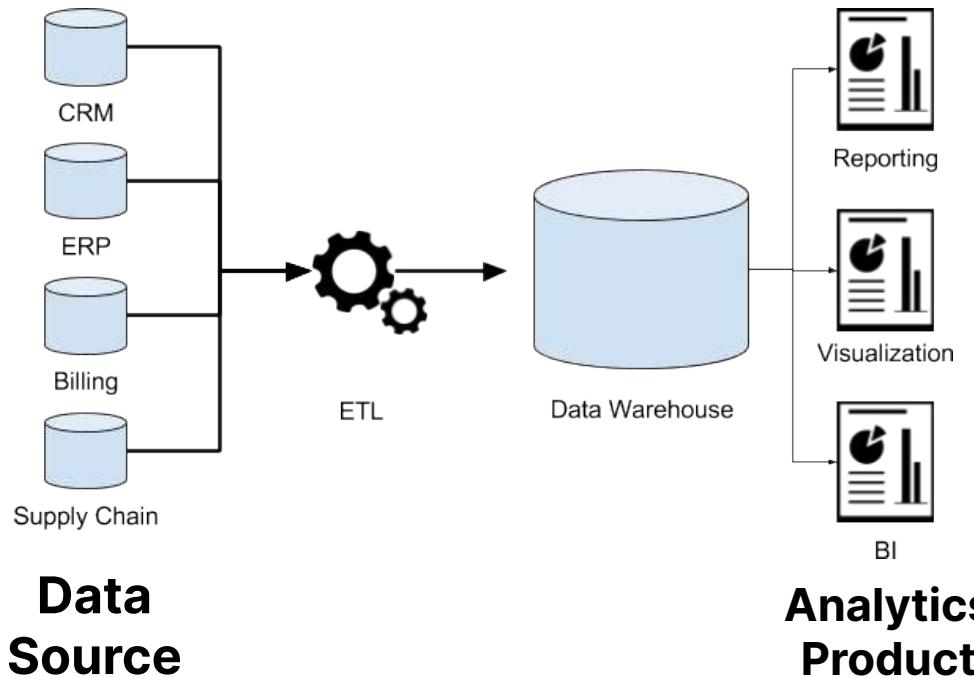


OLTP vs OLAP

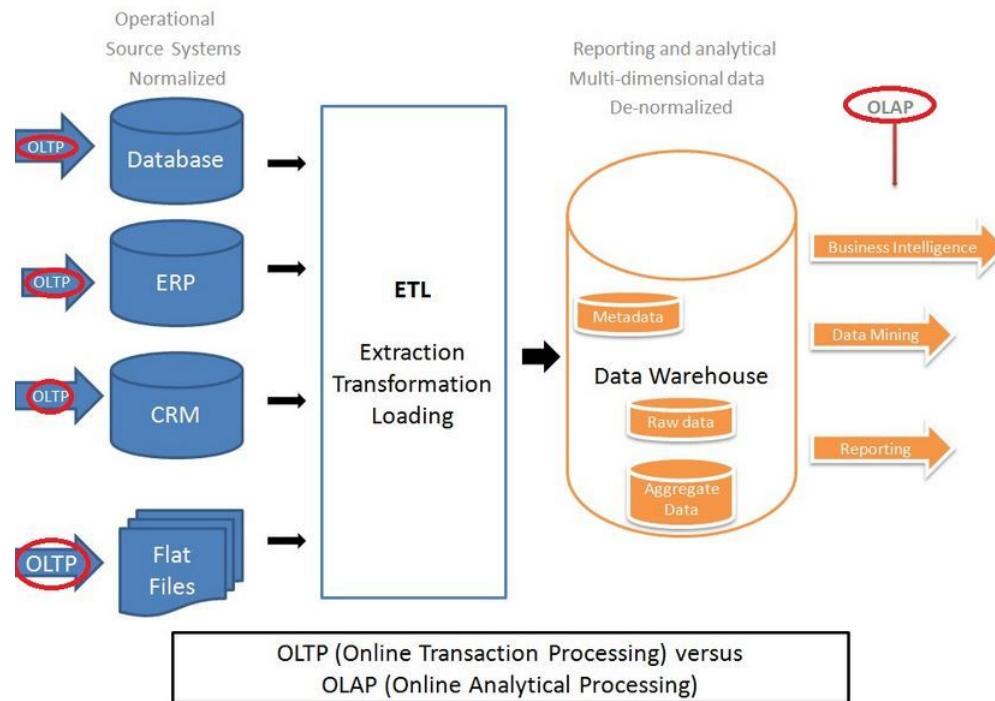




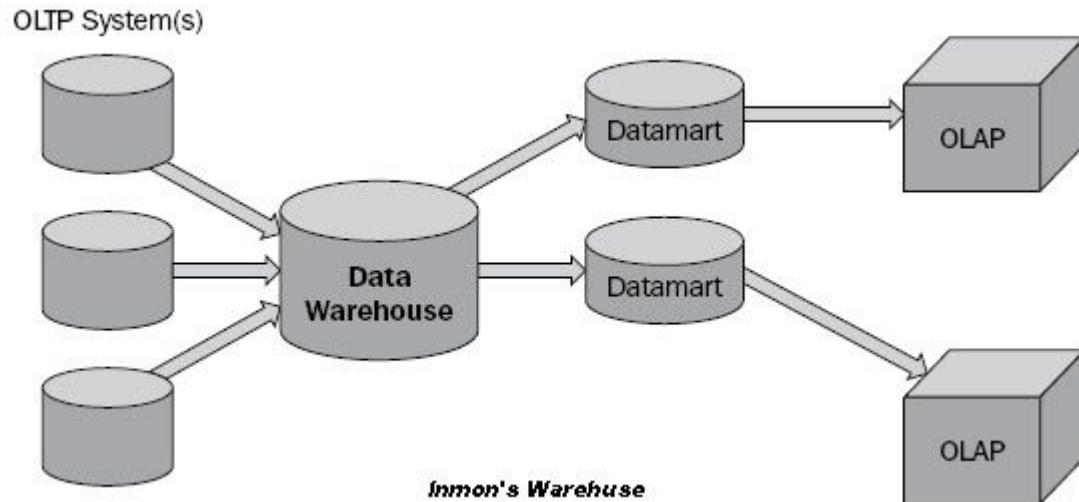
Common Data Architecture



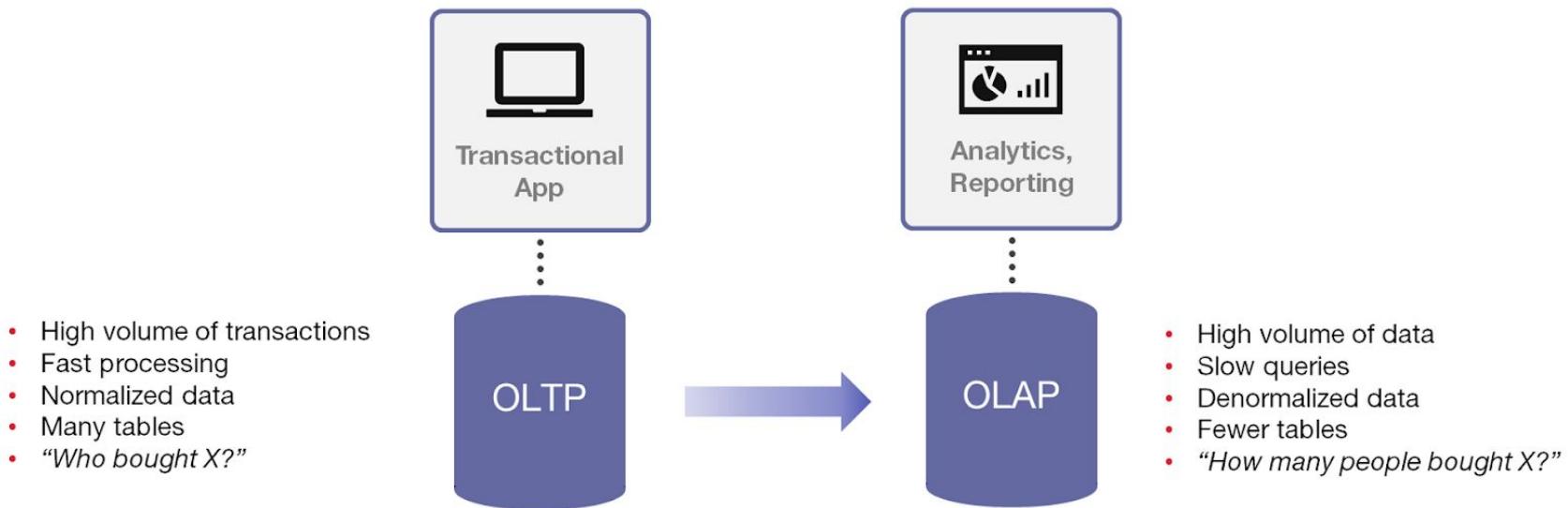
Common Data Architecture



OLTP vs OLAP

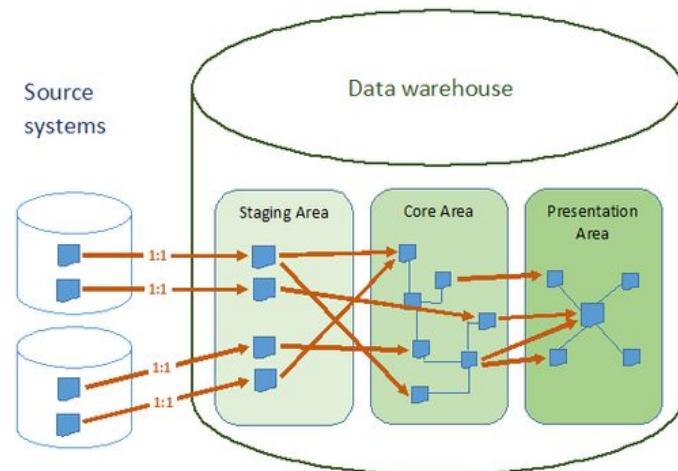


OLTP vs OLAP



Journey to OLAP

processes that happened **before OLAP** is called **ETL (Extract, Transform, Load)** and it happened on the **Data Warehouse System**, forming **Data Mart(s)**



RDBMS



What is RDBMS

RDBMS stands **for Relational Database Management System**

Database : a collection of data that is organized systematically so that it can be easily accessed managed and updated

DBMS : Software that is designed to create and manage databases and perform operations on database

RDBMS : DBMS that stores data on tabular form and the relation between tables is known

DBMS Examples

RDBMS



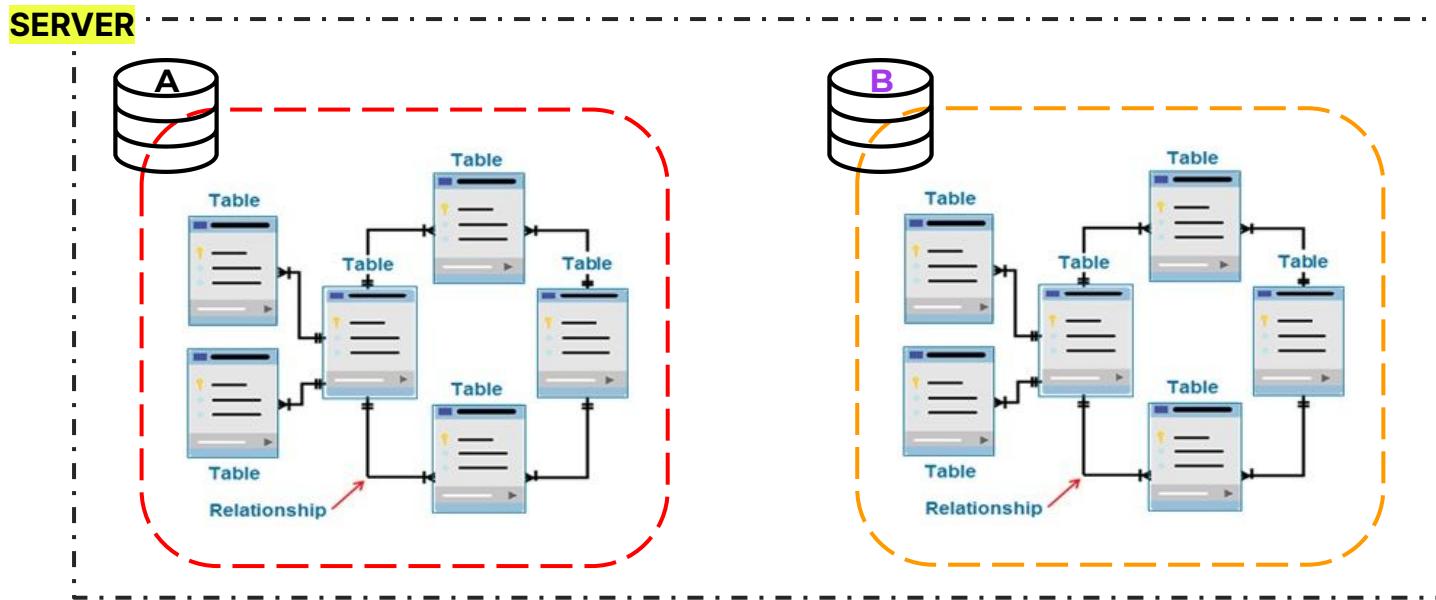
Non-relational DBMS



RDBMS Scheme

In a **server**, there could be several **databases**

In a **databases**, there could be several **tables**



Terminologies in RDBMS

Table : Data that saved with a structure (consists of rows and columns)

Record / Tuple / Row : One entry of data in a table

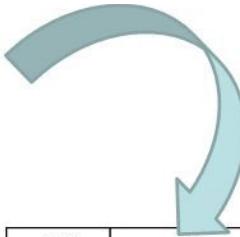
Field / Attribute / Column : Specific Information that's had by a record in a table

Query : A request to retrieve some specific data in a database

SQL : Structured Query Language

RDBMS

- A **relational database** is data represented as a set of tables.
- The tables can be viewed like a spreadsheet
 - **Records** (rows) — data for one item or transaction
 - **Fields** (columns) — represent different elements of data
 - **Key field** is a field that uniquely identifies the record



ISBN	Title	Author	Publisher
0-111-11111-1	89 Years in a Sand Trap	Beck, Fred	Hill and Wang
0-15-500139-6	Business Programming in C	Millspaugh, A. C.	The Dryden Press
0-394-75843-9	Cultural Literacy	Hirsch, E. D. Jr.	Vintage
0-440-22284-2	Five Days in Paris	Steel, Danielle	Dell Publishing
0-446-51251-6	Megatrends	Naisbitt, John	Warner Books
0-446-51652-X	Bridges of Madison County	Waller, Robert James	Warner Books
0-446-60274-4	The Rules	Fein/Schneider	Warner Books
0-451-16095-9	The Stand	King, Stephen	Signet
0-452-26011-6	Song of Solomon	Morrison, Toni	Plume/Penguin
0-517-59905-8	How to Talk to Anyone, Anytime, Anywhere	King, Larry	Crown
0-534-26076-4	A Quick Guide to the Internet	Bradley, Julia Case	Integrated Media Group

RDBMS

SALES				
purchase_number	date_of_purchase	customer_id	item_code	
1	03/09/2016	1	A_1	
2	02/12/2016	2	C_1	
3	15/04/2017	3	D_1	
4	24/05/2017	1	B_2	
5	25/05/2017	4	B_2	
6	06/06/2017	2	B_1	
7	10/06/2017	4	A_2	
8	13/06/2017	3	C_1	
9	20/07/2017	1	A_1	
10	11/08/2017	2	B_1	

RDBMS

Customers					
customer_id	first_name	last_name	email_address	number_of_complaints	
1	John	McKinley	john.mackinley@365careers.com	0	
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2	
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1	
4	Catherine	Winnfield	c.winnfield@365careers.com	0	

Customers					
customer_id	first_name	last_name		email_address	number_of_complaints
1	John	McKinley		john.mackinley@365careers.com	0
2	Elizabeth	McFarlane		e.mcfarlane@365careers.com	2
3	Kevin	Lawrence		kevin.lawrence@365careers.com	1
4	Catherine	Winnfield		c.winnfield@365careers.com	0

SALES			
purchase_number	date_of_purchase	customer_id	item_code
1	03/09/2016	1	A_1
2	02/12/2016	2	C_1
3	15/04/2017	3	D_1
4	24/05/2017	1	B_2
5	25/05/2017	4	B_2
6	06/06/2017	2	B_1
7	10/06/2017	4	A_2
8	13/06/2017	3	C_1
9	20/07/2017	1	A_1
10	11/08/2017	2	B_1

Items						
item_code	item	unit_price_usd	company_id	company	headquarters_phone_number	
A_1	Lamp	20	1	Company A	+1 (202) 555-0196	
A_2	Desk	250	1	Company A	+1 (202) 555-0196	
B_1	Lamp	30	2	Company B	+1 (202) 555-0152	
B_2	Desk	350	2	Company B	+1 (202) 555-0152	
C_1	Chair	150	3	Company C	+1 (229) 853-9913	
D_1	Loudspeakers	400	4	Company D	+1 (618) 369-7392	

Customer_Sales												
purchase_number	date_of_purchase	customer_id	first_name	last_name	email_address	number_of_complaints	item_id	item	unit_price_usd	company	headquarters	phone_number
1	03/09/2016	cust_1	John	McKinley	john.mckinley@365careers.com	0	A_1	Lamp	20	Company A	+1 (202) 555-0196	
2	02/12/2016	cust_2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2	C_1	Chair	150	Company C	+1 (229) 853-9913	
3	15/04/2017	cust_3	Kevin	Lawrence	kevin.lawrence@365careers.com	1	D_1	Loudspeakers	400	Company D	+1 (618) 369-7392	
4	24/05/2017	cust_1	John	McKinley	john.mckinley@365careers.com	0	B_2	Desk	350	Company B	+1 (202) 555-0152	
5	25/05/2017	cust_4	Catherine	Winnfield	c.winnfield@365careers.com	0	B_2	Desk	350	Company B	+1 (202) 555-0152	
6	06/06/2017	cust_2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2	B_1	Lamp	30	Company B	+1 (202) 555-0152	
7	10/06/2017	cust_4	Catherine	Winnfield	c.winnfield@365careers.com	0	A_2	Desk	250	Company A	+1 (202) 555-0196	
8	13/06/2017	cust_3	Kevin	Lawrence	kevin.lawrence@365careers.com	1	C_1	Chair	150	Company C	+1 (229) 853-9913	
9	20/07/2017	cust_1	John	McKinley	john.mckinley@365careers.com	0	A_1	Lamp	20	Company A	+1 (202) 555-0196	
10	11/08/2017	cust_2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2	B_1	Lamp	30	Company B	+1 (202) 555-0152	

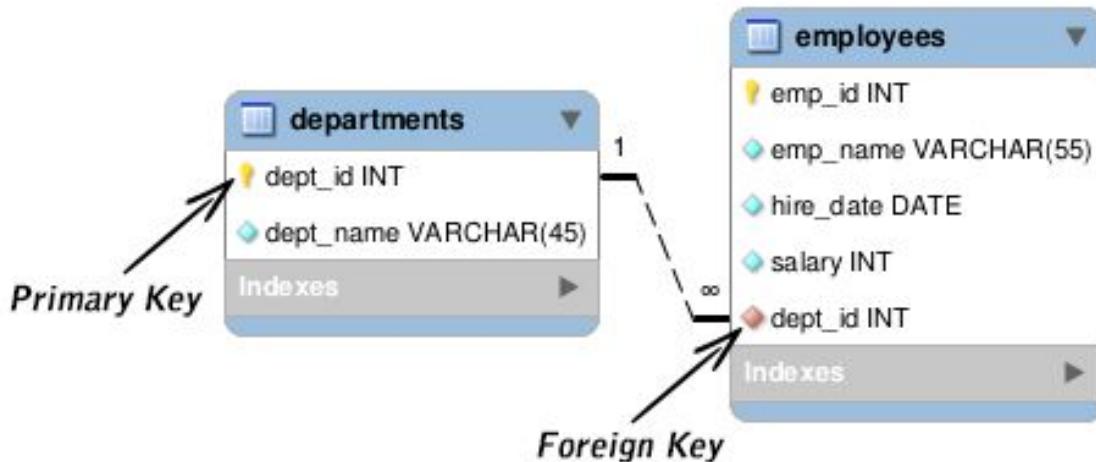
Customers					
customer_id	first_name	last_name		email_address	number_of_complaints
1	John	McKinley		john.mackinley@365careers.com	0
2	Elizabeth	McFarlane		e.mcfarlane@365careers.com	2
3	Kevin	Lawrence		kevin.lawrence@365careers.com	1
4	Catherine	Winnfield		c.winnfield@365careers.com	0

SALES			
purchase_number	date_of_purchase	customer_id	item_code
1	03/09/2016	1	A_1
2	02/12/2016	2	C_1
3	15/04/2017	3	D_1
4	24/05/2017	1	B_2
5	25/05/2017	4	B_2
6	06/06/2017	2	B_1
7	10/06/2017	4	A_2
8	13/06/2017	3	C_1
9	20/07/2017	1	A_1
10	11/08/2017	2	B_1

Items						
item_code	item	unit_price_usd	company_id	company	headquarters_phone_number	
A_1	Lamp	20	1	Company A	+1 (202) 555-0196	
A_2	Desk	250	1	Company A	+1 (202) 555-0196	
B_1	Lamp	30	2	Company B	+1 (202) 555-0152	
B_2	Desk	350	2	Company B	+1 (202) 555-0152	
C_1	Chair	150	3	Company C	+1 (229) 853-9913	
D_1	Loudspeakers	400	4	Company D	+1 (618) 369-7392	

Relational Algebra allow us to retrieve data efficiently

Primary Key vs Foreign Key



Primary Key

Primary Key

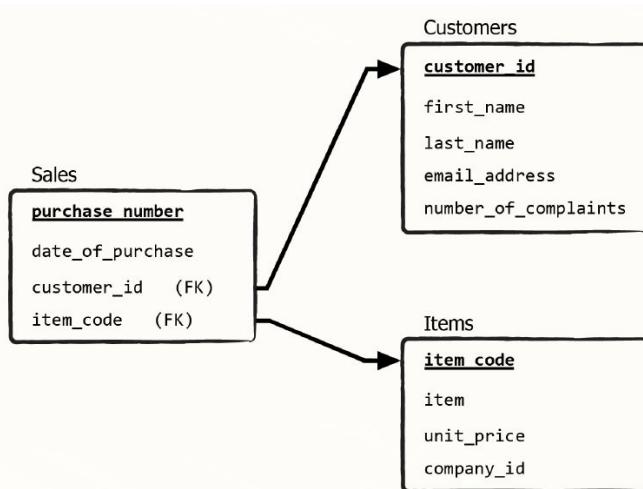
A column (or a set of columns) whose value exists and is unique for every record in a table is called a primary key

- Each table can have one and only one primary key
- In one table, you cannot have 3 or 4 primary keys
- Primary keys are the unique identifiers of a table
- Cannot contain null values!

Foreign Key

Foreign Key

Identifies the relationships between tables, not the tables themselves



Relationships

Relationships

Relationships tell you how much of the data from a **foreign key** field can be seen in the **primary key** column of the table the data is related to and vice versa

Relationships

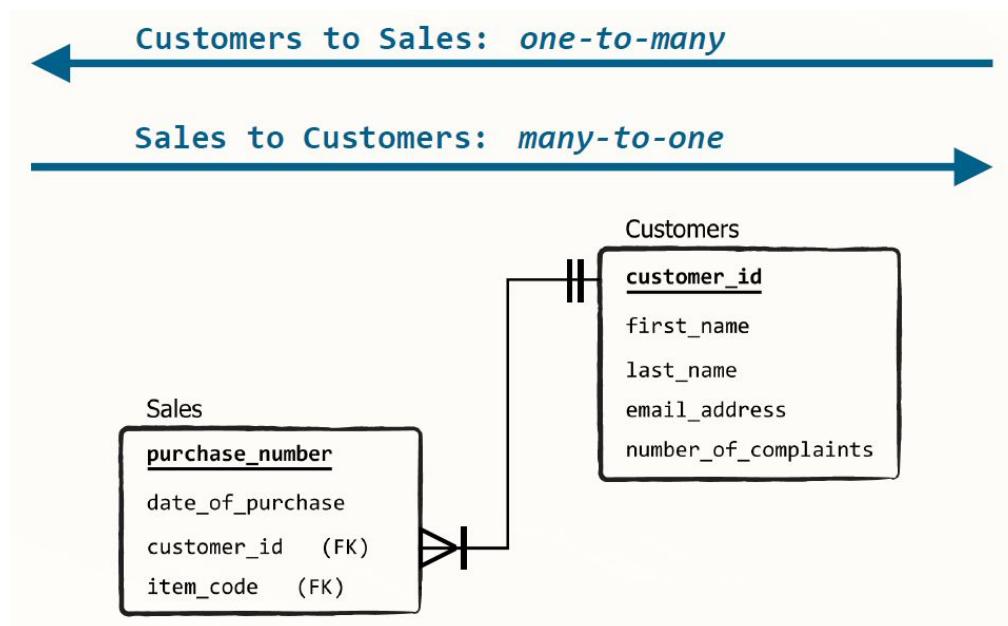
One-to-many type of relationship

One value from the customer_id columns the “Customers” table can be found **many** times in the customer_id column in the “Sales” table

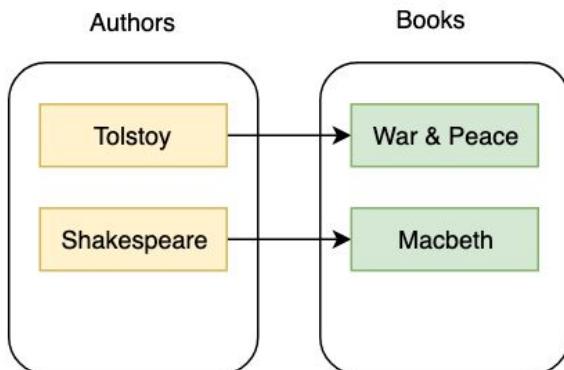
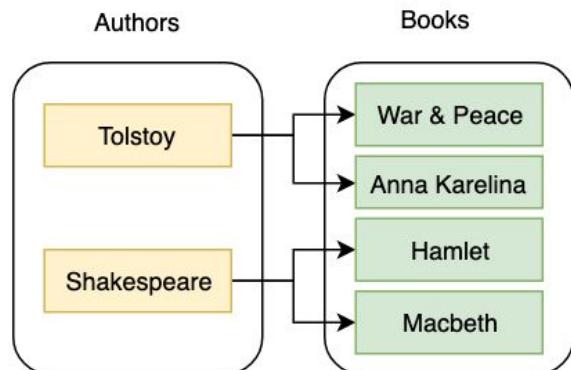
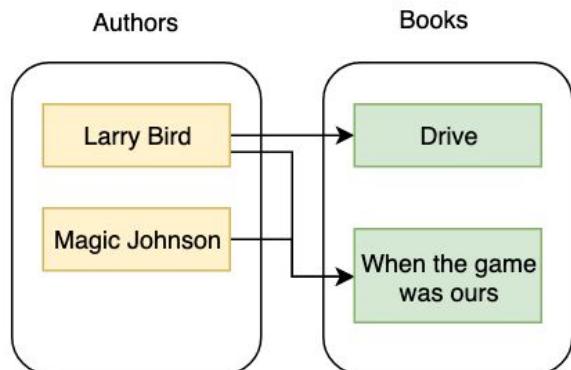
Sales			
purchase_number	date_of_purchase	customer_id	item_code
1	9/3/2016	1	A_1
2	12/2/2016	2	C_1
3	4/15/2017	3	D_1
4	5/24/2017	1	B_2
5	5/25/2017	4	B_2
6	6/6/2017	2	B_1
7	6/10/2017	4	A_2
8	6/13/2017	3	C_1
9	7/20/2017	1	A_1
10	8/11/2017	2	B_1

Customers					
customer_id	first_name	last_name	email_address	number_of_complaints	
1	John	McKinley	john.mackinley@365careers.com	0	
2	Elizabeth	McFarlane	e.mcfarlane@365careers.com	2	
3	Kevin	Lawrence	kevin.lawrence@365careers.com	1	
4	Catherine	Winnfield	c.winnfield@365careers.com	0	

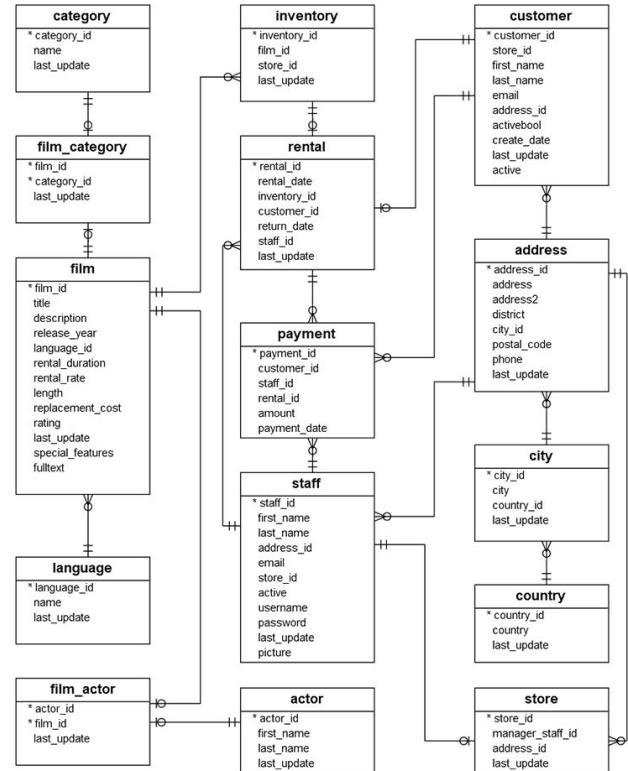
Relationships



Relationships

ONE TO ONE**ONE TO MANY****MANY TO MANY**

Entity Relationship Diagram



SQL Basics



Introduction to SQL

SQL = Structured Query Language

A programming language specifically designed for working with databases

- Create
- Manipulate **DATA** from relational database management systems
- Share

Introduction to SQL

Types of programming:

- Procedural (imperative) -> JAVA
- Declarative (non-procedural) -> SQL

Introduction to SQL

Procedural (imperative) **JAVA**
HOW

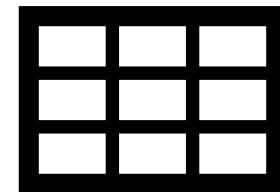
1. Please open the door
2. Go outside
3. Take the bucket I forgot there
4. Bring it back to me

Declarative (non-procedural) **SQL**
WHAT

1. Fetch the bucket, please.

Introduction to SQL

Allows you to write queries that the computer can execute and then provide database insights in return



Query
(a piece of code)

Output

Introduction to SQL

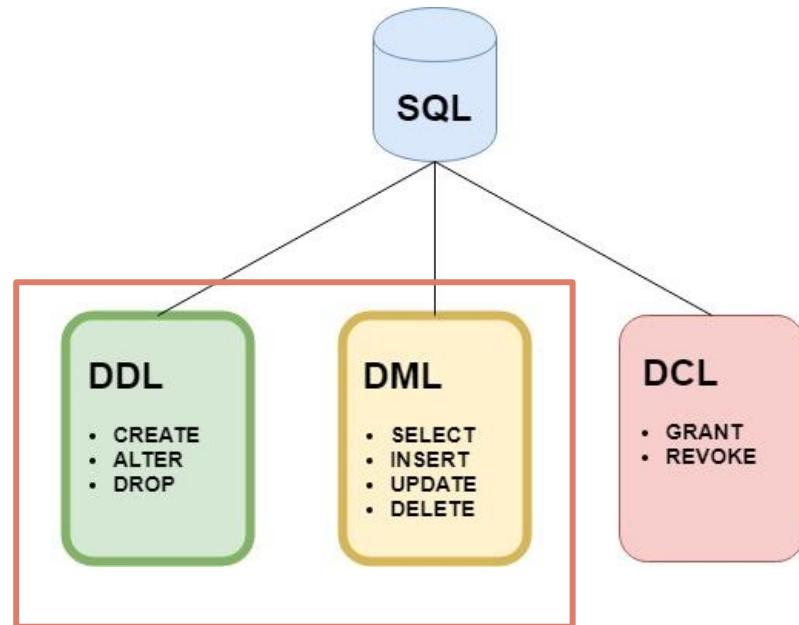
SQL = Structured Query Language

- Intuitive
- Easy-to-learn
- For business problems involving the processing of large amounts of data
- Powerful

SQL Components

SQL consists of **three components** which offer everything required to **manage, maintain** and **use** a database

- Data Definition Language (DDL)
- Data Manipulation Language (DML)
- Data Control Language (DCL)



SQL Environment



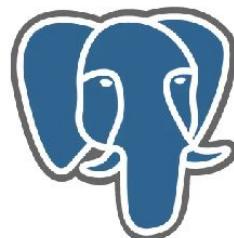
MariaDB®



ORACLE

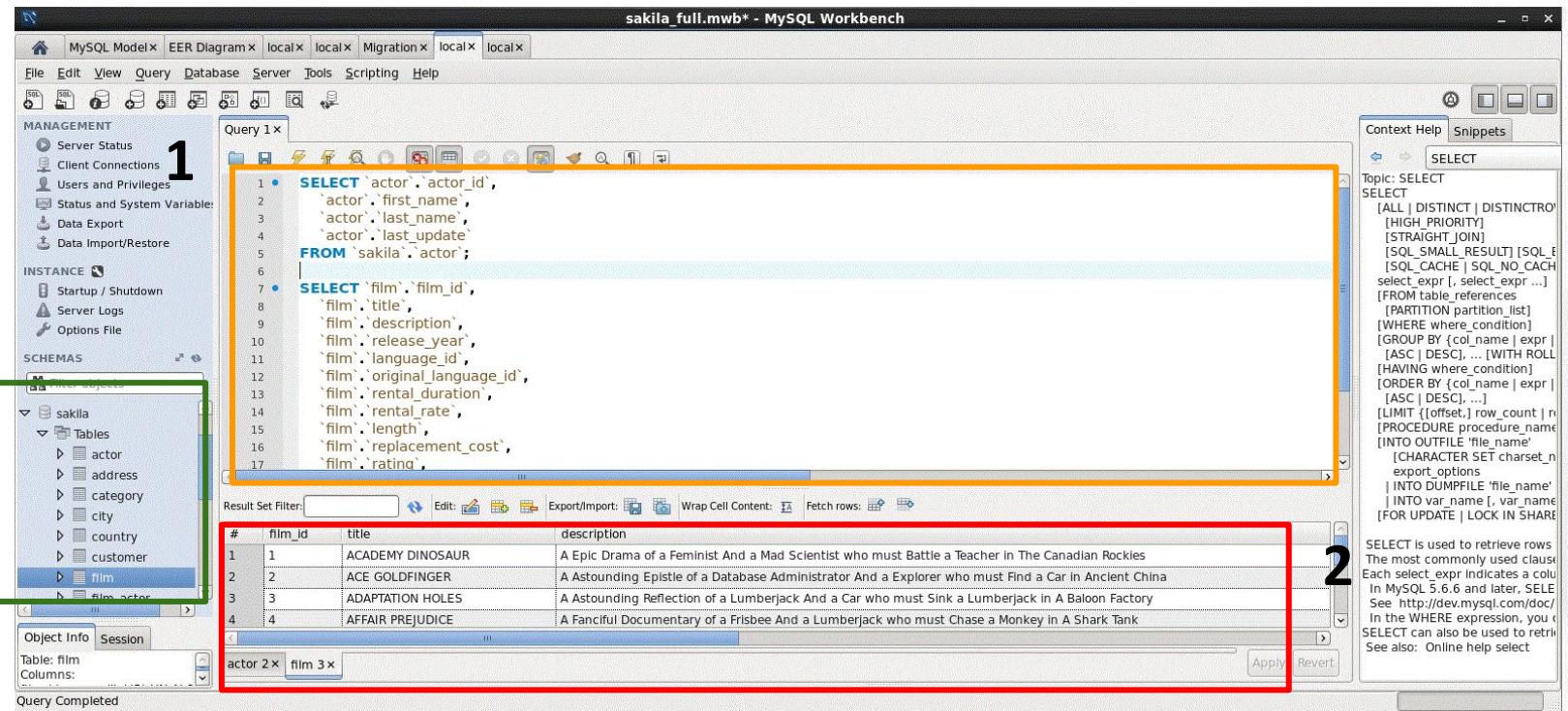


MySQL®



SQL Environment

1



```

1 • SELECT `actor`.`actor_id`,
2   `actor`.`first_name`,
3   `actor`.`last_name`,
4   `actor`.`last_update`
5   FROM `sakila`.`actor`;
6
7 • SELECT `film`.`film_id`,
8   `film`.`title`,
9   `film`.`description`,
10  `film`.`release_year`,
11  `film`.`language_id`,
12  `film`.`original_language_id`,
13  `film`.`rental_duration`,
14  `film`.`rental_rate`,
15  `film`.`length`,
16  `film`.`replacement_cost`,
17  `film`.`rating`;
    
```

2

Result Set Filter: Edit: Export/Import: Wrap Cell Content: Fetch rows:

#	film_id	title	description
1	1	ACADEMY DINOSAUR	A Epic Drama of a Feminist And a Mad Scientist who must Battle a Teacher in The Canadian Rockies
2	2	ACE GOLDFINGER	A Astounding Epistle of a Database Administrator And a Explorer who must Find a Car In Ancient China
3	3	ADAPTATION HOLES	A Astounding Reflection of a Lumberjack And a Car who must Sink a Lumberjack In A Baloon Factory
4	4	AFFAIR PREJUDICE	A Fanciful Documentary of a Frisbee And a Lumberjack who must Chase a Monkey In A Shark Tank

3

Object Info Session

Table: film Columns:

Query Completed

Context Help Snippets

Topic: SELECT

SELECT

[ALL | DISTINCT | DISTINCTROW |
[HIGH_PRIORITY] |
[STRAIGHT_JOIN] |
[SQL_SMALL_RESULT] | [SQL_BIG_RESULT] |
[SQL_CACHE] | SQL_NO_CACHE]
select_expr [, select_expr ...]
[FROM table_references
[PARTITION partition_list]]
[WHERE where_condition]
[GROUP BY {col_name | expr |
[ASC | DESC]} ... [WITH ROLLUP]]
[HAVING where_condition]
[ORDER BY {col_name | expr |
[ASC | DESC]} ...]
[LIMIT {offset[, row_count]} [FOR UPDATE | LOCK IN SHARE_MODE_WAIT]]
[CHARACTER SET charset_name]
[export_options]
| INTO DUMPFILE 'file_name'
| INTO var_name [, var_name]
[FOR UPDATE | LOCK IN SHARE_MODE_WAIT]

SELECT is used to retrieve rows
The most commonly used clause
Each select_expr indicates a column
In MySQL 5.6.6 and later, SELECT can also be used to retrieve rows
See <http://dev.mysql.com/doc/>
In the WHERE expression, you can use the SELECT clause to retrieve rows
See also: Online help select

SQL Commands and Operators



SQL Commands : DDL

This component is used to **define the structure (or schema)** of the database

For tables there are **three main commands** :

CREATE TABLE table_name

To create a table in the database

ALTER TABLE table_name

To add or remove columns from a table in the database

DROP TABLE table_name

To remove a table from the database

SQL Commands : DDL

Example :

```
CREATE TABLE customer (
    customer_id int NOT NULL,
    country varchar(50) NULL
);

ALTER TABLE customer ADD PRIMARY KEY(customer_id);

DROP TABLE customer;
```

SQL Commands : DML

This component is used to manipulate data within a table

There are four main commands :

SELECT

To select rows of data from a table

INSERT

To insert rows of data into a table

UPDATE

To change rows of data in a table

DELETE

To remove rows of data from a table

SQL Commands : DML

Example :

```
SELECT COUNT(customer_id) from customer;  
SELECT COUNT(DISTINCT customer_id) from customer;  
  
INSERT INTO customer  
(customer_id,country)  
VALUES (12000,'Indonesia');  
  
UPDATE customer SET  
country='Republik Indonesia' WHERE  
customer_id=12000;  
  
DELETE FROM customer WHERE customer_id=12000;
```

Additional DML Commands

SELECT DISTINCT Syntax

```
SELECT DISTINCT column1, column2, ...
FROM table_name;
```

BETWEEN Syntax

```
SELECT column_name(s)
FROM table_name
WHERE column_name BETWEEN value1 AND value2;
```

LIKE Syntax

```
SELECT column1, column2, ...
FROM table_name
WHERE columnN LIKE pattern;
```

LIKE Operator	Description
WHERE CustomerName LIKE 'a%'	Finds any values that start with "a"
WHERE CustomerName LIKE '%a'	Finds any values that end with "a"
WHERE CustomerName LIKE '%or%'	Finds any values that have "or" in any position
WHERE CustomerName LIKE '_r%'	Finds any values that have "r" in the second position
WHERE CustomerName LIKE 'a___%'	Finds any values that start with "a" and are at least 3 characters in length
WHERE ContactName LIKE 'a%oo'	Finds any values that start with "a" and ends with "o"

Additional DML Commands

AND Syntax

```
SELECT column1, column2, ...
FROM table_name
WHERE condition1 AND condition2 AND condition3 ...;
```

OR Syntax

```
SELECT column1, column2, ...
FROM table_name
WHERE condition1 OR condition2 OR condition3 ...;
```

NOT Syntax

```
SELECT column1, column2, ...
FROM table_name
WHERE NOT condition;
```

Let's Practice !

https://www.w3schools.com/sql/sql_syntax.asp

SQL Operators

Operator	Description
=	equal
!=, <>	not equal
>	more than
>=	more than or equal
<	less than
<=	less than or equal

Aggregate

GROUP BY

Used in collaboration with the SELECT statement to arrange identical data into groups

GROUP BY

is often used with aggregate functions

Function	Description
AVG	Calculates the average of a set of values
COUNT	Counts rows in a specified table or view
MAX	Gets the maximum value in a set of values
MIN	Gets the minimum value in a set of values
SUM	Calculates the sum of values

Aggregate

https://www.w3schools.com/sql/sql_orderby.asp

SQL Statement:

```
SELECT * FROM Customers
ORDER BY Country;
```

Number of Records: 91

CustomerID	CustomerName	ContactName	Address	City	PostalCode	Country
12	Cactus Comidas para llevar	Patricia Simpson	Cerrito 333	Buenos Aires	1010	Argentina
54	Océano Atlántico Ltda.	Yvonne Moncada	Ing. Gustavo Moncada 8585 Piso 20-A	Buenos Aires	1010	Argentina
64	Rancho grande	Sergio Gutiérrez	Av. del Libertador 900	Buenos Aires	1010	Argentina
20	Ernst Handel	Roland Mendel	Kirchgasse 6	Graz	8010	Austria
59	Piccolo und mehr	Georg Pippes	Geislweg 14	Salzburg	5020	Austria
50	Maison Dewey	Catherine Dewey	Rue Joseph-Bens 532	Bruxelles	B-1180	Belgium
76	Suprêmes délices	Pascale Cartrain	Boulevard Tirou, 255	Charleroi	B-6000	Belgium
15	Comércio Mineiro	Pedro Afonso	Av. dos Lusíadas, 23	São Paulo	05432-043	Brazil

ORDER BY

The ORDER BY keyword is used to sort the result-set in ascending or descending order

ORDER BY

The ORDER BY keyword sorts the records in ascending order by default

To sort the records in descending order, use the DESC keyword

Mini Practice

Questions :

Buatlah query yang dapat membuat output data berupa :

- Daftar customer dengan Country = Spain
- Jumlah customer dengan Country = Germany
- Top 3 City dengan customer terbanyak
- Customer dengan CustomerID = 66
- Customer dengan PostalCode = '05634-030'

Assignment

Assignment 8

Instruksi Assignment 8

using dvrental database you've uploaded to your local postgresql db, please answer these questions

Exercise :

1. Identify the top 10 customers and their email so we can reward them
2. Identify the bottom 10 customers and their emails
3. What are the most profitable movie genres (ratings)?
4. How many rented movies were returned late, early, and on time?
5. What is the customer base in the countries where we have a presence?
6. Which country is the most profitable for the business?
7. What is the average rental rate per movie genre (rating)?

Output : **Topik 15 16 17 - [Nama Lengkap].pdf** file contains :

- written sql query
- screenshot of query result
- u can also visualize the result using ppt chart / python EDA library (optional)

Available from	Until
May 9 at 09.00 PM	May 22 at 11.59PM

Thank you!

Any Questions?

zenius



**Kampus
Merdeka**
INDONESIA JAYA

