



FINAL PROJECT

KALBE NUTRITIONALS – DATA ENGINEER

PRESENTED BY
Ivanoel Rhema C. A.

SEPTEMBER 2023

About Me



Ivanoel Rhema
Data Engineer Intern

A Telecommunication Engineering student with expertise in cybersecurity, networking, electrical and IoT, complemented by strong public communication, project management and technical skills. I am eager to apply my technical knowledge, analytical mindset, and creative thinking to drive impactful solutions.

Currently, I am enthusiastically pursuing a career in cloud computing, with a specific focus on data engineering. This dynamic field fascinates me due to its potential to transform how data is managed and leveraged for various applications.



SEPTEMBER 2023



Experience :

- IEEE ITB Student Branch – Bandung, Indonesia Jun 2023 – Present
Marketing Assistant
- IEEE ITB Student Branch – Bandung, Indonesia Jul 2022 – Apr 2023
Project Officer
- Bengkel Radio IMT "Signum" ITB – Bandung, Indonesia Jan 2023 – Mar 2023
IoT Researcher Intern
- Kaderisasi Wilayah STEI-R 2023 – Bandung, Indonesia May 2023 – Jun 2023
Head of publication and documentation

SEPTEMBER 2023

challenge

1. Create a shell/bash script to check whether directory exists inside a given path.

Pada tugas nomor 1, intern diminta untuk membuat suatu teks shell/bash script yang memiliki fungsi untuk memastikan atau mengecek apakah terdapat suatu directory di path. Ketika script dijalankan maka akan mengeluarkan output yang berupa echo "There is \$name_of_directory Directory Exists!" apabila ditemukan directory yang dimaksud. Akan tetapi, apabila tidak menemukan directory yang dimaksud maka pesan yang keluar berupa echo "\$name_of_directory Directory Not Exists!" lalu akan membuat directory tersebut. Tugas juga memiliki permintaan untuk membuat suatu crontab syntax yang akan menjalankan script tersebut setiap pukul 7 pagi.

```
name_of_directory='data1'
path="kalbe/hdfs/data"
cd $path
case $name_of_directory in
  $path)
    echo "There is $name_of_directory Directory Exists!"
    ;;
  *)
    echo "$name_of_directory Directory Not Exists!"
    mkdir "$name_of_directory"
    echo "Created $name_of_directory directory."
    ;;
esac
```

```
(crontab -l ; echo "0 7 * * * E:/ITB/Kalbe Internship/Final Project/no1.sh") | crontab -
```

Link GitHub :

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/89c8a5dfe6b725b57999bf6940cd3ddb14b052f9/NO.1.zip

challenge

2. Using the question number 1 script, add another condition if directory exists inside the path

Pada tugas nomor 2, intern diminta untuk membuat suatu teks shell/bash script yang melanjutkan tugas pada nomor 1 dengan beberapa condition tambahan. Condition yang dimaksud merupakan suatu command yang bertujuan untuk menyalin/meng-copy file excel yang diminta menuju pada directory yang ditentukan.

Link GitHub

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/1e73cfdf6c9255b46e1cde52678b2a44620abe5/no2.sh

```
name_of_directory='data1'
filename_excel='daily_market_price.xlsx'
source_dir='local/data/market'
cd $source_dir
case $name_of_directory in
  $source_dir)
    echo "There is $name_of_directory Directory Exists!"
    ;;
  *)
    echo "$name_of_directory Directory Not Exists!"
    mkdir "$name_of_directory"
    echo "Created $name_of_directory directory."
    ;;
esac
cp $filename_excel $name_of_directory
```

challenge

3. Complete below Syntax {Highlighted Sentence} to insert data from Python to MySQL.



Pada tugas nomor 3, intern diminta untuk membuat suatu code python untuk melakukan koneksi dengan MySQL menggunakan library MySQL connector untuk meng-insert data MySQL menggunakan Python.

Link GitHub :

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/e89fda03da873ecee584f792991f6aa940738c5d/no%20syntax.py

```
#Melakukan import mysql connector
import mysql.connector

#Melakukan percobaan koneksi
conn = mysql.connector.connect(
    host="local_host",
    user="username",
    password="user_password",
    database="user_database")

#Membuat object cursor sebagai penanda
cursor = conn.cursor()

#Deklarasi SQL Query untuk memasukan record ke DB (KARYAWAN)
sql = "INSERT INTO KARYAWAN (FIRST_NAME, LAST_NAME, AGE, SEX, INCOME) VALUES (%s,%s,%d,%s,%d)"
values = ("Ivan", "Aditama", 21, "Laki", 7500000)

try:
    #Eksekusi SQL Command
    cursor.execute(sql, values)
    #Melakukan perubahan (commit) pada DB
    conn.commit()
except:
    #Roll back apabila ada issue
    conn.rollback()

#Menutup koneksi
conn.close()
```

challenge

4. Convert this instruction into SQL Query Language.

Pada tugas nomor 4, intern diminta untuk membuat query SQL untuk dapat menampilkan item_name yang memiliki jumlah item_total paling tinggi, lalu menghapus atau meng-drop item_name yang memiliki item_total terendah.

Link GitHub :
https://github.com/Ivanoel17/Kalbe_FinalProject/blob/e89fda03da873ecee584f792991f6aa940738c5d/No_4.sql

```
CREATE TABLE Inventory(
    item_code int PRIMARY KEY,
    item_name varchar(50),
    item_price int,
    item_total int
);

INSERT INTO Inventory(item_code,item_name,item_price,item_total)
VALUES (2341,'Promag Tablet',3000,100),(2342,'Hydro Coco 250ML',7000,20),(2343,'Nutrive Benecol 100ML',20000,30),(2344,'Blackmores Vit C 500Mg',9000,100),(2345,'Entrasol Gold 370G',10000,100);

select item_name from Inventory where item_total=(select max(item_total) from Inventory);

update Inventory set item_price = 100000 where item_name = (select item_name from Inventory where item_total= (select max(item_total) from Inventory));

select item_name, item_price from Inventory where item_price = 100000;

insert into Inventory(item_code,item_name,item_price,item_total)
values (2343, 'high_item', 3000,120)

delete from Inventory where item_total = (select max(item_total) from Inventory );

drop table Inventory;
```

	item_name
▶	Entrasol Gold 370G

*Gambar script tidak dapat ditampilkan full karena keterbatasan ruang

challenge

5. Create a Query to display all customer orders where purchase amount is less than 100 or exclude those orders which order date is on or greater than 25 Aug 2022 and customer id is above 2001. Sample table:

customer_orders

SEPTEMBER 2023

Pada tugas nomor 5, intern diminta untuk membuat query SQL untuk dapat menampilkan customer yang melakukan order kurang dari 100 kali atau customer yang melakukan order dengan **order_date** sebelum 25 Agustus 2022 dengan **customer_id** diatas 2001.

Link GitHub :

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/e89fda03da873ecee584f792991f6aa940738c5d/No_5.sql

```
CREATE TABLE customer_orders(
    order_no int PRIMARY KEY,
    purchase_amount int,
    order_date date,
    customer_id int,
    salesman_id int
);

insert into customer_orders(order_no,purchase_amount,order_date,customer_id,salesman_id)
values  (10001,150,'2022-10-05',2005,3002),
        (10009,270,'2022-09-10',2001,3005),
        (10002,65,'2022-10-5',2002,3001),
        (10004,110,'2022-08-17',2009,3003),
        (10007,948,'2022-09-10',2005,3002),
        (10005,2400,'2022-07-27',2007,3001);

select customer_id from customer_orders where purchase_amount < 100 or (order_date<='2022-08-25' and customer_id>2001);
```

	customer_id
>	2002
	2009
	2007

challenge

6. Please explain what is wrong with this picture and give the best solution for this case.

The image displays two side-by-side views. On the left is a screenshot of a database schema named 'dbo_FormN_168'. It lists several columns: ddPlant (nvarchar(max), null), ddPTBU (nvarchar(max), null), ddReqBU (nvarchar(max), null), ddReqLOB (nvarchar(max), null), ddReqLokasi (nvarchar(max), null), ddReqLokasiAstTetap (nvarchar(max), null), ddReqLokasiKategori (nvarchar(max), null), ddReqOrnBU (nvarchar(max), null), ddRequestType (nvarchar(max), null), ddTypeofClaim (nvarchar(max), null), and ddTypeofPurchase (nvarchar(max), null). The 'ddPlant' column is highlighted with a red box. On the right is a user interface for a 'General Information' form. It includes dropdown menus for 'Request Type' (set to 'Modern Trade PO') and 'Plant' (set to 'Belawan'). The 'ddPlant' value from the database schema is also highlighted with a red box, indicating a discrepancy between the database structure and the application's implementation.

Pada tugas nomor 6, intern diminta untuk mengamati dan menganalisis kesalahan yang terjadi pada gambar diatas, sekaligus memberikan solusi untuk memecahkan kesalahan tersebut. Kesalahan terletak pada referensi tabel yang penggunaannya belum sesuai dan tipe data yang tidak sesuai atau tidak konsisten. Hal tersebut dapat terlihat pada tabel atribut ddplant. Dimana masih ada kemungkinan mendapatkan nilai NULL, dimana hal tersebut kurang sesuai. Seharusnya supaya tidak ada nilai NULL maka ditambahkan NOT NULL constraint pada atribut ddplant.



Solusi yang dapat digunakan:

```
ALTER TABLE table_name  
MODIFY COLUMN ddPlant varhcar(max) NOT NULL;
```

[Link GitHub](#)

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/1e73cfdf6c9255b46e1cde52678b2a44620abe5/No.%206.txt

challenge

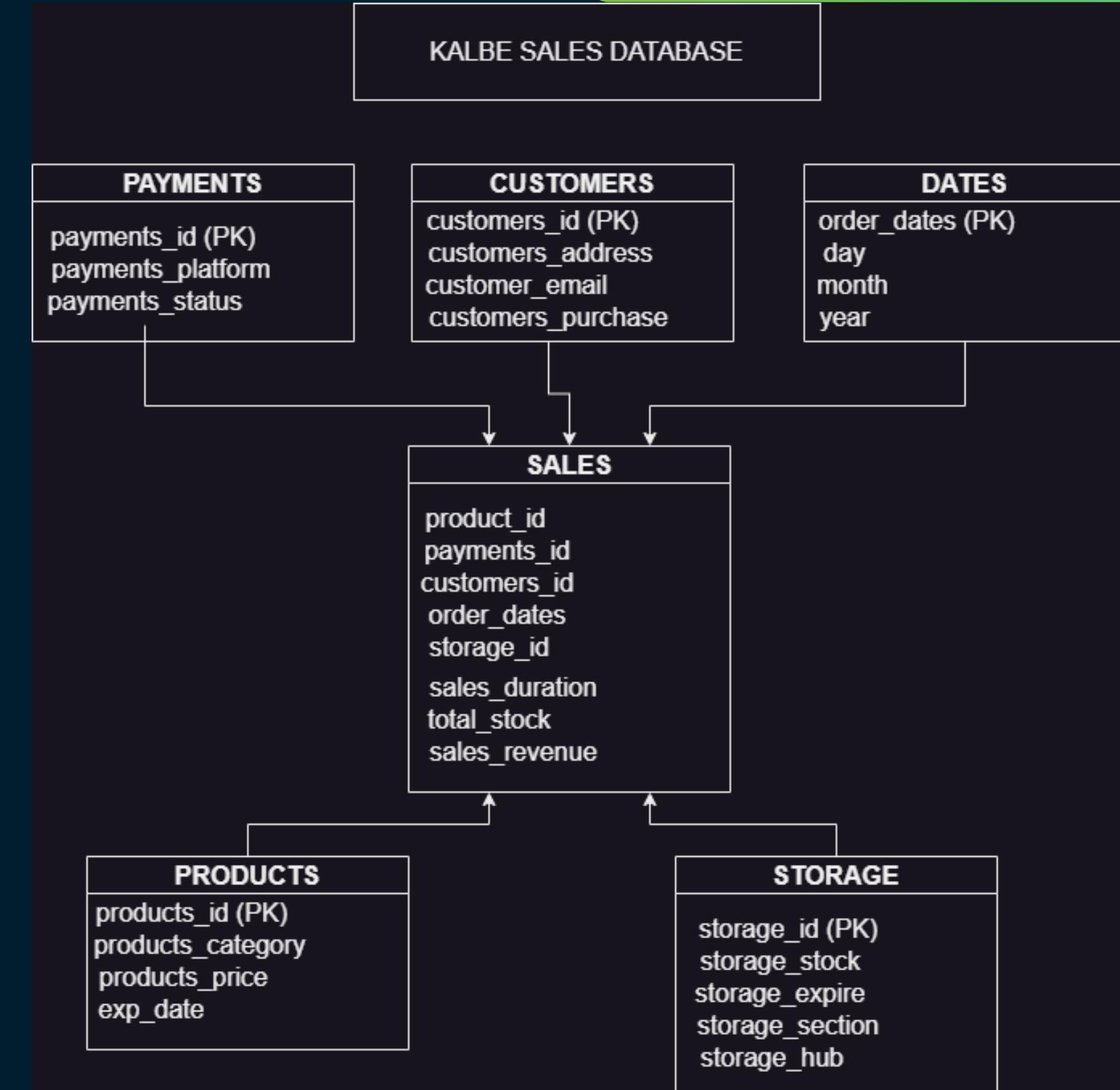
**7. Create a simple star schema for KALBE database
consist of 1 Fact and 5 Dimensions using Physical
Data Model Theory.**



Pada tugas nomor 7, intern diminta untuk membuat star schema untuk KALBE Database yang memuat 1 fact dan 5 dimension menggunakan Pyshical Data Model Theory. Dapat dilihat bahwa database yang ditampilkan merupakan SALES, dimana database SALES terbentuk serta terintegrasi melalui 5 dimension lainnya seperti PAYMENTS, CUSTOMERS, DATES, PRODUCTS dan STORAGE.

Link GitHub

https://github.com/Ivanoel17/Kalbe_FinalProject/blob/1e73cfdf6c9255b46e1cde52678b2a44620abe5/No.7.png



GitHub Link:

https://github.com/Ivanoel17/Kalbe_FinalProject.git

Drive Link:

https://drive.google.com/drive/folders/1SPksrwdd_2GRei_w26rxSisv-XScVTVR?usp=sharing

Video Link:

<https://drive.google.com/file/d/1nB95yQSs0FjmJeaPC5kGe9n5bG6k7SJ4/view?usp=sharing>



THANK YOU!

SEPTEMBER 2023