



# PORTFOLIO

*My*

Rheza Pahlevie





*Welcome*

# **TO MY PORTFOLIO PRESENTATION**

This portfolio contains projects that have been done individually in mini projects and done in groups in course projects.

# LIST OF CONTENTS

01

ABOUT ME

02

EDUCATION

03

WHAT CAN I DO?

04

MY SKILL

05

CERTIFICATIONS

06

MINI PROJECT

07

COURSE PROJECT

# *Introducing* **ABOUT ME**

I am a fifth-semester student majoring in information systems. I have a passion for data, especially as a data analyst. I have a strong interest in UI/UX and also have a deep interest in product management and project management. I have skills in data visualization using Tableau and in creating UI/UX designs and conducting UX research. I have high dedication and excellent teamwork skills and can lead a team by prioritizing collaboration and communication.



**Rheza Pahlevie**  
Data Enthusiast



# EDUCATION



**BRAWIJAYA UNIVERSITY**  
**(2021 - Now)**

Bachelor of Information System  
GPA : 3.69



# WHAT CAN I DO FOR YOU?



Analyze Data



Data Visualization

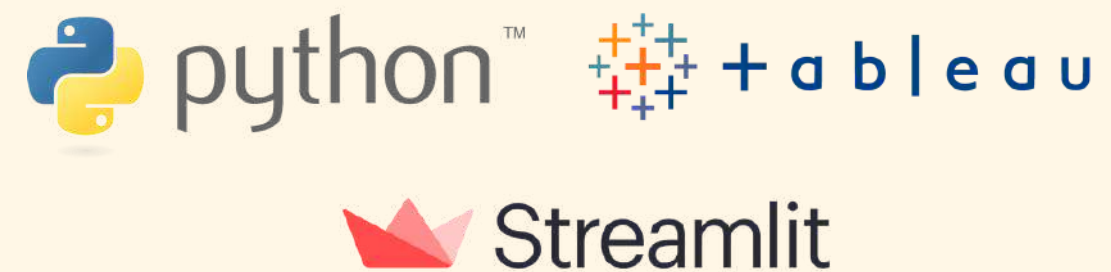


Product Management



Project Management

# MY SKILL



Data Analytics and Visualization



Database and  
Warehousing



Creative Design

# CERTIFICATIONS



[More Details](#)



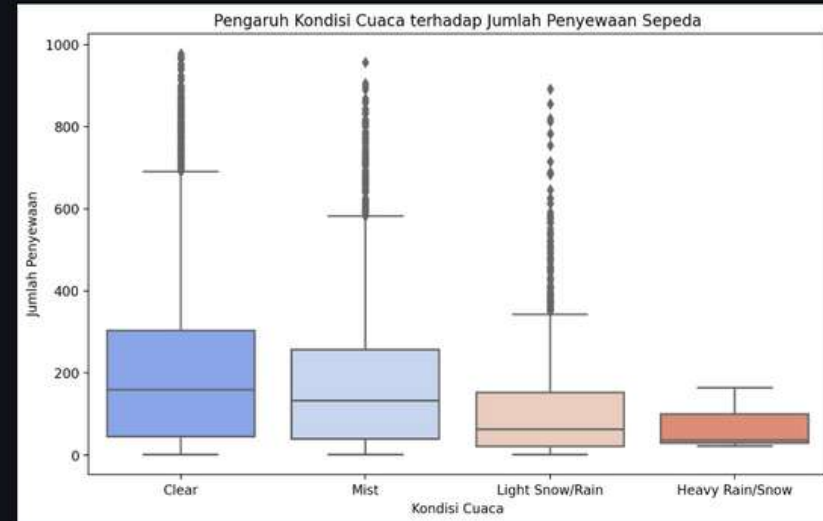


# MINI PROJECT

Projects that Have been Done Individually as a Data Enthusiast



## Pengaruh Kondisi Cuaca terhadap Jumlah Penyewaan Sepeda



## Tren Harian Penyewaan Sepeda



# MINI PROJECT 01

Analyze Data and  
Visualize using Streamlit

(November 2023 - December  
2023)

## Tren Musiman Penyewaan Sepeda (2011-2012)





This project is to perform data analysis on the bicycle rental dataset. The analysis was conducted to answer 2 analysis questions:

- 1. What are the daily and seasonal trends of bicycle rentals over two years (2011 and 2012)?
- 2. What is the effect of weather conditions on the number of bicycle rentals? Is there a significant difference between certain weather conditions?

The data analysis process was performed on the Google Colab platform. The steps of the analysis process included data wrangling, Exploratory Data Analysis, and Data Visualisation using Streamlit.

[Github](#)

```
Menyapkan semua library yang dibutuhkan

[ ] import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

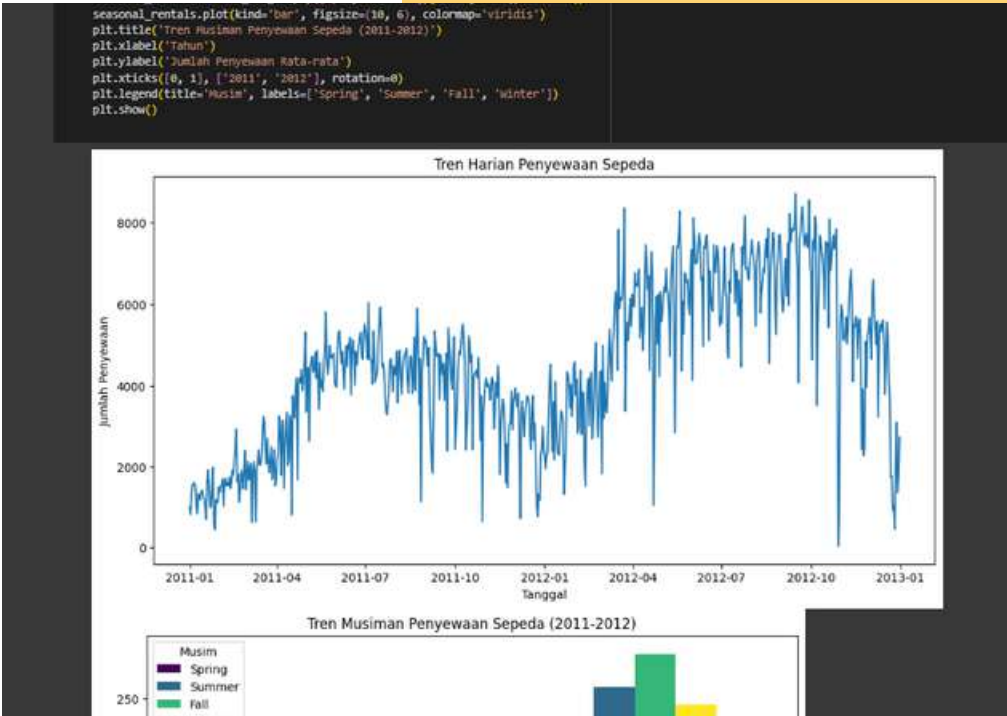
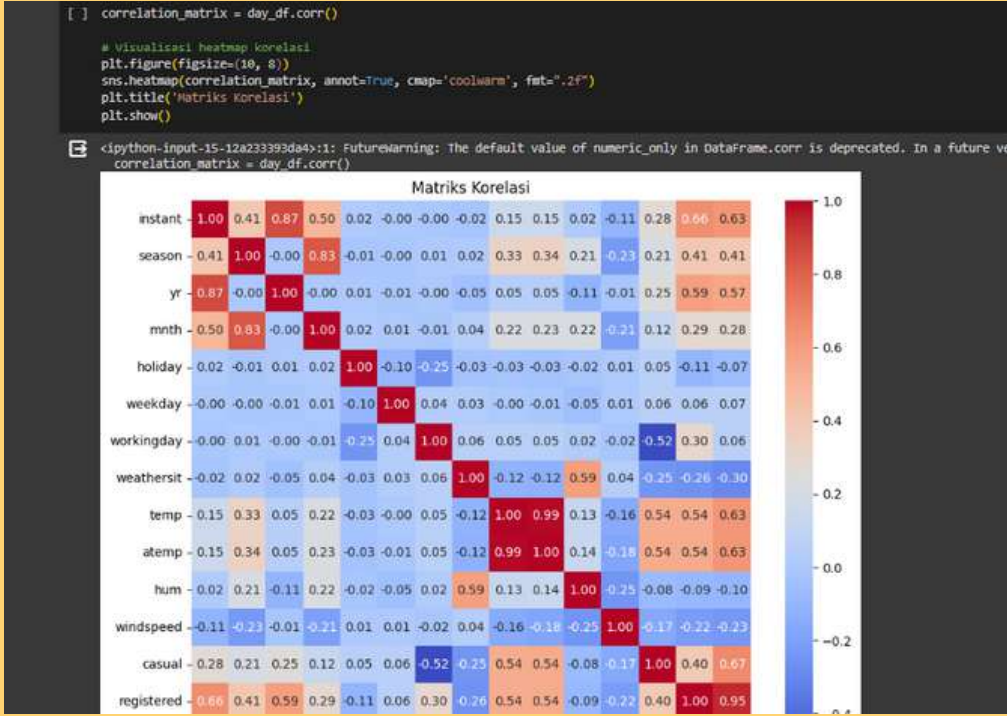
Data Wrangling

Gathering Data

Load Table Day

[ ] day_df = pd.read_csv("https://raw.githubusercontent.com/rhezapahlewie15/Data_Projek_IDCamp/main/day.csv")
day_df.head(10)
```

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed
0	1	2011-01-01	1	0	1	0	6	0	2	0.344167	0.363625	0.805833	0.160446
1	2	2011-01-02	1	0	1	0	0	0	2	0.363478	0.353739	0.696087	0.248539
2	3	2011-01-03	1	0	1	0	1	1	1	0.196364	0.189405	0.437273	0.248309
3	4	2011-01-04	1	0	1	0	2	1	1	0.200000	0.212122	0.590435	0.160296
4	5	2011-01-05	1	0	1	0	3	1	1	0.228957	0.229270	0.436957	0.186900
5	6	2011-01-06	1	0	1	0	4	1	1	0.204348	0.233209	0.518261	0.089565
6	7	2011-01-07	1	0	1	0	5	1	2	0.196522	0.206839	0.498696	0.168726
7	8	2011-01-08	1	0	1	0	6	0	2	0.165000	0.162254	0.535833	0.266804
8	9	2011-01-09	1	0	1	0	0	0	1	0.138333	0.146175	0.434487	0.361950



# MINI PROJECT 02

## Create a Machine Learning to Classify Images

(November 2023 - December 2023)

```
[ ] model = Sequential([
    Conv2D(32, (3, 3), input_shape=(100, 150, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Conv2D(128, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Conv2D(256, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Flatten(),
    Dense(512, activation='relu'),
    Dense(3, activation='softmax') # 3 kelas: rock, paper, scissor
])

[ ] model.summary()
```

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 98, 148, 32)	896
max_pooling2d_4 (MaxPooling2D)	(None, 49, 74, 32)	0
conv2d_5 (Conv2D)	(None, 47, 72, 64)	18496

```
[ ] local_zip = '/content/rockpaperscissors.zip'
zip_ref = zipfile.ZipFile(local_zip, 'r')
zip_ref.extractall('/content/rockpaperscissors')
zip_ref.close()

[ ] base_dir = '/content/rockpaperscissors/rockpaperscissors/rps-cv-images'
rock = os.path.join('/content/rockpaperscissors/rockpaperscissors/rock')
paper = os.path.join('/content/rockpaperscissors/rockpaperscissors/paper')
scissors = os.path.join('/content/rockpaperscissors/rockpaperscissors/scissors')

[ ] train_rock = os.listdir(rock)
train_paper = os.listdir(paper)
train_scissors = os.listdir(scissors)

[ ] print('Total Data Latih Rock :', len(train_rock))
print('Total Data Latih Paper :', len(train_paper))
print('Total Data Latih Scissors :', len(train_scissors))

Total Data Latih Rock : 726
Total Data Latih Paper : 712
Total Data Latih Scissors : 750

[ ] train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=30,
    horizontal_flip=True,
    shear_range=0.2,
    fill_mode='wrap',
    validation_split=0.4)

train_generator = train_datagen.flow_from_directory(
    base_dir,
    target_size=(100, 150),
    shuffle=True,
    class_mode='categorical',
    subset='training')

validation = train_datagen.flow_from_directory(
    base_dir,
    target_size=(100, 150),
    shuffle=True,
    class_mode='categorical',
    subset='validation')

Found 1314 images belonging to 3 classes.
Found 874 images belonging to 3 classes.
```





```
[ ] model.fit(  
    train_generator,  
    steps_per_epoch=25,  
    epochs = 20,  
    validation_data=validation,  
    validation_steps = 5,  
    verbose=2  
)
```

Epoch 1/20  
25/25 - 22s - loss: 1.0046 - accuracy: 0.4900 - val\_loss: 0.5939 - val\_accuracy: 0.8438 - 22s/  
Epoch 2/20  
25/25 - 22s - loss: 0.4059 - accuracy: 0.8487 - val\_loss: 0.2122 - val\_accuracy: 0.9187 - 22s/  
Epoch 3/20  
25/25 - 21s - loss: 0.3612 - accuracy: 0.8792 - val\_loss: 0.3766 - val\_accuracy: 0.9062 - 21s/  
Epoch 4/20  
25/25 - 21s - loss: 0.1873 - accuracy: 0.9388 - val\_loss: 0.2854 - val\_accuracy: 0.9187 - 21s/  
Epoch 5/20  
25/25 - 20s - loss: 0.2415 - accuracy: 0.9150 - val\_loss: 0.2984 - val\_accuracy: 0.9312 - 20s/  
Epoch 6/20  
25/25 - 21s - loss: 0.1972 - accuracy: 0.9377 - val\_loss: 0.1820 - val\_accuracy: 0.9563 - 21s/  
Epoch 7/20  
25/25 - 21s - loss: 0.1338 - accuracy: 0.9550 - val\_loss: 0.2309 - val\_accuracy: 0.9438 - 21s/  
Epoch 8/20  
25/25 - 21s - loss: 0.1340 - accuracy: 0.9500 - val\_loss: 0.2403 - val\_accuracy: 0.9375 - 21s/  
Epoch 9/20



This project is to create a machine learning algorithm to classify images into 3 classes, namely Rock, Paper, and Scissors. This project is built using tools and libraries in Tensorflow. The model architecture in this project is built using Convolutional Neural Network (CNN). The resulting machine learning algorithm can make predictions on uploaded images.

[Github](#)

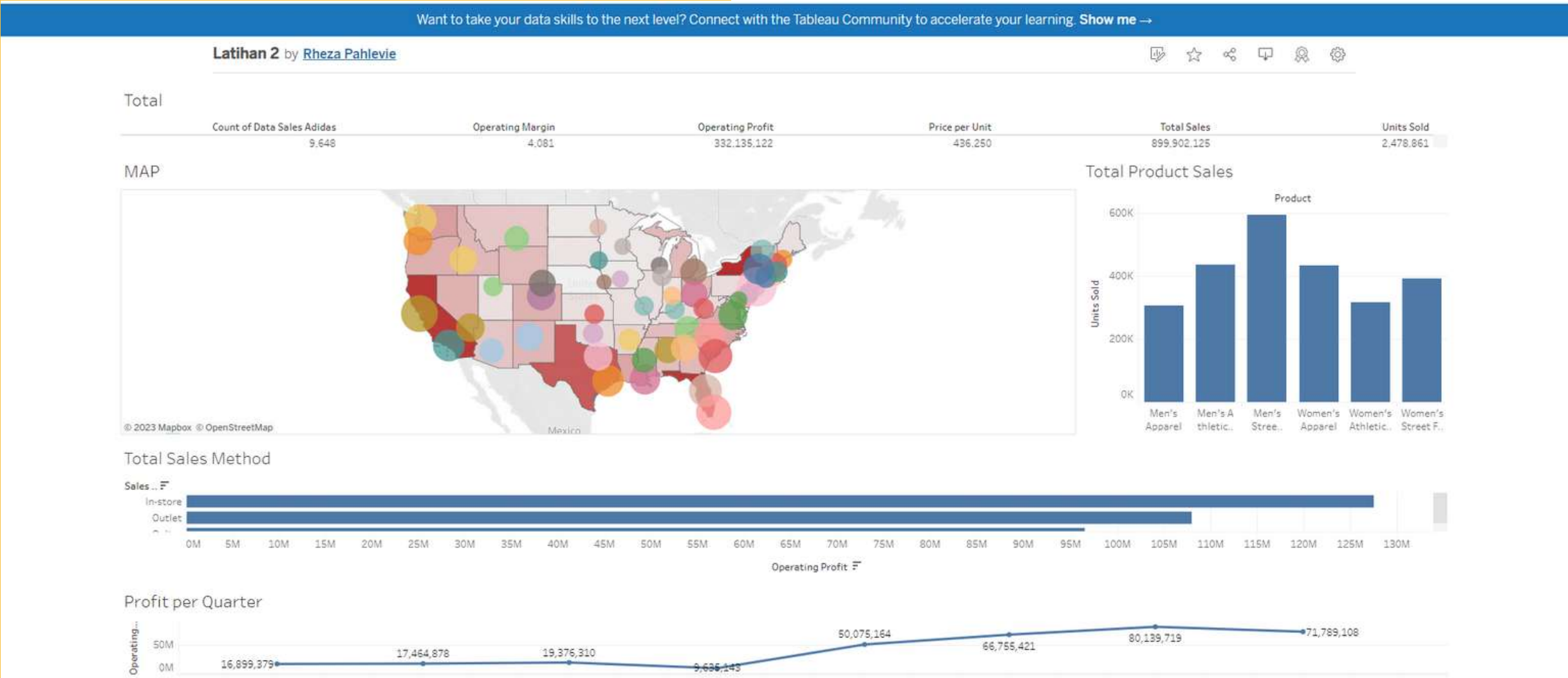
# MINI PROJECT 03

## DASHBOARD ANALYTICS USING TABLEAU

(September 2023 - October 2023)

[Github](#)

This project is to create an analytic dashboard using Tableau. The data used is sales data for the adidas brand in the US. The dashboard displays several information such as, total sales per region, the most sold product category, the type of sales that has the largest total sales, and profit per quarter.





# COURSE PROJECT

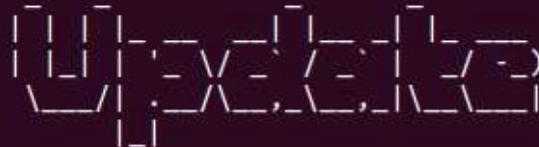
Projects that Have been Done in Groups to Complete Course Assignments as an  
Information Systems Student



```
rheza@rheza-Virtu
```

```
File Edit View Search Terminal Help
```

```
(base) rheza@rheza-VirtualBox:~$ export PATH=/usr/bin:/bin  
(base) rheza@rheza-VirtualBox:~$ conda activate pbd  
(pbd) rheza@rheza-VirtualBox:~$ jupyter notebook
```



Read the migration plan to Notebook 7 to learn about the new features and the ac

[https://jupyter-notebook.readthedocs.io/en/latest/migrate\\_to\\_notebook7.html](https://jupyter-notebook.readthedocs.io/en/latest/migrate_to_notebook7.html)

Please note that updating to Notebook 7 might break some of your extensions.

```
[I 10:37:41.482 NotebookApp] Serving notebooks from local directory: /home/rheza  
[I 10:37:41.483 NotebookApp] Jupyter Notebook 6.5.6 is running at:  
[I 10:37:41.486 NotebookApp] http://localhost:8888/?token=7e5db51ea8201bae2bf114  
[I 10:37:41.486 NotebookApp] or http://127.0.0.1:8888/?token=7e5db51ea8201bae2b  
[I 10:37:41.486 NotebookApp] Use Control-C to stop this server and shut down all  
[C 10:37:41.572 NotebookApp]
```

To access the notebook, open this file in a browser:

file:///home/rheza/.local/share/jupyter/runtime/nbsserver-2283-open.html

Or copy and paste one of these URLs:

```
File Edit View Search Terminal Help
(base) rheza@rheza-VirtualBox:~$ export PATH=/usr/bin:/bin
(base) rheza@rheza-VirtualBox:~$ su hduser
Password:
hduser@rheza-VirtualBox:/home/rheza$ cd
hduser@rheza-VirtualBox:~$ ls -l
total 690008
-rw-r--r--  1 hduser  hadoop           8980 Nov 24 12:52 examples.desktop
drwxr-xr-x 12 hduser  hadoop          4096 Nov 24 13:58 hadoop-3.3.5
-rw-r--r--  1 hduser  hadoop 706533213 Mar 16  2023 hadoop-3.3.5.tar.gz
-rw-r--r--  1 hduser  root           3981 Nov 24 14:04 Mall_Customers2.csv
drwx-----  3 hduser  hadoop          4096 Nov 24 15:30 snap
drwxr-xr-x  4 hduser  hadoop          4096 Nov 24 13:59 tmpdata
hduser@rheza-VirtualBox:~$
```

# COURSE PROJECT 01

# ANALYZE DATA IN BIG DATA ENVIRONMENT

OCTOBER 2023 - NOVEMBER 2023



# COURSE PROJECT 01

This project is the final project of the big data and analytics course. this project performs data analysis in a big data environment. The analysis carried out is to determine customer segmentation using Kmeans Clustering. The dataset is stored in HDFS and the analysis process is carried out on Jupyter Notebook. Customer segmentation is done to determine the right sales strategy to several categories of customers.

## Import Findspark

```
[3]: import findspark
findspark.init()
```

## Inisiasi SparkSession

```
[4]: import pyspark
```

```
[5]: from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("bda").getOrCreate()

23/11/27 15:57:51 WARN Utils: Your hostname, rheza-Vir
2.15 instead (on interface enp0s3)
23/11/27 15:57:51 WARN Utils: Set SPARK_LOCAL_IP if you
```

## Import Library untuk Visualisasi

```
[6]: import seaborn as sns
import matplotlib.pyplot as plt
```

## Mengambil Data dari HDFS

```
[7]: file_path = "hdfs://127.0.1.1:9000/makul/bda/Mall_Customers2.csv"
```

## Membaca Data dan Menampilkan Data

```
[8]: data = spark.read.csv(file_path, header=True, inferSchema=True)
```

```
[9]: data.show(10, truncate=False)
```

CustomerID	Gender	Age	Annual Income (k\$)	Spending Score (1-100)
1	Male	19	15	39
2	Male	21	15	81
3	Female	20	16	6
4	Female	23	16	177

## Memproses Data Menggunakan Vector Assembler dan Sta

```
In [10]: selected_features = data.select("Age", "Annual Income (k$)", "Spending Score (1-100)")
```

```
In [11]: from pyspark.ml.feature import VectorAssembler
from pyspark.ml.feature import StandardScaler
```

```
In [12]: assembler = VectorAssembler(inputCols=["Annual Income (k$)", "Spending Score (1-100)"]
selected_features = assembler.transform(selected_features)
```

```
In [13]: scaler = StandardScaler(inputCol="features", outputCol="scaled_features", withStd=True)
scaler_model = scaler.fit(selected_features)
selected_features = scaler_model.transform(selected_features)
```

## Menggunakan KMeans untuk Pengelompokan Data

```
In [14]: from pyspark.ml.clustering import KMeans
```

```
kmeans = KMeans().setK(5).setSeed(1)
model = kmeans.fit(selected_features)
```

## Melakukan Visualisasi hasil KMeans

```
[18]: import matplotlib.pyplot as plt
import pandas as pd
from sklearn.cluster import KMeans

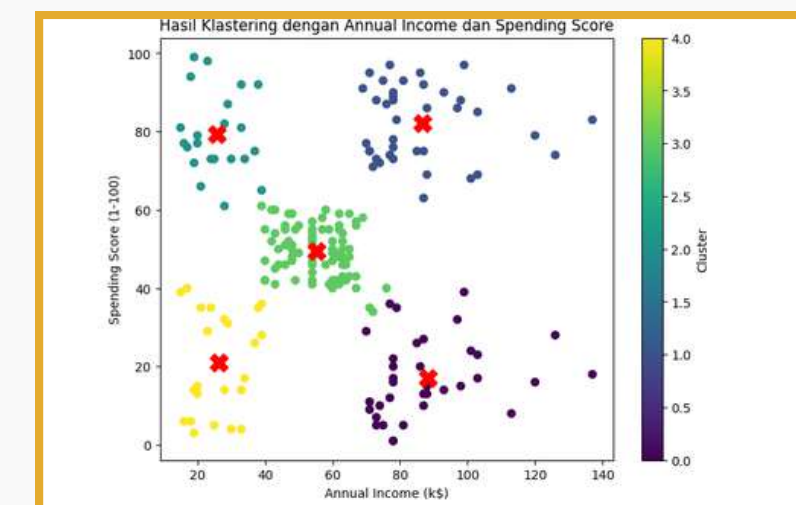
# Mengambil kolom features dan prediction untuk visualisasi
features_and_predictions = predictions.select("features", "prediction")

# Mendefinisikan jumlah kluster (sesuaikan dengan kebutuhan Anda)
num_clusters = 5

# Menampilkan hasil klustering
features_and_predictions.show()

# Konversi ke Pandas DataFrame untuk memvisualisasikan dengan matplotlib
pandas_df = features_and_predictions.toPandas()

# Visualisasi scatter plot
plt.figure(figsize=(8, 6))
scatter = plt.scatter(pandas_df["features"].apply(lambda x: x[0]), pandas_df["prediction"])
plt.title("Hasil Klustering dengan Annual Income dan Spending Score")
plt.xlabel("Annual Income (k$)")
plt.ylabel("Spending Score (1-100)")
```

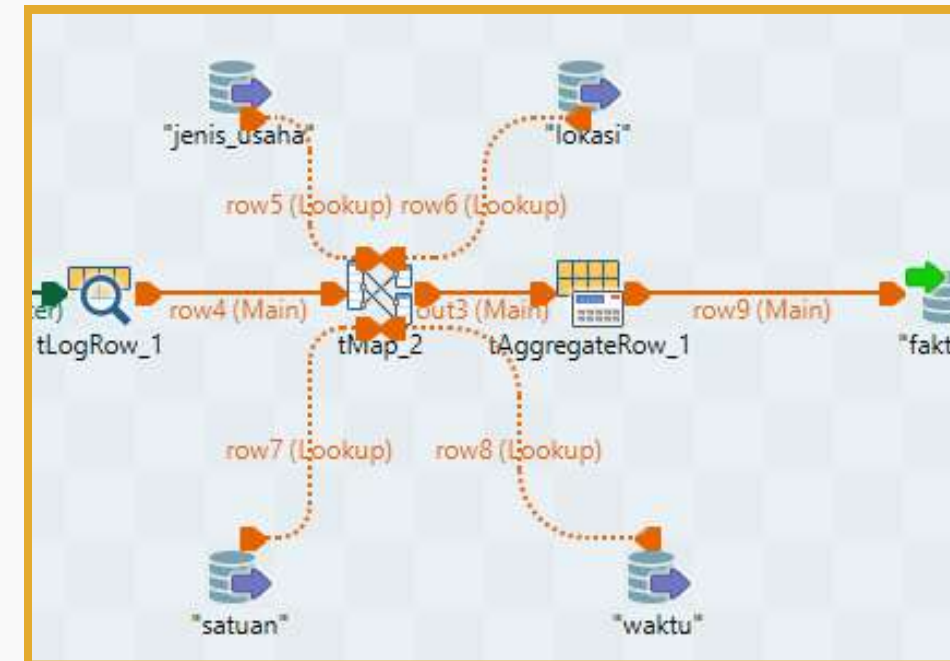


# COURSE PROJECT 02

## DATA WAREHOUSE

March 2023 - May 2023

This project stores data on the number of UMKM in West Java province into a Data Warehouse. Data is transferred using the extract, transform, and load (ETL) process. Then analyse the data in the data warehouse using OLAP.



```
SELECT DENSE_RANK() over(order by jumlah desc) AS ranking, tahun, kota, kategori_jenis_kelamin, kategori_bmi, jumlah FROM waktu as w, lokasi as l, jenis_kelamin as j, bmi as b, fakta as f WHERE w.id_waktu = f.id_waktu AND l.id_lokasi = f.id_lokasi AND j.id_jenis_kelamin = f.id_jenis_kelamin AND b.id_bmi = f.id_bmi AND kota = 'Malang' AND kategori_bmi = 'Normal';
```

☐ Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ] [ Refresh ]

☐ Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

ranking	tahun	kota	kategori_jenis_kelamin	kategori_bmi	jumlah
1	2019	Malang	Perempuan	Normal	10
2	2021	Malang	Laki-laki	Normal	7
3	2019	Malang	Laki-laki	Normal	6
3	2020	Malang	Laki-laki	Normal	6
3	2021	Malang	Perempuan	Normal	6
4	2020	Malang	Perempuan	Normal	5

Showing rows 0 - 3 (4 total, Query took 0.0006 seconds.)

```
SELECT kota, kategori_jenis_kelamin, SUM(jumlah) as jumlah_total FROM waktu as w, lokasi as l, jenis_kelamin as j, bmi as b, fakta as f WHERE w.id_waktu = f.id_waktu AND l.id_lokasi = f.id_lokasi AND j.id_jenis_kelamin = f.id_jenis_kelamin AND b.id_bmi = f.id_bmi AND tahun = 2020 AND kategori_BMI = 'Normal' AND (kota = 'Malang' OR kota = 'Surabaya') GROUP BY kota, kategori_jenis_kelamin ORDER BY JUMLAH_TOTAL DESC;
```

☐ Profiling [ Edit inline ] [ Edit ] [ Explain SQL ] [ Create PHP code ]

☐ Show all | Number of rows: 25 | Filter rows: Search this table

Extra options

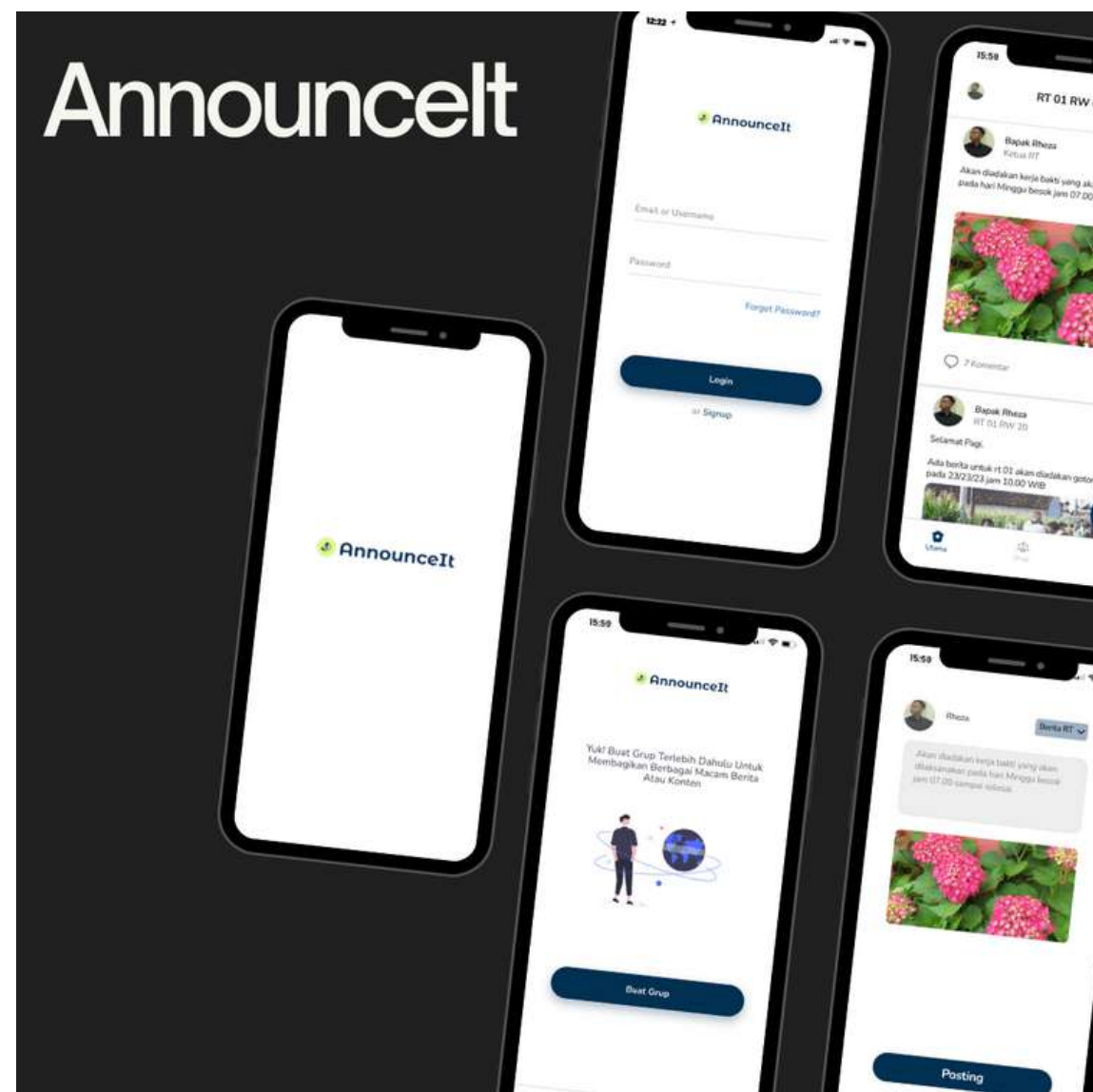
kota	kategori_jenis_kelamin	jumlah_total
Surabaya	Perempuan	7
Malang	Laki-laki	6
Surabaya	Laki-laki	6
Malang	Perempuan	5



# COURSE PROJECT 03

## Announcelt

October 2023 - December 2023



Announcelt is a project from the Information Systems Project Management course. Announce It is a content management system application for the management of broadcasting information / news / announcements of citizens. Announcelt is built using the Agile system. This project focuses on project management in developing applications. The project begins with the creation of a project charter. Then proceed with determining the requirements needed and creating a system design to create a user interface of the application. The application prototype is the output of this project.

# COURSE PROJECT 04

## CityShield

October 2023 - December 2023

CityShield is a project from the E-Government course. CityShield is an SOS application that helps users to immediately get help from the police when experiencing or seeing a criminal act with just one tap. This project focuses on the development of E-Government applications.





# COURSE PROJECT 05

## MyGKM

February 2023 - May 2023

My GKM is an application that makes it easy for Filkom UB students to make purchases from the GKM canteen. My GKM is a project from the user experience design course. This project focuses on exploring user problems to create solutions in the form of application prototypes from the problems experienced by users.





# THANKS FOR READING

Have a Good Day

