

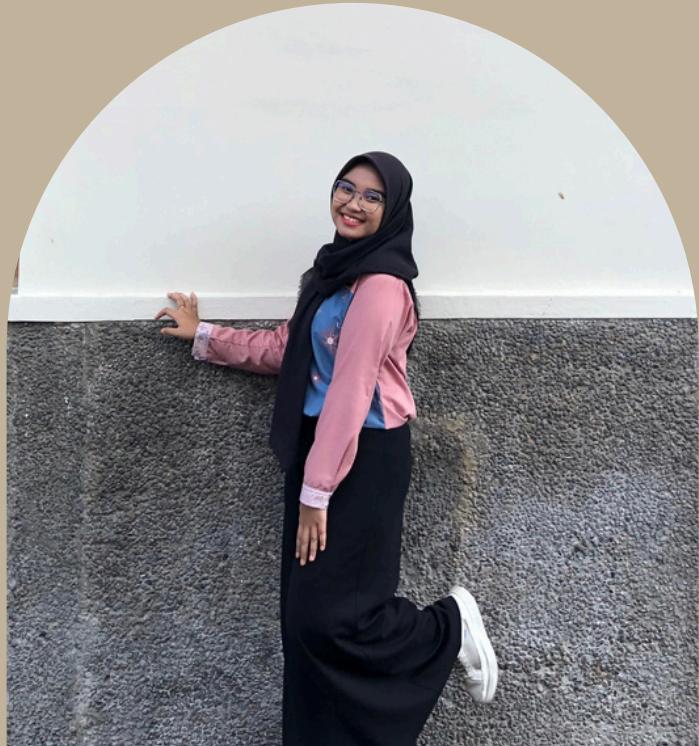
PERSONAL PORTOFOLIO

AULIA TIFFANI RIZKY

DATA ENTRY • SURVEYOR • DATA ENTHUSIAST



TENTANG SAYA



Saya adalah seorang lulusan sarjana matematika dengan pengalaman sebagai data entry pada kerja musiman di Center for Entrepreneurship and Career Development (CENDI) UIN Sunan Kalijaga. Saya memiliki ketertarikan dengan data analyst dan berkeinginan berkarir sebagai data analyst.

BACKGROUND PENDIDIKAN

**SARJANA MATEMATIKA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

Lulus dengan IPK 3.55/4.00 dengan tugas akhir

**“Indeks-Indeks Konektivitas dari Graf Koprime
dari Grup Quaternion Tergeneralisasi”**



SKILLS MICROSOFT EXCEL DAN GOOGLE DATA STUDIO

A	B	C	D	E	F	G
show_id	type	title	director	cast	country	date_added
s1	Movie	The Grand Sed	Don McKellar	Brendan Gleeson, Canada	March 30, 2021	
s2	Movie	Take Care Goo	Girish Joshi	Mahesh Manjre, India	March 30, 2021	
s3	Movie	Secrets of Decu	Josh Webber	Tom Sizemore, United States	March 30, 2021	
s4	Movie	Pink: Staying Ti	Sonia Anderson	Interviews with: United States	March 30, 2021	
s5	Movie	Monster Maker	Giles Foster	Harry Dean Stanton, United Kingdom	March 30, 2021	
s6	Movie	Living With Dinc	Paul Weiland	Gregory Chisholm, United Kingdom	March 30, 2021	
s7	Movie	Hired Gun	Fran Strine	Alice Cooper, United States	March 30, 2021	
s8	Movie	Grease Live!	Thomas Kail, Al	Julianne Hough, United States	March 30, 2021	
s9	Movie	Global Meltdow	Daniel Gilboy	Michael Paré, L	Canada	March 30, 2021
s10	Movie	David's Mother	Robert Allan Ac	Kirstie Alley, Sa	United States	April 1, 2021
s11	Movie	Forest Fairies	Justin G. Dyck	Emily Wilder, A	Canada	April 4, 2021
s12	Movie	Take Care	Liz Tuccillo	Leslie Bibb, Ke	United States	April 10, 2021
s13	Movie	The Night Eats	Dominique Roc	Anders Daniels	France	April 17, 2021
s14	Movie	Resilencia	Jep Barcelona	Rafinha Alcantara	Spain	April 24, 2021

Berdasarkan pivot table, akan dilakukan cleaning data karena terdapat beberapa kolom yang memiliki data null.

Impor data dalam bentuk Microsoft Excel.

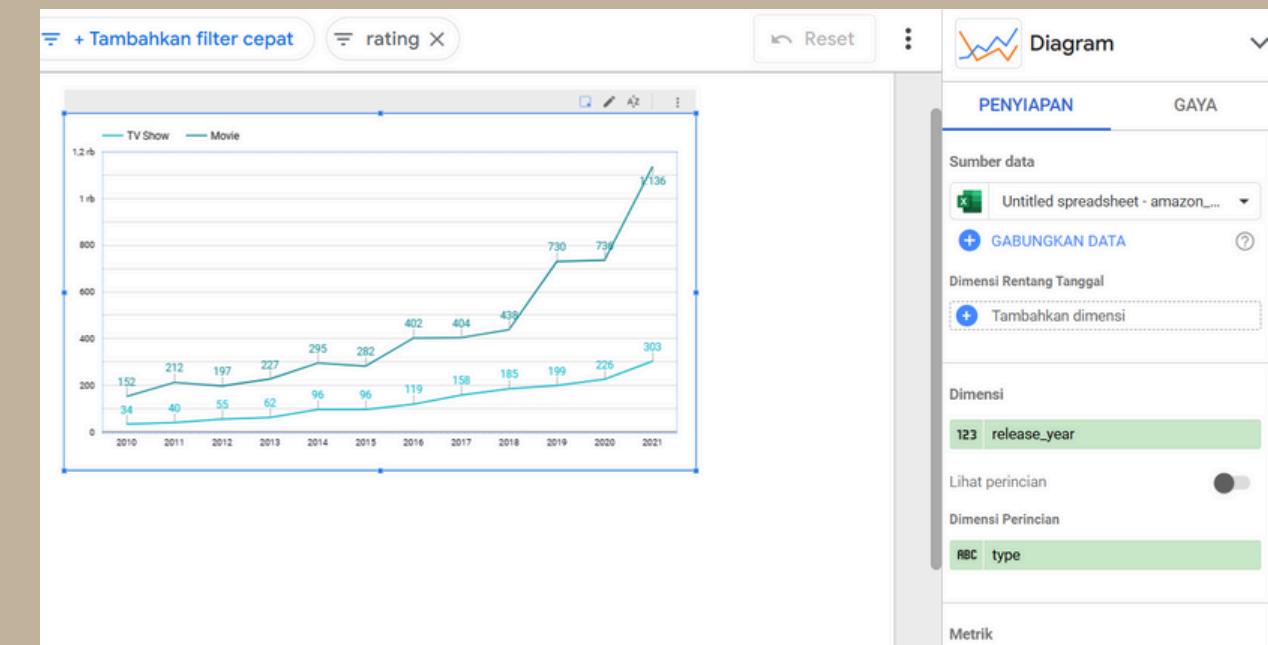
D	E	F	G	H
director	cast	country	date_added	
Don McKellar	Brendan Gleeson, Canada	March 30, 2021		Cari menu
Girish Joshi	Mahesh Manjre, India	March 30, 2021		Potong
Josh Webber	Tom Sizemore, United States	March 30, 2021		Salin
Sonia Anderson	Interviews with: United States	March 30, 2021		Opsi Tempel:
Giles Foster	Harry Dean Stanton, United Kingdom	March 30, 2021		Tempel Spesial...
Paul Weiland	Gregory Chisholm, United Kingdom	March 30, 2021		Sisipkan
Fran Strine	Alice Cooper, United States	March 30, 2021		Hapus
Thomas Kail, Al	Julianne Hough, United States	March 30, 2021		Bersihkan Konten
Daniel Gilboy	Michael Paré, L	Canada	April 1, 2021	Format Sel...
Robert Allan Ac	Kirstie Alley, Sa	United States	April 1, 2021	Lebar Kolom...
Justin G. Dyck	Emily Wilder, A	Canada	April 4, 2021	Sembunyikan
Liz Tuccillo	Leslie Bibb, Ke	United States	April 10, 2021	Munculkan
Dominique Roc	Anders Daniels	France	April 17, 2021	
Jep Barcelona	Rafinha Alcantara	Spain	April 24, 2021	
Sonia Anderson	Elon Musk, Per	United Kingdom	May 2, 2021	
Becca Gleason	Joey King, Jack	United States	June 3, 2021	
Glenn Miller	Marcus Anderson, Kaiwi Lyman, Andre	United States	July 1, 2021	

SKILLS MICROSOFT EXCEL DAN GOOGLE DATA STUDIO

show_id	type	title	release_yea	rating	duration	duration 1	listed_in
s18	TV Show	Zoo Babies	2008	ALL	1 season		1 Kids, Spec
s21	TV Show	Zoboomafoo	2001	TV-Y	1 season		1 Kids
s22	TV Show	Zo Zo Zombie:	2020	TV-Y7	1 season		1 Action, Ani
s3573	TV Show	_DUPE_ The Ma	2016	TV-14	1 season		1 Documenta
s29	TV Show	Yu-Gi-Oh! Zexa	2015	TV-Y7	3 Seasons		3 Animation,
s30	TV Show	Yu-Gi-Oh! ARC	2018	TV-Y7	3 Seasons		3 Action, Adv
s31	TV Show	Yu-Gi-Oh! 5D's	2011	TV-Y7	2 Seasons		2 Anime, Kid
s32	TV Show	Your Bleeped L	2013	TV-PG	1 season		1 TV Shows
s36	TV Show	Young Lions	2002	16+	1 season		1 Drama, Su
s5978	TV Show	"The Paramedic	2021	ALL	1 season		1 Drama
s3571	TV Show	#IMomSoHard	2019	16+	1 season		1 Arts, Enter
s8853	TV Show	#WASHED	2020	18+	1 season		1 Comedy, D
s3567	TV Show	10 Day Yoga fo	2021	13+	1 season		1 Fitness, Sp
s57	TV Show	Yancy Derringe	1959	13+	1 season		1 Action, We
s62	TV Show	WWII in HD	2009	TV-14	1 season		1 Action
s65	TV Show	Would I Lie to Y	2009	TV-PG	1 season		1 Comedy, S
s66	TV Show	Wotakoi: Love i	2018	TV-NR	1 season		1 TV Shows
s68	TV Show	World War II: V	1994	TV-14	1 season		1 Drama
s72	TV Show	World Food Ch	2014	TV-PG	1 season		1 TV Shows
s73	TV Show	WordWorld	2007	TV-G	1 season		1 Kids
s75	TV Show	WordGirl	2010	TV-G	1 season		1 Kids
s76	TV Show	Wonders of Me	2018	TV-G	1 season		1 Documenta

Analisa Data

Visualisasi Data



STUDY CASE FINAL PROJECT 1

at Mini Course Data Analytics with Microsoft Excel and Google Data Studio by RevoU

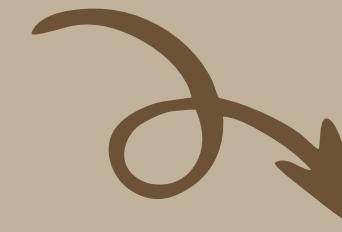
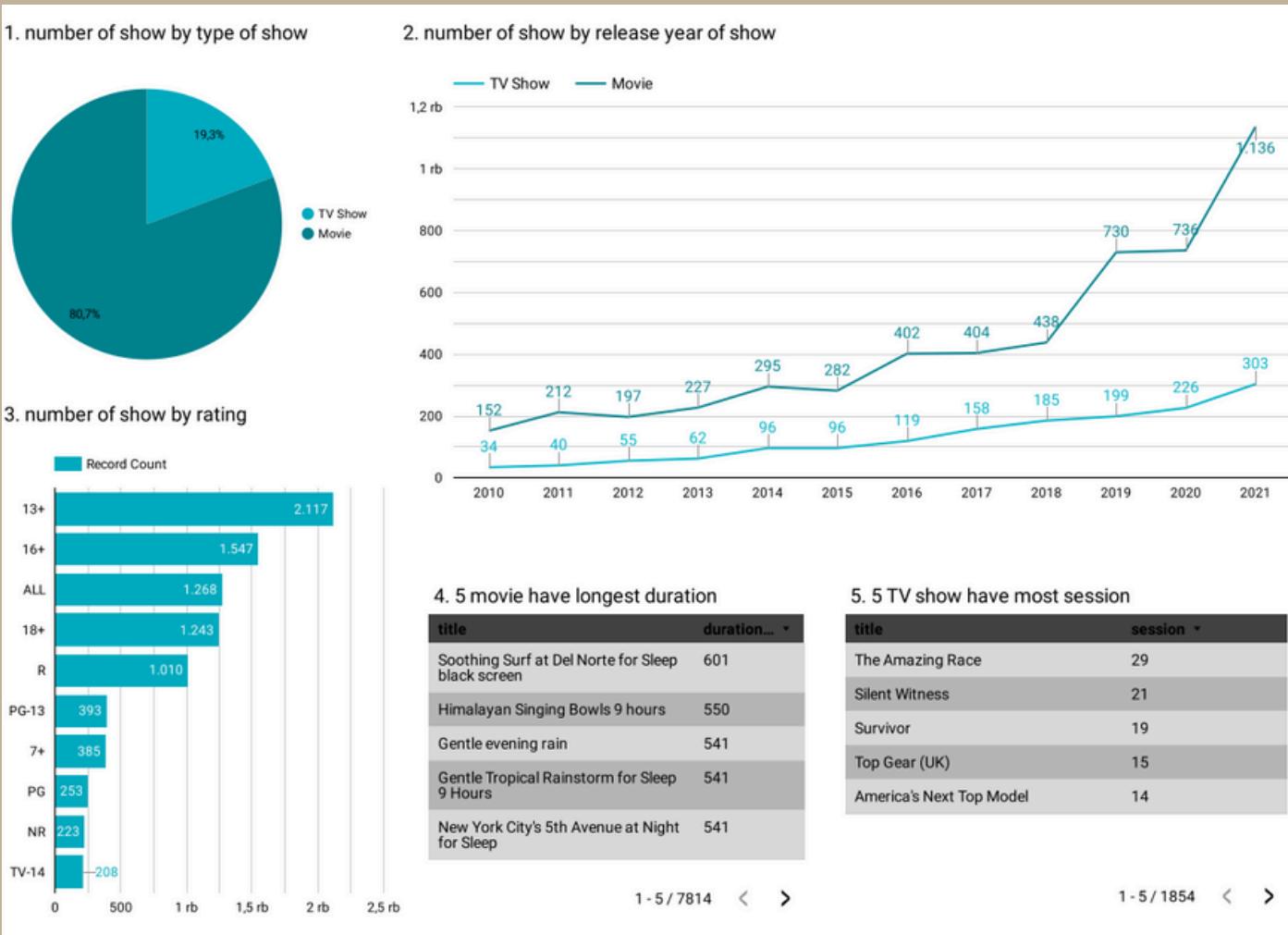
	A	B	C	H	I	J	K	L	M
1	show_id	type	title	release_yea	rating	duration	duration 1	listed_in	description
19	s18	TV Show	Zoo Babies	2008	ALL	1 season		1	Kids, Special In A heart warming
22	s21	TV Show	Zoboomafoo	2001	TV-Y	1 season		1	Kids Chris and Martin
23	s22	TV Show	Zo Zo Zombie:	2020	TV-Y7	1 season		1	Action, Anime, Zo Zo Zombie is
28	s3573	TV Show	_DUPE_ The Ma	2016	TV-14	1 season		1	Documentary, Chronicling iconic
30	s29	TV Show	Yu-Gi-Oh! Zexa	2015	TV-Y7	3 Seasons		3	Animation, Anim It's time to feel t
31	s30	TV Show	Yu-Gi-Oh! ARC	2018	TV-Y7	3 Seasons		3	Action, Adventure Get set to swing
32	s31	TV Show	Yu-Gi-Oh! 5D's	2011	TV-Y7	2 Seasons		2	Anime, Kids With the defeat o
33	s32	TV Show	Your Bleeped U	2013	TV-PG	1 season		1	TV Shows See what an alien
37	s36	TV Show	Young Lions	2002	16+	1 season		1	Drama, Suspense Australian police
38	s5978	TV Show	"The Paramedic	2021	ALL	1 season		1	Drama The tragedy of a
42	s3571	TV Show	#IMomSoHard	2019	16+	1 season		1	Arts, Entertainment Comedians Krist
44	s8853	TV Show	#WASHED	2020	18+	1 season		1	Comedy, Drama 2021 Daytime E
48	s3567	TV Show	10 Day Yoga fo	2021	13+	1 season		1	Fitness, Special If you're looking
58	s57	TV Show	Yancy Derringer	1959	13+	1 season		1	Action, Western Yancy Derringer
63	s62	TV Show	WWII in HD	2009	TV-14	1 season		1	Action The only people
66	s65	TV Show	Would I Lie to Y	2009	TV-PG	1 season		1	Comedy, Speci On the third sea
67	s66	TV Show	Wotakoi: Love i	2018	TV-NR	1 season		1	TV Shows When Narumi, a
69	s68	TV Show	World War II: V	1994	TV-14	1 season		1	Drama At the Tehran ar
73	s72	TV Show	World Food Ch	2014	TV-PG	1 season		1	TV Shows World Food Cha
74	s73	TV Show	WordWorld	2007	TV-G	1 season		1	Kids The loveable Wo
76	s75	TV Show	WordGirl	2010	TV-G	1 season		1	Kids WordGirl and he
77	s76	TV Show	Wonders of Me	2018	TV-G	1 season		1	Documentary, Mexico is rich in



Sumber Data:
Amazon Prime
Movies and TV
Shows.

STUDY CASE FINAL PROJECT 1

at Mini Course Data Analytics with Microsoft Excel and Google Data Studio by RevoU



Visualisasi data dari:

- 1.Jumlah pertunjukan berdasarkan jenis pertunjukan;
- 2.Jumlah pertunjukan berdasarkan tahun rilis pertunjukan;
- 3.Jumlah acara berdasarkan peringkat;
- 4.5 film dengan durasi terpanjang;
- 5.5 acara TV memiliki sesi terbanyak.

CERTIFICATE



SKILLS TABLEAU

transaction_id	transaction_date	transaction_time	store_id	store_location	product_id	transaction_qty	unit_price	Total_Bill
114301	01-06-2023	11.33.29	3	Astoria	45	1	3	
115405	02-06-2023	11.18.24	3	Astoria	45	1	3	
115478	02-06-2023	12.02.45	3	Astoria	45	1	3	
116288	02-06-2023	19.39.47	3	Astoria	45	1	3	
116714	03-06-2023	12.24.57	3	Astoria	45	1	3	
116739	03-06-2023	12.44.17	3	Astoria	45	1	3	
116853	03-06-2023	13.58.14	3	Astoria	45	1	3	
117389	03-06-2023	18.57.29	3	Astoria	45	1	3	
118056	04-06-2023	14.31.43	3	Astoria	45	1	3	
118281	04-06-2023	16.30.33	3	Astoria	45	1	3	
118428	04-06-2023	17.53.30	3	Astoria	45	1	3	
118913	05-06-2023	12.01.03	3	Astoria	45	1	3	
119196	05-06-2023	14.19.59	3	Astoria	45	1	3	
119240	05-06-2023	14.42.16	3	Astoria	45	1	3	
119351	05-06-2023	15.51.35	3	Astoria	45	1	3	
119444	05-06-2023	16.52.40	3	Astoria	45	1	3	
119692	05-06-2023	19.51.09	3	Astoria	45	1	3	
121270	07-06-2023	10.16.40	3	Astoria	45	1	3	

Import data dalam bentuk Microsoft Excel.



STUDY CASE FINAL PROJECT 2

at Mini Course Data Analytics with Tableau by RevoU

The screenshot shows the Tableau Public interface with the following details:

- File Bar:** File, Data, Window, Help.
- Header:** Tableau Desktop Public Edition, Buy Tableau.
- Connections:** Coffee Shop Sales Dataset (Microsoft Excel).
- Sheets:** Coffee Shop (selected), New Union, New Table Extension.
- Data View:** A preview of the "Coffee Shop" sheet showing 20 fields and 149116 rows. The first few rows of data are as follows:

#	Coffee Shop transaction_id	Coffee Shop transaction_date	Coffee Shop transaction_time	#	Coffee Shop store_id	Ab...
87109	07/05/2023 00.00....	30/12/1899 10.01.19		5	Lo...	Coff...
87952	08/05/2023 00.00....	30/12/1899 08.36.15		5	Lo...	sto...
88890	08/05/2023 00.00....	30/12/1899 19.58.16		3	As...	

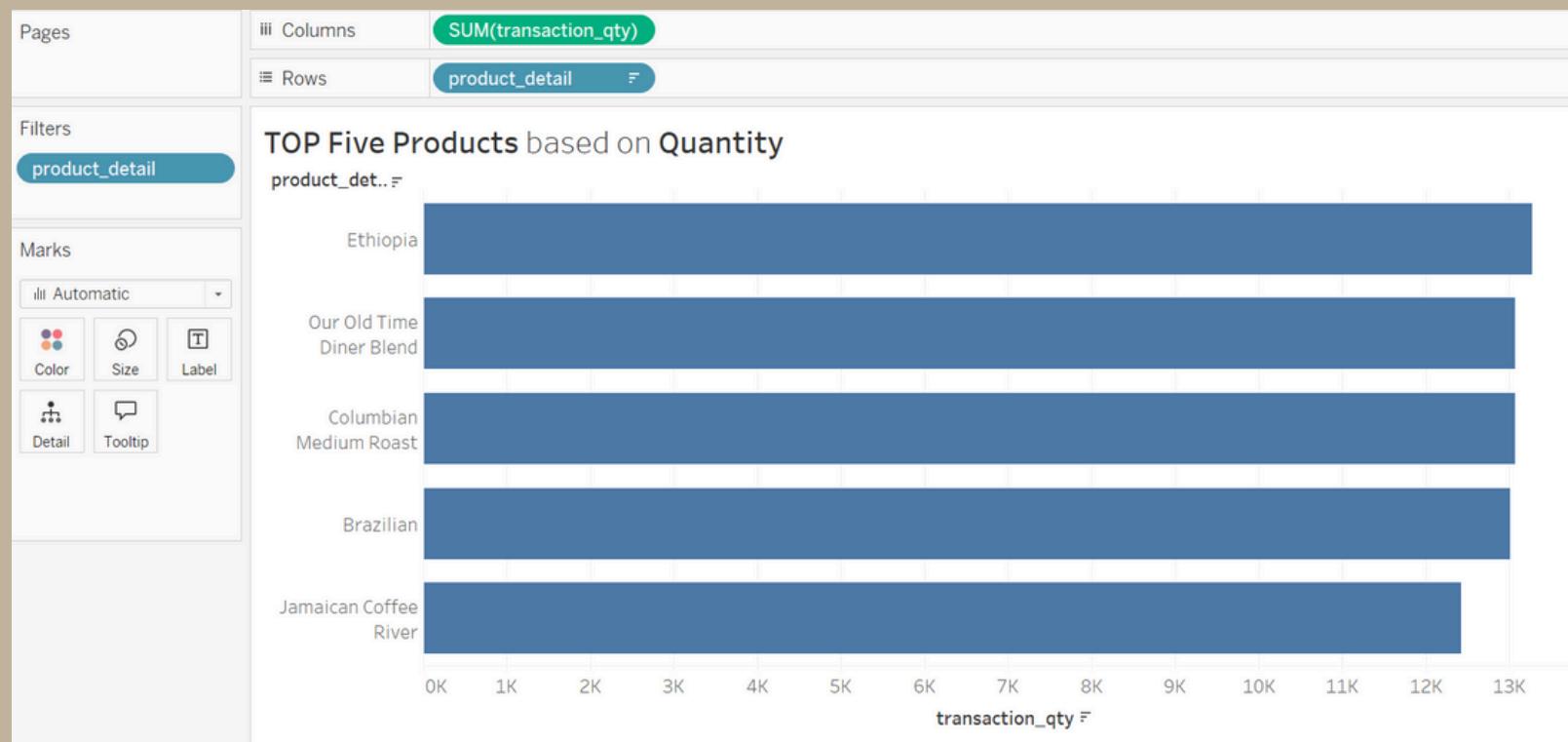
Impor data pada
Tableau Public.



Sumber Data:
Coffee Shop
Sales Dataset.

STUDY CASE FINAL PROJECT 2

at Mini Course Data Analytics with Tableau by RevoU



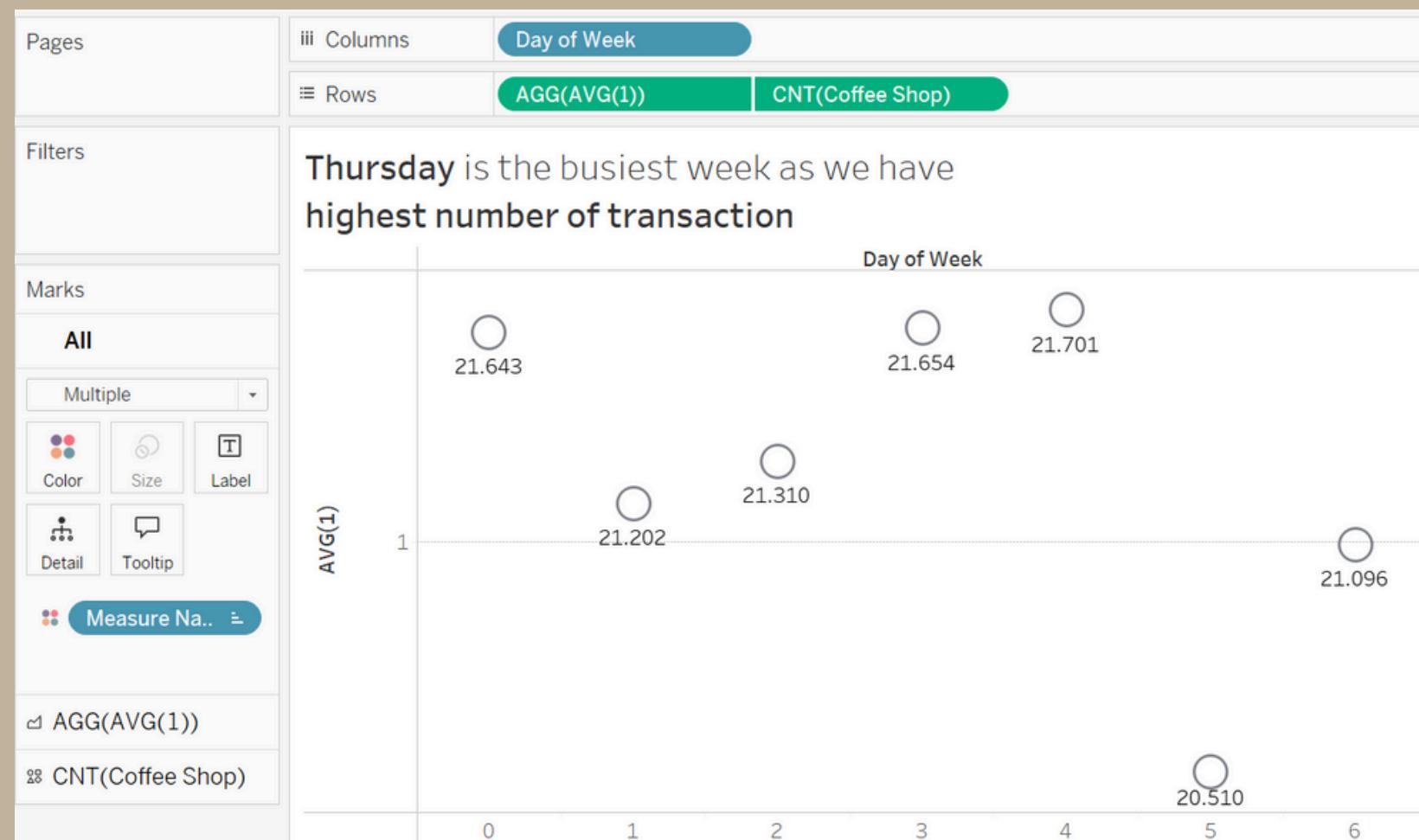
Menginputkan [transaction_qty] dan [product_detail], kemudian mengurutkan [product_detail] berdasarkan SUM[transaction_qty].

STUDY CASE FINAL PROJECT 2

at Mini Course Data Analytics with Tableau by RevoU



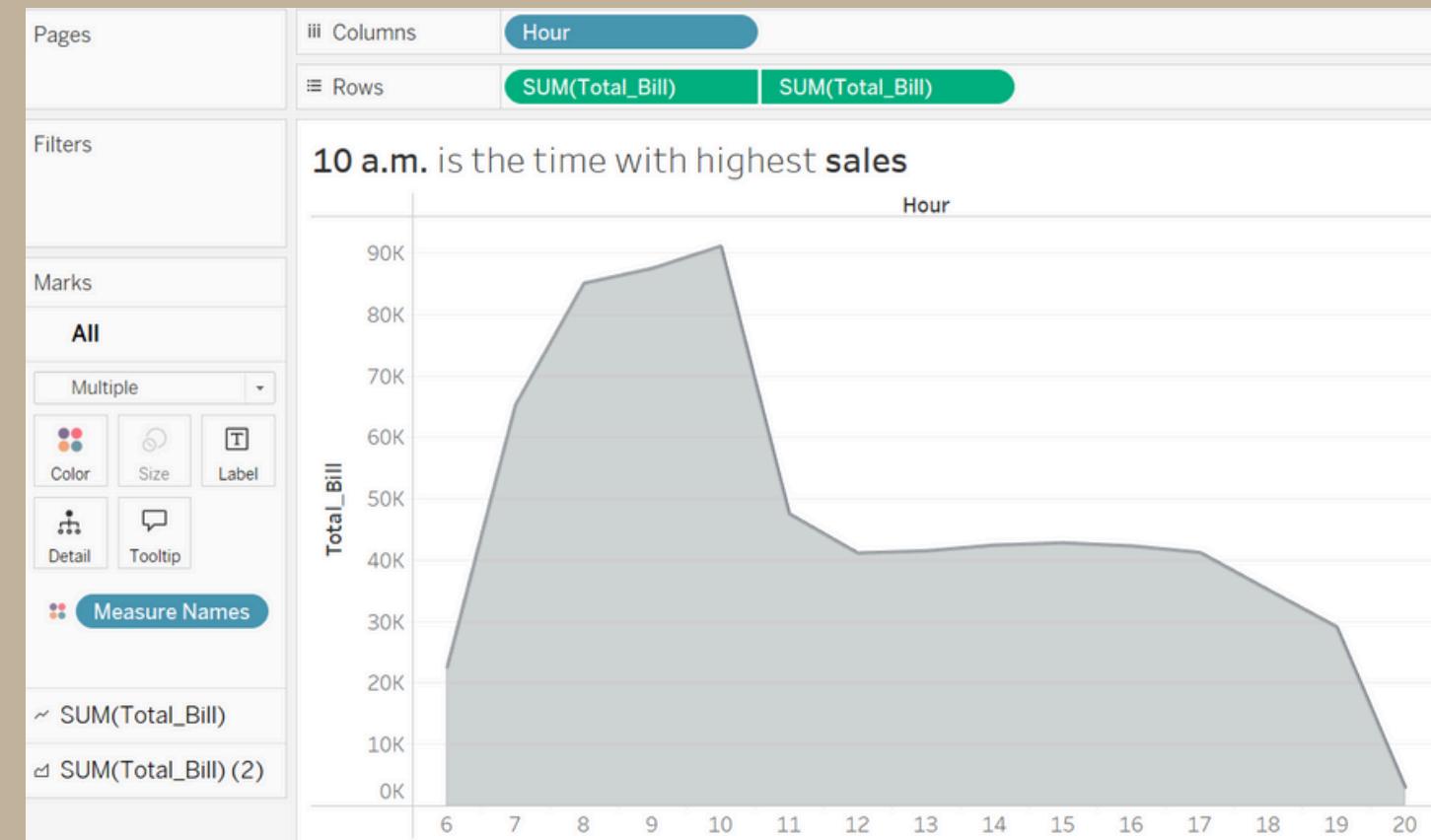
Menginputkan [Day of Week], *Coffee Shop (Count)*, dan mengetik **AVG(1)** pada bagian **Rows**, kemudian menggunakan *dual axis* untuk menggabungkan diagram batang dari **AVG(1)** dengan jumlah transaksi yang sebenarnya.



STUDY CASE FINAL PROJECT 2

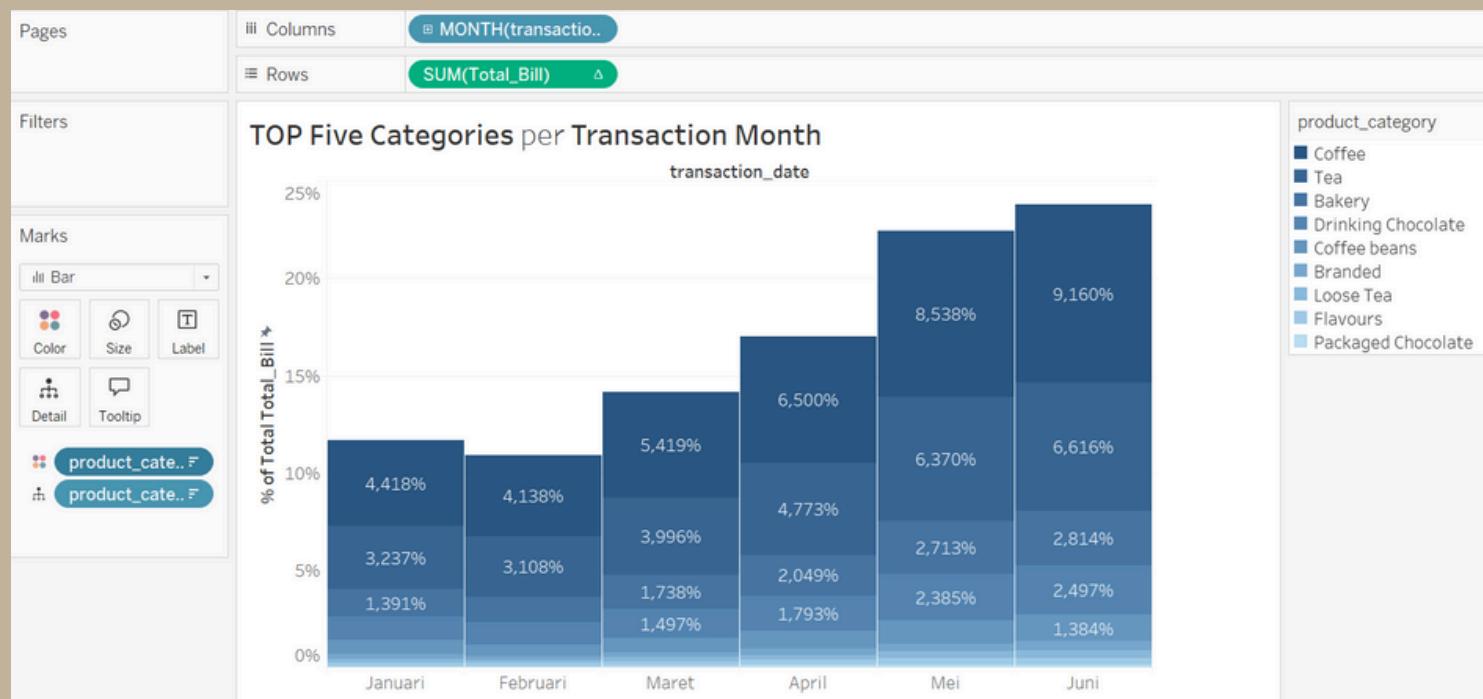
at Mini Course Data Analytics with Tableau by RevoU

Menginputkan [Hour], double SUM[Total_Bill], kemudian menggunakan dual axis untuk menunjukkan SUM[Total_Bill] dalam diagram area dan diagram garis.



STUDY CASE FINAL PROJECT 2

at Mini Course Data Analytics with Tableau by RevoU



Menginputkan [transaction_date], dan pada marks. [transaction_date] berdasarkan bulan, gunakan **Quick Table Calculation** pada SUM[Total_Bill], dan urutkan [product_category] berdasarkan SUM[Total_Bill].



CERTIFICATE



SKILLS DATA ANALYST

Short Class Fundamental Data Analysis



Course Summary

Poin Belajar	Rangkuman
Data Analysis Fundamental	Data analysis menjadi dasar pengumpulan informasi dan validasi bahwa informasi tersebut benar-benar penting. Berdasarkan pengolahan data pada proses ini, informasi yang sudah divalidasi akan menjadi dasar dalam pengambilan keputusan. Manfaat data analysis yakni menghasilkan keputusan yang lebih baik dan memberikan kendali yang lebih luas. Terdapat 4 jenis data analysis yang tepat yakni deskriptif, diagnostik, analisis, dan deskriptif.
Understanding Business Problem	Penentuan akar masalah sangat mempengaruhi ketepatan hasil dalam menjawab persoalan yang dihadapi. Langkah dalam <i>problem solving</i> , yakni memahami latar belakang dan konteks, menentukan <i>stakeholders</i> dan bertanya untuk menambah informasi, serta membuat kerangka berpikir untuk menentukan akar masalah. <i>Problem Solving</i> memiliki 2 tools, yakni <i>5 why, action priority matrix</i> ,
Data Analysis Process	Pendekatan PDCA (Plan, Do, Check, Action) yakni merencanakan dengan mengidentifikasi masalah menggunakan teknik <i>root cause analysis</i> , kemudian menjalankan hal-hal yang telah direncanakan, lalu menjalani pemeriksaan guna menghindari kesalahan terulang, dan melakukan standar baru proses bisnis yang berkelanjutan. Pendekatan ini seperti sebuah lingkaran, setelah sampai pada proses <i>act</i> akan dilakukan evaluasi perencanaan yang telah dibuat, apakah sudah sesuai dengan kebutuhan atau belum sesuai.

#RintisKarirImpian

MySkill

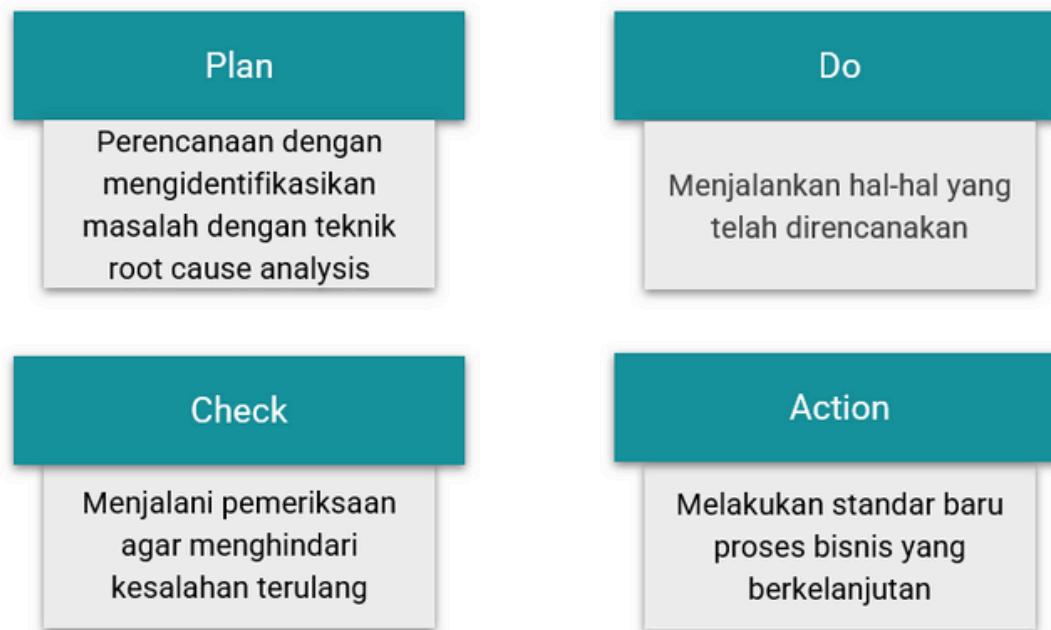
Sumber: MySkill

STUDY CASE FINAL PROJECT 3

at Short Class Data Analysis: Data Analysis Fundamental by MySkill

Framework PDCA

Plan Do Check Action for Data Analyst



Framework Plan, Do,
Check, Action (PDCA)
untuk Data Analyst.

STUDY CASE FINAL PROJECT 3

at Short Class Data Analysis: Data Analysis Fundamental by MySkill

Membuat prediksi
pada test pengujian.



Short Class Fundamental Data Analysis

Task: Implementation of PDCA Framework

Description: Perhatikan case yang tertera di bawah ini. Lalu analisislah menggunakan framework PDCA yang telah kamu pelajari di kelas hari ini. Jawaban dapat kamu tulis pada tabel yang ada di slide berikutnya.



Sebuah perusahaan telekomunikasi yang ada di Indonesia ingin meningkatkan retensi pelanggan. Tentu saja mereka membutuhkan seorang data analyst untuk mengetahui dan memahami pola perilaku pelanggan lainnya yang melakukan retensi.

STUDY CASE FINAL PROJECT 3

at Short Class Data Analysis: Data Analysis Fundamental by MySkill

Implementation of PDCA Framework

PDCA	Analisis Kamu
Plan	Mengidentifikasi masalah dengan menanyakan penyebab permasalahan muncul, menanyakan penyebab pelanggan berpindah, menanyakan penyebab paket internet mahal, menanyakan penyebab jangkauan jaringan buruk dan layanan kurang baik, menanyakan penyebab lambat mengadopsi layanan dan teknologi baru.
Do	Tim analisis data mengumpulkan data penjualan paket internet, data jangkauan jaringan buruk dan layanan yang kurang baik, dan data layanan dan teknologi terbaru. Kemudian, tim mempelajari struktur data, jenis variabel dan kualitas data untuk memahami informasi yang tersedia berdasarkan <i>planning</i> yang sudah dibuat. Kemudian, tim akan menghapus entri yang tidak valid dan melakukan penggabungan data dari berbagai sumber untuk analisis. Dan tim membentuk model prediktif menggunakan metode <i>machine learning</i> dan menganalisis korelasi antara variabel untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap penurunan retensi pelanggan.
Check	Tim analisis data mengevaluasi kinerja model prediktif dengan metrik evaluasi seperti akurasi, presisi, maupun <i>recall</i> . Kemudian, akan dilakukan uji model pada data yang tidak digunakan dalam Pembangunan model guna memastikan kredibilitas dan generalisasi model.
Action	Setelah menganalisis model prediktif, tim analisis data akan merekomendasikan pembaruan layanan dan teknologi baru yang lebih efektif dalam meningkatkan retensi pelanggan kepada tim manajemen perusahaan. Setelah menerapkan rekomendasi, tim analisis data akan mengawasi dampak terhadap kepuasan pelanggan. Jika perlu, tim analisis data mengulangi tahapan pada PDCA guna memperbaiki pengatasan masalah yang akan muncul.



Analisa studi kasus

CERTIFICATE



SKILLS PYTHON

	A	B	C	D	E	F	G	H	I
1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigr	Age	Outcome
2	6	148	72	35	0	33.6	627	50	1
3	1	85	66	29	0	26.6	351	31	0
4	8	183	64	0	0	23.3	672	32	1
5	1	89	66	23	94	28.1	167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	201	30	0
8	3	78	50	32	88	31	248	26	1
9	10	115	0	0	0	35.3	134	29	0
10	2	197	70	45	543	30.5	158	53	1
11	8	125	96	0	0	0	232	54	1
12	4	110	92	0	0	37.6	191	30	0
13	10	168	74	0	0	38	537	34	1
14	10	139	80	0	0	27.1	1.441	57	0
15	1	189	60	23	846	30.1	398	59	1
16	5	166	72	19	175	25.8	587	51	1
17	7	100	0	0	0	30	484	32	1
18	0	118	84	47	230	45.8	551	31	1



Unduh data dalam
Comma Separated
Values (.csv).

Sumber Data:

Diabetes Prediction by MySkill x Lion Parcel

STUDY CASE

FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LIon Parcel

```
[14] data=pd.read_csv('diabetes.csv', sep=',')  
display(data)
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
...
763	10	101	76	48	180	32.9	0.171	63	0
764	2	122	70	27	0	36.8	0.340	27	0
765	5	121	72	23	112	26.2	0.245	30	0
766	1	126	60	0	0	30.1	0.349	47	1
767	1	93	70	31	0	30.4	0.315	23	0



Impor data
pada python.

STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LIon Parcel



Melihat ringkasan
informasi tentang
DataFrame.

```
[13] data.info()

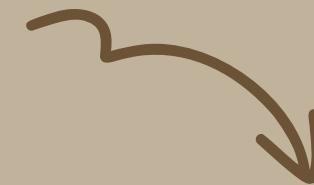
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Pregnancies      768 non-null    int64  
 1   Glucose          768 non-null    int64  
 2   BloodPressure    768 non-null    int64  
 3   SkinThickness    768 non-null    int64  
 4   Insulin          768 non-null    int64  
 5   BMI              768 non-null    float64 
 6   DiabetesPedigreeFunction 768 non-null    float64 
 7   Age              768 non-null    int64  
 8   Outcome          768 non-null    int64  
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
```

STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LIon Parcel

```
[15] data.isnull()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False
...
763	False	False	False	False	False	False	False	False	False
764	False	False	False	False	False	False	False	False	False
765	False	False	False	False	False	False	False	False	False
766	False	False	False	False	False	False	False	False	False
767	False	False	False	False	False	False	False	False	False

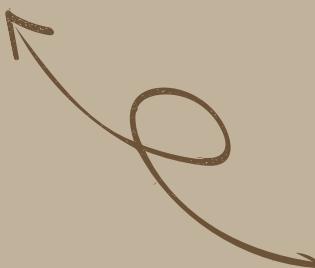


Mendeteksi nilai yang hilang dalam objek DataFrame.

STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LION Parcel

Menghitung jumlah nilai yang hilang dalam setiap kolom pada suatu DataFrame.



```
[16] data.isnull().sum()
```

Pregnancies	0
Glucose	0
BloodPressure	0
SkinThickness	0
Insulin	0
BMI	0
DiabetesPedigreeFunction	0
Age	0
Outcome	0
dtype: int64	

STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LIon Parcel

```
[18] data_x=data.drop('Outcome',axis=1)
display(data_x)
data_y=data['Outcome']
display(data_y)
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age
0	6	148	72	35	0	33.6	0.627	50
1	1	85	66	29	0	26.6	0.351	31
2	8	183	64	0	0	23.3	0.672	32
3	1	89	66	23	94	28.1	0.167	21
4	0	137	40	35	168	43.1	2.288	33
...
763	10	101	76	48	180	32.9	0.171	63
764	2	122	70	27	0	36.8	0.340	27
765	5	121	72	23	112	26.2	0.245	30
766	1	126	60	0	0	30.1	0.349	47
767	1	93	70	31	0	30.4	0.315	23

Menghapus baris atau kolom
dari DataFrame.



STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LION Parcel

Membuat struktur data tabular dua dimensi.



```
[12] new_data=pd.DataFrame({  
    'Pregnancies':[1],  
    'Glucose':[150],  
    'BloodPressure':[72],  
    'SkinThickness':[30],  
    'Insulin':[0],  
    'BMI':[33.5],  
    'DiabetesPedigreeFunction':[0.351],  
    'Age':[50],  
})  
display(new_data)
```

```
new_predict=model.predict(new_data)  
display(new_predict)
```

```
Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age  
0 1 150 72 30 0 33.5 0.351 50  
array([1])
```

STUDY CASE FINAL PROJECT 4

at Short Class Data Science: Data Science Introduction by MySkill x LIon Parcel

```
[23] # Make predictions on the test set
     y_pred = model.predict(x_test)

     # Calculate accuracy (use score method for KNeighborsClassifier)
     accuracy = model.score(x_test, y_test)
     print(accuracy)

     # Calculate confusion matrix (correct spelling)
     confusion_result = confusion_matrix(y_test, y_pred)
     print(confusion_result)

0.7519685039370079
[[146  27]
 [ 36  45]]
```

Membuat prediksi
pada test pengujian.



CERTIFICATE



SKILLS PYTHON

Course Summary

Poin Belajar	Rangkuman
Python Introduction	Bahasa pemrograman yang sering digunakan dalam membangun situs web maupun perangkat lunak, dikembangkan oleh Guido van Rossum dan rilis pada tahun 1991.
Python Syntax	Penggunaan syntax pada python mudah dipelajari oleh pemula, karena memiliki penekanan pada bahasa alami.
Common Python Data Structures	Wadah yang mengelompokkan data berdasarkan jenisnya. Bahasa pemrograman python memiliki 3 data struktur dasar, yakni list, sets, dan tuples.
Conditional Statement	Pengambilan keputusan dalam bahasa pemrograman otomatis dengan pernyataan bersyarat, dengan mengevaluasi kode. Jika terbukti benar, maka program berjalan sesuai kebutuhan.
Looping	Looping terbagi menjadi dua, yakni <i>for</i> dan <i>while</i> . <i>While</i> digunakan ketika pengulangannya tak tentu sehingga tidak dapat diprediksi berapa kali perulangan dieksekusi. Sedangkan <i>for</i> sudah memiliki elemen yang ditentukan, sehingga dapat diprediksi berapa kali perulangan dieksekusi.
Function	Fungsi yang menampung operasi logika dengan output nilai digunakan di baris kode selanjutnya. <i>Function with Return Statement</i> memiliki hasil dan dapat dilihat, sedangkan <i>Function No Return Statement</i> memiliki hasil namun hasilnya tidak terlihat.

#RintisKarirImpian

MySkill

Sumber: MySkill

STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill

```
✓ 0d [98] string_dt = "hello, MySkill Friend! My name is Aulia"  
      print(string_dt)  
  
hello, MySkill Friend! My name is Aulia
```



```
✓ 0d [111] float_dt = 2.23147  
      print(float_dt)  
  
2.23147
```

Data Types in Python

```
✓ 0d [109] int_dt = 64  
      print(int_dt)  
  
64
```



```
✓ 0d [115] is_true = True  
      is_false = False  
      print(is_true)  
  
True
```

STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill

The collage consists of five screenshots from a Python code editor, each demonstrating a different data structure:

- Screenshot 1 (Top Left):** Shows code for creating and manipulating lists. It includes creating an empty list, adding elements, and slicing.
- Screenshot 2 (Top Middle):** Shows code for creating and manipulating dictionaries. It includes adding key-value pairs, changing values, and printing specific items.
- Screenshot 3 (Top Right):** Shows code for creating and manipulating tuples. It includes printing tuple elements and slicing.
- Screenshot 4 (Bottom Left):** Shows code for creating and manipulating sets. It includes union, intersection, difference, and symmetric difference operations.
- Screenshot 5 (Bottom Middle):** Shows examples of set operations: `my_set.difference(my_set_2)` resulting in `{1, 2}`, and `my_set.symmetric_difference(my_set_2)` resulting in `{1, 2, 6}`.



Common Data
Structure

STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill



Conditional
Statement

```
✓ [33] # IF Statement
0d     num = 31
        if num > 0:
            print(num, "is a positive number.")
        print("This statement is true.")

31 is a positive number.
This statement is true.
```

```
✓ [35] # IF Else Statement
0d     num = -53
        if num >= 0:
            print("Positive or Zero")
        else:
            print("Negative number")

Negative number
```

```
✓ [37] # IF, Elif, Else Statement
0d     num = 0
        if num > 0:
            print("Positive number")
        elif num == 0:
            print("Zero")
        else:
            print("Negative number")

Zero
```

```
✓ [38] # IF nested
0d     num = -22
        if num >= 0:
            if num == 0:
                print("zero")
            else:
                print("Positive number")
        else:
            print("Negative number")

Negative number
```

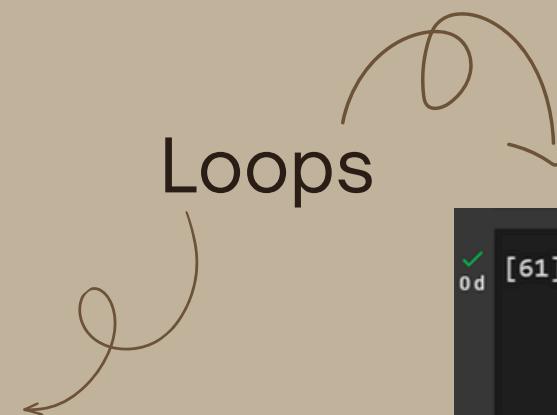
STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill

```
✓ [53] # While Loop
0d     num = 1
0d     odd_nums = []
0d     while num:
0d         if num % 3 != 0:
0d             odd_nums.append(num)
0d         if num >=30:
0d             break
0d         num += 1
0d     print("Odd numbers: ", odd_nums)

Odd numbers: [1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23, 25, 26, 28, 29]

✓ [54] num
0d
0d     30
```



Loops

```
✓ [61] # For Loop
0d     list = [1, 2, 3, 4, 5, 7, 8, 10, 22, 31, 53, 110]
0d     for num in list:
0d         print(num % 5)

1
2
3
4
0
2
3
0
2
1
3
0
```

STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill

```
✓ 0d # Function with Return
    def pow(a):
        return a**2

    print(pow(22))
    power = pow(31)
    print(power)

→ 484
961

✓ [69] power * power

923521

✓ [71] def pow(a):
    print(a**2)

    pow(53)

2809
```

Function



STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill



```
[73] import pandas as pd
      df = pd.read_csv('tips.csv')

[76] df.head(9)

total_bill tip sex smoker day time size
0    16.99  1.01 Female No Sun Dinner 2
1    10.34  1.66 Male   No Sun Dinner 3
2    21.01  3.50 Male   No Sun Dinner 3
3    23.68  3.31 Male   No Sun Dinner 2
4    24.59  3.61 Female No Sun Dinner 4
5    25.29  4.71 Male   No Sun Dinner 4
6     8.77  2.00 Male   No Sun Dinner 2
7    26.88  3.12 Male   No Sun Dinner 4
8    15.04  1.96 Male   No Sun Dinner 2
```

```
[78] df.tail(8)

total_bill tip sex smoker day time size
236     12.60 1.00 Male Yes Sat Dinner 2
237     32.83 1.17 Male Yes Sat Dinner 2
238     35.83 4.67 Female No Sat Dinner 3
239     29.03 5.92 Male No Sat Dinner 3
240     27.18 2.00 Female Yes Sat Dinner 2
241     22.67 2.00 Male Yes Sat Dinner 2
242     17.82 1.75 Male No Sat Dinner 2
243     18.78 3.00 Female No Thur Dinner 2
```

```
[79] df.sample(5)

total_bill tip sex smoker day time size
31      18.35 2.50 Male No Sat Dinner 4
76      17.92 3.08 Male Yes Sat Dinner 2
103     22.42 3.48 Female Yes Sat Dinner 2
90      28.97 3.00 Male Yes Fri Dinner 2
12      15.42 1.57 Male No Sun Dinner 2
```

```
[81] df.info(53)

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 244 entries, 0 to 243
Data columns (total 7 columns):
 #   Column   Non-Null Count  Dtype  
--- 
 0   total_bill    244 non-null   float64
 1   tip          244 non-null   float64
 2   sex          244 non-null   object 
 3   smoker        244 non-null   object 
 4   day           244 non-null   object 
 5   time          244 non-null   object 
 6   size          244 non-null   int64  
dtypes: float64(2), int64(1), object(4)
memory usage: 13.5+ KB
```

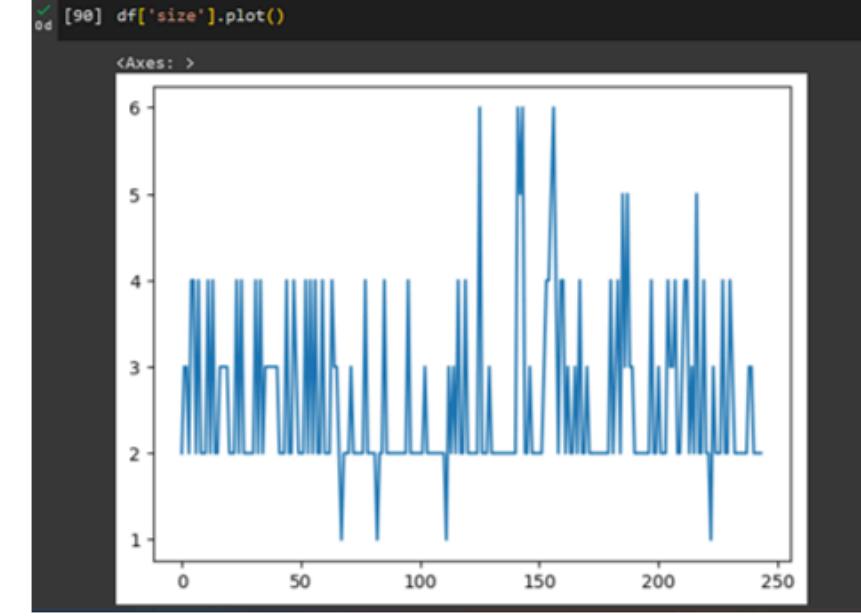
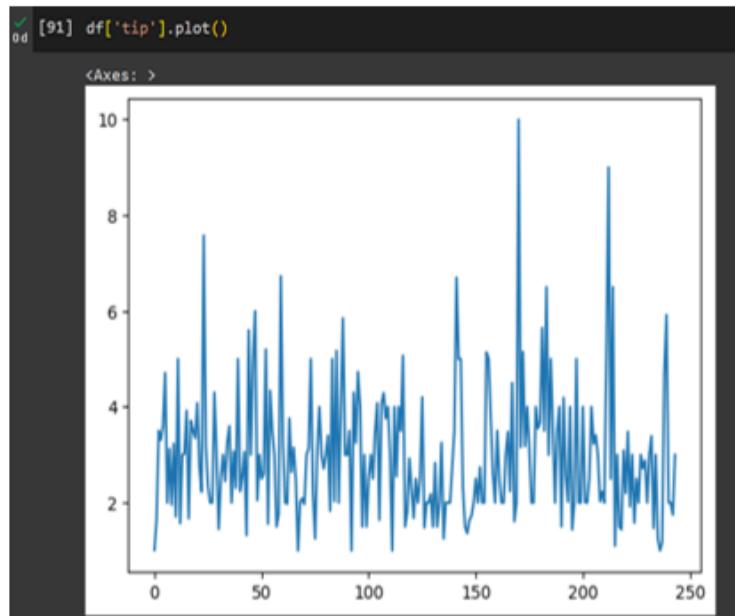
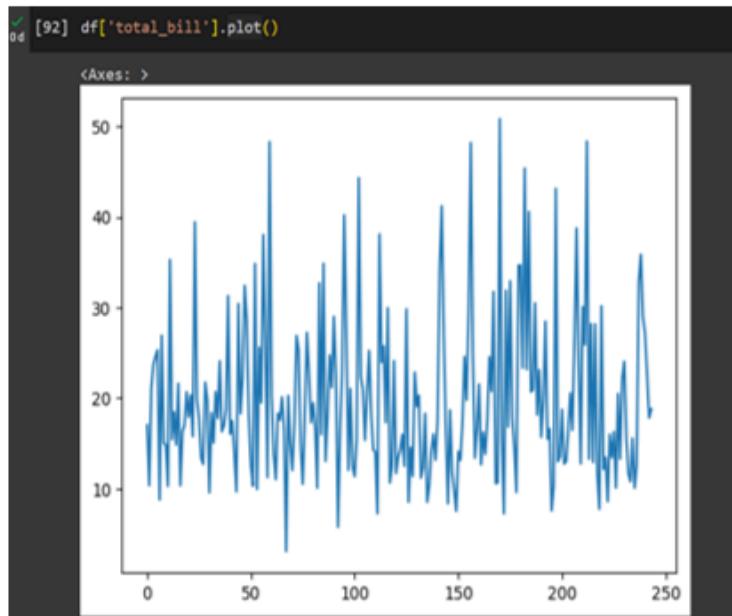
```
[83] df.describe().T

count   mean    std   min    25%   50%   75%   max
total_bill 244.0 19.785943 8.902412 3.07 13.3475 17.795 24.1275 50.81
tip       244.0 2.998279 1.383638 1.00 2.0000 2.900 3.5625 10.00
size      244.0 2.569672 0.951100 1.00 2.0000 2.000 3.0000 6.00
```

Library for
Data Science

STUDY CASE FINAL PROJECT 5

at Short Class Data Analysis: Python Fundamental by MySkill



Library for Data Science (Matplotlib)

CERTIFICATE

MySkill Short Class

April 26th, 2024

This certificate is awarded to:

AULIA TIFFANI RIZKY

For completing short class Data Analysis by MySkill titled:

Python Fundamentals



Angga Fauzan
CEO & Co-Founder MySkill

PT LINIMUDA INSPIRASI NEGERI



143135/DTA/LM/04/2024



CONTACT ME



+62 878 7995 4322



auliatiffani2@gmail.com



<https://www.linkedin.com/in/auliatr22>



@auliaatiffani

