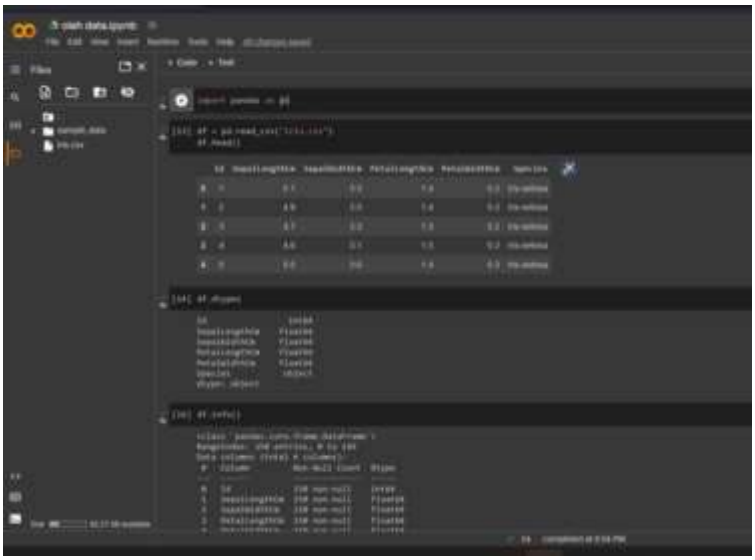
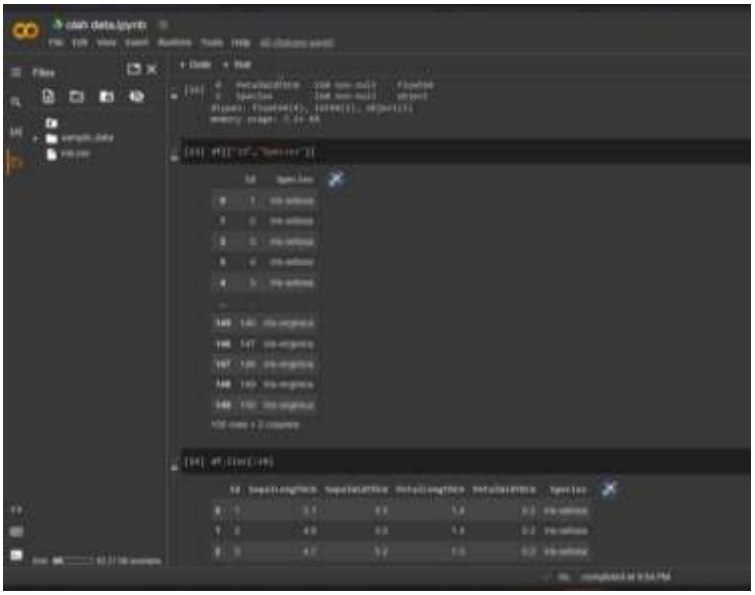
Nama : Novitasari

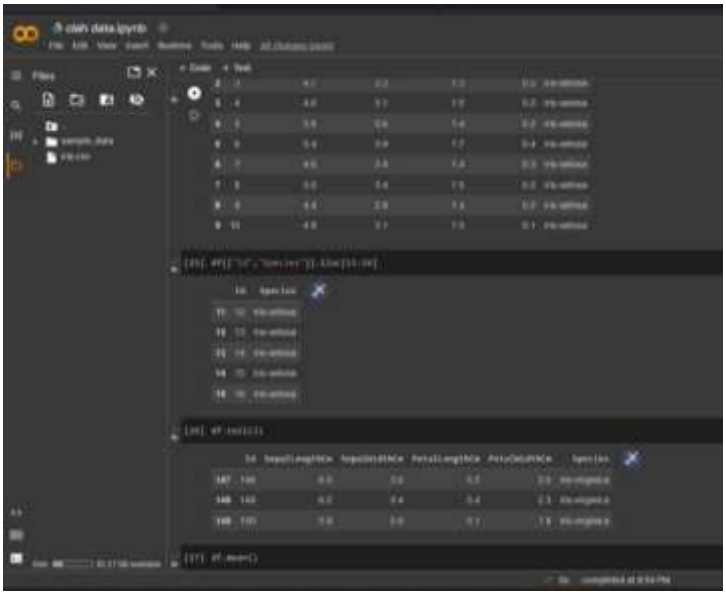
Nim : 20.01.013.012

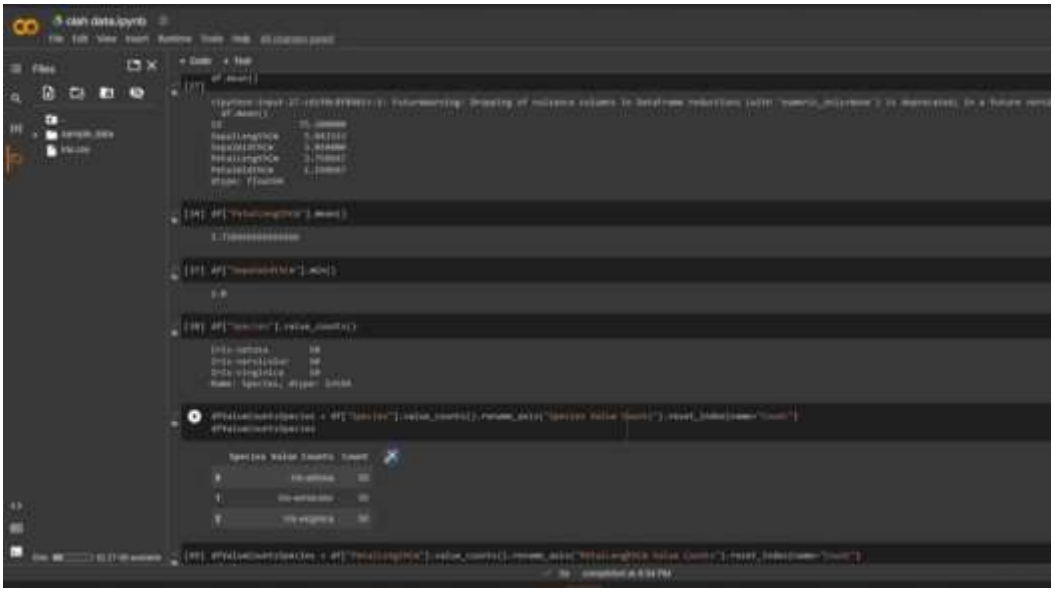
Kelas : Python_D

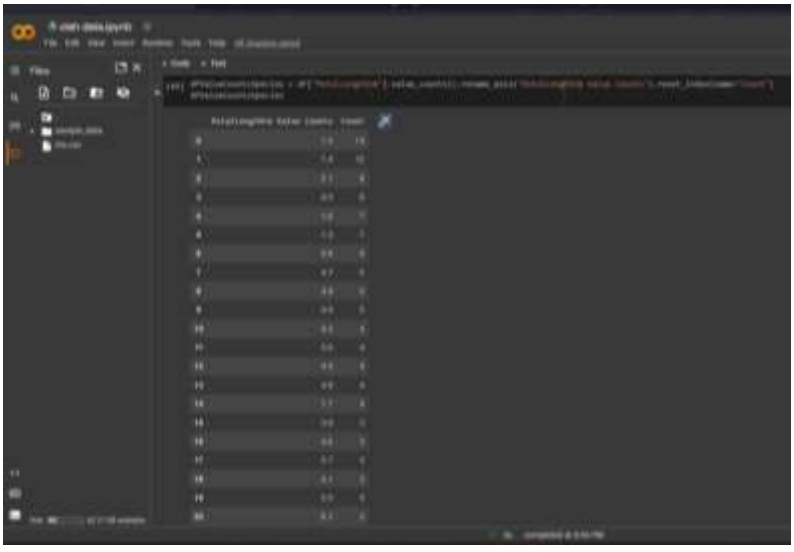
1. Mengolah data iris Database

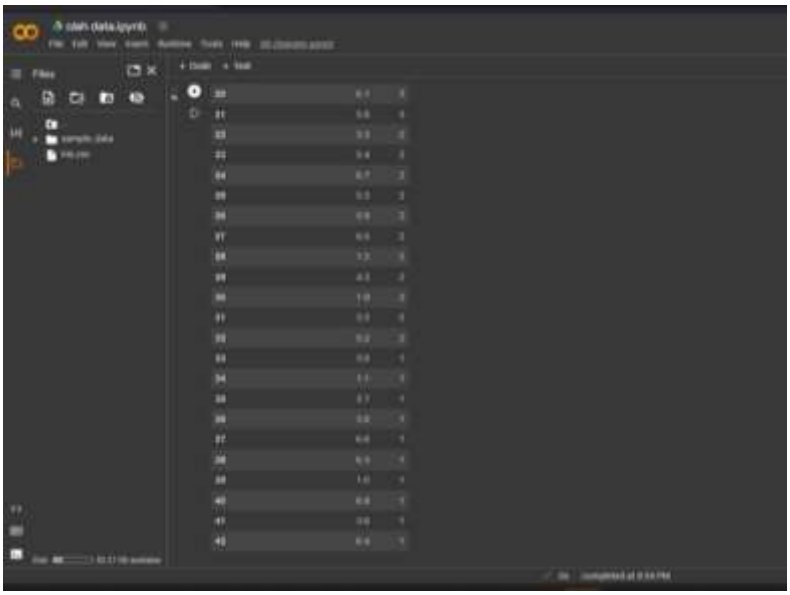




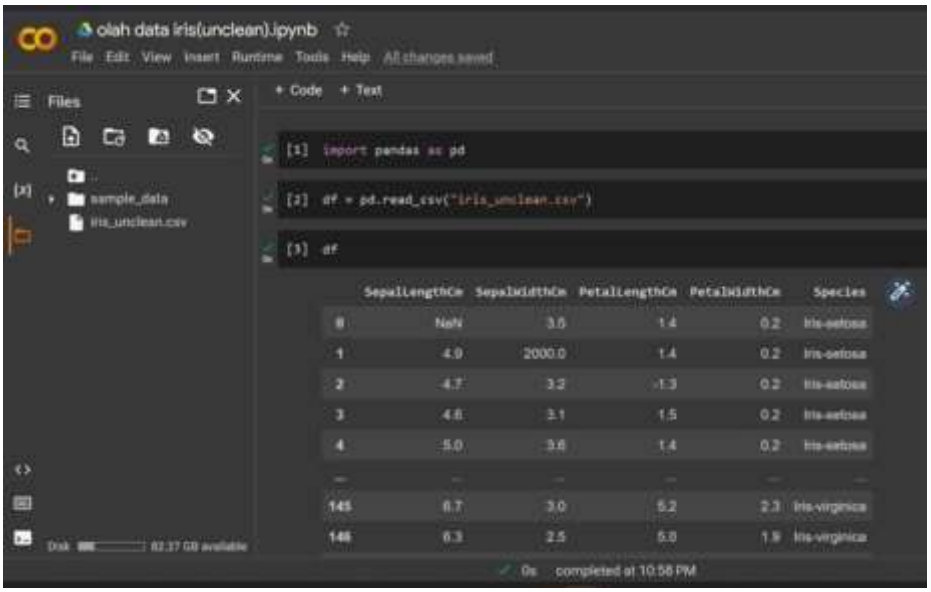


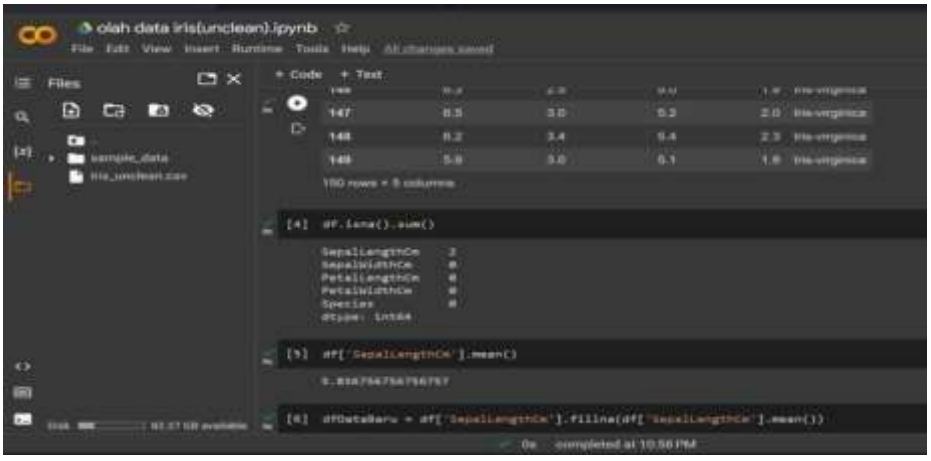


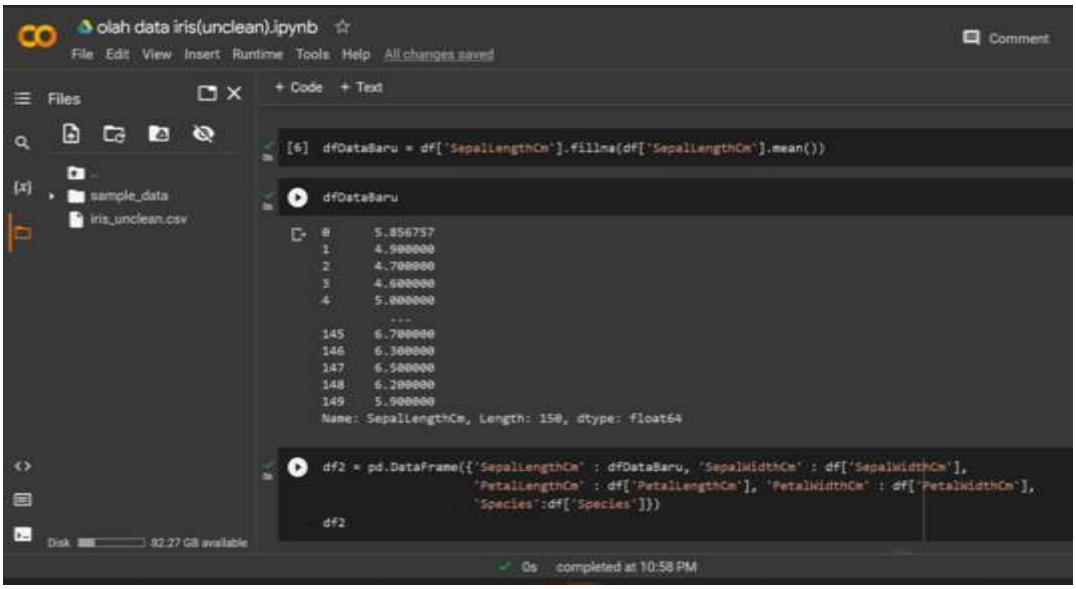


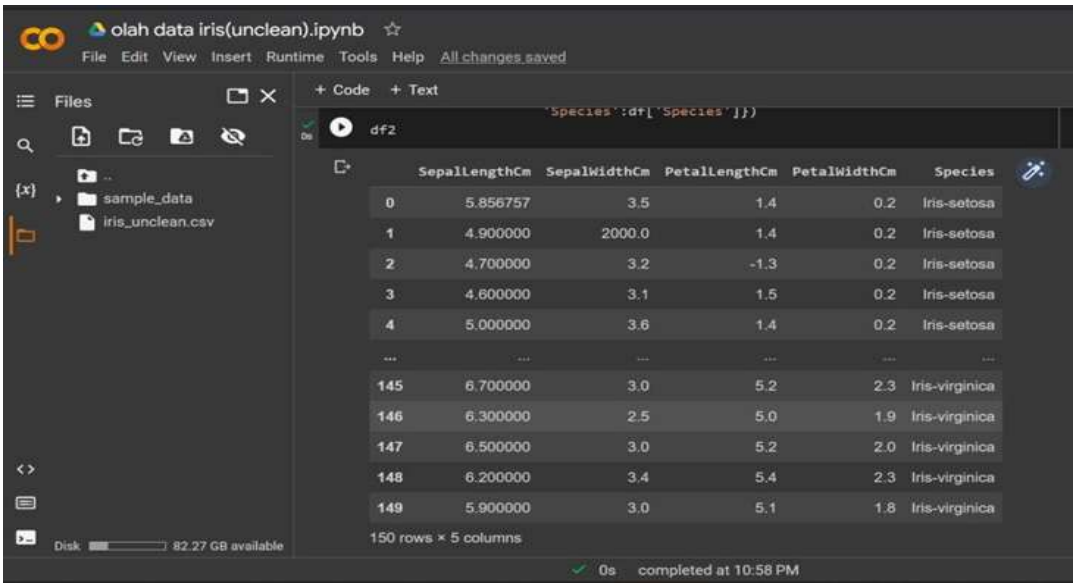


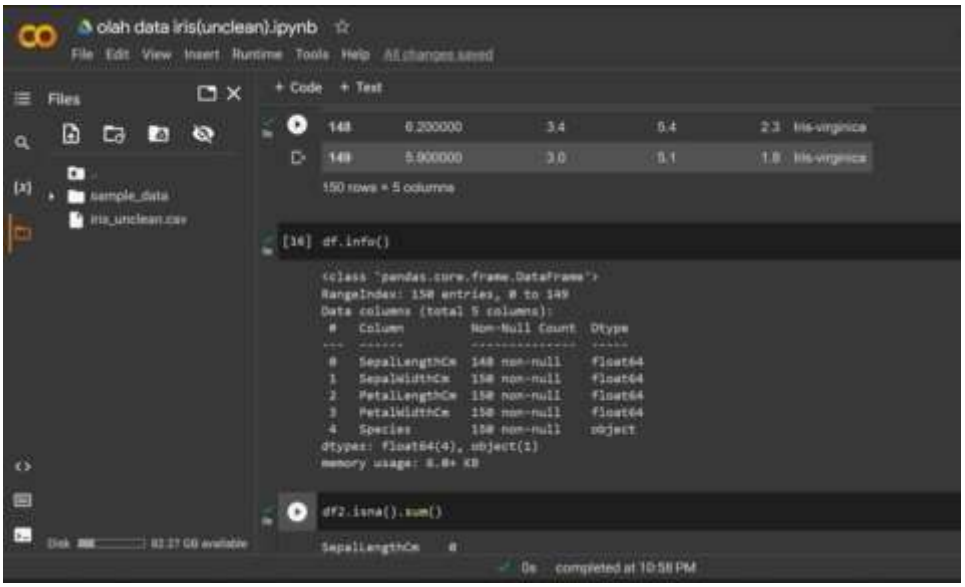
2. Olah Data Iris Unclean

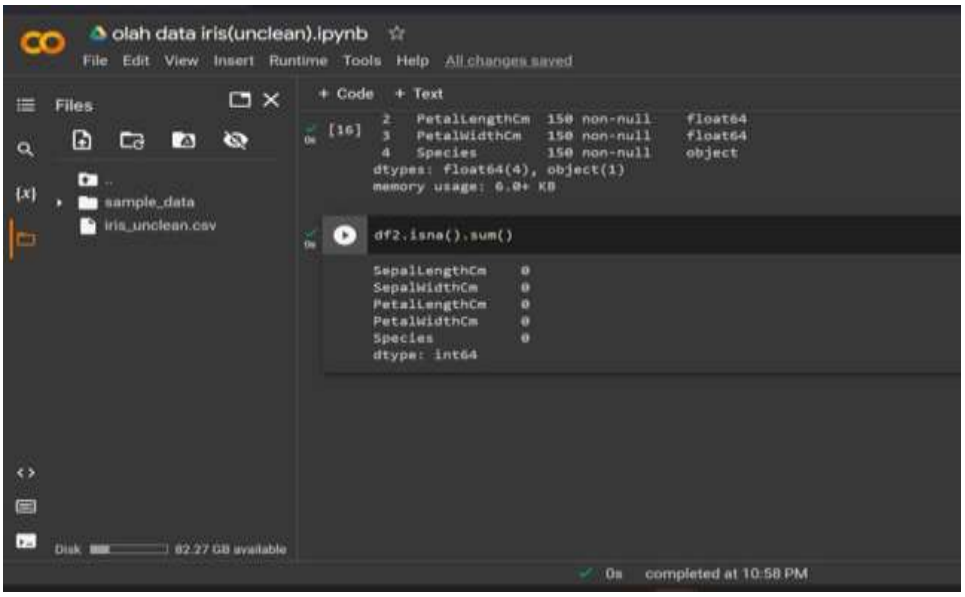




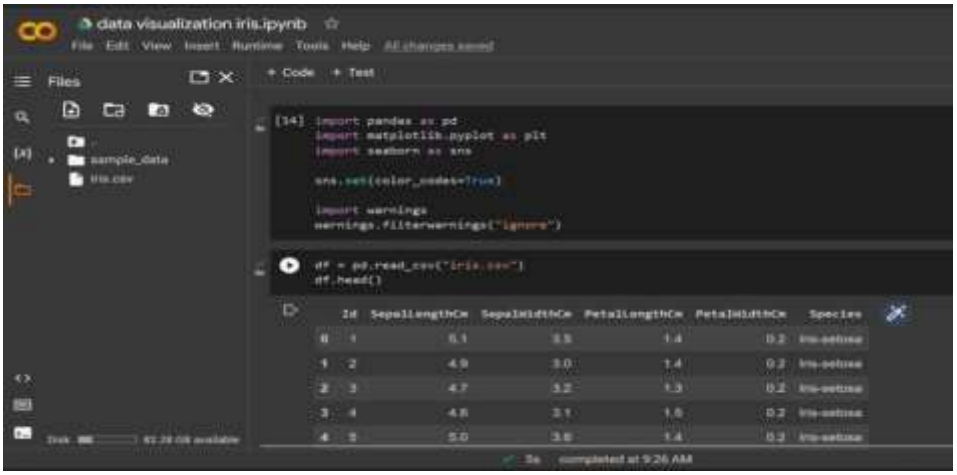

olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
[6] dfDataBaru = df['SepalLengthCm'].fillna(df['SepalLengthCm'].mean())
dfDataBaru
0 5.856757
1 4.900000
2 4.700000
3 4.600000
4 5.000000
145 6.700000
146 6.300000
147 6.500000
148 6.200000
149 5.900000
Name: SepalLengthCm, Length: 150, dtype: float64
df2 = pd.DataFrame({'SepalLengthCm' : dfDataBaru, 'SepalWidthCm' : df['SepalWidthCm'],
'PetalLengthCm' : df['PetalLengthCm'], 'PetalWidthCm' : df['PetalWidthCm'],
'Species':df['Species']})
df2
sample_data
iris_unclean.csv
Disk 82.27 GB available
0s completed at 10:58 PM

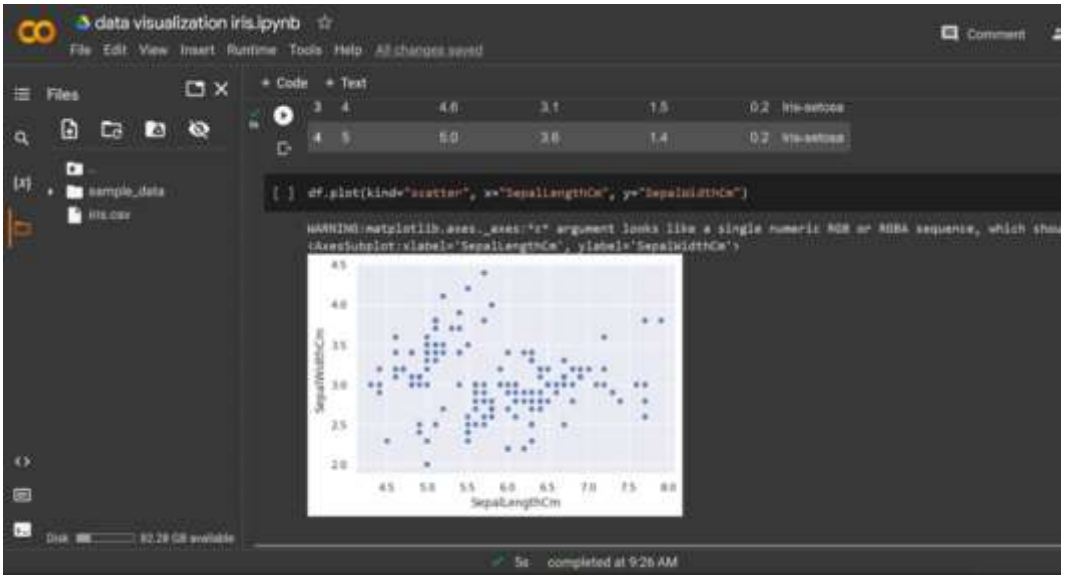

olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
'Species':df['Species']})
df2
SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species
0 5.856757 3.5 1.4 0.2 Iris-setosa
1 4.900000 2000.0 1.4 0.2 Iris-setosa
2 4.700000 3.2 -1.3 0.2 Iris-setosa
3 4.600000 3.1 1.5 0.2 Iris-setosa
4 5.000000 3.6 1.4 0.2 Iris-setosa
... ...
145 6.700000 3.0 5.2 2.3 Iris-virginica
146 6.300000 2.5 5.0 1.9 Iris-virginica
147 6.500000 3.0 5.2 2.0 Iris-virginica
148 6.200000 3.4 5.4 2.3 Iris-virginica
149 5.900000 3.0 5.1 1.8 Iris-virginica
150 rows × 5 columns
sample_data
iris_unclean.csv
Disk 82.27 GB available
0s completed at 10:58 PM

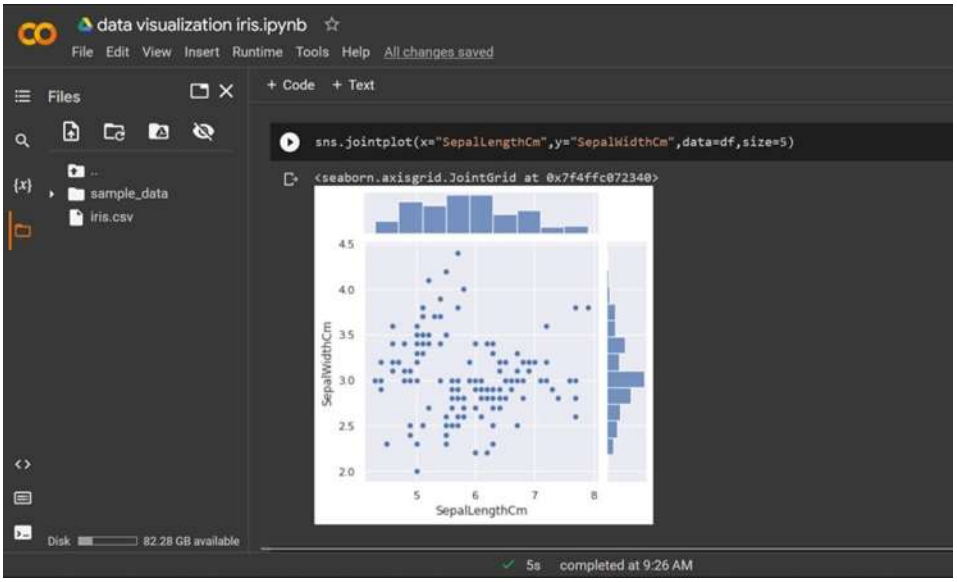

olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
iris_unclean.csv
148 6.200000 3.4 5.4 2.3 Iris-virginica
149 5.900000 3.0 5.1 1.8 Iris-virginica
150 rows × 5 columns
[16] df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
Column Non-Null Count Dtype
0 SepalLengthCm 148 non-null float64
1 SepalWidthCm 150 non-null float64
2 PetalLengthCm 150 non-null float64
3 PetalWidthCm 150 non-null float64
4 Species 150 non-null object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
df2.isna().sum()
SepalLengthCm 0
Disk 82.27 GB available
0s completed at 10:58 PM

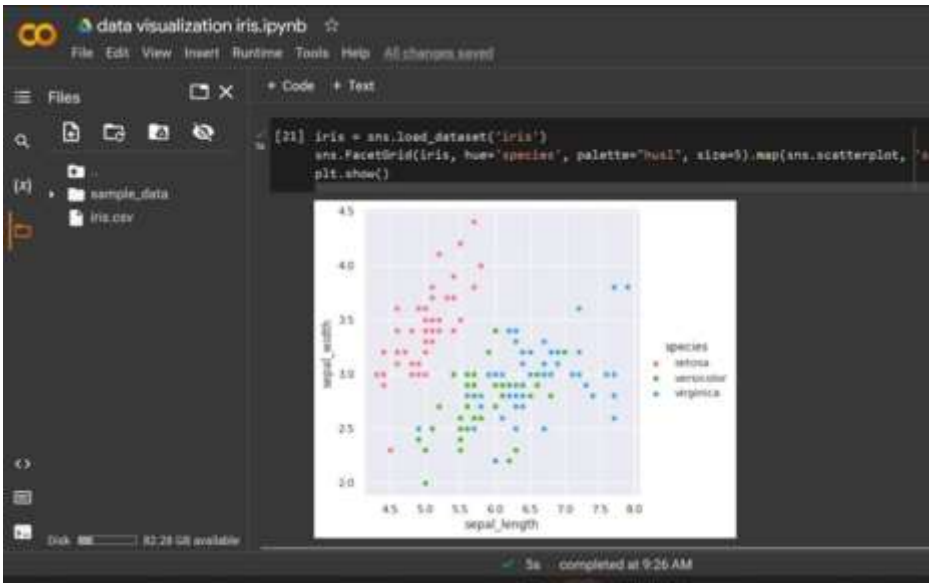

olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
iris_unclean.csv
[16] 2 PetalLengthCm 150 non-null float64
3 PetalWidthCm 150 non-null float64
4 Species 150 non-null object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
df2.isna().sum()
SepalLengthCm 0
SepalWidthCm 0
PetalLengthCm 0
PetalWidthCm 0
Species 0
dtype: int64
Disk 82.27 GB available
0s completed at 10:58 PM

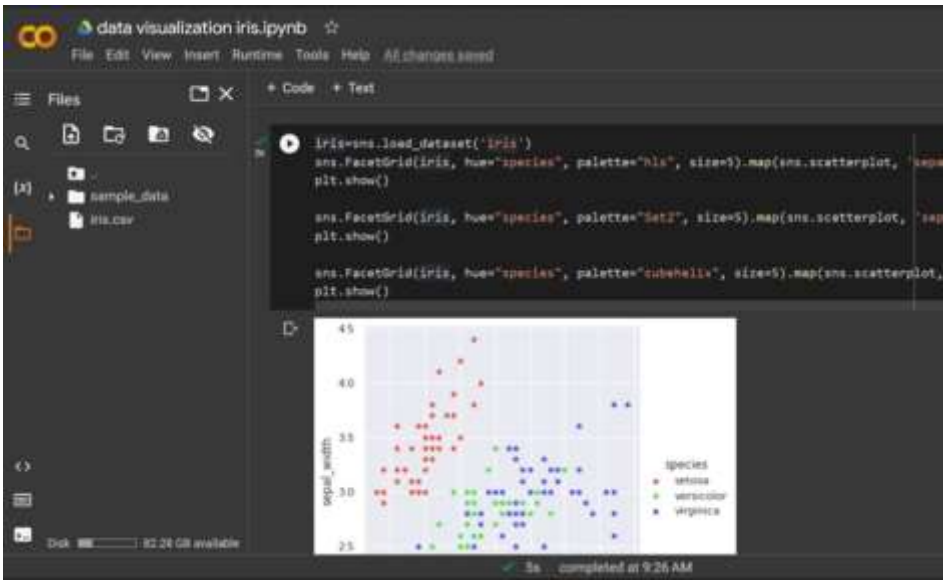
3. Data Visualization Iris





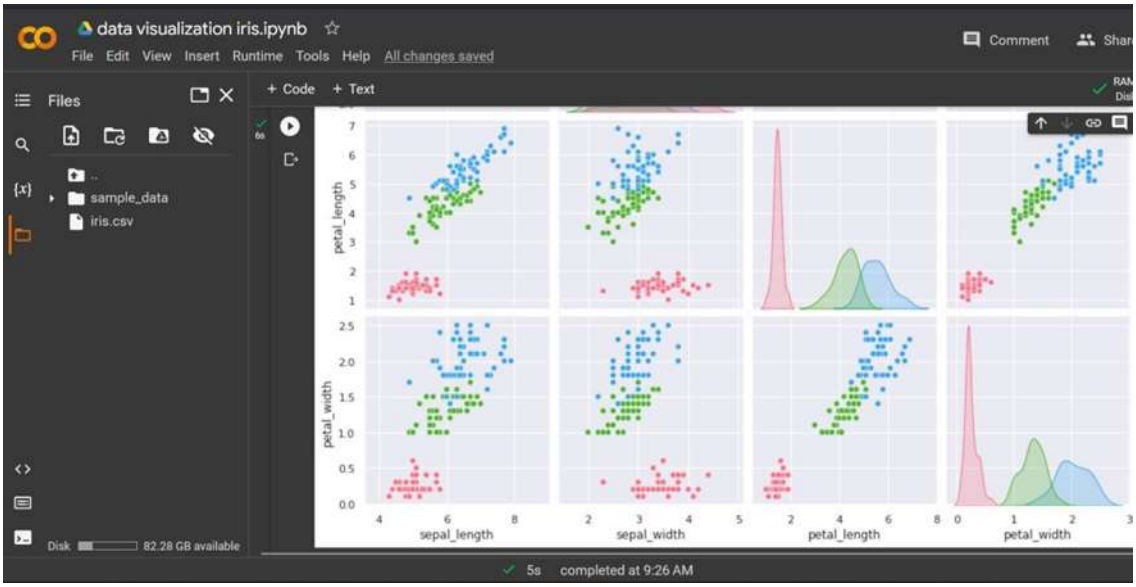



data visualization iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
iris.csv
[21] iris = sns.load_dataset('iris')
sns.FacetGrid(iris, hue='species', palette="husl", size=5).map(sns.scatterplot,
plt.show()
sepal_width
sepal_length
species
setosa
versicolor
virginica
Disk 82.28 GB available
5s completed at 9:26 AM

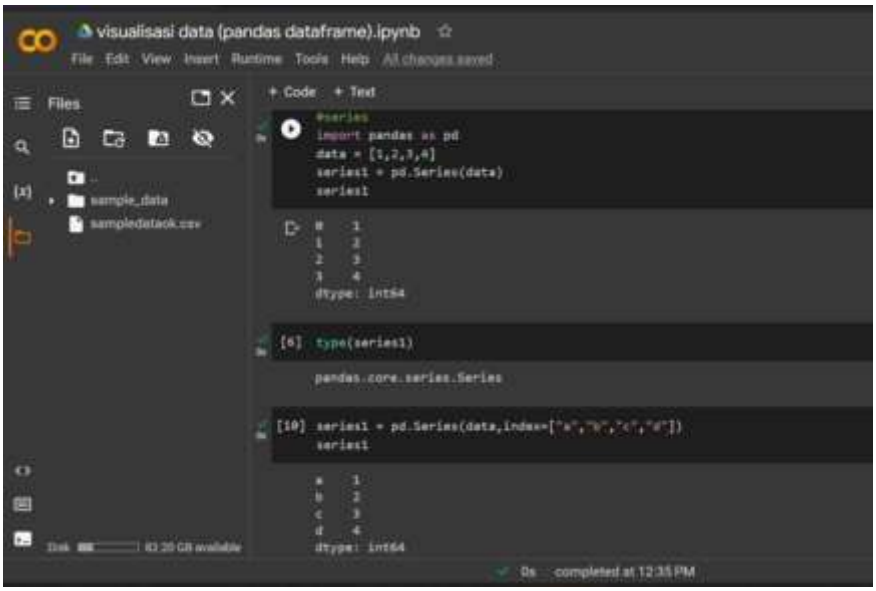

data visualization iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
iris.csv
iris=sns.load_dataset('iris')
sns.FacetGrid(iris, hue="species", palette="hls", size=5).map(sns.scatterplot,
plt.show()
sns.FacetGrid(iris, hue="species", palette="Set2", size=5).map(sns.scatterplot,
plt.show()
sns.FacetGrid(iris, hue="species", palette="cubehelix", size=5).map(sns.scatterplot,
plt.show()
sepal_width
species
setosa
versicolor
virginica
Disk 82.28 GB available
5s completed at 9:26 AM


data visualization iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
iris.csv
sepal_length
sepal_width
species
setosa
versicolor
virginica
Disk 82.28 GB available
5s completed at 9:26 AM

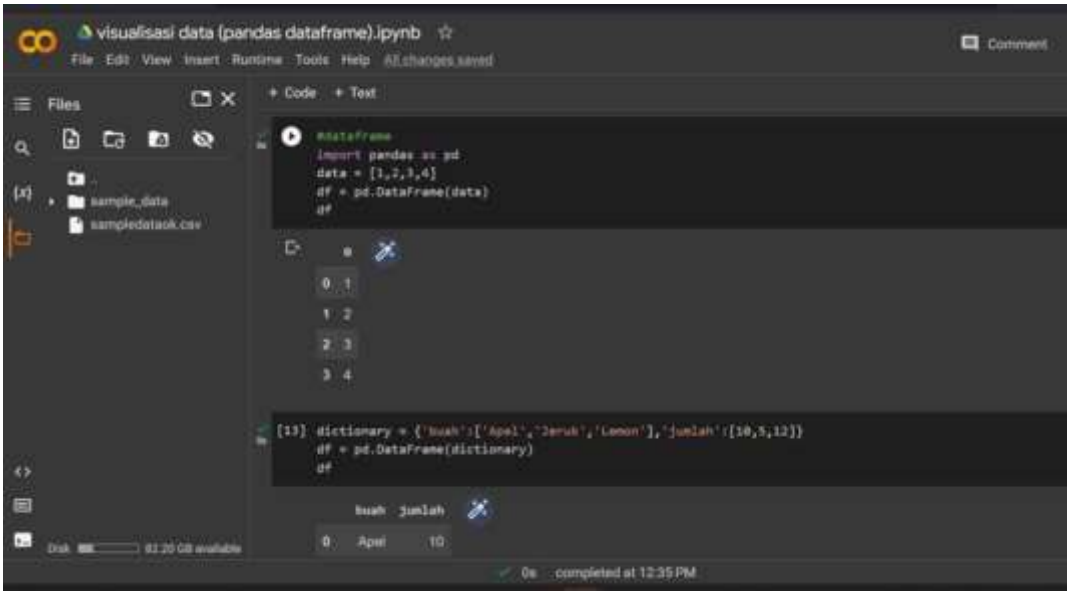

data visualization iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
iris.csv
iris=sns.load_dataset('iris')
sns.pairplot(iris, hue="species", palette="husl", size=3)
<seaborn.axisgrid.PairGrid at 0x7fc8f4f78ac0>
sepal_length
sepal_width
Disk 82.28 GB available
5s completed at 9:26 AM

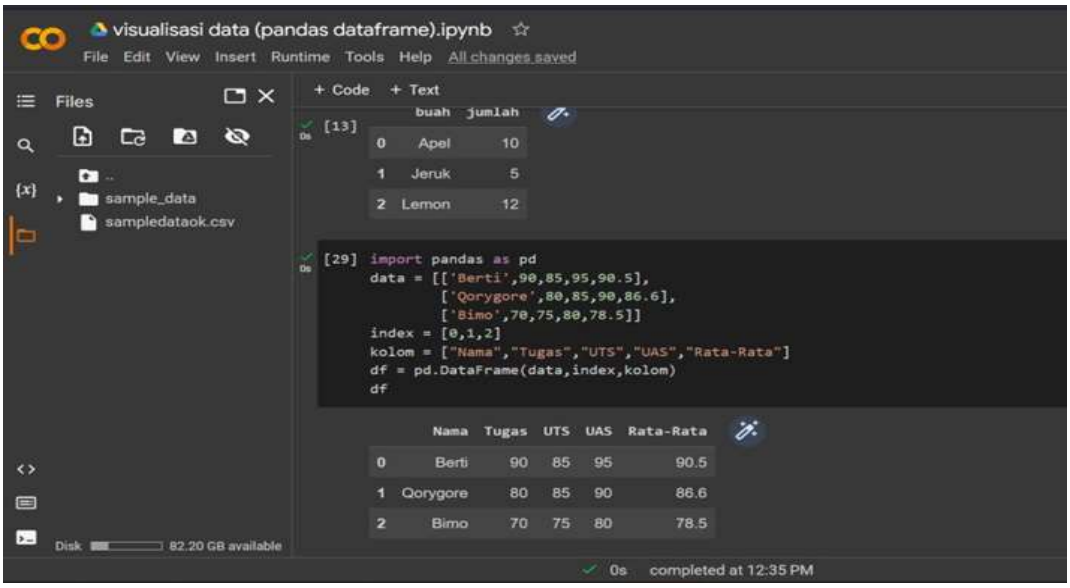


4. visualisasi data (pandas dataframe)
    - Series



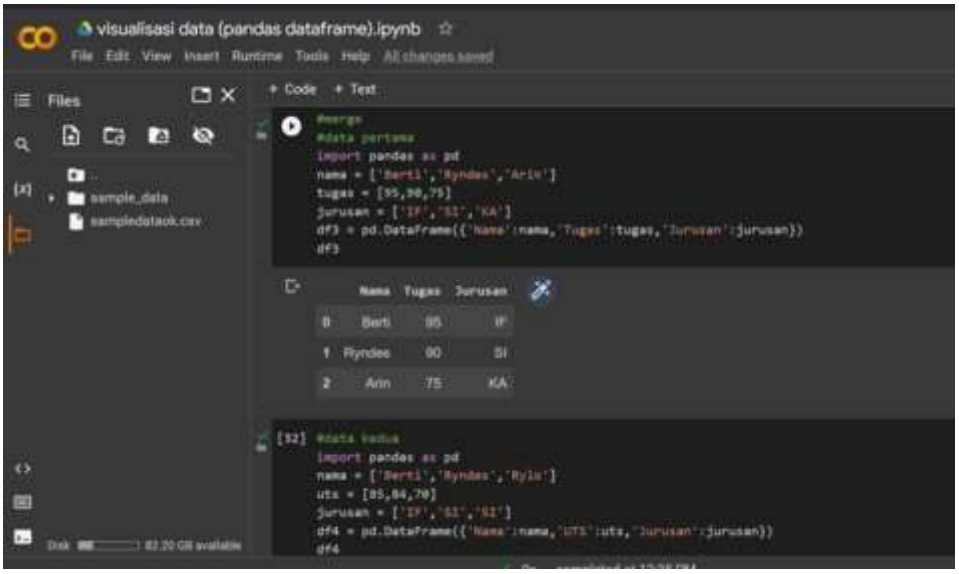
- DataFrame

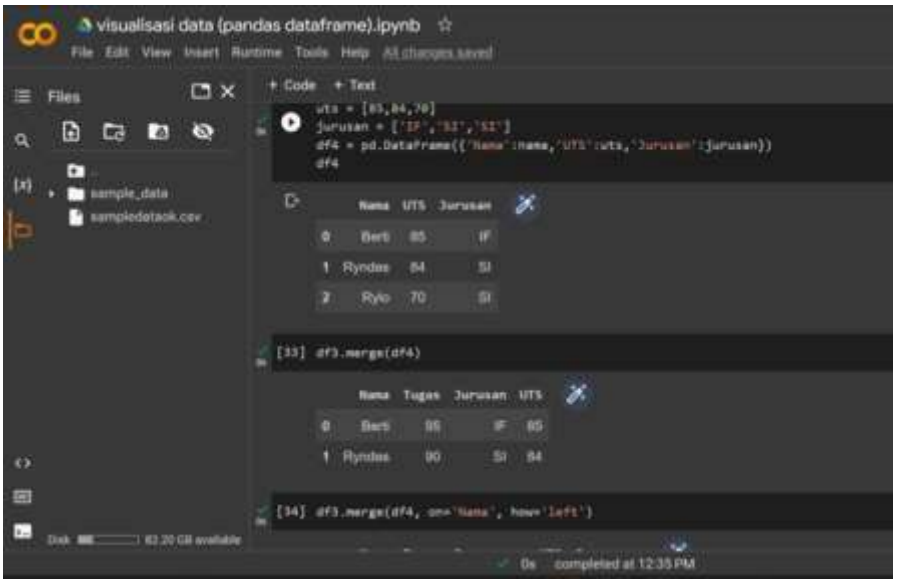

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
sample_data
sampledataok.csv
dataframe
import pandas as pd
data = [1,2,3,4]
df = pd.DataFrame(data)
df
0
0 1
1 2
2 3
3 4
[13] dictionary = {'buah':['Apel','Jeruk','Lemon'],'jumlah':[10,5,12]}
df = pd.DataFrame(dictionary)
df
buah jumlah
0 Apel 10
Disk 82.20 GB available
0s completed at 12:35 PM

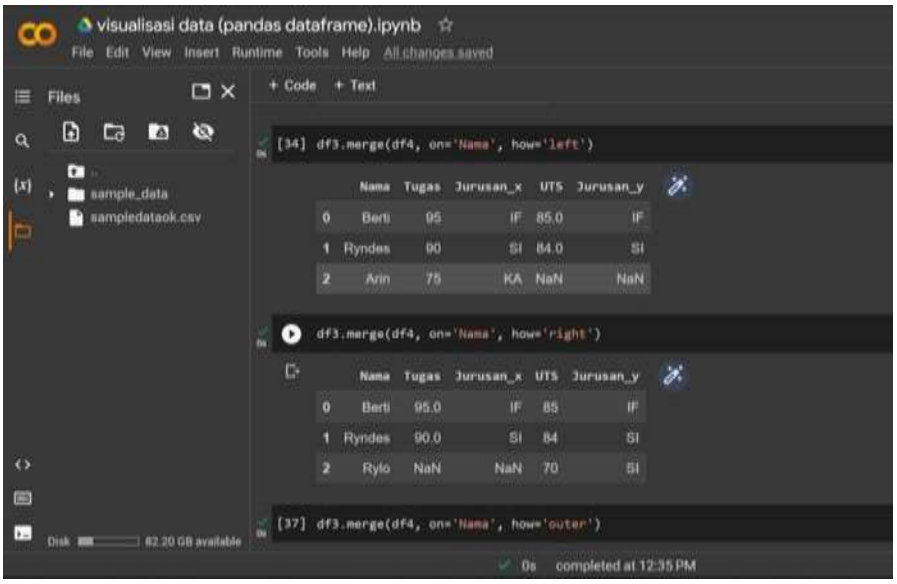

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
[13]
buah jumlah
0 Apel 10
1 Jeruk 5
2 Lemon 12
[29] import pandas as pd
data = [['Berti',90,85,95,90.5],
['Qorygore',80,85,90,86.6],
['Bimo',70,75,80,78.5]]
index = [0,1,2]
kolom = ["Nama","Tugas","UTS","UAS","Rata-Rata"]
df = pd.DataFrame(data,index,kolom)
df
Nama Tugas UTS UAS Rata-Rata
0 Berti 90 85 95 90.5
1 Qorygore 80 85 90 86.6
2 Bimo 70 75 80 78.5
Disk 82.20 GB available
0s completed at 12:35 PM

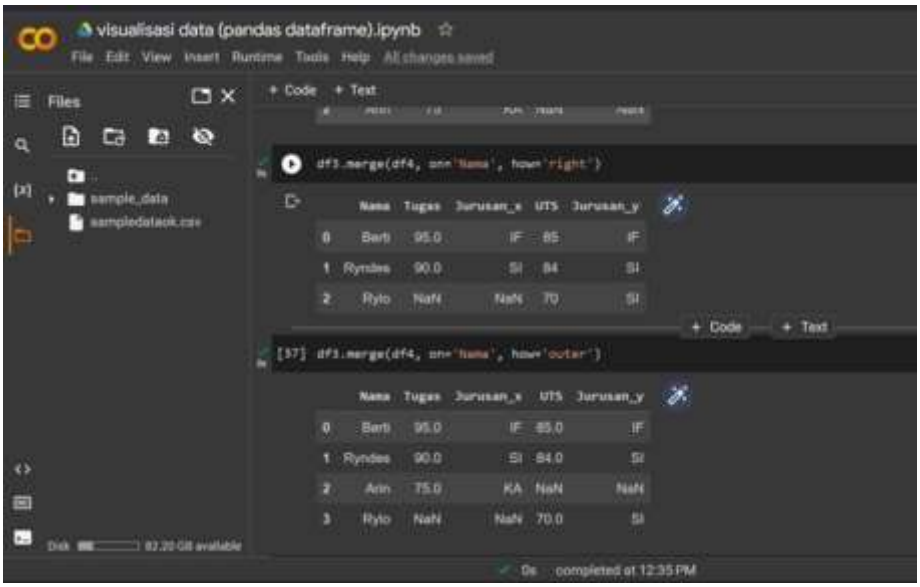

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
sample_data
sampledataok.csv
[29]
Nama Tugas UTS UAS Rata-Rata
0 Berti 90 85 95 90.5
1 Qorygore 80 85 90 86.6
2 Bimo 70 75 80 78.5
[30] import pandas as pd
nama = ['Berti','Qorygore','Bimo']
tugas = [90,80,70]
uts = [85,85,75]
uas = [95,90,80]
ratarata = [90.5,86.6,78.5]
df2 = pd.DataFrame({'Nama':nama, 'Tugas':tugas, 'UTS':uts, 'UAS':uas, 'Rata-Rata':ratarata})
df2
Nama Tugas UTS UAS Rata-Rata
0 Berti 90 85 95 90.5
1 Qorygore 80 85 90 86.6
2 Bimo 70 75 80 78.5
Disk 82.20 GB available
0s completed at 12:35 PM

- Merge

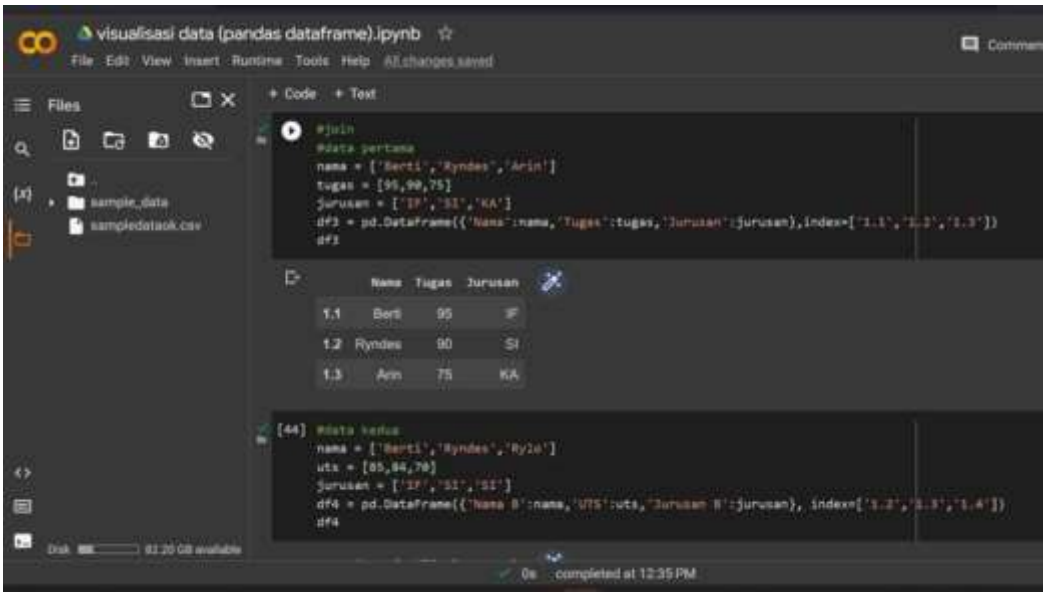

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
#merge
#data pertama
import pandas as pd
nama = ['Berti','Ryndes','Arin']
tugas = [95,90,75]
jurusan = ['IF','SI','KA']
df3 = pd.DataFrame({'Nama':nama,'Tugas':tugas,'Jurusan':jurusan})
df3
Nama Tugas Jurusan
0 Berti 95 IF
1 Ryndes 90 SI
2 Arin 75 KA
[32] #data kedua
import pandas as pd
nama = ['Berti','Ryndes','Ryla']
uts = [85,84,70]
jurusan = ['IF','SI','SI']
df4 = pd.DataFrame({'Nama':nama,'UTS':uts,'Jurusan':jurusan})
df4
Disk 82.20 GB available

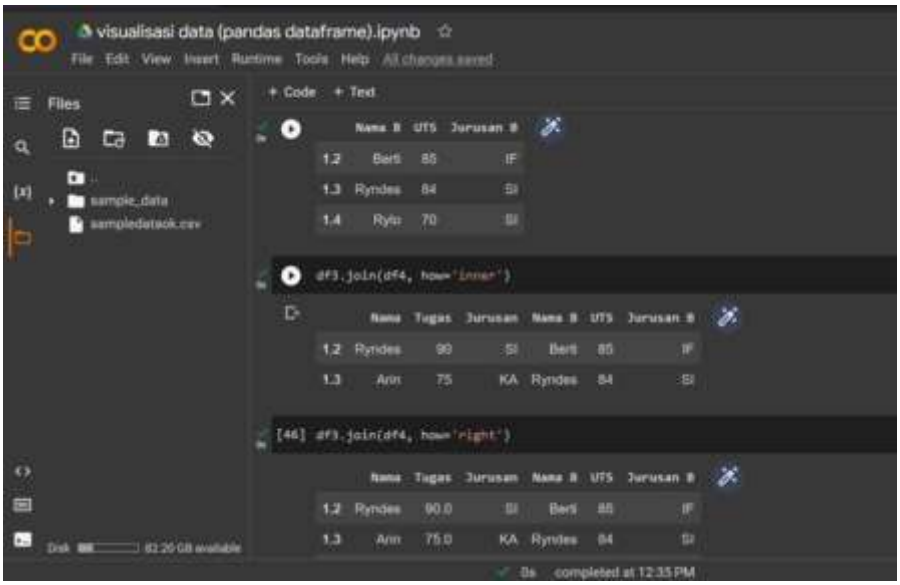

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
uts = [85,84,70]
jurusan = ['IF','SI','SI']
df4 = pd.DataFrame({'Nama':nama,'UTS':uts,'Jurusan':jurusan})
df4
Nama UTS Jurusan
0 Berti 85 IF
1 Ryndes 84 SI
2 Rylo 70 SI
[33] df3.merge(df4)
Nama Tugas Jurusan UTS
0 Berti 95 IF 85
1 Ryndes 90 SI 84
[34] df3.merge(df4, on='Nama', how='left')
Disk 82.20 GB available
0s completed at 12:35 PM

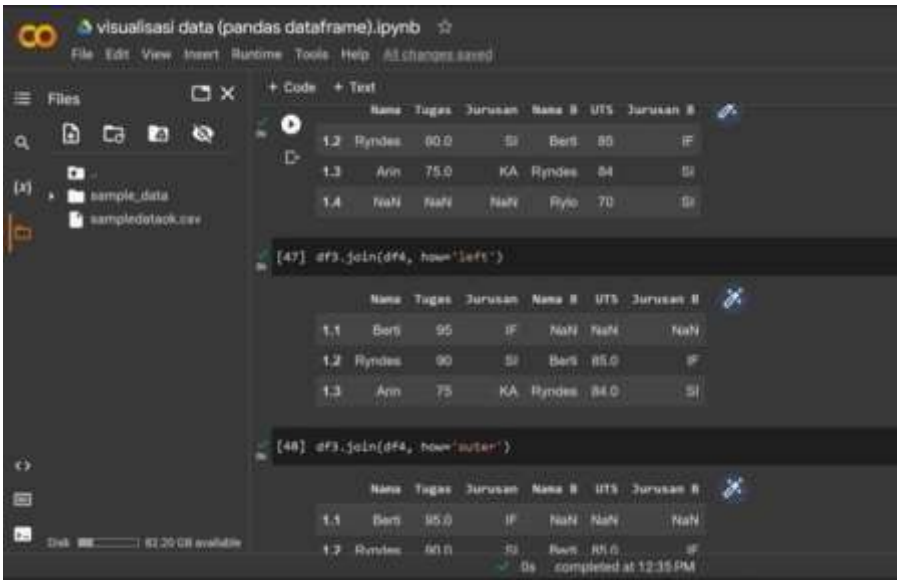

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
[34] df3.merge(df4, on='Nama', how='left')
Nama Tugas Jurusan_x UTS Jurusan_y
0 Berti 95 IF 85.0 IF
1 Ryndes 90 SI 84.0 SI
2 Arin 75 KA NaN NaN
df3.merge(df4, on='Nama', how='right')
Nama Tugas Jurusan_x UTS Jurusan_y
0 Berti 95.0 IF 85 IF
1 Ryndes 90.0 SI 84 SI
2 Rylo NaN NaN 70 SI
[37] df3.merge(df4, on='Nama', how='outer')
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
df3.merge(df4, on='Nama', how='right')
Nama Tugas Jurusan_x UTS Jurusan_y
0 Berti 95.0 IF 85 IF
1 Ryndes 90.0 SI 84 SI
2 Ryto NaN NaN 70 SI
[37] df3.merge(df4, on='Nama', how='outer')
Nama Tugas Jurusan_x UTS Jurusan_y
0 Berti 95.0 IF 85.0 IF
1 Ryndes 90.0 SI 84.0 SI
2 Arin 75.0 KA NaN NaN
3 Ryto NaN NaN 70.0 SI
0s completed at 12:35 PM

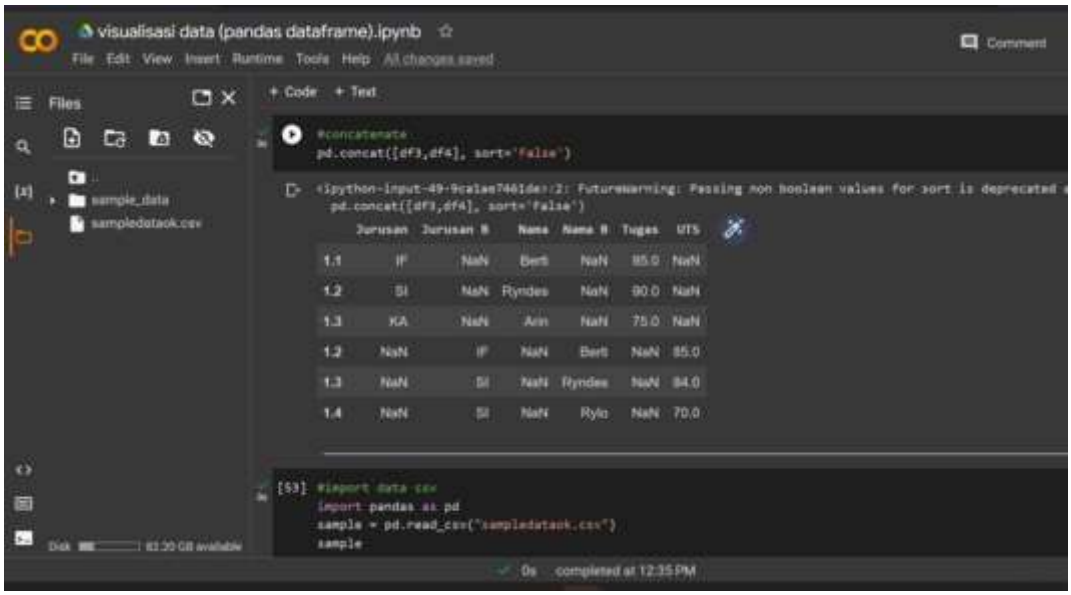
- Join

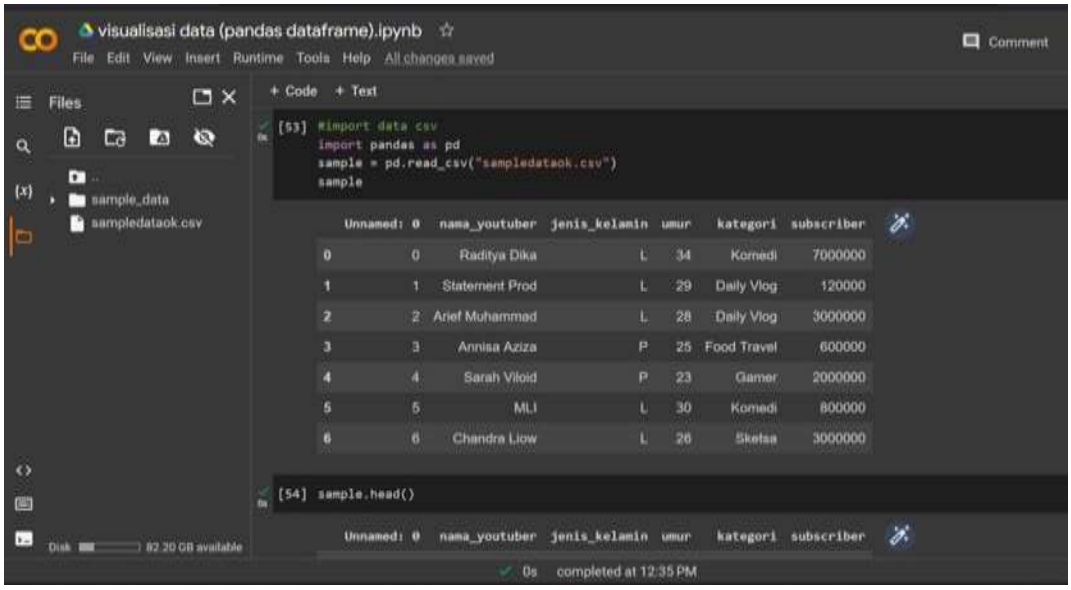

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
#join
#data pertama
nama = ['Berti','Ryndes','Arin']
tugas = [95,90,75]
jurusan = ['IF','SI','KA']
df3 = pd.DataFrame({'Nama':nama,'Tugas':tugas,'Jurusan':jurusan},index=['1.1','1.2','1.3'])
df3
Nama Tugas Jurusan
1.1 Berti 95 IF
1.2 Ryndes 90 SI
1.3 Arin 75 KA
[44] #data kedua
nama = ['Berti','Ryndes','Ryto']
uts = [85,84,70]
jurusan = ['IF','SI','SI']
df4 = pd.DataFrame({'Nama B':nama,'UTS':uts,'Jurusan B':jurusan}, index=['1.2','1.3','1.4'])
df4
0s completed at 12:35 PM

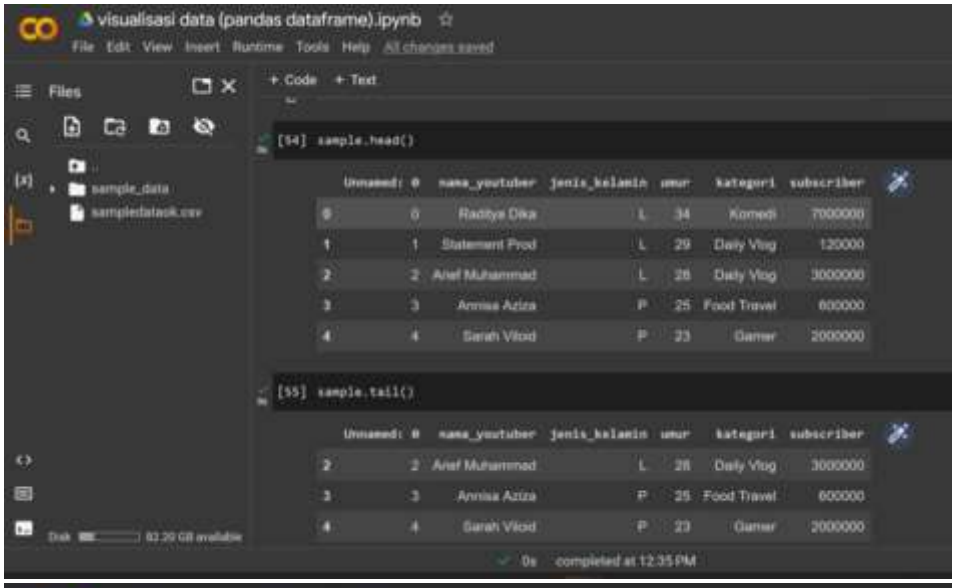


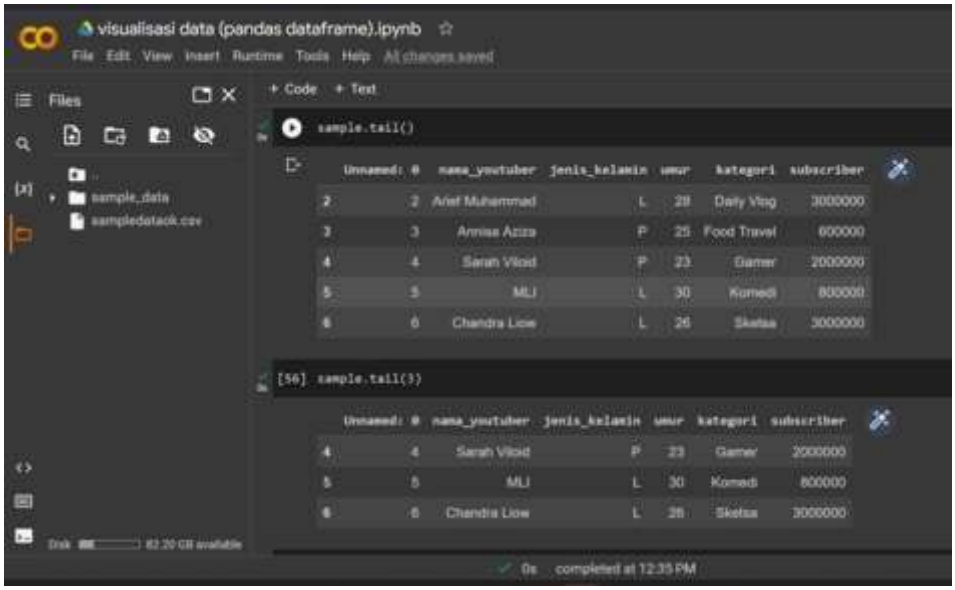


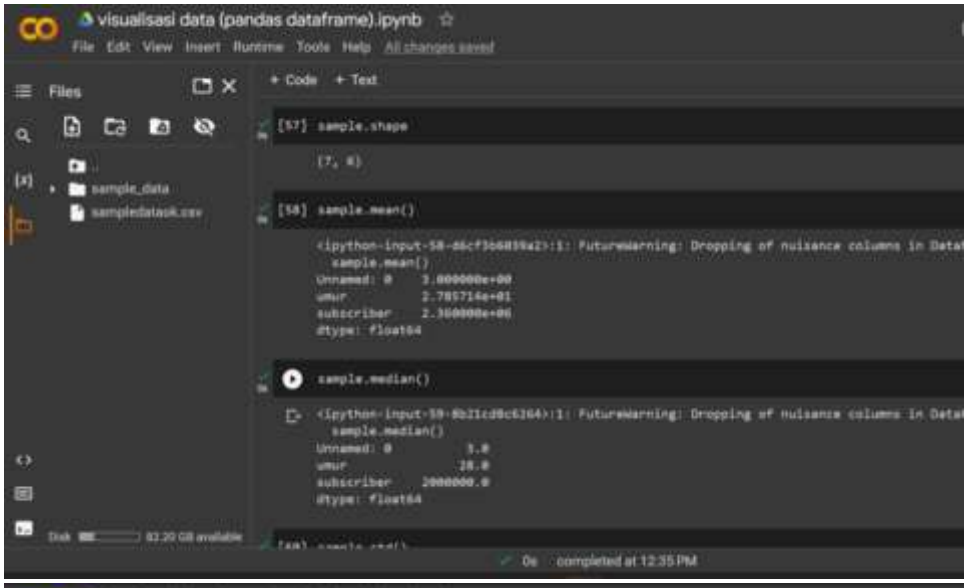
- Concatenate

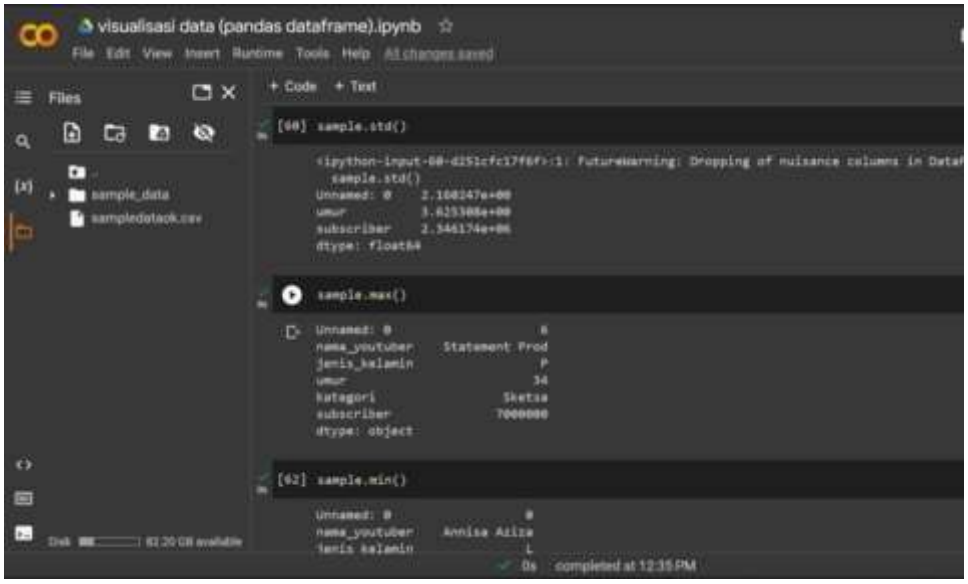

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
sample_data
sampledataok.csv
#concatenate
pd.concat([df3,df4], sort='false')
<ipython-input-49-9ca1ae7461de>:2: FutureWarning: Passing non boolean values for sort is deprecated
pd.concat([df3,df4], sort='false')
Jurusan Jurusan B Name Name B Tugas UTS
1.1 IF NaN Berti NaN 85.0 NaN
1.2 SI NaN Ryndes NaN 90.0 NaN
1.3 KA NaN Arin NaN 75.0 NaN
1.2 NaN IF NaN Berti NaN 85.0
1.3 NaN SI NaN Ryndes NaN 84.0
1.4 NaN SI NaN Rylo NaN 70.0
[53] #import data csv
import pandas as pd
sample = pd.read_csv("sampledataok.csv")
sample
0s completed at 12:35 PM

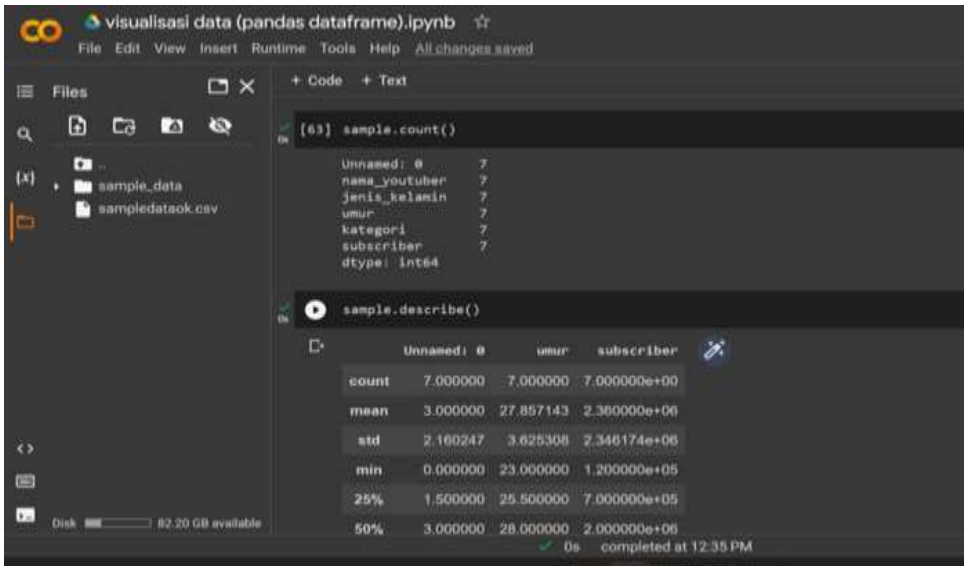

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Files
+ Code + Text
sample_data
sampledataok.csv
[53] #import data csv
import pandas as pd
sample = pd.read_csv("sampledataok.csv")
sample
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
0 0 Radityа Dika L 34 Komedi 7000000
1 1 Statement Prod L 29 Daily Vlog 120000
2 2 Arief Muhammad L 28 Daily Vlog 3000000
3 3 Annisa Aziza P 25 Food Travel 600000
4 4 Sarah Viloid P 23 Gamer 2000000
5 5 MLI L 30 Komedi 800000
6 6 Chandra Liow L 26 Sketsa 3000000
[54] sample.head()
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
0s completed at 12:35 PM

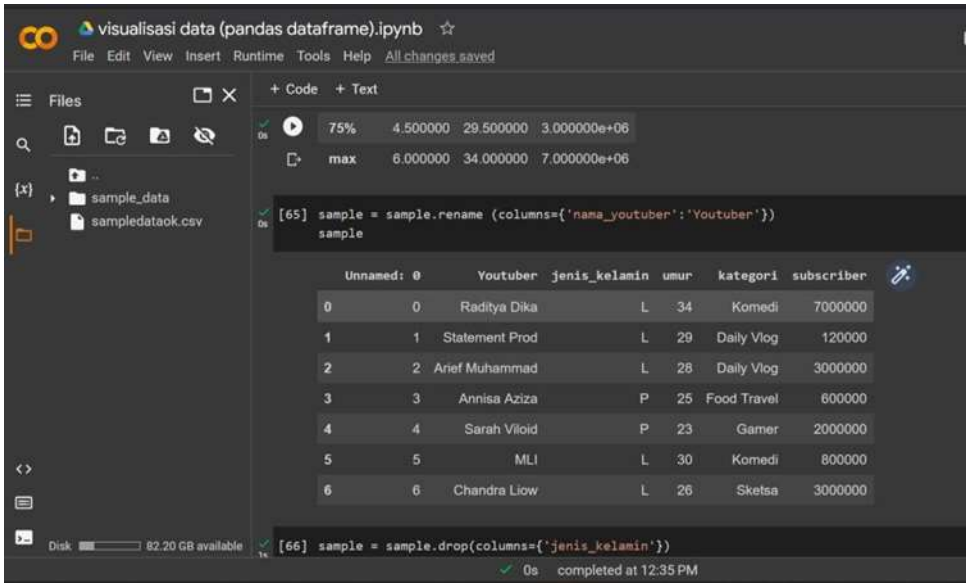

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
[54] sample.head()
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
0 0 Raditya Dika L 34 Komedi 7000000
1 1 Statement Prod L 29 Daily Vlog 120000
2 2 Arief Muhammad L 28 Daily Vlog 3000000
3 3 Annisa Aziza P 25 Food Travel 600000
4 4 Sarah Vllod P 23 Gamer 2000000
[55] sample.tail()
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
2 2 Arief Muhammad L 28 Daily Vlog 3000000
3 3 Annisa Aziza P 25 Food Travel 600000
4 4 Sarah Vllod P 23 Gamer 2000000
sample_data
sampledatacsv.csv
Disk 82.20 GB available
0s completed at 12:35 PM

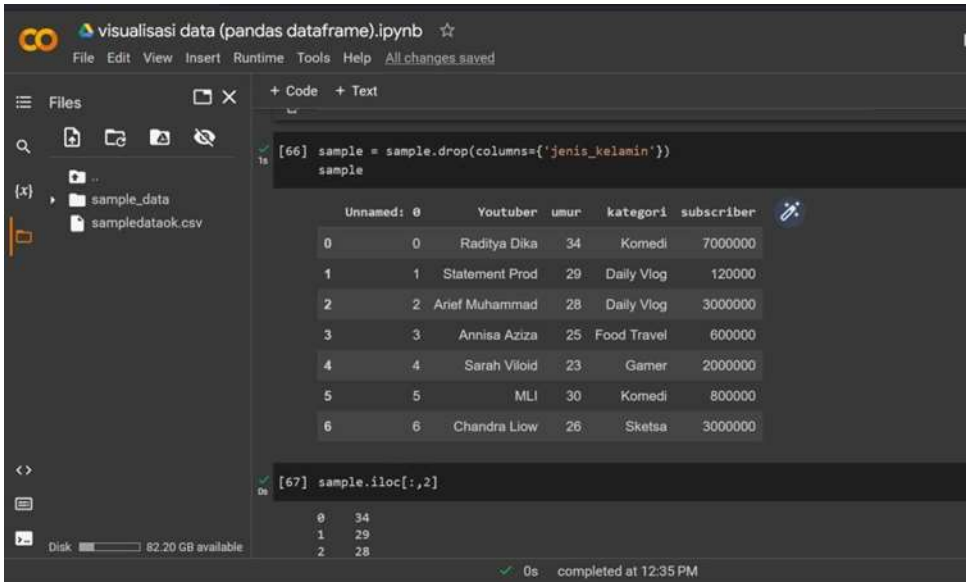

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample.tail()
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
2 2 Arief Muhammad L 28 Daily Vlog 3000000
3 3 Annisa Aziza P 25 Food Travel 600000
4 4 Sarah Vllod P 23 Gamer 2000000
5 5 MLI L 30 Komedi 800000
6 6 Chandra Liow L 26 Skelsa 3000000
[56] sample.tail(3)
Unnamed: 0 nama_youtuber jenis_kelamin umur kategori subscriber
4 4 Sarah Vllod P 23 Gamer 2000000
5 5 MLI L 30 Komedi 800000
6 6 Chandra Liow L 26 Skelsa 3000000
sample_data
sampledatacsv.csv
Disk 82.20 GB available
0s completed at 12:35 PM

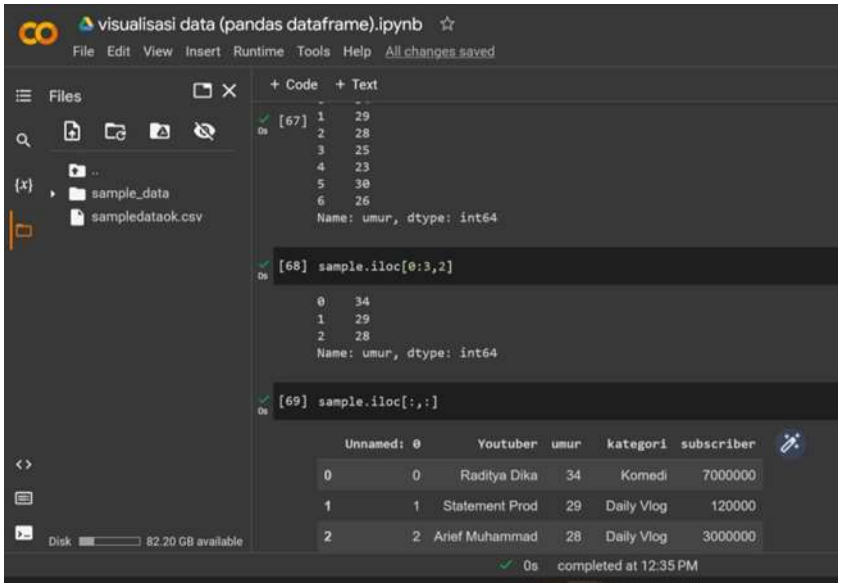

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
[57] sample.shape
(7, 6)
[58] sample.mean()
<ipython-input-58-d6cf3b6839a2>:1: FutureWarning: Dropping of nuisance columns in Data
sample.mean()
Unnamed: 0 3.000000e+00
umur 2.785714e+01
subscriber 2.360000e+06
dtype: float64
sample.median()
<ipython-input-59-8b31cd8c6364>:1: FutureWarning: Dropping of nuisance columns in Data
sample.median()
Unnamed: 0 3.0
umur 28.0
subscriber 2000000.0
dtype: float64
sample_data
sampledataok.csv
Disk 82.20 GB available
0s completed at 12:35 PM

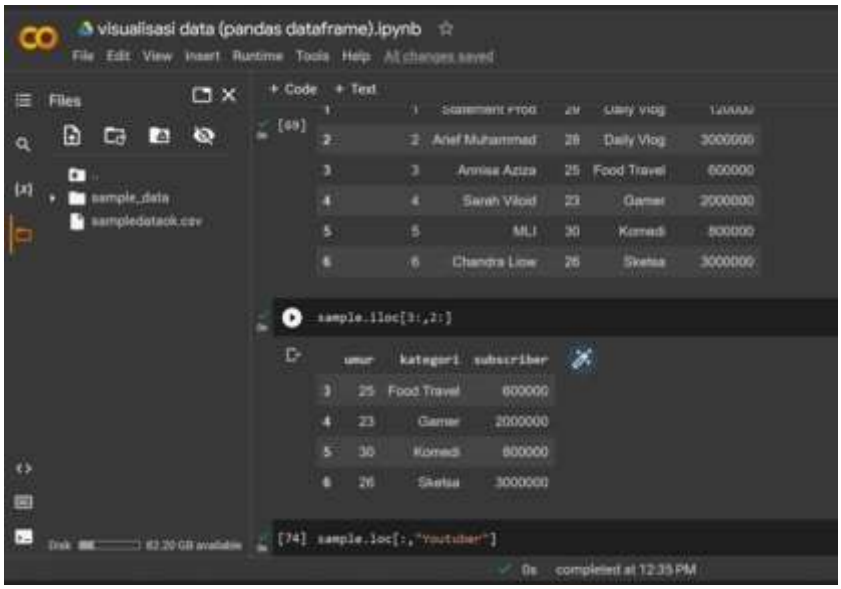

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
[60] sample.std()
<ipython-input-60-d251cfc17f6f>:1: FutureWarning: Dropping of nuisance columns in Data
sample.std()
Unnamed: 0 2.160247e+00
umur 3.625308e+00
subscriber 2.346174e+06
dtype: float64
sample.max()
Unnamed: 0 6
nama_youtuber Statement Prod
jenis_kelamin P
umur 34
kategori Sketsa
subscriber 7000000
dtype: object
[62] sample.min()
Unnamed: 0 0
nama_youtuber Annisa Aziza
jenis_kelamin L
sample_data
sampledataok.csv
Disk 82.20 GB available
0s completed at 12:35 PM

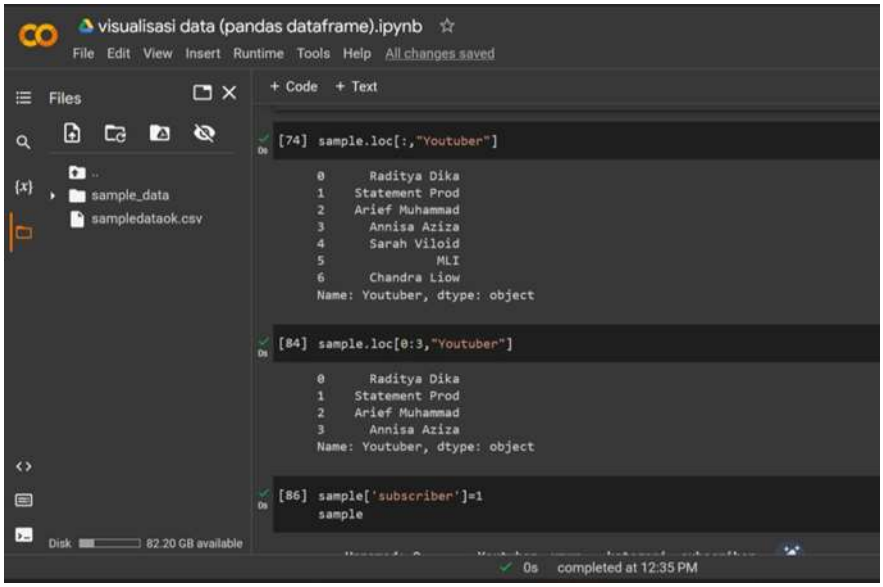

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
[63] sample.count()
Unnamed: 0 7
nama_youtuber 7
jenis_kelamin 7
umur 7
kategori 7
subscriber 7
dtype: int64
sample.describe()
Unnamed: 0 umur subscriber
count 7.000000 7.000000 7.000000e+00
mean 3.000000 27.857143 2.360000e+06
std 2.160247 3.625308 2.346174e+06
min 0.000000 23.000000 1.200000e+05
25% 1.500000 25.500000 7.000000e+05
50% 3.000000 28.000000 2.000000e+06
sample_data
sampledataok.csv
Disk 82.20 GB available
0s completed at 12:35 PM

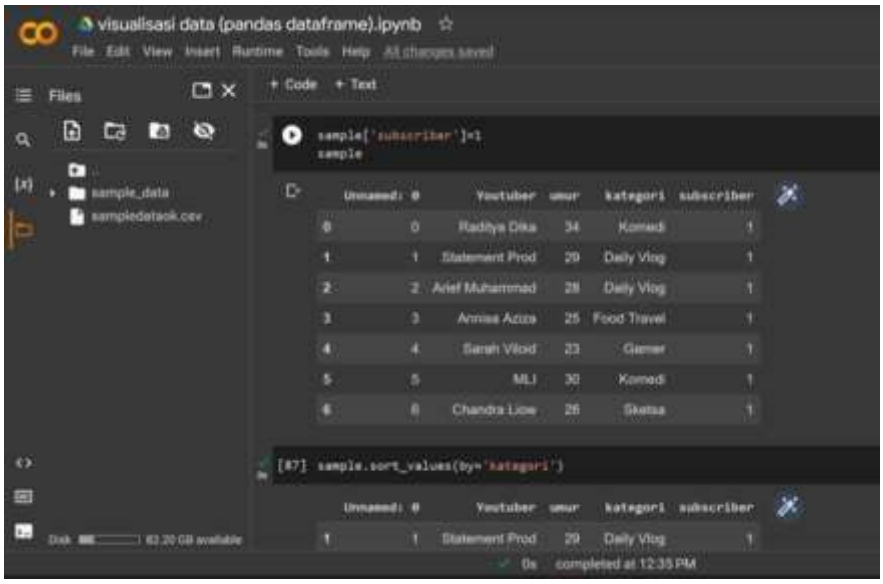

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
75% 4.500000 29.500000 3.000000e+06
max 6.000000 34.000000 7.000000e+06
[65] sample = sample.rename (columns={'nama_youtuber':'Youtuber'})
sample
Unnamed: 0 Youtuber jenis_kelamin umur kategori subscriber
0 0 Raditya Dika L 34 Komedi 7000000
1 1 Statement Prod L 29 Daily Vlog 120000
2 2 Arief Muhammad L 28 Daily Vlog 3000000
3 3 Annisa Aziza P 25 Food Travel 600000
4 4 Sarah Viloid P 23 Gamer 2000000
5 5 MLI L 30 Komedi 800000
6 6 Chandra Liow L 26 Sketsa 3000000
[66] sample = sample.drop(columns={'jenis_kelamin'})
Disk 82.20 GB available
0s completed at 12:35 PM

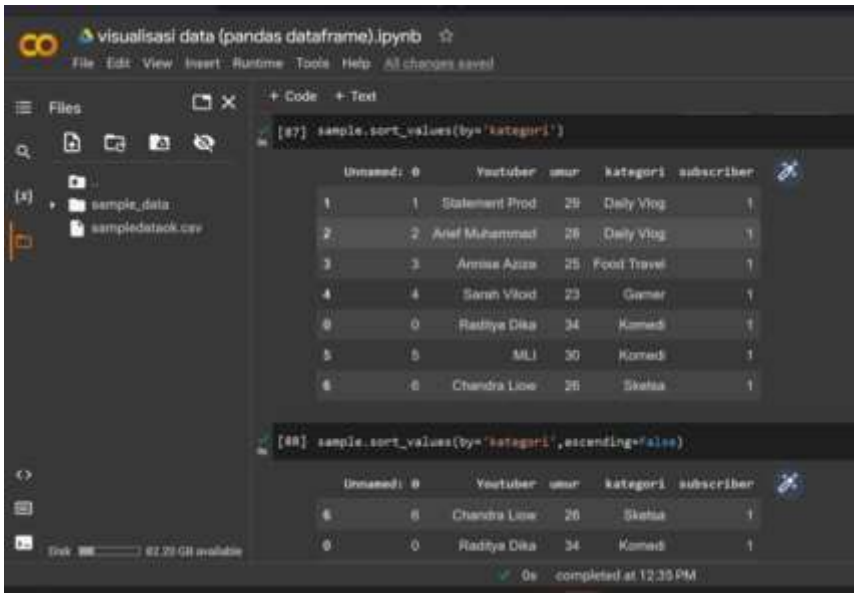

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
[66] sample = sample.drop(columns={'jenis_kelamin'})
sample
Unnamed: 0 Youtuber umur kategori subscriber
0 0 Raditya Dika 34 Komedi 7000000
1 1 Statement Prod 29 Daily Vlog 120000
2 2 Arief Muhammad 28 Daily Vlog 3000000
3 3 Annisa Aziza 25 Food Travel 600000
4 4 Sarah Viloid 23 Gamer 2000000
5 5 MLI 30 Komedi 800000
6 6 Chandra Liow 26 Sketsa 3000000
[67] sample.iloc[:,2]
0 34
1 29
2 28
Disk 82.20 GB available
0s completed at 12:35 PM

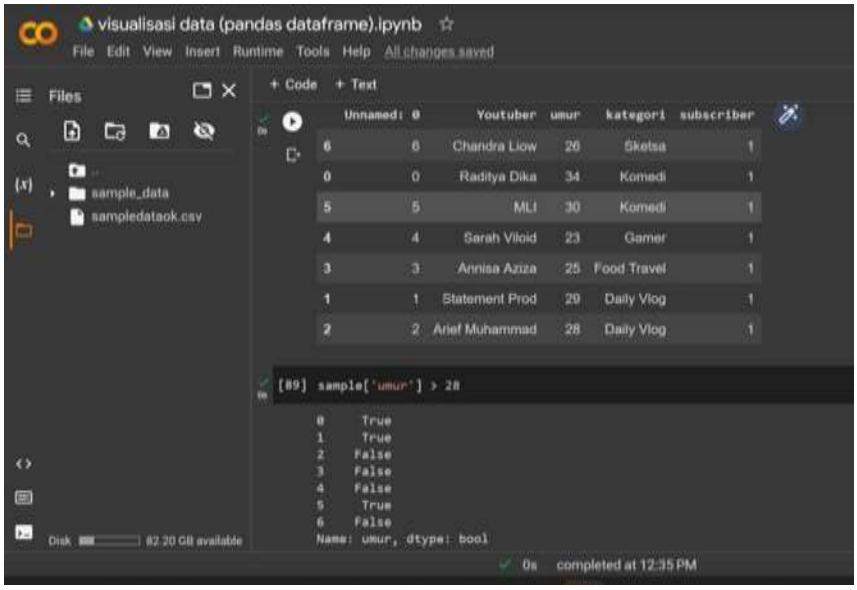

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
sample_data
sampledataok.csv
+ Code + Text
[67] 1 29
2 28
3 25
4 23
5 30
6 26
Name: umur, dtype: int64
[68] sample.iloc[0:3,2]
0 34
1 29
2 28
Name: umur, dtype: int64
[69] sample.iloc[:,:]
Unnamed: 0 Youtuber umur kategori subscriber
0 0 Raditya Dika 34 Komedi 7000000
1 1 Statement Prod 29 Daily Vlog 120000
2 2 Arief Muhammad 28 Daily Vlog 3000000
Disk 82.20 GB available
0s completed at 12:35 PM

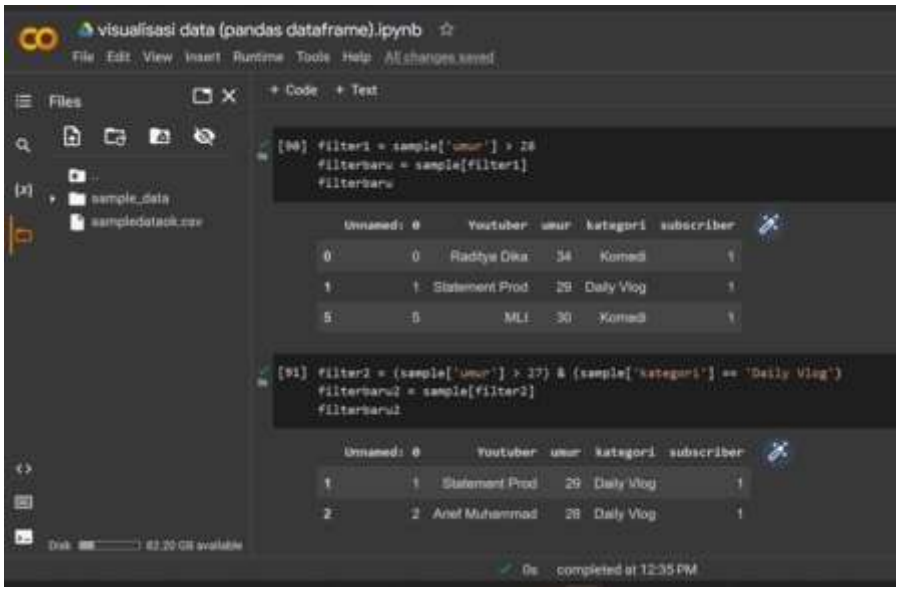

visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
sample_data
sampledataok.csv
+ Code + Text
[69]
2 2 Arief Muhammad 28 Daily Vlog 3000000
3 3 Annisa Aziza 25 Food Travel 600000
4 4 Sarah Vlog 23 Gamer 2000000
5 5 MLI 30 Komedi 800000
6 6 Chandra Liow 26 Sketsa 3000000
sample.iloc[3:,2:]
umur kategori subscriber
3 25 Food Travel 600000
4 23 Gamer 2000000
5 30 Komedi 800000
6 26 Sketsa 3000000
[74] sample.loc[:,"Youtuber"]
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
[74] sample.loc[:,"Youtuber"]
0 Raditya Dika
1 Statement Prod
2 Arief Muhammad
3 Annisa Aziza
4 Sarah Viloid
5 MLI
6 Chandra Liow
Name: Youtuber, dtype: object
[84] sample.loc[0:3,"Youtuber"]
0 Raditya Dika
1 Statement Prod
2 Arief Muhammad
3 Annisa Aziza
Name: Youtuber, dtype: object
[86] sample['subscriber']=1
sample
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledataok.csv
sample['subscriber']=1
sample
Unnamed: 0 Youtuber umur kategori subscriber
0 0 Raditya Dika 34 Komedi 1
1 1 Statement Prod 29 Daily Vlog 1
2 2 Arief Muhammad 28 Daily Vlog 1
3 3 Annisa Aziza 25 Food Travel 1
4 4 Sarah Viloid 23 Gamer 1
5 5 MLI 30 Komedi 1
6 6 Chandra Liow 26 Skena 1
[87] sample.sort_values(by='kategori')
Unnamed: 0 Youtuber umur kategori subscriber
1 1 Statement Prod 29 Daily Vlog 1
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
sample_data
sampledataok.csv
+ Code + Text
[87] sample.sort_values(by='kategori')
Unnamed: 0 Youtuber umur kategori subscriber
1 1 Statement Prod 29 Daily Vlog 1
2 2 Arief Muhammad 28 Daily Vlog 1
3 3 Annisa Aziza 25 Food Travel 1
4 4 Sarah Viloid 23 Gamer 1
0 0 Raditya Dika 34 Komedi 1
5 5 MLI 30 Komedi 1
6 6 Chandra Liow 26 Sketsa 1
[88] sample.sort_values(by='kategori',ascending=False)
Unnamed: 0 Youtuber umur kategori subscriber
6 6 Chandra Liow 26 Sketsa 1
0 0 Raditya Dika 34 Komedi 1
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
sample_data
sampledataok.csv
+ Code + Text
Unnamed: 0 Youtuber umur kategori subscriber
6 6 Chandra Liow 26 Sketsa 1
0 0 Raditya Dika 34 Komedi 1
5 5 MLI 30 Komedi 1
4 4 Sarah Viloid 23 Gamer 1
3 3 Annisa Aziza 25 Food Travel 1
1 1 Statement Prod 29 Daily Vlog 1
2 2 Arief Muhammad 28 Daily Vlog 1
[89] sample['umur'] > 28
0 True
1 True
2 False
3 False
4 False
5 True
6 False
Name: umur, dtype: bool
Disk 82.20 GB available
0s completed at 12:35 PM


visualisasi data (pandas dataframe).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
sample_data
sampledatacsv.csv
[90] filter1 = sample['umur'] > 28
filterbaru = sample[filter1]
filterbaru
Unnamed: 0 Youtuber umur kategori subscriber
0 0 Raditya Dika 34 Komedi 1
1 1 Statement Prod 29 Daily Vlog 1
5 5 MLI 30 Komedi 1
[91] filter2 = (sample['umur'] > 27) & (sample['kategori'] == 'Daily Vlog')
filterbaru2 = sample[filter2]
filterbaru2
Unnamed: 0 Youtuber umur kategori subscriber
1 1 Statement Prod 29 Daily Vlog 1
2 2 Arief Muhammad 28 Daily Vlog 1
Disk 82.20 GB available
0s completed at 12:35 PM