

Tugas 1: Judul tugas – Statistik Deskriptif dan Probabilitas

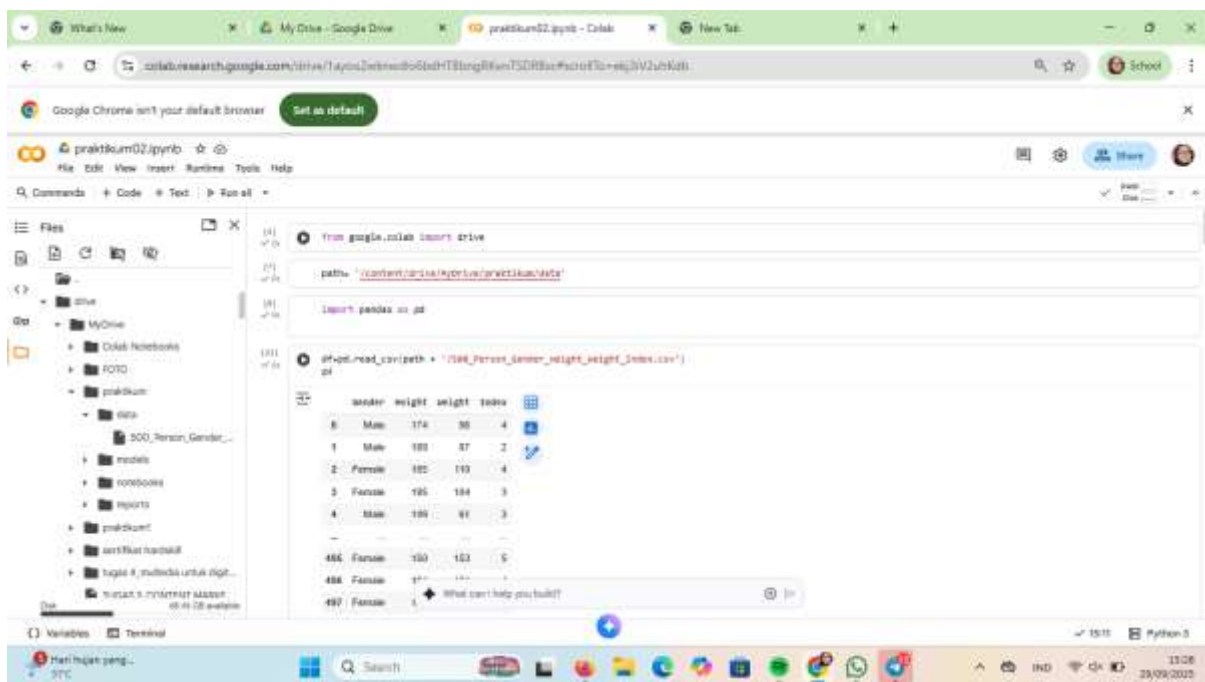
Siti aisah - 0110222129¹

¹ Teknik Informatika, STT Terpadu Nurul Fikri, Depok

*E-mail: siti22129ti@student.nurulfikri.ac.id

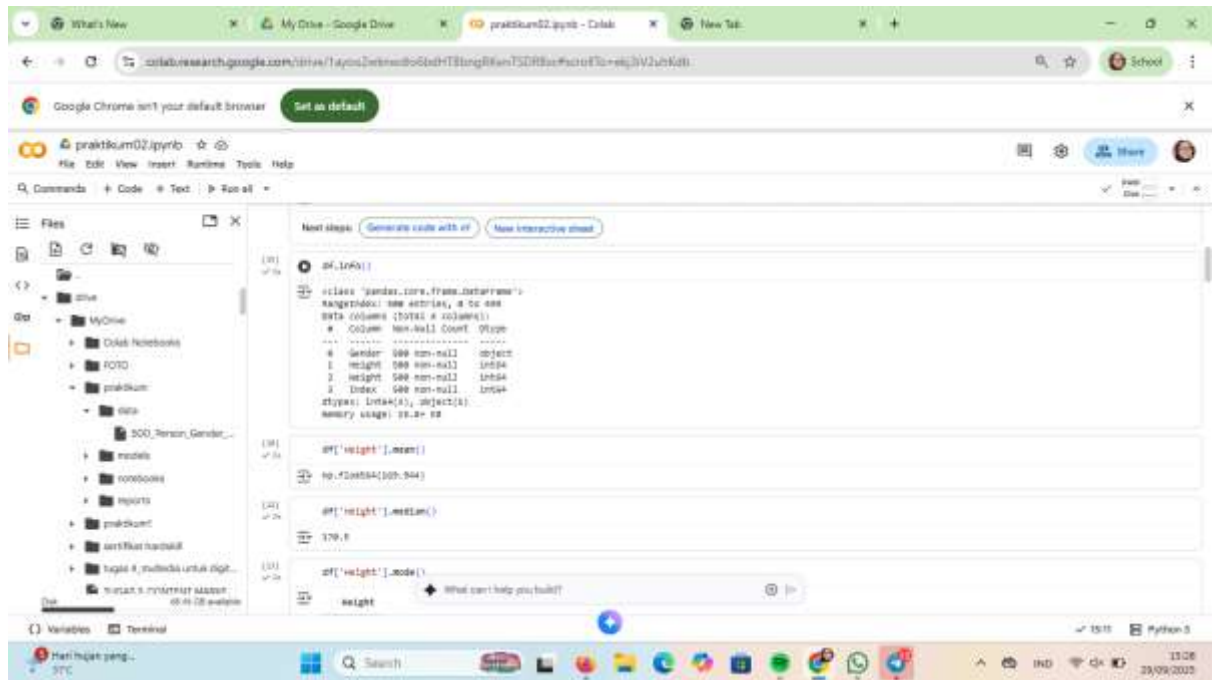
Abstract. Pembelajaran *Machine Learning* merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada pengembangan algoritma dan model statistik untuk memungkinkan sistem komputer belajar dari data dan membuat prediksi atau keputusan secara otomatis tanpa pemrograman eksplisit..

1. Praktikum 1



1.1 . Membaca file CSV

2. melihat informasi umum data



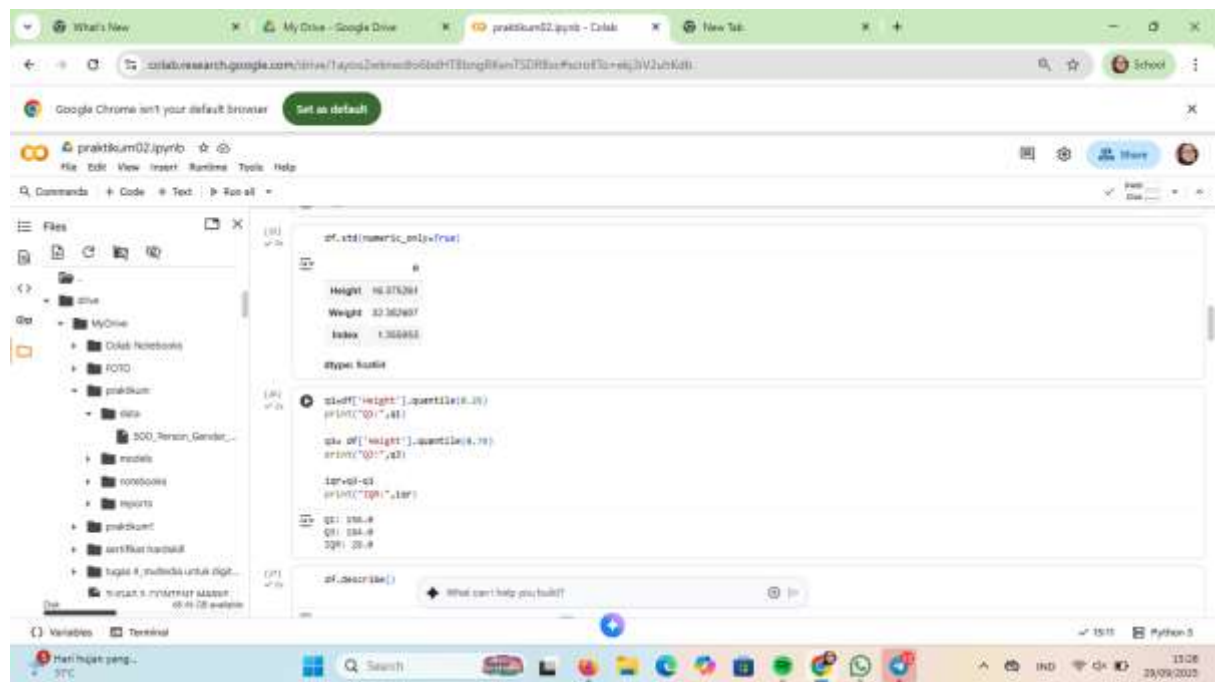
1.2 Hasil melihat informasi umum data

3. Menghitung Nilai-Nilai Sentral (Mean, Median, Modus)



1.3 hasil dari . Menghitung Nilai-Nilai Sentral (Mean, Median, Modus)

4.menghitung ukuran persebaran (variasi &standar deviasi)



```
(0) df['numeric_only=True']
Out[0]:
height    16.375261
weight    32.362807
index      1.368253
dtype: float64

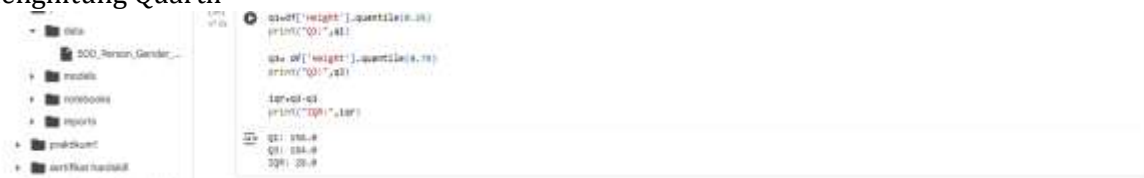
(1) df['height'].quantile(0.25)
Out[1]: 16.375261
df['height'].quantile(0.75)
Out[2]: 32.362807
Iqr=ql-ql
Out[3]: 16.375261

(2) df.describe()
```

The screenshot shows a Jupyter Notebook interface with a file explorer on the left and a code editor on the right. The code in the editor calculates the variance and standard deviation for the 'height' variable. The output shows the variance as 16.375261 and the standard deviation as 32.362807. The IQR is calculated as 16.375261.

1.4 hasil dari .menghitung ukuran persebaran (variasi &standar deviasi)

5.menghitung Quartil



```
(0) df['height'].quantile(0.25)
Out[0]: 16.375261
df['height'].quantile(0.75)
Out[1]: 32.362807
Iqr=ql-ql
Out[2]: 16.375261

(1) df.describe()
```

The screenshot shows a Jupyter Notebook interface with a file explorer on the left and a code editor on the right. The code in the editor calculates the quartiles for the 'height' variable. The output shows the first quartile (Q1) as 16.375261 and the third quartile (Q3) as 32.362807. The IQR is calculated as 16.375261.

1.5 hasil dari menghitung Quartil

6. Menghitung Statistik Deskriptif Otomatis



```
(0) df.describe()
```

	height	weight	index
count	100.000000	100.000000	100.000000
mean	16.375261	32.362807	1.368253
std	16.375261	32.362807	1.368253
min	140.000000	50.000000	0.000000
25%	150.000000	30.000000	2.000000
50%	170.000000	30.000000	4.000000
75%	104.000000	130.000000	5.000000
max	100.000000	100.000000	5.000000

The screenshot shows a Jupyter Notebook interface with a file explorer on the left and a code editor on the right. The code in the editor calculates the automatic descriptive statistics for the 'height' variable. The output shows a table with the following statistics: count, mean, std, min, 25%, 50%, 75%, and max for both 'height' and 'weight' variables.

1.6 hasil dari menghitung statistic deskriptif otomatis

7. menghitung korelasi



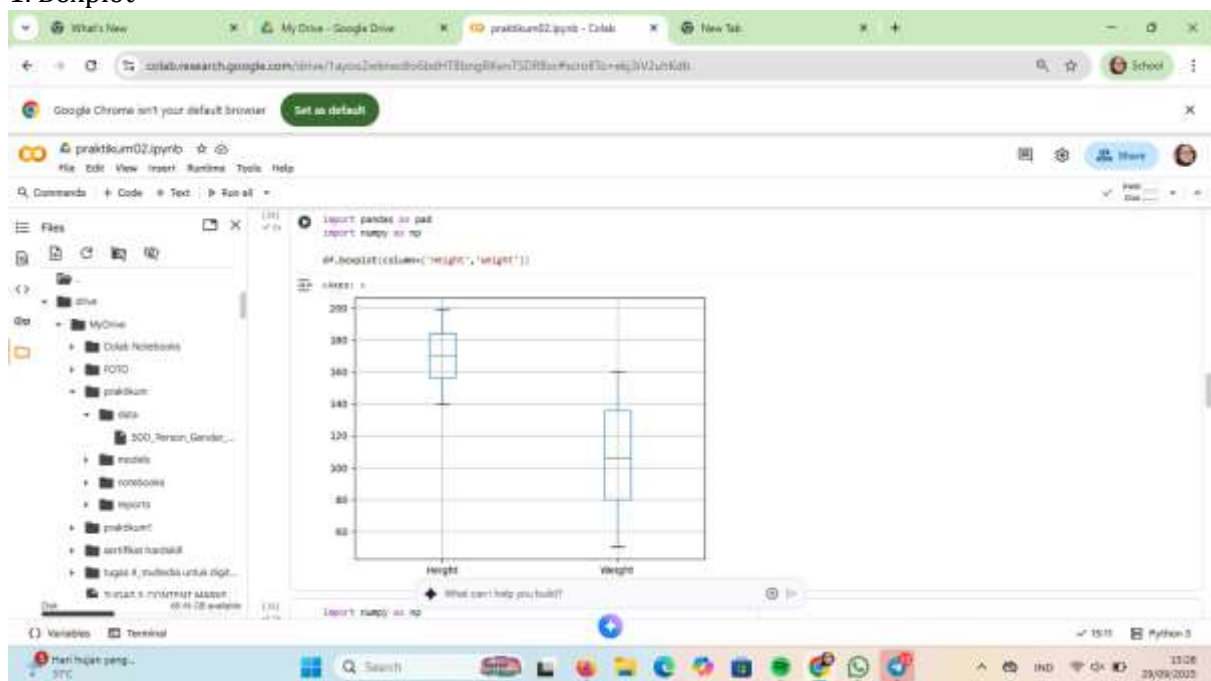
```
praktikum02.ipynb
File Edit View Insert Runtime Tools Help
Commands Code Text Run all
Files
drive
MyDrive
correlation_matrix_df.corr(numeric_only=True)
print(correlation_matrix)

          height  weight  score
height  1.000000  0.666444  0.422123
weight  0.666444  1.000000  0.664265
score   0.422123  0.664265  1.000000
```

1.7 hasil menghitung korelasi

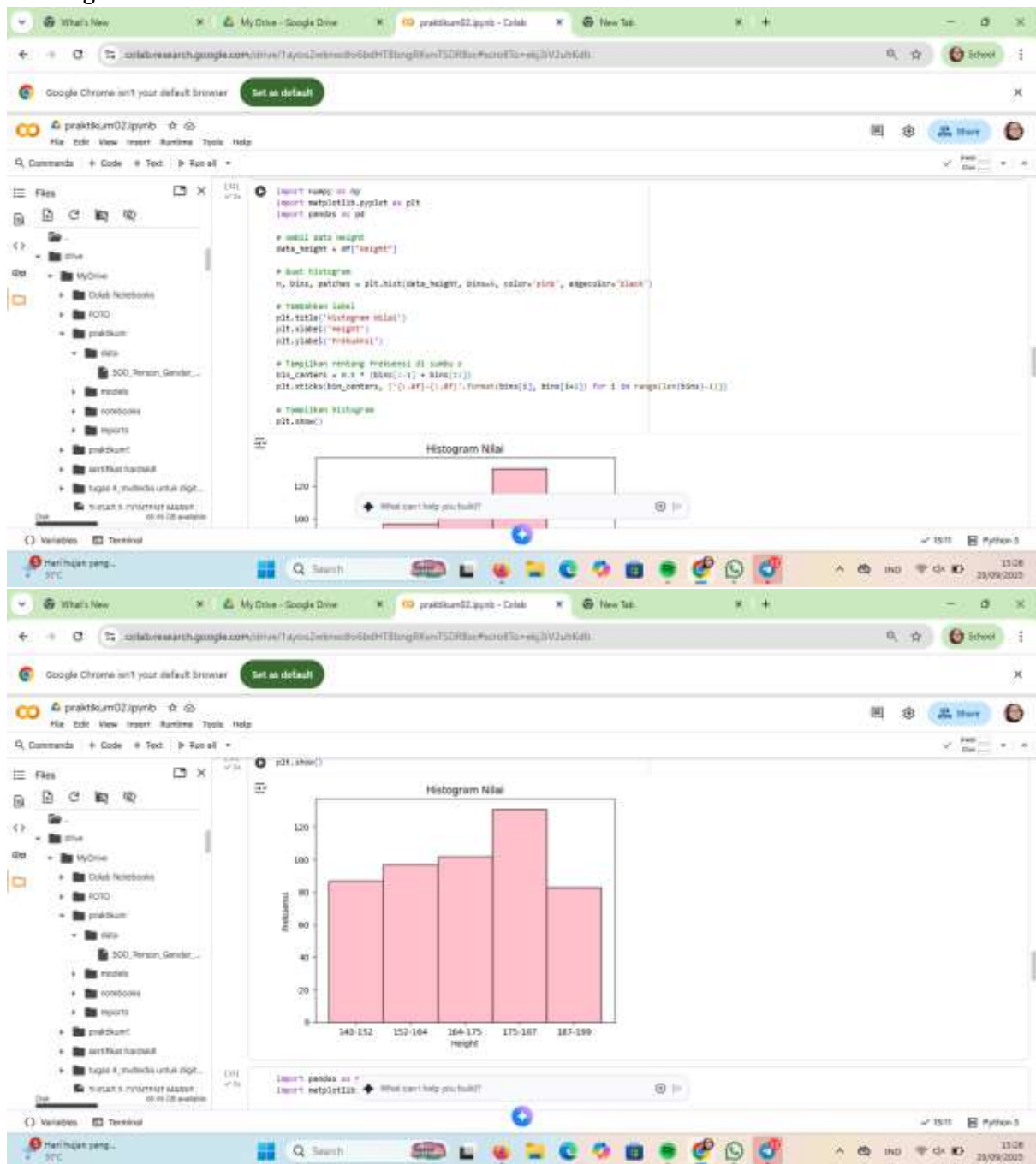
8. visualisasi data

1. Boxplot



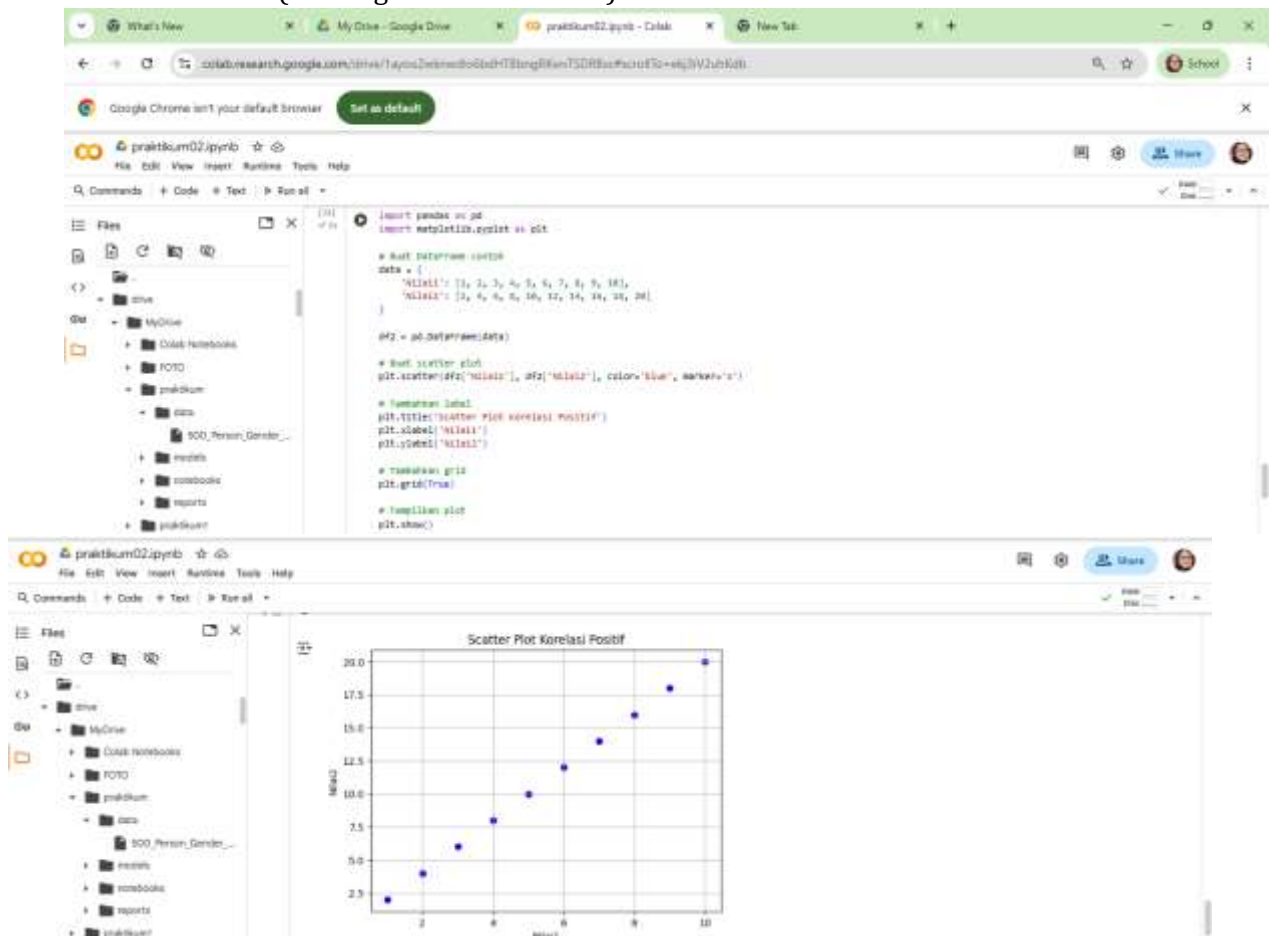
1.1 hasil dari boxplot

2. Histogram

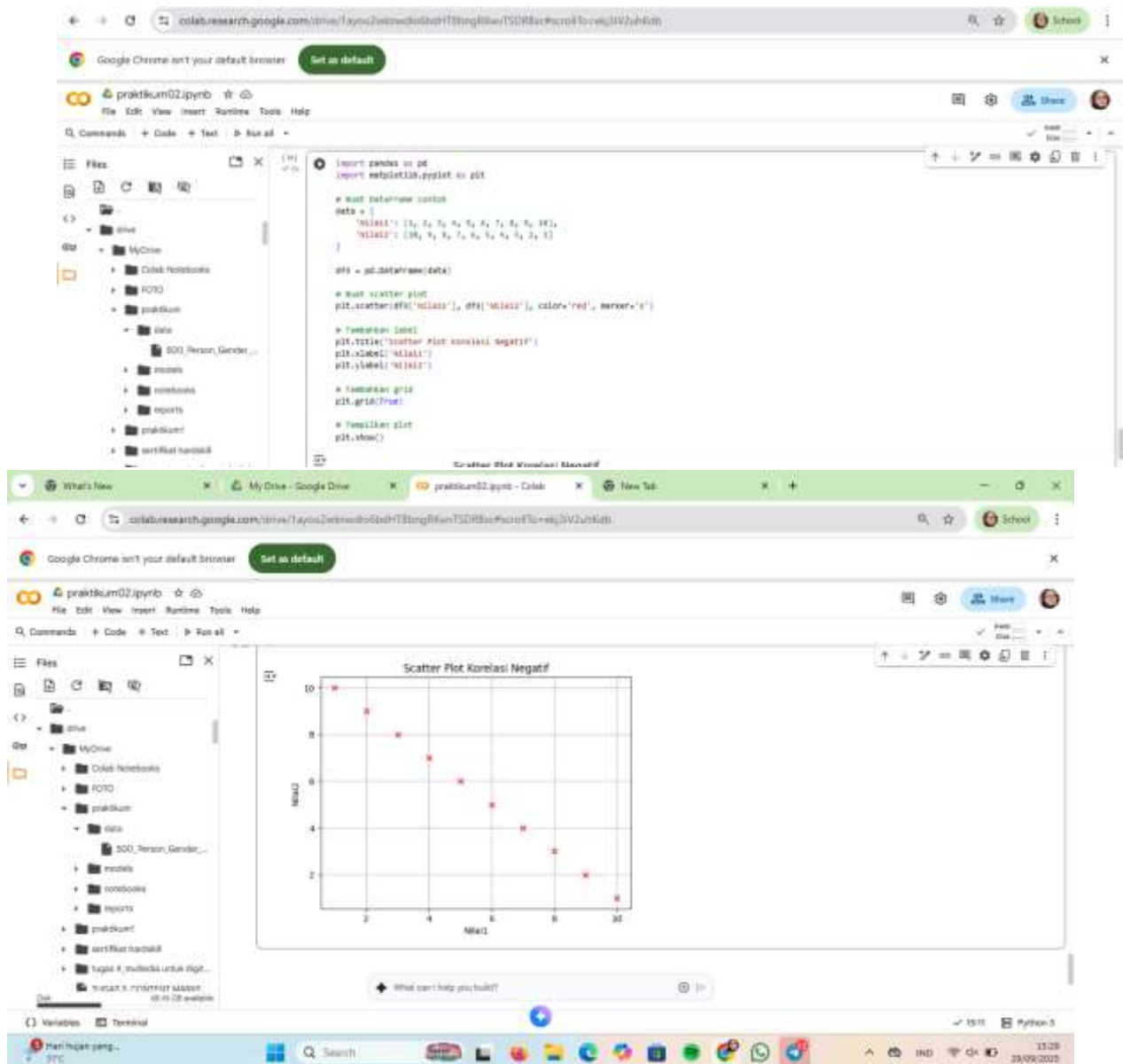


1.2 hasil histogram

3.Scatter Plot (Hubungan Antar Variabel)

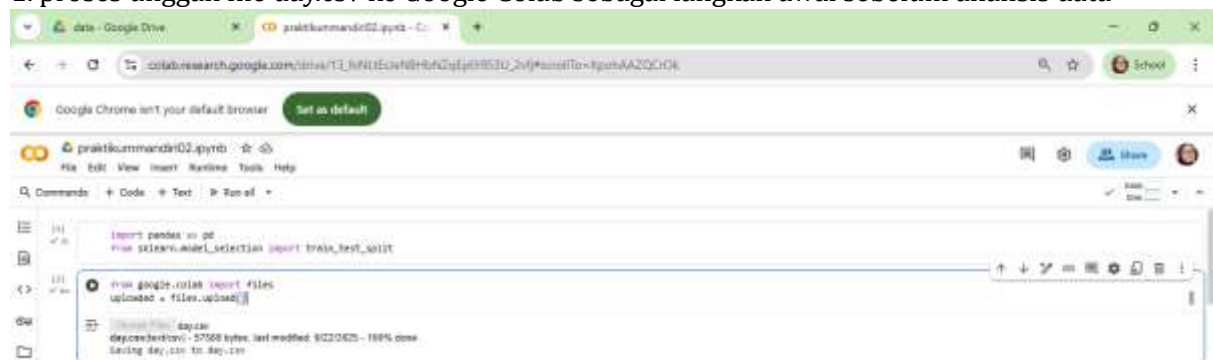


1.3 hasil dari .Scatter Plot (Hubungan Antar Variabel)



TUGAS PRAKTIKUM MANDIRI 2

1. proses unggah file day.csv ke Google Colab sebagai langkah awal sebelum analisis data



2. hasil pembacaan data day.csv dari google colab

```
print('jumlah total data:', len(df))
print(df.head())
```

```
jumlah total data: 711
Instant      dayofweek  season  yr  mth  holiday  weekday  workingday
0      1  2013-01-01      1    0    1    0      0      0
1      2  2013-01-02      1    0    1    0      0      0
2      3  2013-01-03      1    0    1    0      0      1
3      4  2013-01-04      1    0    1    0      1      1
4      5  2013-01-05      1    0    1    0      1      1

weather(sat)      temp      atemp      hum      windspeed      casual      registered
0      2  6.344157  6.383215  6.899221  6.348446  212  404
1      3  6.364775  6.377149  6.898877  6.348229  212  406
2      1  6.378344  6.384405  6.857173  6.348189  218  4109
3      3  6.388880  6.371212  6.839425  6.348796  208  3494
4      1  6.338957  6.329278  6.436157  6.338990  22  1518

cnt
0  554
1  591
2  5348
3  1532
4  1498
```

3. tampilan jumlah (Training, Validation, dan Testing)

```
train, test = train_test_split(df, test_size=0.1, random_state=42)

train, val = train_test_split(train, test_size=0.1, random_state=42)

print('jumlah data training:', len(train))
print(train.head())

print('jumlah data validasi:', len(val))
print(val.head())

print('jumlah data testing:', len(test))
print(test.head())
```

```
jumlah data training: 629
Instant      dayofweek  season  yr  mth  holiday  weekday  workingday
387      0    2013-10-19      4    1    10      0      0      1
388      1    2013-05-17      1    0    5      0      0      1
389      0    2013-11-02      4    0    11      0      0      1
390      2    2013-04-21      2    0    4      0      5      1
620      5    2013-06-21      1    1    6      0      0      1

weather(sat)      temp      atemp      hum      windspeed      casual      registered
447      2  6.561212  6.578306  6.815886  6.134954  793  4671
383      1  6.619089  6.602958  6.794621  6.385218  382  4237
386      1  6.577586  6.596112  6.712756  6.382953  178  1818
311      2  6.516657  6.523784  6.733931  6.219523  177  1390
530      1  6.773586  6.734121  6.873758  6.102542  664  4819

cnt
```


Referensi:

- Munir, S., Seminar, K. B., Sudradjat, Sukoco, H., & Buono, A. (2022). The Use of Random Forest Regression for Estimating Leaf Nitrogen Content of Oil Palm Based on Sentinel 1-A Imagery. *Information*, 14(1), 10. <https://doi.org/10.3390/info14010010>
- Seminar, K. B., Imantho, H., Sudradjat, Yahya, S., Munir, S., Kaliana, I., Mei Haryadi, F., Noor Baroroh, A., Supriyanto, Handoyo, G. C., Kurnia Wijayanto, A., Ijang Wahyudin, C., Liyantono, Budiman, R., Bakir Pasaman, A., Rusiawan, D., & Sulastri. (2024). PreciPalm: An Intelligent System for Calculating Macronutrient Status and Fertilizer Recommendations for Oil Palm on Mineral Soils Based on a Precision Agriculture Approach. *Scientific World Journal*, 2024(1). <https://doi.org/10.1155/2024/1788726>

link github: <https://github.com/Sitiaisah1604/machine-learning/tree/main/praktikum1>