

## Ada 3 Tahap yang perlu dilakukan untuk menganalisis data yakni

### 1. Understanding data

Tujuannya untuk mengetahui tipe data, berapa banyak data, nilai2 penting data misal rata2, standar deviasi, persentil dll → memahami data yang akan dieksplorasi

### 2. Preprocessing Data

Tujuannya untuk menyiapkan data yang akan diolah, yakni misalnya mengubah bentuk data yang semula string menjadi float/intejer

### 3. EDA/Exploratory Data Analysis dan Data Visualization

Tujuannya untuk mengeksplor data menggunakan chart, mengetahui hubungan antar data lalu diplotkan bersama dll

## UNDERSTANDING DATA

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
%matplotlib inline
import seaborn as sns
import folium
import folium.plugins
from sklearn import preprocessing
from statistics import mode
import statsmodels.api as sm
import matplotlib.pyplot as plt
data = pd.read_excel("C:/Users/LENOVO/Data_Nilai_Pegawai.xlsx")
```

Merupakan fungsi-fungsi yang akan dipanggil, dan juga perintah membaca data excel

```
data.head()
```

	ID PERSON	Tingkat Pendidikan Akhir	Usia	Satuan Kerja	Bulan	Nilai
0	1001	S3	61	DINAS B	Juli	7.88436
1	1002	SLTA	59	DINAS A	Juli	6.95183
2	1003	S2	57	BADAN D	Juli	8.18452
3	1004	SD	56	DINAS B	Juli	6.55612
4	1005	D3	54	DINAS B	Juli	6.34694

data.head() merupakan perintah untuk melihat 5 data teratas

```
data.shape
```

```
(30, 6)
```

data.shape merupakan perintah untuk mengetahui seberapa banyak kolom dan baris, untuk kasus ini maka ada 30 kolom dan 6 baris

```
data.dtypes
```

```
ID PERSON          int64
Tingkat Pendidikan Akhir  object
Usia               int64
Satuan Kerja        object
Bulan              object
Nilai              float64
dtype: object
```

data.dtypes gunanya untuk mengetahui jenis data, integer → bilangan bulat, objek → tulisan, float → pecahan

Untuk kasus ini ternyata:

ID PERSON : integer (bilangan bulat)

Tingkat Pendidikan Akhir : object karena berisi tulisan yaitu: SD, SLTP, SLTA, ...

Usia : integer

Satuan Kerja: objek karena Dinas A, Badan B dsb

Bulan: objek, Juli semua

Nilai: float(pecahan), nilai dari masing2 person(pegawai)

```
#untuk mengetahui banyaknya data yang tidak terisi(missing value) pada features dan %nya terhadap total data
total = data.isnull().sum().sort_values(ascending=False)
percent = ((data.isnull().sum()*100)/data.isnull().count()).sort_values(ascending=False)
missing_data = pd.concat([total, percent], axis=1, keys=['Total', 'Percent'],
                          sort=False).sort_values('Total', ascending=False)
missing_data.head(16)
```

	Total	Percent
Nilai	0	0.0
Bulan	0	0.0
Satuan Kerja	0	0.0
Usia	0	0.0
Tingkat Pendidikan Akhir	0	0.0
ID PERSON	0	0.0

Diatas merupakan perintah untuk mengetahui missing data, yang ternyata tidak ada data yang hilang

```
data.describe()
```

	ID PERSON	Usia	Nilai
count	30.000000	30.000000	30.000000
mean	1015.500000	40.233333	7.185791
std	8.803408	10.753294	0.484566
min	1001.000000	25.000000	6.346940
25%	1008.250000	32.500000	6.964457
50%	1015.500000	41.000000	7.179420
75%	1022.750000	47.000000	7.452400
max	1030.000000	61.000000	8.403630

Merupakan perintah untuk mengetahui hal penting terkait data yakni rata-rata, standar deviasi, minimal nilai data , persentil 25%, persentil 50% dan persentil 75 % dan nilai data maksimal.

Disini hanya 3 data yang dapat dilihat karena data object tidak dapat terbaca (masih harus diproses dulu)

Tahap 1 selesai: dapat dipahami bahwa data yang akan kita olah merupakan data nilai pegawai yang berisi feature yaitu ID person, Tingkat Pendidikan Akhir, Satuan Kerja, Bulan dan Nilai pegawai, (6 baris), terhadap 30 kolom(sebanyak 30 pegawai). Tidak ada data yang hilang. Namun, ada data yang harus diubah terlebih dahulu yaitu Tingkat Pendidikan Akhir, Satuan Kerja dan Bulan.

## PREPROCESSING DATA

```
data.rename(  
    columns={  
        'ID PERSON':'ID',  
        'Tingkat Pendidikan Akhir' : 'TPA',  
        'Satuan Kerja' : 'SK'  
    },  
    inplace=True  
)  
data.head()
```

	ID	TPA	Usia	SK	Bulan	Nilai
0	1001	S3	61	DINAS B	Juli	7.88436
1	1002	SLTA	59	DINAS A	Juli	6.95183
2	1003	S2	57	BADAN D	Juli	8.18452
3	1004	SD	56	DINAS B	Juli	6.55612
4	1005	D3	54	DINAS B	Juli	6.34694

Merupakan perintah untuk mengubah judul kolom(feature), hal ini berguna untuk proses pemanggilan, karena di notebook jupyter tidak bisa memanggil nama dengan spasi

```
SK ={'DINAS A':1, 'DINAS B':2, 'DINAS F':3,  
     'DINAS P':4, 'BADAN B':5,  
     'BADAN D':6 }  
data['SK'] = data.SK.map(SK)  
data.dtypes
```

```
ID          int64  
TPA          object  
Usia         int64  
SK           int64  
Bulan        object  
Nilai        float64  
dtype: object
```

Perintah untuk mengubah data Satuan Kerja(SK) yang berisi dinas A, dinas B dsb (objek) menjadi bilangan bulat(intejer)

DINAS A :1

DINAS B :2

DINAS F :3

DINAS P :4

BADAN B :5

BADAN D: 6

```
TPA ={'SD':1, 'SLTP':2, 'SLTA':3,  
      'D3':4, 'D4':5,  
      'S1':6, 'S2':7, 'S3':8}  
data['TPA'] = data.TPA.map(TPA)  
data.dtypes
```

```
ID          int64  
TPA          int64  
Usia         int64  
SK           int64  
Bulan        object  
Nilai        float64  
dtype: object
```

Sama halnya dengan diatas, mengubah objek Tingkat Pendidikan Akhir(TPA) menjadi intejer.

```
data.head()
```

	ID	TPA	Usia	SK	Bulan	Nilai
0	1001	8	61	2	Juli	7.88436
1	1002	3	59	1	Juli	6.95183
2	1003	7	57	6	Juli	8.18452
3	1004	1	56	2	Juli	6.55612
4	1005	4	54	2	Juli	6.34694

Sehingga akhirnya kita dapat data seperti diatas, untuk Bulan tidak perlu diubah karena isinya sama yakni Juli semua.

```
#untuk mengetahui korelasi antar feature
data.corr().style.background_gradient(cmap='coolwarm')
```

	ID	TPA	Usia	SK	Nilai
ID	1	0.33707	-0.978945	-0.0800524	0.146537
TPA	0.33707	1	-0.29064	-0.219014	0.5143
Usia	-0.978945	-0.29064	1	0.0803054	-0.127943
SK	-0.0800524	-0.219014	0.0803054	1	-0.158226
Nilai	0.146537	0.5143	-0.127943	-0.158226	1

Ternyata ada hubungan antara nilai kompetensi pegawai dan Tingkat Pendidikan Akhir(TPA). Maps ini menunjukkan keterkaitan Antara satu feature dengan feature lain(semakin mendekati 1/semakin merah warna mapnya, maka keterkaitan semakin dekat)

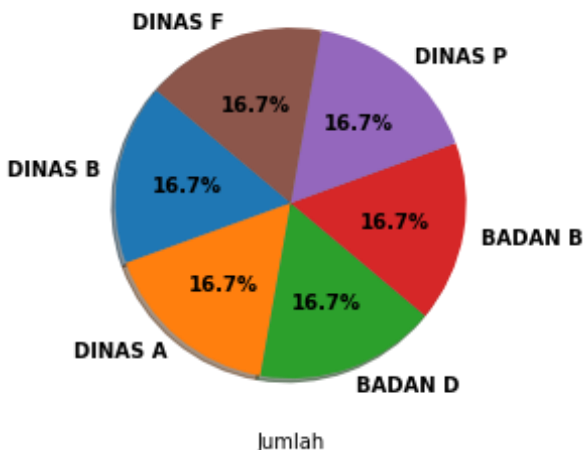
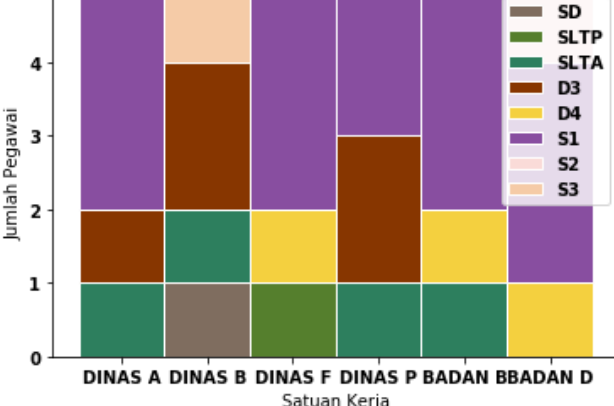
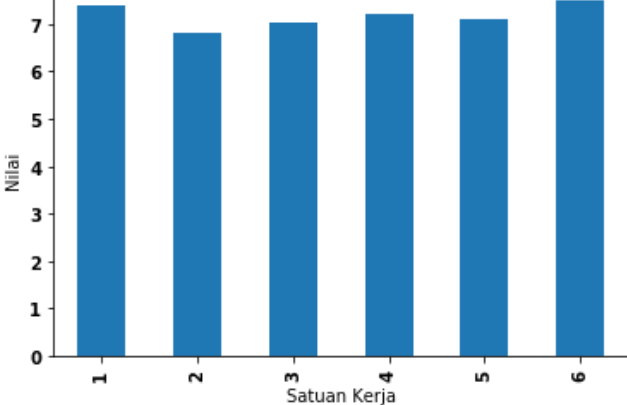
```
data.describe()
```

	ID	TPA	Usia	SK	Nilai
count	30.000000	30.000000	30.000000	30.000000	30.000000
mean	1015.500000	4.966667	40.233333	3.500000	7.185791
std	8.803408	1.586219	10.753294	1.737021	0.484566
min	1001.000000	1.000000	25.000000	1.000000	6.346940
25%	1008.250000	4.000000	32.500000	2.000000	6.964457
50%	1015.500000	6.000000	41.000000	3.500000	7.179420
75%	1022.750000	6.000000	47.000000	5.000000	7.452400
max	1030.000000	8.000000	61.000000	6.000000	8.403630

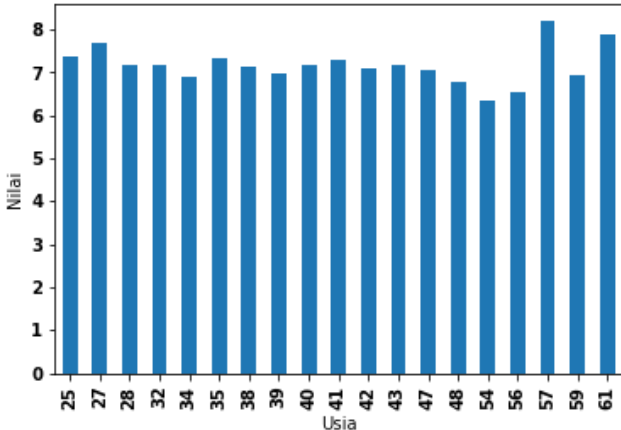
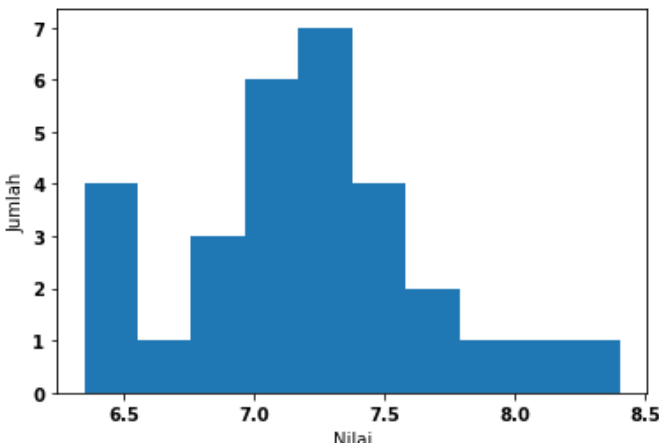
Data diatas menunjukkan nilai mean, std, persentil 25%, persentil 50%, persentil 75%, dan maksimial. Misalnya data usia rata-rata yakni 40 tahun dengan std 10.7(std: menunjukkan rentang antar data umur cukup jauh), sedangkan rata-rata nilai 7.18 dengan std 0.48(menunjukkan rentang antar data nilai tidak begitu jauh).

## EDA/ Exploratory Data Analysis and Visualization

Berikut hasil eksplorasi data dan visualisasi. Cara memplot dan eksplorasi data akan saya lampirkan di bagian paling akhir agar pada bagian ini khusus untuk menjawab 2 pertanyaan yang diajukan.

No	Gambar	Keterangan																																																															
1	 <p>A pie chart titled 'Jumlah' showing the distribution of 30 employees across six units. Each unit represents 16.7% of the total. The units are: DINAS F (brown), DINAS P (purple), BADAN B (red), BADAN D (green), DINAS A (orange), and DINAS B (blue).</p> <table><thead><tr><th>Satuan Kerja</th><th>Persentase</th></tr></thead><tbody><tr><td>DINAS F</td><td>16.7%</td></tr><tr><td>DINAS P</td><td>16.7%</td></tr><tr><td>BADAN B</td><td>16.7%</td></tr><tr><td>BADAN D</td><td>16.7%</td></tr><tr><td>DINAS A</td><td>16.7%</td></tr><tr><td>DINAS B</td><td>16.7%</td></tr></tbody></table>	Satuan Kerja	Persentase	DINAS F	16.7%	DINAS P	16.7%	BADAN B	16.7%	BADAN D	16.7%	DINAS A	16.7%	DINAS B	16.7%	Gambar disamping menunjukkan jumlah pegawai dari setiap satuan kerja. Dapat diketahui total pegawai ada 30 orang, dapat terlihat bahwa setiap satuan kerja yang jumlahnya ada 6 berisi 5 pegawai (dari pie chart disamping).																																																	
Satuan Kerja	Persentase																																																																
DINAS F	16.7%																																																																
DINAS P	16.7%																																																																
BADAN B	16.7%																																																																
BADAN D	16.7%																																																																
DINAS A	16.7%																																																																
DINAS B	16.7%																																																																
2	 <p>A stacked bar chart showing the number of employees by unit and education level. The units are: DINAS A, DINAS B, DINAS F, DINAS P, BADAN B, and BADAN D. The education levels are: SD (brown), SLTP (green), SLTA (dark green), D3 (orange), D4 (yellow), S1 (purple), S2 (pink), and S3 (light orange).</p> <table><thead><tr><th>Satuan Kerja</th><th>SD</th><th>SLTP</th><th>SLTA</th><th>D3</th><th>D4</th><th>S1</th><th>S2</th><th>S3</th></tr></thead><tbody><tr><td>DINAS A</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>DINAS B</td><td>1</td><td>0</td><td>1</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>DINAS F</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>DINAS P</td><td>0</td><td>1</td><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>BADAN B</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>BADAN D</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr></tbody></table>	Satuan Kerja	SD	SLTP	SLTA	D3	D4	S1	S2	S3	DINAS A	1	1	0	1	0	0	0	0	DINAS B	1	0	1	2	0	0	0	0	DINAS F	0	1	0	0	1	0	0	0	DINAS P	0	1	0	2	0	0	0	0	BADAN B	0	1	0	0	1	0	0	0	BADAN D	0	0	0	0	0	1	0	0	Gambar disamping menunjukkan keterkaitan antara banyaknya jumlah pegawai pada masing-masing satuan kerja berdasarkan tingkat pendidikan akhir. Misalnya dinas A memiliki satu pegawai lulusan SLTA, satu pegawai lulusan D3 dan tiga pegawai lulusan S1.
Satuan Kerja	SD	SLTP	SLTA	D3	D4	S1	S2	S3																																																									
DINAS A	1	1	0	1	0	0	0	0																																																									
DINAS B	1	0	1	2	0	0	0	0																																																									
DINAS F	0	1	0	0	1	0	0	0																																																									
DINAS P	0	1	0	2	0	0	0	0																																																									
BADAN B	0	1	0	0	1	0	0	0																																																									
BADAN D	0	0	0	0	0	1	0	0																																																									
3	 <p>A bar chart showing the average value for each unit. The units are numbered 1 to 6. The values are approximately: 1 (7.5), 2 (6.8), 3 (7.1), 4 (7.3), 5 (7.2), and 6 (7.6).</p> <table><thead><tr><th>Satuan Kerja</th><th>Nilai</th></tr></thead><tbody><tr><td>1</td><td>7.5</td></tr><tr><td>2</td><td>6.8</td></tr><tr><td>3</td><td>7.1</td></tr><tr><td>4</td><td>7.3</td></tr><tr><td>5</td><td>7.2</td></tr><tr><td>6</td><td>7.6</td></tr></tbody></table>	Satuan Kerja	Nilai	1	7.5	2	6.8	3	7.1	4	7.3	5	7.2	6	7.6	Gambar disamping memperlihatkan nilai rata-rata dari masing-masing satuan kerja, dapat terlihat Badan D dan Dinas A memiliki nilai rata-rata pegawai yang paling tinggi diantara yang lain.																																																	
Satuan Kerja	Nilai																																																																
1	7.5																																																																
2	6.8																																																																
3	7.1																																																																
4	7.3																																																																
5	7.2																																																																
6	7.6																																																																

4	<table border="1"><caption>Data for Gambar 4: Jumlah Pegawai per Tingkat Pendidikan Akhir</caption><thead><tr><th>Tingkat Pendidikan Akhir</th><th>Jumlah</th></tr></thead><tbody><tr><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td></tr><tr><td>3</td><td>4</td></tr><tr><td>4</td><td>5</td></tr><tr><td>5</td><td>3</td></tr><tr><td>6</td><td>14</td></tr><tr><td>7</td><td>1</td></tr><tr><td>8</td><td>1</td></tr></tbody></table>	Tingkat Pendidikan Akhir	Jumlah	1	1	2	1	3	4	4	5	5	3	6	14	7	1	8	1	Gambar disamping menunjukkan jumlah pegawai terbanyak ternyata merupakan lulusan S1 yakni 14 orang, ddikuti lulusan D3 sebanyak 5 orang.																						
Tingkat Pendidikan Akhir	Jumlah																																									
1	1																																									
2	1																																									
3	4																																									
4	5																																									
5	3																																									
6	14																																									
7	1																																									
8	1																																									
5	<table border="1"><caption>Data for Gambar 5: Nilai Rata-rata per Tingkat Pendidikan Akhir</caption><thead><tr><th>Tingkat Pendidikan Akhir</th><th>Nilai</th></tr></thead><tbody><tr><td>D3</td><td>7.0</td></tr><tr><td>D4</td><td>6.8</td></tr><tr><td>S1</td><td>7.3</td></tr><tr><td>S2</td><td>8.2</td></tr><tr><td>S3</td><td>7.8</td></tr><tr><td>SD</td><td>6.5</td></tr><tr><td>SLTA</td><td>6.9</td></tr><tr><td>SLTP</td><td>7.2</td></tr></tbody></table>	Tingkat Pendidikan Akhir	Nilai	D3	7.0	D4	6.8	S1	7.3	S2	8.2	S3	7.8	SD	6.5	SLTA	6.9	SLTP	7.2	Gambar disamping menunjukkan pegawai lulusan S2 dan S3 memiliki nilai rata-rata paling tinggi. Namun perlu diingat bahwa nilai tersebut diambil dari 1 data(lihat gambar 4 diatas), sedangkan nilai rata-rata S1 diambil dari 14 data sehingga standar deviasinya sangat jauh berbeda. Begitupun dengan nilai pegawai lulusan SLTP yang hanya 1 orang/1 data.																						
Tingkat Pendidikan Akhir	Nilai																																									
D3	7.0																																									
D4	6.8																																									
S1	7.3																																									
S2	8.2																																									
S3	7.8																																									
SD	6.5																																									
SLTA	6.9																																									
SLTP	7.2																																									
6	<table border="1"><caption>Data for Gambar 6: Matriks Korelasi Antar Fitur</caption><thead><tr><th></th><th>ID</th><th>TPA</th><th>Usia</th><th>SK</th><th>Nilai</th></tr></thead><tbody><tr><th>ID</th><td>1.00</td><td>0.34</td><td>-0.98</td><td>0.29</td><td>0.15</td></tr><tr><th>TPA</th><td>0.34</td><td>1.00</td><td>-0.29</td><td>0.26</td><td>0.51</td></tr><tr><th>Usia</th><td>-0.98</td><td>-0.29</td><td>1.00</td><td>-0.28</td><td>-0.13</td></tr><tr><th>SK</th><td>0.29</td><td>0.26</td><td>-0.28</td><td>1.00</td><td>0.18</td></tr><tr><th>Nilai</th><td>0.15</td><td>0.51</td><td>-0.13</td><td>0.18</td><td>1.00</td></tr></tbody></table>		ID	TPA	Usia	SK	Nilai	ID	1.00	0.34	-0.98	0.29	0.15	TPA	0.34	1.00	-0.29	0.26	0.51	Usia	-0.98	-0.29	1.00	-0.28	-0.13	SK	0.29	0.26	-0.28	1.00	0.18	Nilai	0.15	0.51	-0.13	0.18	1.00	Gambar disamping menunjukkan keterkaitan antar feature. Bisa dilihat Tingkat Pendidikan Akhir dan Nilai yaitu 0.51 (cukup ada keterkaitan) meskipun tidak terlalu besar (keterkaitan sangat besar jika mendekati nilai 1). Hal ini dapat dipahami dari gambar 5 diatas yaitu semakin tinggi tingkat pendidikan akhir, bisa jadi semakin tinggi pula nilai pegawainya. Namun, pegawai lulusan SLTA dan SLTP (tingkat pendidikan rendah) pun juga cukup memiliki nilai pegawai yang tidak terlalu rendah.				
	ID	TPA	Usia	SK	Nilai																																					
ID	1.00	0.34	-0.98	0.29	0.15																																					
TPA	0.34	1.00	-0.29	0.26	0.51																																					
Usia	-0.98	-0.29	1.00	-0.28	-0.13																																					
SK	0.29	0.26	-0.28	1.00	0.18																																					
Nilai	0.15	0.51	-0.13	0.18	1.00																																					
7	<table border="1"><caption>Data for Gambar 7: Jumlah Pegawai per Usia</caption><thead><tr><th>Usia</th><th>Jumlah</th></tr></thead><tbody><tr><td>25</td><td>5</td></tr><tr><td>27</td><td>1</td></tr><tr><td>28</td><td>1</td></tr><tr><td>32</td><td>1</td></tr><tr><td>34</td><td>1</td></tr><tr><td>35</td><td>1</td></tr><tr><td>38</td><td>1</td></tr><tr><td>39</td><td>2</td></tr><tr><td>40</td><td>1</td></tr><tr><td>41</td><td>2</td></tr><tr><td>42</td><td>3</td></tr><tr><td>43</td><td>2</td></tr><tr><td>47</td><td>3</td></tr><tr><td>48</td><td>1</td></tr><tr><td>54</td><td>1</td></tr><tr><td>56</td><td>1</td></tr><tr><td>57</td><td>1</td></tr><tr><td>59</td><td>1</td></tr><tr><td>61</td><td>1</td></tr></tbody></table>	Usia	Jumlah	25	5	27	1	28	1	32	1	34	1	35	1	38	1	39	2	40	1	41	2	42	3	43	2	47	3	48	1	54	1	56	1	57	1	59	1	61	1	Jumlah pegawai yang berumur 25 tahun cukup banyak yakni berjumlah 5 orang, diikuti pegawai yang berumur 42 dan 47 orang sebanyak 3.
Usia	Jumlah																																									
25	5																																									
27	1																																									
28	1																																									
32	1																																									
34	1																																									
35	1																																									
38	1																																									
39	2																																									
40	1																																									
41	2																																									
42	3																																									
43	2																																									
47	3																																									
48	1																																									
54	1																																									
56	1																																									
57	1																																									
59	1																																									
61	1																																									

8	 <table><caption>Data for Gambar 8: Nilai vs Usia</caption><thead><tr><th>Usia</th><th>Nilai</th></tr></thead><tbody><tr><td>25</td><td>7.5</td></tr><tr><td>27</td><td>7.8</td></tr><tr><td>28</td><td>7.2</td></tr><tr><td>32</td><td>7.2</td></tr><tr><td>34</td><td>6.8</td></tr><tr><td>35</td><td>7.4</td></tr><tr><td>38</td><td>7.2</td></tr><tr><td>39</td><td>7.0</td></tr><tr><td>40</td><td>7.2</td></tr><tr><td>41</td><td>7.4</td></tr><tr><td>42</td><td>7.1</td></tr><tr><td>43</td><td>7.2</td></tr><tr><td>47</td><td>7.1</td></tr><tr><td>48</td><td>6.8</td></tr><tr><td>54</td><td>6.4</td></tr><tr><td>56</td><td>6.6</td></tr><tr><td>57</td><td>8.2</td></tr><tr><td>59</td><td>7.0</td></tr><tr><td>61</td><td>7.8</td></tr></tbody></table>	Usia	Nilai	25	7.5	27	7.8	28	7.2	32	7.2	34	6.8	35	7.4	38	7.2	39	7.0	40	7.2	41	7.4	42	7.1	43	7.2	47	7.1	48	6.8	54	6.4	56	6.6	57	8.2	59	7.0	61	7.8	<p>Gambar disamping memperlihatkan usia 57 dan 61 memiliki nilai pegawai yang paling tinggi. Kemungkinan karena sudah pegawai senior dan tentu saja memiliki pengalaman lebih dibanding dengan yang lain.</p>
Usia	Nilai																																									
25	7.5																																									
27	7.8																																									
28	7.2																																									
32	7.2																																									
34	6.8																																									
35	7.4																																									
38	7.2																																									
39	7.0																																									
40	7.2																																									
41	7.4																																									
42	7.1																																									
43	7.2																																									
47	7.1																																									
48	6.8																																									
54	6.4																																									
56	6.6																																									
57	8.2																																									
59	7.0																																									
61	7.8																																									
9	 <table><caption>Data for Gambar 9: Jumlah vs Nilai</caption><thead><tr><th>Nilai Range</th><th>Jumlah</th></tr></thead><tbody><tr><td>6.5 - 7.0</td><td>4</td></tr><tr><td>7.0 - 7.5</td><td>3</td></tr><tr><td>7.5 - 8.0</td><td>7</td></tr><tr><td>8.0 - 8.5</td><td>4</td></tr><tr><td>8.5 - 9.0</td><td>2</td></tr><tr><td>9.0 - 9.5</td><td>1</td></tr></tbody></table>	Nilai Range	Jumlah	6.5 - 7.0	4	7.0 - 7.5	3	7.5 - 8.0	7	8.0 - 8.5	4	8.5 - 9.0	2	9.0 - 9.5	1	<p>Jika dilihat dari data nilai bahwa pegawai memiliki nilai 7.25 paling banyak diantara yang lain.</p>																										
Nilai Range	Jumlah																																									
6.5 - 7.0	4																																									
7.0 - 7.5	3																																									
7.5 - 8.0	7																																									
8.0 - 8.5	4																																									
8.5 - 9.0	2																																									
9.0 - 9.5	1																																									

Pertanyaan:

1. Berikan rekomendasi bagaimana anda dapat membandingkan performa satu pegawai dengan pegawai yang lain?
2. Dapatkah anda menyimpulkan siapa pegawai yang memiliki performa yang sangat baik dan sebaliknya?

Jawab:

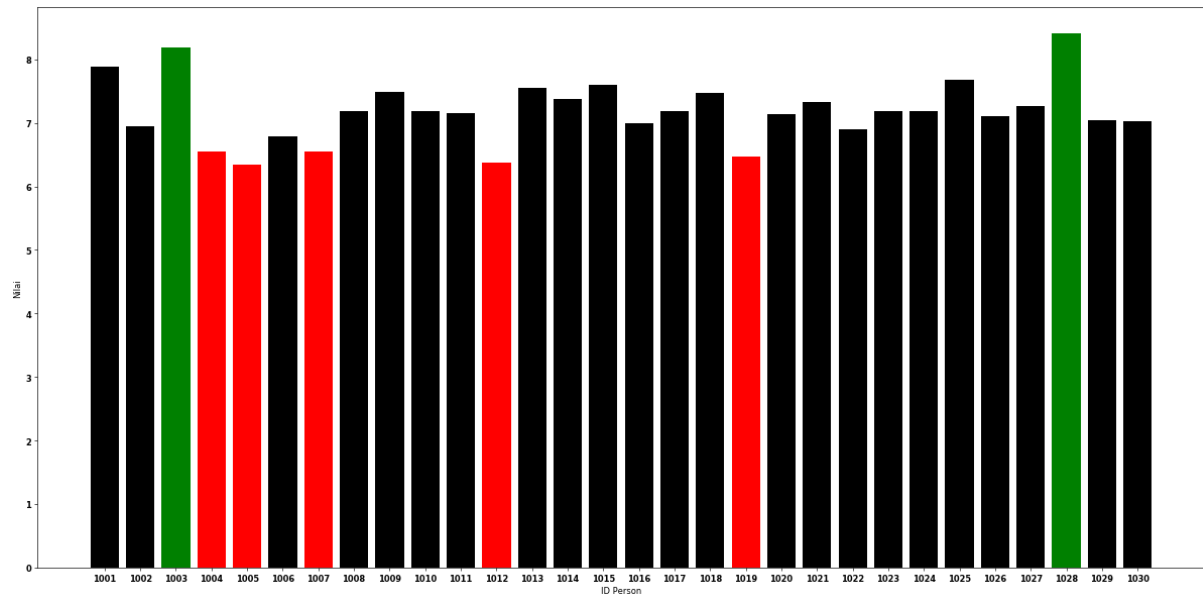
1. Berdasarkan hal-hal yang sudah saya visualisasikan diatas, saya merekomendasikan performa pegawai berdasarkan:
  - a. lulusan/tingkat pendidikan akhir semakin tinggi cukup sejalan dengan nilai performa pegawai. Hal ini dapat dilihat dari tingkat pendidikan akhir S1, S2 dan S3 yang memiliki nilai paling tinggi daripada tingkat pendidikan akhir yang lebih rendah seperti SD, SLTA.
  - b. Usia matang yakni 57 keatas mungkin memiliki nilai performansi yang cukup tinggi mengingat tingkat kesenioran dan kematangan dalam berkarir. Dalam kasus ini usia 57 dan 61 memiliki nilai performansi paling tinggi.
  - c. Badan D dan Dinas A memiliki nilai rata-rata performansi pegawai paling tinggi, namun Dinas B memiliki nilai rata-rata performansi pegawai paling rendah. Hal ini terjadi karena Badan D dan Dinas A diisi oleh pegawai lulusan S1 yang memiliki nilai cukup tinggi,



sedangkan Dinas B diisi oleh pegawai lulusan SD yang nilainya sangat rendah sehingga berkontribusi langsung kepada nilai rata-rata Dinas B.

2. Pegawai yang memiliki performansi paling baik adalah pegawai yang memiliki ID Person 1028(lulusan S1) dan 1003(lulusan S2) karena nilainya sangat tinggi (lebih dari 8), sedangkan pegawai yang memiliki performansi paling rendah(sekitar 6) yaitu pegawai yang memiliki ID Person 1004(lulusan SD), 1005(lulusan D3),1007(lulusan SLTA), 1012(lulusan D4) dan 1019(lulusan D4).

10.



## LAMPIRAN

Cara memplot data di Jupyter Notebook

Urutan per gambar dari gambar 1 sampai 10

1.

```
import pandas as pd
import matplotlib.pyplot as plt

product_data = ['DINAS B', 'DINAS A', 'BADAN D', 'BADAN B', 'DINAS P', 'DINAS F']
bug_data = [5, 5, 5, 5, 5, 5]

plt.pie(bug_data, labels=product_data, autopct='%1.1f%%', shadow=True, startangle=140)
plt.xlabel('Jumlah')

plt.show()
```

2.

```

# libraries
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import rc
import pandas as pd

# y-axis in bold
rc('font', weight='bold')

# Values of each group
SD = (0, 1, 0, 0, 0, 0)
SLTP = (0, 0, 1, 0, 0, 0)
SLTA = (1, 1, 0, 1, 1, 0)
D3 = (1, 2, 0, 2, 0, 0)
D4 = (0, 0, 1, 0, 1, 1)
S1 = (3, 0, 3, 2, 3, 3)
S2 = (0, 0, 0, 0, 0, 1)
S3 = (0, 1, 0, 0, 0, 0)

# Heights of bars1 + bars2
bars2 = np.add(SD, SLTP).tolist()
bars3 = np.add(bars2, SLTA).tolist()
bars4 = np.add(bars3, D3).tolist()
bars5 = np.add(bars4, D4).tolist()
bars6 = np.add(bars5, S1).tolist()
bars7 = np.add(bars6, S2).tolist()
bars8 = np.add(bars7, S3).tolist()

# The position of the bars on the x-axis
r = [0,1,2,3,4,5]

# Names of group and bar width
names = ['DINAS A', 'DINAS B', 'DINAS F', 'DINAS P', 'BADAN B', 'BADAN D']
barWidth = 1

# Create brown bars
plt.bar(r, SD, color='#7f6d5f', edgecolor='white', width=barWidth)
# Create green bars (middle), on top of the first ones
plt.bar(r, SLTP, bottom=SD, color='#557f2d', edgecolor='white', width=barWidth)
# Create green bars (top)
plt.bar(r, SLTA, bottom=bars2, color='#2d7f5e', edgecolor='white', width=barWidth)
plt.bar(r, D3, bottom=bars3, color='#873600', edgecolor='white', width=barWidth)
plt.bar(r, D4, bottom=bars4, color='#f4d03f', edgecolor='white', width=barWidth)
plt.bar(r, S1, bottom=bars5, color='#884ea0', edgecolor='white', width=barWidth)
plt.bar(r, S2, bottom=bars6, color='#fadbd8', edgecolor='white', width=barWidth)
plt.bar(r, S3, bottom=bars7, color='#f5cba7', edgecolor='white', width=barWidth)
plt.legend (('SD', 'SLTP', 'SLTA', 'D3', 'D4', 'S1', 'S2', 'S3'))

# Custom X axis
plt.xticks(r, names, fontweight='bold')
plt.xlabel("Satuan Kerja")
plt.ylabel('Jumlah Pegawai')

# Show graphic

```

3.

```

data = data.groupby(['SK'])['Nilai'].mean()
data.plot.bar()
plt.xlabel("Satuan Kerja")
plt.ylabel('Nilai')

```

4.

```

import seaborn as sns
sns.countplot(data['TPA']).set(xlabel='Tingkat Pendidikan Akhir', ylabel='Jumlah')

```

5.

```
data = data.groupby(['TPA'])['Nilai'].mean()
data.plot.bar()
plt.xlabel("Tingkat Pendidikan Akhir")
plt.ylabel('Nilai')
```

6.

```
import seaborn as sns
corr = data.corr()
sns.heatmap(corr,annot=True, fmt='.2f')
```

7.

```
import seaborn as sns
sns.countplot(data['Usia']).set(xlabel='Usia', ylabel='Jumlah')
data['Usia'].value_counts()
```

8.

```
data = data.groupby(['Usia'])['Nilai'].mean()
data.plot.bar()
plt.xlabel("Usia")
plt.ylabel('Nilai')
```

9.

```
data['Nilai'].hist(grid=False).set(xlabel='Nilai', ylabel='Jumlah')
```

10.

```

import matplotlib.pyplot as plt
width_in_inches = 20
height_in_inches = 10
dots_per_inch = 60

plt.figure(
    figsize=(width_in_inches, height_in_inches),
    dpi=dots_per_inch)
x = ['1001', '1002', '1003', '1004', '1005', '1006', '1007', '1008',
      '1009', '1010', '1011', '1012', '1013', '1014', '1015', '1016',
      '1017', '1018', '1019', '1020', '1021', '1022', '1023', '1024',
      '1025', '1026', '1027', '1028', '1029', '1030']
y = data['Nilai']

plt.xlabel("ID Person")
plt.ylabel('Nilai')

plt.bar(x, y, color=['black', 'black', 'green', 'red', 'red',
                    'black', 'red', 'black', 'black', 'black',
                    'black', 'red', 'black', 'black', 'black',
                    'black', 'black', 'black', 'red', 'black',
                    'black', 'black', 'black', 'black', 'black',
                    'black', 'black', 'green', 'black', 'black'])

plt.tight_layout()

```