In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
pd.__version__
```

Out[2]:
```
'1.0.1'
```

In [3]:
```
df = pd.read_csv('C:/Users/Baihaki/Downloads/datamining-master/datamining-master/Uas/dataset_soal No. 4.txt',
                 delimiter=',')
```

In [4]:
```
df
```

Out[4]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian
0	22	1	0	2	0	0
1	26	2	0	1	0	1
2	26	2	1	1	0	0
3	28	1	0	2	0	0
4	22	2	0	1	0	1
...	...	...	...	...	...	...
75	27	2	1	1	0	0
76	33	4	0	1	0	1
77	29	2	1	2	0	1
78	25	1	2	0	0	1
79	24	2	2	1	0	0

80 rows × 6 columns

In [6]:
```
import math
dis = []
for i in range(80):
    dis.append(math.sqrt((float(df.iloc[i]['Usia'])-29)**2+
                         (float(df.iloc[i]['Kelahiran_ke-'])- 2)**2+
                         (float(df.iloc[i]['Waktu_Kelahiran'])-0)**2+
                         (float(df.iloc[i]['Tekanan_darah'])-2)**2))
```

In [7]:
```
df['dis'] = dis
df
```

Out[7]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
0	22	1	0	2	0	0	7.071068
1	26	2	0	1	0	1	3.162278
2	26	2	1	1	0	0	3.316625
3	28	1	0	2	0	0	1.414214
4	22	2	0	1	0	1	7.071068
...	...	...	...	...	...	...	...
75	27	2	1	1	0	0	2.449490
76	33	4	0	1	0	1	4.582576
77	29	2	1	2	0	1	1.000000
78	25	1	2	0	0	1	5.000000
79	24	2	2	1	0	0	5.477226

80 rows × 7 columns

In [9]:
```
df.sort_values('dis')
```

Out[9]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
54	29	2	0	1	1	1	1.000000
77	29	2	1	2	0	1	1.000000
67	29	2	0	1	1	0	1.000000
59	30	2	1	2	1	1	1.414214
3	28	1	0	2	0	0	1.414214
...	...	...	...	...	...	...	...
61	19	1	0	1	0	1	10.099505
26	18	1	1	2	1	1	11.090537
31	40	1	0	1	1	1	11.090537
25	18	1	0	1	0	0	11.090537
70	17	1	0	0	0	1	12.206556

80 rows × 7 columns

In [11]:
```
y = df.sort_values('dis').head(5)
y
```

Out[11]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
54	29	2	0	1	1	1	1.000000
77	29	2	1	2	0	1	1.000000
67	29	2	0	1	1	0	1.000000
59	30	2	1	2	1	1	1.414214
3	28	1	0	2	0	0	1.414214

In [12]:
```
z = y["Caesarian"]
z
```

Out[12]:
```
54    1
77    1
67    0
59    1
3     0
Name: Caesarian, dtype: int64
```

In [14]:
```
np.mean(z)
```

Out[14]:
```
0.6
```

In [14]:
```
df.to_excel("C:/Users/Baihaki/Downloads/datamining-master/datamining-master/Uas/JawabaNo4b.xls")
```

In [ ]:
```

```