In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [3]:
```
pd.__version__
```

Out[3]:
```
'1.0.1'
```

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

In [5]:
```
df
```

Out[5]:

| | 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'])-30)**2+
                         (float(df.iloc[i]['Kelahiran_ke-'])- 1)**2+
                         (float(df.iloc[i]['Waktu_Kelahiran'])-0)**2+
                         (float(df.iloc[i]['Tekanan_darah'])-1)**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 | 8.062258 |
| **1** | 26 | 2 | 0 | 1 | 0 | 1 | 4.123106 |
| **2** | 26 | 2 | 1 | 1 | 0 | 0 | 4.242641 |
| **3** | 28 | 1 | 0 | 2 | 0 | 0 | 2.236068 |
| **4** | 22 | 2 | 0 | 1 | 0 | 1 | 8.062258 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **75** | 27 | 2 | 1 | 1 | 0 | 0 | 3.316625 |
| **76** | 33 | 4 | 0 | 1 | 0 | 1 | 4.242641 |
| **77** | 29 | 2 | 1 | 2 | 0 | 1 | 2.000000 |
| **78** | 25 | 1 | 2 | 0 | 0 | 1 | 5.477226 |
| **79** | 24 | 2 | 2 | 1 | 0 | 0 | 6.403124 |

80 rows × 7 columns

In [8]:
```
df.sort_values('dis')
```

Out[8]:

| | Usia | Kelahiran_ke- | Waktu_Kelahiran | Tekanan_darah | Kelainan_jantung | Caesarian | dis |
|---|---|---|---|---|---|---|---|
| **27** | 30 | 1 | 0 | 1 | 0 | 0 | 0.000000 |
| **38** | 31 | 1 | 0 | 1 | 0 | 0 | 1.000000 |
| **67** | 29 | 2 | 0 | 1 | 1 | 0 | 1.414214 |
| **54** | 29 | 2 | 0 | 1 | 1 | 1 | 1.414214 |
| **59** | 30 | 2 | 1 | 2 | 1 | 1 | 1.732051 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **41** | 19 | 1 | 0 | 1 | 0 | 1 | 11.000000 |
| **61** | 19 | 1 | 0 | 1 | 0 | 1 | 11.000000 |
| **25** | 18 | 1 | 0 | 1 | 0 | 0 | 12.000000 |
| **26** | 18 | 1 | 1 | 2 | 1 | 1 | 12.083046 |
| **70** | 17 | 1 | 0 | 0 | 0 | 1 | 13.038405 |

80 rows × 7 columns

In [9]:
```
y = df.sort_values('dis').head(5)
y
```

Out[9]:

| | Usia | Kelahiran_ke- | Waktu_Kelahiran | Tekanan_darah | Kelainan_jantung | Caesarian | dis |
|---|---|---|---|---|---|---|---|
| **27** | 30 | 1 | 0 | 1 | 0 | 0 | 0.000000 |
| **38** | 31 | 1 | 0 | 1 | 0 | 0 | 1.000000 |
| **67** | 29 | 2 | 0 | 1 | 1 | 0 | 1.414214 |
| **54** | 29 | 2 | 0 | 1 | 1 | 1 | 1.414214 |
| **59** | 30 | 2 | 1 | 2 | 1 | 1 | 1.732051 |

In [10]:
```
z = y["Caesarian"]
z
```

Out[10]:
```
27    0
38    0
67    0
54    1
59    1
Name: Caesarian, dtype: int64
```

In [11]:
```
np.mean(z)
```

Out[11]:
```
0.4
```

In [14]:
```
df.to_excel("C:/Users/Baihaki/Downloads/datamining-master/datamining-master/Uas/JawabaNo4.xls")
```

In [ ]:
```

```