Nama : Muhammad Azfar Wisnu
Kelas : 2IA02
NPM : 50420803
Tugas : Praktikum DGX

Instruksi Praktikum mahasiswa Tekrek : Teknik Industri, Teknik Mesin, Agroteknologi dan FTSP

- Transformasikan data pada kolom strength menjadi kolom baru bernama 'Category', dengan ketentuan :
  1. jika nilai >= 65 = 'Hard'
  2. jika nilai >= 40 = 'Medium'
  3. Jika nilai < 40 = 'Light'
- Lakukan analisis histogram untuk pengaruh kolom age terhadap kolom strength, apakah yang dapat Anda simpulkan ?
- Cari tahu komposisi concrete yang memiliki kekuatan tertinggi dan terendah
- Berikan kesimpulan akhir anda terhadap pengaruh komposisi concrete berdasarkan dataset yang digunakan

Muhammad azfar wisnu - 2ia02 - 50420803

```
[7] # Load data train dan test ke dalam pandas dataframe
    # dataset : https://gitlab.com/andreass.bayu/file-directory/-/raw/main/new_concreate.csv
    concrete = pd.read_csv("https://gitlab.com/andreass.bayu/file-directory/-/raw/main/new_concreate.csv")
```

```
concrete.head()
```

	cement	slag	ash	water	superplastic	coarseagg	fineagg	age	strength
0	141.3	212.0	NaN	203.5	NaN	971.8	748.5	28	29.89
1	168.9	42.2	124.3	158.3	10.8	1080.8	796.2	14	23.51
2	250.0	NaN	95.7	187.4	5.5	956.9	861.2	28	29.22
3	266.0	114.0	NaN	228.0	NaN	932.0	670.0	28	45.85
4	154.8	183.4	NaN	193.3	9.1	1047.4	696.7	28	18.29

```
[41] #cleaning
     concrete['ash'] = concrete['ash'].fillna(0)
     concrete['slag'] = concrete['slag'].fillna(0)
     concrete['superplastic'] = concrete['superplastic'].fillna(0)
     concrete.isnull().sum()
```

```
[41] cement          0
     slag            0
     ash             0
     water           0
     superplastic    0
     coarseagg       0
     fineagg         0
     age             0
     strength        0
     dtype: int64
```

Transformasikan data pada kolom strength menjadi kolom baru bernama 'Category', dengan ketentuan

```
[12] dataArray = []
     level = ""

     for val in concrete['strength']:
       if(val >= 65):
         level = "Hard"
       elif(val >= 40):
         level = "Medium"
       elif(val < 40):
         level = "Light"

       dataArray.append(level)

     print(dataArray)
```

[12]
```
['Light', 'Light', 'Light', 'Medium', 'Light', 'Light', 'Light', 'Light', 'Light', 'Light', 'Light', 'Li
```

[18]
```
df = pd.DataFrame(concrete)
newData = {
    'Category': dataArray,
}

df.assign(Category = dataArray)
```

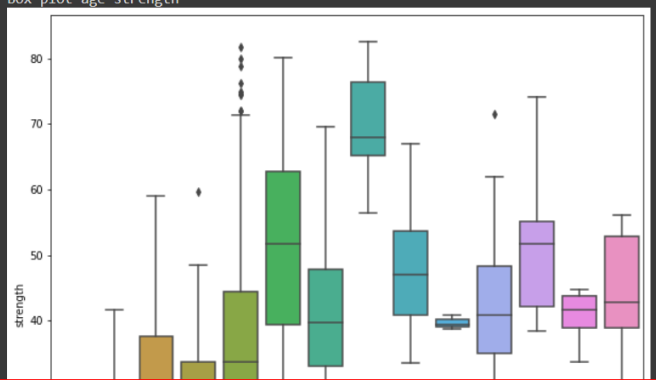
	cement	slag	ash	water	superplastic	coarseagg	fineagg	age	strength	Category
0	141.3	212.0	NaN	203.5	NaN	971.8	748.5	28	29.89	Light
1	168.9	42.2	124.3	158.3	10.8	1080.8	796.2	14	23.51	Light
2	250.0	NaN	95.7	187.4	5.5	956.9	861.2	28	29.22	Light
3	266.0	114.0	NaN	228.0	NaN	932.0	670.0	28	45.85	Medium
4	154.8	183.4	NaN	193.3	9.1	1047.4	696.7	28	18.29	Light
...	...	...	...	...	...	...	...	...	...	...
1025	135.0	NaN	166.0	180.0	10.0	961.0	805.0	28	13.29	Light
1026	531.3	NaN	NaN	141.8	28.2	852.1	893.7	3	41.30	Medium
1027	276.4	116.0	90.3	179.6	8.9	870.1	768.3	28	44.28	Medium
1028	342.0	38.0	NaN	228.0	NaN	932.0	670.0	270	55.06	Medium

✓ 0 d selesai pada 11.47

Lakukan analisis histogram untuk pengaruh kolom age terhadap kolom strength, apakah yang dapat Anda simpulkan ?

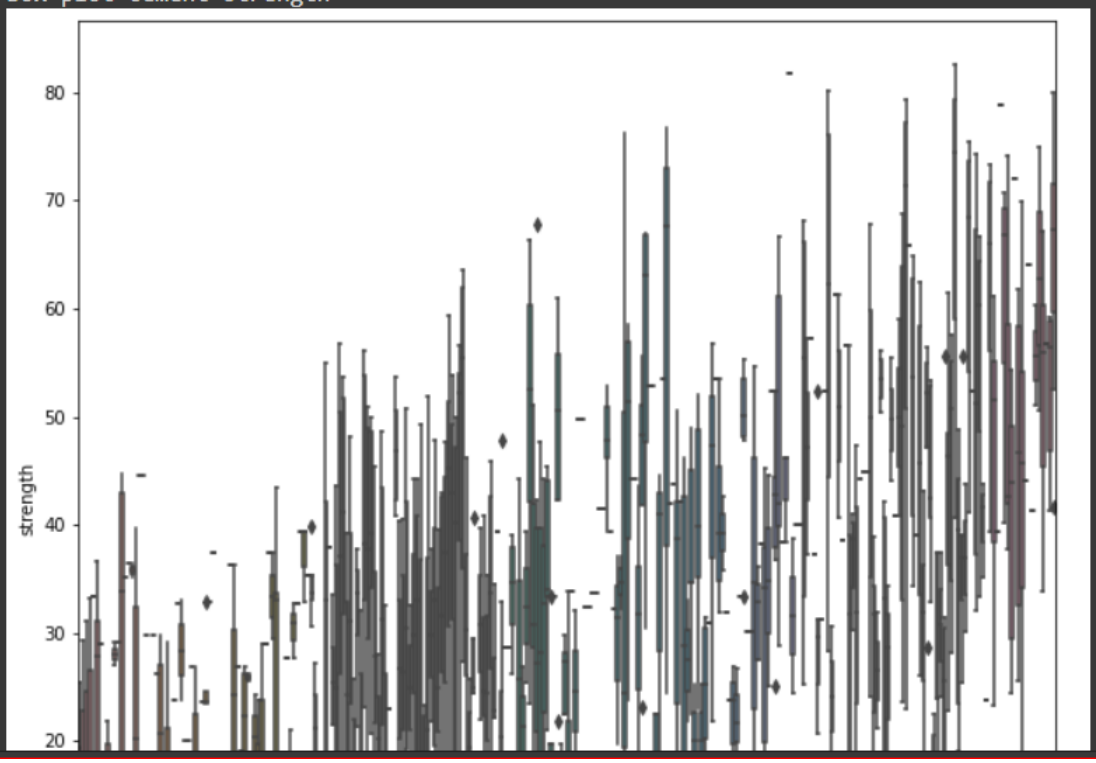
[22]
```
#strength dengan menggunakan age
print("Box plot", "age", "strength")
fig = plt.figure(figsize=(10,10))
sns.boxplot(x="age", y="strength", data=concrete)
plt.show()
```

```
Box plot age strength
```

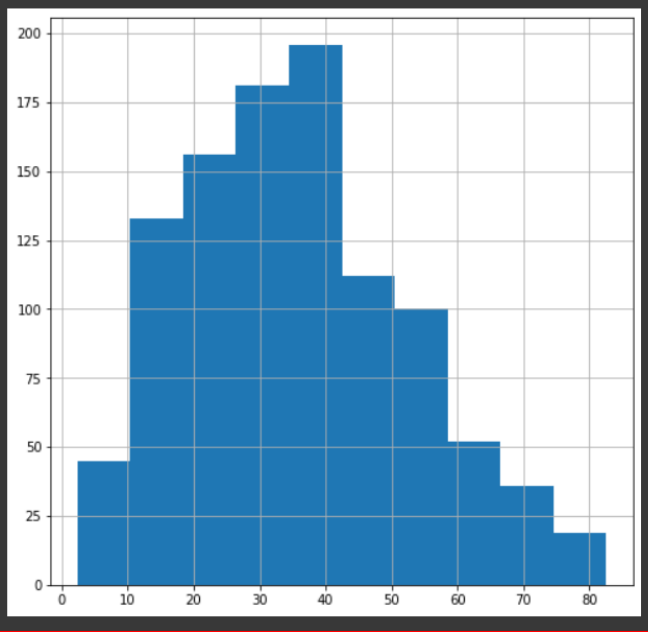


```
[26] #strength dengan tanpa menggunakan age
     print("Box plot", "cement", "strength")
     fig = plt.figure(figsize=(10,10))
     sns.boxplot(x="cement", y="strength", data=concrete)
     plt.show()
```

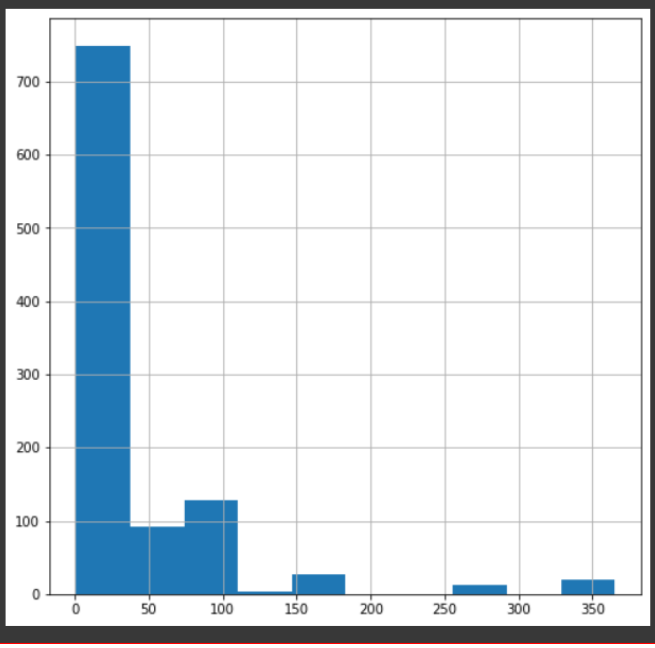
```
Box plot cement strength
```



```
[39] concrete['strength'].hist(figsize=(8,8))
     plt.show()
```




```
concrete['age'].hist(figsize=(8,8))
plt.show()
```



Cari tahu komposisi concrete yang memiliki kekuatan tertinggi dan terendah

```
[38] concrete.min()
```

```
cement          102.00
slag             11.00
ash              24.50
water           121.80
superplastic      1.70
coarseagg       801.00
fineagg         594.00
age               1.00
strength          2.33
dtype: float64
```

```
[37] concrete.max()
```

```
cement          540.0
slag            359.4
ash             200.1
water           247.0
superplastic     32.2
coarseagg      1145.0
fineagg         992.6
age             365.0
strength         82.6
```

kesimpulan: semua panduan tercampur satu menjadi sebuah object yang ada pada data