

Nama : Elmo Alistair Heriyanto

NPM : 12118220

Kelas : 4KA17

```
+ Code + Text

1. Silahkan modifikasi kode operasi yang ada menggunakan library perhitungan berbasis GPU (Library Cupy)

# import libraries
import pandas as pd
import cupy as cp
import seaborn as sns
import matplotlib.pyplot as plt
from pandas.plotting import autocorrelation_plot
from scipy import stats
plt.style.use("ggplot")
import warnings
warnings.filterwarnings("ignore")
from scipy import stats

2. Bacalah dataset yang berada tersimpan url https://raw.githubusercontent.com/supasonicx/ATA-praktikum-01/main/concrete.csv

[ ] df = pd.read_csv('https://raw.githubusercontent.com/supasonicx/ATA-praktikum-01/main/concrete.csv')
df.head()
```

	cement	slag	ash	water	superplastic	coarseagg	fineagg	age	strength
0	141.3	212.0	0.0	203.5	0.0	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

```
+ Code + Text

[ ] df.head()
```

2	250.0	0.0	95.7	187.4		5.5	956.9	861.2	28	29.22
3	266.0	114.0	0.0	228.0		0.0	932.0	670.0	28	45.85
4	154.8	183.4	0.0	193.3		9.1	1047.4	696.7	28	18.29

```
[ ] df.shape

(1030, 9)

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1030 entries, 0 to 1029
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  -
0   cement          1030 non-null  float64
1   slag            1030 non-null  float64
2   ash             1030 non-null  float64
3   water           1030 non-null  float64
4   superplastic    1030 non-null  float64
5   coarseagg       1030 non-null  float64
6   fineagg         1030 non-null  float64
7   age             1030 non-null  int64
8   strength        1030 non-null  float64
dtypes: float64(8), int64(1)
memory usage: 72.5 KB
```

3. Periksa dataset apakah terdapat data yang bernilai null dengan menggunakan fungsi `.isnull()`

```
df.isnull().sum()
```

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

```
[ ] df.duplicated().sum()
```

```
25
```

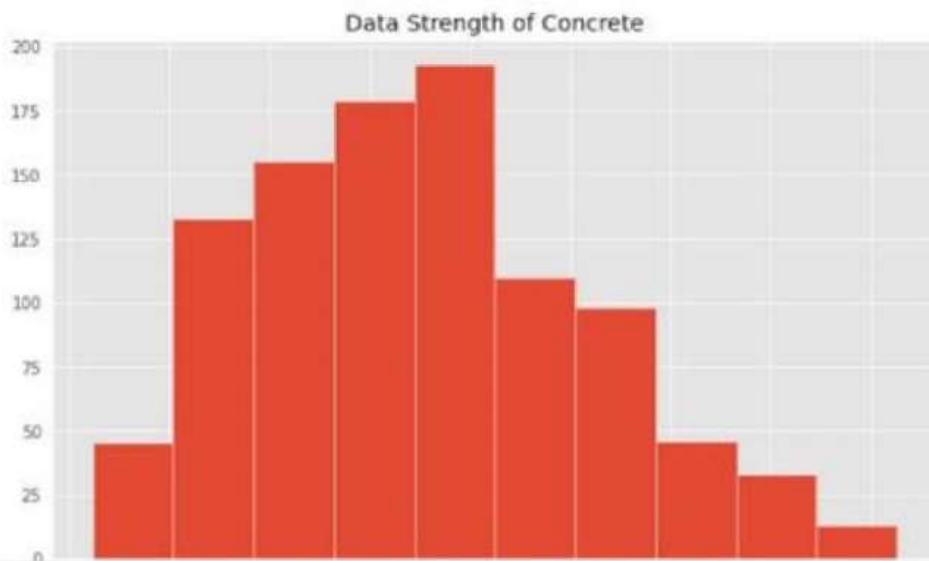
```
[ ] df.drop_duplicates(inplace=True)
```

```
[ ] df.duplicated().sum()
```

```
0
```

4. Buatlah sebuah histogram dari data kolom 'strength'.

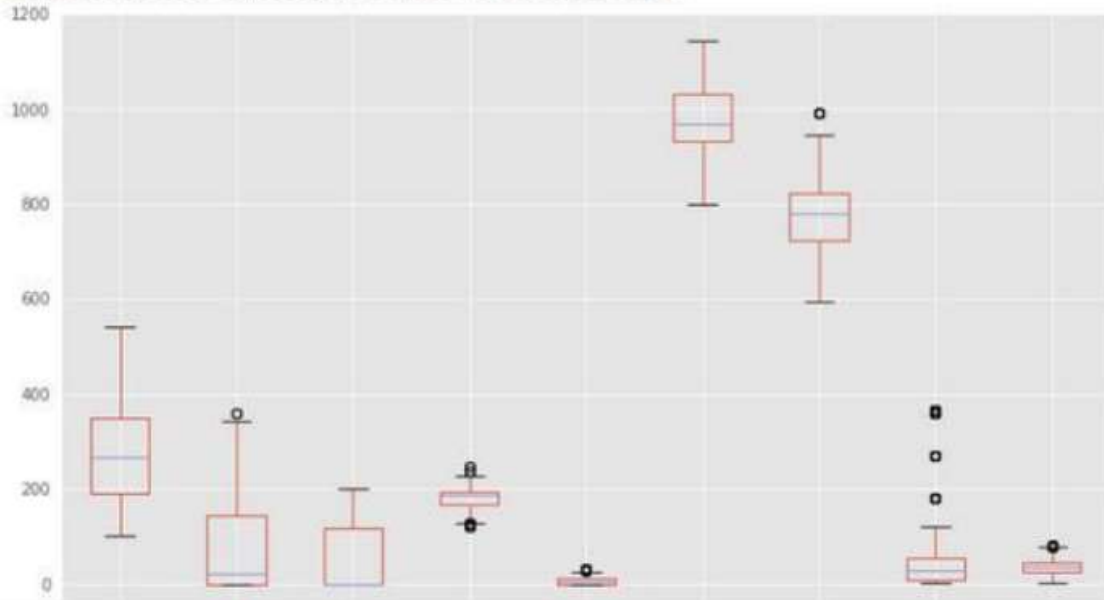
```
[ ] plt.figure(figsize=(10,6))
plt.hist(df['strength'])
plt.title("Data Strength of Concrete")
plt.xlabel('Strength')
plt.show()
```



5. Buatlah diagram boxplot dari dataset yang ada.

```
plt.figure(figsize=(12,7))
df.boxplot()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f62bf6449d0>
```



6. Hitung karakteristik statistik (standar deviasi, variance, mean, median) dari masing-masing kolom data.

```
std = pd.DataFrame(df.std())
var = pd.DataFrame(df.var())
mean = pd.DataFrame(df.mean())
median = pd.DataFrame(df.median())
data = pd.DataFrame(std)
data[1] = var
data[2] = mean
data[3] = median
data.columns = ['Standar Deviasi', 'Variance', 'Mean', 'Median']
data
```

	Standar Deviasi	Variance	Mean	Median
cement	104.344261	10887.724744	278.631343	265.0
slag	86.170807	7425.407978	72.043483	20.0
ash	64.207969	4122.663231	55.536318	0.0
water	21.339334	455.367179	182.075323	185.7
superplastic	5.919967	35.046006	6.033234	6.1
coarseagg	77.579667	6018.604691	974.376816	968.0
fineagg	80.340435	6454.585440	772.688259	780.0

7. Buatlah correlation map dari dataset tersebut.

```
[ ] sns.set({'figure.figsize':(10,8)})
sns.heatmap(df.corr(), annot=True)
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f62b54b9ed0>



8. Hitung covariance dari kolom data yang diminta

```
[ ] df.var()
```

```
cement      10887.724744
slag        7425.407978
ash         4122.663231
water       455.367179
superplastic  35.046006
coarseagg   6018.604691
fineagg     6454.585440
age         4062.110923
strength    265.195212
dtype: float64
```

9. Hitung pearson correlation dan spearsman correlation dari kolom data yang diminta

```
[ ] print('Pearson Correlation Dari Variabel Age dan Strength Pada Data Kekuatan Beton')
print(stats.pearsonr(df['age'],df['strength']))
print('-'*80)
print('Spearman Correlation Dari Variabel Age dan Strength Pada Data Kekuatan Beton = ')
print(stats.spearmanr(df['age'],df['strength']))
```

```
Pearson Correlation Dari Variabel Age dan Strength Pada Data Kekuatan Beton
(0.3373669316044723, 3.5719493008704364e-28)
```

```
Spearman Correlation Dari Variabel Age dan Strength Pada Data Kekuatan Beton
```

10. Hitung nilai hipotesis testing untuk kolom age dan strength.

```
[ ] print('Berikut adalah rata-rata dari data age(umur) pada dataset kekuatan beton = ',cp.mean(df.age))
    print('Berikut adalah rata-rata dari data strength(kekuatan) pada dataset kekuatan beton = ',cp.mean(df.strength))

    print('-'*80)

    result = stats.ttest_ind(df['age'],df['strength'])
```

```
Berikut adalah rata-rata dari data age(umur) pada dataset kekuatan beton = 45.85671641791045
Berikut adalah rata-rata dari data strength(kekuatan) pada dataset kekuatan beton = 35.25037810945276
-----
```

```
[ ] if result[1] < result[0]:
    print('Hipotesis Diterima')
else:
    print('Hipotesis Ditolak')
```

```
Hipotesis Diterima
```

Uji Hipotesis :

H0 = Rata-rata Umur dan kekuatan memiliki kedudukan yang sama atau setara pada dataset kekuatan beton

H1 = Rata-rata Umur dan kekuatan memiliki kedudukan yang tidak sama atau tidak setara pada dataset kekuatan beton Hasil : Hipotesis Diterima