⭐ Regression Model to Predict Cement Compressive Strength


```
# import library
import pandas as pd
import numpy as np
```

```
# import data
cement = pd.read_csv('https://github.com/ybifoundation/Dataset/raw/main/Concrete%20Compressive%20Strength.csv')
```

```
# view data
import numpy as np
```

```
# info of data
import numpy as np
```

```
# summary statistics
```

	Cement (kg in a m^3 mixture)	Blast Furnace Slag (kg in a m^3 mixture)	Fly Ash (kg in a m^3 mixture)	Water (kg in a m^3 mixture)	Superplasticizer (kg in a m^3 mixture)	Coarse Aggregate (kg in a m^3 mixture)	Fine Aggregate (kg in a m^3 mixture)	Age (day)	Concr Compress Strength(MI megapasca
count	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000(
mean	281.165631	73.895485	54.187136	181.566359	6.203112	972.918592	773.578883	45.662136	35.817
std	104.507142	86.279104	63.996469	21.355567	5.973492	77.753818	80.175427	63.169912	16.705
min	102.000000	0.000000	0.000000	121.750000	0.000000	801.000000	594.000000	1.000000	2.331
25%	192.375000	0.000000	0.000000	164.900000	0.000000	932.000000	730.950000	7.000000	23.707
50%	272.900000	22.000000	0.000000	185.000000	6.350000	968.000000	779.510000	28.000000	34.442
75%	350.000000	142.950000	118.270000	192.000000	10.160000	1029.400000	824.000000	56.000000	46.136
max	540.000000	359.400000	200.100000	247.000000	32.200000	1145.000000	992.600000	365.000000	82.599

```
# check for missing value
cement.describe()
```

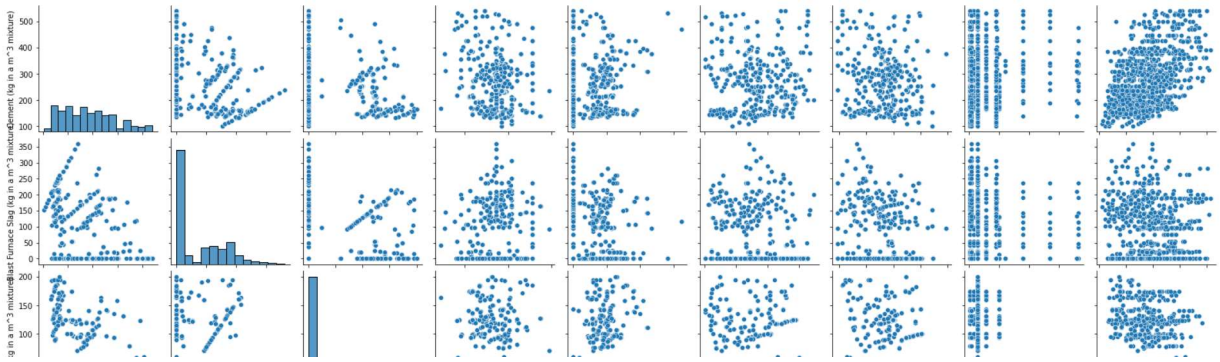
	Cement (kg in a m^3 mixture)	Blast Furnace Slag (kg in a m^3 mixture)	Fly Ash (kg in a m^3 mixture)	Water (kg in a m^3 mixture)	Superplasticizer (kg in a m^3 mixture)	Coarse Aggregate (kg in a m^3 mixture)	Fine Aggregate (kg in a m^3 mixture)	Age (day)	Concr Compress Strength(M megapasca
count	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000000	1030.000
mean	281.165631	73.895485	54.187136	181.566359	6.203112	972.918592	773.578883	45.662136	35.817
std	104.507142	86.279104	63.996469	21.355567	5.973492	77.753818	80.175427	63.169912	16.705
min	102.000000	0.000000	0.000000	121.750000	0.000000	801.000000	594.000000	1.000000	2.331

```
# check for categories
cement.nunique()
```

```
Cement (kg in a m^3 mixture)                        280
Blast Furnace Slag (kg in a m^3 mixture)            187
Fly Ash (kg in a m^3 mixture)                       163
Water (kg in a m^3 mixture)                         205
Superplasticizer (kg in a m^3 mixture)              155
Coarse Aggregate (kg in a m^3 mixture)              284
Fine Aggregate (kg in a m^3 mixture)                304
Age (day)                                            14
Concrete Compressive Strength(MPa, megapascals)     938
dtype: int64
```

```
# visualize pairplot
import seaborn as sns
sns.pairplot(cement)
```

```
<seaborn.axisgrid.PairGrid at 0x7f2eb8c9fbb0>
```



```
# columns name
cement.columns
```

```
Index(['Cement (kg in a m^3 mixture)',
       'Blast Furnace Slag (kg in a m^3 mixture)',
       'Fly Ash (kg in a m^3 mixture)', 'Water (kg in a m^3 mixture)',
       'Superplasticizer (kg in a m^3 mixture)',
       'Coarse Aggregate (kg in a m^3 mixture)',
       'Fine Aggregate (kg in a m^3 mixture)', 'Age (day)',
       'Concrete Compressive Strength(MPa, megapascals) '],
      dtype='object')
```

```
# define y
y=cement['Concrete Compressive Strength(MPa, megapascals) ']
```

```
# define X
X=cement[['Cement (kg in a m^3 mixture)',
'Blast Furnace Slag (kg in a m^3 mixture)',
'Fly Ash (kg in a m^3 mixture)', 'Water (kg in a m^3 mixture)',
'Superplasticizer (kg in a m^3 mixture)',
'Coarse Aggregate (kg in a m^3 mixture)',
'Fine Aggregate (kg in a m^3 mixture)', 'Age (day)']]
```

```
# split data
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y,train_size=0.7,random_state=2559)
```

```
# verify shape
X_train.shape,X_test.shape,y_train.shape,y_test.shape
```

```
((721, 8), (309, 8), (721,), (309,))
```

```
# select model
from sklearn.linear_model import LinearRegression
model=LinearRegression()
```

```
# train model
model.fit(X_train,y_train)
```

```
LinearRegression()
```

```
# predict with model
y_pred=model.predict(X_test)
```

```
# model evaluation
from sklearn.metrics import mean_absolute_error,mean_absolute_percentage_error,mean_squared_error
```

```
# model MAE
mean_absolute_error(y_test,y_pred)
```

```
7.814891951068712
```

```
# model MAPE
mean_absolute_percentage_error(y_test,y_pred)
```

```
0.28040027489426594
```

```
# model MSE
mean_squared_error(y_test,y_pred)
```

```
102.62674212692517
```

```
# future prediction
X.sample()
```

	Cement (kg in a m^3 mixture)	Blast Furnace Slag (kg in a m^3 mixture)	Fly Ash (kg in a m^3 mixture)	Water (kg in a m^3 mixture)	Superplasticizer (kg in a m^3 mixture)	Coarse Aggregate (kg in a m^3 mixture)	Fine Aggregate (kg in a m^3 mixture)	Age (day)
917	148.0	175.0	0.0	171.0	2.0	1000.0	828.0	28

```
# define X_new
X_new=X.sample()
X_new
```

	Cement (kg in a m^3 mixture)	Blast Furnace Slag (kg in a m^3 mixture)	Fly Ash (kg in a m^3 mixture)	Water (kg in a m^3 mixture)	Superplasticizer (kg in a m^3 mixture)	Coarse Aggregate (kg in a m^3 mixture)	Fine Aggregate (kg in a m^3 mixture)	Age (day)
392	333.0	17.5	163.0	167.0	17.9	996.0	652.0	28

```
# predict for X_new
model.predict(X_new)
```

```
array([48.77143516])
```