

```
In [1]: %matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import plotly.express as px
from IPython.display import display

In [2]: df = pd.read_excel("Random Price Dataset Trial Assignment.xlsx")
df.head()
```

Out[2]:

	Date	Domestic Market (Contract) Blow Molding, Low	Spot/Export Blow Molding	Spot, Domestic	WTISPLC	MCOILBRENTU	GASREGM	IMPCH	EXPCH	PRUBBUSDM	...	Russia_import	South_Africa_import	Turkey
0	2000-01-01	41.0	NaN	NaN	27.18	25.51	1.289	6902.1	863.1	29.207387	...	NaN	NaN	NaN
1	2000-02-01	41.0	NaN	NaN	29.35	27.78	1.377	6584.4	972.7	33.391099	...	NaN	NaN	NaN
2	2000-03-01	45.0	NaN	NaN	29.89	27.49	1.516	6424.1	1330.5	30.941913	...	NaN	NaN	NaN
3	2000-04-01	47.0	NaN	NaN	25.74	22.76	1.465	7070.5	1227.5	31.930148	...	NaN	NaN	NaN
4	2000-05-01	47.0	NaN	NaN	28.78	27.74	1.487	7850.2	1526.3	31.201702	...	NaN	NaN	NaN

5 rows × 50 columns

```
In [3]: df.isnull().sum()
```

```

Out[3]: Date
0
Domestic Market (Contract) Blow Molding, Low
0
Spot/Export Blow Molding
91
Spot, Domestic
51
WTISPLC
0
MCOILBRENTU
0
GASREGM
0
IMPCH
1
EXPCH
1
PRUBBUSDM
1
WPUFD4111
1
PCU325211325211
1
PCU32611332611301
1
WPU0915021625
44
PCU3252132521
48
MHHNGSP
0
WPU072205011
61
PCU32611132611115
61
PCU32611332611301.1
72
PCU32611132611112
61
WPU0915021622
43
Producer Price Index by Industry: Plastics Material and Resins Manufacturing: Thermoplastic Resins and Plastics Materials
0
Australia _export
38
Canada_export
22
Saudi_export
43
Usa_export
22
India_export
34
Russia_export
68
South_Africa_export
22
Turkey
4
Brazil
1
France_export
36
Germeny_export
98
United Kingdome_export
23
China_export
68
Australia _import
67
Canada_import
22
Saudi_import
43
Usa_import
22
India_import
34
Russia_import
68
South_Africa_import
22
Turkey_import
08

```

```

Brazil_import
08
France_import
36
Germany_import
98
United Kingdom_import
22
China_import
68
Japan_import
23
South_korea_import
76
dtype: int64

```

1

1

1

1

2

In [4]: df.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 276 entries, 0 to 275
Data columns (total 50 columns):
#   Column
Non-Null Count  Dtype
---  -
0   Date
276 non-null    datetime64[ns]
1   Domestic Market (Contract) Blow Molding, Low
276 non-null    float64
2   Spot/Export Blow Molding
185 non-null    float64
3   Spot, Domestic
125 non-null    float64
4   WTISPLC
276 non-null    float64
5   MCOILBRENTU
276 non-null    float64
6   CACBREM

```

In [ ]: `###df = df.drop(['Date'], axis = 1)`

In [5]: df['Date']=pd.to\_datetime(df['Date'])

In [6]: df.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 276 entries, 0 to 275
Data columns (total 50 columns):
#   Column
Non-Null Count  Dtype
---  -
0   Date
276 non-null    datetime64[ns]
1   Domestic Market (Contract) Blow Molding, Low
276 non-null    float64
2   Spot/Export Blow Molding
185 non-null    float64
3   Spot, Domestic
125 non-null    float64
4   WTISPLC
276 non-null    float64
5   MCOILBRENTU
276 non-null    float64
6   CACBREM

```

In [7]: df = df.fillna(df.mean())

```

C:\Users\DELL\AppData\Local\Temp\ipykernel_10752\114435927.py:1: FutureWarning: DataFrame.mean and DataFrame.median with numeri
c_only=None will include datetime64 and datetime64tz columns in a future version.
  df = df.fillna(df.mean())
C:\Users\DELL\AppData\Local\Temp\ipykernel_10752\114435927.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reduc
tions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns befor
e calling the reduction.
  df = df.fillna(df.mean())

```

```
In [8]: df.isnull().sum()
```

localhost:8889/notebooks/01ML/LAZY PRE DL1.ipynb

```

Brazil_import
0
France_import
0
Germany_import
0
United Kingdom_import
0
China_import
0
Japan_import
0
South_korea_import
76
dtype: int64

```

2

```
In [9]: df = df.drop(['Australia_export', 'Saudi_export', 'South_korea_import'], axis=1)
```

```
In [10]: df.isnull().sum()
```

```

Out[10]: Date                                0
Domestic Market (Contract) Blow Molding, Low  0
Spot/Export Blow Molding                      0
Spot, Domestic                               0
WTISPLC                                       0
MCOILBRENTU                                 0
GASREGM                                      0
IMPCH                                        0
EXPCH                                        0
PRUBBUSDM                                   0
WPUFD4111                                   0
PCU325211325211                            0
PCU32611332611301                          0
WPU0915021625                              0
PCU3252132521                              0
MHHNGSP                                     0
WPU072205011                               0
PCU32611132611115                          0
PCU32611332611301.1                        0
PCU32611132611112                          0
WPU0915021622                              0
Producer Price Index by Industry: Plastics Material and Resins Manufacturing: Thermoplastic Resins and Plastics Materials  0
Canada_export                               0
Usa_export                                 0
India_export                              0
Russia_export                             0
South_Africa_export                       0
Turkey                                   0
Brazil                                   0
France_export                             0
Germany_export                           0
United Kingdom_export                   0
China_export                             0
Australia_import                        0
Canada_import                           0
Saudi_import                            0
Usa_import                              0
India_import                            0
Russia_import                           0
South_Africa_import                     0
Turkey_import                           0
Brazil_import                           0
France_import                           0
Germany_import                           0
United Kingdom_import                   0
China_import                             0
Japan_import                             0
dtype: int64

```

```
In [11]: from sklearn import preprocessing
```

```
In [12]: Standardisation = preprocessing.StandardScaler()
```

```

In [13]: import lazypredict
from sklearn.model_selection import train_test_split
from lazypredict.Supervised import LazyRegressor

```

```
In [14]: X = df.drop(["Domestic Market (Contract) Blow Molding, Low"], axis=1)
Y = df["Domestic Market (Contract) Blow Molding, Low"]
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size = 0.2, random_state = 64)
reg = LazyRegressor(verbose=0, ignore_warnings=False, custom_metric=None)
models, pred = reg.fit(X_train, X_test, y_train, y_test)
models
```

100%|██████████| 42/42 [00:06<00:00, 6.63it/s]

Out[14]:

	Adjusted R-Squared	R-Squared	RMSE	Time Taken
Model				
ExtraTreesRegressor	0.87	0.98	2.89	0.42
TransformedTargetRegressor	0.84	0.97	3.22	0.01
LinearRegression	0.84	0.97	3.22	0.02
HistGradientBoostingRegressor	0.83	0.97	3.28	0.53
LassoCV	0.82	0.97	3.43	0.27
LassoLarsCV	0.82	0.97	3.43	0.13
LassoLarsIC	0.82	0.97	3.43	0.05
Ridge	0.82	0.97	3.45	0.03
HuberRegressor	0.81	0.97	3.47	0.09
RandomForestRegressor	0.81	0.97	3.50	1.00
GradientBoostingRegressor	0.80	0.97	3.58	0.64
ElasticNetCV	0.79	0.97	3.64	0.33
BayesianRidge	0.79	0.97	3.66	0.05
LGBMRegressor	0.79	0.96	3.71	0.10
PoissonRegressor	0.78	0.96	3.74	0.08
RidgeCV	0.78	0.96	3.76	0.02
SGDRegressor	0.78	0.96	3.80	0.02
AdaBoostRegressor	0.76	0.96	3.90	0.26
BaggingRegressor	0.72	0.95	4.26	0.11
LinearSVR	0.70	0.95	4.38	0.02
Lasso	0.68	0.95	4.51	0.03
LassoLars	0.68	0.95	4.51	0.03
XGBRegressor	0.65	0.94	4.76	0.22
KNeighborsRegressor	0.60	0.93	5.09	0.15
ElasticNet	0.58	0.93	5.22	0.02
TweedieRegressor	0.53	0.92	5.51	0.04
GammaRegressor	0.50	0.92	5.65	0.05
ExtraTreeRegressor	0.50	0.92	5.65	0.02
PassiveAggressiveRegressor	0.42	0.90	6.12	0.02
OrthogonalMatchingPursuitCV	0.42	0.90	6.13	0.03
DecisionTreeRegressor	0.39	0.90	6.26	0.02
OrthogonalMatchingPursuit	0.37	0.90	6.35	0.02
RANSACRegressor	0.14	0.86	7.42	0.29
SVR	-0.16	0.81	8.64	0.02
NuSVR	-0.47	0.76	9.73	0.03
LarsCV	-0.93	0.68	11.14	0.08
MLPRegressor	-3.84	0.21	17.66	0.49
DummyRegressor	-5.13	-0.00	19.87	0.01
QuantileRegressor	-5.14	-0.01	19.89	0.48
GaussianProcessRegressor	-59.39	-8.88	62.38	0.04
KernelRidge	-80.13	-12.28	72.30	0.02
Lars	-302060.26	-49427.21	4411.64	0.02

```
In [ ]:
```



