

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from typing import Tuple, List, Optional, Dict
from sklearn.model_selection import train_test_split, cross_val_score, GridSearchCV
from sklearn.preprocessing import StandardScaler, MinMaxScaler, RobustScaler,
PowerTransformer, OneHotEncoder, LabelEncoder
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.impute import SimpleImputer
from sklearn.pipeline import Pipeline
from sklearn.compose import ColumnTransformer
from sklearn.feature_selection import VarianceThreshold, SelectKBest, f_classif, f_regression, RFE
from sklearn.decomposition import PCA
from sklearn.linear_model import LogisticRegression, LinearRegression, LassoCV
from sklearn.svm import SVC, SVR
from sklearn.neighbors import KNeighborsClassifier, KNeighborsRegressor
from sklearn.tree import DecisionTreeClassifier, DecisionTreeRegressor
from sklearn.ensemble import RandomForestClassifier, RandomForestRegressor,
GradientBoostingClassifier, GradientBoostingRegressor
from sklearn.naive_bayes import GaussianNB
from sklearn.cluster import KMeans
from sklearn.metrics import (accuracy_score, precision_score, recall_score, f1_score,
roc_auc_score, roc_curve,
confusion_matrix, classification_report, mean_squared_error, mean_absolute_error, r2_score)

```

pd.read_csv()	np.log()	MeanShift()
pd.read_excel()	np.log1p()	mean_squared_error()
pd.read_json()	np.sqrt()	mean_absolute_error()
pd.read_html()	np.cbrt()	r2_score()
pd.read_sql()	PowerTransformer()	mean_absolute_percentage_err
pd.read_parquet()	BoxCox() (from scipy.stats)	or()
pd.read_pickle()	df.corr()	mean_squared_log_error()
np.loadtxt()	df.cov()	accuracy_score()
np.genfromtxt()	sns.heatmap()	precision_score()
np.load()	SelectKBest()	recall_score()
df.head()	chi2()	f1_score()
df.info()	f_classif()	classification_report()
df.describe()	mutual_info_classif()	confusion_matrix()

df.isnull()	Lasso() / LassoCV()	GradientBoostingClassifier()
df.isna()	Ridge()	GaussianNB()
df.fillna()	ElasticNet()	BernoulliNB()
df.dropna()	RandomForestClassifier()	MultinomialNB()
df.duplicated()	RandomForestRegressor()	XGBClassifier()
df.drop_duplicates()	GradientBoostingClassifier()	KMeans()
df.replace()	OneHotEncoder()	AgglomerativeClustering()
df.astype()	LabelEncoder()	DBSCAN()
SimpleImputer()	OrdinalEncoder()	roc_auc_score()
KNNImputer()	pd.cut()	roc_curve()
LabelEncoder()	pd.qcut()	train_test_split()
OneHotEncoder()	PolynomialFeatures()	cross_val_score()
OrdinalEncoder()	PowerTransformer()	StratifiedKFold()
ColumnTransformer()	Normalizer()	KFold()
Pipeline()	df['col'].dt.year	GridSearchCV()
StandardScaler()	df['col'].dt.month	RandomizedSearchCV()
MinMaxScaler()	df['col'].dt.day	plt.plot()
RobustScaler()	LinearRegression()	plt.scatter()
MaxAbsScaler()	Ridge()	plt.bar()
Normalizer()	Lasso()	plt.hist()
PowerTransformer()	ElasticNet()	plt.boxplot()
QuantileTransformer()	SVR()	plt.show()
df.describe()	DecisionTreeRegressor()	sns.heatmap()
df.quantile()	RandomForestRegressor()	sns.pairplot()
np.percentile()	GradientBoostingRegressor()	sns.countplot()
zscore() (from scipy.stats)	KNeighborsRegressor()	sns.boxplot()
IsolationForest()	XGBRegressor()*	sns.distplot() (deprecated)
LocalOutlierFactor()	LogisticRegression()	sns.histplot()
df.skew()	SVC()	sns.violinplot()
mutual_info_regression()	KNeighborsClassifier()	sns.scatterplot()
VarianceThreshold()	DecisionTreeClassifier()	sns.lineplot()
SelectFromModel()	RandomForestClassifier()	

