In [1]:
```python
# Multiple linear regression

# Importing the libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

In [2]:
```python
# Importing the dataset
dataset = pd.read_csv("/home/forbidden_devil/Machine Le
arning A-Z/Part 2 - Regression/Section 5 - Multiple Lin
ear Regression/50_Startups.csv")
X = dataset.iloc[:, :-1].values
y = dataset.iloc[:, 4].values
```

In [3]:
```python
# Encoding categorical data

from sklearn.preprocessing import LabelEncoder, OneHotE
ncoder
labelencoder = LabelEncoder()
X[:, 3] = labelencoder.fit_transform(X[:, 3])
onehotencoder = OneHotEncoder(categorical_features = [3
])
X = onehotencoder.fit_transform(X).toarray()

# Avoiding the dummy variable trap
X = X[:, 1:]
```

In [4]:
```python
# Splitting the dataset into the Training set and Test
set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size = 0.2, random_state = 0)
```

In [5]:
```python
# Fitting Multiple Linear Regression to the Training se
t
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

Out[5]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs
=1, normalize=False)

In [6]:
```python
# Predicting the Test set results
y_pred = regressor.predict(X_test)
regressor.score(X_test, y_test)
```

Out[6]: 0.9347068473282922