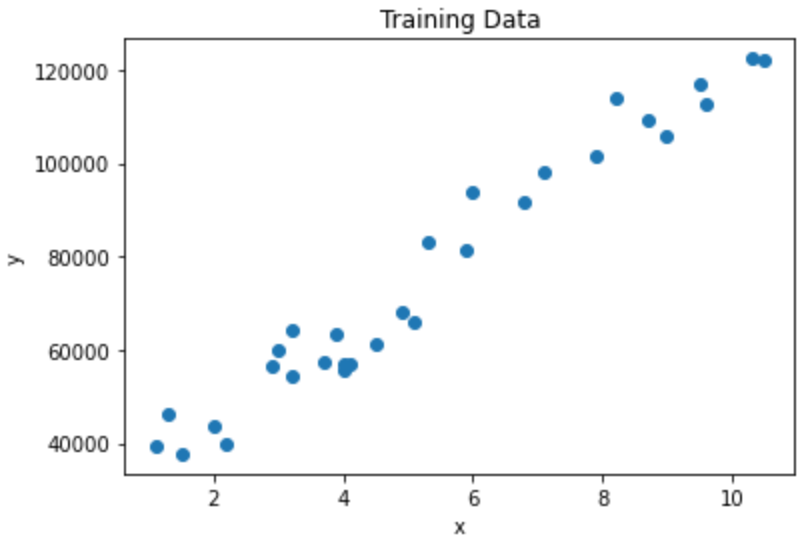
Making the import

In [1]:
```
import numpy as np
from sklearn.linear_model import LinearRegression
import pandas as pd
import matplotlib.pyplot as plt
plt.rcParams['figure.figsize'] = (6.0, 4.0)
from sklearn.metrics import mean_squared_error # for calculating mean_squared error
```

Preprocessing Input data

In [2]:
```
data = pd.read_csv('Salary_Data.csv')
X = data.iloc[:, 0]
Y = data.iloc[:, 1]
plt.scatter(X, Y)
plt.xlabel('x')
plt.ylabel('y')
plt.title('Training Data')
plt.show()
X = np.array(X).reshape((-1, 1))
```



Building the model

In [3]:
```
model = LinearRegression()
model.fit(X, Y)
print('theta0 = ', model.intercept_, 'theta1 = ', model.coef_)
```

```
theta0 =  25792.20019866871 theta1 =  [9449.96232146]
```

Making predictions for Training Data

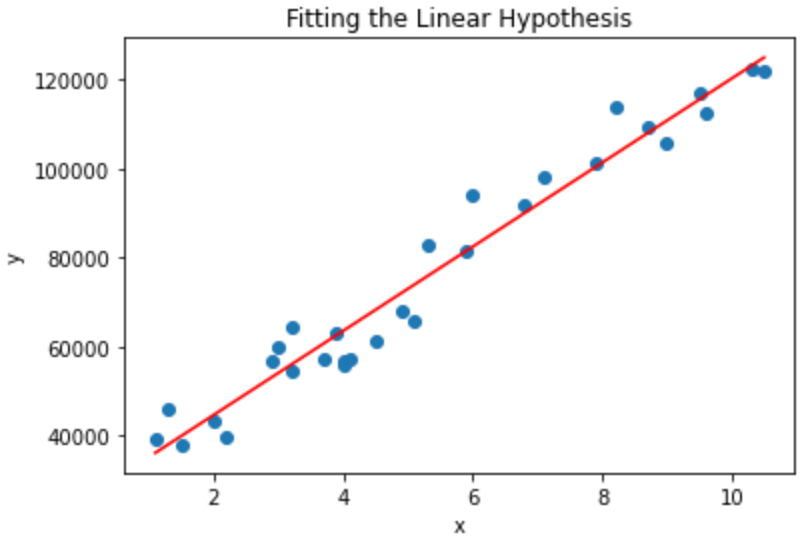
In [4]:
```
Y_pred = model.predict(X)
```

Computing mean_squared error

In [5]:
```
print('RMSE for Regression=>',np.sqrt(mean_squared_error(Y,Y_pred)))
plt.scatter(X, Y)
plt.plot(X, Y_pred, color = 'red')
plt.xlabel('x')
plt.ylabel('y')
```

```
plt.title('Fitting the Linear Hypothesis')
plt.show()
```

```
RMSE for Regression=> 5592.043608760662
```



Making predictions

In [6]:

```
test_X = 1.2
Y_pred = model.predict(np.array(test_X).reshape((-1, 1)))
```

predict() is equivalent to below

In [7]:

```
Y_pred2 = model.intercept_ + model.coef_ * test_X
print("Salary Prediction for Experience of ", str(test_X), " years =", Y_pred, Y_pred2)
```

```
Salary Prediction for Experience of  1.2  years = [37132.15498441] [37132.15498441]
```

In [ ]: