In [24]:
```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
```

In [25]:
```
# X = Home Size; y = KW hrs/Month
X = np.array([1290,1350,1470,1600,1710,1840,1980,2230,2400,2930])
y = np.array([1182,1172,1264,1493,1571,1711,1804,1840,1956,1954])
```

In [26]:
```
X = X.reshape(-1, 1)
```

In [27]:
```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.20,random_state=0)
```

In [28]:
```
from sklearn.linear_model import LinearRegression
lr=LinearRegression()
lr.fit(X_train,y_train)
```

Out[28]:
```
LinearRegression()
```

In [29]:
```
y_pred=lr.predict(X_test)
y_pred
```

Out[29]:
```
array([1394.09419127, 1855.93937011])
```

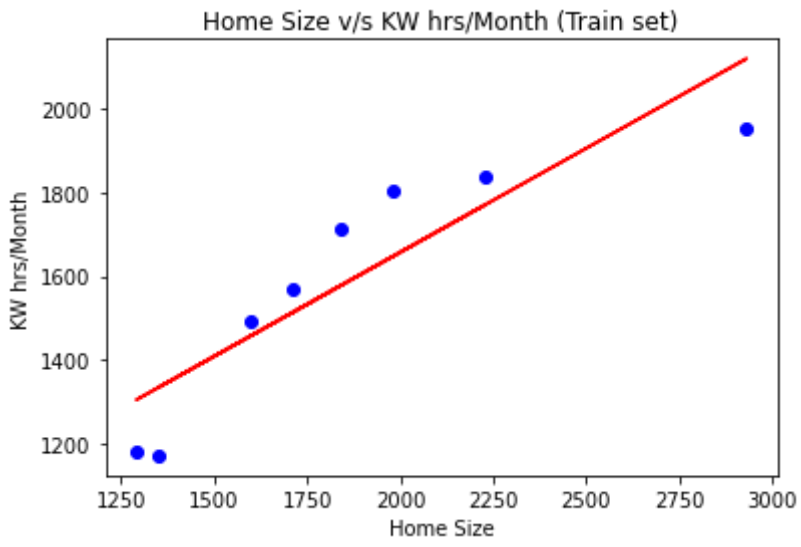
In [33]:
```
X_test
```

Out[33]:
```
array([[1470],
       [2400]])
```

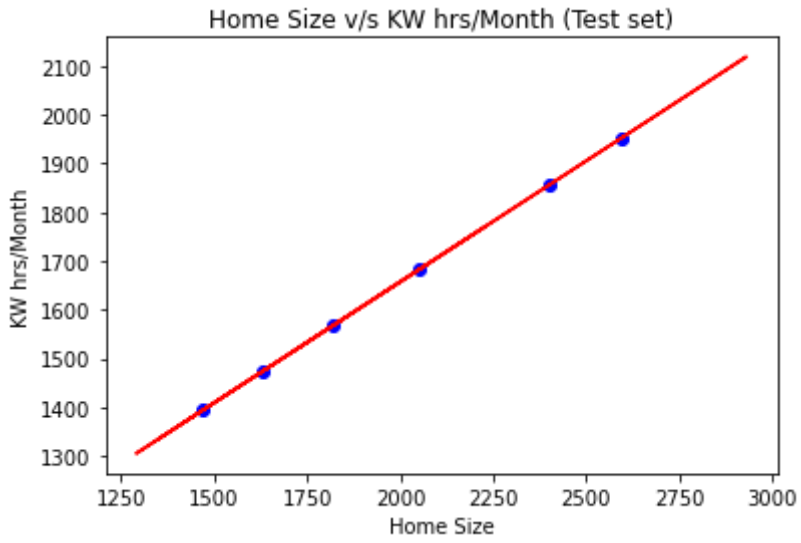
In [41]:
```
test=np.array([1470,2400,1630,2592,1820,2052])
test = test.reshape(-1,1)
```

In [42]:
```
test_pred = lr.predict(test)
```

In [43]:
```
plt.scatter(X_train,y_train,color='blue')
plt.plot(X_train,lr.predict(X_train),color='red')
plt.title('Home Size v/s KW hrs/Month (Train set)')
plt.xlabel('Home Size')
plt.ylabel('KW hrs/Month')
plt.show()
```



In [45]:
```
plt.scatter(test,test_pred,color='blue')
plt.plot(X_train,lr.predict(X_train),color='red')
plt.title('Home Size v/s KW hrs/Month (Test set)')
plt.xlabel('Home Size')
plt.ylabel('KW hrs/Month')
plt.show()
```



In [47]:
```
from sklearn import metrics
print('RMSE:',np.sqrt(metrics.mean_absolute_error(y_test,y_pred)))
print('R2:',np.sqrt(metrics.r2_score(y_test,y_pred)))
```

```
RMSE: 10.727413974491324
R2: 0.9420710130928927
```

In [ ]:

In [ ]: