```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
```

```python
df = pd.read_csv("lunap.csv")
```

```python
print(df.head())
```

```
         date  sales  customer  cost_of_operation  value_of_sales  \
    0  1.1.22    150       120               5000           35000
    1  2.1.22    200       140               5000           35000
    2  3.1.22    250       160               5000           35000
    3  4.1.22    300       180               5000           35000
    4  5.1.22    350       200               5000           35000

       profit_from_sales
    0              15000
    1              18000
    2              20000
    3              22000
    4              24500
```

```python
X = df.drop(['value_of_sales','date'], axis=1)
```

```python
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

```python
n_clusters = 3
```

```python
kmeans = KMeans(n_clusters=n_clusters, random_state=42)
kmeans.fit(X_scaled)
```

```
    /usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning: The default value of `n_init` will change fr
      warnings.warn(
```

```
    ▼              KMeans
    KMeans(n_clusters=3, random_state=42)
```

```python
df['cluster'] = kmeans.labels_
```

```python
print("Centros dos Clusters:")
print(scaler.inverse_transform(kmeans.cluster_centers_))
```

```
    Centros dos Clusters:
    [[ 231.81818182   152.72727273  5000.         37045.45454545]
     [ 414.28571429   225.71428571  5000.         31400.        ]
     [ 425.          230.          5000.         62450.        ]]
```

```python
plt.figure(figsize=(10, 6))

for cluster in range(n_clusters):
    cluster_data = df[df['cluster'] == cluster]
    plt.scatter(cluster_data['customer'], cluster_data['profit_from_sales'], label=f'Cluster {cluster}')

plt.title('Clusters de Dados de Habitação na Califórnia')
plt.xlabel('customer')
plt.ylabel('profit_from_sales')
plt.legend()
plt.show()
```

Clusters de Dados de Habitação na Califórnia