Exp 3: Model planning and building.

AIM:

To build a linear regression model predicting sales or advertising budgets and to apply k means clustering to group data based on advertising features.

Algorithm:

1. Load and explore the dataset.
2. Split data into features (TV, Radio, Newspaper and target sales).
3. Split into training and testing sets.
4. Train a linear regression model and evaluate it
5. Scale features and apply k means clustering
6. Visualize clusters of TV budgets vs sales plot.
7. Apply k means clustering (with 3 clusters) on ~~the~~ scaled features.
8. Add cluster labels to the data set and analyse clusters by plotting TV budget.

Code:

```
import pandas as pd.
import matplotlib.pyplot as plt
import seaborn as sns.
from sklearn.linear_model import LinearRegression.
from sklearn.model_selection import train_test_split.
df = pd.read_csv("advertising.csv")
print(df.head())
print(df.describe())
x = df[['TV', 'Radio', 'Newspaper']]
y = df['Sales']
# Split data.
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2
                                                    random_state)
# train model
model = LinearRegression.
model.fit(x_train, y_train)
y_pred = model.predict(x_test)
mse = mean_squared_error(y_test, y_pred)
print("Linear Regression MSE:", mse)
plt.figure(figsize=(8,5))
sns.scatterplot(x=y_test, y=y_pred)
```

Results:

The linear regression model with calculate mean square error, showing reasonable fit between actual and predicted sales.