9. Feature Encoding (Student Performance Dataset)

```
# Loading the Dataset
        import pandas as pd

        df = pd.DataFrame({
          'gender': ['Male', 'Female', 'Female', 'Male'],
          'parental_education': ['Bachelor', 'Master', 'High School', 'Associate'],
          'math_score': [85, 90, 78, 88],
          'reading_score': [80, 95, 75, 85]
        })

        print("Original Dataset:\n", df)
    #Output:
        Original Dataset:
           gender parental_education  math_score  reading_score
        0    Male         Bachelor        85          80
        1  Female           Master        90          95
        2  Female      High School        78          75
        3    Male        Associate        88          85

# Label Encoding
        from sklearn.preprocessing import LabelEncoder
        label_encoder = LabelEncoder()
        df['gender'] = label_encoder.fit_transform(df['gender'])
        print("\nAfter Label Encoding:\n", df)
      #Output:
        After Label Encoding:
          gender parental_education  math_score  reading_score
```

0	1	Bachelor	85	80
1	0	Master	90	95
2	0	High School	78	75
3	1	Associate	88	85

One-Hot Encoding

```
df = pd.get_dummies(df, columns=["parental_education"])
print("\nAfter One-Hot Encoding:\n", df)
```

Output:

After One-Hot Encoding:

	gender	math_score	reading_score	parental_education_Associate	parental_education_Bachelor	parental_education_High School	parental_education_Master
0	1	85	80	0	1	0	0
1	0	90	95	0	0	0	1
2	0	78	75	0	0	1	0
3	1	88	85	1	0	0	0

Preparing Dataset for Machine Learning

- Separating features (X) and target variable (y)

```
X = df.drop(columns=["math_score"])
y = df["math_score"]

print("\nFeatures (X):\n", X)
print("\nTarget Variable (y):\n", y)
```

Output:

Features (X):

	gender	reading_score	parental_education_Associate	parental_education_Bachelor	parental_education_High School	parental_education_Master
0	1	80	0	1	0	0
1	0	95	0	0	0	1
2	0	75	0	0	1	0
3	1	85	1	0	0	0

Target Variable (y):

0 85

1 90

2 78

3 88

Name: math_score, dtype: int64