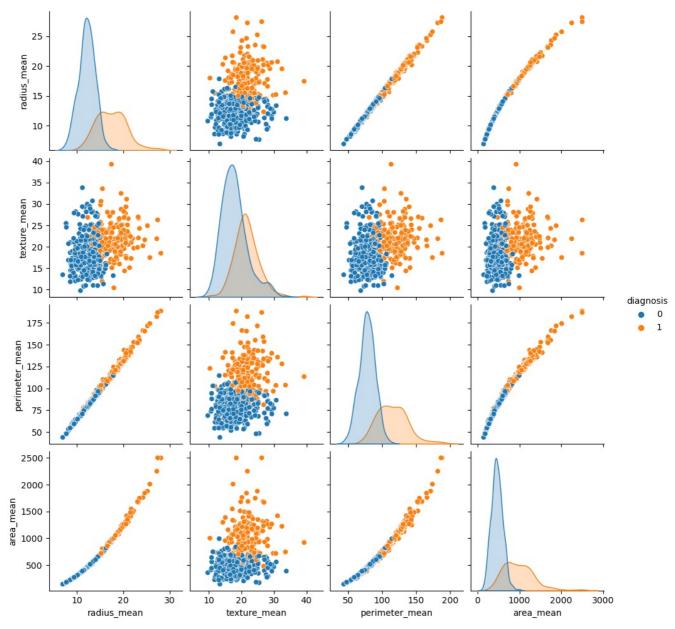
```
In [15]: import pandas as pd
           import seaborn as sns
           import matplotlib.pyplot as plt
           # Load the dataset
In [16]:
           data = pd.read_csv(r"C:\Users\Bhawana gupta\Downloads\breast cancer prediction dataset\data.csv")
In [17]:
           data
                      id diagnosis radius_mean texture_mean perimeter_mean area_mean smoothness_mean compactness_mean concavity_mean
             0
                 842302
                                         17.99
                                                      10.38
                                                                    122.80
                                                                               1001.0
                                                                                               0.11840
                                                                                                                 0.27760
                                                                                                                                 0.30010
                 842517
                                M
                                         20.57
                                                      17.77
                                                                    132.90
                                                                               1326.0
                                                                                               0.08474
                                                                                                                 0.07864
                                                                                                                                 0.08690
             2 84300903
                                М
                                         19.69
                                                      21.25
                                                                    130.00
                                                                               1203.0
                                                                                               0.10960
                                                                                                                 0.15990
                                                                                                                                 0.19740
             3
               84348301
                                         11.42
                                                      20.38
                                                                     77.58
                                                                                386.1
                                                                                               0.14250
                                                                                                                 0.28390
                                                                                                                                 0.24140
             4
               84358402
                                         20.29
                                                                    135.10
                                                                               1297.0
                                                                                               0.10030
                                                                                                                 0.13280
                                                                                                                                 0.19800
                                M
                                                      14.34
           564
                 926424
                                M
                                         21.56
                                                      22.39
                                                                    142.00
                                                                               1479.0
                                                                                               0.11100
                                                                                                                 0.11590
                                                                                                                                 0.24390
                 926682
                                M
                                         20.13
                                                                    131.20
                                                                               1261.0
                                                                                               0.09780
                                                                                                                 0.10340
                                                                                                                                 0.14400
           565
                                                      28.25
           566
                 926954
                                M
                                         16.60
                                                      28.08
                                                                    108.30
                                                                                858.1
                                                                                               0.08455
                                                                                                                 0.10230
                                                                                                                                 0.09251
           567
                 927241
                                         20.60
                                                      29.33
                                                                    140.10
                                                                               1265.0
                                                                                               0.11780
                                                                                                                 0.27700
                                                                                                                                 0.35140
                                                                                181.0
                                                                                                                                 0.00000
           568
                  92751
                                В
                                          7.76
                                                      24.54
                                                                     47.92
                                                                                               0.05263
                                                                                                                 0.04362
          569 rows × 33 columns
4
In [18]:
           # Encode the target variable
           data['diagnosis'] = data['diagnosis'].map({'M':1, 'B':0})
In [19]: # Display the first few rows of the dataset
           print(data.head())
                     id diagnosis radius_mean texture_mean perimeter_mean
                                                                                      area_mean \
           0
                842302
                                            17.99
                                                            10.38
                                                                             122.80
                                                                                          1001.0
                842517
                                            20.57
                                                            17.77
                                                                             132.90
                                                                                          1326.0
                                  1
           1
              84300903
                                                                             130.00
                                                            21.25
           2
                                  1
                                            19.69
                                                                                          1203.0
           3
              84348301
                                  1
                                            11.42
                                                            20.38
                                                                              77.58
                                                                                           386.1
                                                                                          1297.0
              84358402
                                                                             135.10
           4
                                  1
                                            20.29
                                                            14.34
              smoothness_mean
                                 compactness_mean concavity_mean concave points_mean
           0
                       0.\overline{1}1840
                                           0.\overline{27760}
                                                              0.3001
                                                                                     0.\overline{14710}
                       0.08474
                                           0.07864
                                                              0.0869
                                                                                     0.07017
           1
                       0.10960
                                                              0.1974
                                           0.15990
           2
                                                                                     0 12790
           3
                       0.14250
                                           0.28390
                                                              0.2414
                                                                                     0.10520
           4
                       0.10030
                                           0.13280
                                                              0.1980
                                                                                     0.10430
                    texture_worst
                                    perimeter worst area worst
                                                                     smoothness worst
              . . .
           0
                             17.33
                                               184.60
                                                            2019.0
                                                                                \overline{0.1622}
              . . .
                             23.41
                                               158.80
                                                            1956.0
                                                                                0.1238
           1
              . . .
           2
                             25.53
                                               152.50
                                                            1709.0
                                                                                0.1444
              . . .
           3
                             26.50
                                                98.87
                                                             567.7
                                                                                0.2098
              . . .
           4
                             16.67
                                               152.20
                                                            1575.0
                                                                                0.1374
              . . .
              compactness worst concavity worst concave points worst symmetry worst \
           0
                           0.6656
                                              0.7119
                                                                      0.2654
                          0.1866
                                              0.2416
                                                                      0.1860
                                                                                         0.2750
           1
                                                                      0.2430
                           0.4245
                                              0.4504
           2
                                                                                        0.3613
           3
                          0.8663
                                              0.6869
                                                                      0.2575
                                                                                        0.6638
           4
                          0.2050
                                              0.4000
                                                                      0.1625
                                                                                        0.2364
              fractal_dimension_worst Unnamed: 32
           0
                                0.11890
                                                   NaN
           1
                                0.08902
                                                   NaN
           2
                                0.08758
                                                   NaN
           3
                                0.17300
                                                   NaN
           4
                                0.07678
                                                   NaN
           [5 rows x 33 columns]
In [20]: # Pairplot to visualize relationships between features
           sns.pairplot(data, hue='diagnosis', vars=['radius_mean', 'texture_mean', 'perimeter_mean', 'area_mean'])
           plt.show()
```



In [21]: # Heatmap to visualize correlations between features
plt.figure(figsize=(12, 10))
sns.heatmap(data.corr(), annot=True, fmt='.2f')
plt.show()

```
- 1.0
                     diagnosis -0.04.00.79.49.79.70.36.60.70.70.330.00.5-0.00.56.5-0.00.29.29.4-0.00.09.70.40.78.70.40.59.60.79.40.3
          radius_mean -0.0 0.73.00 32.00 99.10.50 60.80.150.30 650.10 60.740.20 20.19.350.10.0 0.99.30.10.40.50.70.16.01
         texture_mean - .10.42.31.00.38.32.00.24.30.29.07.08.28.39.28.26.00.19.14.16.00.05.30.90.36.34.08.28.30.30.10.12
       perimeter_mean -0.070.74.00.331.00.99.20.50.70.80.180.26.690.09.60.7-0.20.26.20.40.08.00.970.30.90.90.16.46.50.70.19.09
                                                                                                                                       - 0.8
            area_mean -0.10.7 D.90.30.99.00.18.50.60.80.150.20.70.00.7 D.80.10.20.2 D.30.00.0).90.20.96.90.1 D.39.50.70.14.00
     compactness mean -0.00.60.50.24.56.50.61.00.88.80.60.50.50.50.40.14.70.50.64.20.50.50.54.20.50.50.80.80.80.50.6
                            .70.68.30.70.69.5<mark>0.88.00.9</mark>0.50.34.68.08.60.62.10.60.69.68.18.45.69.30.70.68.40.75.88.80.40.5
       concavity mean -0.05
  concave points_mean -0.0-0.78.80.20.88.80.50.88.9.2.00.46.10.70.00.70.69.08.49.44.60.10.20.80.20.86.80.49.60.78.90.38.3
                                                                                                                                       - 0.6
       symmetry mean -0.00.38.19.00.18.19.56.60.50.44.00.48.30.19.30.20.19.40.34.39.49.38.19.09.20.18.48.40.48.40.48.70.44
radius_se -0.14.57.68.28.60.70.30.50.68.70.30.00.00.20.970.90.16.36.38.50.24.20.70.19.70.70.14.29.38.58.09.00
            texture_se -0.00.00.10.39.09.07.07.07.05.08.07.18.16.21.00.27.10.40.28.19.28.40.28.10.40.10.08.07.09.07.09.07.12.18.0
          perimeter_se -0.1 0.5 0.6 0.2 8.6 0.7 0.3 0.5 0.6 0.7 0.3 0.0 0.9 0.2 1.0 0.9 0.1 0.4 0.3 0.5 0.2 0.2 0.7 0.2 0.7 0.1 0.3 0.4 0.5 0.1 0.0 0
                                                                                                                                       0.4
               area_se -0.18.50.70.20.74.80.25.46.62.69.220.0<mark>0.95</mark>.10.94.00.08.28.20.42.18.13.70.20.76.80.18.28.39.54.07.02
        smoothness_se--).1@.07.22.0D.20.17.3B.14.10.09.19.40.16.40.19.01.200.39.27.3B.4D.40.20.07.20.10.3D.06.06.10.10.10
       compactness_se +0.00.29.20.19.26.20.30.740.60.49.40.50.30.28.40.28.34.00.80.74.39.80.20.14.26.20.28.68.69.48.28.5
          concavity_se -0.06.25.19.14.2B.2D.25.57.69.44.34.45.35.19.36.27.2.10.30.77.33.77.19.10.2B.19.17.45.66.44.20.4
     concave points_se--).08.4D.38.16.4D.3D.38.64.68.6D.39.34.5D.28.56.4D.30.74.71.00.3D.6D.36.09.39.34.2D.4D.5D.6D.14.3
                                                                                                                                       - 0.2
          fractal_dimension_se-0.03.040.050.00.002.28.50.45.26.38.69.28.28.24.18.40.80.70.60.31.00.04.00.00.00.00.10.39.38.20.10.5
          radius_worst -0.00.78.90.30.90.90.20.54.60.80.19.20.70.10.70.70.20.20.19.30.10.00.30.99.90.20.48.50.70.20.00
          texture_worst -0.0<mark>6.46.30.9</mark>0.30.29.04.29.30.29.09.05.19.40.20.20.00.14.10.09.06.09.34<mark>.00</mark>.30.39.28.36.30.36.20.2
       perimeter_worst -\,000.78.90.30.90.90.24.50.78.80.20.20.20.70.10.70.70.20.26.28.30.10.00.99.31.00.90.24.50.60.80.20.1
                                                                                                                                       - 0 0
            area_worst -).10.78.90.30.94.90.28.51.60.80.18.2).70.00.78.80.18.20.19.340.10.0).90.30.98.00.20.44.50.70.20.08
     compactness_worst -0.00.59.40.29.46.39.40.80.76.60.40.46.29.09.34.28.06.68.48.45.06.39.48.36.58.44.51.00.89.80.60.81
       concavity_worst -0.02.66.58.30.56.5 D.40.82.88.70.48.30.30.00.42.30.06.64.66.50.04.38.57.37.62.54.50.89.00.80.58.
  concave points_worst -0.0-0.79.70.30.70.70.50.80.86.09.48.18.50.12.50.54.10.48.44.60.08.2.2.79.30.80.70.70.50.80.86.00.50.5
                                                                                                                                        -0.2
       symmetry worst -0.09.40.16.10.19.14.39.50.40.38.70.30.00.16.10.00.10.28.20.14.39.10.24.20.20.20.49.60.50.50.00.5
fractal_dimension_worst 0.06.32.00.12.00.00.05.0.69.50.69.50.37.40.70.050.06.09.02.10.59.44.30.08.59.09.22.14.08.60.80.69.50.54.00
                                                concavity_mean
                                                   concave points_mean
                                                         fractal_dimension_mean
                                                               texture_se
                                                                           compactness_se
                                                                                 concave points_se
                                                                                                            concavity_worst
                                                                                                                  symmetry_worst
                                                                                                                        Unnamed: 32
                              adius_mean
                                 texture_mean
                                    perimeter mean
                                       area_mean
                                          smoothness_mean
                                             compactness mean
                                                      symmetry mean
                                                            radius_se
                                                                  perimeter_se
                                                                     area se
                                                                        smoothness se
                                                                               concavity_se
                                                                                    symmetry se
                                                                                       fractal_dimension_se
                                                                                          radius_worst
                                                                                             texture worst
                                                                                                perimeter_worst
                                                                                                    area_worst
                                                                                                      smoothness_worst
                                                                                                         compactness_worst
                                                                                                               concave points_worst
                                                                                                                     fractal dimension worst
```

```
In [22]: # Check for missing values
print(data.isnull().sum())

# Replace missing values with the mean of the respective columns
data.fillna(data.mean(), inplace=True)

# Verify that there are no more missing values
print(data.isnull().sum())
```

```
diagnosis
                                     Θ
         radius mean
                                     0
         texture mean
         perimeter mean
                                     0
         area_mean
                                     0
         smoothness mean
                                     0
         compactness mean
         concavity_mean
                                     0
         concave points_mean
                                     0
         symmetry mean
         fractal_dimension_mean
                                     0
         radius_se
                                     0
         texture se
         {\tt perimeter\_se}
                                     0
                                     Θ
         area_se
         smoothness se
                                     0
         compactness se
                                     0
         concavity se
         concave points_se
                                     0
         symmetry_se
         fractal dimension se
         radius_worst
                                     0
         texture worst
                                    0
         perimeter worst
                                    0
         area worst
                                    0
         smoothness_worst
         compactness worst
         concavity worst
                                    0
         concave points_worst
                                     0
         symmetry worst
                                     0
         fractal dimension worst
         Unnamed: 32
                                   569
         dtype: int64
         id
                                     0
                                     0
         diagnosis
         radius mean
                                     0
         texture mean
                                     0
         perimeter mean
         area mean
                                     0
         smoothness mean
                                     0
         compactness mean
                                     0
         concavity mean
                                     0
         concave points mean
                                     0
         symmetry_mean
                                     0
         fractal_dimension_mean
                                     0
         radius se
         texture se
                                     0
         perimeter_se
                                     0
         area se
         smoothness_se
                                     0
         compactness se
                                     0
         concavity se
                                     0
         concave points_se
         symmetry_se
                                     0
         fractal dimension se
         radius worst
                                     0
         texture_worst
                                     0
         perimeter_worst
         area worst
         smoothness_worst
                                     0
         compactness worst
                                     0
                                     0
         concavity worst
         concave points_worst
                                     0
         symmetry_worst
                                     0
         fractal_dimension_worst
                                     0
                                   569
         Unnamed: 32
         dtvpe: int64
In [23]: from sklearn.model_selection import train_test_split
         from sklearn.preprocessing import StandardScaler
In [24]: # Split the data into features and target
         X = data.drop(['id', 'diagnosis', 'Unnamed: 32'], axis=1)
         y = data['diagnosis']
In [25]: # Split the data into training and testing sets
         X train, X test, y train, y test = train test split(X, y, test size=0.2, random state=42)
In [26]: # Standardize the features
         scaler = StandardScaler()
         X train = scaler.fit transform(X train)
         X_test = scaler.transform(X_test)
In [27]: from sklearn.ensemble import RandomForestClassifier
         from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
```

```
In [28]: # Initialize the model
         model = RandomForestClassifier(n estimators=100, random state=42)
         # Train the model
         model.fit(X train, y train)
Out[28]: v
                   RandomForestClassifier
         RandomForestClassifier(random_state=42)
In [29]: # Make predictions
         y_pred = model.predict(X_test)
In [30]: # Evaluate the model
         print("Accuracy:", accuracy_score(y_test, y_pred))
         print("Classification Report:\n", classification_report(y_test, y_pred))
         print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
         Accuracy: 0.9649122807017544
         Classification Report:
                        precision
                                     recall f1-score
                                                        support
                    0
                            0.96
                                      0.99
                                                0.97
                                                            71
                            0.98
                                      0.93
                                                0.95
                                                            43
                                                0.96
                                                           114
             accuracy
                            0.97
                                      0.96
                                                0.96
                                                           114
            macro avg
         weighted avg
                            0.97
                                      0.96
                                                0.96
                                                           114
         Confusion Matrix:
          [[70 1]
          [ 3 40]]
In [31]: import seaborn as sns
         import matplotlib.pyplot as plt
         from sklearn.metrics import confusion_matrix
In [32]: # Confusion matrix
         cm = confusion matrix(y test, y pred)
In [33]:
         plt.figure(figsize=(8, 6))
         sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['Benign', 'Malignant'], yticklabels=['Benign',
         plt.xlabel('Predicted')
         plt.ylabel('Actual')
         plt.title('Confusion Matrix')
         plt.show()
                                      Confusion Matrix
                                                                                        70
                                                                                       - 60
                              70
                                                               1
```

