Assignment 2
Aditya verma
Date 4-6-25

Create a linear regression model to predict life expectancy from immunization factors (Polio, Diphtheria, Hepatitis B). Interpret the coefficients - which vaccine has the strongest association with life expectancy? Use Data set of life expectancy shared earlier .

```
#Create a linear regression model to predict life expectancy from immunization factors (Polio, Diphtheria,
Hepatitis B). Interpret the coefficients - which vaccine has the strongest association with life expectancy? Use
Data set of life expectancy shared earlier .

immunization_features1 = df[['Schooling','Incomecompositionofresources']]
immunization_features = df[['Polio', 'Diphtheria', 'HepatitisB']]
immunization_target = df['Lifeexpectancy']
immunization_X_train, immunization_X_test, immunization_y_train, immunization_y_test = train_test_split(
    immunization_features1, immunization_target, test_size=0.2, random_state=42
)
```

0.0s Python

```
from sklearn.linear_model import LinearRegression
immunization_lr = LinearRegression()
immunization_lr.fit(immunization_X_train, immunization_y_train)
immunization_y_pred = immunization_lr.predict(immunization_X_test)
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
immunization_rmse = np.sqrt(mean_squared_error(immunization_y_test, immunization_y_pred))
immunization_mae = mean_absolute_error(immunization_y_test, immunization_y_pred)
immunization_r2 = r2_score(immunization_y_test, immunization_y_pred)
print("\nImmunization Linear Regression Results:")
print(f"RMSE: {immunization_rmse:.2f}")
print(f"MAE: {immunization_mae:.2f}")
print(f"R2 Score: {immunization_r2:.2f}")
```

0.0s Python

Immunization Linear Regression Results:
RMSE: 6.13
MAE: 4.33
R2 Score: 0.57

```python
coefficients = immunization_lr.coef_
intercept = immunization_lr.intercept_

print("\nImmunization Linear Regression Coefficients:")
immunization_feature_names = immunization_features1.columns
print("\nImmunization Feature Coefficients:")
for feature, coef in zip(immunization_feature_names, coefficients):
    print(f"{feature}: {coef:.4f}")

print("\nImmunization Feature INTERCEPTS:",intercept)

# Interpretation of coefficients
print("\nInterpretation of Coefficients:")
for feature, coef in zip(immunization_feature_names, coefficients):
    if coef > 0:
        print(f"A unit increase in {feature} is associated with an increase in life expectancy by {coef:.4f} years.")
    else:
        print(f"A unit increase in {feature} is associated with a decrease in life expectancy by {-coef:.4f} years.")
```

✓ 0.0s Python

Immunization Linear Regression Coefficients:

Immunization Feature Coefficients:
Schooling: 1.2522
Incomecompositionofresources: 16.5652

Immunization Feature INTERCEPTS: 43.710149059244344

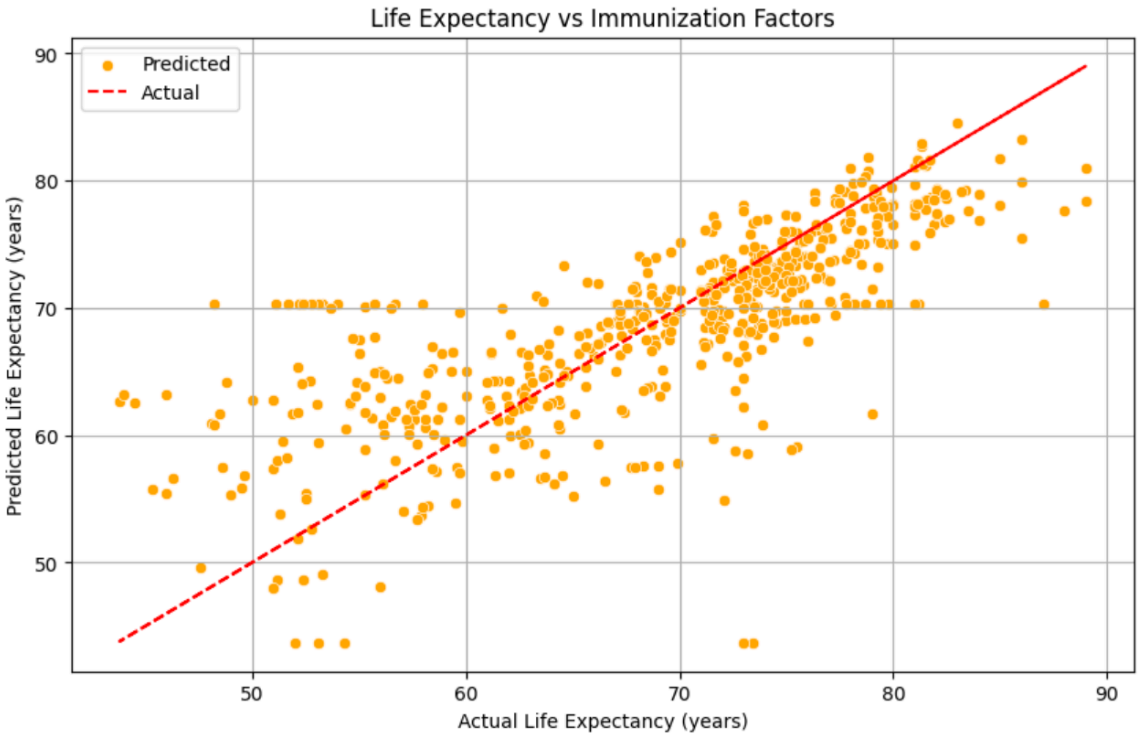
Interpretation of Coefficients:
A unit increase in Schooling is associated with an increase in life expectancy by 1.2522 years.
A unit increase in Incomecompositionofresources is associated with an increase in life expectancy by 16.5652 years.

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

sns.scatterplot(x=immunization_y_test, y=immunization_y_pred, label='Predicted', color='orange')

plt.plot(immunization_y_test, immunization_y_test, color='red', linestyle='--', label='Actual')  # Diagonal line for actual values
plt.grid(True)
plt.title('Life Expectancy vs Immunization Factors')
plt.xlabel('Actual Life Expectancy (years)')
plt.ylabel('Predicted Life Expectancy (years)')
plt.legend()
```

✓ 0.2s Py


Life Expectancy vs Immunization Factors
Predicted
Actual
Predicted Life Expectancy (years)
Actual Life Expectancy (years)
90
80
70
60
50
50
60
70
80
90