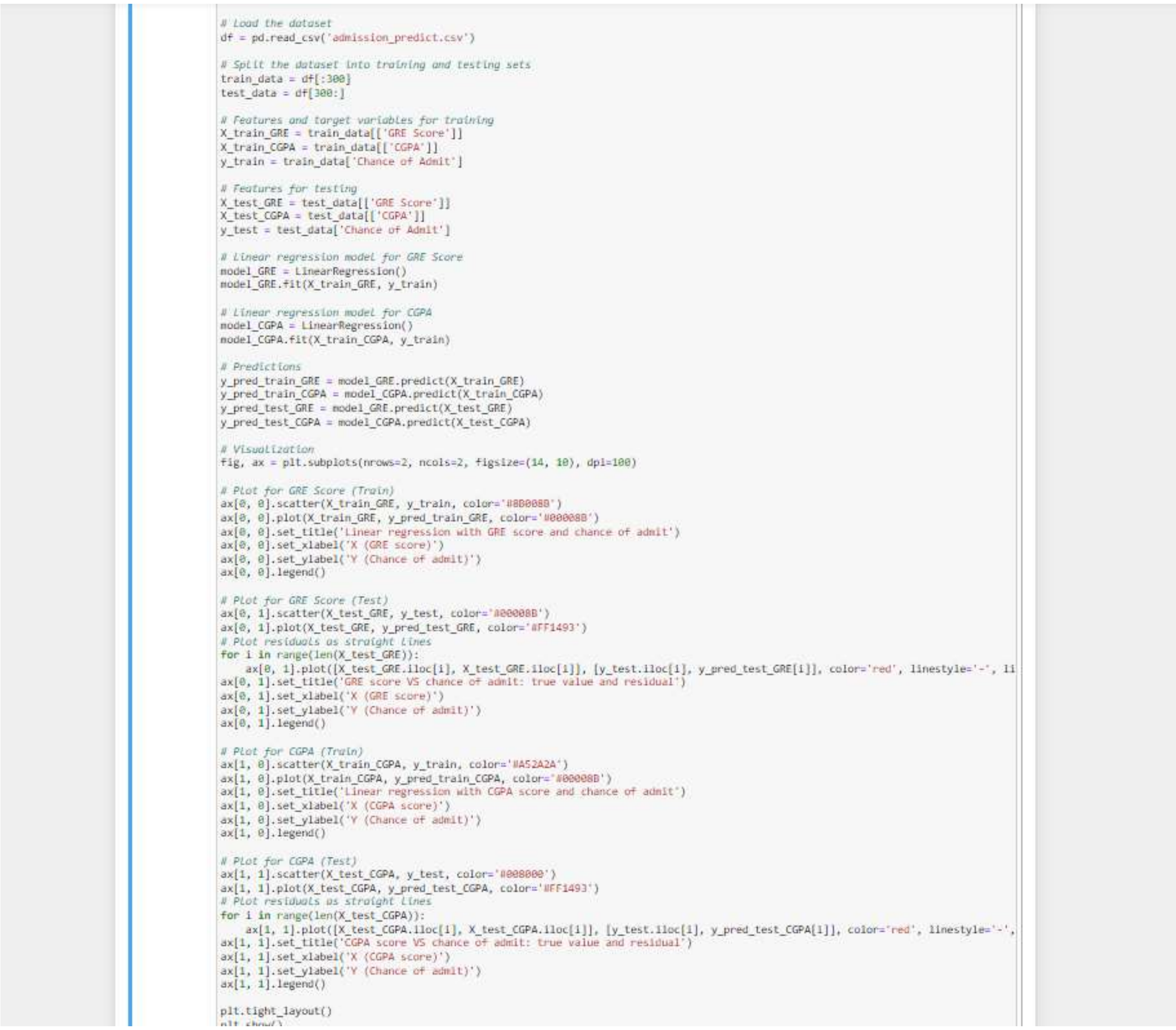
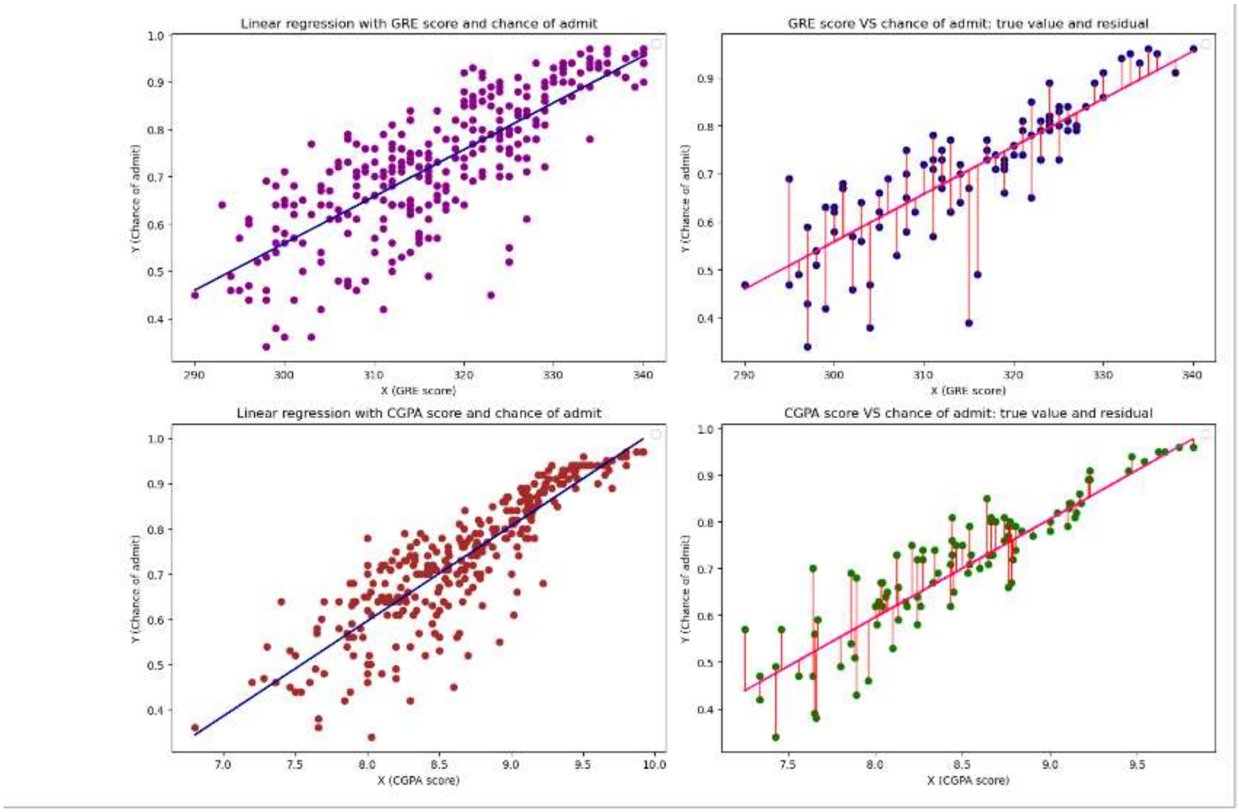
Credit Task 5.2C: Linear Regression

The dataset contains several parameters which are considered important during the application for Masters Programs. The parameters included are:
1. GRE Scores (out of 340)
2. TOEFL Scores (out of 120)
3. University Rating (out of 5)
4. Statement of Purpose and Letter of Recommendation Strength (out of 5)
5. Undergraduate GPA (out of 10)
6. Research Experience (either 0 or 1)
7. Chance of Admit (ranging from 0 to 1)



```
# Load the dataset
df = pd.read_csv('admission_predict.csv')

# Split the dataset into training and testing sets
train_data = df[:300]
test_data = df[300:]

# Features and target variables for training
X_train_GRE = train_data[['GRE Score']]
X_train_CGPA = train_data[['CGPA']]
y_train = train_data['Chance of Admit']

# Features for testing
X_test_GRE = test_data[['GRE Score']]
X_test_CGPA = test_data[['CGPA']]
y_test = test_data['Chance of Admit']

# Linear regression model for GRE Score
model_GRE = LinearRegression()
model_GRE.fit(X_train_GRE, y_train)

# Linear regression model for CGPA
model_CGPA = LinearRegression()
model_CGPA.fit(X_train_CGPA, y_train)

# Predictions
y_pred_train_GRE = model_GRE.predict(X_train_GRE)
y_pred_train_CGPA = model_CGPA.predict(X_train_CGPA)
y_pred_test_GRE = model_GRE.predict(X_test_GRE)
y_pred_test_CGPA = model_CGPA.predict(X_test_CGPA)

# Visualization
fig, ax = plt.subplots(nrows=2, ncols=2, figsize=(14, 10), dpi=100)

# Plot for GRE Score (Train)
ax[0, 0].scatter(X_train_GRE, y_train, color='#8B008B')
ax[0, 0].plot(X_train_GRE, y_pred_train_GRE, color='#00008B')
ax[0, 0].set_title('Linear regression with GRE score and chance of admit')
ax[0, 0].set_xlabel('X (GRE score)')
ax[0, 0].set_ylabel('Y (Chance of admit)')
ax[0, 0].legend()

# Plot for GRE Score (Test)
ax[0, 1].scatter(X_test_GRE, y_test, color='#00008B')
ax[0, 1].plot(X_test_GRE, y_pred_test_GRE, color='#FF1493')
# Plot residuals as straight lines
for i in range(len(X_test_GRE)):
    ax[0, 1].plot([X_test_GRE.iloc[i], X_test_GRE.iloc[i]], [y_test.iloc[i], y_pred_test_GRE[i]], color='red', linestyle='-', li
ax[0, 1].set_title('GRE score VS chance of admit: true value and residual')
ax[0, 1].set_xlabel('X (GRE score)')
ax[0, 1].set_ylabel('Y (Chance of admit)')
ax[0, 1].legend()

# Plot for CGPA (Train)
ax[1, 0].scatter(X_train_CGPA, y_train, color='#A52A2A')
ax[1, 0].plot(X_train_CGPA, y_pred_train_CGPA, color='#00008B')
ax[1, 0].set_title('Linear regression with CGPA score and chance of admit')
ax[1, 0].set_xlabel('X (CGPA score)')
ax[1, 0].set_ylabel('Y (Chance of admit)')
ax[1, 0].legend()

# Plot for CGPA (Test)
ax[1, 1].scatter(X_test_CGPA, y_test, color='#008000')
ax[1, 1].plot(X_test_CGPA, y_pred_test_CGPA, color='#FF1493')
# Plot residuals as straight lines
for i in range(len(X_test_CGPA)):
    ax[1, 1].plot([X_test_CGPA.iloc[i], X_test_CGPA.iloc[i]], [y_test.iloc[i], y_pred_test_CGPA[i]], color='red', linestyle='-',
ax[1, 1].set_title('CGPA score VS chance of admit: true value and residual')
ax[1, 1].set_xlabel('X (CGPA score)')
ax[1, 1].set_ylabel('Y (Chance of admit)')
ax[1, 1].legend()

plt.tight_layout()
plt.show()
```



Utilizing the "Admission Predict" dataset, the algorithm applies linear regression analysis to forecast the "Chance of Admit" according to the "CGPA" and "GRE Score." Pandas is used to import the dataset first, after which it is divided into training (the first 300 rows) and testing (the remaining rows) groups. Next, using "GRE Score" and "CGPA" as features, distinct Linear Regression models are developed.

Training and testing data predictions are generated with the taught models. The findings are then plotted on distinct subplots for the training and testing sets for the GRE Score and CGPA, showing actual values against anticipated using matplotlib.

Relative to the anticipated values, or residuals, are likewise displayed as red lines on the test set plots. The charts illustrate the fit of the models and the size of prediction errors, offering insights into how effectively the models predict the "Chance of Admit" based on the "GRE Score" and "CGPA."