

Source Code

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score

# Sample match data (Home team, Away team, Home goals, Away goals, Possession
ratio, Outcome)
data = {
    'HomeTeam': ['TeamA', 'TeamB', 'TeamC', 'TeamA', 'TeamB'],
    'AwayTeam': ['TeamB', 'TeamC', 'TeamA', 'TeamC', 'TeamA'],
    'HomeGoals': [2, 1, 3, 2, 1],
    'AwayGoals': [1, 2, 1, 2, 3],
    'PossessionHome': [55, 48, 60, 50, 47],
    'PossessionAway': [45, 52, 40, 50, 53],
    'Outcome': [1, 0, 1, 0, 0] # 1 = Home Win, 0 = Away Win
}

# Create a DataFrame
df = pd.DataFrame(data)

# Features: HomeGoals, AwayGoals, PossessionHome, PossessionAway
X = df[['HomeGoals', 'AwayGoals', 'PossessionHome', 'PossessionAway']]

# Target variable: Outcome
y = df['Outcome']
```

```

# 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)

# Create a logistic regression model
model = LogisticRegression()

# Train the model
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)

# Output the result
print(f'Predicted outcomes: {y_pred}')
print(f'Accuracy: {accuracy * 100:.2f}%')

# Plotting the outcomes
match_ids = range(1, len(y_pred) + 1) # Match IDs for plotting

# Plot predicted vs actual outcomes
plt.figure(figsize=(10, 5))
plt.bar(match_ids, y_test, width=0.4, label='Actual', align='center', alpha=0.6)
plt.bar(match_ids, y_pred, width=0.4, label='Predicted', align='edge', alpha=0.6)

```

```
# Adding labels and title
plt.xlabel('Match ID')
plt.ylabel('Outcome (1 = Home Win, 0 = Away Win)')
plt.title('Predicted vs Actual Outcomes of Matches')
plt.legend()

# Display the plot
plt.show()
```

OUTPUT

Predicted outcomes: [0 0]
Accuracy: 100.00%



