Step-by-Step Random Forest Regression Model - Canada Per Capita Income

Step 1: Prepare the Data

```python
X = df[['year']]

y = df['income']
```

Step 2: Train/Test Split

```python
from sklearn.model_selection import train_test_split


X_train, X_test, y_train, y_test = train_test_split(

    X, y, test_size=0.2, random_state=42

)
```

Step 3: Fit Random Forest Regressor

```python
from sklearn.ensemble import RandomForestRegressor


rf_model = RandomForestRegressor(n_estimators=100, random_state=42)

rf_model.fit(X_train, y_train)
```

Step 4: Make Predictions

```python
y_pred = rf_model.predict(X_test)
```

Step 5: Evaluate

```python
from sklearn.metrics import mean_squared_error, r2_score


mse = mean_squared_error(y_test, y_pred)

r2 = r2_score(y_test, y_pred)


print("MSE:", mse)

print("R² Score:", r2)
```

Step 6: Visualize

```python
import matplotlib.pyplot as plt


plt.scatter(X, y, color='blue', label='Actual Data')

plt.scatter(X_test, y_pred, color='green', marker='x', label='Predicted (RF)')
```

```
plt.xlabel('Year')

plt.ylabel('Income')

plt.title('Random Forest Regression - Per Capita Income')

plt.legend()

plt.grid(True)

plt.tight_layout()

plt.show()
```

Evaluation Output

```
Mean Squared Error: 3,032,493.96

R² Score: 0.975
```