Step 1: Import Libraries

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
```

Load dataset

```
city_data = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/internal/IPC_crimes_(City-wise)_2019-2021.csv')
state_data = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/internal/IPC_crimes_(State-UT-wise)_2019-2021.csv')
```

Convert the '2019' column in state_data to numeric (if necessary)

```
state_data['2019'] = pd.to_numeric(state_data['2019'], errors='coerce')
```

Handle missing values (if any)

```
state_data.dropna(inplace=True)
```

Step 4: Feature Selection

Select features and target for the model

```
# Combine relevant features or use them as they are for the model
# Here we will focus on predicting the 2021 crime rates

X = city_data[['2019', '2020', 'Population', 'RateofCognizableCrimes']]
y = city_data['2021']
```

Step 5: Split the Data

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
```

Step 6: Standardize the Data

```
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
```

Step 7: Model Training - Random Forest Regressor

```
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
```

```
RandomForestRegressor
RandomForestRegressor(random_state=42)
```

Step 8: Model Evaluation

```
y_pred = model.predict(X_test)
```

Calculate Mean Squared Error and R2 Score

```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
print(f'R2 Score: {r2}')
```

```
Mean Squared Error: 749366.4741090909
R2 Score: 0.8764539986576807
```

Step 9: Making Predictions

```
new_data = [[5000, 4500, 15.0, 350.0]]  # Example input data
new_data_scaled = scaler.transform(new_data)
prediction = model.predict(new_data_scaled)

print(f'Predicted Crime Rate for 2021: {prediction[0]}')
```

```
Predicted Crime Rate for 2021: 5061.65
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:465: UserWarning: X does not have valid feature names, but StandardScaler wa
  warnings.warn(
```