



# Correlation of unemployment and crime data for G7 nations + India



presented by  
Name: Jeevika Chopra







# Introduction & Motivation

- 1. Objective:** Explore unemployment-crime correlations in G7 countries and India for informed policy-making.
- 2. Global Focus:** Study G7 and India for diverse economic, cultural, and governance insights beyond regions.
- 3. Statistical Analysis:** Use data analysis to find patterns, guiding evidence-based policymaking for targeted interventions.
- 4. Temporal Perspective:** Spanning 2000-2020, observe historical trends amidst evolving socio-economic conditions.







# Unemployment Dataset

Source: International Labour Organisation  
[Link to Unemployment data](#)

# Crime Dataset

Source: UN Office on Drugs and Crime's International  
Homicide Statistics database  
[Link to crime data](#)



# Questions

**Correlation Inquiry:** What is the statistical correlation between unemployment rates and crime rates in the selected global entities over the past two decades?

**Comparative Analysis:** How do patterns and trends in the relationship between unemployment and crime differ among G7 countries, the European Union, and India, considering their diverse economic structures and cultural contexts?

**Policy Implications:** What evidence-based policy recommendations can be drawn from the analysis to address the interconnected dynamics of unemployment and crime, taking into account the evolving socio-economic conditions influenced by technological advancements and globalization?

# Methodology used in Analysis

- 1. About Dataset :

- Unemployment Data:

Having country names with unemployment rate from 1960 to 2021

Data type : zipped CSV

- Crime Data:

Having country names with crime rates from 1960 to 2022

Data type : zipped CSV

## 2. Making ETL Pipeline :

- install dependencies

- extraction.py : Responsible for extracting data from the original sources.

- transform.py : Responsible for removing null and empty values and transformed into a consistent format.



# Methodology used in the analysis

- load.py: This code takes charge of storing the processed data in a CSV file.
- pipeline.py: This module brings together the extract, transform, and load components into a cohesive automated pipeline

## 3. Data visualisation & Analysis:

- Import libraries
- Visualise graphs made and analyse results

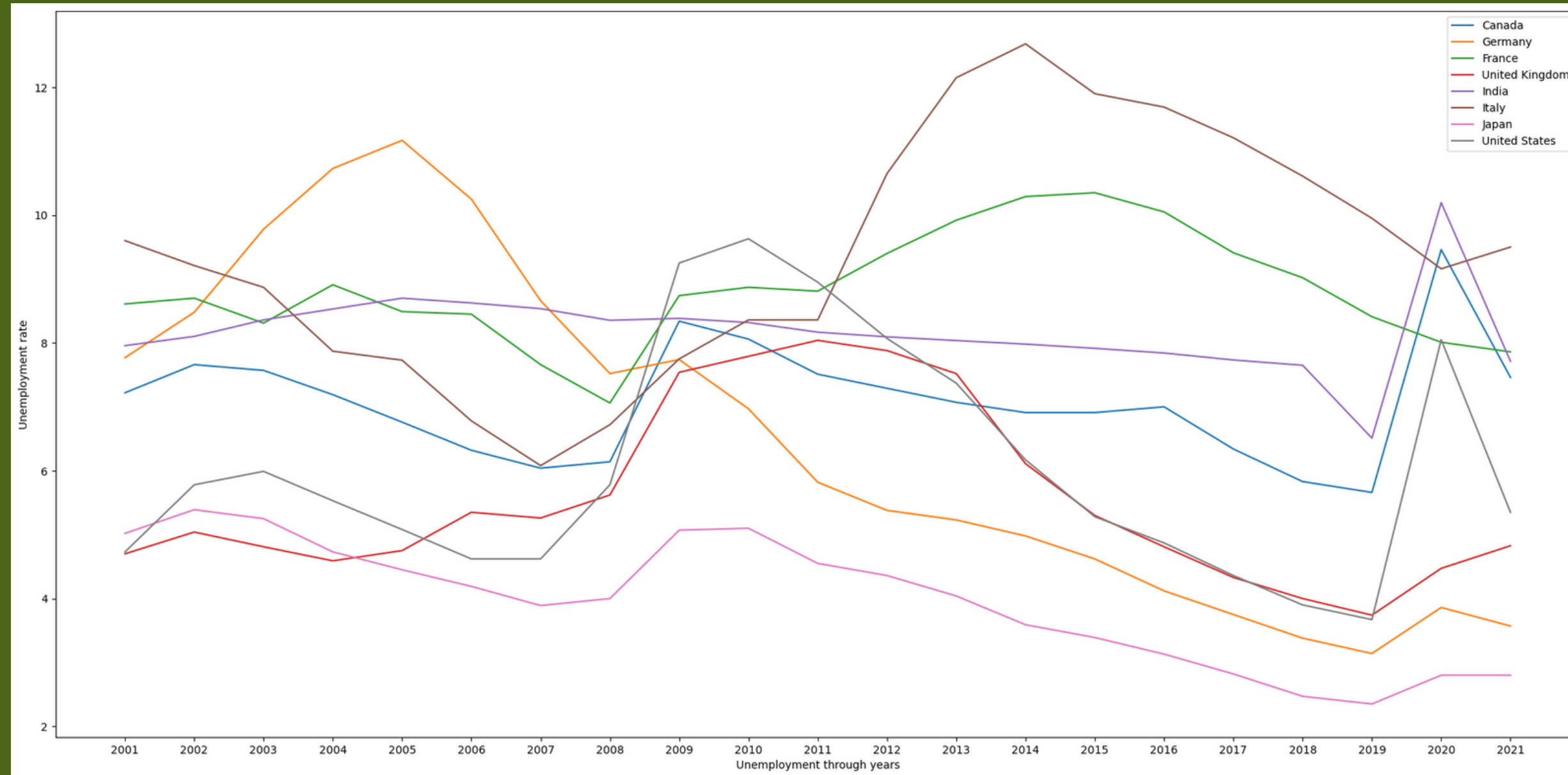
## 4. Problems encountered :

- Multiple null values and empty rows in crime dataset.
- Shape of both datasets were different
- Transformed into 1 shape
- Same countries needed to be taken into account by performing set intersection





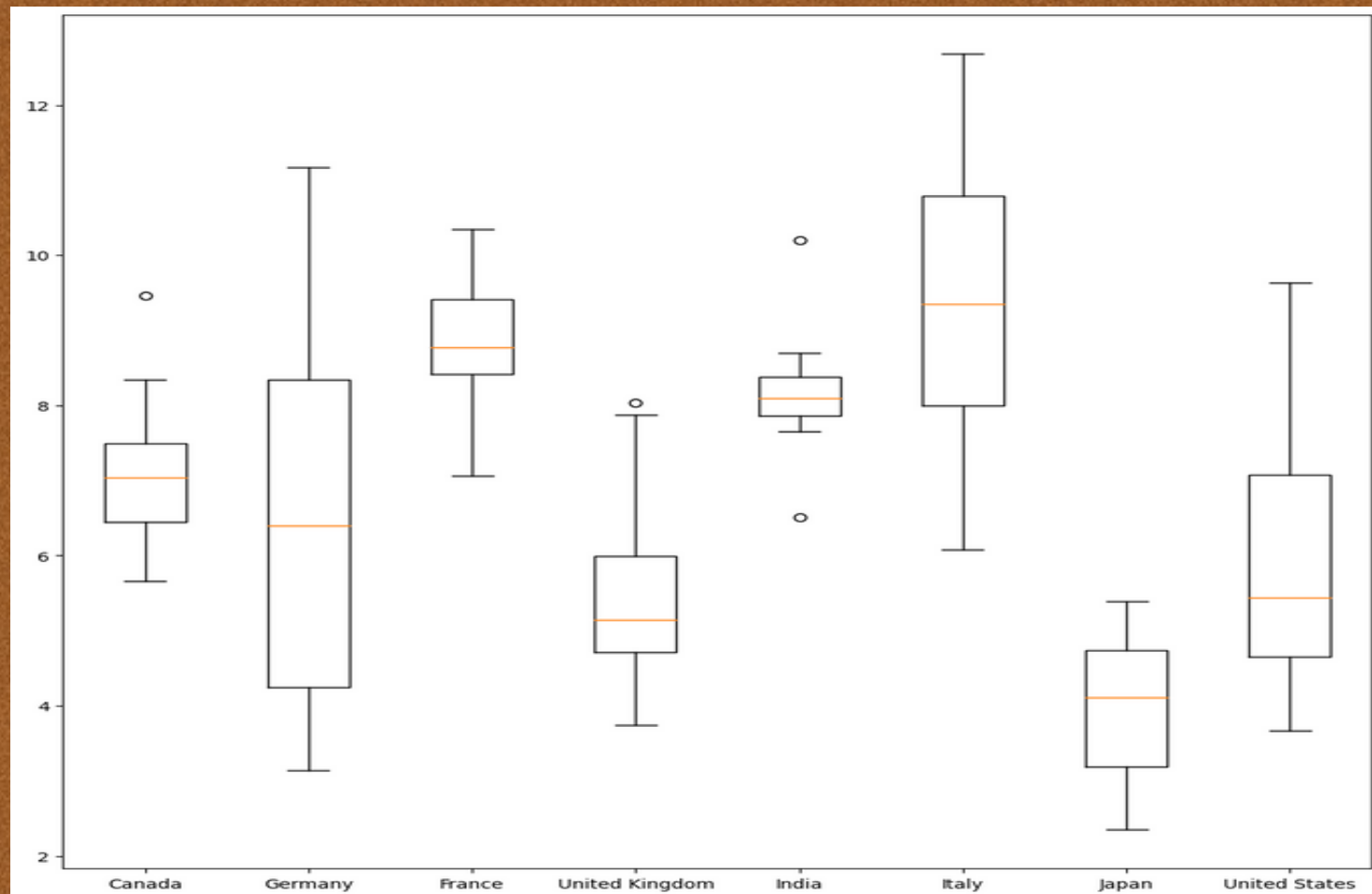
# Exploration Of Datasets



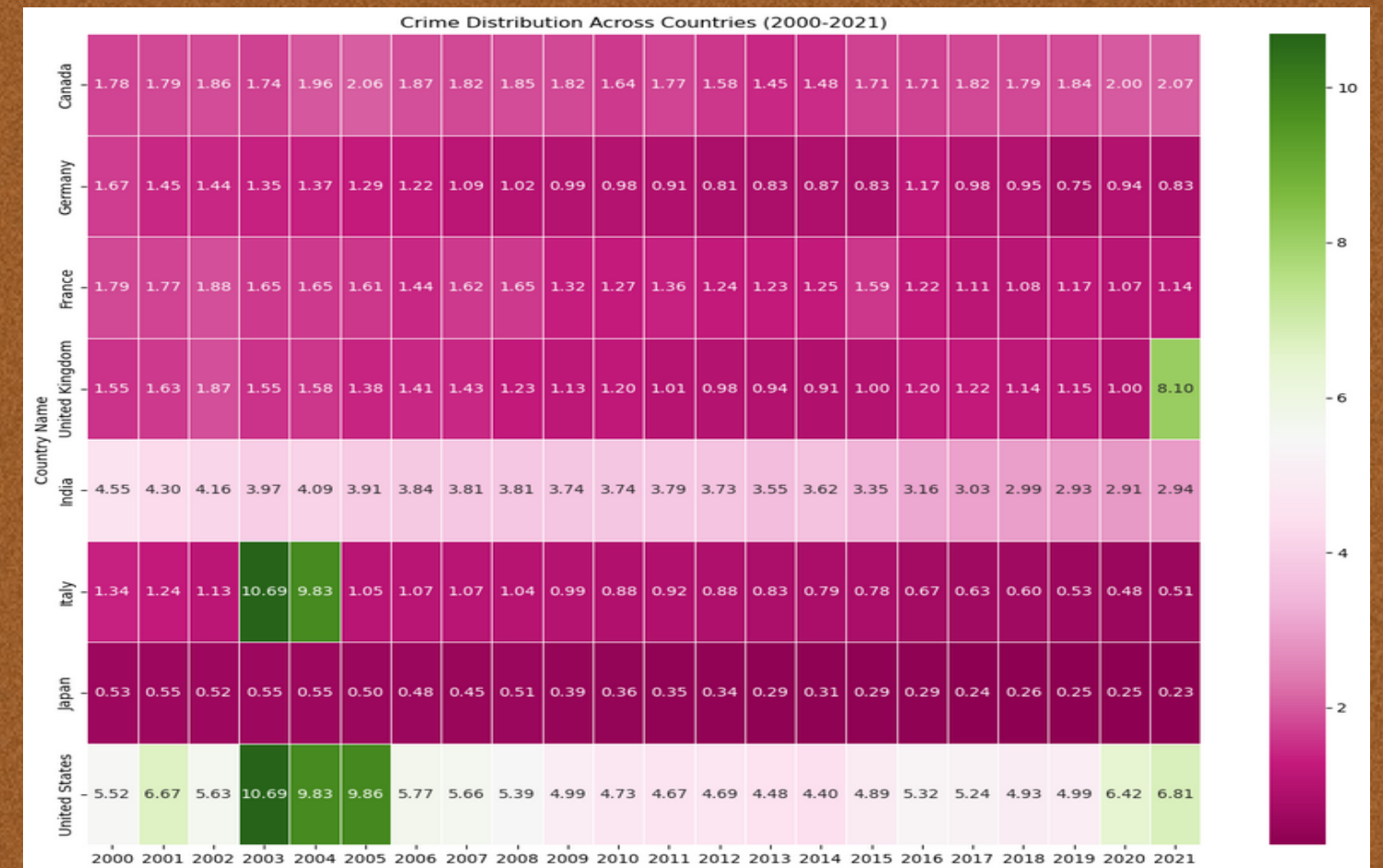
The line plot illustrates notable reductions in unemployment for Germany and Japan, volatility in Italy, an unusual peak in 2020 for most countries except Italy and France, and India experiencing fluctuations in 2019-2020 without significant change over two decades.



# Exploration Of Datasets



The box plot shows minimal spread in unemployment for Canada, France, and India, with outliers indicating unusual extremes; Japan has the lowest rates, the UK and US are left-skewed, and Italy exhibits the highest median unemployment rate.



Heatmap reveals the highest crime rates in the US in 2003 persisting in subsequent years, with India ranking second; 2021 exhibits an anomalous surge in the UK, possibly an outlier, while Italy records unexpected crime rates in 2003 and 2004."



# Results

Correlation analysis for all countries over the years

**1. Germany and Japan:**

**High Positive Linear Correlation**

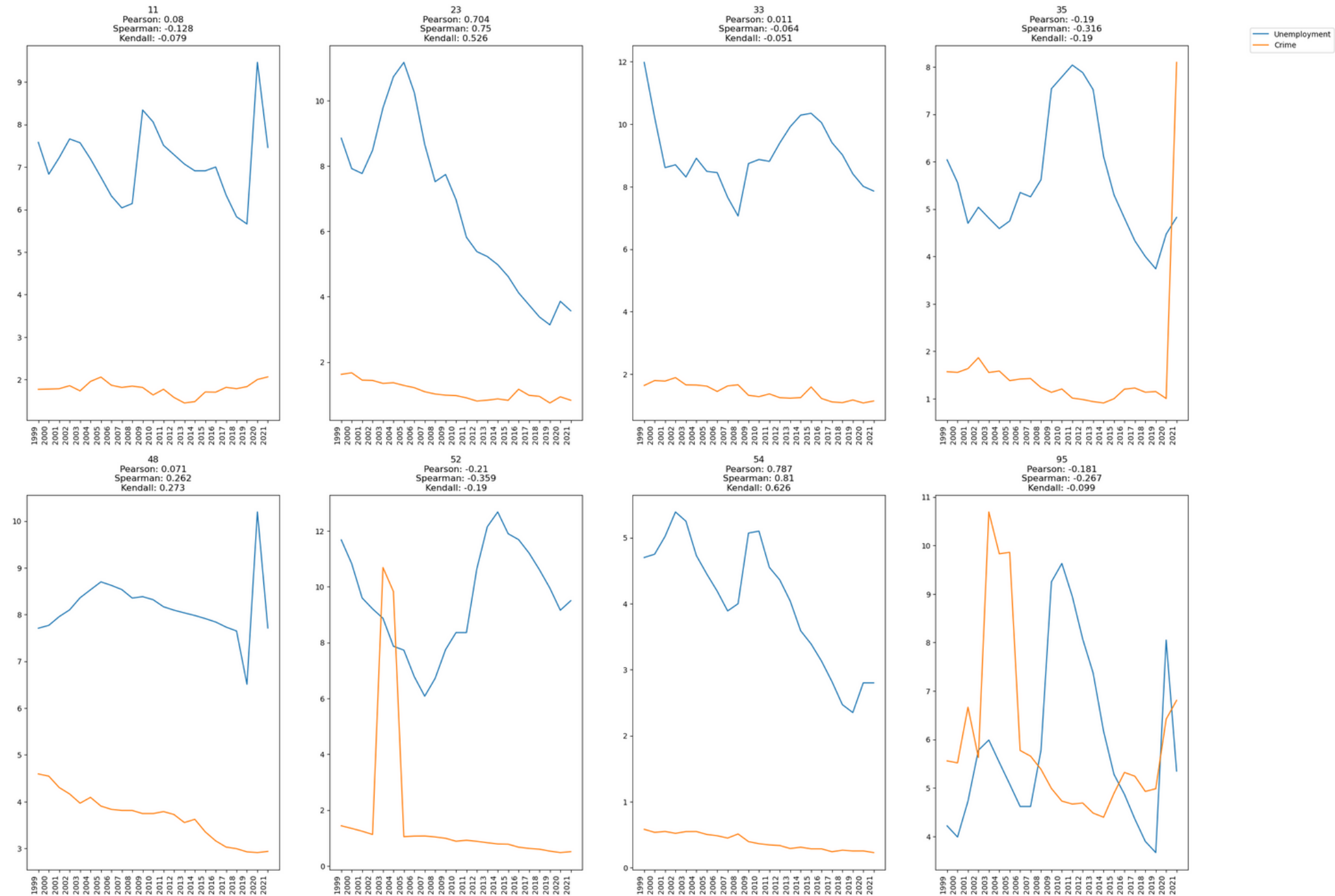
**2. UK, India, Italy:**

**Moderate Spearman Correlation**

**3. Canada, France, US:**

**Less Correlation**

## 1. Unemployment and Crime Relationship for Different Countries over the years





# Results

## Correlation analysis for different years

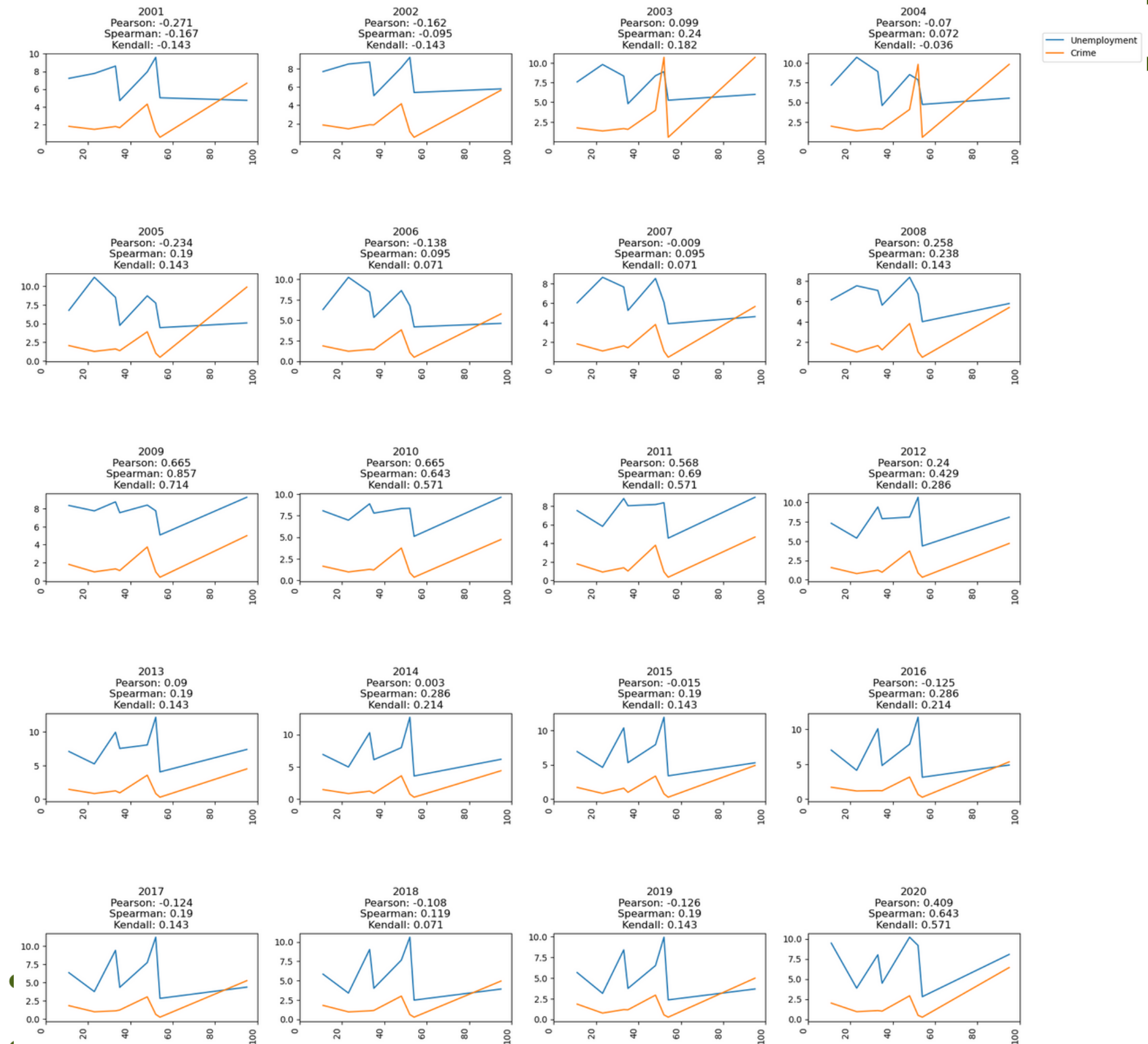
- High Pearson Correlation (2009, 2010, 2011, 2012, 2020):

Moderately Strong Linear Relationship.

- Weak Relationship (2001-2008, 2013-2019):

Weak Correlation

## 2.Unemployment and Crime Relationship for countries in specific year





# Conclusion

- . **Global Variances:** Germany and Japan show a strong link, the UK, India, and Italy have a moderate connection, while Canada, France, and the US exhibit less correlation between unemployment and crime.
- . **Temporal Trends:** Years 2009, 2010, 2011, 2012, 2020 show a strong link, while from 2001 to 2008 and 2013 to 2019, the connection is weak, indicating changing dynamics.
- . **Economic Paradox:** Low unemployment may hide poverty, and high unemployment can coexist with development, emphasizing the complexity of understanding crime dynamics.



# Limitations

1. Causation Complexity: Correlation  $\neq$  causation; unseen variables complicate the relationship between unemployment and crime.
2. Data Reliability Concerns: Findings rely on accurate data; unaccounted seasonal unemployment, unregistered female unemployment, and passive job seekers pose challenges.
3. Crime Data Validity: Homicide data varies; differing assessments and validation processes result in nuanced interpretations.
4. Granularity and Regional Oversight: Lack of specificity in crime types and regional variations within countries may limit insights.



# Future Work

1. Multifactorial Analysis: Examine social policies, education, and economic conditions collectively impacting crime for a nuanced view.

2. Policy Evaluation: Assess diverse policies' influence on unemployment-crime dynamics for effective strategy development.

3. Machine Learning Insights: Use regression trees or neural networks for complex pattern revelation beyond traditional statistics.

4. **Global Comparative Studies:** Expand analysis beyond G7 and India for a broader cross-cultural perspective on unemployment and crime.



