

Data Analysis Report

Michael V Cumbo

December 27, 2023

Abstract

This document presents a comprehensive analysis of the data pertaining to trade unions.

1 Introduction

This paper was written in response to the United Auto Workers strike and the SAG-AFTRA strike of 2023. The goal of this paper is to contextualize the state of trade union power in the United States, blending data analysis with a literature review. The findings in this paper also contextualize union power in a select number of nation-states, adding perspective to the modes of influence individual working-class individuals may have within those states.

2 Methodology

Libraries Used in the Analysis

This analysis utilized several R packages, each contributing unique functions essential for data management, manipulation, visualization, and database interaction. Below is a description of each package and its role in our analysis:

- **tidyverse:** An aggregation of several data manipulation packages, `tidyverse` simplifies many aspects of data analysis. It includes packages like `ggplot2` for data visualization, `dplyr` for data manipulation, and `readr` for data import. Its unified data philosophy makes data science tasks more straightforward and efficient.
 - Installation: `install.packages("tidyverse")`
 - Documentation: tidyverse.org
- **RSQLite:** This package provides a database interface and SQLite driver for R. It allows for seamless integration of SQLite database capabilities within R, enabling the storage, management, and retrieval of large datasets efficiently.
 - Installation: `install.packages("RSQLite")`
 - Documentation: [RSQLite on CRAN](https://www.RSQLite.org)
- **DBI:** The `DBI` package defines a common interface between R and database management systems. It is crucial for establishing database connections and executing database queries.
 - Installation: `install.packages("DBI")`
 - Documentation: [DBI on CRAN](https://dbi.r-lib.org)
- **ggplot2:** A part of the `tidyverse`, `ggplot2` is a powerful and flexible tool for creating elegant data visualizations in R. It is based on the Grammar of Graphics and allows for building plots iteratively.
 - Installation: `install.packages("ggplot2")`

- Documentation: ggplot2.tidyverse.org
- **dplyr:** Also within the **tidyverse**, **dplyr** is used for data manipulation. It provides a set of verbs like filter, select, mutate, and summarize, making data manipulation tasks more intuitive and readable.
 - Installation: `install.packages("dplyr")`
 - Documentation: dplyr.tidyverse.org
- **forcats:** This package, part of the **tidyverse**, is designed for handling categorical variables (factors in R). It provides functions for reordering factor levels, collapsing levels, and changing the display of factor levels.
 - Installation: `install.packages("forcats")`
 - Documentation: forcats.tidyverse.org
- **GGally:** An extension of **ggplot2**, **GGally** provides additional functions and utilities to enhance the capability of **ggplot2**, especially for creating complex multi-plot layouts.
 - Installation: `install.packages("GGally")`
 - Documentation: **GGally** on CRAN
- **stringr:** This package, also a part of the **tidyverse**, simplifies the process of working with strings (text data). It provides consistent and easy-to-use functions for string manipulation.
 - Installation: `install.packages("stringr")`
 - Documentation: stringr.tidyverse.org

These libraries collectively provided the comprehensive toolkit necessary for our data analysis, from data manipulation and querying to visualization and string processing.

2.1 Data Collection

Data was sourced from the International Labor Organization, OECD datasets, and Harvard datasets.

2.2 Data Preparation

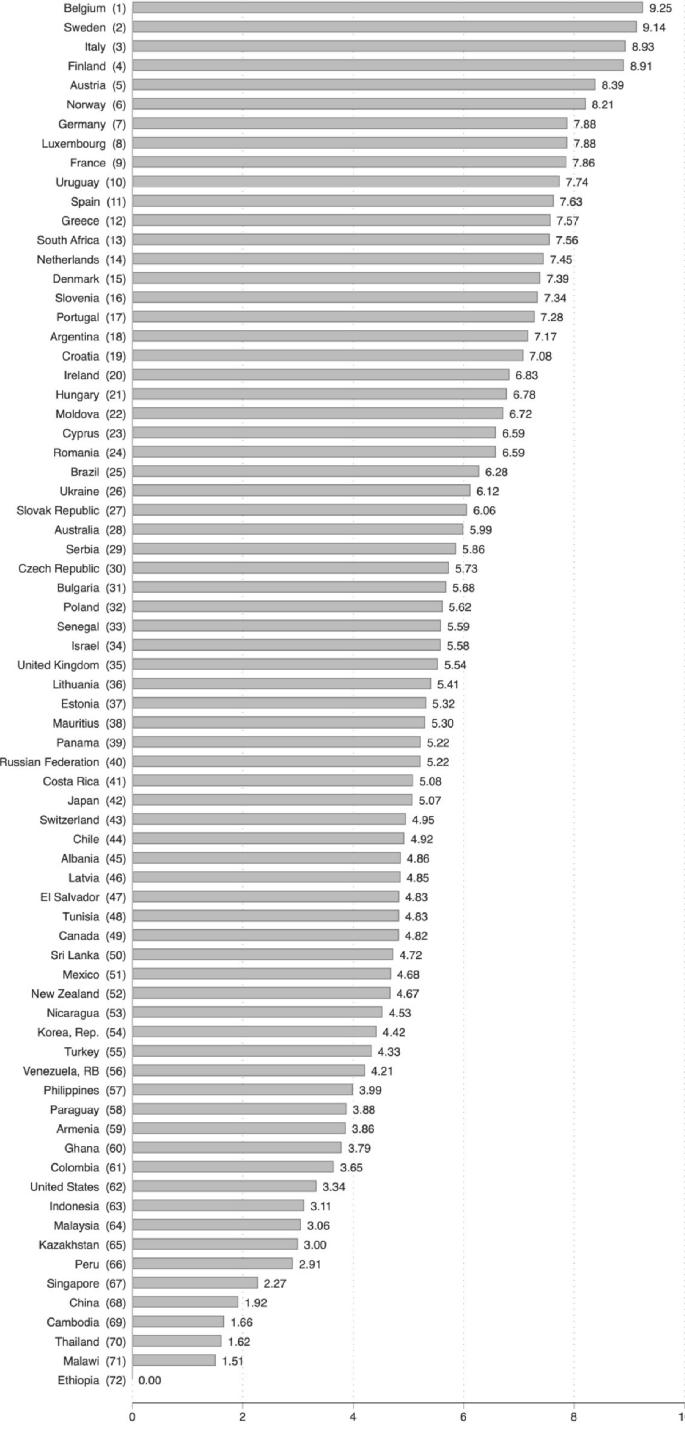
Data was sourced from the appropriate online databases and stored into a SQLite database. The data cleaning required the changing of column names to make the data more readable. Data was extracted from the databases into a tibble and prepared to be analysed using the following code:

```

1 # R code snippet
2 setwd("~/workbook")
3 con <- dbConnect(RSQLite::SQLite(), "trade_union_data.db")
4 dbListTables(con)
5 tudr <- as_tibble(dbReadTable(con, "TUDR"))
6 cbcr <- as_tibble(dbReadTable(con, "CBCR"))
7 collective_bargaining <- as_tibble(dbReadTable(con, "CollectiveBargaining"))
8 trade_union_density <- as_tibble(dbReadTable(con, "TradeUnionDensity"))
9 workplace_rights <- as_tibble(dbReadTable(con, "WorkplaceRights"))
10 state_union_membership_density <-
11   as_tibble(dbReadTable(con, "State_Union_Membership_Density_1964-2021"))
12 dbDisconnect(con)

```

⁰Source: Metten, A. (2021). *Rethinking Trade Union Density: A new index for Measuring Union Strength*. Industrial Relations Journal, 52(6), 528–549. <https://doi.org/10.1111/irj.12347>



Note: The ranking summarizes the average value over all years of the CLF for all countries in the sample and lists them in descending order.

Figure 1: Histogram depicting the distribution of the collective labour force index. The incomplete data availability pushes the Index down to around 72 countries, so the ranking matters less than the Index itself. The picture it paints is the freedom a working person has in the labor force(Union Power). This study is an assessment of union power. It indicates union power by combining and weighing the following stats: Trade Union Density, Collective Bargaining Coverage, Labour Force Participation Rate, Employment in agriculture, Democracy, Core labor rights ratification, hiring and firing constraints, and Hours Regulation. The United States is 62(3.32 index rating) in the Index, and China's place is 68 (1.92 index rating). The data is only recorded from 2010 to 2016.

3 Discussion

For the United States, it is essential to note that the 2020s have seen a growing labor movement, a slow but consistent growth in strikes and labor action, which typically leads to increases in labor union density and collective bargaining coverage (see West Virginia in the 1970s as shown in figure 4, which saw a sharp increase of ten percent union membership due to wildcat strikes). Because of this, there is a strong possibility that the quality of life could improve greatly for working-class people in the US at the end of this decade.

China, when breaking down the stats, has a density of forty percent; however, there is only one union that exists in China, the All-China Federation of Trade Unions (ACFTU), a state-controlled union that allows company chiefs to elect the leadership of the union. It is not really a union, as it is an apparatus of state control. So, things like strikes and collective bargaining agreements are off the table for your average Chinese citizen. See this book for more details on how the ACFTU operates: Fu, Diana (2017). *Mobilizing Without the Masses: Control and Contention in China*. Cambridge University Press.

Unions, and thus the quality of life in the United States, have been decreasing over the past 50 years. The US, for the first time in a century, has a National Labour Relations Board that is friendly toward unions. There are things to look forward to coming to the US (including but not limited to a maglev connecting Baltimore to New York and DC.), and the more active our unions get, the better our quality of life will get. China, however, doesn't have access to mechanisms that can truly redistribute wealth as it is created. Even if the ACFTU wanted to, it would be up to what the company chiefs allow.

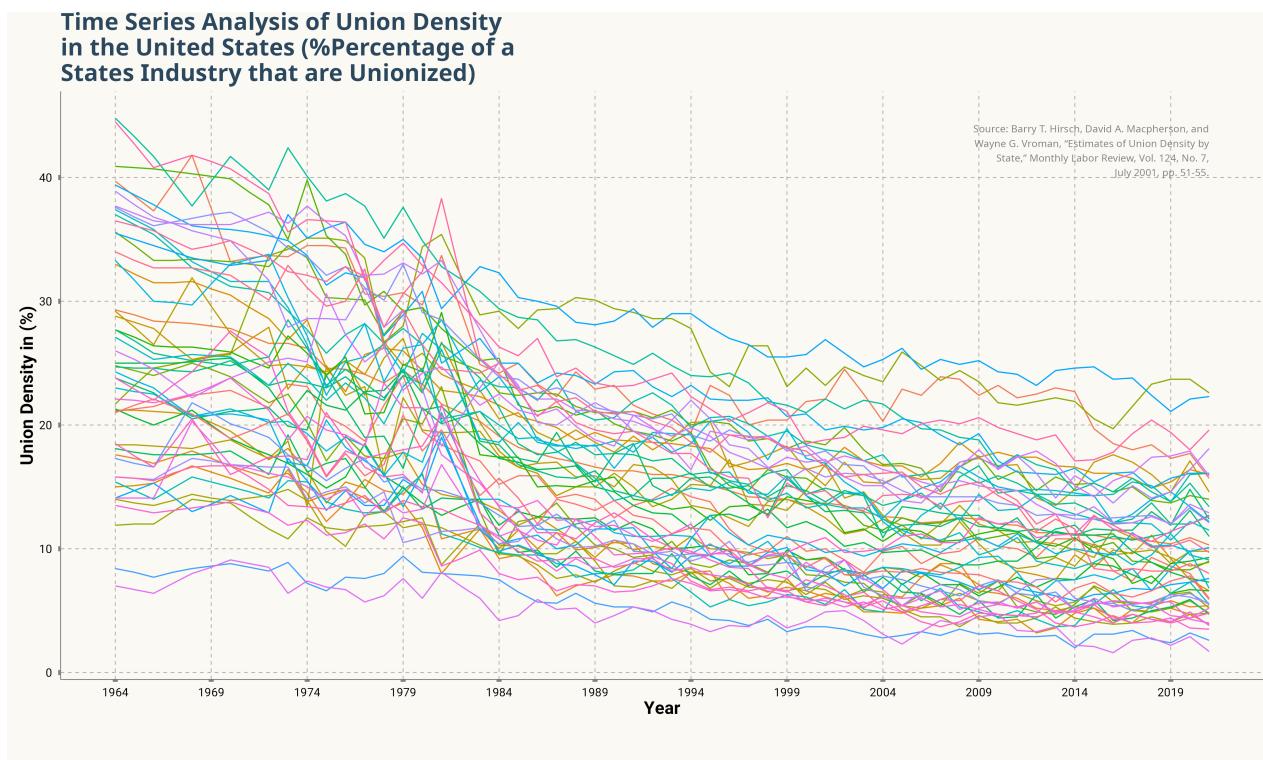


Figure 2: [Detailed caption for Figure 1]

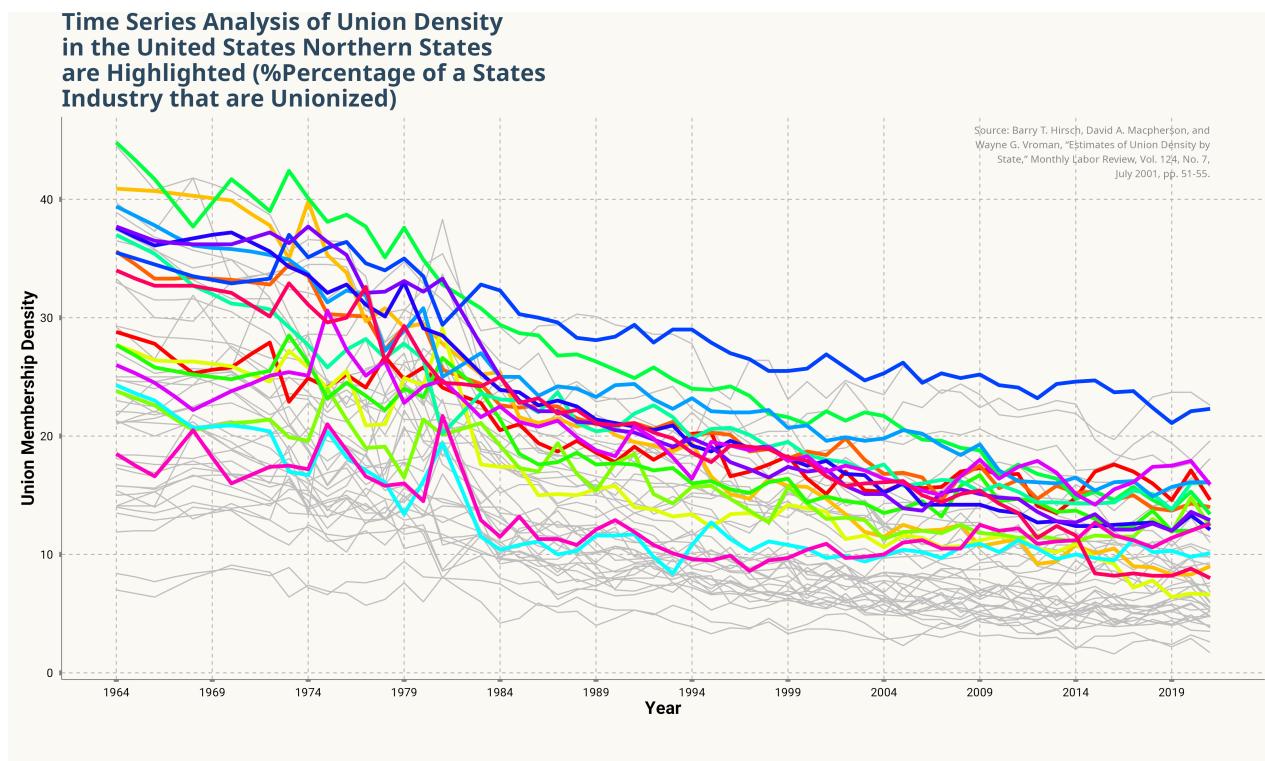


Figure 3: [Detailed caption for Figure 2]

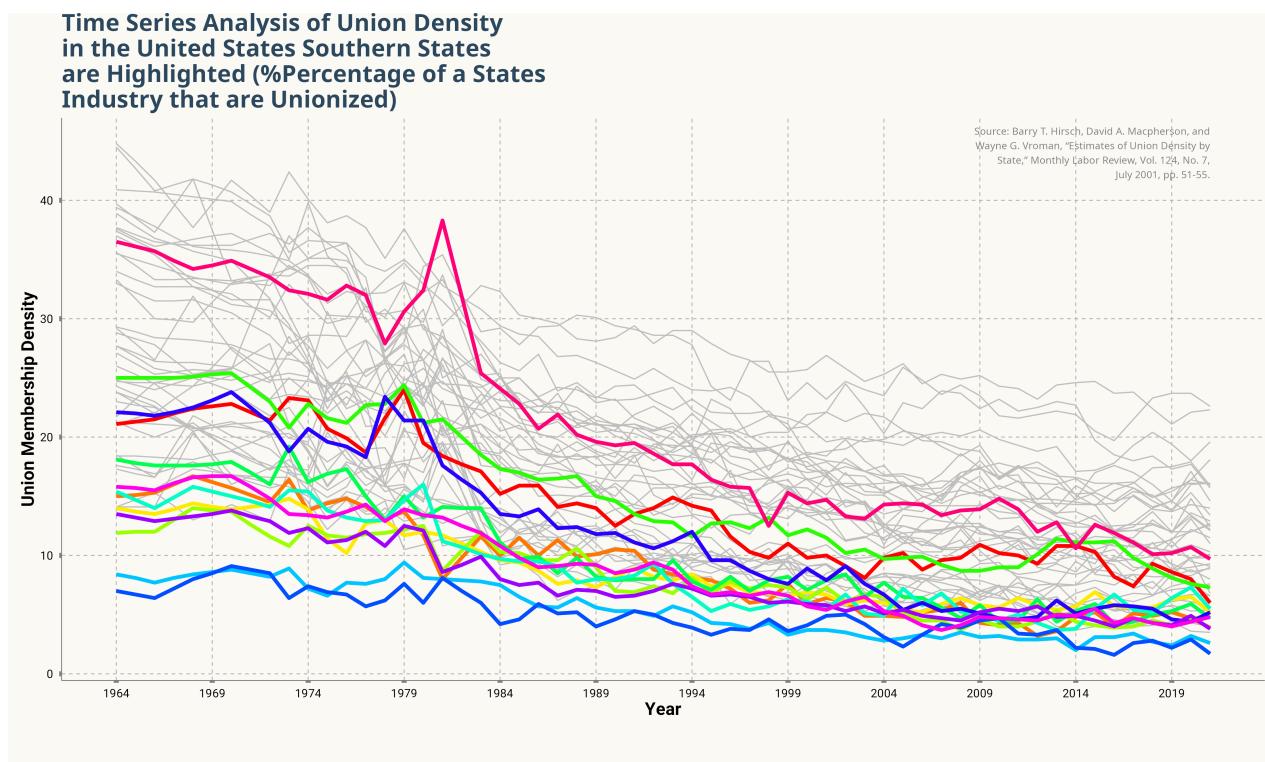


Figure 4: [Detailed caption for Figure 3]

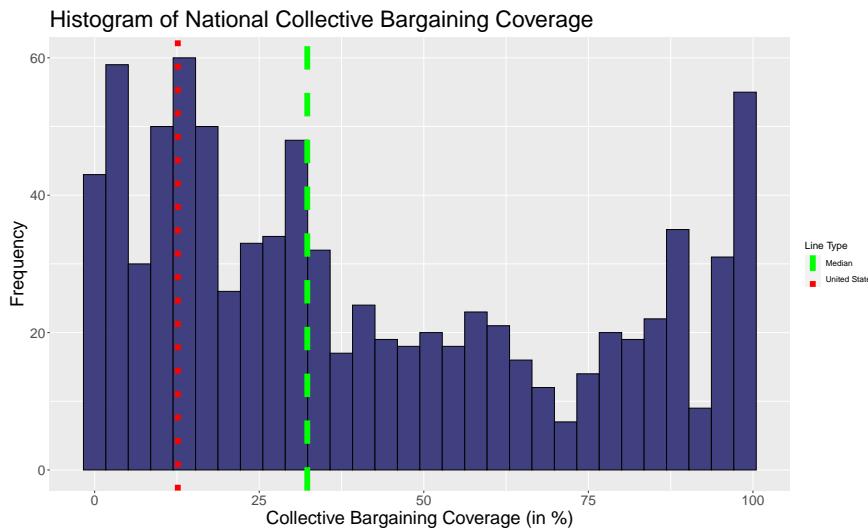


Figure 5: Collective Bargaining Coverage 2000-2019 Histogram depicting the frequency distribution of collective bargaining coverage measured in percent recorded by the International Labor Organization database. The data is grouped by country, highlighting the predominance of collective bargaining coverage in the United States compared to the rest of the world.

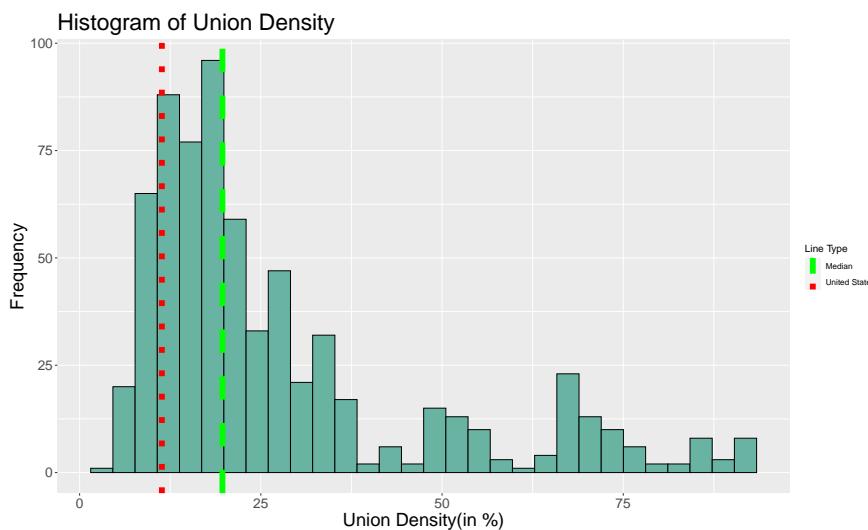


Figure 6: Histogram depicting the frequency distribution of Union Density in percent. The percentage measures a country's total unionized industries; the higher the percentage, the more of that country's workforce is unionized. This data is grouped by country; the mean of the United States and the mean of the total population are marked on this graph.

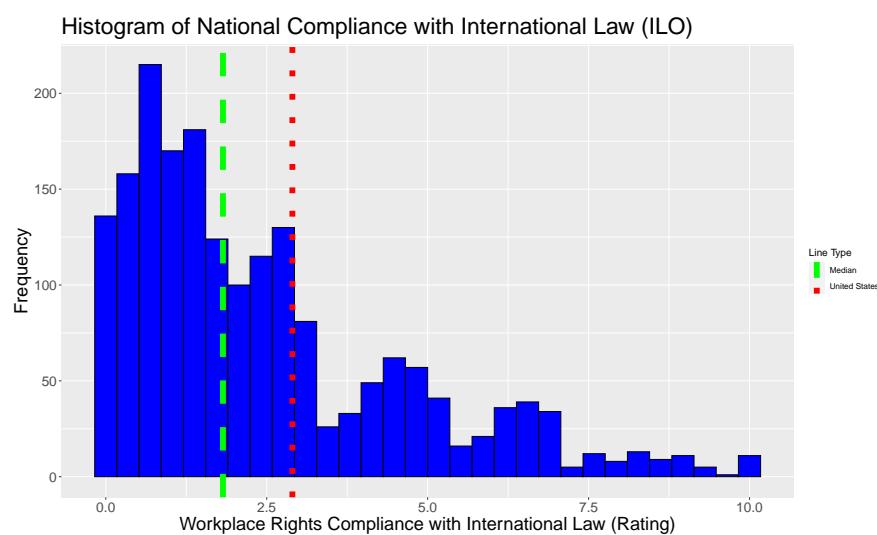


Figure 7: Histogram depicting the frequency distribution of Compliance with international law bargaining coverage as scale from 0 to 10 with 10 being the most out of compliance with international law a country could be, and 0 being completely incompliance with international labor law recorded by the International Labor organization. This data is grouped by country, the united states place is marked in the along with the mean of all countries.

4 Conclusion

Summarize the main findings and their relevance to the initial objectives of your study.

5 References

1. Visser, J. (2021). OECD/AIAS database on Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) [Data set]. OECD. Retrieved from <https://www.oecd.org/employment/ictwss-database.htm>
2. International Labour Organization. (2022). Trade union density rate () – Annual (Id: ILR_TUMT_NOC_RT_A) [Data set]. Industrial Relations Data (IRdata). Retrieved from https://www.ilo.org/shinyapps/bulkexplorer30/?lang=en&id=ILR_TUMT_NOC_RT_A
3. International Labour Organization. (2022). Collective bargaining coverage rate () – Annual (Id: ILR_CBCT_NOC) [Data set]. Industrial Relations Data (IRdata). Retrieved from https://www.ilo.org/shinyapps/bulkexplorer36/?lang=en&id=ILR_CBCT_NOC_RT_A
4. International Labour Organization. (2023). SDG indicator 8.8.2 - Level of national compliance with labour rights (freedom of association and collective bargaining) based on ILO textual sources and national legislation – Annual (Id: SDG_0882_NOC_RT_A) [Data set]. SDG Labour Market Indicators (ILOSDG). Retrieved from https://www.ilo.org/shinyapps/bulkexplorer30/?lang=en&id=ILR_TUMT_NOC_RT_A