



## Data Article

# Pillars of the Global Innovation Index by income level of economies: longitudinal data (2011–2022) for researchers' use

Gonçalo Rodrigues Brás

*IN+, LARSyS, Center for Innovation, Technology and Policy Research, 'Instituto Superior Técnico', 'Universidade de Lisboa', Lisbon, Portugal*

*Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, Coimbra, Portugal*

*ISLA - Instituto Superior de Gestão e Administração de Santarém, Santarém, Portugal*

*DEGEIT, University of Aveiro, Aveiro, Portugal*

*DINÂMIA'CET-Iscte, Centre for Socioeconomic and Territorial Studies, Lisbon, Portugal*

## ARTICLE INFO

### Article history:

Received 13 October 2022

Revised 2 December 2022

Accepted 7 December 2022

Available online 12 December 2022

Dataset link: [Global Innovation Index: panel data \(2011–2022\) \(Original data\)](#)

### Keywords:

Innovation

Innovation data

Income level

National innovation systems

Panel data

## ABSTRACT

It is widely known that the Global Innovation Index reports are of unique value for research purposes. The aim of this work is to provide a panel data file with all pillars of the Global Innovation Index from 2011 until 2022, covering all available economies (149 in total) by income level. After the secondary data was gathered, it was reshaped in an exhaustive process that involved directly importing it from databases or manual insertion. Based on successive Global Innovation Index reports and World Bank data, this work attempts to provide a whole set of data on the incomes of world economies by using Gross Domestic Product per capita based on purchasing power parity (constant 2017 international \$ and current international \$) and Gross National Income per capita in current U.S. dollars (Atlas method). A descriptive analysis is also provided of data and inferences drawn based on the income differences between economies. The data compilation shared here has a singular relevance as it makes a large amount of structured information easier to access. Moreover, data from subsequent years or even from new entries of economies in the Global Innovation Index reports could be added to the data file. As a practical implica-

E-mail address: [goncalo.bras@tecnico.ulisboa.pt](mailto:goncalo.bras@tecnico.ulisboa.pt)

<https://doi.org/10.1016/j.dib.2022.108818>

2352-3409/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

tion, this work should be considered a reliable tool for quantitative research directly or indirectly related with innovation topics (policies, ecosystems, technologies, programmes, among others), as it reduces the time-consuming process of gathering data.

© 2022 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

---

## Specifications Table

Subject	Business, Management and decision sciences/Management of Technology and Innovation
Specific subject area	The panel data describes the pillars of Global Innovation Index by economic income level from 2011 until 2022
Type of data	Tables and Figures
How the data were acquired	Secondary data were automatically imported from an excel file provided by the Global Innovation Index (GII) (2013-2022) and the remaining secondary data were manually inserted based on GII reports (2011-2012). The income of the world's economies from 2011 to 2021 (last year available) was obtained through the World Bank database.
Data format	Raw and Analysed
Description of data collection	Data were reshaped from wide to long Panel data. Excel commands were helpful, but several manual insertions were required in the excel data files to harmonise the names of the economies. Additionally, differences were identified between the excel data files downloaded from the World Intellectual Property Organization website and the GII reports. More specifically, in 2015 and 2016 some economies presented values in the excel data file that were not included in the final GII reports. In order to meet criteria for the inclusion vs. exclusion of economies from the report, only the economies presented in GII reports were maintained.
Data source location	Global Innovation Index <ul style="list-style-type: none"> <li>• World Intellectual Property Organization</li> <li>• City: Geneva</li> <li>• Country: Switzerland</li> </ul> URL: <a href="https://www.globalinnovationindex.org/analysis-indicator">https://www.globalinnovationindex.org/analysis-indicator</a> Income <ul style="list-style-type: none"> <li>- World Bank</li> <li>- City: Washington, D.C.</li> <li>- Country: United States</li> <li>- URLs: <a href="https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD">https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD</a></li> </ul> <a href="https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD">https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD</a> <a href="https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD">https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD</a>
Data accessibility	Repository name: Mendeley Data Data identification number: 10.17632/cvkdzr8tv3.4. Direct URL to data: <a href="https://data.mendeley.com/datasets/cvkdzr8tv3/4">https://data.mendeley.com/datasets/cvkdzr8tv3/4</a>

---

## Value of the Data

- The different dimensions of innovation are now emergent hot topics and researchers have used Global Innovation Index reports as a reliable source for their studies. However, shaping a panel data file to structure information is a time-consuming process, particularly a panel data model covering several years.
- These data are useful because just one excel file brings together very harmonised data on the seven GII pillars (scores/values and rank-ordered) by income, covering all the available economies.

- These data are particularly relevant for researchers directly or indirectly interested in developing quantitative studies about innovation, national innovation systems, innovation ecosystems, innovation policies or innovation technology topics. Such data can obviously be combined with data from other research topics, particularly for quantitative studies. This work aims to act as a starting point to leverage research on innovation and related topics by offering a panel database and therefore bypassing the time-consuming process of initial data collection.
- The panel data are ready for other researchers to use in studies directly or indirectly related with innovation topics. Unlike the successive reports and databases from different years of the GII, the nomenclature of the economies has now been uniformised into a single file from 2011 until 2022.

## 1. Objective

It is known that successive GII reports have been used for research purposes, particularly in domains such as innovation, national innovation systems, innovation ecosystems, innovation policies or innovation technology. The first edition of the GII report was in 2007 and the data has since been publicly available. The data has been subjected to a number of technical and methodological changes over the years, adapting its metrics to the evolution of innovation ecosystems and following some trends in innovation domains. Despite these changes, since 2011 there has been a clear attempt to stabilise the metrics over time and only minor year on year changes have been made. However, there are no known database containing the compilation of GII pillars (scores/values and rank-ordered), at least for the purposes of comparing year on year data (from 2011). This work strives to address this gap by providing a panel data file with all GII pillars from 2011 until 2022 for researchers seeking data on GII scores/values or the economies' GII' rankings.

## 2. Data Description

The analysis was conducted using the panel framework for the period 2011 to 2022 with the seven GII pillars covering 149 economies by income – Gross National Income (GNI) per capita in current U.S. dollars, Gross Domestic Product (GDP) per capita based on purchasing power parity (constant 2017 international \$) and GDP per capita based on purchasing power parity (current international \$). In addition to the overall value of GII, the seven GII pillars are: institutions (I), human capital and research (HC&R), infrastructure (Inf), market sophistication (MS), business sophistication (BS), knowledge and technology outputs (K&TO), and creative outputs (CO). All of them are typical index variables, ranging from 0 to 100 and ranked from 1 (highest performance) to N economies (lowest performance), depending on the economies covered in each year. Considering only these score variables from the GII for the referred period and economies, unbalanced panel data were obtained with 1597 complete observations for each group variable. Some descriptive statistics are available in [Table 1](#).

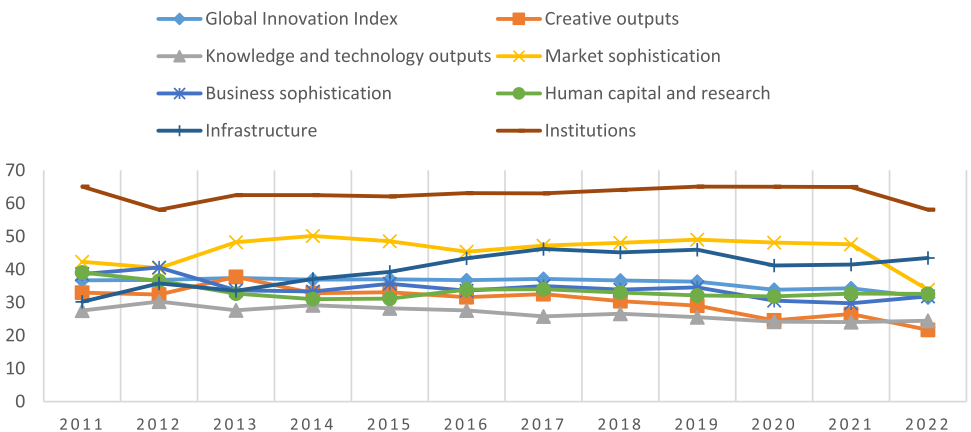
Briefly summarising the measures of central tendency and variability presented in [Table 1](#), the mean and median measures of the variable 'Institutions' (I) present far higher values in relation to other variables; it also presents greater variability in comparison to other variables with a standard deviation of 16.042. Regarding univariate normality, based on kurtosis and skewness criteria defended by Huck [1] in which normality may be indicated up to the absolute value of 1, [Table 1](#) suggests the absence of severe deviations from normality in data-generating process. [Fig. 1](#) shows the mean values (averages) of all variables between 2011–2022.

The mean values of all variables except 'market sophistication' and 'infrastructure' remain very stable between 2011 and 2021. Nevertheless, [Fig. 1](#) also shows a steady decline in some variables in the same period, namely: 'knowledge and technology outputs', 'business sophistication', 'creative outputs', 'human and capital research', and 'global innovation index'. Overall,

**Table 1**

Main descriptive statistics (score variables from GII pillars).

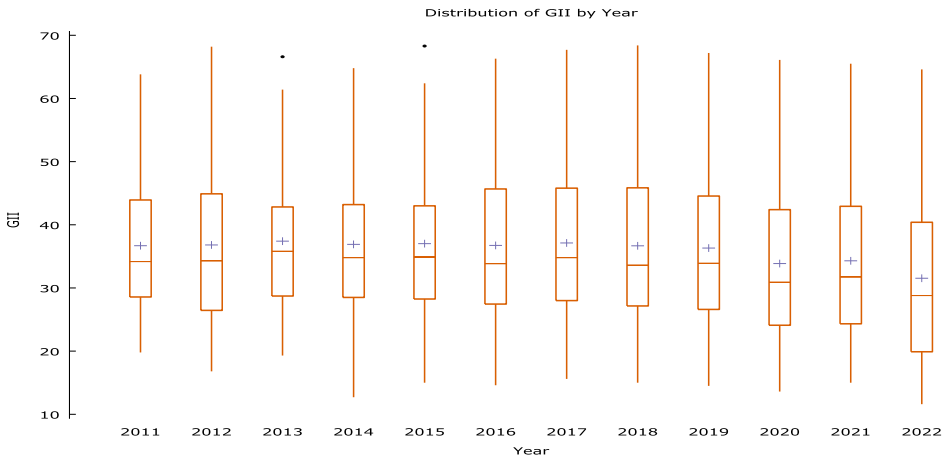
Variable	Mean	Median	Minimum	Maximum
<i>I</i>	62.745	61.100	15.400	95.900
<i>HC&amp;R</i>	33.354	31.400	0.70000	74.700
<i>Inf</i>	40.124	39.500	6.20000	69.900
<i>MS</i>	45.775	44.900	4.40000	88.600
<i>BS</i>	34.272	31.800	8.60000	79.100
<i>K&amp;TO</i>	26.806	23.900	1.60000	74.900
<i>CO</i>	30.529	29.400	0.30000	73.700
<i>GII (overall)</i>	35.963	33.600	11.600	68.400
Variable	Std. Dev.	C.V.	Skewness	Ex. Kurtosis
<i>I</i>	16.042	0.25567	0.13247	-0.62350
<i>HC&amp;R</i>	15.105	0.45286	0.36922	-0.71777
<i>Inf</i>	13.617	0.33938	0.11777	-0.97227
<i>MS</i>	12.827	0.28022	0.33207	0.58611
<i>BS</i>	12.564	0.36659	0.68020	-0.14120
<i>K&amp;TO</i>	13.649	0.50918	0.80723	0.15657
<i>CO</i>	13.774	0.45118	0.22578	-0.35896
<i>GII (overall)</i>	12.068	0.33556	0.57835	-0.52049
Variable	5% Perc.	95% Perc.	IQ range	Missing obs.
<i>I</i>	39.000	90.420	23.850	191
<i>HC&amp;R</i>	11.700	60.810	22.750	191
<i>Inf</i>	19.490	62.610	22.450	191
<i>MS</i>	26.090	68.410	15.300	191
<i>BS</i>	17.600	58.110	17.150	191
<i>K&amp;TO</i>	8.9000	54.810	17.650	191
<i>CO</i>	8.5900	54.000	19.750	191
<i>GII (overall)</i>	19.799	58.510	16.450	191



**Fig. 1.** Evolution of GII and GII pillars (mean values by year).

despite the reduced volatility in all GII pillars between 2011 and 2022, special note goes to the sharp decrease in the 'institutions' and 'market sophistication' pillars between 2021 and 2022. In fact, this significant decrease can be explained by methodological differences in these two pillars in the 2021 and 2022 GII editions. Turning our focus to the main innovation variable, Fig. 2 shows the distribution of GII by year.

Fig. 2 shows that between 2011 and 2019 the central tendency indicators of GII (median and mean values) remain stable followed by a slight downward trend. Regarding the dispersion



**Fig. 2.** Distribution of GII (2011-2022) by quartiles and mean values.



**Fig. 3.** Average values of GII pillars (2011-2021) per income level.

domain, we can see that 2013, 2014, and 2015 exhibit much more homogenous values than the remaining years, where the dispersion is more visible (particularly a right-skewed distribution corresponding to a longer upper tail).

Despite some limitations for country comparisons over time, Gross National Income per capita in current U.S. dollars (Atlas method) is useful as the World Bank (WB) defines different thresholds by year classifying economies into four groups: low-income (1), lower middle-income (2), upper middle-income (3), high-income (4). These thresholds are not invariant in time and also present income level per economy and year in the panel database; the thresholds are also presented in a separate spreadsheet. Hence, Fig. 3 is based on the mean values (averages) for the seven GII pillars between 2011 and 2021 grouped by income level of economies.

Considering the mean values (averages), Fig. 3 shows that high-income economies clearly outperform the other groups of economies in each GII pillar. It also seems clear that low-income

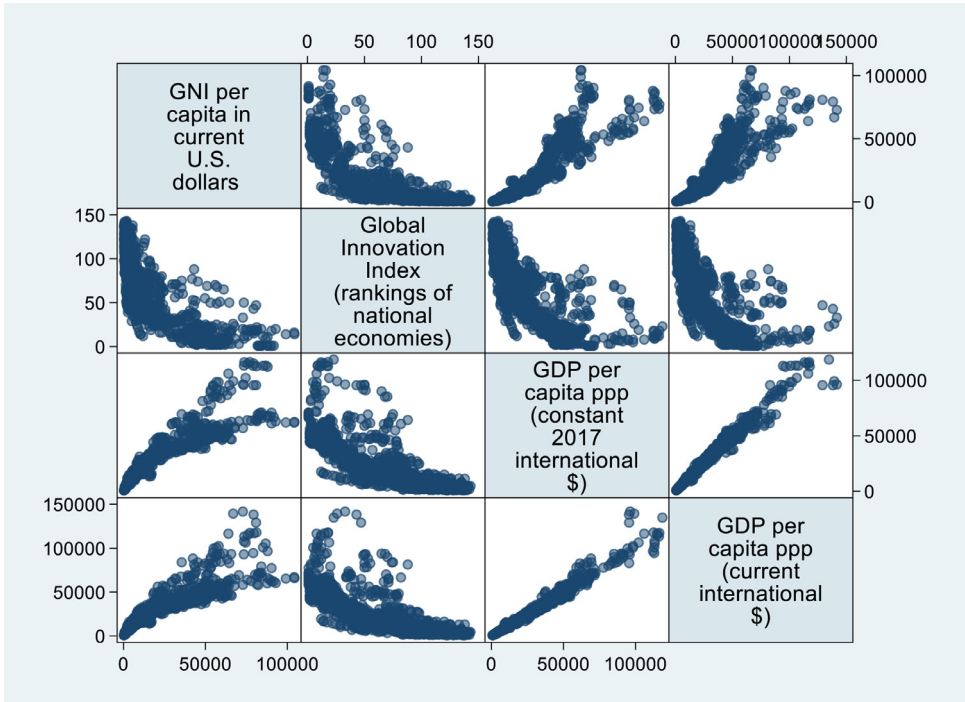


Fig. 4. Distribution of ranked economies in GII per income.

economies and lower middle-income economies present similar mean values (averages) for all the GII pillars.

In addition to the score for each economy, the GII also report ranks economies and thus indicates the performance of innovation systems from best to worst. The GII economies' rankings can also be analysed by income and a description given of some generic patterns or trends in the income measure published by the WB – 2021 was the last available year. Thereafter, when GII rankings are crossed with the income level of economies, the number of observations goes down to 1462 between 2011 and 2021. As the income level of economies is only available until 2021 and Iraq and Mauritania were introduced in the 2022 GII edition, the number of economies falls to 147. Fig. 4 describes the distribution of ranked national economies in GII by income (per capita).

Fig. 4 clearly shows that higher-income economies (per capita) are more likely to achieve the top positions of GII than the lower-income economies (per capita) between 2011 and 2021.

### 3. Experimental Design, Materials and Methods

#### 3.1. Secondary data reliability

The 'World Intellectual Property Organization' (WIPO) now publishes the GII report in partnership with the Portulans Institute, with corporate and academic partners and in collaboration with the GII Advisory Board. However, the GII project started with Professor Soumitra Dutta and the first edition brought INSEAD together with a British magazine called World Business back in 2007. INSEAD has continued to be part of subsequent editions with several partnerships, the most notable of which is the Confederation of Indian Industry; the WIPO's first support of the

project in 2011 is also noteworthy. The 2012 GII report was published by INSEAD and WIPO, and Cornell University joined this partnership to publish editions from 2013 to 2020. Studies on national innovation systems are frequently supported by GII reports [2–5].

The International Bank for Reconstruction and Development, also known as WB, was founded in 1944 and has proved a reliable and reputable source for a wide range of studies across the globe. When GDP per capita based on constant and current PPPs is sourced from the WB, it favours year and country comparisons and is therefore used in various longitudinal studies with panel data models [6–8]; on the other hand, GNI per capita in current U.S. dollars using the Atlas method is often chosen as a measure to define the income level of economies and applied to a range of scientific areas [9–13].

### 3.2. Institutional methods and data comparability

Irrespective of the institutions that have contributed to the different GII reports over time, particularly since 2011, there has been a clear effort to harmonise the data gathered and to standardise methodology. For instance, there are few differences between the seven pillars of the 2011 edition (80 indicators) and the most recent edition (81 indicators in 2022). Although the designation of the sixth pillar changed from 2011 to 2022, the factors did not. Fig. 5 depicts the GII pillars and their corresponding factors by comparing 2011 and 2022.

On the one hand, it is true that the methods used for gathering some indicators may vary slightly from year to year, that some indicators measuring the same factor changed, and even that some new factors appeared (for instance, the introduction of the ecological sustainability factor in the ‘infrastructure’ pillar) while others were removed. However, the main structure of factors has remained almost unchanged since 2011. Therefore, the scores/values from 2011 to 2022 tend to be comparable and, despite slight differences year on year, each GII report continues to be an admirable effort to capture the innovation ecosystem performance and to track the global innovation trends.

On the other hand, it is not possible to confirm there is no common method bias despite the fact that data collection was obtained from the same economies, in the same measurement context, using the same factor context and similar factor characteristics. Hence, a panel data set using ordinal (rank-ordered) scores per year is also provided as a normalisation process that eliminates all potential invariance issues from the GII scores/values.

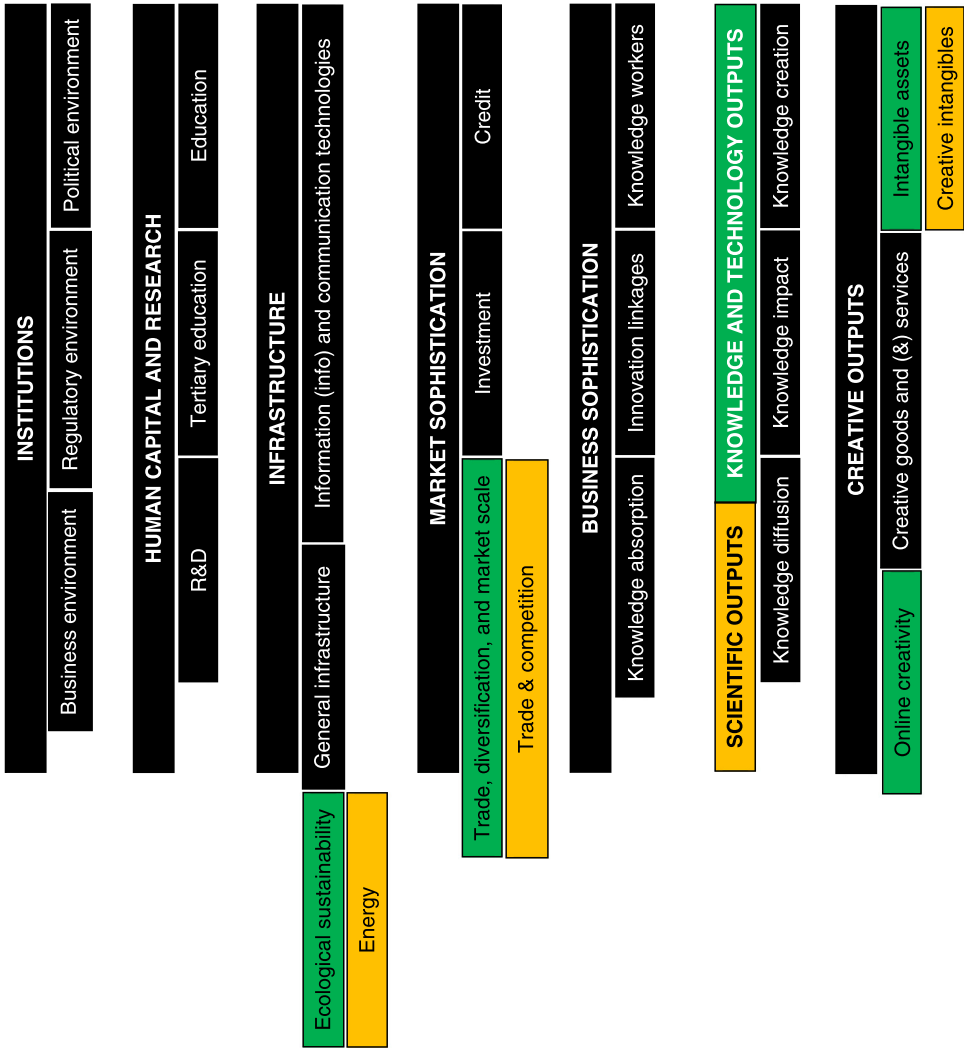
Lastly, whether by constant or current international \$, GDP per capita based on purchasing power parity (PPP) is a widely used measure for country comparisons over time. GNI per capita in current U.S. dollars using the Atlas method has the advantage of defining income groups but is limited for country comparisons over time. The three income measures were collected from the WB database.

### 3.3. Data collection

The data for each edition were collected from the GII excel files available from 2013 to 2022 using the World Intellectual Property Organization website [14]. The data from 2011 and 2012 were based on GII reports [15,16] and introduced manually. Excel commands helped merge all these data into one single file shaped into a panel data file in long format. This was a time-consuming process because some economies had been given different designations<sup>1</sup> in the excel files and the names of others changed<sup>2</sup>. Moreover, some economies were not included in the GII

<sup>1</sup> Examples include ‘Cabo Verde’ and ‘Cape Verde’, ‘United Republic of Tanzania (the)’ and ‘Tanzania, United Republic of’, ‘Venezuela, Bolivarian Republic of’ and ‘Venezuela, Bolivarian Rep.’ or ‘Hong Kong (China)’ and ‘Hong Kong, China’ and ‘Hong Kong’ among many others.

<sup>2</sup> The Former Yugoslav Republic (FYR) of Macedonia until 2018 and thereafter North Macedonia; Swaziland until 2016 and thereafter Eswatini.



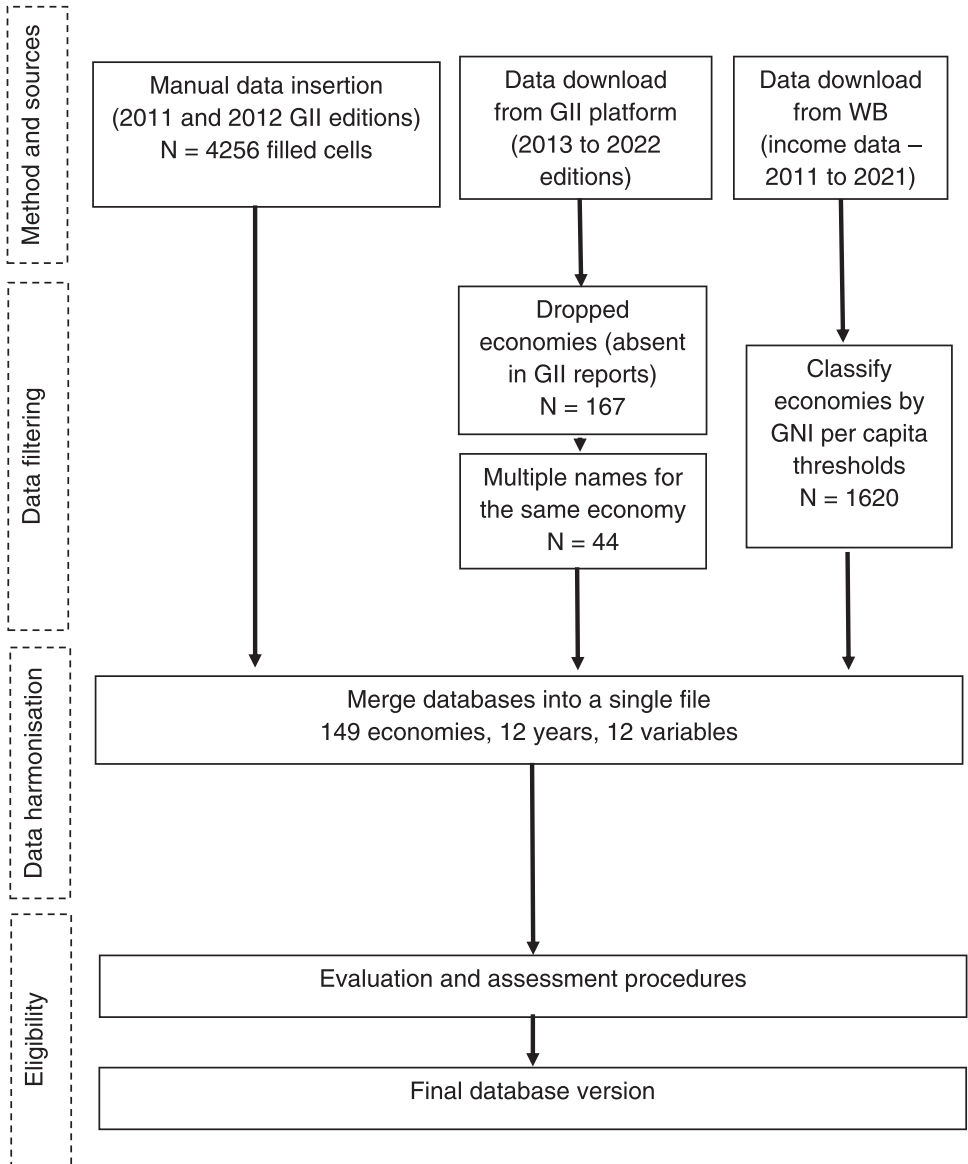
Notes:  Common pillars/factors between 2011 and 2022;  Pillars/factors in 2022

Pillars/factors in 2011

Fig. 5. GII pillars and factors (2011 and 2022).

reports despite having values in excel data files (particularly in 2015 and 2016). Despite understanding the criteria for the inclusion vs exclusion of economies from the GII reports, the excel files imported for 2015 and 2016 had to be corrected – please see economies dropped from analysis for these years in the worksheet entitled ‘incomplete data - removed’; these economies were not included in the analysis (2011-2022) as they only have repeated values for 2015 and 2016 and incomplete data. After these procedures, only economies presented annually in GII reports were considered in this analysis. The income data were imported directly from the WB database. The GNI thresholds in each year were considered and the values were retained by country over different years in order to cluster economies by income.





**Fig. 6.** Information flow diagram of the data collection.

Evaluation and assessment procedures give consistency through comparisons between the data downloaded and GII reports where a gap was found, i.e., although Morocco did not show any rank-ordered score in the download file, it was present in the 2014 GII report. At this stage, several data controls were performed through random sampling and amendments made using some Excel commands to reach a final panel database [17]. Finally, Stata 16 software and Excel were used for some of the data description in the previous main point. Fig. 6 summarises the data collection.

## Ethics Statements

Since this dataset is used for scientific purposes and only contains public information from Global Innovation Index (2011–2022), it meets all terms and conditions required from the mentioned institution.

## CRedit Authorship Contribution Statement

**Gonçalo Rodrigues Brás:** Conceptualization, Data curation, Methodology, Writing-Reviewing and Editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

## Data Availability

[Global Innovation Index: panel data \(2011–2022\) \(Original data\)](#) (Mendeley Data).

## Acknowledgements

The author would like to thank all people and institutions engaged in and committed to publishing GII since its first edition. In addition, he notes this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- [1] S.W. Huck, *Reading statistics and research*, Pearson, Boston, MA, 2012.
- [2] A.J.C. Fernandes, R.G. Rodrigues, J.J. Ferreira, National innovation systems and sustainability: What is the role of the environmental dimension? *J. Cleaner Prod.* 347 (2022) 1–10, doi:[10.1016/j.jclepro.2022.131164](#).
- [3] B. Jankowska, A. Matysek-Jędrzych, K. Mroczek-Dąbrowska, Efficiency of national innovation systems: Poland and Bulgaria in the context of the Global Innovation Index, *Comparat. Econ. Res. Centr. East. Eur.* 20 (3) (2017) 77–94, doi:[10.1515/ceer-2017-0021](#).
- [4] N.F. Crespo, C.F. Crespo, Global innovation index: Moving beyond the absolute value of ranking with a fuzzy-set analysis, *J. Bus. Res.* 69 (11) (2016) 5265–5271, doi:[10.1016/j.jbusres.2016.04.123](#).
- [5] K.-H. Huarng, T.H.-K. Yu, Analysis of Global Innovation Index by structural qualitative association, *Technol. Forecast. Soc. Change* 182 (2022) 1–7, doi:[10.1016/j.techfore.2022.121850](#).
- [6] J.D. Alba, D.H. Papell, Purchasing power parity and country characteristics: Evidence from panel data tests, *J. Dev. Econ.* 83 (1) (2007) 240–251, doi:[10.1016/j.jdeveco.2005.09.006](#).
- [7] B. Lenaerts, M. Demont, The global burden of chronic and hidden hunger revisited: New panel data evidence spanning 1990–2017, *Clim. Change Global Food Secur.* 28 (2021) 1–10, doi:[10.1016/j.gfs.2020.100480](#).
- [8] N. Cachanosky, A. Padilla, A panel data analysis of Latin American populism, *Constitut. Politi. Econ.* 31 (3) (2020) 329–343, doi:[10.1007/s10602-020-09302-w](#).
- [9] T.C. Nokeri, *Classifying Economic Data Applying Logistic Regression*, in: T.C. Nokeri (Ed.), *Econometrics and Data Science: Apply Data Science Techniques to Model Complex Problems and Implement Solutions for Economic Problems*, Apress, Berkeley, CA, 2022, pp. 97–128, doi:[10.1007/978-1-4842-7434-7\\_5](#).
- [10] A. Al-Ansari, M.A. Nazir, Association of body mass index and gross national income with caries experience in children in 117 countries, *Acta Odontol. Scand.* 78 (4) (2020) 303–308, doi:[10.1080/00016357.2019.1704054](#).
- [11] M. Peck, M. Gacic-Dobo, M.S. Diallo, Y. Nedelec, S.V. Sodha, A.S. Wallace, Global Routine Vaccination Coverage, 2018, *MMWR, MMWR Morb. Mortal. Wkly. Rep.* 68 (42) (2019) 937–942, doi:[10.15585/mmwr.mm6842a1](#).
- [12] B. Tekeş, Y. Üzümcüoğlu, C. Hoe, T. Özkan, The relationship between Hofstede's cultural dimensions, Schwartz's cultural values, and obesity, *Psychol. Rep.* 122 (3) (2019) 968–987, doi:[10.1177/0033294118777965](#).

- [13] V.M. Matyushok, N.M. Baranova, L.V. Sorokin, Human capital impact for sustainable economic growth, in: S. Sahdev, R.B. Singh, M. Kumar (Eds.), *Geocology of Landscape Dynamics*, Springer Singapore, Singapore, 2020, pp. 21–36, doi:10.1007/978-981-15-2097-6\_3.
- [14] World Intellectual Property Organization, 2022 Global Innovation Index Data, v10, Url: <https://www.globalinnovationindex.org/analysis-indicator>, (accessed 30 September 2022).
- [15] INSEAD, The Global Innovation Index 2011: Accelerating Growth and Development. [https://www.wipo.int/edocs/pubdocs/en/economics/gii/gii\\_2011.pdf](https://www.wipo.int/edocs/pubdocs/en/economics/gii/gii_2011.pdf), 2011 (accessed 01 September 2022).
- [16] INSEAD and WIPO, The Global Innovation Index 2012: Stronger Innovation Linkages for Global Growth. [https://www.wipo.int/edocs/pubdocs/en/economics/gii/gii\\_2012.pdf](https://www.wipo.int/edocs/pubdocs/en/economics/gii/gii_2012.pdf), 2012 (accessed 03 September 2022).
- [17] G.R. Brás, Global Innovation Index: panel data (2011–2022), 2022 Mendeley Data, v4, doi:10.17632/cvkdzr8tv3.4.