

spiR, An R package to access social progress data

Thierry Warin^{1, 2}

DOI: 1 Professor, HEC Montral (CANADA) 2 Principal Investigator, CIRANO (Montreal, Canada)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted:

Published:

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

‘spiR’ is an R wrapper to access the Social Progress Index datasets easily. The spiR package aims to make two contributions: (1) to give access to new societal and economic data and (2) to integrate easily in a researcher’s workflow through the R language.

Providing new open data in an integrated way highlights some other principles: open data, open code, open and reproducible science. The novelty of the data, collected and computed by the Social Progress Imperative, whose advisory board is chaired by Professor Michael Porter (Harvard Business School), helps inform better policy decisions to accelerate efforts to drive equitable, inclusive and thriving societies.

Updated quarterly, the R package contains 72 metadata, covers 194 countries from 2014 and is rich of 1434 observations. The update is done by collecting, cleaning and joining different datasets resulting in a unique cross-section times-series final dataset. The scripts to do perform these tasks have to be run outside of the package itself since the primary sources are not always in the same location or format. Hence, the value added by this package to propose a stable final product with a panel perspective as a result of the prior data treatment. spiR is also expanding by providing access to other data sources to augment geographic granularity while making sure data comparisons are consistent.

Through its functions, spiR is very easy to use for a researcher in social science, for instance.

Statement of need

The spiR package was developed to facilitate access to new indicators measuring social progress while being in the open data spirit. ‘spiR’ is an R package to access the Social Progress Index datasets easily. It is inspired by several other initiatives in the open data space (Kim et al., 2019).

This R package provides a collection of new indicators to compute complexity metrics that offer easy access to these data for social scientists (Vargas, 2020). Indeed, in the past decade, a new field has emerged: Economic Complexity. Economic Complexity is a data-driven approach that may inform the territorial development policies with evidence-based metrics. An interdisciplinary analytical framework by essence, Economic Complexity stands at the crossroads between evolutionary economics and institutional economics (Cimoli & Dosi, 1995; Hirschman, 1958; Teece, Rumelt, Dosi, & Winter, 1994). It was initially designed to analyze the determinants of economic development. Driven by an inferential approach, the need for new indicators and new data has occurred. In our current technological “state of the art”, it is crucial for social scientists to take advantage of the new possibilities. Social scientists can leverage technologies to access new kinds of data (structured and unstructured) and compute new indicators. The goal is the same: to better understand and inform individuals in their interactions with societies and societies themselves.

The differences with the previous literature are essentially twofold. This literature proposes to consider different dimensions (industries and product spaces) at the theoretical level, shifting the focus away from only aggregate variables (Hidalgo & Hausmann, 2009; Tacchella, Cristelli, Caldarelli, Gabrielli, & Pietronero, 2012). The second difference is the result of researchers having better access to computing power and data. Indeed, concerning the previous literature, Economic Complexity is by nature data-intensive and computing-power intensive. With the current acceleration in computing power access and open data initiatives worldwide, Economic Complexity can now be used to provide complementary analyses to more conventional macroeconomic methods. The shift in granularity allowed by this new access to data allows researchers to answer several policy-relevant questions quantitatively.

In 2015, The 17 United Nations' Sustainable Development Goals were adopted. 'spiR' is a wrapper of several open datasets published by the Social Progress Imperative (<https://www.socialprogress.org/>), including the Social Progress Index (a synthetic measure of human development across the world). 'spiR' 's goal is to provide data in a stable environment to help policymakers and researchers prioritize actions that accelerate social progress across the world in the context of the Sustainable Development Goals. The Social Progress Index proposes a new perspective on social challenges and needed efforts to accelerate social progress with the Sustainable Development Goals. In this context, 'spiR' aims to allow an easy connection to social scientists to this new way of measuring social progress.

The index measures the quality of life for 98% of the world's population. In its current version, the R client provides access to global data. In different versions, this R client will include different geographical levels: states, regions, cities and sometimes communities.

Three overarching dimensions are (1) Basic Human Needs, (2) Foundations of well-being and (3) Opportunity. Within each dimension, four components further divide the indicators into thematic categories. The index consists of 51 social and environmental variables, covering the years 2014 to 2019.

spiR covers 72 metadata (mostly economic and societal variables) and contains 1434 observations for each metadata from 2014 to 2019 across 194 countries. These data are easily accessible through three functions: (1) to select the country, (2) to select the indicator(s) and (3) to collect the data in a directly usable data frame.

Acknowledgments

The author would like to thank the Center for Interuniversity Research and Analysis of Organizations (CIRANO, Montreal) for its support, as well as Thibault Senegas, Marine Leroi and Martin Paquette. The usual caveats apply.

References

- Cimoli, M., & Dosi, G. (1995). Technological Paradigms, Patterns of Learning and Development: An Introductory Roadmap. *Journal of Evolutionary Economics*, 5(3), 243–268.
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575. <https://doi.org/10.1073/pnas.0900943106>
- Hirschman, A. O. (1958). *The strategy of economic development*. Yale University Press.
- Kim, H., D'Orazio, V., Brandt, P., Looper, J., Salam, S., Khan, L., & Shoemate, M. (2019). UTDEventData: An R package to access political event data. *Journal of Open Source Software*, 4(36), 1322. <https://doi.org/10.21105/joss.01322>

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A New Metrics for Countries' Fitness and Products' Complexity. *Scientific Reports*, 2(1), 1–7. <https://doi.org/10.1038/srep00723>

Teece, D. J., Rumelt, R., Dosi, G., & Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, 23(1), 1–30. [https://doi.org/10.1016/0167-2681\(94\)90094-9](https://doi.org/10.1016/0167-2681(94)90094-9)

Vargas, M. (2020). Economic complexity: Computational Methods for Economic Complexity. *Journal of Open Source Software*, 5(46), 1866. <https://doi.org/10.21105/joss.01866>