

Associate Editor Letter

Tutorial-Articles: The Importance of Data and Code Sharing

Artigos-tutoriais: A Importância do Compartilhamento de Dados e Códigos



Henrique Castro Martins^{1,2}

■ ABSTRACT

Context: this document is designed to be along with those that are in the first edition of the new section of the *Journal of Contemporary Administration* (RAC): the tutorial-articles section. **Objective:** the purpose is to present the new section and discuss relevant topics of tutorial-articles. **Method:** I divide the document into three main parts. First, I provide a summary of the state of the art in open data and open code at the current date that, jointly, create the context for tutorial-articles. Second, I provide some guidance to the future of the section on tutorial-articles, providing a structure and some insights that can be developed in the future. Third, I offer a short R script to show examples of open data that, I believe, can be used in the future in tutorial-articles, but also in innovative empirical studies. **Conclusion:** finally, I provide a short description of the first tutorial-articles accepted for publication in this current RAC's edition.

Keywords: tutorial-article; open data; R script; empirical research; open science.

■ RESUMO

Contexto: este documento foi escrito para compor a primeira edição da nova seção da *Revista de Administração Contemporânea* (RAC): a seção de artigos-tutoriais. **Objetivo:** o objetivo deste artigo é apresentar a nova seção e discutir tópicos relevantes a um artigo-tutorial. **Método:** este documento é dividido em três partes principais. Primeiro, oferece-se um resumo das mais importantes práticas de dados abertos e materiais abertos atualmente que, conjuntamente, criam o contexto ideal para artigos-tutoriais. Em seguida, oferecem-se diretrizes para o futuro da seção e algumas ideias que podem ser desenvolvidas no futuro. Em seguida, oferece-se um protocolo de pesquisa em R com exemplos de bases de dados abertas, que, acredita-se, podem ser relevantes para artigos-tutoriais futuros, mas também para estudos empíricos diversos. **Conclusão:** finalmente, ofereço uma breve descrição dos artigos-tutoriais aceitos na presente seção da RAC.

Palavras-chave: artigos-tutoriais; dados abertos, R script; pesquisa empírica; ciência aberta.

1. Pontifícia Universidade Católica do Rio de Janeiro, IAG - Escola de Negócios, Rio de Janeiro, RJ, Brazil.
2. Associate Editor for Tutorials and Open Data of the Journal of Contemporary Administration (RAC).

Como citar: Martins, H. C. (2021). Tutorial-articles: The importance of data and code sharing. *Revista de Administração Contemporânea*, 25(1), e200212. <https://doi.org/10.1590/1982-7849rac2021200212>

INTRODUCTION

It is almost 2021, and the discussion about open research data, transparency, credibility, and research funding has been extensive in the past few years in the academic research communities of business and management (Babin, Ortinau, Herrmann, & Lopez, 2020; Beugelsdijk, Van Witteloostuijn, & Meyer, 2020). The list of journals debating and increasing practices related to what is known as open science (OS) grows by the minute and, at the moment, includes *Management Science*, *Journal of International Business Studies*, *Journal of Marketing*, *Journal of Business Research*, *Strategic Management Journal*, *Organization Science*, *Journal of Management*, and many others. Lists like this one are not a novelty anymore. They merely represent a movement that started in other fields, for instance, medical, psychology, and computer science fields, that now is becoming more frequent in management.

In times where data is abundant — in marketing (Sheth & Kellstadt, 2020), in finance (Begenau, Farboodi, & Veldkamp, 2018), and in accounting (Bhimani & Willcocks, 2014) — and transparency is key (Beugelsdijk et al., 2020; Mendes-da-Silva, 2019; Mendes-da-Silva & Leal, 2020), articles with open data, open code, and reused data become essential to the research community. All these practices have numerous benefits to researchers (increasing visibility, reputation, and citation), to students (decrease learning efforts and time), and to the general society (increase transparency and decrease the cost of science) (Drachen, Ellegaard, Larsen, & Dorch, 2016; McKiernan et al., 2016; Pampel & Dallmeier-Tiessen, 2014; Piwowar & Vision, 2013).

On top of that, there is what RAC is calling tutorial-articles (Martins & Mendes-da-Silva, 2020). Though tutorial-articles are not typically listed among the ‘core’ OS practices (Banks et al., 2019; McKiernan et al., 2016), they are essential to transparency and credibility. The concept of an article providing a tutorial of some sorts is not new and can be traced back to the study of Hax and Majluf (1982) when they say “the objective of a tutorial article is to describe an important technique or an application area for ... readers who are non-experts in the field” (Hax & Majluf, 1982, p. 50). In a more digital world, with more powerful technological resources and big data, tutorial-articles certainly play a role in academic research. Although the management field does not have a history on tutorial-articles, in other areas such as psychology and computing sciences, tutorial-articles are far from being a novelty (see, for instance, the *Journal of Mathematical Psychology*).

Building upon OS concepts, the purpose of this document is fourfold. First, I provide a summary of what I understand as the state of the art of OS, as it is in late

2020/early 2021, which creates the context for tutorial-articles. Second, I provide some guidance to the future of the new RAC’s section on tutorial-articles, providing a structure and some insights that can be developed in the future. Third, I show examples of open data available in different places and produced by various sources that can be used in management research studies and, perhaps more importantly, in future tutorial-articles submitted to RAC. Alongside these examples, I provide a short R script with the commands necessary to create the outputs (i.e., to collect the data required to create the graphs) included in this document. Fourth, I provide a short description of the tutorial-articles published in the current RAC’s edition, which inaugurates the tutorial-articles section of RAC. The last section contains some final thoughts about OS and tutorial-articles.

THE CURRENT STATE OF THE ART IN OPEN DATA AND OPEN CODE

Data sharing in different fields

I would like to start this section by highlighting the differences between fields in open science practices to this date. In a recent study published in the journal *Scientometrics*, Rousi and Laakso (2020) analyze 120 A-Journals in the fields of (a) neuroscience, (b) physics, and (c) business operations research. They concluded the following:

The results affirm that considerable differences between research fields remain when it comes to policy existence, strength, and specificity. The findings revealed that one of the most important factors influencing the dimensions of what, where and when of research data policies was whether the journal’s scope included specific data types related to life sciences which have established methods of sharing through community-endorsed public repositories (Rousi & Laakso, 2020, p. 131).

This extract only reinforces what all researchers in the management field intuitively know: we are far behind fields as neuroscience and medicine (i.e., life science fields) as we still do not have widely accepted protocols about data sharing. However, the adoption of data sharing policies is not but increasing in the management field over the last few years too, and, like anything that is in its infancy, the upside potential is tremendous (Banks et al., 2019).

Code sharing

In a recent editorial of the *Critical Finance Review*, Spiegel (2019) draws our attention to the fact that codes

and research protocols are usually not well structured. Thus, if journals start demanding the publication of codes, authors would have additional demands before publication (demands that add up to other highly demanding efforts such as writing the research per se, but also the replies to referees, applications, exams, and many others). I believe he has a point. Though intermediate codes and research protocols are usually not yet available on a large scale to the public, the general feeling is that they are indeed not well structured in management research. Authors typically do not have the necessary training, and there are no guidelines for writing clean codes. To help in that arena, in [Martins \(2020\)](#), I provide a basic first step with generic questions intended to assist researchers in reviewing their codes before submitting to RAC.

On top of that, it is not uncommon that students reach some intermediate estimate that cannot be replicated by the supervisor because the student did not track all modifications and steps made throughout the process of reaching such an estimate. This is a problem at the micro-level because the student has to rerun old estimations, but also at the macro-level since it hampers the replication of published studies. Perhaps more relevant, the practice of not writing clean codes can lead to ingenuous mistakes in the use of data, and, therefore, to incorrect results and implications, fueling the low credibility cycle that management and business research face ([Aguinis, Cascio, & Ramani, 2017](#)).

I see two primary practices helping to overcome these difficulties. First, by definition, a tutorial-article needs to share clean codes written by more experienced authors with the main goal of showing how to estimate a particular topic focusing (but not exclusively) on the issues faced by beginners. Thus, tutorial-articles can help authors to improve their codes by mirroring the best practices of more experienced researchers. Second, the best practices seem to be pointing to an equilibrium where open data and open materials receive a DOI number and, thus, can be cited. This is perhaps the best argument to make authors start writing clean codes and sharing them with colleagues. On top of that, it is never late to say that previous research shows that articles that shared code receive more citations than articles that did not share them ([Vandewalle, 2012](#)). A final comment on the same topic: code sharing is mandatory in Nature-journals ([Kousta, Pastrana, & Swaminathan, 2019](#)) and, by December 2020, is going to be mandatory in RAC too.

Data management

Accompanying the discussion about code sharing is the need for structured data management protocols ([Beugelsdijk et al., 2020](#)). There are two layers where this discussion occurs. At the macro-level, we see cross-national council (for instance, the [European Research Council, 2019](#))

discussing topics such as the property rights of data, and ethical and legal issues of shared data, the implementation of the FAIR principles, and providing guidelines to the whole community of data creators and data users.

However, it is at the micro-level that I would like to make some more nuanced comments. Similar to the code sharing issues commented by [Spiegel \(2019\)](#), more often than not, management researchers do not have clear protocols on how to manage research data. Spreadsheets can quickly become unbearable to anyone except the researcher that started handling the data. As open data becomes more and more critical in our community, we need to pay more attention to how we handle data. Some simple steps include (a) never saving changes in datasets, only in codes and research protocols, (b) use, as often as possible, the names of the original variables as defined by the primary data source to facilitate replication by other researchers, (c) document all micro-decisions made along the research process with comments and notes in the code (especially those micro-decisions that usually are not discussed in the main text, such as merging and appending, generic manipulation of data, and several others), and (d) even if the research data is not directly linked to a particular study, use public depositories (which are free) to store and to facilitate controlling the version of a specific dataset.

These are simple steps that, I believe, will help authors of future tutorial-articles published by RAC to comply with the requirements of a tutorial-article. These simple steps also have positive externalities for all remaining types of articles. Finally, I must add that I find highly relevant a tutorial-article written with the primary purpose of discussing data management; this shall be a valuable contribution to our community.

P-hacking, HARKing (hypothesizing after the results are known), and positive publication bias

A significant set of issues that recent editorials and articles are paying more attention to are those related to p-hacking, and HARKing (hypothesizing after the results are known), which leads to positive publication bias ([Aguinis et al., 2017](#); [Babin et al., 2020](#); [Harvey, 2017](#)). Considering a lack of a universal definition, one could say that p-hacking is the practice of changing systematically a model or the variables used in a model to find p-values that are significant at typical thresholds such as the 1% or 5%, “leading” the results to be “significant” (emphasis on the quote-unquote here, since this significance is likely to be found by chance). Along similar lines, HARKing is the practice of writing (or changing) the theory after knowing the results to increase the fit between theory and results. While one practice does

not exclude the other, both lead to what is commonly called as positive publication bias, which is the empirical observation that published papers, more often than not, contain significant results and hinder those results that are non-significant.

In the 2016 Presidential Address of the American Finance Association, [Harvey \(2017\)](#) discusses the bias in t-statistics of published articles investigating the factors that explain stock return in the field of financial economics. More specifically, he shows the density of t-statistics of reported coefficients of published articles in this field, and he observes that the higher frequencies of reported coefficients occur when t-stat equals at least 2. This is the threshold that represents p-values typically below 10%, meaning that the coefficient is interpreted as significant. [Harvey \(2017\)](#) concludes that only a very small portion of published articles contains t-stats below 2, leading him to acknowledge the existence of publication bias.

Arguably, financial economics is a narrow field. However, it is a field that usually provides several batteries of robustness tests, which some are not significant. Thus, I suspect that if one makes the same analysis of [Harvey \(2017\)](#) in other fields, the bias might be even higher since other fields typically provide fewer robustness tests. My point here is that, in 2020, it seems undebatable that academic research has a severe publication bias, and lots of non-significant results never had the opportunity to come to the public. As [Aguinis, Cascio and Ramani \(2017\)](#) point out, if we want to restore credibility in academic research, this is a problem we need to tackle.

I believe tutorial-articles can play a role in tackling this problem. As I discuss later in this document, tutorial-articles are usually written by more experienced authors to help beginners in the application of a particular research technique. Thus, the margins for p-hacking and HARKing decrease as authors, editors, reviewers, and the public are better equipped with tools to replicate the results quickly and to assess if either a particular published research outcome is trustworthy or it was found by chance. Though no tutorial-article is *ex-ante* free of error, the fact that it passes through the typical reviewing process (where other researchers review the empirical part of the tutorial-article) decreases the likelihood that the user is making a mistake. Furthermore, after we build some years of experience and a tutorial-articles collection is in place with accumulated knowledge, tutorial-articles should become an excellent tool to discuss and improve anyone's research, which should lead to systematic fewer errors in academic research.

Moreover, despite the relevancy of management research to society, it seems that we face some public relations problem, in the sense that part of the society does

not recognize the real potential impact of our research, either because of lack of trust or because academic research is too opaque ([Aguinis, Banks, Rogelberg, & Cascio, 2020](#)). Though tutorial-articles are far from being a silver bullet to this problem, there are several reasons to believe that they help mitigate these issues, since tutorial-articles help those less familiar with the technical aspects of research, create a body of shareable knowledge that can be applied in several areas, and improve the communication and, therefore, the trust in a research's results.

FAIR principles

In the inner circles of open science, a dominant topic is the FAIR principles of data sharing ([Go Fair, 2016](#); [Wilkinson et al., 2016](#)). Despite its importance, a recent survey shows that more than 50% of respondents have never heard about the FAIR principles before ([Fane, Ayris, Hahnel, Hrynaskiewicz, Baynes, & Farrell, 2019](#)). This is not surprising since most journals still have flexible guidance to data sharing (some even do not have guidance at all). As a consequence, on the author-level, authors still struggle to share their data and usually do not have a personal protocol to operationalize it online.

The FAIR principles postulate that data must be Findable, Accessible, Interoperable, and Reusable. These principles are mainly focused on the machine side of the ecosystem in order to promote easy ways for users to collect research data ([Wilkinson et al., 2016](#)). Nevertheless, journals and editors must recognize these principles to facilitate as much as possible data sharing at the individual level. Authors also can benefit by understanding the underlying concepts of data sharing in order to demand from journals FAIR data. In that arena, journals seem to be more and more turning to public depositories such as Dataverse, Mendeley Data, Figshare, OSF, Zenodo, among others, to share published articles' data. For instance, the *American Journal of Political Science* is consolidating data in the Harvard Dataverse, while the *North American Journal of Economics and Finance* is using Mendeley data.

Thus, in my view, the bottom line on the FAIR principles to authors is: (a) authors should understand that they exist and what their purposes are, and (b) demand their operationalization from journals and data creators as a whole. Importantly, tutorial-articles meet FAIR principles since they provide open data and protocols in a public depository.

WHAT IS A TUTORIAL-ARTICLE?

In [Martins and Mendes-da-Silva \(2020\)](#), we mentioned that: a tutorial-article must explain straightforwardly and help researchers to execute a technique that is relevant to the field, in our case, the management field perceived broadly. In this context, simplicity and didactic matter, as beginners are a considerable part of readers. But focus, depth, and high-quality execution also matter, as the tutorial-article is intended to be used in future empirical studies, so the technique must be free of errors and easy to implement.

This is not a straightforward balance; that is why a tutorial-article usually discusses one topic only.

The list of potential topics to be discussed by tutorial-articles, nevertheless, is extensive. The topic must fill the needs of being relevant, insightful, and be of great help to other researchers. Below, I include a list of topics that, I believe, are worth exploring in future tutorial-articles. By no means is this list exhaustive, and there are several potential topics candidates for a tutorial-article not included in Table 1.

Table 1. Topics worth exploring in future tutorial-articles at RAC.

Topic	Comments
Statistical analysis (in a broad sense)	A tutorial-article that deals with statistical-related issues, such as those discussed in Harvey (2017) and Babin, Ortinau, Herrmann and Lopez (2020) , seems to be quite relevant to our community.
Factor analysis and principal component analyses, response item theory, Guttman's analyses, Cronbach's alpha, etc.	This is a topic highly relevant to research in management, and very often researchers have questions about rotation techniques, differences between an exploratory and confirmatory factor analysis, etc.
Replication studies	With the increased volume of open data around the world, data reuse becomes a topic of greater relevance. Thus, a tutorial-article discussing how to replicate a published study seems worth exploring.
Semantic analysis, sentiment analysis, etc.	A tutorial-article with applications, particularly but not exclusively, to marketing and organizational research about semantic, sentiment analyses, also seems to be a real contribution to this literature.
Social network analysis, graph theory, etc.	Tutorial-articles dealing with these topics in an environment of big data seem worth exploring.
Endogeneity issues, instrumental variables, causal models, matching techniques, regression discontinuity designs, synthetic control models, and randomized control trials	These topics usually demand high-skilled treatments and corrections and non-intuitive techniques to mitigate biases and find proper estimations. Thus, I believe that these topics represent a fruitful arena to tutorial-articles. Schiozer, Mourad and Martins (2021) is an unparalleled first step, but so much more is needed here.
Linear models, limited dependent variables, panel data techniques, etc.	These methods are used throughout several types of management articles. Econometric books are often too technical, making non-experts less likely to use them or more likely to make mistakes. Thus, tutorial-articles guiding the use of these methods to an audience of non-experts, given that they provide a real contribution over an econometric book, are always welcome.
Time-series modeling, time-series forecast, volatility modeling, AR and MA models, etc.	These topics are frequently used in financial markets, and tutorials dealing with the micro-decisions on these models seem beneficial to the community.
Machine learning, techniques to handle big data, etc.	Big data is already a current topic, so we all need to be able to manage it. Thus, tutorial-articles that handle the specificities of big data in management research are also welcome.
Simulation methods and optimization methods	While frequently used in operations research, these techniques are sometimes misinterpreted by non-experts. Thus, a tutorial-article showing the correct application in a management-related example seems beneficial to the community.
Data visualization	Data visualization is one of the primary issues in big data analytics. Thus, a tutorial-article exploring how to visualize the data frequently used in management research effectively would contribute to the community.
Accounting-related issues, earnings management, disclosure quality assessment, etc.	These are accounting-related issues that demand several micro-decisions, leading to significant variability and difficulties in estimations. Thus, tutorial-articles about these topics are quite interesting to those researchers working at the interface between management and accounting.

Note. This table contains a non-exhaustive list of potential topics that future tutorial-articles submitted to RAC may discuss. The table also includes some reasons why tutorial-articles on these topics will benefit the management research community.

A short description of a tutorial-article structure

As mentioned before, tutorial-articles are built upon simplicity and didacticism, but also depth and high-quality

execution. Thus, tutorial-articles require a structure that facilitates these characteristics. In my opinion, a tutorial-article requires the following sections. A well-designed motivation about the topic analyzed, focusing on the real implications it has to the community. A short but sound review of the method/technique, dealing with the technical aspects that users

need to understand. An example of a real research application, providing the protocol (i.e., the code) to estimate the method or to use the technique. Moreover, a section discussing the traps that users may face and providing guidance about how to use the code correctly is always welcome. At the code level, authors are invited to make as many comments as necessary to the correct usage of it and, when appropriate, a readme.txt file. In the operational parts of the code (i.e., where authors use software's packages or functions), the code needs to be as clean as possible: it should be consistent, educational, easy to follow, and avoid any unnecessary complexity.

EXAMPLES OF DATA USEFUL TO FUTURE TUTORIAL-ARTICLES

In this section, I show how to download data that, I believe, might be useful to management researchers and data that researchers can use in future tutorial-articles. I am more familiar with datasets used in quantitative research (more particularly those used in finance studies), but it is possible to find datasets with qualitative data, for instance in [SAGE Research Methods \(2020\)](#). The code used to download the

data and to generate the graphs in this section is available at RAC's Dataverse ([Harvard Dataverse, 2020](#)). Please, keep in mind that the purpose of this section is to illustrate the use of some data that is open; it is not a tutorial of how to collect nor how to visualize data.

Example 1: World Bank Data

Let's start with one of the leading datasets providers in the world: the World Bank. It publishes several inter-related datasets, which one can consult in [The World Bank \(2020\)](#). The list of datasets is enormous and contains country-level data for climate change, education, environment, gender, infrastructure, poverty, science and technology, social development, social protection and labor, urban development, and many more.

The R script shows how to download these datasets. To illustrate, I download and plot the data for the first of the 17 Sustainable Development Goals (the 2030 agenda of World Bank). Figure 1 shows how Brazil is dealing with poverty in the last few years. The trend is quite favorable; although we see some increase in poverty since 2014.

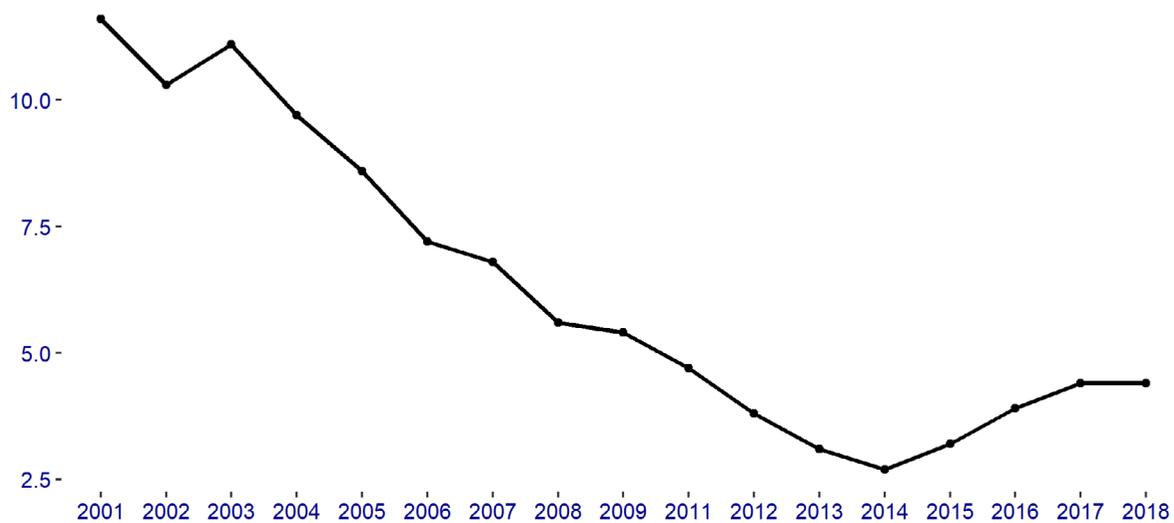


Figure 1. Evolution of the poverty headcount at \$1.90 a day (% of the population) in Brazil.

This figure plots the evolution of the poverty headcount at \$1.90 a day (% of the population) in Brazil from 2001 to 2018. The data is downloaded from [The World Bank \(2020\)](#) dataset (the variable indicator is SI.POV.DDAY). The evolution shows a consistent decrease in the population percentage marked as poor from 2003 to 2014 (the percentage decrease from roughly 12% to 2.5%). Since 2014, it shows a slight increase in the number of the population marked as poor (the percentage increased from roughly 2.5% to around 4.5%).

Example 2: Google Trends

For marketing research, one source of relevant information is Google Trends. For example, researchers

can use this type of data to improve the 'nowcasting' of the market's demand for goods or services ([Carrière-Swallow & Labbé, 2013](#)). To illustrate, I am searching for the words 'Administração,' 'Curso técnico,' 'Graduação,' 'MBA,'

'Mestrado,' and 'Doutorado' in Brazil for the period 2010-2020 (Fig. 2), and for the words 'Management,' 'MBA,' 'Master,' and 'PhD' in the United States (Fig. 3). Although a more profound discussion of the topic is out of the scope of this document, it seems that the demand for 'Administração' and 'MBA' is decreasing since 2010 in Brazil. On the other hand, 'Curso técnico' appears to be a more searched option for Brazilians in the last few years. The search for 'Management' and 'MBA' also seems to be decreasing since 2010. We observe similar trends in the searches for related terms in the United States, showing an apparent decrease in the public interest for these courses in this country. Also, notice the seasonality of these series, suggesting the data is fruitful to tutorial-articles dealing with this topic.

Example 3: Economic Policy Uncertainty datasets

Another example of a useful dataset can be found in [Baker, Bloom and Davis \(2016\)](#): the Economic Policy Uncertainty. The data is described at <https://www.policyuncertainty.com/> (retrieved in August

10th, 2020) and is available for several countries, including Brazil. In Figure 4, I show three graphs with data from this dataset: the Economic Policy Uncertainty index, the World Uncertainty index, and the Geopolitical Risk index. These indexes are measured using newspapers, counting words related to uncertainty and risk. Though a more detailed investigation is out of the scope of this document, it seems that COVID-19 rose the uncertainty across the globe since March 2020.

More useful datasets

Alongside all articles with open data published by RAC, which can be found at RAC's Dataverse, some other datasets are worth mentioning and could be used in future tutorial-articles. For instance, [Perlin, Kirch and Vancin \(2019\)](#) provide an open-source algorithm to download corporate data from B3, including financial data and corporate events. Along similar lines, the BETS (Brazilian economic time series) ([Costa, Speranza, & Costa, 2018](#)) R-package provides the time series of several economic and financial indexes of Brazil.

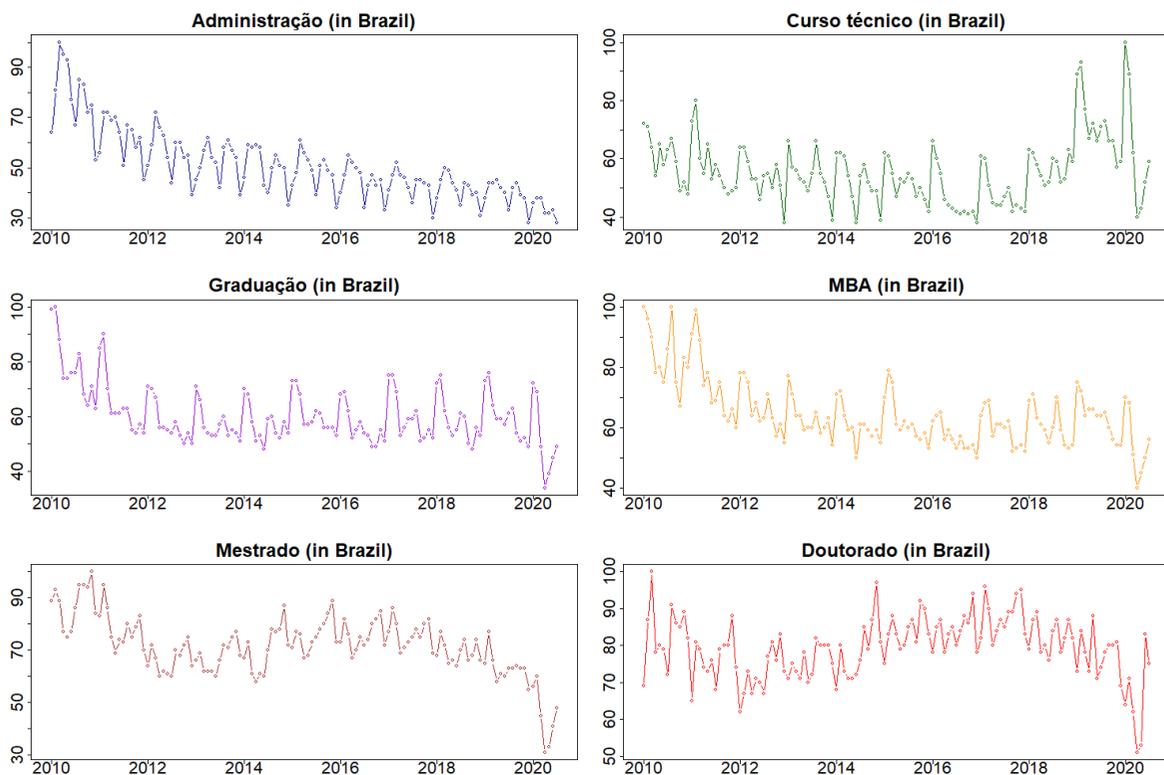


Figure 2. Evolution of Google searches in Brazil for six different terms related to management courses.

Source: elaborated by the author. This figure shows the evolution of Google searches in Brazil for six different terms related to management courses from 2010 until mid-2020. There are six plots, where the searched term is 'Administração,' 'Curso técnico,' 'Graduação,' 'MBA,' 'Mestrado,' and 'Doutorado,' respectively. The figure shows a decrease in searches for 'Administração,' 'Graduação,' 'MBA,' and 'Mestrado.' There is no apparent change in search for 'Doutorado.' There is a slight increase in the search for 'Curso técnico' in early 2020. All plots show considerable seasonality in searches.

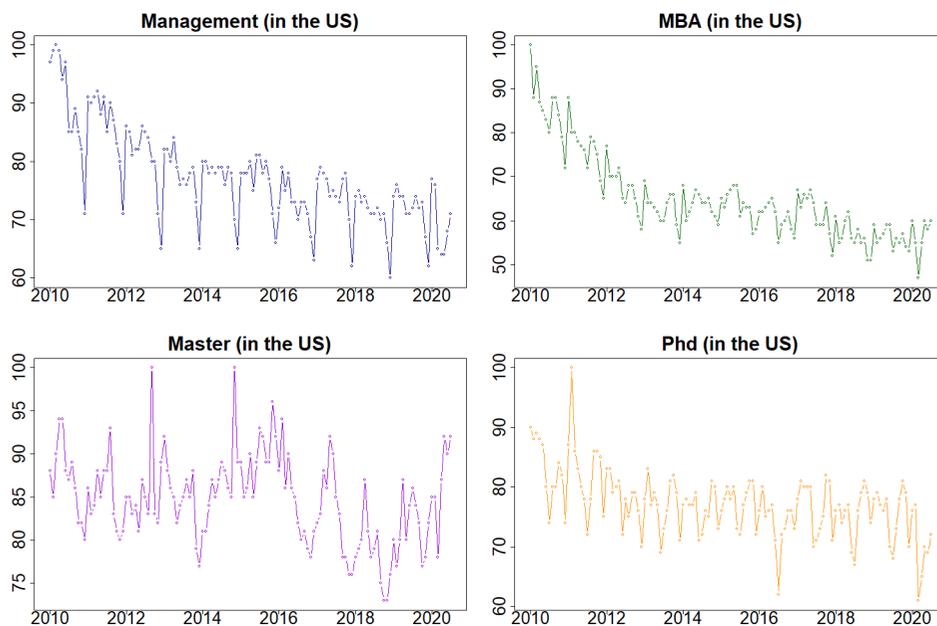


Figure 3. Evolution of Google searches in the U.S. for four different terms related to management courses.

Source: elaborated by the author. This figure shows the evolution of Google searches in the United States for four different terms related to management courses from 2010 until mid-2020. There are four plots, where the word searched is 'Management,' 'MBA,' 'Master,' and 'PhD,' respectively. The figure shows a decrease in searches of 'Management,' 'MBA,' and 'PhD.' There is no apparent change in the search for 'Master.' All plots show considerable seasonality in searches.

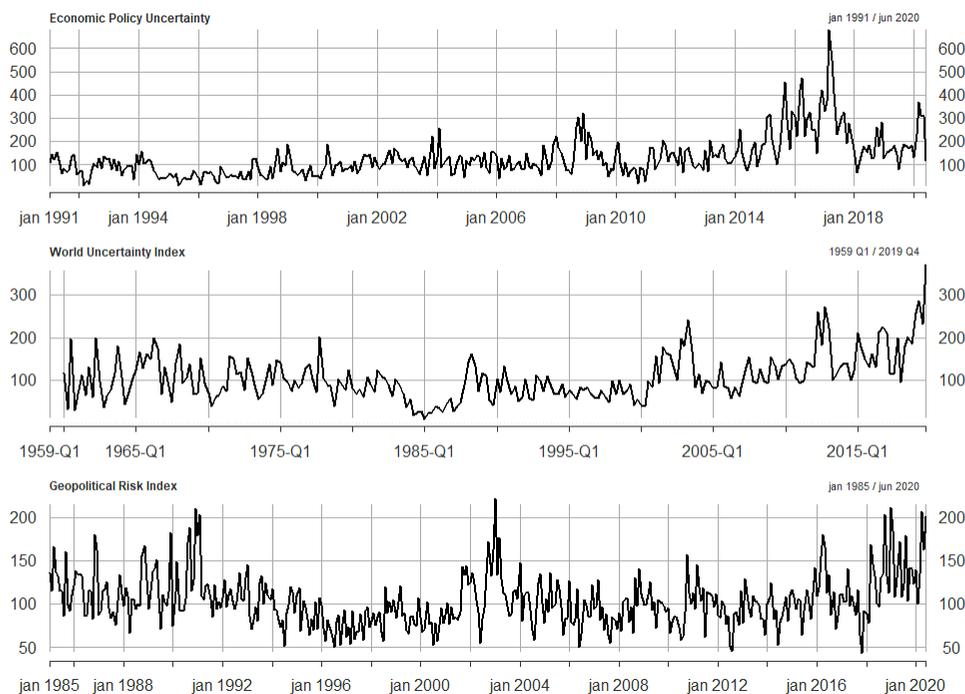


Figure 4. Evolution of economic policy uncertainty, world uncertainty, and geopolitical risk for Brazil.

Source: elaborated by the author based on indexes described in Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636. <https://doi.org/10.1093/qje/qjw024>; This figure shows the evolution of the indexes Economic Policy Uncertainty, World Uncertainty, and Geopolitical Risk for Brazil, each in a different plot. The periods are respectively from Jan./1991 to Jun./2020, first quarter-1959 to fourth-quarter 2019, and Jan./1985 to Jun./2020. The plots show significant variability in the three indexes during all periods. It shows a local increase in uncertainty in early 2020.

Moreover, there are already several open datasets published in the data journal named *Data-in-Brief* (<https://www.journals.elsevier.com/data-in-brief>; retrieved in August 10th, 2020). For instance, an article by [Rossoni and Gonçalves \(2019\)](#) contains a dataset with board social capital of Brazilian companies, while an article by [Paulo and Camões \(2019\)](#) includes data about ecological fiscal transfers in the state of Minas Gerais. *Data-in-Brief* is a multidisciplinary journal that publishes data articles. It contains all types of datasets from several fields. Thus, this is an interesting source of data that can be used in future tutorial-articles.

On top of data, some journals that keep accounts in public depositories have articles with open data. For instance, an article by [Campoli, Alves Júnior, Rossato and Rebelatto \(2020\)](#), published at the *Socio-Economic Planning Sciences*, contains data on social indicators of Brazilians' states from 2004 to 2014, while an article by [Barboza and Vasconcelos \(2018\)](#), published at the *North American Journal of Economics and Finance*, contains open data related to BNDES and investment in Brazil.

Finally, there are international projects that manage open datasets with all sorts of information. One example is *Our world in data* project, which contains data about global indicators, such as demographics, food and agriculture, education, environment, innovation, and technological change. The full list of Brazilian data is available at [Our World in Data \(2020\)](#). Another example is Kaggle.com ([Kaggle, 2020](#)), which contains several different types of datasets, including COVID tweets, flight delays, air pollution data, and several others.

MANUSCRIPTS PUBLISHED IN THIS SPECIAL EDITION

In this new section, we have five tutorial-articles accepted for publication. All tutorial-articles passed through the traditional double-blind review, where during the reviewing process, authors and reviewers do not know the identity of each other. By the time of publication, reviewers were invited to identify themselves. Those that accepted have their names included in the tutorial-article first page. This process is based on open science practices, which is the open peer-review process. Before anything, I would like to thank all the reviewers for their excellent work in providing detailed reports. Their job was not easy since we (myself included), as a community of management academics, did not have experience in evaluating tutorial-articles. I would like to thank the reviewers that accepted to identify themselves (in alphabetical order): Charles Kirschbaum, Christine da Silva Schröder, Cristiano Machado Costa, Eduardo Kayo, Júlio César Araújo da Silva Junior, Marcelo S. Perlin,

Pablo Cristini Guedes, Pedro Raffy Vartanian, Ricardo Ratner Rochman, but also to those reviewers that preferred not to disclose their identity.

The first tutorial-article is written by Rafael Felipe Schiozer, Frederico Abou Mourad, and Theo Cotrim Martins, and the title is *A tutorial on the use of differences-in-differences in management, finance, and accounting*. [Schiozer et al. \(2021\)](#) provide a discussion about some problems that researchers face when defining their empirical design and offer a solution using the differences-in-differences (DiD) method. They provide an R script, a Stata do-file, and a dataset. They also show a DiD application using the 2008 subprime crisis as an exogenous shock. Their contribution to the field is impressive, and their article has countless empirical applications to corporate finance scholars and scholars that wish to estimate shock-based causal models.

The second tutorial-article is written by Naielly Lopes Marques, Carlos de Lamare Bastian-Pinto, and Luiz Eduardo Teixeira Brandão, and the title is *A tutorial for modeling real options lattices from project cash flows*. [Marques, Bastian-Pinto and Brandão \(2021\)](#) show how to model real options lattices using a project's cash flow as dividends. They also provide a code where authors can modify values at their will in order to make their own estimations. This is an impressive contribution to the field, with real implications for stakeholders inside but also outside academia.

The third tutorial-article comes from Marcelo Perlin, Mauro Mastella, Daniel Vancin, and Henrique Ramos, and is titled *A GARCH tutorial in R*. [Perlin, Mastella, Vancin and Ramos \(2021\)](#) discuss and provide all steps necessary to make quality and correct estimations of GARCH models. In their code, users can download the required financial data from an open dataset to go through all estimates. Still, this data can be used in all sorts of studies using stock prices of Brazilian companies. Furthermore, they provide several additional protocols, including a fitness test to select the most appropriate GARCH model and to make sound simulations using financial data.

The fourth tutorial-article is written by Luciano Rossoni, Clayton Gonçalves, Mônica Silva, and Alex Gonçalves, and is titled *Mapping organizational culture schemas based on correlational class analysis: A tutorial*. [Rossoni, Gonçalves, Silva and Gonçalves \(2021\)](#) describe the use of correlational class analysis to map the culture schemes of organizations. Their tutorial is useful to separate organizations into cultural schematic groups and to understand the outcomes of culture as a whole. Their tutorial is highly relevant for scholars studying in the fields of organizational culture and organizational structures in a broad sense.

The fifth tutorial-article is written by Humberto Lopes and Marlusa Gosling and the title is *Cluster analysis in practice: Dealing with outliers in managerial research*. Lopes and Gosling (2021) discuss the use of cluster analysis in management research and present a code to understand and estimate alternative techniques for dealing with outliers. They offer a dataset to help users to execute the code and also provide an appendix, where they complement the main text with an example that expands the selected number of clusters. Their tutorial is an excellent contribution to researchers that need a tool to deal with outliers in empirical research.

FINAL WORDS

The purpose of this document is to launch the new section of RAC that will continuously publish tutorial-articles. I discuss some Open Science practices related to tutorial-articles and provide reasoning about why this specific type of article can help to mitigate some of the well-known problems in the management research field (but

other fields too). I also provide some topics that, through my lenses, are worth exploring in the future. But I should say that, by no means, the list of Table 1 is exhaustive, and there are certainly several topics not included in this list that could be the topic of a tutorial-article. Then, I provide some examples of open data sets, and an R script to download them, that could be used in future tutorial-articles.

Finally, I would like to emphasize that this document presents RAC's edition of quantitative tutorial-articles. At the current date, there is Lanka, Lanka, Rostron and Singh (2019) call for papers inviting the submission of qualitative tutorial-articles. Moreover, the publication of tutorial-articles is, from now on, an ongoing process in RAC. So I take these last few words to invite the submission of new tutorial-articles to be considered for publication in the next editions of RAC.

REFERENCES

- Aguinis, H., Banks, G. C., Rogelberg, S. G., & Cascio, W. F. (2020). Actionable recommendations for narrowing the science-practice gap in open science. *Organizational Behavior and Human Decision Processes*, 158, 27-35. <https://doi.org/10.1016/j.obhdp.2020.02.007>
- Aguinis, H., Cascio, W. F. & Ramani, R. S. (2017). Science's reproducibility and replicability crisis: International business is not immune. *Journal of International Business Studies*, 48, 653–663. <https://doi.org/10.1057/s41267-017-0081-0>
- Babin, B. J., Ortinau, D. J., Herrmann, J. L., & Lopez, C. (2020). Science is about corroborating empirical evidence, even in academic business research journals. *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2020.06.002>
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636. <https://doi.org/10.1093/qje/qjw024>
- Banks, G. C., Field, J. G., Oswald, F. L., O'Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 questions about open science practices. *Journal of Business and Psychology*, 34(3), 257-270. <https://doi.org/10.1007/s10869-018-9547-8>
- Barboza, R. de M., & Vasconcelos, G. F. R. (2019). Measuring the aggregate effects of the Brazilian Development Bank on investment. *The North American Journal of Economics and Finance*, 47, 223-236. <https://doi.org/10.1016/j.najef.2018.12.013>
- Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97, 71-87. <https://doi.org/10.1016/j.jmoneco.2018.05.01>
- Beugelsdijk, S., Van Witteloostuijn, A., & Meyer, K. E. (2020). A new approach to data access and research transparency (DART). *Journal of International Business Studies*, 51, 887–905. <https://doi.org/10.1057/s41267-020-00323-z>
- Bhimani, A., & Willcocks, L. (2014). Digitisation, 'big data' and the transformation of accounting information. *Accounting and Business Research*, 44(4), 469-490. <https://doi.org/10.1080/00014788.2014.910051>
- Campoli, J. S., Alves Júnior, P. N., Rossato, F. G. F. D. S., & Rebelatto, D. A. D. N. (2020). The efficiency of bolsa familia program to advance toward the millennium development goals (MDGs): A human development indicator to Brazil. *Socio-Economic Planning Sciences*, 71(C). <https://doi.org/10.1016/j.seps.2019.100748>

- Carrière-Swallow, Y., & Labbé, F. (2013). Nowcasting with google trends in an emerging market. *Journal of Forecasting*, 32(4), 289-298. <https://doi.org/10.1002/for.1252>
- Costa, P., Speranza, T., & Costa, J. (2018). *Brazilian Economic Time Series* (Package 'BETS'; Version 0.4.9) [Package]. Retrieved from <https://github.com/nmecsyst/BETS>
- Drachen, T. M., & Ellegaard, O., Larsen, A. V., & Dorch, S. B. F. (2016). Sharing data increases citations. *LIBER Quarterly*, 26(2), 67-82. <http://doi.org/10.18352/lq.10149>
- European Research Council. (2019). *Open research data and data management plans: Information for ERC grantees*. Retrieved from https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf
- Fane, B., Ayris, P., Hahnel, M., Hrynaszkiewicz, I., Baynes, G., & Farrell, E. (2019): The state of open data report 2019. *Digital Science*. <https://doi.org/10.6084/m9.figshare.9980783.v2>
- Go Fair. (2016). *FAIR principles*. Retrieved from <https://www.go-fair.org/fair-principles/>
- Harvard Dataverse. (2020). *Journal of Contemporary Administration dataverse*. Retrieved from <https://dataverse.harvard.edu/dataverse/rac>
- Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, 72(4), 1399-1440. <https://doi.org/10.1111/jofi.12530>
- Hax, A. C., & Majluf, N. S. (1982). Competitive cost dynamics: The experience curve. *INFORMS Journal on Applied Analytics*, 12(5), 50-61. <http://dx.doi.org/10.1287/inte.12.5.50>
- Kaggle (2020). *Datasets*. Retrieved from <https://www.kaggle.com/datasets>
- Kousta, S., Pastrana, E., & Swaminathan, S. (2019). Three approaches to support reproducible research. *Science Editor*, 42(3) 77-82. Retrieved from <https://www.csscienceeditor.org/article/three-approaches-to-support-reproducible-research/>
- Lanka, E., Lanka, S., Rostron, A., & Singh, P. (2019). Research methods in qualitative management research. *Zenodo*. <http://doi.org/10.5281/zenodo.3934265>
- Lopes, H. E. G., & Gosling, M. S. (2021). Cluster analysis in practice: Dealing with outliers in managerial research. *RAC - Revista de Administração Contemporânea*, 25(1), e200081. <https://doi.org/10.1590/1982-7849rac2021200081>
- Marques, N. L., Bastian-Pinto, C. L., & Brandão, L. E. T. (2021). A tutorial for modeling real options lattices from project cash flows. *RAC - Revista de Administração Contemporânea*, 25(1), e200093. <https://doi.org/10.1590/1982-7849rac2021200093>
- Martins, H. C. (2020). Checklist evaluation for tutorial-articles: Clean code. *Zenodo*. <http://doi.org/10.5281/zenodo.3785747>
- Martins, H. C., & Mendes-da-Silva, W. (2020). The new section of the Journal of Contemporary Administration: Tutorial articles. *RAC - Revista de Administração Contemporânea*, 24(3), 275-282. <https://doi.org/10.1590/1982-7849rac2020200044>
- McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., & Yarkoni, T. (2016). Point of view: How open science helps researchers succeed. *ELife*, 5, e16800. <https://doi.org/10.7554/eLife.16800>
- Mendes-Da-Silva, W., & Cerqueira Leal, C. (2020). Salami science in the age of open data: Déjà lu and accountability in management and business research. *RAC - Revista de Administração Contemporânea*, 25(1), e200194. <https://doi.org/10.1590/1982-7849rac2021200194>
- Mendes-Da-Silva, W. (2019). Have we been transparent enough? Challenges in replicability and credibility in business research. *RAC - Revista de Administração Contemporânea*, 23(5), 1-6. <https://doi.org/10.1590/1982-7849rac2019190306>
- Our World in Data (2020). Brazil. Retrieved from <https://ourworldindata.org/country/brazil>
- Pampel, H., & Dallmeier-Tiessen, S. (2014). Open research data: From vision to practice. In S. Bartling, S. Friesike (Eds.), *Openingscience* (pp.213-224). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-00026-8_14
- Paulo, F. L. L. de, & Camões, P. J. S. (2019). Dataset on ecological fiscal transfers and municipal protected areas in the state of Minas Gerais, Brazil. *Data In Brief*, 27, 104601. <https://doi.org/10.1016/j.dib.2019.104601>
- Perlin, M. S., Mastella, M., Vancin, D. F., & Ramos, H. P. (2021). A GARCH tutorial with R. *RAC - Revista de Administração Contemporânea*, 25(1), e200088. <https://doi.org/10.1590/1982-7849rac2021200088>
- Perlin, M. S., Kirch, G., & Vancin, D. (2019). Accessing financial reports and corporate events with GetDFPData. *Brazilian Review of Finance*, 17(3), 85-108. <http://dx.doi.org/10.12660/rbfin.v17n3.2019.78654>
- Piowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, e175. <https://doi.org/10.7717/peerj.175>
- Rossoni, L., & Gonçalves, A. F. (2019). Board social capital and structure, ownership and financial variables of Brazilian companies: A three levels dataset integrating directors, board networks and firm characteristics. *Data in Brief*, 26, 104502. <https://doi.org/10.1016/j.dib.2019.104502>
- Rossoni, L., Gonçalves, C. P., Silva, M. P. da., & Gonçalves, A. F. (2021). Mapping organizational culture schemas based on correlational class analysis: A tutorial. *RAC - Revista de Administração Contemporânea*, 25(1), e200096. <https://doi.org/10.1590/1982-7849rac2021200096>
- Rousi, A. M., & Laakso, M. (2020). Journal research data sharing policies: A study of highly-cited journals in neuroscience, physics, and operations research. *Scientometrics*, 124, 131-152. <https://doi.org/10.1007/s11192-020-03467-9>
- SAGE Research Methods. (2020). Datasets. Retrieved from <https://methods.sagepub.com/datasets>
- Schiozer, R. F., Mourad, F. A., & Martins, T. C. (2021). A tutorial on the use of differences-in-differences in management, finance, and accounting. *RAC - Revista de Administração Contemporânea*, 25(1), e200067. <https://doi.org/10.1590/1982-7849rac2021200067>

- Sheth, J., & Kellstadt, C. H. (2020). Next frontiers of research in data driven marketing: Will techniques keep up with data tsunami? *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2020.04.050>
- Spiegel, M. (2019). Replication? Do you even have access to the data? *Critical Finance Review*, 8(1-2), 11-13. <https://doi.org/10.1561/104.00000079>
- The World Bank (2020). *Indicators*. Retrieved from <https://data.worldbank.org/indicator>
- Vandewalle, P. (2012). Code sharing is associated with research impact in image processing. *Computing in Science and Engineering*, 14(4), 42–47. <https://doi.org/10.1109/MCSE.2012.63>

- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L. B. S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Growth, P., Goble, C., Grethe, J. S., Heringa, J., Hoehn, P. A. C., Hoof, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., Schaik, R. van, Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Lei, J. V. D, Mulligen, E. van, Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

Authorship

Henrique Castro Martins*

Pontifícia Universidade Católica do Rio de Janeiro, IAG - Escola de Negócios.
Rua Marquês de São Vicente, nº. 255, 22451-900, Rio de Janeiro, RJ, Brazil.

E-mail address: hcm@iag.puc-rio.br

 <https://orcid.org/0000-0002-3186-4245>

* Corresponding Author

Plagiarism Check

The RAC maintains the practice of submitting all documents approved for publication to the plagiarism check, using specific tools, e.g.: iThenticate.

Conflict of Interests

The authors have stated that there is no conflict of interest.

Copyrights

RAC owns the copyright to this content.

Data Availability

All data and materials were made publicly available through the Harvard Dataverse platform and can be accessed at:



Henrique Castro Martins, 2020, "Replication Data for: Tutorial-articles: the importance of data and code sharing", Harvard Dataverse, V1. <https://doi.org/10.7910/DVN/95LWQN>
