



Implementing Increased Transparency, and Reproducibility in Economics - The short version -

Lars Vilhuber
Cornell University

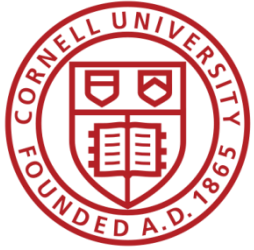
The opinions expressed in this talk are solely the authors, and do not represent the views of the U.S. Census Bureau, the American Economic Association, or any of the funding agencies.



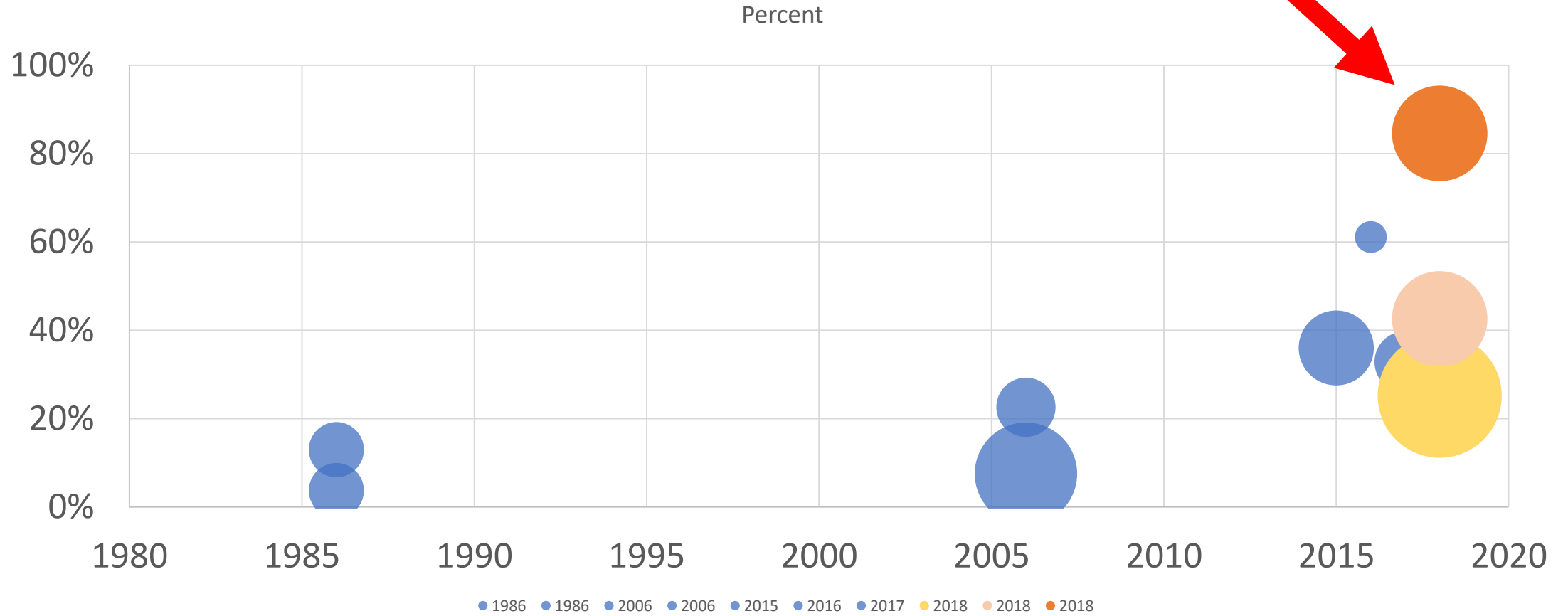
Progress

- Replication archives and Data (Code) Availability policies
 - Since 1985 in economics
 - More broadly since early 2000s





Results?





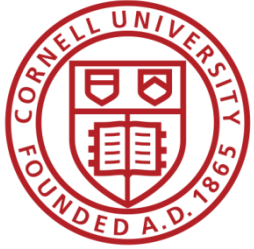
In a nutshell

- **40%** use restricted-access data
- **25%** use public-use data and are mostly or completely reproducible
- **25%** use public-use data and are only partially reproducible
- **10%** fail to yield useful results

It's only $\frac{1}{2}$ full!

Hey, it's not empty!





Second round (2012-)

- **Greater enforcement of data (and code) availability**
 - 2015, AJ Political Science
 - 2016, Data Editor for ASA Software Section
 - 2016, Statistical review added Science
 - 2017: AEA appoints Data Editor, with mandate to do similar activities (also EJ, Restud)



Action: Data citations and metadata

What is **FAIR**?

- **F**indable,
- **A**ccessible,
- **I**nteroperable, and
- **R**e-usable

A screenshot of the FORCE11 website. The logo at the top left consists of a blue circular icon with a white dot in the center, followed by the text "FORCE11" in a bold, sans-serif font. Below the logo is the tagline "The Future of Research Communications and e-Scholarship". A dark grey navigation bar contains three items: "ABOUT" with a downward arrow, "COMMUNITY" with a downward arrow, and "CODE OF CON". Below the navigation bar is a breadcrumb trail: "FORCE11 » Groups » The FAIR Data Principles". The main heading is "THE FAIR DATA PRINCIPLES" in bold. Below it is a sub-heading "JOIN IN THE DISCUSSION - LEA FAIR Data Principles". The word "Preamble" is displayed in a larger font, followed by the text "One of the grand challenges of data-intensiv".

FORCE11
The Future of Research Communications and e-Scholarship

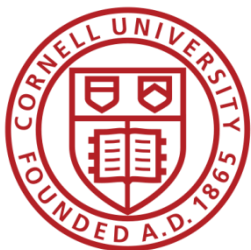
ABOUT ▾ COMMUNITY ▾ CODE OF CON

FORCE11 » Groups » The FAIR Data Principles

THE FAIR DATA PRINCIPLES

JOIN IN THE DISCUSSION - LEA
FAIR Data Principles

Preamble
One of the grand challenges of data-intensiv



The Future of Research Communications and e-Scholarship

Search

English

ABOUT ▾ COMMUNITY ▾ CODE OF CONDUCT GROUPS RESOURCES ▾ NEWS + BLOGS ▾ EVENTS ▾ PUBLIC

perceived criteria of importance.

1. Importance

Data should be considered legitimate, citable products of research. Data should be accorded the same importance in the scholarly record as citable research objects, such as publications[1].



2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and le attribution to all contributors to the data, recognizing that a single style or of attribution may not be applicable to all data[2].

3. Evidence

In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited[3].

4. Unique Identification

A data citation should include a persistent method for identification that is actionable, globally unique, and widely used by a community[4].

5. Access

Data citations should facilitate access to the data themselves and to such metadata, documentation, code, and other materials as are necessary for

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014
[<https://www.force11.org/group/joint-declaration-data-citation-principles-final>].



AMERICAN ECONOMIC ASSOCIATION

American Economic Review



The *American Economic Review* is a general-interest economics journal. Established in 1911, the AER is among the nation's oldest and most respected scholarly journals in economics.

American Economic Review: Insights



AER: Insights is designed to be a top-tier, general-interest economics journal publishing papers of the same quality and importance as those in the AER, but devoted to publishing papers with important insights that can be conveyed succinctly.

Journal of Economic Literature



The *Journal of Economic Literature* (JEL), first published in 1969, is designed to help economists keep abreast of and synthesize the vast flow of literature.

Journal of Economic Perspectives



The *Journal of Economic Perspectives* (JEP) fills the gap between the general interest press and academic economics journals.

American Economic Journal: Applied Economics



American Economic Journal: Applied Economics publishes papers covering a range of topics in applied economics, with a focus on empirical microeconomic issues.

American Economic Journal: Economic Policy



American Economic Journal: Economic Policy publishes papers covering a range of topics, the common theme being the role of economic policy in economic outcomes.

American Economic Journal: Macroeconomics



American Economic Journal: Macroeconomics focuses on studies of aggregate fluctuations and growth, and the role of policy in that context.

American Economic Journal: Microeconomics



American Economic Journal: Microeconomics publishes papers focusing on microeconomic theory; industrial organization; and the microeconomic aspects of international trade, political economy, and finance.



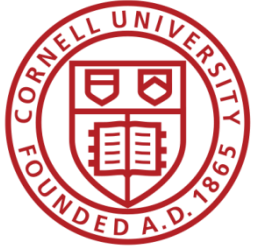
Current efforts at the AEA

- **Pre-emptively improve code archives**
 - By conducting reproducibility checks when we can
 - By working with groups that conduct reproducibility checks when we cannot
- **Better archives**
 - Greater transparency of the code and data archives
- **Better provenance tracking**
 - Leave code where it is when appropriate
 - Leave data where it is almost always
 - Display that information



AEA Pre-Publication Verification

- Every paper that receives a “conditional acceptance” is verified
 - *Data citations*
 - *Quality of README*
 - *Quality of code*
 - *Reproducibility of code*
 - *Quality of metadata in the repository*



Who is doing that?

- Earlier reproducibility work: **Flavio Stanchi** (now at AirBnb), **Sylverie Herbert** (on the market), **Hautahi Kingi** (Impaq)
- Current lead graduate students: **David Wasser** (until Dec 2019), **Meredith Welch** (since Jan 2020)
- Current and past undergraduate students: Alexia Ge, Anthony Peraza, Craig Schulman, Elijah B. Ruiz, Gabriel Bond, Jason S. Katz, **Jeong Hyun Lee**, **Jiayin Song**, John Park, **Joshua Passel**, Kirubeal T. Wondimu, Linchen Zhang, **Louis Liu**, **Luis Lopez Cabrera**, Luke O'Leary, **Mary-Jo Ajiduah**, **Naomi Li**, Nicholas Swan, Nishat Peuly, **Ryan Ali**, Samuel Frey, Siyang (Elaine) Yu, **Steve Yeh**, **Weilun Shi**, William Hernandez, **Yanyun (Iris) Chen**, Yuan-Hsuan (Sharon) Lin, **Zebang Xu**, Xing Su, Jiazhen Tan, Xueshi Su, Vendela Norman, Anderson Park, *Nehedin Juarez, Rubal Mistry, Syon Verma, William Silverman, Zechariah Karsana*
- Other graduate students: Aviv Caspi, Leah Kim

AEA Data and Code Guidance



**AMERICAN
ECONOMIC
ASSOCIATION**

Guidance for authors wishing to create data and code supplements, and for replicators.

Unofficial guidance on various topics by the AEA Data Editor

These web pages provide unofficial and developing guidance on the implementation of the American Economic Association (AEA)'s Data and Code Availability Policy. We also provide links to **generic guidance** being developed by a loose collective ("guild") of data editors and people in a similar role at various social science journals.

 Follow @aeadata

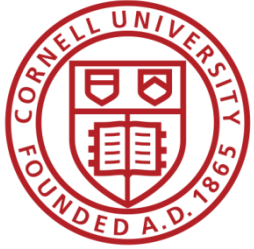
Order in which AEA authors should read these resources:

1. Start with the **official Data and Code Availability Policy**
2. Look for general guidance at the **Social Science Data Editors pages**
3. Read the **AEA's FAQ**
4. Look for any guidance specific to the AEA at the **Unofficial AEA Data and Code Guidance**
5. Last but not least, have a look at the **draft FAQ on this site**

Comments are welcome, please file them as **issues** in our Github repo.

Guidance on creating replicable data and program archives

How should researchers create replicable data and program archives? How




Current efforts at the AEA

- **Pre-emptively improve code archives**
 - By conducting reproducibility checks when we can
 - By working with groups that conduct reproducibility checks when we cannot
- **Better archives**
 - Greater transparency of the code and data archives
- **Better provenance tracking**
 - Leave code where it is when appropriate
 - Leave data where it is almost always
 - Display that information



Full-featured repository

OPEN ICPSR Find Data Share Data openICPSR Repositories ▾ GO Sign Up Sign In

 AMERICAN ECONOMIC ASSOCIATION [AEA Deposit Instructions](#) [Browse AEA Deposits](#) [Contact](#)

Depositing Data in the AEA Data and Code Repository

The *American Economic Association journals* require authors to deposit data and materials with a community-recognized or general repositories. The *AEA Data and Code Repository at ICPSR* serves that purpose. Please see the AEA's [Data and Code Availability Policy](#) and data citation guidance at the [Sample References](#) page for more details. **Authors are required to include a citation pointing to the deposit in the reference section of the final version of the article sent to the AEA.** The *openICPSR* repository automatically generates a citation when the data are "published."

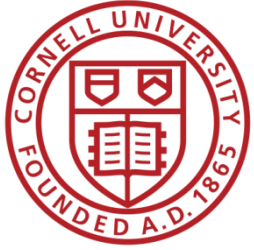
Deposits should include all data, annotated program code, command files, and documentation that is needed to replicate the findings from the authors' submitted article.

- **Data** should be comprehensively documented (see ICPSR's [Guide to Social Science Data Preparation and Archiving, 5th Edition](#) for guidance). The **author** is responsible for removing identifying information from the data to protect [confidentiality](#). Neither the AEA nor ICPSR review submissions for disclosure risk.
- **Program** code and command files should be annotated to facilitate replication and ensure clear correspondence between code and figures, tables, and analyses in the published article.
- Authors retain ownership and copyright to the data and code. Authors are required to affirm that they have the right to publish and redistribute the material. However,
 - ICPSR requires a license for distribution of data.
 - An **open license** is required by the AEA, in order to allow others to re-use the data and code, in particular for replication. Authors can select from several license options, including CC-BY 4.0 for data and Modified BSD for software and code. If an author would like to use multiple licenses or create a customized license, she should select the "Other" license option and upload a LICENSE file alongside the data and documentation.

By depositing in the AEA Data and Code Repository, the depositors allow the AEA staff to add keywords and other metadata which are important for proper indexing in linking. Any other changes are subject to the license chosen for the materials.

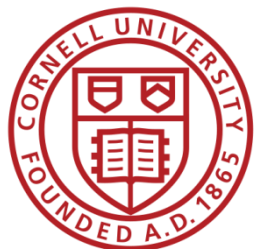
[View more extensive \(unofficial\) guidance.](#)

[Start Your Deposit](#)



Migrating archives

- Between Oct 11 and Oct 13, 2019, the staff at openICPSR ingested **2,552 historical AEA supplements**, increasing the size of the openICPSR repository **by a factor of 3**.

[Find Data](#) / [Imperial Russian Factory Database, 1894-1908](#)

Imperial Russian Factory Database, 1894-1908

Principal Investigator(s): Amanda Gregg, Middlebury College**Version:** V1AMERICAN
ECONOMIC
ASSOCIATION

| Name | File Type | | Last Modified |
|------------------------------------|---|--------|---------------------|
| 1894MicroData.xlsx | application/vnd.openxmlformats-officedocument.spreadsheetml.sheet | 4.5 MB | 08/08/2019 11:01:AM |

Project Citation:

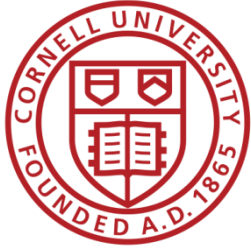
Gregg, Amanda. Imperial Russian Factory Database, 1894-1908. Nashville, TN: American Economic Association [publisher], 2020. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2020-01-29. <https://doi.org/10.3886/E110681V1>

| | | | |
|--|---------------------|---------|---------------------|
| AG_Corp_CleaningandDatabaseCompiler.do | text/x-stata-syntax | 23.4 KB | 08/08/2019 11:02:AM |
|--|---------------------|---------|---------------------|

Related Publications

The following publications are supplemented by the data in this project.

- Gregg, Amanda. "Factory Productivity and the Concession System of Incorporation in Late Imperial Russia, 1894-1908." *American Economic Review* 110, no. 2 (February 2020): 401-27. <https://doi.org/10.1257/aer.20151656>.



... and findability relies on metadata

Google

Imperial Russian Factory

Updated Date

Download Format

Usage Rights

Free

2 datasets found



Imperial Russian Factory
Database, 1894-1908

www.openicpsr.org
www.da-ra.de

Updated Jan 29, 2020



Data from: Антиосманские
выступления болгар и русско-
турецкие войны второй...

explore.openaire.eu

Updated 24 нояб. 2015 г.



Not seeing a result you expected?
[Learn](#) how you can add new
datasets to our index.



Imperial Russian Factory Database, 1894-1908

[Explore at openICPSR Self-Deposit Archive](#)

[Explore at www.da-ra.de](http://www.da-ra.de)

Unique identifier

<https://doi.org/10.3886/E110681V1>

Dataset updated Jan 29, 2020

Dataset provided by

[Middlebury College](#)

Authors

Amanda Gregg

License

[Attribution 4.0 \(CC BY 4.0\)](#)

License information was derived automatically

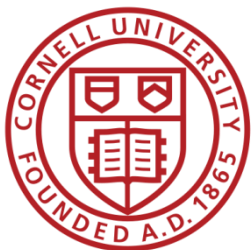
Area covered

European Russia (Russian Empire)



Current efforts at the AEA

- **Pre-emptively improve code archives**
 - By conducting reproducibility checks when we can
 - By working with groups that conduct reproducibility checks when we cannot
- **Better archives**
 - Greater transparency of the code and data archives
- **Better provenance tracking**
 - Leave code where it is when appropriate
 - Leave data where it is almost always
 - Display that information



The Future of Research Communications and e-Scholarship

Search

English

ABOUT ▾ COMMUNITY ▾ CODE OF CONDUCT GROUPS RESOURCES ▾ NEWS + BLOGS ▾ EVENTS ▾ PUBLIC

perceived criteria of importance.

1. Importance

Data should be considered legitimate, citable products of research. Data should be accorded the same importance in the scholarly record as citable research objects, such as publications[1].

DC¹

Data Citation Principles

2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and le attribution to all contributors to the data, recognizing that a single style or of attribution may not be applicable to all data[2].

3. Evidence

In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited[3].

4. Unique Identification

A data citation should include a persistent method for identification that is actionable, globally unique, and widely used by a community[4].

5. Access

Data citations should facilitate access to the data themselves and to such metadata, documentation, code, and other materials as are necessary for

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014
[<https://www.force11.org/group/joint-declaration-data-citation-principles-final>].



Every manuscript is checked

- What datasets are used
- Are they cited?
- Is there additional information access?
 - → URL leads to exact document
 - → URL leads to application
 - → other access procedures



Le labyrinthe des juges, par [Guillaume Beck](#)



Every manuscript is checked

- What datasets are used
- Are they cited?
- Is there additional information on access?
- Is there license/ data use information?
 - → Should the author provide the data?
 - → Is the author allowed to provide data?



Social science “guild”



Data and Code Guidance by Data Editors

Guidance for authors wishing to create data and code supplements, and for replicators.

Authors: Lars Vilhuber

This project is maintained by [social-science-data-editors](#)

Disclaimer

Unofficial guidance on various topics by Social Science Data Editors

Guidance on creating replicable data and program archives

This guidance is for the author wanting to create a replication archive.

See [Requested information](#) for the information the Data Editor may request from you, prior to the acceptance of your paper for publication.

Guidance on testing replicability of code

This guidance has two audiences:

- the author wanting to verify whether her code passes muster as a replicable archive
- the replicator wanting to verify the replicability of such an archive

See [Verification guidance](#)

FAQ

See our growing [FAQ](#). If you have questions or answers to add, please notify us by creating a [new issue](#).

[https://
social-science-
-data-editors.
github.io/
guidance/](https://social-science-data-editors.github.io/guidance/)

Thank you!

DOI: to come

Extra slides



Evolving Journal and Data Infrastructure

**Data Publishers struggle
with this!**



Example 1: OECD

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT **OECD.Stat**

1. Gross domestic product (GDP) ⁱ

[Customise](#) [Export](#) [Draw chart](#) [My Queries](#)

Please refer to the dataset [Gross domestic product \(GDP\), 2019 archive](#) to access longer time series based
Please note that OECD reference year from 2010 to 2015 changed on Tuesday 3rd of December, 2019.

[Transaction](#)

- B1_GA: Gross domestic product (output approach)
- B1G_P119: Gross value added at basic prices, excluding FISIM
- B1G: Gross value added at basic prices, total activity
 - B1G: Gross value added at basic prices, total activity
 - B1GVA: Agriculture, forestry and fishing (ISIC rev4)
 - B1GVB_E: Industry, including energy (ISIC rev4)
 - B1GVB_E: Industry, including energy (ISIC rev4)
 - B1GVC: of which energy (ISIC rev4)
 - B1GVF: Construction (ISIC rev4)
 - B1GVG_I: Distributive trade, repairs, transport, (ISIC rev4)
 - B1GVJ: Information and communication (ISIC rev4)
 - B1GVK: Financial and insurance activities (ISIC rev4)
 - B1GVL: Real estate activities (ISIC rev4)
 - B1GVM_N: Prof., scientific, techn., admin., sup (ISIC rev4)

URL does not always change!



Example 1: OECD

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT **OECD.Stat**

Short-Term Labour Market Statistics

Customise Export Draw chart My Queries

| | Subject | Measure | Frequency | Unit | | | |
|----------------|--|--------------------------------------|-----------|------------|---------|---------|---------|
| | | | | Q4-2016 | Q1-2017 | Q2-2017 | Q3-2017 |
| → Country | | | | | | | |
| Australia | Employment rate, Aged 15-64, All persons | Level, rate or quantity series, s.a. | Quarterly | 72.3 | 72.4 | 72.8 | |
| Austria | | | | 71.7 | 71.8 | 72.2 | |
| Belgium | | | | 63.3 | 62.4 | 62.9 | |
| Canada | | | | 72.9 | 73.3 | 73.4 | |
| Chile | | | | 62.4 | 62.4 | 62.6 | |
| Czech Republic | | | | 72.7 | 73.1 | 73.3 | |
| Denmark | | | | rev4) 72.0 | 72.0 | 72.0 | |

URL does not always change!
(but sometimes it does)



Example 1: OECD

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

OECD.Stat

Short-Term Labour Market Statistics : *Unemployed Population*

Customise Export Draw chart My Queries

| | Subject | Unemployed population, Aged 15 and over, All person | | | |
|----------------|-----------|---|---------|---------|---|
| | Measure | Level, rate or quantity series, s.a. | | | |
| | Frequency | Quarterly | | | |
| | Unit | | | | |
| | Time | Q2-2017 | Q3-2017 | Q4-2017 | Q |
| Country | | | | | |
| Australia | | 721 | 718 | 717 | |
| Austria | | 249 | 245 | 242 | |
| Belgium | | 363 | 355 | 319 | |
| Canada | | 1 275 | 1 215 | 1 185 | |
| Chile | | 595 | 581 | 609 | |
| Czech Republic | | 164 | 147 | 132 | |
| Denmark | | 174 | 177 | 160 | |
| Estonia | | 48 | 38 | 38 | |
| Finland | | 227 | 221 | 222 | |

URL does not always change!
(and then it doesn't...)



Example 2: Academic data publisher



[Home](#)

[Methodology](#)

[Media](#)

[Research & Applications](#)

[About Us](#)

EPU Indices

[All Country-Level Data](#)

| | |
|--|-----------------------------|
| Global | USA |
| Australia | Brazil |
| Canada | Chile |
| China | Colombia |
| Croatia New | France |
| Germany | Greece |
| Hong Kong | India |
| Ireland | Italy |
| Japan | South Korea |

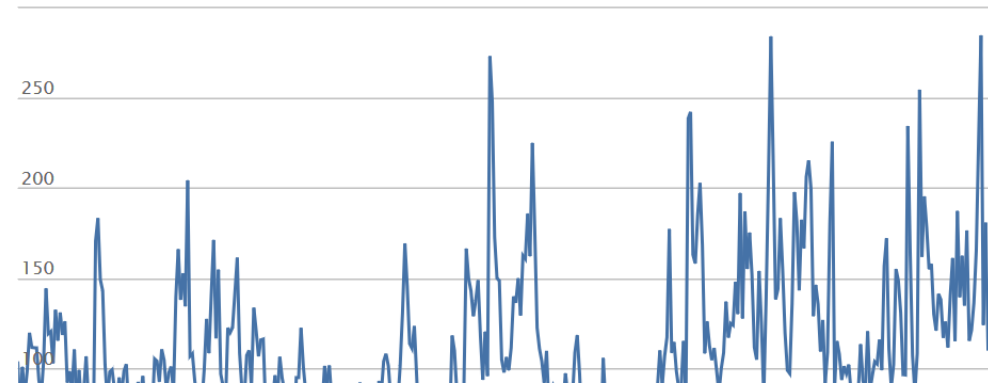
Economic Policy Uncertainty Index

We develop indices of economic policy uncertainty for countries around the world.

Monthly US Economic Policy Uncertainty Index



Zoom [1m](#) [3m](#) [6m](#) [1y](#) [7y](#) [All](#)





Example 2: Academic data publisher



https://www.policyuncertainty.com/index.html

Go

SEP DEC JAN

103 captures

18 Aug 2012 - 14 Dec 2019



2018 2019 2020



ECONOMIC POLICY UNCERTAINTY

Home

Methodology

Media

Research & Applications

About Us

EPU Indices

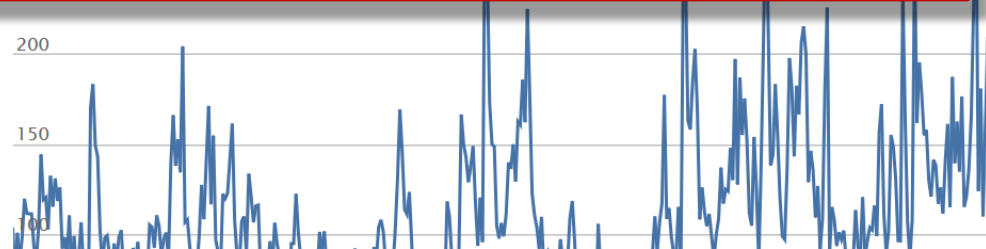
All Country-Level Data

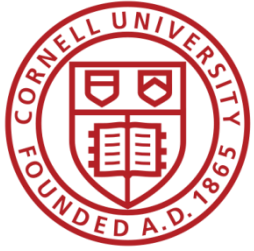
- [Global](#)
- [Australia](#)
- [Canada](#)
- [China](#)
- [Croatia](#)
- [France](#)
- [Germany](#)
- [Greece](#)
- [Hong Kong](#)
- [India](#)
- [Ireland](#)
- [Italy](#)
- [Japan](#)
- [South Korea](#)

Economic Policy Uncertainty Index


We develop indices of economic policy uncertainty for countries around the world.

© 2012-2018 by Economic Policy Uncertainty





Example 2: Academic data publisher-new!

 **ECONOMIC POLICY UNCERTAINTY**

[Home](#) [Methodology](#) [Media](#) [Research & Applications](#) [About Us](#)

EPU Indices


[All Country-Level Data](#)

[Global](#) [USA](#)

Economic Policy Uncertainty Index

We develop indices of economic policy uncertainty for countries around the world.

[Monthly US Economic Policy Uncertainty Index](#)



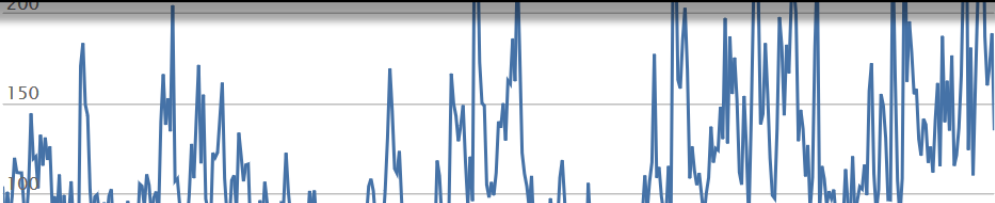
This work is licensed under a [Creative Commons Attribution 4.0 International License](#)

[Germany](#) [Greece](#)

[Hong Kong](#) [India](#)

[Ireland](#) [Italy](#)

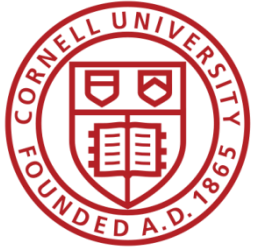
[Japan](#) [South Korea](#)



The chart displays the Economic Policy Uncertainty Index for the USA from approximately 2008 to 2023. The y-axis ranges from 100 to 200. The index shows significant volatility, with major peaks around 2009, 2014, 2016, and 2020, and a notable dip around 2012. The index generally fluctuates between 100 and 150, with the most recent period showing a sharp increase towards 200.

Response

Thanks, but I'll stick with what I've been doing for at least 20 years. At some point I might figure out the right license, but it's been working so far. And your inference is correct, the authors can use the data but not redistribute it. In this specific case, there is no reason for them to do so because the data are freely available to everyone.



Example 3: FRED (St. Louis Fed)

A screenshot of the FRED (Federal Reserve Economic Data) website. The header includes the FRED logo, "ECONOMIC RESEARCH FEDERAL RESERVE BANK OF ST. LOUIS", and a search bar. Below the header is a navigation menu with links for "FRED® Economic Data", "Information Services", "Publications", "Working Papers", "Economists", and "About". A "MY ACCOUNT" link is in the top right. Below the navigation is a breadcrumb trail: "Categories > Prices > Consumer Price Indexes (CPI and PCE)".

FRED ECONOMIC DATA | ST. LOUIS FED | ECONOMIC RESEARCH FEDERAL RESERVE BANK OF ST. LOUIS

MY ACCOUNT

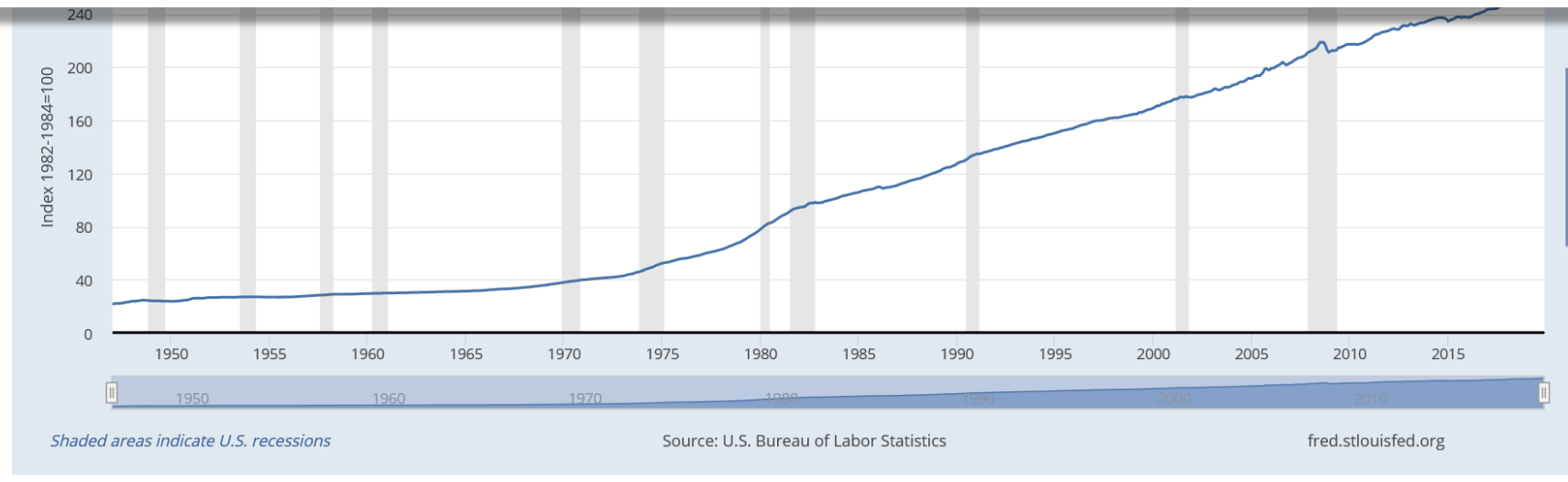
Search FRED

FRED® Economic Data Information Services Publications Working Papers Economists About St. Louis Fed Home

Categories > Prices > Consumer Price Indexes (CPI and PCE)

Suggested Citation:

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/CPIAUCSL>, February 3, 2020.





Example 4: German Restricted-access



RESEARCH DATA CENTRE (FDZ)
of the German Federal Employment Agency (BA)
at the Institute for Employment Research (IAB)

[Home](#) | [Newsletter](#) | [Jobs](#) | [Contact](#) | [Data Privacy](#) | [Imprint](#)



| Data Version | DOI (Link to Description of Data Version) | Availability (yyyy-mm-dd) |
|------------------------------|---|---------------------------|
| BHP 7518 v1 (current) | 10.5164/IAB.BHP7518.de.en.v1 | 2020-01-13 |
| BHP 7517 v1 | 10.5164/IAB.BHP7517.de.en.v1 | 2018-12-12 |
| BHP 7516 v1 | 10.5164/IAB.BHP7516.de.en.v1 | 2018-04-11 |

External data

[Data Archive](#)

[Data Access](#)

[Campus Files](#)

[Publications](#)

[Events](#)

[Projects of FDZ users](#)

[FDZ Projects](#)

[Complaint point of the RatSWD](#)

[Figures of the FDZ](#)

employees, both in total and broken down by gender, age, occupational status, qualification and nationality. Means and medians of wages for full-time employees are given, too. Additional datasets providing information about (gross) worker flows and about foundations and closures of establishments are available on request.

Data Versions

Old versions are only available for replication studies and only in justified exceptional cases for new Projects.

| Data Version | DOI (Link to Description of Data Version) | Availability (yyyy-mm-dd) |
|------------------------------|---|---------------------------|
| BHP 7518 v1 (current) | 10.5164/IAB.BHP7518.de.en.v1 | 2020-01-13 |



Example 4: German Restricted-access



RESEARCH DATA CENTRE (FDZ)
of the German Federal Employment Agency (BA)
at the Institute for Employment Research (IAB)

[Home](#) | [Newsletter](#) | [Jobs](#) | [Contact](#) | [Data Privacy](#) | [Imprint](#)



| Data Version | DOI (Link to Description of Data Version) | Availability (yyyy-mm-dd) |
|------------------------------|---|---------------------------|
| BHP 7518 v1 (current) | 10.5164/IAB.BHP7518.de.en.v1 | 2020-01-13 |
| BHP 7517 v1 | 10.5164/IAB.BHP7517.de.en.v1 | 2018-12-12 |
| BHP 7516 v1 | 10.5164/IAB.BHP7516.de.en.v1 | 2018-04-11 |

External data

[Data Archive](#)

[Data Access](#)

[Campus Files](#)

[Publications](#)

[Events](#)

[Projects of FDZ users](#)

[FDZ Projects](#)

[Complaint point of the RatSWD](#)

[Figures of the FDZ](#)

employees, both in total and broken down by gender, age, occupational status, qualification and nationality. Means and medians of wages for full-time employees are given, too. Additional datasets providing information about (gross) worker flows and about foundations and closures of establishments are available on request.

Data Versions

Old versions are only available for replication studies and only in justified exceptional cases for new Projects.

| Data Version | DOI (Link to Description of Data Version) | Availability (yyyy-mm-dd) |
|------------------------------|---|---------------------------|
| BHP 7518 v1 (current) | 10.5164/IAB.BHP7518.de.en.v1 | 2020-01-13 |



Example 4: German Restricted-access

Establishment History Panel (BHP) – Version 7518 v1

DOI: 10.5164/IAB.BHP7518.de.en.v1

Summary

Data source:

Data Access



The IAB Establishment Panel is available via the following ways of access:

- On-site use at the FDZ. Further information on Applying for [on-site use](#).
- Remote data Access. Further information on Applying for [remote data access](#).

nationality. Means and medians of wages for full-time employees are given, too. Additional datasets providing information about (gross) worker flows and about foundations and closures of establishments are available on request.

Dataset Descriptions and Frequencies

German

- DOI: [10.5164/IAB.FDZD.2001.de.v1](#)
-  [FDZ-Datenreport 01/2020](#)
-  [Fallzahlen und Labels](#)

English

- DOI: [10.5164/IAB.FDZD.2001.en.v1](#)