

# BEYOND THE NUMBERS

## THE ECONOMIC DATA ECOSYSTEM

FEDERAL RESERVE BANK OF ST. LOUIS

November 7-9, 2018

### Wednesday, November 7, 2018

12:00 p.m. – 4:30 p.m.	<p><b>Pre-Conference Workshop:</b> <b>Creating Data Literacy Lesson Plans</b> Diego Mendez-Carbajo, Charissa Jefferson, Katrina Stierholz Mary Suiter</p> <p>(River Room) <i>Registration for this workshop is full.</i></p>
5:00 p.m. – 7:00 p.m.	<p><b>Opening Reception</b></p> <p>Welcome David Wheelock, Group Vice President and Deputy Director of Research Federal Reserve Bank of St. Louis</p> <p><b>Fed-ucation: The Economics of Trade</b> B. Ravikumar, Senior Vice President and Deputy Director of Research, Federal Reserve Bank of St. Louis</p> <p>(Gateway Auditorium)</p>

### Exhibitors and Data Checkups

During Wednesday's reception and throughout the conference, please visit the information provider exhibit tables. On Thursday and Friday, providers will also offer "Data Checkups" to answer specific questions or give hands-on demonstrations. Each breakout session time slot will have three 15-minute checkup sessions. Each session is first-come, first-served and can accommodate up to 3 people; sign up via the links below.

Thursday, November 8, 10 a.m.-10:45 a.m.	<a href="#">Census data in FRED</a>
Thursday, November 8, 11 a.m.-11:45 a.m.	<a href="#">OECD</a>
Thursday, November 8, 11:45 a.m.-12:30 p.m.	<a href="#">FRED</a>
Thursday, November 8, 1:45 p.m.-2:30 p.m.	<a href="#">Transparency.Treasury.Gov Data</a>
Thursday, November 8, 2:30 p.m.- 3:15 p.m.	<a href="#">World Bank</a>
Thursday, November 8, 3:30 p.m.-4:15 p.m.	<a href="#">FRED</a>
Friday, November 9, 10 a.m.-10:45 a.m.	<a href="#">Census</a>
Friday, November 9, 10:45 a.m.-11:30 a.m.	<a href="#">St. Louis Fed Economic Education</a>

Thursday, November 8, 2018

8:00 am – 8:45 am	Continental Breakfast		
8:45 am – 9:45 am	<p><b>Keynote Address: Census: A One Stop Shop for All Your Data Needs</b> Stephanie Lee Studds, Division Chief, Economic Indicators Division, U.S. Census Bureau (Gateway Auditorium)</p>		
9:45 am – 10:00 am	15 min Break		
10:00 am – 10:45 am Breakout Sessions A	<p><b>How Does the World Bank Tell Stories with Data?</b> Malarvizhi Veerappan (Session A1 – Gateway Auditorium)</p>	<p><b>Developing Data Literacy Lessons</b> Diego Mendez-Carbajo, Charissa Jefferson, and Katrina Stierholz (Session A2 – Tennessee Room)</p>	
10:45 am – 11:00 am	15 min Break		
11:00 am – 11:45 am Breakout Sessions B	<p><b>Not a Mystery Inside an Enigma: An Overview of Russian Government Data</b> Natasha Arguello and Ekaterina Zatsepina (Session B1 – Gateway Auditorium)</p>	<p><b>The Philly Manufacturing Business Outlook Survey Turns 50! Does Everything Improve With Age?</b> Mike Trebing (Session B2 – St. Louis Room)</p>	
11:45 am – 12:30 pm Breakout Sessions C	<p><b>Working with International Financial Data Sets: Potential Issues and Solutions</b> Todd Hines and Shikha Sharma (Session C1 – Gateway Auditorium)</p>	<p><b>G.17 and H.6 Statistical Releases at the Federal Reserve Board of Governors</b> Kimberly Bayard, Brian Preslopsky, and Heidi Markovitz (Session C2 – St. Louis Room)</p>	
12:30 pm – 1:45 pm	Birds of a Feather Lunch		
1:45 pm – 3:15 pm Breakout Workshops*	<p><b>Working with Data in Stata</b> James Ng (Workshop A – Missouri Room)</p> <p><i>Attendees are encouraged to bring a laptop with Stata installed.*</i></p>	<p><b>What Can We Do About Preserving Business Sources</b> Jennifer Boettcher (Workshop B – Gateway Auditorium)</p>	<p><b>R for Economics</b> Christine Murray (Workshop C – Tennessee Room)</p> <p><i>Attendees are encouraged to bring a laptop with R and RStudio installed.*</i></p>
3:15 pm – 3:30 pm	15 min Break		
3:30 pm – 4:15 pm Breakout Sessions D	<p><b>Evidence-Based Research with the Census Bureau Data Linkage Infrastructure</b> Kate McNamara (Session D1 – Gateway Auditorium)</p>	<p><b>Why This Restriction? Data, Compliance, and the Law</b> Brett Carrier (Session D2a – Tennessee Room)</p> <p><b>Federal Financial Data Solutions for Librarians, Researchers and Students</b> Alicia Shroff (Session D2b – Tennessee Room)</p>	
4:15 pm – 5:15 pm	<p><b>Plenary Session</b> From “Skip the Numbers” to “Great Stuff”: A Data Education Project Kristin Fontichiaro and Wendy Stephens (Gateway Auditorium)</p>		

Friday, November 9, 2018

8:00 am – 8:45 am	Continental Breakfast	
8:45 am – 9:45 am	<p align="center"><b>Plenary Session</b>  All About You, Up For Sale: How Data Brokers Like Cambridge Analytica  Construct Consumer Identities  Wendy Stephens  (Gateway Auditorium)</p>	
10:00 am – 10:45 am Breakout Sessions E	<p align="center"><b>IP for Government Data: Understanding Open Access vs. Public Domain</b>  Jennifer Boettcher  (Session E1 – Gateway Auditorium)</p>	<p align="center"><b>FREDcast: The Survey of Unprofessional Forecasters</b>  Michael T. Owyang  (Session E2 – St. Louis Room)</p>
10:45 am – 11:30 am Breakout Sessions F	<p align="center"><b>Preservation of Born-Digital Government Information Concerning Economics</b>  PEGI Project Representatives Marie  Concannon, Lynda Kellam, Katrina Stierholz  (Session F1 – Gateway Auditorium)</p>	<p align="center"><b>Patron Driven Data Acquisitions: Prizes Perils and Pitfalls</b>  Jim Church  (Session F2 – St. Louis Room)</p>
11:30 a.m. – 1:00 pm	<p align="center"><b>Lunch &amp; Data Producers Q&amp;A Panel</b>  Representatives from Census, the Federal Reserve Board and Banks, the BLS, the World Bank,  the OECD, and FRED  (River Room)  <b>Conference Close</b></p>	

Blue sessions highlight government information; green sessions highlight Federal Reserve information; and orange sessions highlight international information.

\*Attendees without laptops are welcome to attend these workshops. Please contact the conference planners via email (or in person during the conference) if you have any questions. Attendees may also choose to visit the Inside the Economy Museum on the first floor of the Bank at this time.

# Abstracts

## **Pre-Conference Workshop: Creating Data Literacy Lesson Plans** (Wednesday, noon)

Presenter(s): Diego Mendez-Carbajo, Professor of Economics, Illinois Wesleyan University, Charissa Jefferson, Business and Data Librarian, California State University—Northridge, Katrina Stierholz, VP & Director of Library and Information Services, Federal Reserve Bank of St. Louis, Mary Suiter, AVP & Director of Economic Education, Federal Reserve Bank of St. Louis

Abstract: Do your patrons struggle to understand economic data? Are you looking for resources to make your instruction more effective? Attend this half-day workshop designed for librarians who are interested in helping the profession move forward with data literacy. You will learn how to create instructional lessons using the ACRL Framework for Information Literacy Threshold Concepts focused on data topics. Lessons created using this approach are designed to engage students in active learning using the ACRL information literacy frames and actual economic data. *Registration for this workshop is full. For more information, visit the workshop site [here](#).*

## **BREAKOUT SESSIONS**

### **Presentation: Developing Data Literacy Lessons** (Thursday, 10 a.m., Session A2)

Presenter(s): Diego Mendez-Carbajo, Professor of Economics, Illinois Wesleyan University, Charissa Jefferson, Business and Data Librarian, California State University—Northridge, Katrina Stierholz, VP & Director of Library and Information Services, Federal Reserve Bank of St. Louis

Abstract: In order to address a demonstrated need for data literacy instruction, the presenters are undertaking a lesson-writing project targeted at librarians. Our goal is to help the library community create a set of lessons that develop students' data literacy skills grounded in the ACRL Framework for Information Literacy Threshold Concepts and use the FRED (Federal Reserve Economic Data) portal. We propose to introduce the topic of mapping the ACRL Frames to data-specific skills. For example, the frame of "searching as strategic exploration" can be articulated into concrete knowledge practices by using data categories, filtering tags, or sources to facilitate data discovery. The final product will be a set of edited, peer-reviewed lessons on data literacy. These lessons will be added to the open access St. Louis Fed's award-winning Economic Education portal, to Project Cora, and to other librarian-focused education repositories. This presentation will provide the rationale for the proposed work, a framework and plan for executing the development of lessons. Participants will complete a brief hands-on activity from a draft lesson.

### **Presentation: Evidence-Based Research with the Census Bureau Data Linkage Infrastructure** (Thursday, 3:30 p.m, Session D1)

Presenter(s): Kate McNamara, US Census Bureau

Abstract: This presentation discusses the use of administrative data (demographic, socioeconomic, businesses) in the experience of the U.S. Census Bureau and touches on a vision for the future. Until recently, administrative records have been incorporated and available for researchers' use in a limited way via the Census Bureau's research data centers (RDCs). Now the Census Bureau is expanding its Data Linkage Infrastructure to improve access for evaluators, program and policy analysts, and researchers. The new initiative improves the process for accessing these data, increases the discoverability of data with inventories and metadata, and explores new technology for data linkage. The data linkage infrastructure includes economic and demographic data such as Longitudinal Employer-Household Dynamics (LEHD) data, HUD experimental study data, Medicaid enrollment, SNAP administrative records, and more that can be linked to Census Bureau survey data, such as the American Community Survey. Curators will benefit from hearing about our infrastructure and data discoverability, and researchers will learn how to access our data for new projects. Examples will be provided to highlight challenges and successes.

### **Presentation: Federal Financial Data Solutions for Librarians, Researchers and Students – Alicia Shroff** (Thursday, 3:30 p.m., Session D2b)

Presenter(s): Alicia Shroff, Federal Reserve Bank of St. Louis Treasury Agency Outreach

This presentation will provide an overview of the federal financial data solutions available to librarians, researchers, and students interested in government, policy, and fiscal accounting. The U.S. Department of the Treasury, Bureau of the Fiscal Service designed and administers the [transparency.treasury.gov](https://www.transparency.treasury.gov) site to provide usable, easy-to-understand information about the billions of financial transactions processed by Treasury each year. This site was established as part of the federal government's initiative to create more transparency and openness in the federal government. Transparency fosters accountability, promotes effective and efficient operations, encourages citizen participation, and helps to increase trust and confidence in the federal government.

**Presentation: FREDcast: The Survey of Unprofessional Forecasters** (Friday, 10 a.m., Session E2)

Presenter(s): Michael T. Owyang, Federal Reserve Bank of St. Louis

**Presentation: G.17 and H.6 Statistical Releases at the Federal Reserve Board of Governors** (Thursday, 11:45 a.m., Session C2)

Presenter(s): Kimberly Bayard, Brian Preslopsky, and Heidi Markovitz, Board of Governors of the Federal Reserve System

**Presentation: How Does the World Bank Tell Stories with Data?** (Thursday, 10 a.m., Session A1)

Presenter(s): Malarvizhi Veerappan, World Bank Group

**Presentation: IP for Government Data: Understanding Open Access vs. Public Domain** (Friday, 10 a.m., Session E1)

Presenter(s): Jennifer C. Boettcher, Georgetown University

Abstract: Public domain and open data policies and how they are made. Current status of open data policies in the Federal government are changing with new laws. What is HR4174/S4047 and what does it say and mean? What are possible changes related to HR4174/S4047 and other trends in government data policies regarding access to that statistical data? Attendees will come away with a better understand of federal policies and laws regarding data.

**Presentation: Not a Mystery inside an Enigma: An Overview of Russian Government Data** (Thursday, 11 a.m., Session B1)

Presenter(s): Natasha Arguello, University of Texas at San Antonio, and Ekaterina Zatsepina, McKinsey & Company

Abstract: This session will provide an overview of the official statistical information collected in the Russian Federation, with a special focus on economic data. Today, Russia dominates the news cycle, but many misconceptions persist. In less than 30 years, Russia has undergone fundamental changes, transitioning from a Soviet empire to what Senator John McCain called “a gas station masquerading as a country” to a market economy and omnipresent geopolitical force. Building on the legacy of centralized planning, Russia collects detailed statistical information on economic, trade, demographic, social, environmental, and other topics. The Russian government adopted IMF’s Special Data Dissemination Standard (SDDS) to provide economic and financial data to the public. Efforts have been undertaken to make official statistics more transparent and accessible, but the vast majority of data still exists only in Russian. The presenters will discuss the Russian- and English-language versions of the Federal Service for State Statistics (Rosstat) and Russian State Statistics Service websites and provide tips for navigating and obtaining data.

**Presentation: Patron Driven Data Acquisitions: Prizes Perils and Pitfalls** (Friday, 10:45 a.m., Session F2)

Presenter(s): Jim Church, University of California Berkeley

Abstract: The presentation will chronicle the successes, failures, and lessons learned from a patron driven data acquisitions program piloted by the University of California Berkeley Library in collaboration with the University D-Lab. The pilot was funded from January 2015 to 2018 resulting in over \$150,000 worth of patron initiated data purchases using a combination of selector and central library collections funding. Applications were received from the departments of Economics, Agricultural and Resource Economics, Political Science, Chinese Studies, Linguistics, Biosciences, and Public Health. Over 70% of requests were from economics-related disciplines and the majority were successfully funded. The results of the pilot were mixed. Users were generally satisfied with the outcome, as the program enabled them to acquire data sources unavailable to them previously. Yet numerous pitfalls and difficulties were encountered along the way, including complex and time-consuming contract and license negotiations, oversights about data confidentiality, and unexpected storage and computing constraints. Of particular concern were vendor demands that data be needlessly limited to specific categories of users or hosted in restricted environments. Patrons also expressed frustration about lengthy time frames which were hampered by difficulties coordinating complex workflows between multiple campus and library units. The future of the program is under consideration due workload, technology, and communications issues as well campus budget constraints. Economics data service providers interested in embarking in similar programs will come away with first-hand experience, cautions, and advice about undertaking similar initiatives.

**Presentation: The Philly Manufacturing Business Outlook Survey Turns 50! Does Everything Improve With Age?** (Thursday, 11 a.m., Session B2)

Presenter(s): Mike Trebing, Federal Reserve Bank of Philadelphia

Abstract: Every month since May 1968, the Federal Reserve Bank of Philadelphia has been publishing the *Manufacturing Business Outlook Survey* (MBOS). Economists, the media, and investors carefully watch the regional survey because it is widely believed to have an influential impact on views about the national economy, and, consequently, has often moved financial markets. My presentation reviews the factors that have contributed to the success of the MBOS, including, timeliness, longevity, a consistent and simple set of questions over time, and a carefully selected sample of reporters.

Over the survey's 50-year history, technology has not only revolutionized the process of producing goods but the collecting and processing of survey data has also changed, so it makes for an interesting "behind the numbers" case study. Further, the history and insights gained by the MBOS provide a useful perspective on the challenges of survey collection in this era where credibility (of the survey taker), privacy, and data security all interplay. I will review some of the other Philly Fed data products that have endured the test of time, namely, the Survey of Professional Forecasters, the Livingston Survey, and our coincident indicators for the 50 states.

**Presentation: Preservation of Born-Digital Government Information Concerning Economics** (Friday, 10:45 a.m., Session F1)

Presenter(s): Marie Concannon, University of Missouri; Lynda M. Kellam, University of North Carolina at Greensboro; Katrina Stierholz, Federal Reserve Bank of St. Louis

Abstract: The PEGI project (Preservation of Electronic Government Information, [www.pegiproject.org](http://www.pegiproject.org)) is a grant-funded initiative to document challenges and discover new efforts for ensuring future access to digital government information. As avid users of federal and state data, economic researchers' input is crucial to the work. In this session we will review factors that have put economic information at risk for loss, and tell about projects aimed at reversing the trend. We will encourage participants to share data access experiences (for better or worse) and tell us what should be prioritized for preservation. PEGI will draw on thoughts collected from this session and others like it to prepare its culminating product: an interdisciplinary white paper describing a national agenda for collaborative action.

**Presentation: Why This Restriction? Data, Compliance, and the Law – Brett Currier** (Thursday, 3:30 p.m., Session D2a)

Presenter(s): Brett Currier, Federal Reserve Bank of Kansas City

**Presentation: Working with International Financial Data Sets: Potential Issues and Solutions** (Thursday, 11:45 a.m., Session C1)

Presenter(s): Todd Hines, Stanford University GSB Library and Shikha Sharma, MIT Dewey Library

Abstract: Finance researchers will often need to merge data from two or more data sets in order to perform statistical analysis. Historically, such projects have primarily used data sets covering companies in the United States. With an increasing number of projects involving international companies, linking data is becoming more complex. Moreover, very little documentation of recommended best practices exists in this area. Some of the topics covered are expected to be: recommended international company identifiers, the difficulties of linking international company daily stock prices to financial statement data, discussion of how far back data are available and the strengths and weaknesses of the most commonly available databases for this type of research. We also plan to illustrate many of these issues by walking through an international-company-linking example.

## WORKSHOPS

**Workshop: R for Economics** (Thursday, 1:45 p.m., Workshop C)

Presenter(s): Christine Murray, Bates College

Abstract: R is software for statistical analysis, data manipulation, and data visualization. Because it is open source, and because of the many user-contributed packages that extend its functions, R is being increasingly adopted by researchers in many fields, including economics. Those who learn R now will be ahead of the curve as use spreads. Undoubtedly of use to researchers, R can also be an effective teaching tool in economics, combined with the RStudio integrated development environment and proper instructional support. This hands-on workshop will focus on three of R's strengths in data access, reproducibility, and visualization; participants will create a reusable script to create publication-ready charts with economic data pulled from the web. Participants will be able to navigate the RStudio interface, install packages related to economics, access economic data from FRED and other sources, use R documentation to understand R code, and create charts. Further, participants will be able to discuss the relative merits of R for research and teaching compared to other common data analysis tools such as Stata or Excel, will be familiar with how R is currently used in economics, and will leave the workshop with a plan for further learning. No knowledge of statistical methods or programming is expected, but familiarity with economic data will enhance learning.

**Workshop: What Can We Do About Preserving Business Sources** (Thursday, 1:45 p.m., Workshop B)

Presenter(s): Jennifer C. Boettcher, Georgetown University

Abstract: Who knew space would so valuable that we are throwing away our business heritage? How do we as business and economic librarians have a responsibility to make sure future researchers can access these seemingly ephemeral published sources. We will be the responsible and explore the possibilities. This workshop will be dominated by discussions in small groups about formats, services, and innovations. We will build on the improv rule of "yes, and." Then

report back and come up with an agreement on principles. I can give you the structure and topics, if you want to explore this workshop.

**Workshop: Working with Data in Stata** (Thursday, 1:45 p.m., Workshop A)

Presenter(s): James Ng, Center for Digital Scholarship, Hesburgh Libraries, University of Notre Dame

Abstract: Stata is widely used in economics and other social sciences. It is commercial and thus not free but with many of the hallmarks of open source software. It is code-driven and aides reproducible research. In this workshop, participants will learn how to work with and analyze data using Stata. Under the guidance of the instructor, participants will perform basic exploratory data analysis in Stata. Participants will get some practice writing Stata code, an indispensable part of reproducible research. Goal: By the end of this workshop, participants will gain a working knowledge of how Stata is used by researchers and be able to perform basic tasks in Stata. They will become familiar with the strengths and limitations of Stata. Prerequisites: An interest to learn how to use Stata. This is not a workshop on statistical methods, so no knowledge of statistics is required. Programming experience is helpful but not required. Tech requirements: A laptop with Stata. Please [contact the conference committee](#) if you would like to participate but do not already have access to Stata.

## PLENARIES

**Plenary Session: All About You, Up For Sale: How Data Brokers Like Cambridge Analytica Construct Consumer Identities** (Thursday, 4:15 p.m.)

Presenter(s): Wendy Stephens, Assistant Professor of Library Media, Jacksonville State University

Abstract: When revelations about political operatives leveraging profiles and connections scraped from Facebook grabbed headlines this spring, the use of personal and demographic information has long determined everything from the mail order catalogs homes receive to the prices listed in those catalogs. Beyond targeting advertising, collocation of these types of information points can have long-term ramifications for our health, wealth and security. Track the evolution of increasingly sophisticated prediction models, consider how the data we share determines how we are marketed to, and increase your personal agency when it comes to data sharing.

**Plenary Session: From “Skip the Numbers” to “Great Stuff”: A Data Education Project** (Friday, 8:45 a.m.)

Presenter(s): Kristin Fontichiaro, Clinical Associate Professor, University of Michigan School of Information and Wendy Stephens, Assistant Professor of Library Media, Jacksonville State University

Abstract: “I just tell them to skip the numbers.” We’ve heard high school librarians dismiss data this way when they discussed scaffolding support for student research projects among themselves. Yet teens can gain significant agency in critically “reading” and “writing” with numerical analysis. Our belief in the importance of that data and its implications for our citizenry led us on a multi-year journey, funded by the Institute for Museum and Library Services, offering model lessons for how we might parse out and integrate statistics, visualization, data and privacy, Big Data, and prosocial data projects like citizen science. In this talk, we’ll share what we learned throughout the process of creating two collected volumes, three virtual conferences, and 27 webinars, identifying a range of opportunities for librarians and educators to inculcate practical data literacy skills, providing tools and strategies designed to work within the existing high school curriculum. <http://dataliteracy.si.umich.edu>