Statipedia: A wiki platform for collaboration across agencies

Joint Statistical Meetings
Miami, July 31, 2011

Peter B. Meyer
Office of Productivity and Technology, BLS

James A. Buszuwski
Office of Compensation and Working Conditions, BLS

Michael Messner
Office of Water, EPA

...with lots of advice from others

...not representing the agency, just the named authors

www.bls.gov
Vision: share methodology information online

Statipedia is a wiki for statistical staff across agencies

- An online workspace
- With methodology source material
- Pilot project
- Not for the public (outside .gov)
- Not for sensitive content (e.g. early data or PII)

- Modeled on Wikipedia and on Intellipedia at intel agencies
Knowledge pooling

We can

- Share definitions
- Copy one another’s designs, procedures, software
- Develop materials together easily
- Enable new staff to figure out more

→ Development of science and technology goes faster with easily shared knowledge

Platform is cheap and robust to budget uncertainty
Can help us get things done cheaply
Looks like Wikipedia.
User can edit a page directly.
A user can see the history of changes to a wiki page.
We can define common terms

Industry value added

The phrase *industry value added* means the *value added output* by an industry in a certain time period, usually a year. "Industry value added" is synonymous with *industry value added*: "gross product by industry", "gross product originating by industry", and "gross product originating".[1]

An industry’s value added output is the value of its *Gross output* minus the value of its *Intermediate inputs*. Gross output is the sum of sales or receipts and other operating income, commodity taxes paid, and not inventory growth. The value of intermediate inputs is the value of goods and services purchased from other industries or imported and used up or consumed.[2]

Industry value added is estimated for most industries from expenditure data not industry income, because expenditures data is thought to be more accurately estimated.[3]

Future improvement

* follow up the sources to the source cited here

References

1. ↑ "Box: Gross Product Originating: Definition and Relationship to Gross Domestic Product" from BEA web site
2. ↑ "Box: Gross Product Originating: Definition and Relationship to Gross Domestic Product" from BEA web site
3. ↑ "Box: Gross Product Originating: Definition and Relationship to Gross Domestic Product" from BEA web site

Categories: BEA | NIPA
We can describe and search classification systems.
Can share
source code
Can add value to methodology documents

- searching
- links and footnotes that can be clicked for more info
- equations can be copied
Hyperlinks can go to definitions, summaries of sources, and new bibliographies.
Can find pages by “category”
“BLS” is a category, with subcategories

BLS

Category: BLS

This category is for pages related to the Bureau of Labor Statistics.

Subcategories

This category has the following 2 subcategories, out of 2 total.

B
  ▪ Statisticians at BLS

C
  ▪ CPS

Pages in category "BLS"

The following 59 pages are in this category, out of 59 total.

7
  ▪ 790 program

A
  ▪ ATUS
  ▪ American Time Use Survey
  ▪ At the agencies: Availability of Journal Archives

B
  ▪ Template: At the agencies
  ▪ Federal Economic Statistics Advisory Committee
  ▪ Office of Productivity and Technology
  ▪ BLS Handbook of Methods
  ▪ BLS Handbook of Methods chapter 15 on International Price Indexes
  ▪ At the agencies: Bayesian Statistics
  ▪ Benefit Incidence
  ▪ At the agencies: Browsers
  ▪ Bureau of Labor Statistics

C
  ▪ C&S

C cont.
  ▪ Census of Fatal Occupational Injuries (CFOI)
  ▪ At the agencies: Cloud Computing
  ▪ Collective Bargaining Agreements (statistical)
  ▪ Commodities and Services Survey
  ▪ Compensation and Working Conditions
  ▪ Confidential Information Protection and Statistical Act (CIPSEA)
  ▪ At the agencies: Confidentiality and Data Dissemination
  ▪ Current Employment Statistics

D
  ▪ Diffusion index
  ▪ Discouraged workers
  ▪ Displaced workers
  ▪ Duration of unemployment

E
  ▪ At the agencies: Editing Wikipedia
  ▪ At the agencies: Email address structure
  ▪ Employed persons
  ▪ Employment Situation
We can develop joint information on the international data institutions, methods, and findings.
Can share info on seminars in DC area

| Date         | Time   | Location          | Who                       | Topic and notes
|--------------|--------|-------------------|---------------------------|------------------|
| Mar 9        | 4:00-5:45 | Georgetown/DC Area Economometrics Workshops & ICC 550 | Konrad Menzel (EU) | FDI and Wages: Evidence from a Linked Employer-Employee Analysis
| Mar 8        | 10:30 AM | NIST campus, Gaithersburg | Wanjie Chen | NIST Fundamentals of Uncertainty Analysis Short Course
| Mar 23       | 10:30 AM | BLS economics seminars | Bruce Fallick (Federal Reserve) | Nonemployment Duration and the Consequences of Job Separations
| Mar 1        | 1:30 PM  | Georgetown/DC Area Economometrics Workshops & ICC 550 | Christian Hansen (Chicago Booth) | --
| Feb 23       | 10:30 AM | BLS economics seminars | John Earle (GMU) | --
| Feb 22       | 4:00-5:00 | Georgetown/DC Area Economometrics Workshops | Ivan Fernandez-Val (EU) | --
### Conference submission due dates

<table>
<thead>
<tr>
<th>Conference</th>
<th>When</th>
<th>Where</th>
<th>Dates submissions due</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011 Joint Statistical Meetings</td>
<td>July 30 - August 4, 2011</td>
<td>Miami Beach Convention Center</td>
<td>2011-02-01</td>
</tr>
<tr>
<td>FedCASIC Workshop</td>
<td>22-24 March 2011</td>
<td>Wash DC, at BLS</td>
<td>2011-02-01 (?) 50 word abstract Dan Gillman</td>
</tr>
<tr>
<td>Federal Forecasters Conference</td>
<td>21 April 2011</td>
<td>Wash DC, at BLS</td>
<td>2011-02-01</td>
</tr>
<tr>
<td>International Symposium on Forecasting</td>
<td>June 26-29, 2011</td>
<td>Prague, The University of Economics</td>
<td>2011-03-04</td>
</tr>
<tr>
<td>Wikimania 2011</td>
<td>Aug 4-7 2011 (plus meetings Aug 2-4)</td>
<td>Haifa, Israel</td>
<td>2011-04-30 Call for participation</td>
</tr>
<tr>
<td>ICES IV</td>
<td>11-14, June 2012</td>
<td>Montreal, Canada</td>
<td>2011-03-01 for invited sessions; 2011-08-31 for topic-contributed paper, poster session submissions</td>
</tr>
<tr>
<td>SHOT (Society for the History of Technology) 2011</td>
<td>3-6 Nov 2011</td>
<td>Cleveland</td>
<td>2011-spring (not set at this writing)</td>
</tr>
<tr>
<td>2012 Society of Labor Economists</td>
<td>4-5 May 2012</td>
<td>Hyatt Regency Chicago</td>
<td>2011-fall</td>
</tr>
<tr>
<td>ASSA Annual Meeting 2012 (economists)</td>
<td>6-8 Jan 2012</td>
<td>Chicago</td>
<td>2011-01-31? Not stated yet on website</td>
</tr>
<tr>
<td>Predictive Analytics World including workshops</td>
<td>2011-03-13 (March 13-19 2011)</td>
<td>San Francisco</td>
<td>Didn’t see how to submit an abstract</td>
</tr>
</tbody>
</table>
Can compare procedures across agencies

This is one of many “At the agencies” pages
Can share simple training materials in the wiki

Tiny tutorial for Matlab

From a Unix shell one can just type `matlab` as a command on any computer that has it, and start to type interactive statements such as those below. One could also put them in a file with the `.m` extension to run them from within `matlab` with `run file.m` or from the shell with `matlab < file.m` This tutorial covers very little but you can see something of the language.

% The percent sign begins comments.
% The statements below can be typed interactively one per line to get
% clear responses from Matlab. There's no need to type the comment part at the
% end of the lines. Make sure to use upper and lower case in the
% same way as in the statements shown.

A=[1 2;3 4]  % defines matrix A as a 2x2 with first line [1 2]
B=A'        % transpose
B=A+A       % sum, element by element
Ainv=inv(A) % takes inverse of a matrix
A*Ainv      % calculates and prints the result of a matrix multiplication
B=[A;A]     % stacked so B has twice as many rows as A
B=[A A]     % the A's are side by side. B has twice as many columns as A.
B=A(1,1)    % B is a scalar now, the upper left element of A
B=A'*A      % matrix multiplication
B=A(:,1)    % B is set to first row of A
B=A.*A      % element by element multiplication
B=B./A      % element by element division
A=zeros(3,3) % special definition of a matrix of zeros
B=ones(3,1) % defines a matrix of ones
A=eye(5)    % defines identity matrix
B=A(1:2,1:3) % takes part of matrix
more on      % may not be needed; prevents help screen from scrolling off
help *       % shows sample of the help available

Categories: Software | Tutorials
And detailed training materials
Recent changes to the wiki
New issues

→ Quick technical discussion
Discussion of critical perspectives
User and administrative norms

- **Scale up:** Benefits rise as communities of interest emerge
- **Meet open technical standards**
  - HTML, TEX, wiki-text links, hooks/extensions
  - Co-evolution: Today’s community, technology, platform ➔ Tomorrow’s community, technology, platform
- **Serve and empower staff; Invite voluntary participation**
- **Encourage users to:**
  - Anchor discussions to sources, evidence, theory
    - Enable drilling down to sources
  - Address broad/open audiences
    - not organization-specific
Anticipated long run effects on knowledge management

More efficient scientific communities
- Shared source material, more reference points
- Mutual awareness and peer review
- Build community skills with open-source tools

Reduce duplication of effort
Specialists serve a larger audience
- Users can find them on this platform

Ease training and turnover

Less email? ➔ “discoverability” not “dissemination”

TOGETHER, the agencies have vast, diverse expertise and capability, economies of scale and scope, knowledge of data, and great computer resources
Conclusions and future steps

Statipedia is a wiki for statistical staff across agencies

- With methodology source material
- Not for the public – but could include non-.gov
- Not for sensitive content
- Content is growing

- In future
  - Can adopt and develop more tools and content
  - Share source code online better (not just on wiki)
  - Make it quickly and easily available
  - to new employees; faster logon
Contact Information

Meyer.Peter@bls.gov
202-691-5678

Buszuwski.James@bls.gov
202-691-6238

Messner.Michael@epa.gov