

Esri User Conference, July 12-15, San Diego California

Harvard Data Visualization Project

Spatio-Temporal Visualization of Global
Phenomena: 1850 to the Present

Harvard Business School

Geoffrey Jones

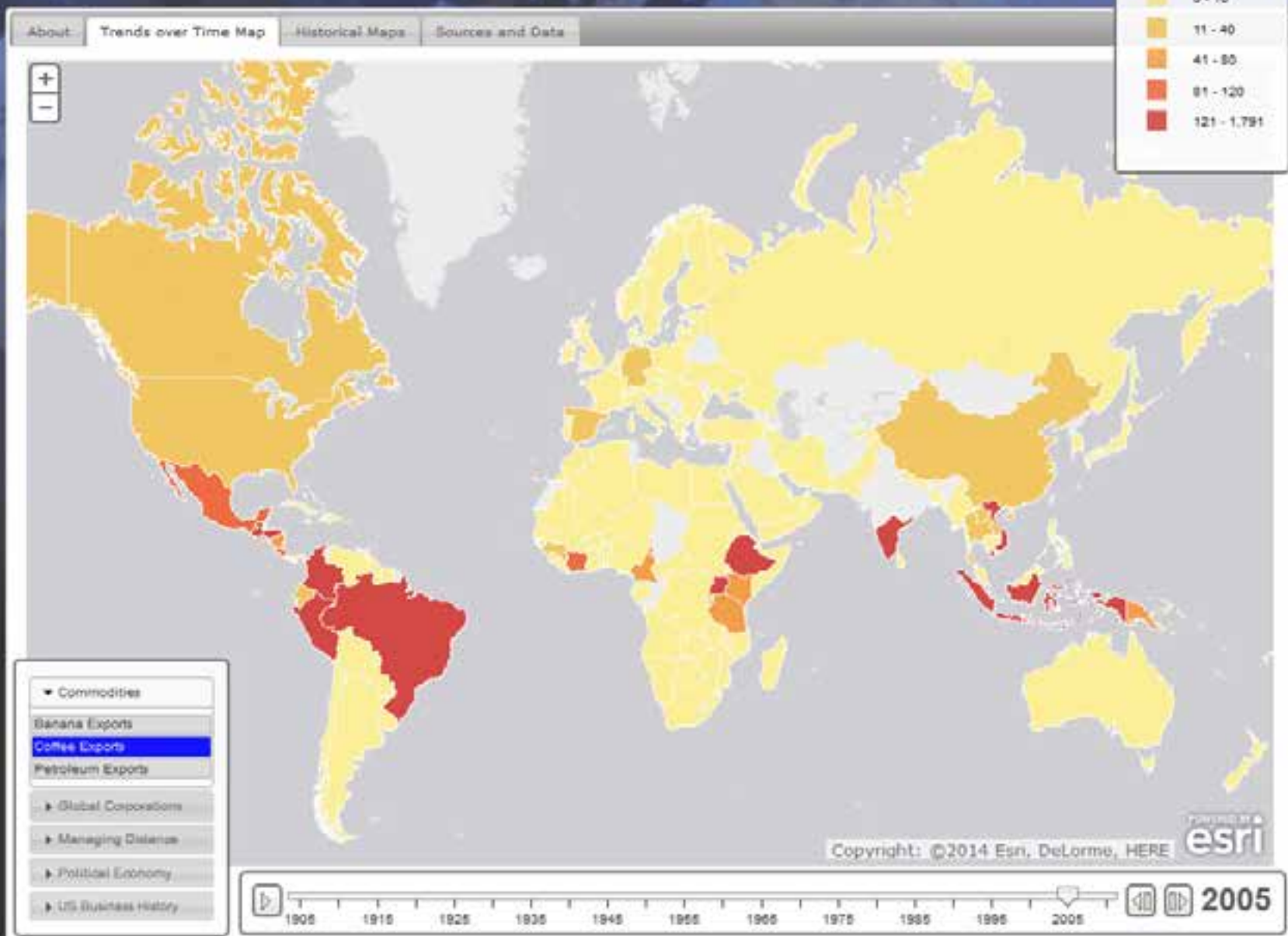
Walter Friedman, Patrick Clapp, Jeffrey Cronin, HBS Educational Technology Services

Harvard Center for Geographic Analysis

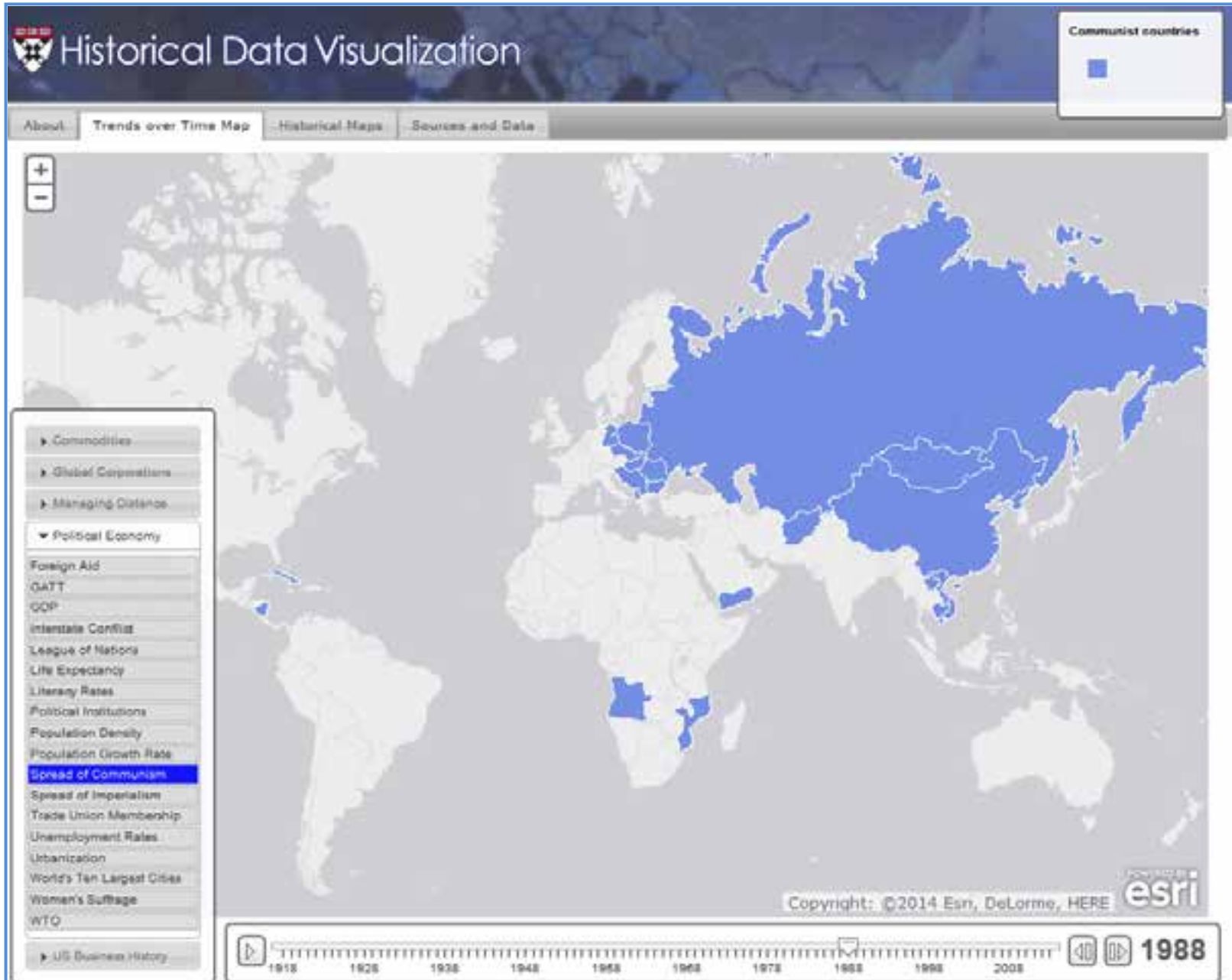
Giovanni Zambotti, Ben Lewis, Wendy Guan

HARVARD BUSINESS SCHOOL

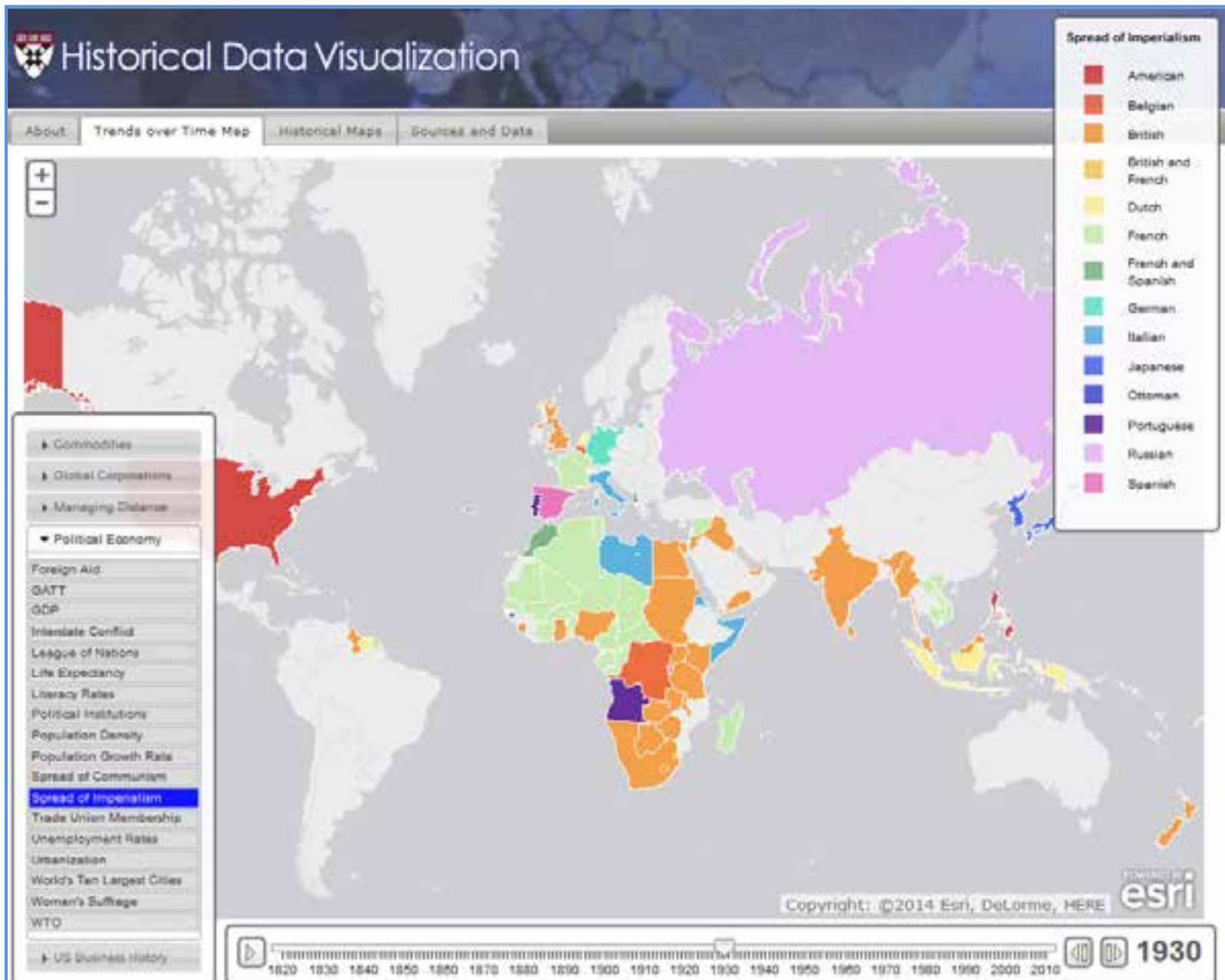




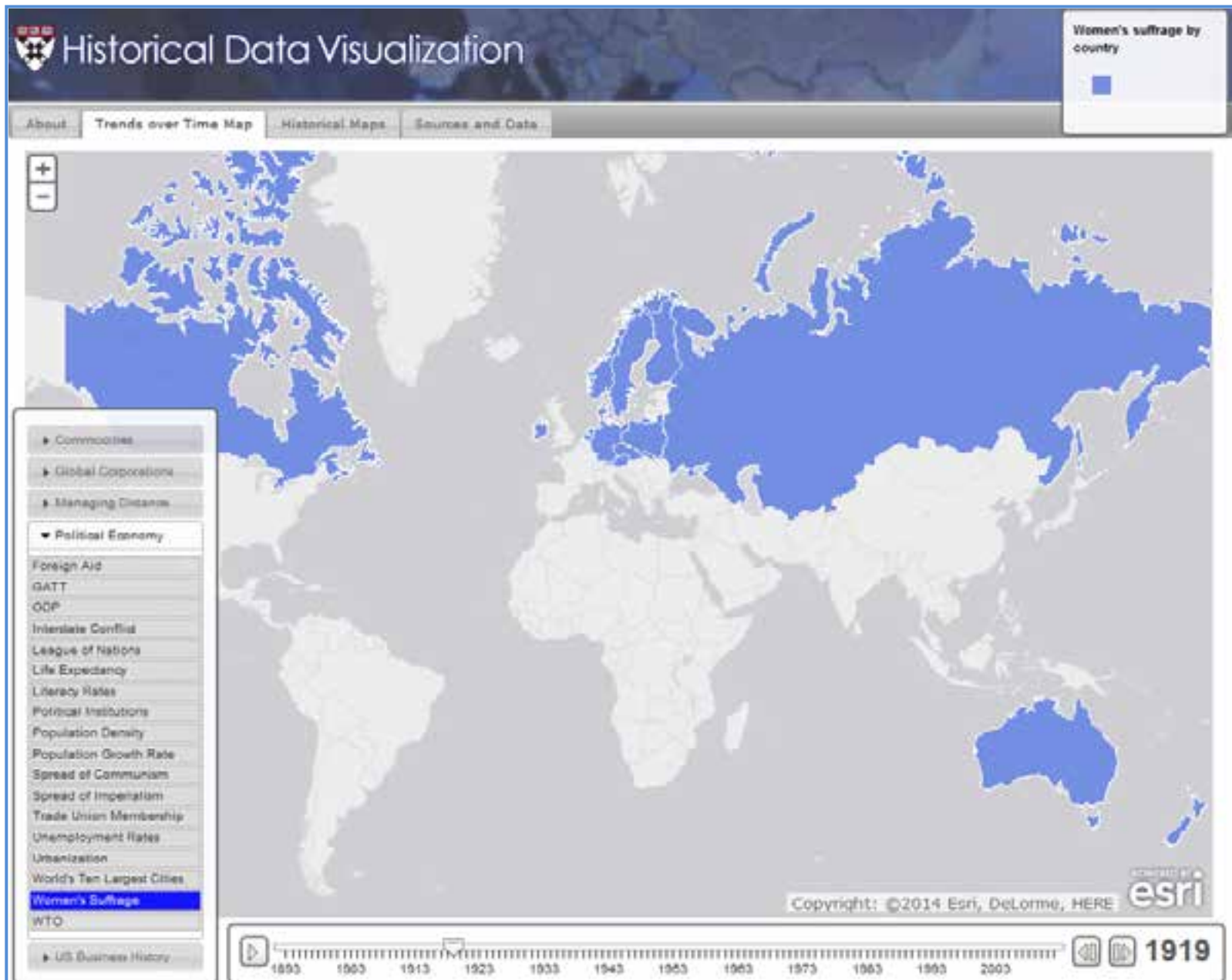
Communism



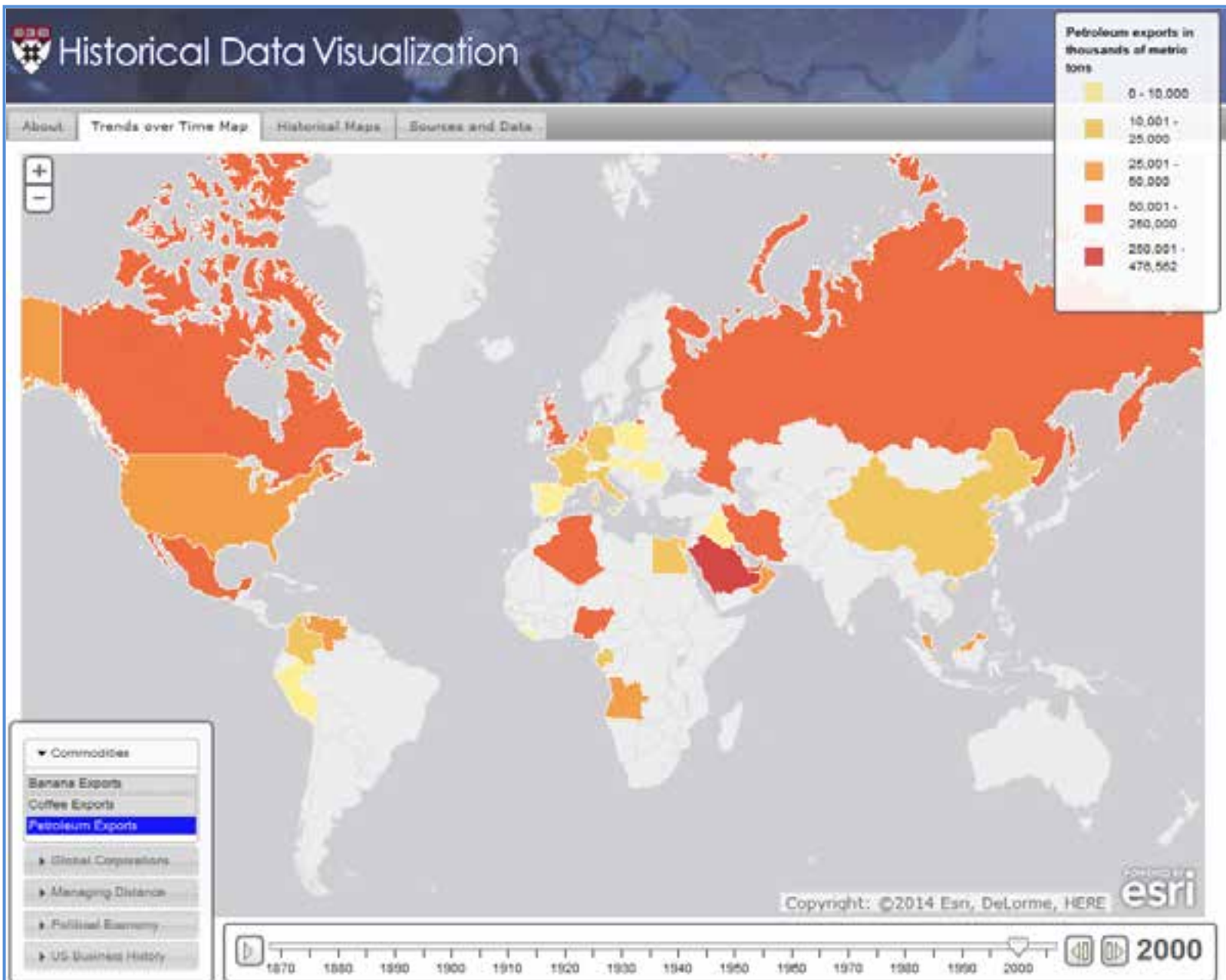
Spread of Imperialism



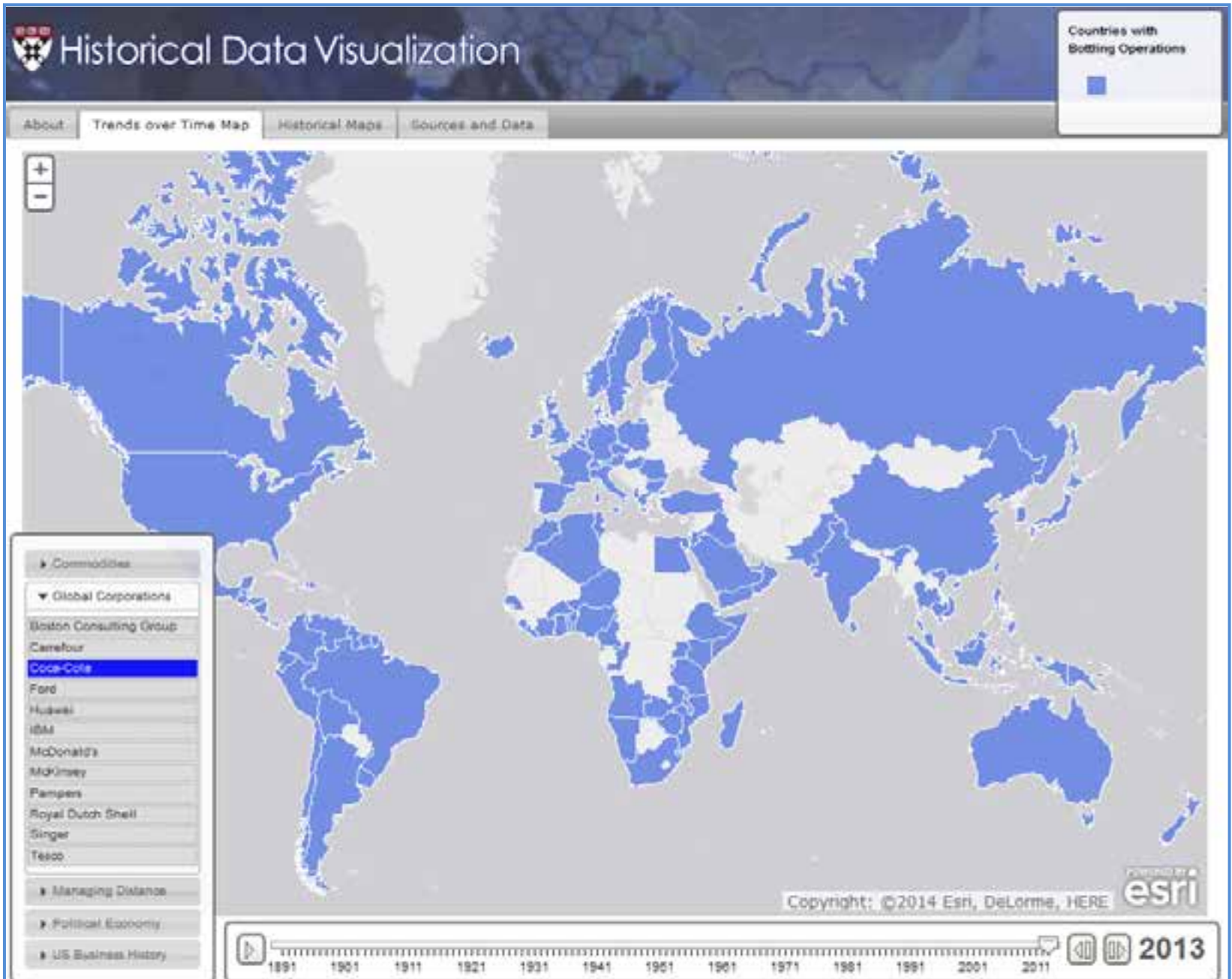
Gender Relations



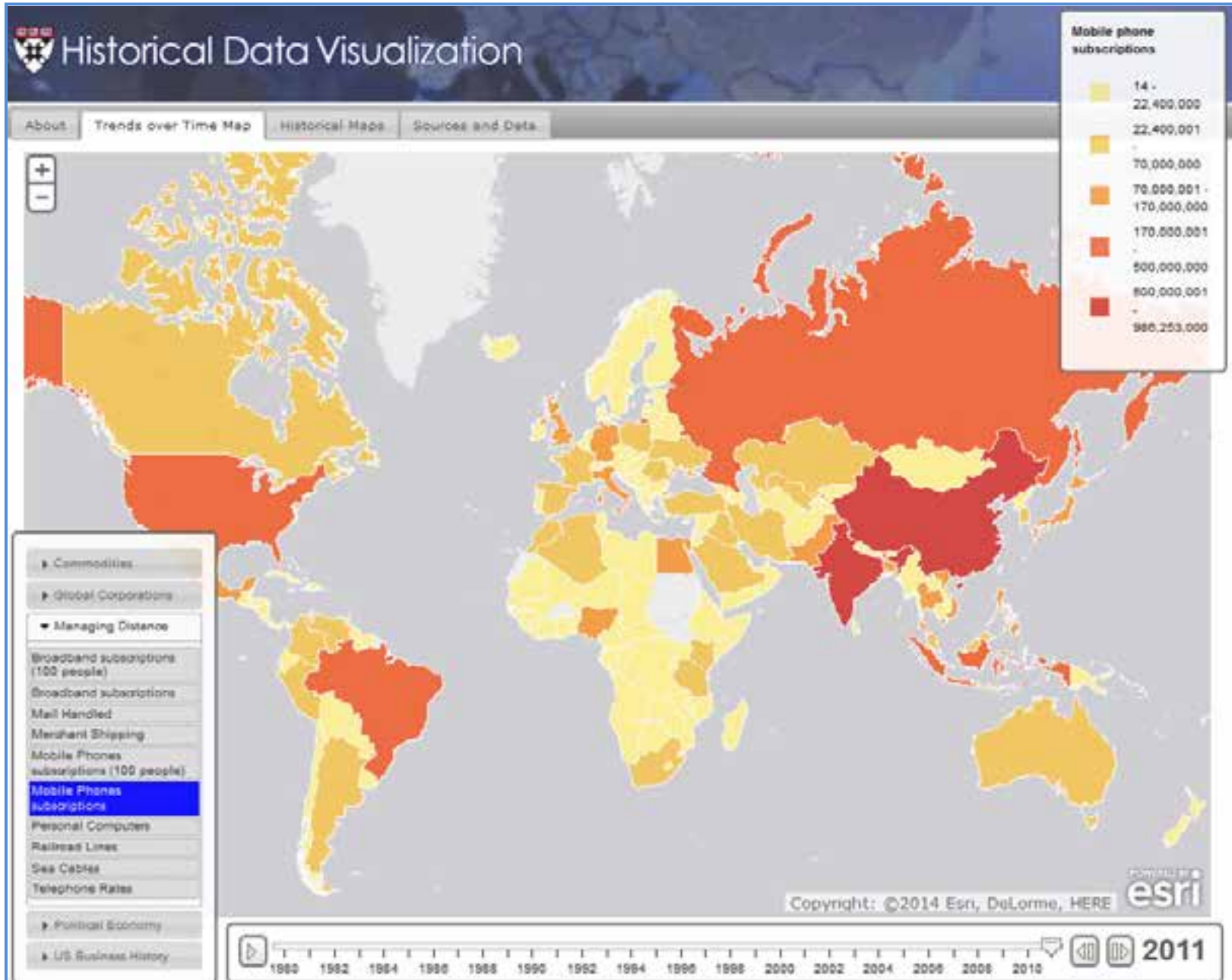
Category: Commodities



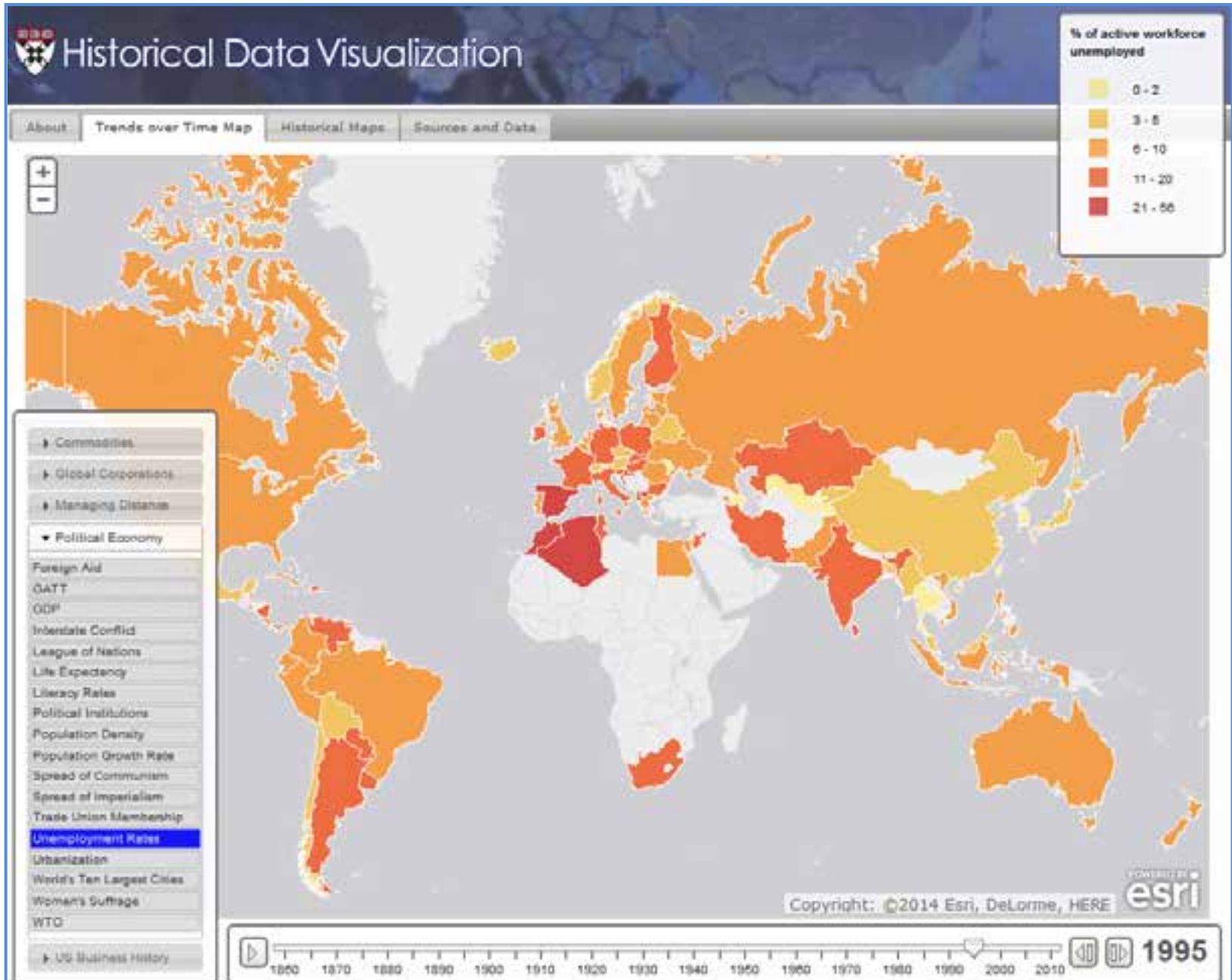
Category: Global Corporations



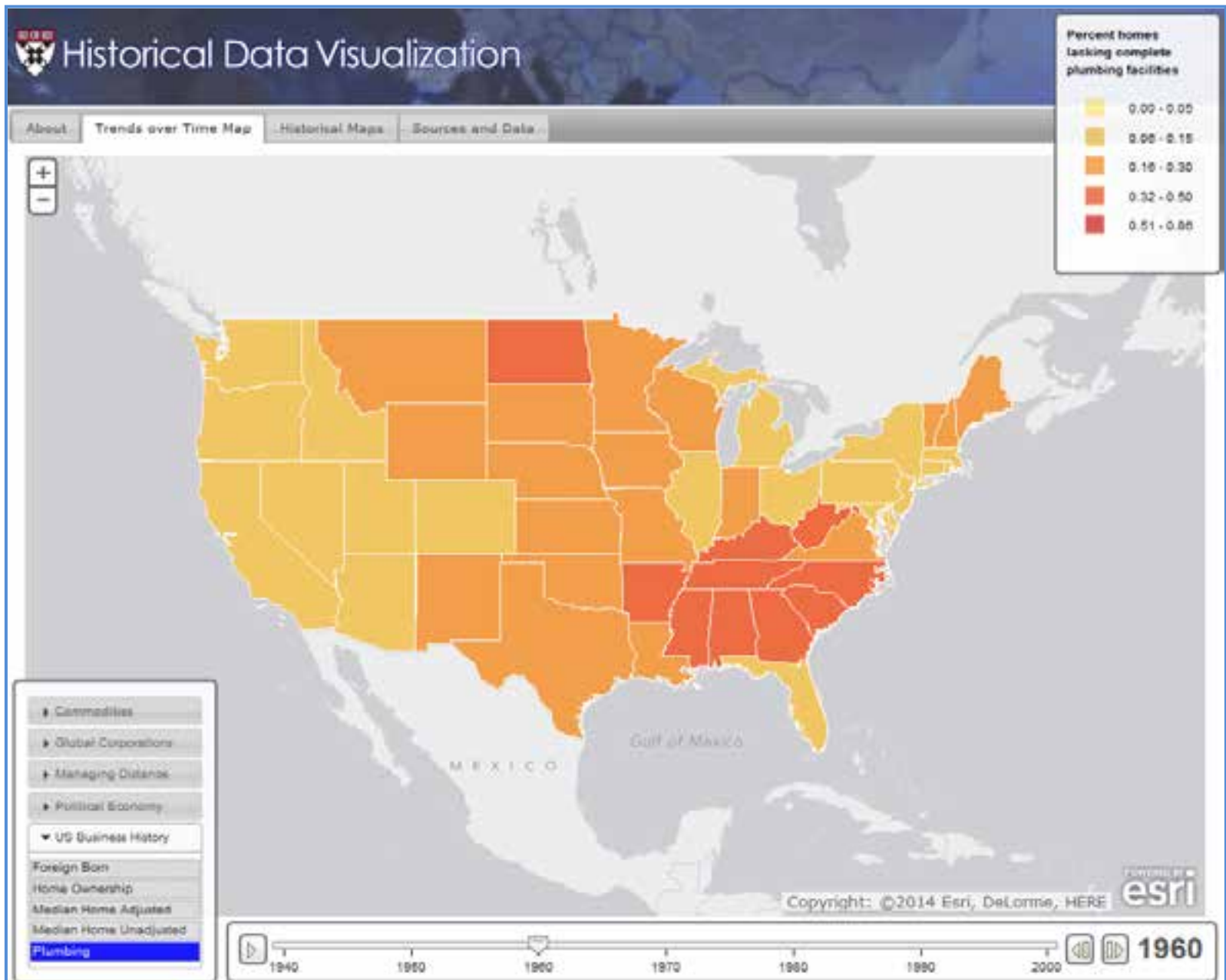
Category: Managing Distance



Category: Political Economy



Category: U.S. Business History



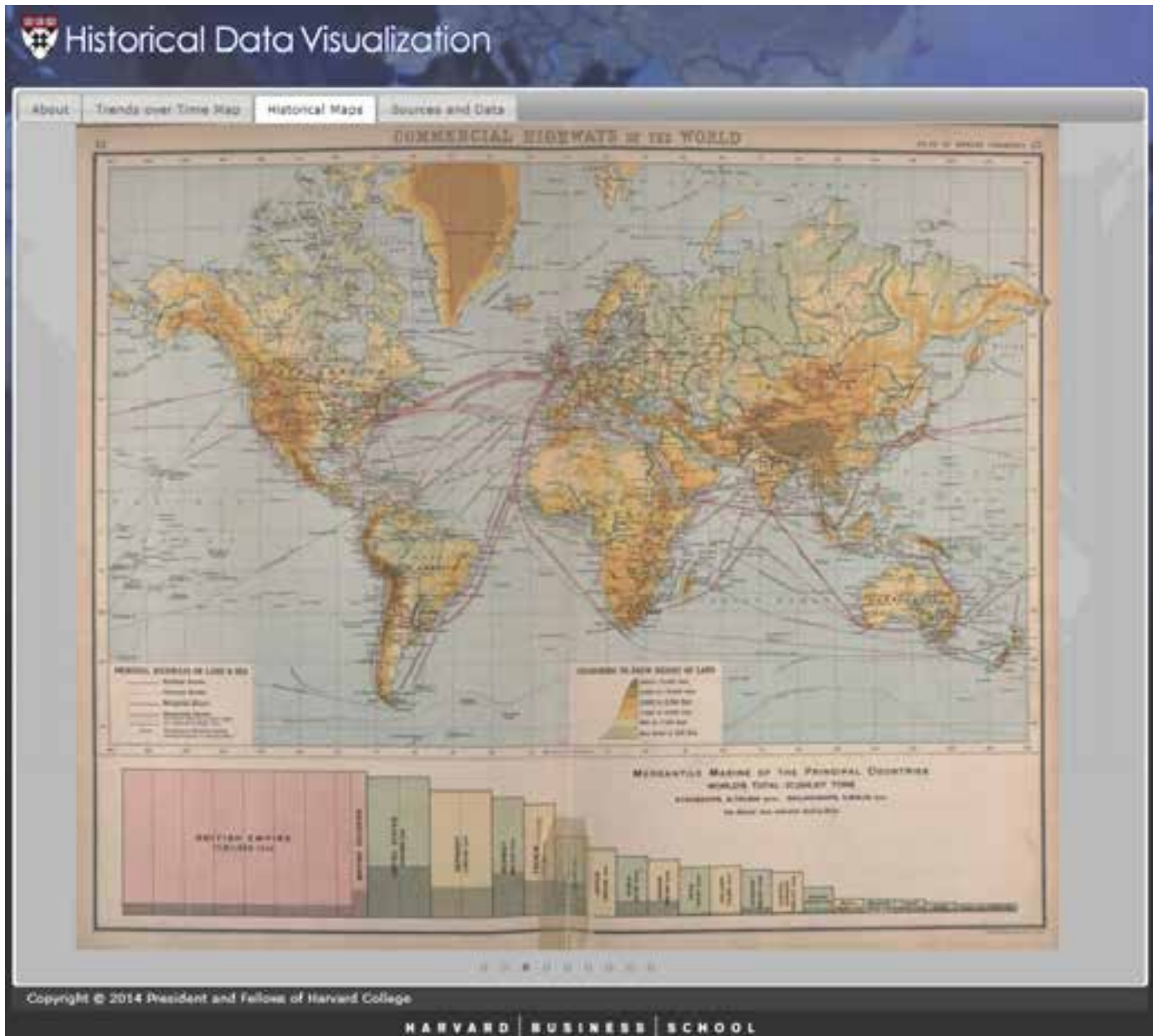
Data for Download and Analysis

Historical Data Visualization

About Trends over Time Map Historical Maps Sources and Data

Data Set	Source	Notes	Excel Downloads
Banana Exports	FAO (Food and Agricultural Organization of the United Nations), www.fao.org ; Food and Agricultural Organization of the United Nations, Yearbook of Food and Agricultural Statistics, various years; Institut International d'Agriculture, Annuaire International de Statistique Agricole, various years.		Download
Coffee Exports	FAO (Food and Agricultural Organization of the United Nations), www.fao.org ; Food and Agricultural Organization of the United Nations, Yearbook of Food and Agricultural Statistics, various years; Institut International d'Agriculture, Annuaire International de Statistique Agricole, various years.	United States coffee production from 1930-2007 was entirely based in Hawaii.	Download
Petroleum Exports	E.R. Mitchell, International Historical Statistics: Europe 1750-2005, New York: Palgrave-MacMillan, 2007; E.R. Mitchell, International Historical Statistics: the Americas 1750-2005, New York: Palgrave-MacMillan, 2007; E.R. Mitchell, International Statistics: Africa, Asia, Oceania 1750-2005, New York: Palgrave-MacMillan, 2007; United Nations, International Trade Statistics, various years.		Download
Boston Consulting Group	BCG website, http://www.bcg.com/about_bcg/history/default.asp		Download
Carrefour	http://carrefourusa.com/index.php?option=com_content&view=article&id=113&Itemid=44&lang=en http://news.bbc.co.uk/2/hi/business/4325567.stm http://www.bbc.co.uk/news/business-10403268 http://www.carrefour.com.sg/aboutus.aspx http://www.carrefour.ro/despre-grupa-carrefour.html http://www.carrefourjordan.com/index.php/en/jordan-license http://www.thejakartaglobe.com/archive/rench-supermarket-carrefour-to-ent-singapore/242981/		Download
Coca-Cola	Courtesy of the Coca-Cola Company, heritage Communications		Download
Ford	M. Wilkins, & P.R. Hill, American Business Abroad: Ford on Six Continents; Ford Motor Company Annual Reports, various years; Directory of American Firms Operating in Foreign Countries, various editions; OneSource database.		Download
Huawei	FedExia, accessed 9/5/2013		Download
IBM	J. Connolly, History of Computing in Europe; IBM Annual Reports, various years; Directory of American Firms Operating in Foreign Countries, various editions; OneSource		Download
McDonald's	McDonald's PR newswire, various others		Download
McInerney	McInerney	Georgia from 1978 - 1991 is recorded as Soviet Union	Download
Pampers	Proctor and Gamble		Download
Royal Dutch Shell	Jan Luitzen van Zanden, A History of Royal Dutch Shell, vol. 4 (Oxford: Oxford University Press, 2007), pp. 1.3.1-1.3.3; Royal Dutch Shell, Annual Report, various years.		Download
Singer	Godley, Andrew, "Selling the Sewing Machine Around the World: Singer's International Marketing Strategies, 1850-1920", Enterprise & Society: The International Journal of Business History, Vol. 7, No. 2, pp. 266-314, 2006 Available at SSRN: http://ssrn.com/abstract=951768		Download
Tesco	Tesco Website, http://www.tescopl.com/index.asp?pageid=8		Download
Broadband subscriptions (100 people)	The World Bank, http://data.worldbank.org/indicator/IT.NET.BBND.P1?display=default		Download
Broadband subscriptions	The World Bank, http://data.worldbank.org/indicator/IT.NET.BBND?display=default		Download
Mail Handled	E.R. Mitchell, International Historical Statistics: Europe 1750-2005, New York: Palgrave-MacMillan, 2007; E.R. Mitchell, International Historical Statistics: the Americas 1750-2005, New York: Palgrave-MacMillan, 2007; E.R. Mitchell, International Statistics: Africa, Asia, Oceania 1750-2005, New York: Palgrave-MacMillan, 2007; Universal Postal Union, www.upu.int	Values for China from 1900-1935 include data from Taiwan. Values for Austria-Hungary are a combination of values obtained for Austria and Hungary from 1870-1915.	Download
Merchant Shipping	Heraklis, Odessa, A History of Greek-Owned Shipping: The Making of an International Tramp Fleet, 1830 to Present Day, London, Routledge, 1986.	1993 Data was reported in 1992, 2003 Data was reported in 2000.	Download

Historical Maps



Building the System

- Step 1: Defined the system requirements
 - Web-based, public facing
 - Interactive map
 - Interactive time bar
 - Scanned historic maps
 - Access to data downloads and metadata for all layers
 - Able to support many layers and categories

Building the System

- Step 2: Defined the components needed
 - ArcGIS Server for GIS map publishing services
 - ArcGIS Server time map client
 - Hosting on Amazon EC2
 - ArcGIS desktop for data organization
 - Zhukov dataset for historic nation state boundaries

Overview of System Development

1. Develop data creation workflow
2. Install and configure server software
3. Develop web client application
4. Load data to system. Check for errors, fix, reload
5. Test system for performance and stability. Make adjustments as needed.

Data Creation Work Flow (1)

1. Create table of historical country ID from Zhukov dataset
2. Provided a table to historian team to join their variables to
3. Historians provide GIS team 1 spreadsheet for each map layer with country IDs and variables
4. GIS group uses ArcGIS to join tables to master table and then extract a dataset for each layer
5. New layer is generalized

Data Creation Work Flow (2)

6. Data layers loaded to ArcGIS server on Amazon
7. Historian team reviews temporal data on web application for accuracy and clarity of symbolization, passes changes back to GIS team
8. GIS team updates data until it is correct
9. Once a layer is final a video is produced of it

Zhukov Dataset (1875-1998)

	A	B	C	D	E	F	G	H	I	J	K	L
1	SP_ID	CNTRY_CODE	COUNTRY	OID	year	MAP	MAP_CCOI	year	numid	MAP_CNTRY	bbb	scode
2	2	206923.0000000000	Afghanistan	0	7001925	d1921-1939	206923	1925	700	Afghanistan	AFG	AFG
3	3	206924.0000000000	Albania	1	3391925	d1921-1939	206924	1925	339	Albania	ALB	ALB
4	10	206932.0000000000	Argentina	2	1601925	d1921-1939	206932	1925	160	Argentina	ARG	ARG
5	12	206937.0000000000	Australia	3	9001925	d1921-1939	206937	1925	900	Australia	AUL	AUL
6	13	206938.0000000000	Austria	4	3051925	d1921-1939	206938	1925	305	Austria	AUS	AUS
7	18	206945.0000000000	Belgium	5	2111925	d1921-1939	206945	1925	211	Belgium	BEL	BEL
8	23	206951.0000000000	Bolivia	6	1451925	d1921-1939	206951	1925	145	Bolivia	BOL	BOL
9	25	206954.0000000000	Brazil	7	1401925	d1921-1939	206954	1925	140	Brazil	BRA	BRA
10	28	206957.0000000000	Bulgaria	8	3551925	d1921-1939	206957	1925	355	Bulgaria	BIUL	BIUL
11	34	206963.0000000000	Canada	9	201925	d1921-1939	206963	1925	20	Canada	CAN	CAN
12	39	206969.0000000000	Chile	10	1551925	d1921-1939	206969	1925	155	Chile	CHL	CHL
13	245	206970.0000000000	China	11	7101925	d1921-1939	206970	1925	710	China	CHN	CHN
14	43	206974.0000000000	Colombia	12	1001925	d1921-1939	206974	1925	100	Colombia	COL	COL
15	47	206979.0000000000	Costa Rica	13	941925	d1921-1939	206979	1925	94	Costa Rica	COS	COS
16	48	206981.0000000000	Cuba	14	401925	d1921-1939	206981	1925	40	Cuba	CUB	CUB
17	50	206984.0000000000	Denmark	16	3901925	d1921-1939	206984	1925	390	Denmark	DEN	DEN
18	53	206987.0000000000	Dominican Republic	17	421925	d1921-1939	206987	1925	42	Dominican Republic	DOM	DOM
19	54	206988.0000000000	Ecuador	18	1301925	d1921-1939	206988	1925	130	Ecuador	ECU	ECU
20	55	206989.0000000000	Egypt	19	6511925	d1921-1939	206989	1925	651	Egypt	EGY	EGY
21	56	206990.0000000000	El Salvador	20	921925	d1921-1939	206990	1925	92	El Salvador	SAL	SAL
22	58	206992.0000000000	Estonia	21	3661925	d1921-1939	206992	1925	366	Estonia	EST	EST
23	59	206993.0000000000	Ethiopia	22	5301925	d1921-1939	206993	1925	530	Ethiopia	ETH	ETH
24	63	206997.0000000000	Finland	23	3751925	d1921-1939	206997	1925	375	Finland	FIN	FIN
25	64	206998.0000000000	France	24	2201925	d1921-1939	206998	1925	220	France	FRN	FRN
26	243	207005.0000000000	Germany	25	2551925	d1921-1939	207005	1925	255	Germany	GMY	GMY
27	71	207008.0000000000	Greece	26	3501925	d1921-1939	207008	1925	350	Greece	GRC	GRC
28	76	207013.0000000000	Guatemala	27	901925	d1921-1939	207013	1925	90	Guatemala	GUA	GUA
29	81	207018.0000000000	Honduras	28	911925	d1921-1939	207018	1925	91	Honduras	HON	HON
30	82	207020.0000000000	Hungary	29	3101925	d1921-1939	207020	1925	310	Hungary	HUN	HUN
31	86	207025.0000000000	Iran	30	6301925	d1921-1939	207025	1925	630	Iran	IRN	IRN

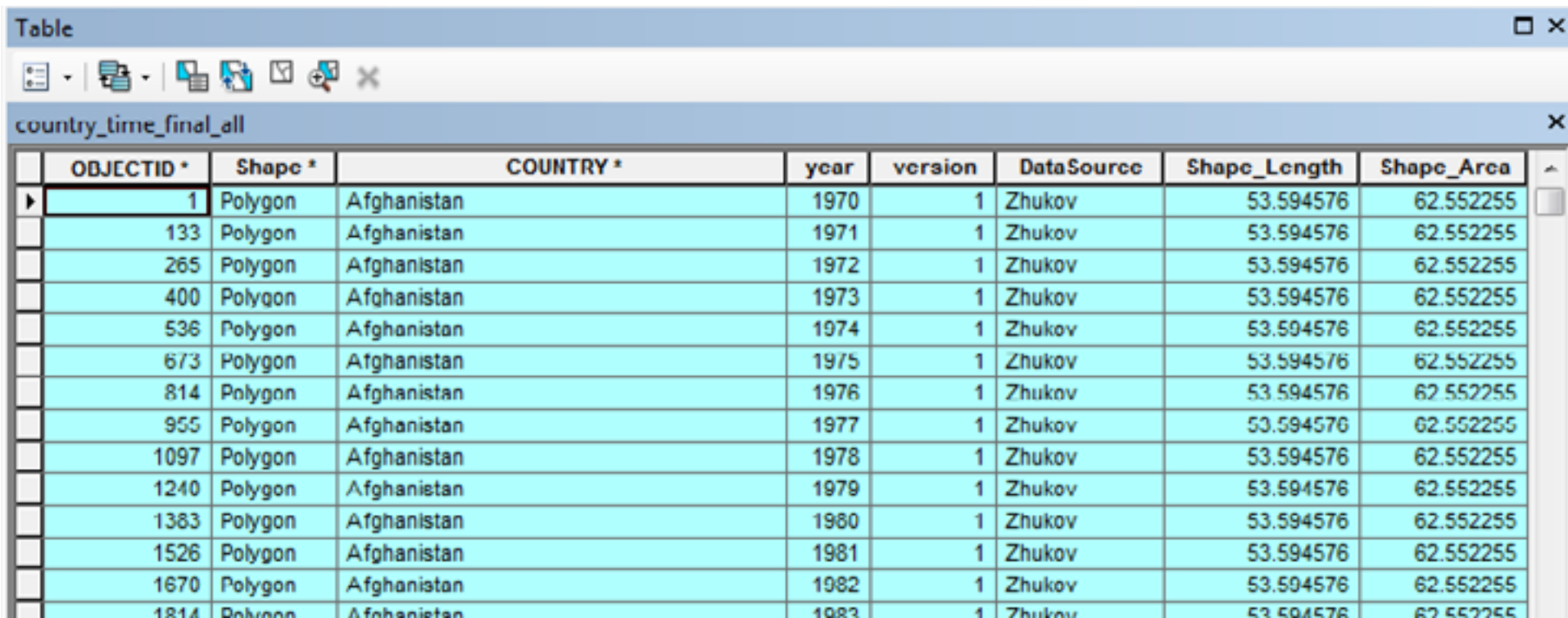
<http://hdl.handle.net/1902.1/21595>

Used Zhukov Data to Communicate between the Historian Team and GIS Team

1. Merged all Zhukov shapefiles into a single Feature Class
2. Exported the resulting Feature Class to a text file
3. Provided this large table to historians
4. Historians would then pull records they need for a given topic, add their data to those records, give modified table to GIS team
5. GIS team then uses ObjectID as the unique ID for joining business school data back to spatial features for mapping

Table of country extraction

- Merge Zhukov dataset into single FeatureClass (Global Zhukov Feature Class)
- Export the new Feature Class attributes to text



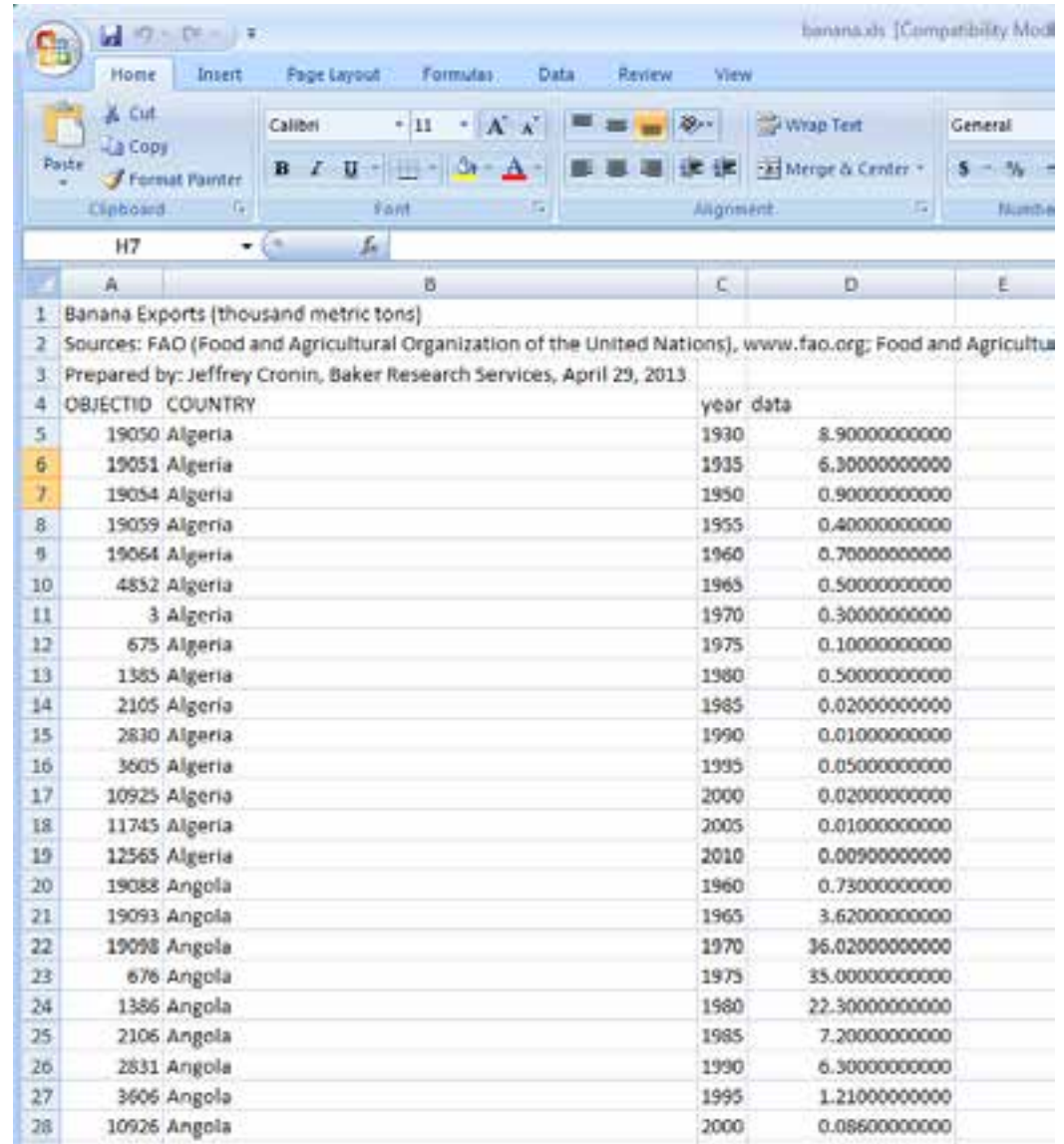
Table

country_time_final_all

OBJECTID *	Shape *	COUNTRY *	year	version	DataSource	Shape_Length	Shape_Area
1	Polygon	Afghanistan	1970	1	Zhukov	53.594576	62.552255
133	Polygon	Afghanistan	1971	1	Zhukov	53.594576	62.552255
265	Polygon	Afghanistan	1972	1	Zhukov	53.594576	62.552255
400	Polygon	Afghanistan	1973	1	Zhukov	53.594576	62.552255
536	Polygon	Afghanistan	1974	1	Zhukov	53.594576	62.552255
673	Polygon	Afghanistan	1975	1	Zhukov	53.594576	62.552255
814	Polygon	Afghanistan	1976	1	Zhukov	53.594576	62.552255
955	Polygon	Afghanistan	1977	1	Zhukov	53.594576	62.552255
1097	Polygon	Afghanistan	1978	1	Zhukov	53.594576	62.552255
1240	Polygon	Afghanistan	1979	1	Zhukov	53.594576	62.552255
1383	Polygon	Afghanistan	1980	1	Zhukov	53.594576	62.552255
1526	Polygon	Afghanistan	1981	1	Zhukov	53.594576	62.552255
1670	Polygon	Afghanistan	1982	1	Zhukov	53.594576	62.552255
1814	Polygon	Afghanistan	1983	1	Zhukov	53.594576	62.552255

Historians country/year/ID table

- Historians pulled the country/year/ID records they need for each subject matter layer (Banana Export, etc.) and associate relevant historic data with the records, returning the a set of spreadsheets to CGA, a spreadsheet for each layer to be created.



The screenshot shows an Excel spreadsheet with the following content:

1 Banana Exports (thousand metric tons)
2 Sources: FAO (Food and Agricultural Organization of the United Nations), www.fao.org; Food and Agriculture
3 Prepared by: Jeffrey Cronin, Baker Research Services, April 29, 2013

OBJECTID	COUNTRY	year	data
19050	Algeria	1930	8.90000000000000
19051	Algeria	1935	6.30000000000000
19054	Algeria	1950	0.90000000000000
19059	Algeria	1955	0.40000000000000
19064	Algeria	1960	0.70000000000000
4852	Algeria	1965	0.50000000000000
3	Algeria	1970	0.30000000000000
675	Algeria	1975	0.10000000000000
1385	Algeria	1980	0.50000000000000
2105	Algeria	1985	0.02000000000000
2830	Algeria	1990	0.01000000000000
3605	Algeria	1995	0.05000000000000
10925	Algeria	2000	0.02000000000000
11745	Algeria	2005	0.01000000000000
12565	Algeria	2010	0.00900000000000
19088	Angola	1960	0.73000000000000
19093	Angola	1965	3.62000000000000
19098	Angola	1970	36.02000000000000
676	Angola	1975	35.00000000000000
1386	Angola	1980	22.30000000000000
2106	Angola	1985	7.20000000000000
2831	Angola	1990	6.30000000000000
3606	Angola	1995	1.21000000000000
10926	Angola	2000	0.08600000000000

ArcGIS to join back

- GIS team used ArcGIS to join the tables from historians to the master table, then extract out a dataset for each layer with just the polygons required to render that time layer.

The image shows two overlapping windows. The left window is ArcGIS, displaying a table with the following columns: OBJECTID, Shape, COUNTRY, year, version, DataSource, Shape_Length, and Shape_Area. The right window is Microsoft Excel, showing a spreadsheet with columns A, B, C, and D. The spreadsheet contains data for 'Banana Exports (thousand metric tons)' for various years and countries, with a header row for 'year data'.

OBJECTID	Shape	COUNTRY	year	version	DataSource	Shape_Length	Shape_Area
19059	Polygon	Algeria	1950	3	Zhukov	11.260656	213.523661
19058	Polygon	Algeria	1964	3	Zhukov	11.260656	213.523661
19057	Polygon	Algeria	1953	3	Zhukov	11.260656	213.523661
19056	Polygon	Algeria	1952	3	Zhukov	11.260656	213.523661
19055	Polygon	Algeria	1951	3	Zhukov	11.260656	213.523661
19054	Polygon	Algeria	1950	3	Zhukov	11.260656	213.523661
19053	Polygon	Algeria	1945	3	Zhukov	11.260656	213.523661
19052	Polygon	Algeria	1940	3	Zhukov	11.260656	213.523661
19051	Polygon	Algeria	1938	3	Zhukov	11.260656	213.523661
19050	Polygon	Algeria	1930	3	Zhukov	11.260656	213.523661
19049	Polygon	Algeria	1925	3	Zhukov	11.260656	213.523661
19048	Polygon	Algeria	1920	3	Zhukov	11.260656	213.523661
19047	Polygon	Algeria	1915	3	Zhukov	11.260656	213.523661
19046	Polygon	Algeria	1913	3	Zhukov	11.260656	213.523661
19045	Polygon	Algeria	1910	3	Zhukov	11.260656	213.523661
19044	Polygon	Algeria	1905	3	Zhukov	11.260656	213.523661
19043	Polygon	Algeria	1900	3	Zhukov	11.260656	213.523661
19042	Polygon	Algeria	1895	3	Zhukov	11.260656	213.523661
19041	Polygon	Algeria	1890	3	Zhukov	11.260656	213.523661

	A	B	C	D
1	Banana Exports (thousand metric tons)			
2	Sources: FAO (Food and Agricultural Organization of the United Nations), www.fao.org; Food and			
3	Prepared by: Jeffrey Crutwin, Baker Research Services, April 25, 2013			
4	OBJECTID	COUNTRY	year	data
5	19050	Algeria	1930	8.9000000000
6	19051	Algeria	1935	6.3000000000
7	19054	Algeria	1950	0.9000000000
8	19053	Algeria	1955	0.4000000000
9	19064	Algeria	1960	0.7000000000
10	4852	Algeria	1965	0.5000000000
11	3	Algeria	1970	0.3000000000
12	675	Algeria	1975	0.1000000000
13	1385	Algeria	1980	0.5000000000
14	2105	Algeria	1985	0.0200000000

Geometry Generalization



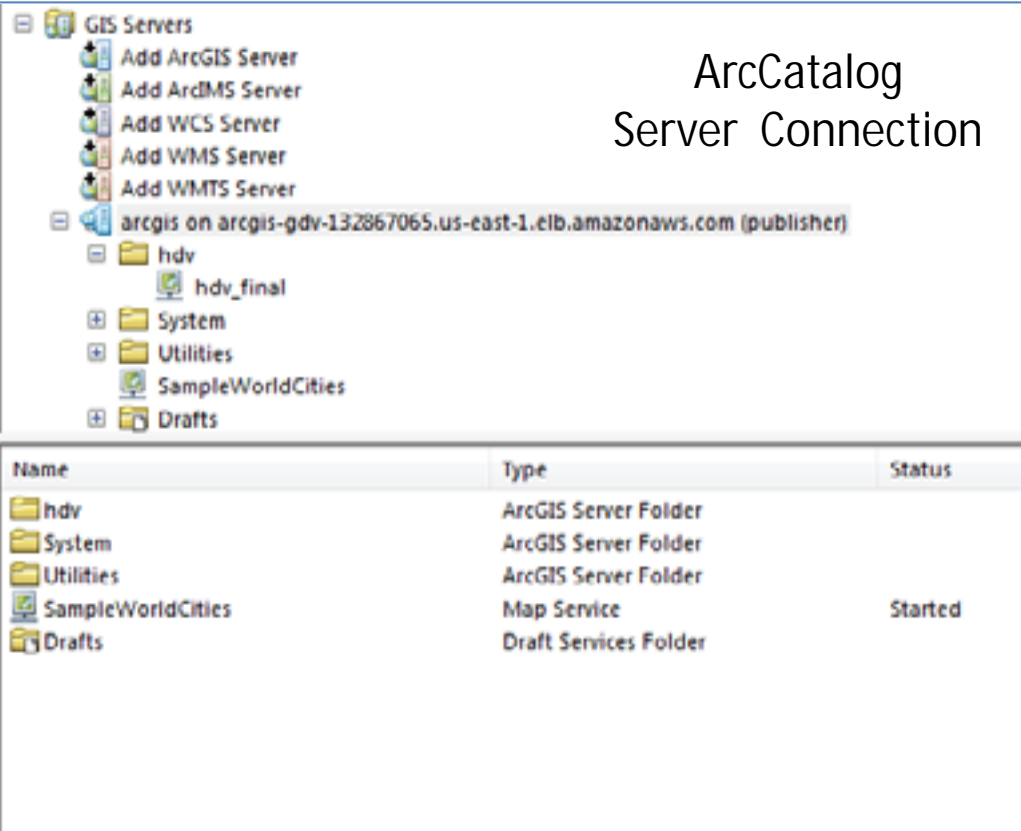
Using a tolerance of 10 kilometers,
dataset went from 7.9 GB to 886 MB.

Setting up ArcGIS Server on Amazon EC2

- Create an Amazon account
- Request ESRI to connect your Amazon account with their AWS AMI stack
- Create an Access Key ID and Secret access key on your AWS (need it for connection with ArcGIS Server and Cloud Builder)
- Install Cloud Builder on your desktop ArcGIS Server on AWS in order to initialize the AMI. (You will use Cloud Builder to manage your AMI.)
- Set up auto-scaling on your AWS if needed

Connecting Desktop and Server

ArcCatalog Server Connection



The screenshot shows the ArcCatalog interface. On the left, a tree view displays the 'GIS Servers' folder, which contains several 'Add' options (ArcGIS, ArcIMS, WCS, WMS, WMTS) and a connection to 'arcgis on arcgis-gdv-132867065.us-east-1.elb.amazonaws.com (publisher)'. Under this connection, there are folders for 'hdv', 'System', 'Utilities', 'SampleWorldCities', and 'Drafts'. The 'hdv' folder is expanded to show 'hdv_final'. On the right, a table lists the server details:

Name	Type	Status
hdv	ArcGIS Server Folder	
System	ArcGIS Server Folder	
Utilities	ArcGIS Server Folder	
SampleWorldCities	Map Service	Started
Drafts	Draft Services Folder	

ArcGIS REST Service Directory



The screenshot shows the ArcGIS REST Services Directory. The breadcrumb path is 'Home > services > hdv > hdv_final (MapServer)'. The page title is 'hdv/hdv_final (MapServer)'. Below the title, there are links for 'View In:' (ArcGIS JavaScript, ArcGIS.com Map, Google Earth, ArcMap, ArcGIS Explorer) and 'View Footprint In:' (ArcGIS.com Map). The 'Service Description' is 'HDV - update 2/24/2014'. The 'Map Name' is 'Layers'. There are links for 'Legend' and 'All Layers and Titles'. The 'Layers' section lists the following services:

- Commodities (0)
 - tanania (1)
 - coffee (2)
 - petroleum (3)
- Global Corporations (4)
 - tesa (5)
 - cerrefour (6)
 - cocacola (7)
 - ford (8)
 - framel (9)
 - ibm (10)
 - mcdonalds (11)
 - mckinsey (12)
 - pampers (13)
 - ryakutchshel (14)
 - singer (15)
 - tesco (16)
- Managing Distance (17)
 - broadsband100 (18)
 - broadsband50 (19)
 - mailandrei (20)
 - merchandise (21)
 - mobilephonesub100 (22)
 - mobilephonesub (23)
 - personalcomputers (24)

Spatio-Temporal Web Client Development

Built from scratch in JavaScript using:

- **jQuery UI** <https://jqueryui.com/>

and

- **ArcGIS JS API**
<https://developers.arcgis.com/javascript/>

Getting data right

- The GIS team loaded new tables to the web server and the historians checked it for accuracy and communicated necessary changes back to GIS team.
- Spatio-temporal data is tricky and getting numbers in books properly represented by accurate maps takes a lot of work!

Getting performance right

- Installed ArcGIS Server on Amazon EC2 and turned on auto –scaling
- We generalized country polygons by 10 KM (reduced size by 900%)
- Created video version of all layers, stored them in Vimeo

Thoughts for the future...

- Ideally the historian team and GIS team works together in the same office rather than separate offices.
- A dataset like Zhukov should be hosted like an open source project and improved over time by the community.
- Still no robust (open or closed source) system for animating spatio-temporal data flexibly. Let's build one!

Thank you

- We welcome your feedback, commentary, and corrections.
- For content related issues, please contact wfriedman@hbs.edu, or etg@hbs.edu
- GIS related issues, please contact gzambotti@cga.harvard.edu

Credits:

- Geoffrey Jones – Isidor Straus Professor of Business History
- Walter Friedman – Director, Harvard Business History Initiative, Lecturer of Business Administration
- Patrick Clapp – Harvard Business School Library, Information Research Specialist
- Jeffrey Cronin - Harvard Business School Library, Information Research Specialist
- Giovanni Zambotti – Harvard GIS Specialist
- Ben Lewis – Harvard GIS Project Manager