WEB SCRAPING FOR THE CODING-CHALLENGED

FDLP WEBINAR

May 23rd, 2019

Carl p. Olson, Librarian III, Coordinator for government publications

Albert S. Cook Library, Towson University

TU TOWSON UNIVERSITY

Today's Agenda

- What is Data-Scraping?
- What Does One Do with Spreadsheet Data?
- Hard and Easy ways to scrape data;
- Data-Scraping HTML with Google Sheets;
- Further Information.

What is Data Scraping

Web-scraping is a (typically automated) process which transfers content from online documents to an interactive format, such as Excel or CSV, for analysis, aggregation, or further computation.

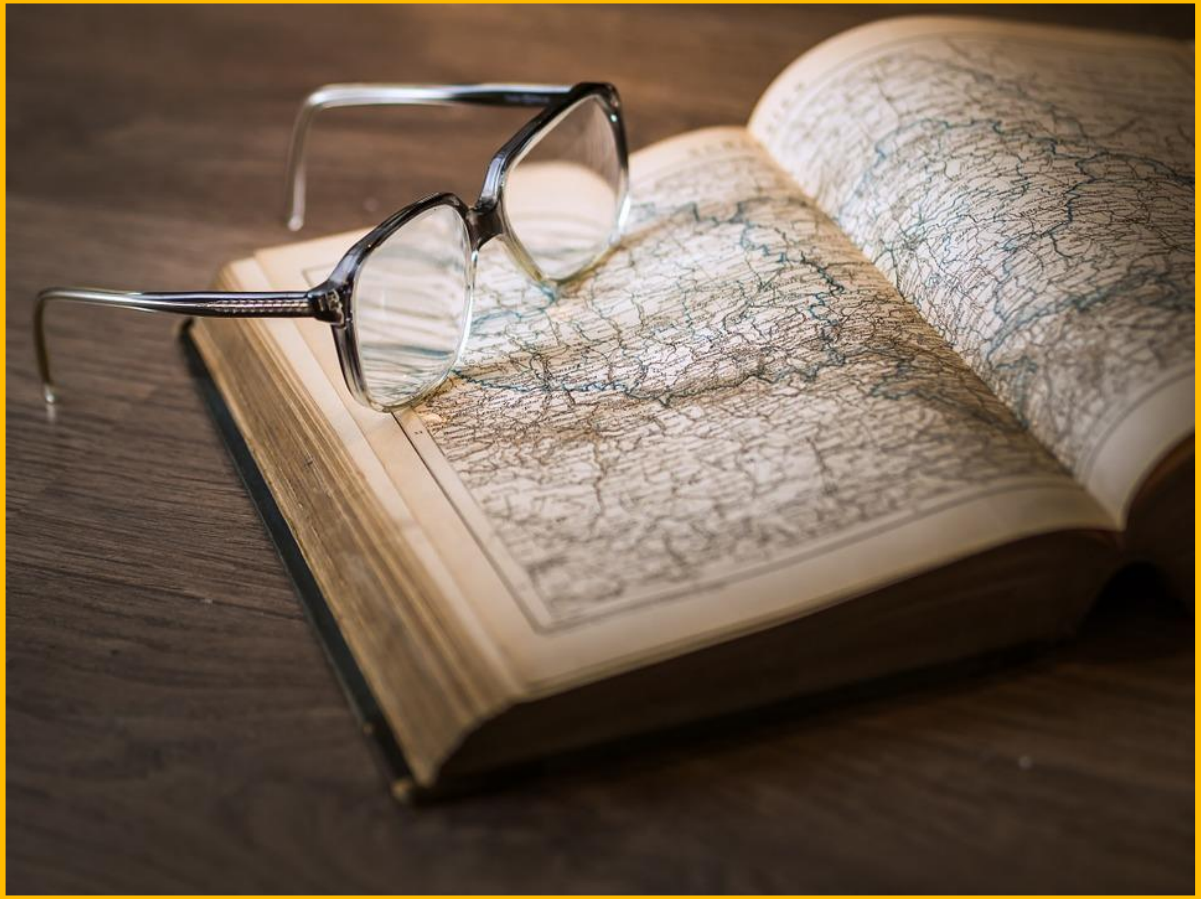
What is Data Scraping

- ❑ Web-scraping is as old as the Web itself;
- ❑ Web-scraping: "content-harvesting lite."
- ❑ Now used by business analysts, journalists, and researchers;
- ❑ Ccoding-challenged professionals on a deadline.

Why Do People Scrape Data?

TOWSON UNIVERSITY

- Directories;
- Employment listings;
- Products & pricing;
- Web addresses;
- Site maps;
- Annual reports;
- Data tables from reports;
- Documents!



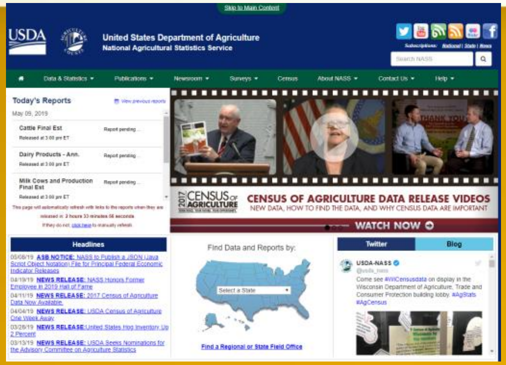
The Easiest Data-Scraping:

Many Federal Sites offer data in spreadsheets:

- ❑ BLS;
- ❑ BEA;
- ❑ FRED;
- ❑ FBI;
- ❑ DOA;
- ❑ CENSUS;
- ❑ CDC;
- ❑ NCES;
- ❑ BTS;

Criminal Justice Information Services (CJIS)

The FBI's Criminal Justice Information Services Division, or CJIS, is a high-tech hub in the hills of West Virginia that provides a range of state-of-the-art tools and services to law enforcement, national security and intelligence community partners, and the general public.



Skip to Main Content

USDA United States Department of Agriculture
National Agricultural Statistics Service

Subscriptions: National | State | News

Today's Reports — View previous reports

May 09, 2019

Cattle Final Est Released at 3:00 pm ET	Report pending ...
Dairy Products - Ann. Released at 3:00 pm ET	Report pending ...
Milk Cows and Production Final Est Released at 3:00 pm ET	Report pending ...

This page will automatically refresh with links to the reports when they are released in: 2 hours 33 minutes 56 seconds

If they do not, click here to manually refresh.

2017 CENSUS OF AGRICULTURE — CENSUS OF AGRICULTURE DATA RELEASE VIDEOS
NEW DATA, HOW TO FIND THE DATA, AND WHY CENSUS DATA ARE IMPORTANT
WATCH NOW

Headlines

- 05/08/19 ASB NOTICE: NASS to Publish a JSON (Java Script Object Notation) File for Principal Federal Economic Indicator Releases
- 04/19/19 NEWS RELEASE: NASS Honors Former Employee in 2019 Hall of Fame
- 04/11/19 NEWS RELEASE: 2017 Census of Agriculture Data Now Available
- 04/04/19 NEWS RELEASE: USDA Census of Agriculture One Week Away
- 03/28/19 NEWS RELEASE: United States Hog Inventory Up 2 Percent
- 03/13/19 NEWS RELEASE: USDA Seeks Nominations for the Advisory Committee on Agriculture Statistics

Find Data and Reports by:

Select a State

Find a Regional or State Field Office

Twitter | Blog

USDA-NASS @usda_nass

Come see #WICensusdata on display in the Wisconsin Department of Agriculture, Trade and Consumer Protection building lobby. #AgStats #AgCensus

bea | Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

April 19, 2019

Gross Domestic Product by Industry: Fourth Quarter and Annual 2018

Wholesale trade; mining; and information were the leading contributors to the increase in U.S. economic growth in the fourth quarter of 2018. Overall, 15 of 22 industry groups contributed to the 2.2 percent increase in real GDP in the fourth quarter.

Current Release

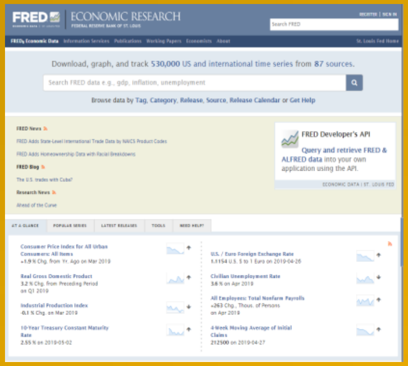
Principal Federal Economic Indicators

Gross Domestic Product	Q4 2018	+2.2%
Personal Income	February 2019	+0.2%
International Trade in Goods & Services (Trade Balance)	February 2019	-$49.4B
U.S. International Transactions (Current Account Balance)	Q4 2018	-$134.4B

Explore Data by Topic | U.S. Economy at a Glance | View Current Releases

Noteworthy

- National Data Catches Up in April
- Updated Digital Economy Data
- New Deputy Director
- Upcoming Release Schedule
- Updated Health Care Statistics
- County GDP Project



FRED ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

Download, graph, and track 530,000 US and international time series from 87 sources.

Search FRED data e.g., gdp, inflation, unemployment

Browse data by Tag, Category, Release, Source, Release Calendar or Get Help

FRED News

FRED Adds State-Level International Trade Data by NAICS Product Codes

FRED Adds Homeownership Data with Racial Breakdowns

FRED Blog

The U.S. trades with Cuba?

Research News

Ahead of the Curve

FRED Developer's API

Query and retrieve FRED & ALFRED data into your own application using the API.

ECONOMIC DATA | ST. LOUIS FED

Consumer Price Index for All Urban Consumers: All Items
+1.9 % Chg. from Yr. Ago on Mar 2019

Real Gross Domestic Product
3.2 % Chg. from Preceding Period on Q1 2019

Industrial Production Index
-0.1 % Chg. on Mar 2019

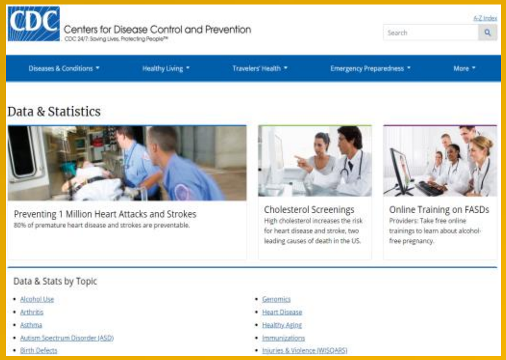
10-Year Treasury Constant Maturity Rate
2.55 % on 2019-05-02

U.S. / Euro Foreign Exchange Rate
1.1154 U.S. $ to 1 Euro on 2019-04-26

Civilian Unemployment Rate
3.6 % on Apr 2019

All Employees: Total Nonfarm Payrolls
+263 Chg., Thous. of Persons on Apr 2019

4-Week Moving Average of Initial Claims
212500 on 2019-04-27



CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

Data & Statistics

Preventing 1 Million Heart Attacks and Strokes
80% of premature heart disease and strokes are preventable.

Cholesterol Screenings
High cholesterol increases the risk for heart disease and stroke, two leading causes of death in the US.

Online Training on FASDs
Providers: Take free online trainings to learn about alcohol-free pregnancy.

Data & Stats by Topic

- Alcohol Use
- Arthritis
- Asthma
- Autism Spectrum Disorder (ASD)
- Birth Defects
- Genomics
- Heart Disease
- Healthy Aging
- Immunizations
- Injuries & Violence (WISQARS)

UNITED STATES DEPARTMENT OF LABOR

BUREAU OF LABOR STATISTICS

Databases, Tables & Calculators by Subject

On This Page:

- Inflation & Prices
- Employment
- Unemployment
- Employment Projections
- Pay & Benefits
- Spending & Time Use
- Productivity
- Workplace Injuries
- Occupational Requirements
- Regional Resources
- International
- Historical News Release Tables
- Maps
- Calculators
- Public Data API

Inflation & Prices

Database Name	Special Notice	Top Picks	Data Finder	One Screen	Multi-Screen	Tables	Text Files
Prices - Consumer							
All Urban Consumers (Current Series) (Consumer Price Index - CPI)	●	★	●	●	●	●	●
Urban Wage Earners and Clerical Workers (Current Series) (Consumer Price Index - CPI)	●	★	●	●	●	●	●
All Urban Consumers (Chained CPI) (Consumer Price Index - CPI)		★	●	●	●	●	●
Average Price Data (Consumer Price Index - CPI)	●	★	●	●	●		●

Google: site:gov filetype:xlsx [kw]

Why Do People Scrape Data?

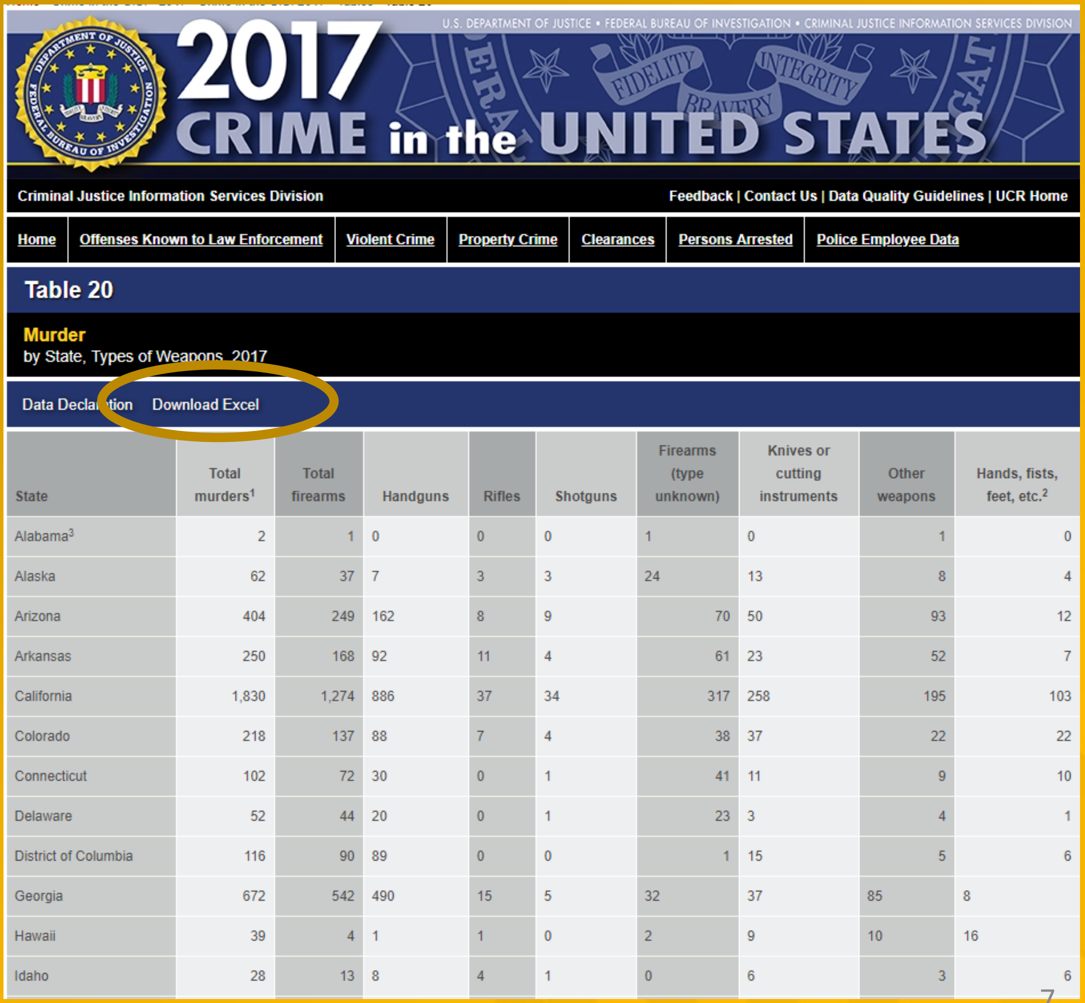
TOWSON UNIVERSITY

FBI, **Crime in the U.S.,** 2017;

Murders in the U.S.;

By State;

By Type of Weapon



U.S. DEPARTMENT OF JUSTICE • FEDERAL BUREAU OF INVESTIGATION • CRIMINAL JUSTICE INFORMATION SERVICES DIVISION

2017 CRIME in the UNITED STATES

Criminal Justice Information Services Division

Feedback | Contact Us | Data Quality Guidelines | UCR Home

Home | Offenses Known to Law Enforcement | Violent Crime | Property Crime | Clearances | Persons Arrested | Police Employee Data

Table 20

Murder

by State, Types of Weapons, 2017

Data Declaration Download Excel

State	Total murders[1]	Total firearms	Handguns	Rifles	Shotguns	Firearms (type unknown)	Knives or cutting instruments	Other weapons	Hands, fists, feet, etc.[2]
Alabama[3]	2	1	0	0	0	1	0	1	0
Alaska	62	37	7	3	3	24	13	8	4
Arizona	404	249	162	8	9	70	50	93	12
Arkansas	250	168	92	11	4	61	23	52	7
California	1,830	1,274	886	37	34	317	258	195	103
Colorado	218	137	88	7	4	38	37	22	22
Connecticut	102	72	30	0	1	41	11	9	10
Delaware	52	44	20	0	1	23	3	4	1
District of Columbia	116	90	89	0	0	1	15	5	6
Georgia	672	542	490	15	5	32	37	85	8
Hawaii	39	4	1	1	0	2	9	10	16
Idaho	28	13	8	4	1	0	6	3	6

Why Do People Scrape Data?

"Interviewing" data:

- Autosum
- Transpose;
- Ranking;
- Ratios.

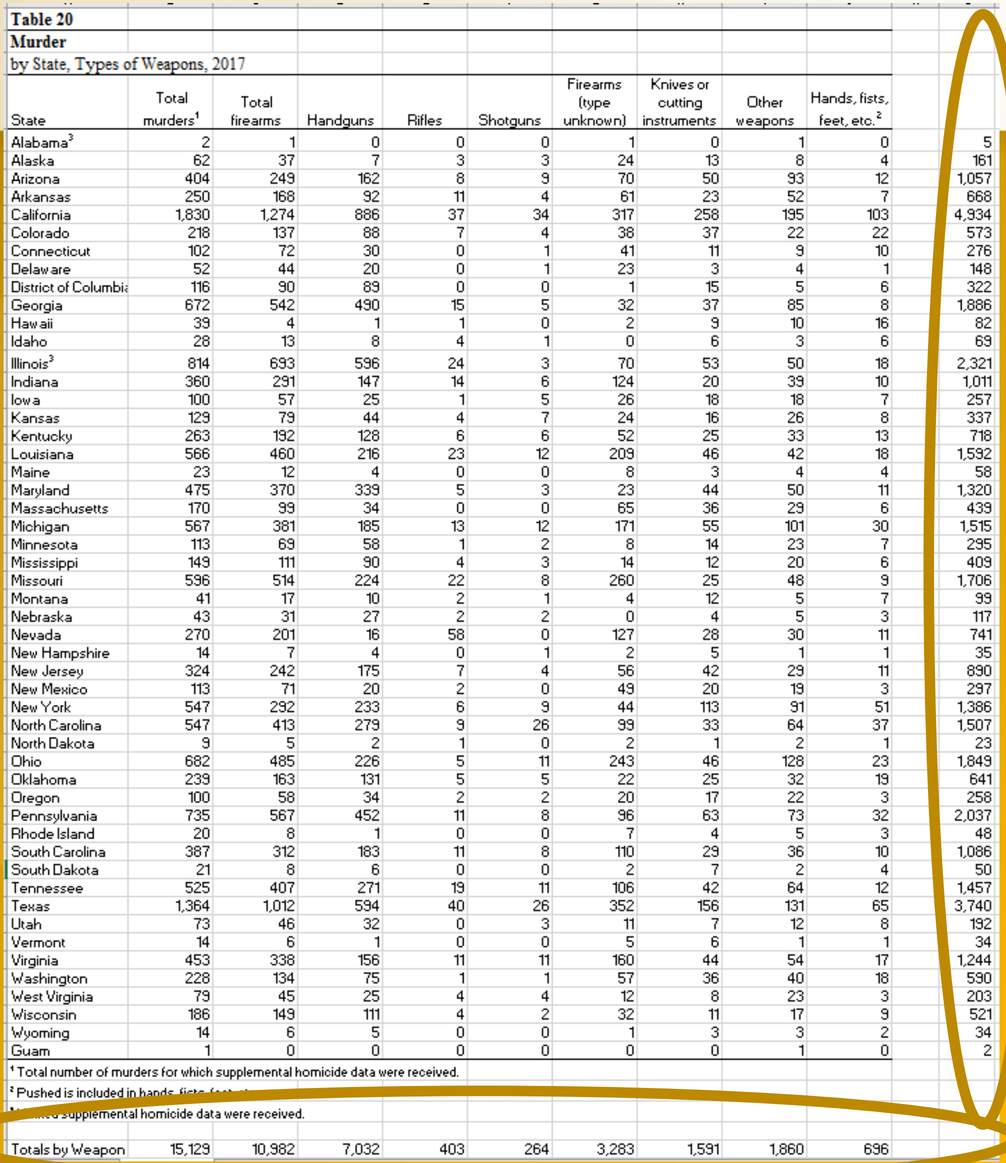


Table 20

Murder

by State, Types of Weapons, 2017

State	Total murders[1]	Total firearms	Handguns	Rifles	Shotguns	Firearms (type unknown)	Knives or cutting instruments	Other weapons	Hands, fists, feet, etc.[2]	
Alabama[3]	2	1	0	0	0	1	0	1	0	5
Alaska	62	37	7	3	3	24	13	8	4	161
Arizona	404	249	162	8	9	70	50	93	12	1,057
Arkansas	250	168	92	11	4	61	23	52	7	668
California	1,830	1,274	886	37	34	317	258	195	103	4,934
Colorado	218	137	88	7	4	38	37	22	22	573
Connecticut	102	72	30	0	1	41	11	9	10	276
Delaware	52	44	20	0	1	23	3	4	1	148
District of Columbi	116	90	89	0	0	1	15	5	6	322
Georgia	672	542	490	15	5	32	37	85	8	1,886
Hawaii	39	4	1	1	0	2	9	10	16	82
Idaho	28	13	8	4	1	0	6	3	6	69
Illinois[3]	814	693	596	24	3	70	53	50	18	2,321
Indiana	360	291	147	14	6	124	20	39	10	1,011
Iowa	100	57	25	1	5	26	18	18	7	257
Kansas	129	79	44	4	7	24	16	26	8	337
Kentucky	263	192	128	6	6	52	25	33	13	718
Louisiana	566	460	216	23	12	209	46	42	18	1,592
Maine	23	12	4	0	0	8	3	4	4	58
Maryland	475	370	339	5	3	23	44	50	11	1,320
Massachusetts	170	99	34	0	0	65	36	29	6	439
Michigan	567	381	185	13	12	171	55	101	30	1,515
Minnesota	113	69	58	1	2	8	14	23	7	295
Mississippi	149	111	90	4	3	14	12	20	6	409
Missouri	596	514	224	22	8	260	25	48	9	1,706
Montana	41	17	10	2	1	4	12	5	7	99
Nebraska	43	31	27	2	2	0	4	5	3	117
Nevada	270	201	16	58	0	127	28	30	11	741
New Hampshire	14	7	4	0	1	2	5	1	1	35
New Jersey	324	242	175	7	4	56	42	29	11	890
New Mexico	113	71	20	2	0	49	20	19	3	297
New York	547	292	233	6	9	44	113	91	51	1,386
North Carolina	547	413	279	9	26	99	33	64	37	1,507
North Dakota	9	5	2	1	0	2	1	2	1	23
Ohio	682	485	226	5	11	243	46	128	23	1,849
Oklahoma	239	163	131	5	5	22	25	32	19	641
Oregon	100	58	34	2	2	20	17	22	3	258
Pennsylvania	735	567	452	11	8	96	63	73	32	2,037
Rhode Island	20	8	1	0	0	7	4	5	3	48
South Carolina	387	312	183	11	8	110	29	36	10	1,086
South Dakota	21	8	6	0	0	2	7	2	4	50
Tennessee	525	407	271	19	11	106	42	64	12	1,457
Texas	1,364	1,012	594	40	26	352	156	131	65	3,740
Utah	73	46	32	0	3	11	7	12	8	192
Vermont	14	6	1	0	0	5	6	1	1	34
Virginia	453	338	156	11	11	160	44	54	17	1,244
Washington	228	134	75	1	1	57	36	40	18	590
West Virginia	79	45	25	4	4	12	8	23	3	203
Wisconsin	186	149	111	4	2	32	11	17	9	521
Wyoming	14	6	5	0	0	1	3	3	2	34
Guam	1	0	0	0	0	0	0	1	0	2
Totals by Weapon	15,129	10,982	7,032	403	264	3,283	1,591	1,860	696	

[1] Total number of murders for which supplemental homicide data were received.

[2] Pushed is included in hands, fists, feet, etc.

[3] [illegible] supplemental homicide data were received.

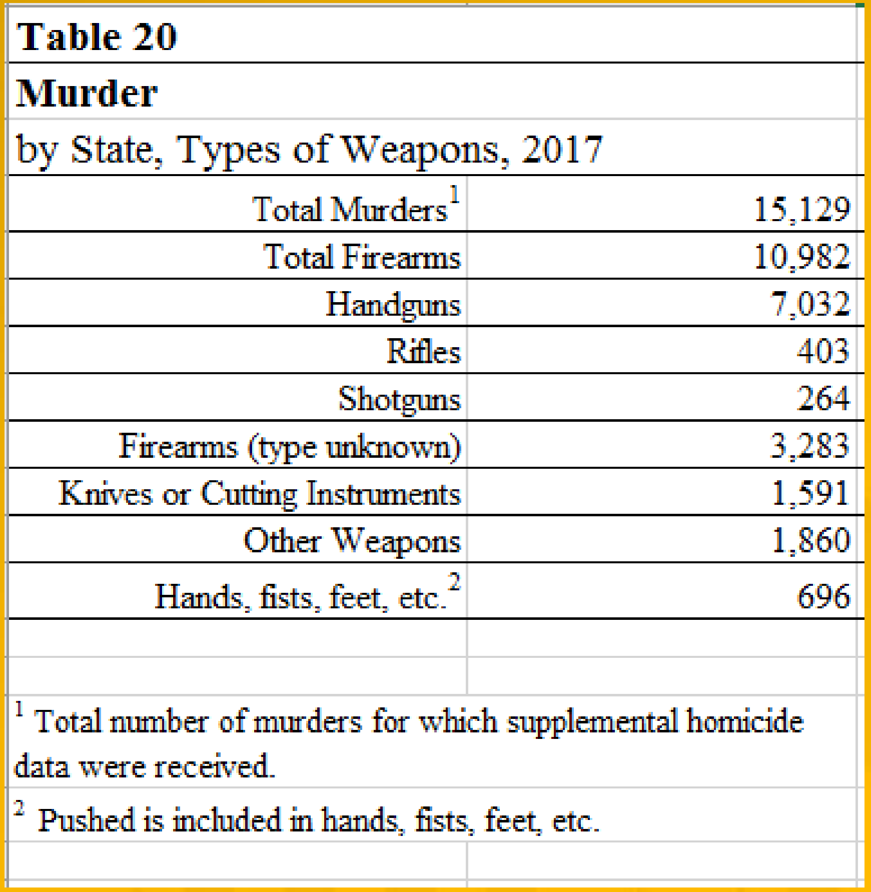


Table 20

Murder

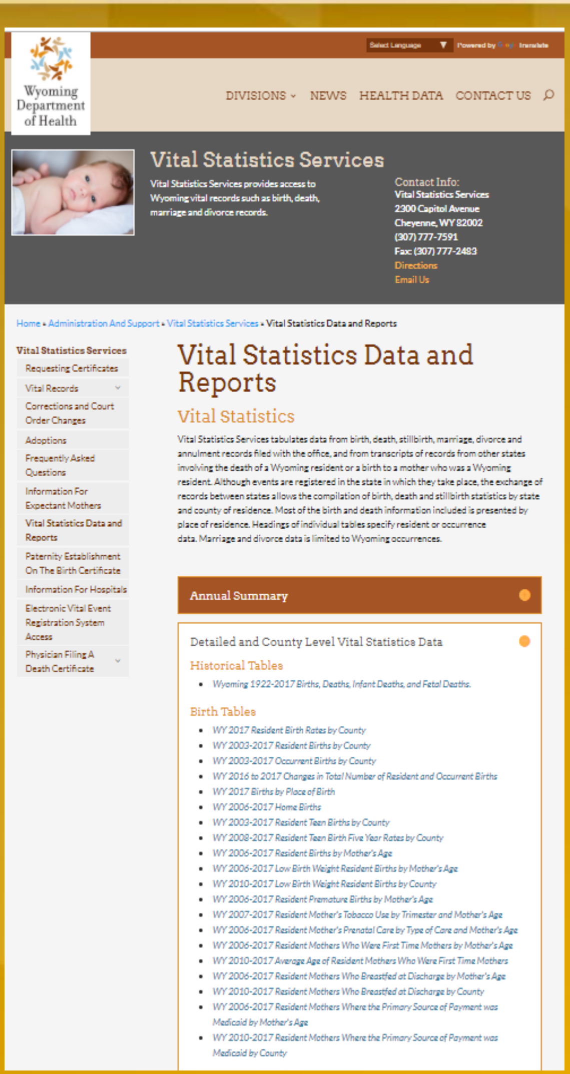
by State, Types of Weapons, 2017

Total Murders[1]	15,129
Total Firearms	10,982
Handguns	7,032
Rifles	403
Shotguns	264
Firearms (type unknown)	3,283
Knives or Cutting Instruments	1,591
Other Weapons	1,860
Hands, fists, feet, etc.[2]	696

[1] Total number of murders for which supplemental homicide data were received.

[2] Pushed is included in hands, fists, feet, etc.

What If It Isn't Online in XLS or CSV?


Select Language
Powered by Google Translate
Wyoming Department of Health
DIVISIONS
NEWS
HEALTH DATA
CONTACT US
Vital Statistics Services
Requesting Certificates
Vital Records
Corrections and Court Order Changes
Adoptions
Frequently Asked Questions
Information For Expectant Mothers
Vital Statistics Data and Reports
Paternity Establishment On The Birth Certificate
Information For Hospitals
Electronic Vital Event Registration System Access
Physician Filing A Death Certificate

Vital Statistics Services

Vital Statistics Services provides access to Wyoming vital records such as birth, death, marriage and divorce records.

Contact Info:
Vital Statistics Services
2300 Capitol Avenue
Cheyenne, WY 82002
(307) 777-7591
Fax: (307) 777-2483
Directions
Email Us

Home » Administration And Support » Vital Statistics Services » Vital Statistics Data and Reports

Vital Statistics Data and Reports

Vital Statistics

Vital Statistics Services tabulates data from birth, death, stillbirth, marriage, divorce and annulment records filed with the office, and from transcripts of records from other states involving the death of a Wyoming resident or a birth to a mother who was a Wyoming resident. Although events are registered in the state in which they take place, the exchange of records between states allows the compilation of birth, death and stillbirth statistics by state and county of residence. Most of the birth and death information included is presented by place of residence. Headings of individual tables specify resident or occurrence data. Marriage and divorce data is limited to Wyoming occurrences.

Annual Summary

Detailed and County Level Vital Statistics Data

Historical Tables

- *Wyoming 1922-2017 Births, Deaths, Infant Deaths, and Fetal Deaths.*

Birth Tables

- *WY 2017 Resident Birth Rates by County*
- *WY 2003-2017 Resident Births by County*
- *WY 2003-2017 Occurrent Births by County*
- *WY 2016 to 2017 Changes in Total Number of Resident and Occurrent Births*
- *WY 2017 Births by Place of Birth*
- *WY 2006-2017 Home Births*
- *WY 2003-2017 Resident Teen Births by County*
- *WY 2008-2017 Resident Teen Birth Five Year Rates by County*
- *WY 2006-2017 Resident Births by Mother's Age*
- *WY 2006-2017 Low Birth Weight Resident Births by Mother's Age*
- *WY 2010-2017 Low Birth Weight Resident Births by County*
- *WY 2006-2017 Resident Premature Births by Mother's Age*
- *WY 2007-2017 Resident Mother's Tobacco Use by Trimester and Mother's Age*
- *WY 2006-2017 Resident Mother's Prenatal Care by Type of Care and Mother's Age*
- *WY 2006-2017 Resident Mothers Who Were First Time Mothers by Mother's Age*
- *WY 2010-2017 Average Age of Resident Mothers Who Were First Time Mothers*
- *WY 2006-2017 Resident Mothers Who Breastfed at Discharge by Mother's Age*
- *WY 2010-2017 Resident Mothers Who Breastfed at Discharge by County*
- *WY 2006-2017 Resident Mothers Where the Primary Source of Payment was Medicaid by Mother's Age*
- *WY 2010-2017 Resident Mothers Where the Primary Source of Payment was Medicaid by County*

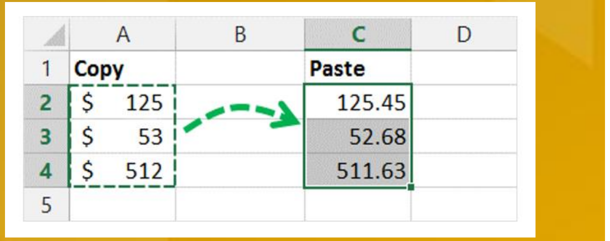
- Many agencies offer no Excel version;
  - Older documents;
  - Variable state, industry or agency standards;
  - Variable enforcement and compliance;
  - Smaller departments, sections or offices;
  - Federal councils, commissions or contractors.

What is the Hardest way?

TOWSON UNIVERSITY

The hardest way is to transcribe data by hand;

Next hardest is to copy and paste into Excel



	A	B	C	D
1	Copy		Paste	
2	$ 125		125.45	
3	$ 53		52.68	
4	$ 512		511.63	
5				

Result from Data.Census.Gov:

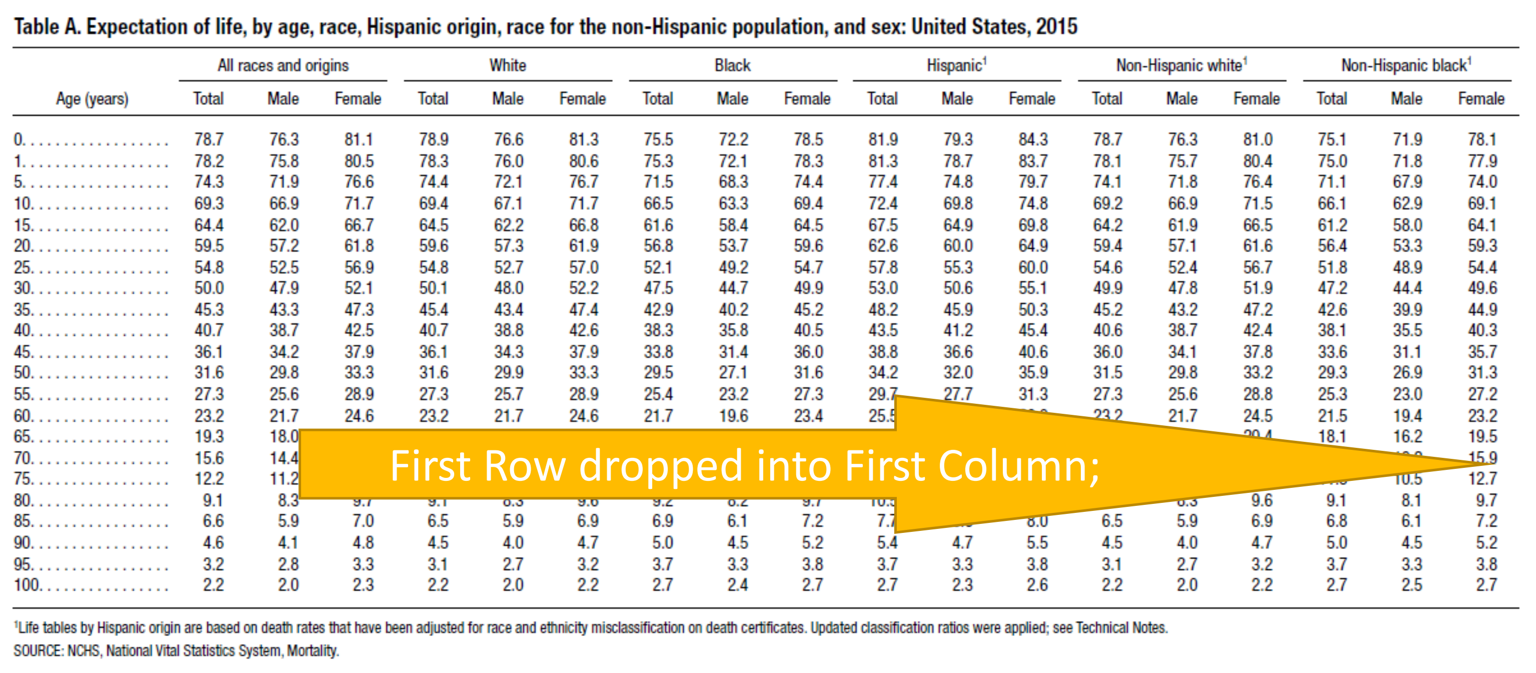


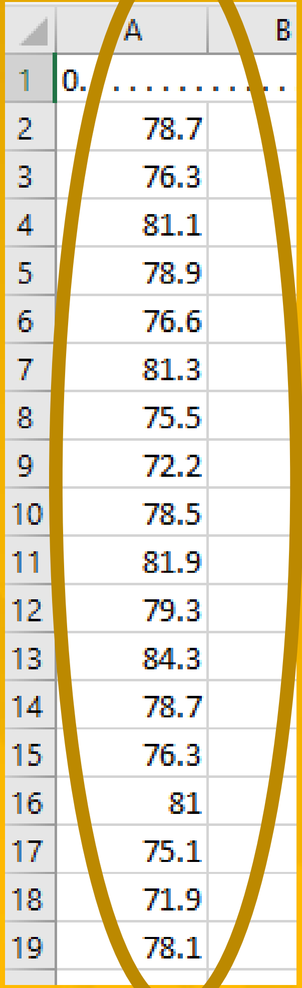
Table A. Expectation of life, by age, race, Hispanic origin, race for the non-Hispanic population, and sex: United States, 2015

	All races and origins			White			Black			Hispanic[1]			Non-Hispanic white[1]			Non-Hispanic black[1]		
Age (years)	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
0.	78.7	76.3	81.1	78.9	76.6	81.3	75.5	72.2	78.5	81.9	79.3	84.3	78.7	76.3	81.0	75.1	71.9	78.1
1.	78.2	75.8	80.5	78.3	76.0	80.6	75.3	72.1	78.3	81.3	78.7	83.7	78.1	75.7	80.4	75.0	71.8	77.9
5.	74.3	71.9	76.6	74.4	72.1	76.7	71.5	68.3	74.4	77.4	74.8	79.7	74.1	71.8	76.4	71.1	67.9	74.0
10.	69.3	66.9	71.7	69.4	67.1	71.7	66.5	63.3	69.4	72.4	69.8	74.8	69.2	66.9	71.5	66.1	62.9	69.1
15.	64.4	62.0	66.7	64.5	62.2	66.8	61.6	58.4	64.5	67.5	64.9	69.8	64.2	61.9	66.5	61.2	58.0	64.1
20.	59.5	57.2	61.8	59.6	57.3	61.9	56.8	53.7	59.6	62.6	60.0	64.9	59.4	57.1	61.6	56.4	53.3	59.3
25.	54.8	52.5	56.9	54.8	52.7	57.0	52.1	49.2	54.7	57.8	55.3	60.0	54.6	52.4	56.7	51.8	48.9	54.4
30.	50.0	47.9	52.1	50.1	48.0	52.2	47.5	44.7	49.9	53.0	50.6	55.1	49.9	47.8	51.9	47.2	44.4	49.6
35.	45.3	43.3	47.3	45.4	43.4	47.4	42.9	40.2	45.2	48.2	45.9	50.3	45.2	43.2	47.2	42.6	39.9	44.9
40.	40.7	38.7	42.5	40.7	38.8	42.6	38.3	35.8	40.5	43.5	41.2	45.4	40.6	38.7	42.4	38.1	35.5	40.3
45.	36.1	34.2	37.9	36.1	34.3	37.9	33.8	31.4	36.0	38.8	36.6	40.6	36.0	34.1	37.8	33.6	31.1	35.7
50.	31.6	29.8	33.3	31.6	29.9	33.3	29.5	27.1	31.6	34.2	32.0	35.9	31.5	29.8	33.2	29.3	26.9	31.3
55.	27.3	25.6	28.9	27.3	25.7	28.9	25.4	23.2	27.3	29.7	27.7	31.3	27.3	25.6	28.8	25.3	23.0	27.2
60.	23.2	21.7	24.6	23.2	21.7	24.6	21.7	19.6	23.4	25.5	[illegible]	[illegible]	23.2	21.7	24.5	21.5	19.4	23.2
65.	19.3	18.0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	20.4	18.1	16.2	19.5
70.	15.6	14.4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	15.9
75.	12.2	11.2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	10.5	12.7
80.	9.1	8.3	9.7	9.1	8.3	9.6	9.2	8.2	9.7	10.3	[illegible]	[illegible]	[illegible]	8.3	9.6	9.1	8.1	9.7
85.	6.6	5.9	7.0	6.5	5.9	6.9	6.9	6.1	7.2	7.7	[illegible]	8.0	6.5	5.9	6.9	6.8	6.1	7.2
90.	4.6	4.1	4.8	4.5	4.0	4.7	5.0	4.5	5.2	5.4	4.7	5.5	4.5	4.0	4.7	5.0	4.5	5.2
95.	3.2	2.8	3.3	3.1	2.7	3.2	3.7	3.3	3.8	3.7	3.3	3.8	3.1	2.7	3.2	3.7	3.3	3.8
100.	2.2	2.0	2.3	2.2	2.0	2.2	2.7	2.4	2.7	2.7	2.3	2.6	2.2	2.0	2.2	2.7	2.5	2.7

[1] Life tables by Hispanic origin are based on death rates that have been adjusted for race and ethnicity misclassification on death certificates. Updated classification ratios were applied; see Technical Notes.

SOURCE: NCHS, National Vital Statistics System, Mortality.

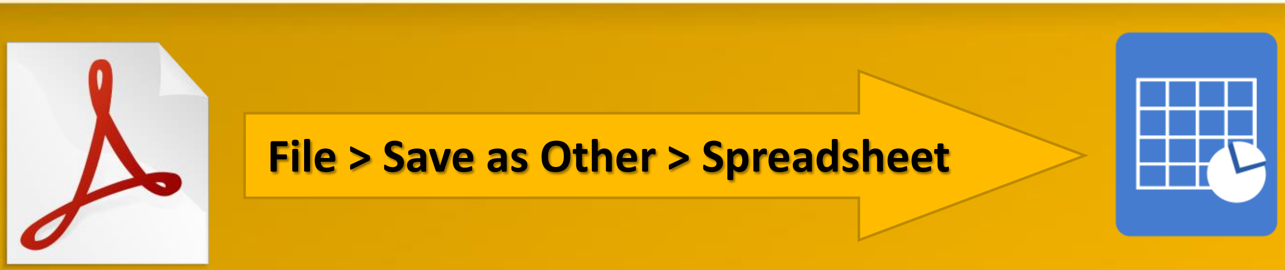
First Row dropped into First Column;



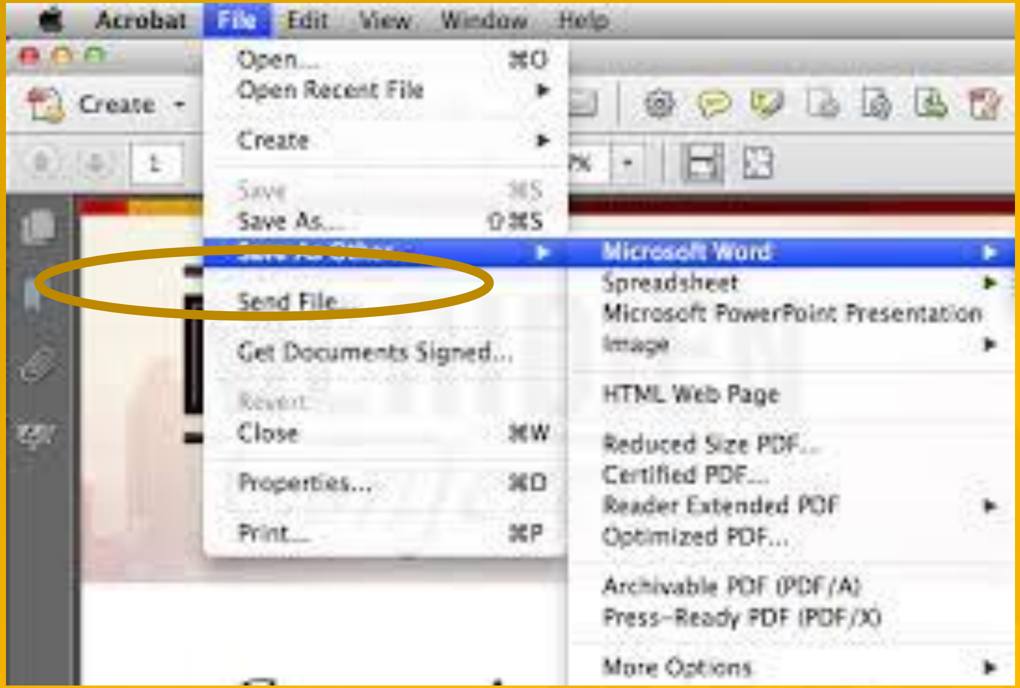
	A	B
1	0.	
2	78.7	
3	76.3	
4	81.1	
5	78.9	
6	76.6	
7	81.3	
8	75.5	
9	72.2	
10	78.5	
11	81.9	
12	79.3	
13	84.3	
14	78.7	
15	76.3	
16	81	
17	75.1	
18	71.9	
19	78.1	

What Next?

- Adobe Acrobat can export PDF to Excel;
- Easy as File, Save as Other → xlsx or csv
- Grey-out in Adobe Reader;
- Works only in Adobe Acrobat Pro 10.0;

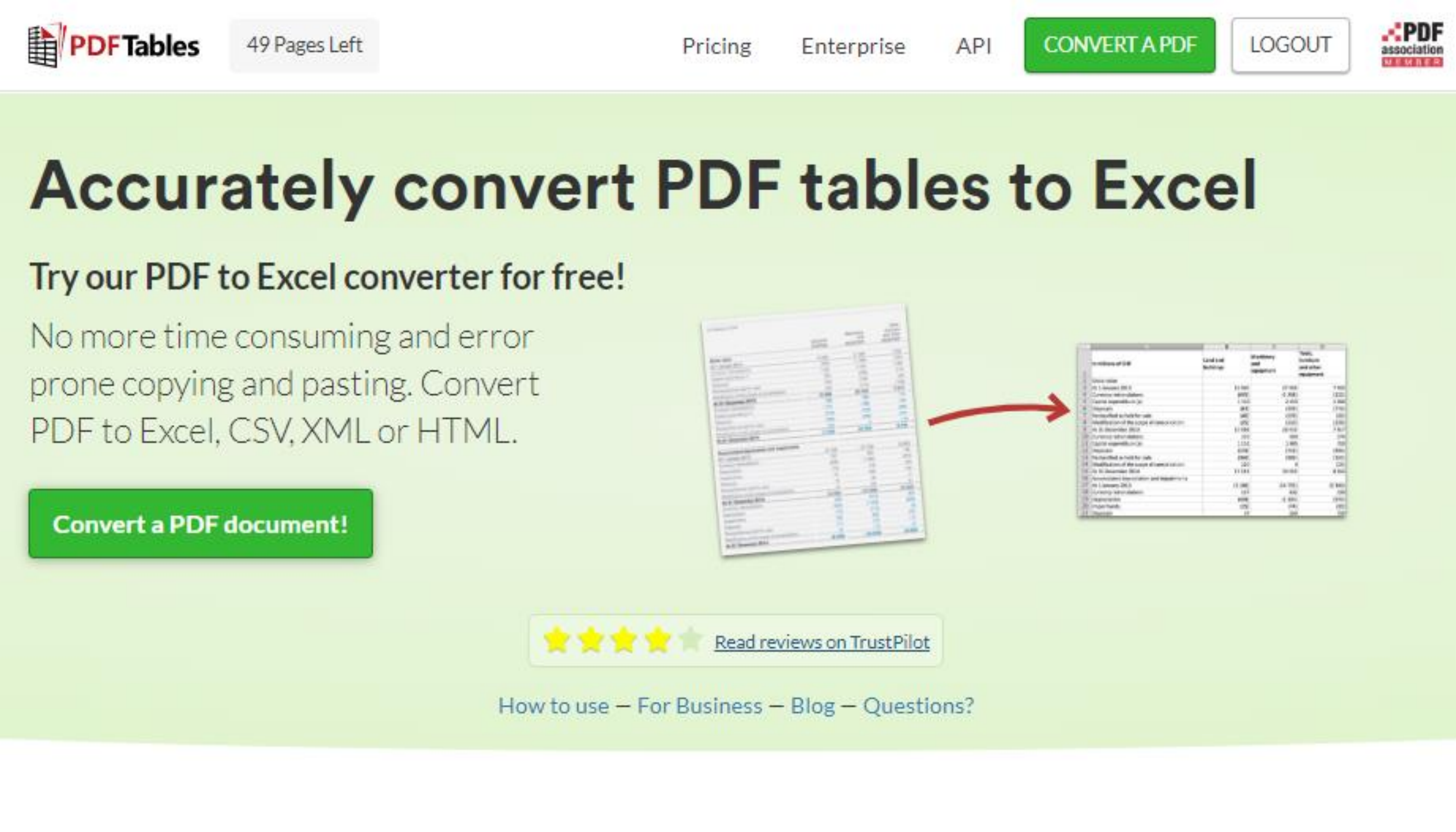


File > Save as Other > Spreadsheet



Can Anyone Export PDF to Excel?

- PDF Tables exports to Excel;
- Is it quick? YES;
- Is it easy? YES;
- Is it free? No…
- Well…Is it accurate?
- That depends…



Accurately convert PDF tables to Excel

Try our PDF to Excel converter for free!

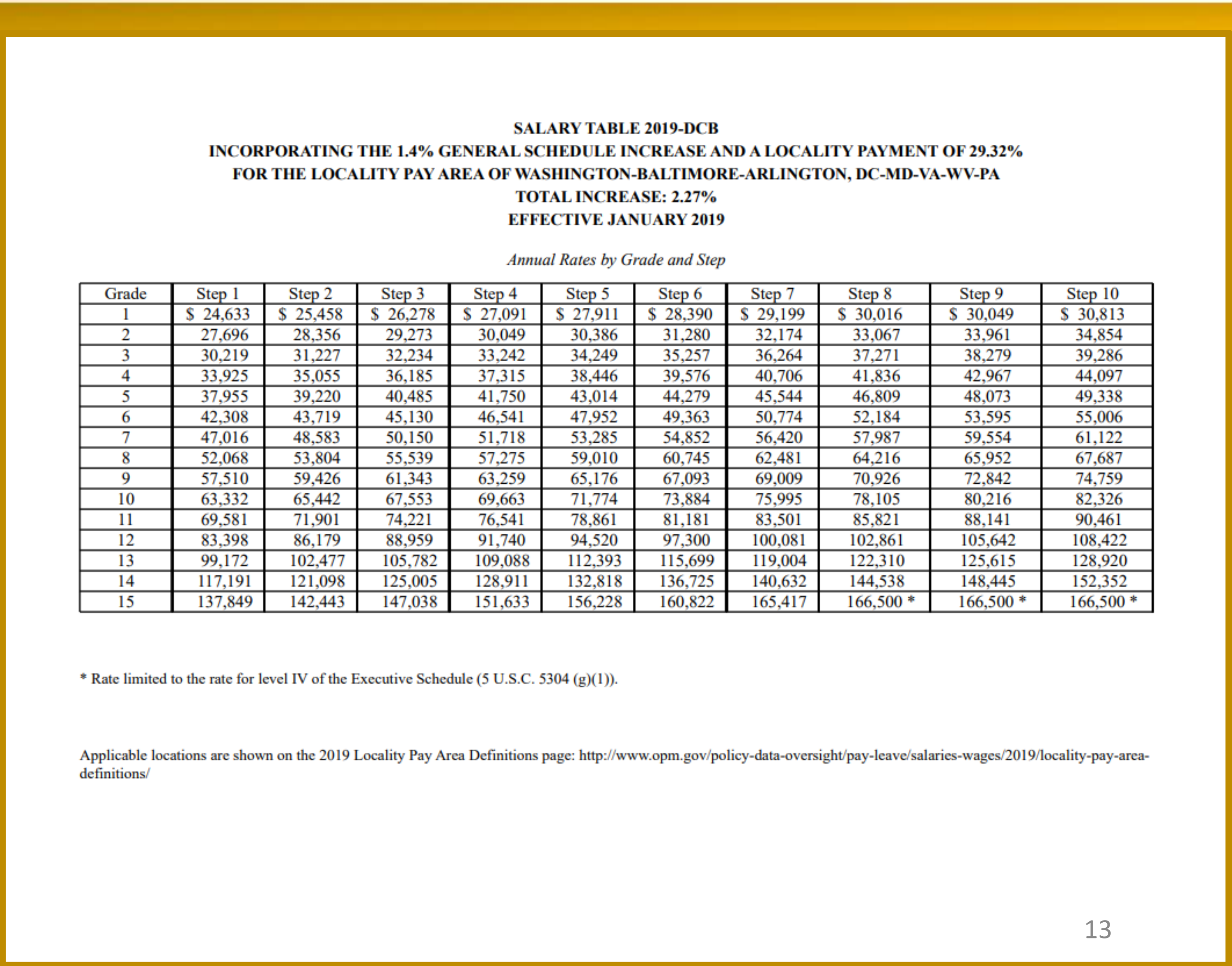
No more time consuming and error prone copying and pasting. Convert PDF to Excel, CSV, XML or HTML.

Convert a PDF document!
Read reviews on TrustPilot
How to use – For Business – Blog – Questions?

http://pdftables.com

PDF Tables – One page, One Table

- SALARY TABLE
- 2019-DCB
- One page document;
- One data table.



SALARY TABLE 2019-DCB
INCORPORATING THE 1.4% GENERAL SCHEDULE INCREASE AND A LOCALITY PAYMENT OF 29.32%
FOR THE LOCALITY PAY AREA OF WASHINGTON-BALTIMORE-ARLINGTON, DC-MD-VA-WV-PA
TOTAL INCREASE: 2.27%
EFFECTIVE JANUARY 2019

Annual Rates by Grade and Step

Grade	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10
1	$ 24,633	$ 25,458	$ 26,278	$ 27,091	$ 27,911	$ 28,390	$ 29,199	$ 30,016	$ 30,049	$ 30,813
2	27,696	28,356	29,273	30,049	30,386	31,280	32,174	33,067	33,961	34,854
3	30,219	31,227	32,234	33,242	34,249	35,257	36,264	37,271	38,279	39,286
4	33,925	35,055	36,185	37,315	38,446	39,576	40,706	41,836	42,967	44,097
5	37,955	39,220	40,485	41,750	43,014	44,279	45,544	46,809	48,073	49,338
6	42,308	43,719	45,130	46,541	47,952	49,363	50,774	52,184	53,595	55,006
7	47,016	48,583	50,150	51,718	53,285	54,852	56,420	57,987	59,554	61,122
8	52,068	53,804	55,539	57,275	59,010	60,745	62,481	64,216	65,952	67,687
9	57,510	59,426	61,343	63,259	65,176	67,093	69,009	70,926	72,842	74,759
10	63,332	65,442	67,553	69,663	71,774	73,884	75,995	78,105	80,216	82,326
11	69,581	71,901	74,221	76,541	78,861	81,181	83,501	85,821	88,141	90,461
12	83,398	86,179	88,959	91,740	94,520	97,300	100,081	102,861	105,642	108,422
13	99,172	102,477	105,782	109,088	112,393	115,699	119,004	122,310	125,615	128,920
14	117,191	121,098	125,005	128,911	132,818	136,725	140,632	144,538	148,445	152,352
15	137,849	142,443	147,038	151,633	156,228	160,822	165,417	166,500 *	166,500 *	166,500 *

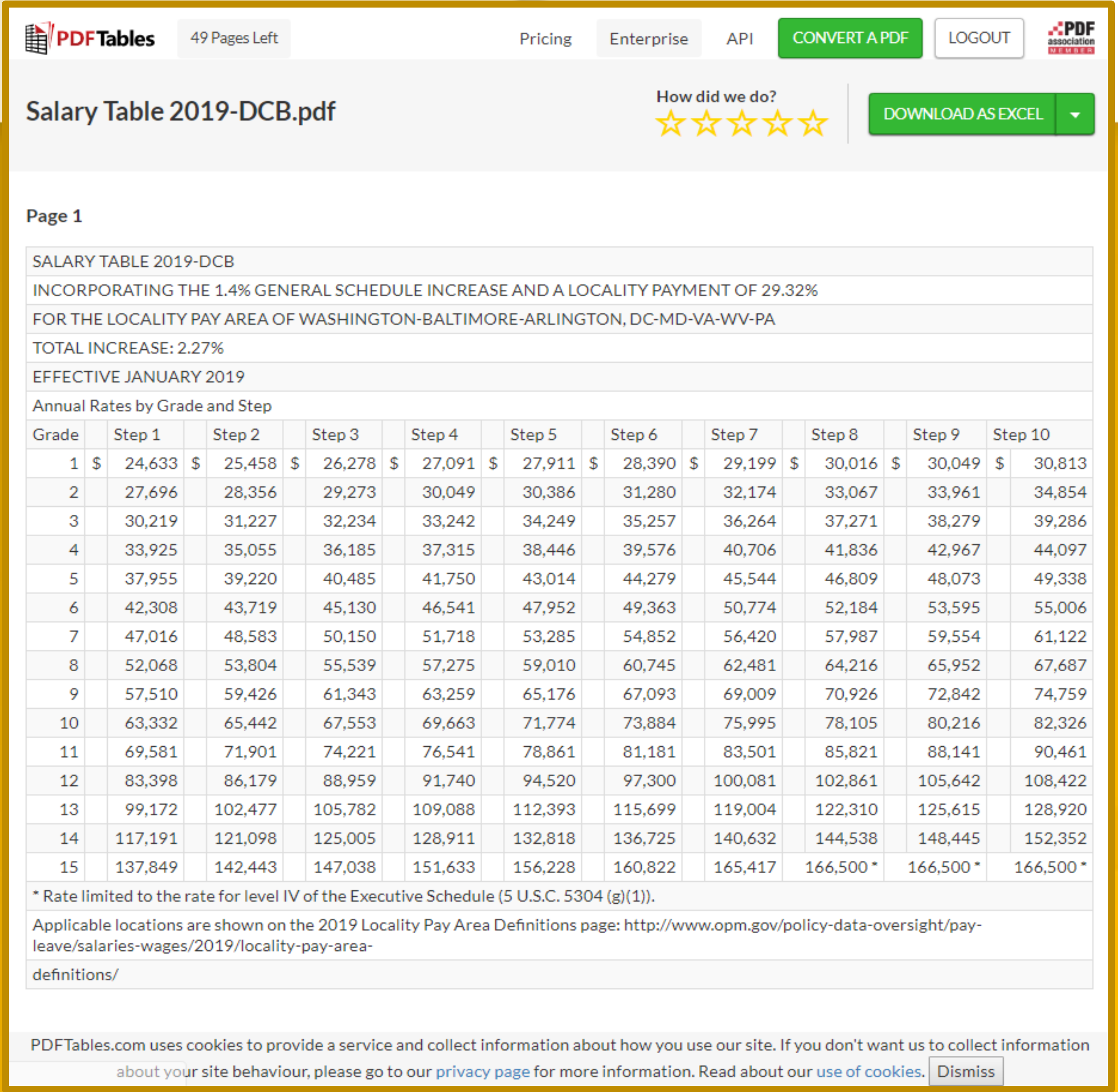
* Rate limited to the rate for level IV of the Executive Schedule (5 U.S.C. 5304 (g)(1)).

Applicable locations are shown on the 2019 Locality Pay Area Definitions page: http://www.opm.gov/policy-data-oversight/pay-leave/salaries-wages/2019/locality-pay-area-definitions/

PDF Tables – One Page, One Table

TOWSON UNIVERSITY

- SALARY TABLE
- 2019-DCB
- Imported to PDF Tables;
- Preview;
- Download to Excel.



Salary Table 2019-DCB.pdf

Page 1

SALARY TABLE 2019-DCB

INCORPORATING THE 1.4% GENERAL SCHEDULE INCREASE AND A LOCALITY PAYMENT OF 29.32%

FOR THE LOCALITY PAY AREA OF WASHINGTON-BALTIMORE-ARLINGTON, DC-MD-VA-WV-PA

TOTAL INCREASE: 2.27%

EFFECTIVE JANUARY 2019

Annual Rates by Grade and Step

Grade	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10
1	$ 24,633	$ 25,458	$ 26,278	$ 27,091	$ 27,911	$ 28,390	$ 29,199	$ 30,016	$ 30,049	$ 30,813
2	27,696	28,356	29,273	30,049	30,386	31,280	32,174	33,067	33,961	34,854
3	30,219	31,227	32,234	33,242	34,249	35,257	36,264	37,271	38,279	39,286
4	33,925	35,055	36,185	37,315	38,446	39,576	40,706	41,836	42,967	44,097
5	37,955	39,220	40,485	41,750	43,014	44,279	45,544	46,809	48,073	49,338
6	42,308	43,719	45,130	46,541	47,952	49,363	50,774	52,184	53,595	55,006
7	47,016	48,583	50,150	51,718	53,285	54,852	56,420	57,987	59,554	61,122
8	52,068	53,804	55,539	57,275	59,010	60,745	62,481	64,216	65,952	67,687
9	57,510	59,426	61,343	63,259	65,176	67,093	69,009	70,926	72,842	74,759
10	63,332	65,442	67,553	69,663	71,774	73,884	75,995	78,105	80,216	82,326
11	69,581	71,901	74,221	76,541	78,861	81,181	83,501	85,821	88,141	90,461
12	83,398	86,179	88,959	91,740	94,520	97,300	100,081	102,861	105,642	108,422
13	99,172	102,477	105,782	109,088	112,393	115,699	119,004	122,310	125,615	128,920
14	117,191	121,098	125,005	128,911	132,818	136,725	140,632	144,538	148,445	152,352
15	137,849	142,443	147,038	151,633	156,228	160,822	165,417	166,500 *	166,500 *	166,500 *

* Rate limited to the rate for level IV of the Executive Schedule (5 U.S.C. 5304 (g)(1)).

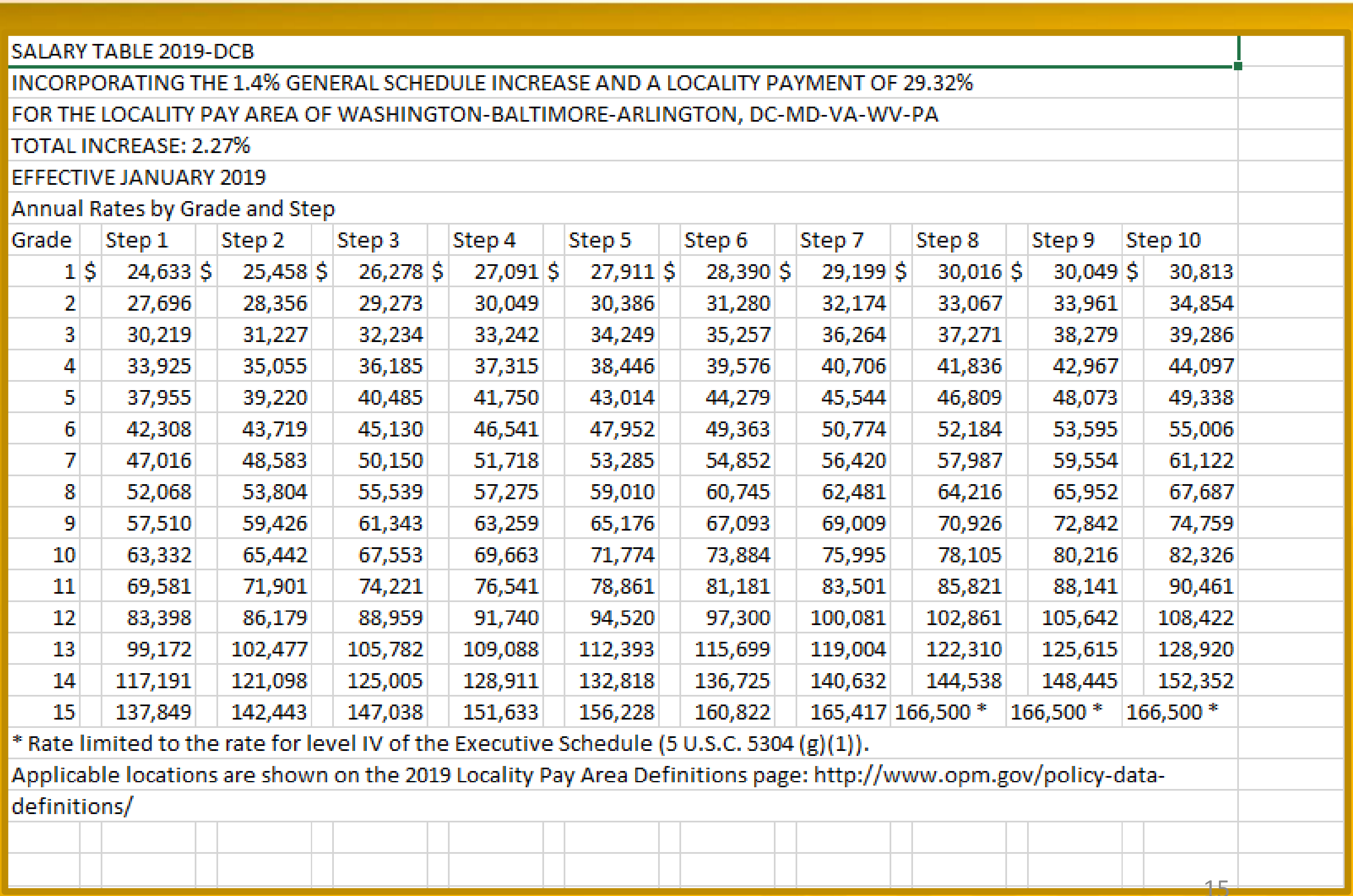
Applicable locations are shown on the 2019 Locality Pay Area Definitions page: http://www.opm.gov/policy-data-oversight/pay-leave/salaries-wages/2019/locality-pay-area-

definitions/

PDFTables.com uses cookies to provide a service and collect information about how you use our site. If you don't want us to collect information about your site behaviour, please go to our privacy page for more information. Read about our use of cookies. Dismiss

PDF Tables – One Page, One Table

- SALARY TABLE
- 2019-DCB
- Output to Excel;
- Very Pretty;
- Amenable to edits and analyses.



SALARY TABLE 2019-DCB

INCORPORATING THE 1.4% GENERAL SCHEDULE INCREASE AND A LOCALITY PAYMENT OF 29.32%

FOR THE LOCALITY PAY AREA OF WASHINGTON-BALTIMORE-ARLINGTON, DC-MD-VA-WV-PA

TOTAL INCREASE: 2.27%

EFFECTIVE JANUARY 2019

Annual Rates by Grade and Step

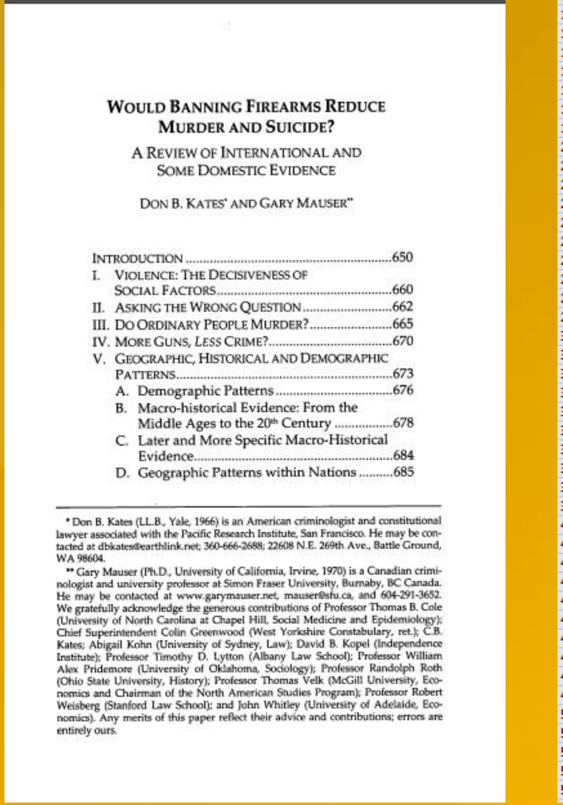
Grade	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10
1	$ 24,633	$ 25,458	$ 26,278	$ 27,091	$ 27,911	$ 28,390	$ 29,199	$ 30,016	$ 30,049	$ 30,813
2	27,696	28,356	29,273	30,049	30,386	31,280	32,174	33,067	33,961	34,854
3	30,219	31,227	32,234	33,242	34,249	35,257	36,264	37,271	38,279	39,286
4	33,925	35,055	36,185	37,315	38,446	39,576	40,706	41,836	42,967	44,097
5	37,955	39,220	40,485	41,750	43,014	44,279	45,544	46,809	48,073	49,338
6	42,308	43,719	45,130	46,541	47,952	49,363	50,774	52,184	53,595	55,006
7	47,016	48,583	50,150	51,718	53,285	54,852	56,420	57,987	59,554	61,122
8	52,068	53,804	55,539	57,275	59,010	60,745	62,481	64,216	65,952	67,687
9	57,510	59,426	61,343	63,259	65,176	67,093	69,009	70,926	72,842	74,759
10	63,332	65,442	67,553	69,663	71,774	73,884	75,995	78,105	80,216	82,326
11	69,581	71,901	74,221	76,541	78,861	81,181	83,501	85,821	88,141	90,461
12	83,398	86,179	88,959	91,740	94,520	97,300	100,081	102,861	105,642	108,422
13	99,172	102,477	105,782	109,088	112,393	115,699	119,004	122,310	125,615	128,920
14	117,191	121,098	125,005	128,911	132,818	136,725	140,632	144,538	148,445	152,352
15	137,849	142,443	147,038	151,633	156,228	160,822	165,417	166,500 *	166,500 *	166,500 *

* Rate limited to the rate for level IV of the Executive Schedule (5 U.S.C. 5304 (g)(1)).

Applicable locations are shown on the 2019 Locality Pay Area Definitions page: http://www.opm.gov/policy-data-definitions/

PDF Tables –Long Scholarly Article

A Little Awkward



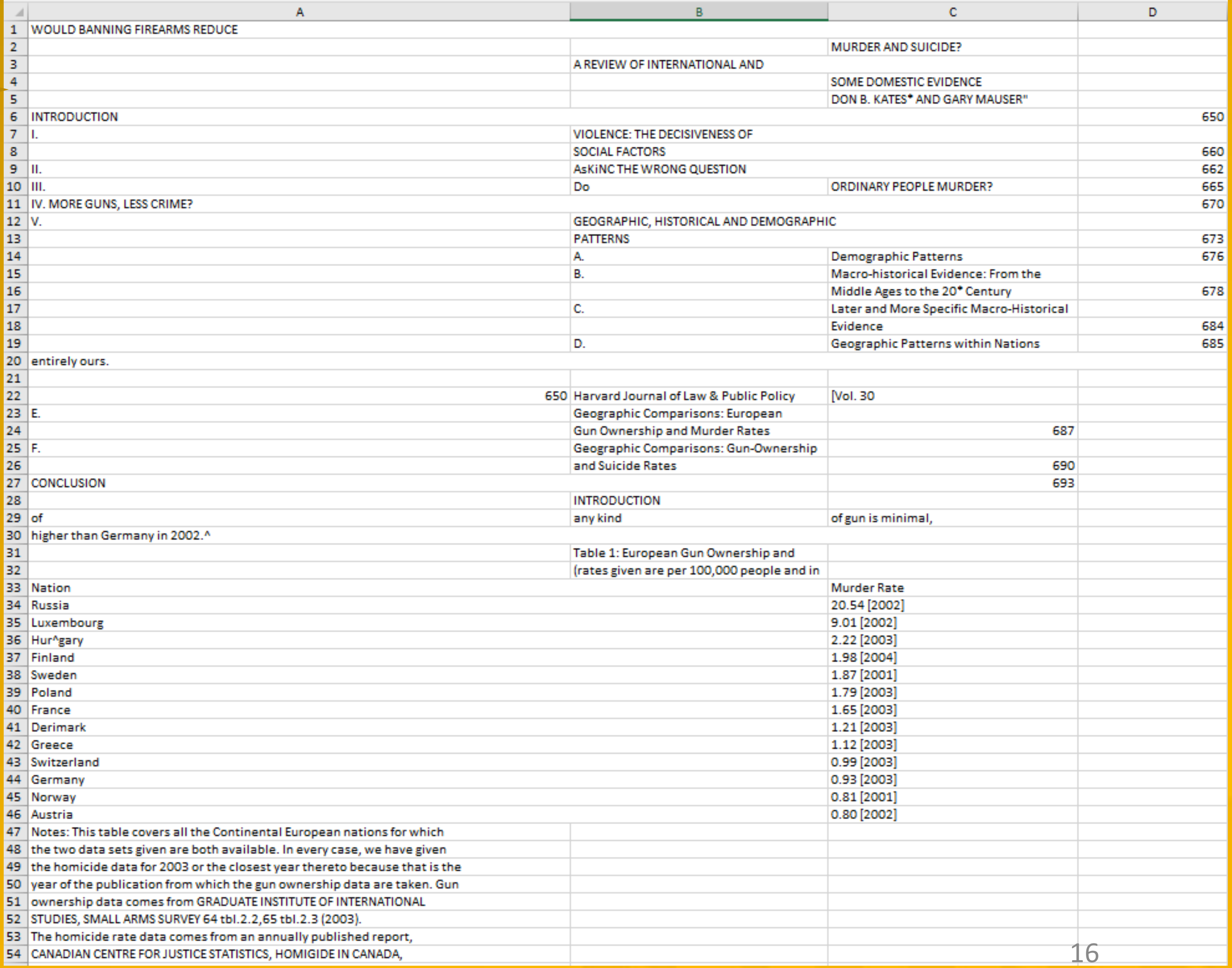
WOULD BANNING FIREARMS REDUCE MURDER AND SUICIDE?

A REVIEW OF INTERNATIONAL AND SOME DOMESTIC EVIDENCE

DON B. KATES* AND GARY MAUSER**

* Don B. Kates (LL.B., Yale, 1966) is an American criminologist and constitutional lawyer associated with the Pacific Research Institute, San Francisco. He may be contacted at dbkates@earthlink.net; 360-666-2688; 22608 N.E. 269th Ave., Battle Ground, WA 98604.

** Gary Mauser (Ph.D., University of California, Irvine, 1970) is a Canadian criminologist and university professor at Simon Fraser University, Burnaby, BC Canada. He may be contacted at www.garymauser.net, mauser@sfu.ca, and 604-291-3652. We gratefully acknowledge the generous contributions of Professor Thomas B. Cole (University of North Carolina at Chapel Hill, Social Medicine and Epidemiology); Chief Superintendent Colin Greenwood (West Yorkshire Constabulary, ret.); C.B. Kates; Abigail Kohn (University of Sydney, Law); David B. Kopel (Independence Institute); Professor Timothy D. Lytton (Albany Law School); Professor William Alex Pridemore (University of Oklahoma, Sociology); Professor Randolph Roth (Ohio State University, History); Professor Thomas Velk (McGill University, Economics and Chairman of the North American Studies Program); Professor Robert Weisberg (Stanford Law School); and John Whitley (University of Adelaide, Economics). Any merits of this paper reflect their advice and contributions; errors are entirely ours.

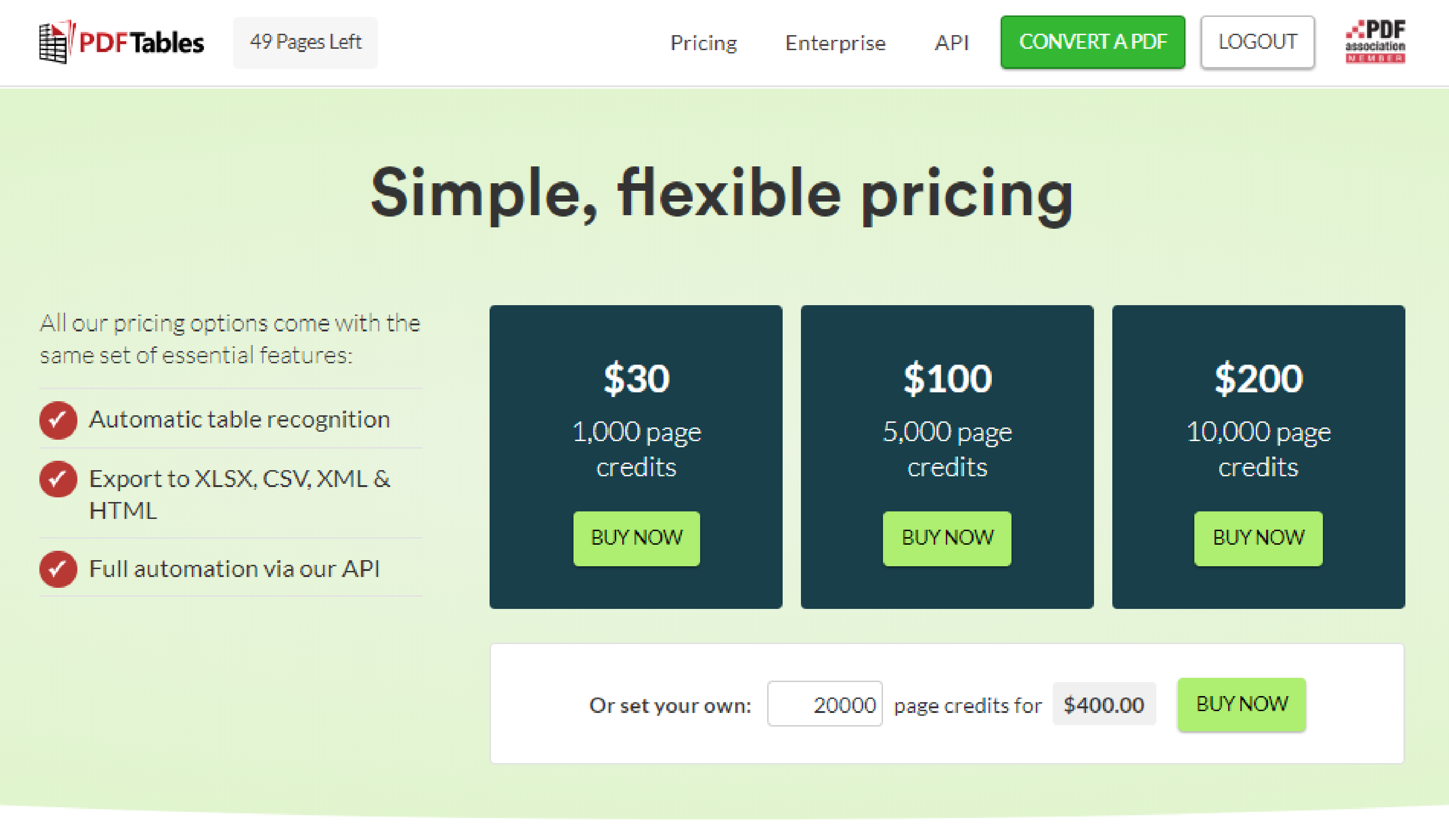


	A	B	C	D
1	WOULD BANNING FIREARMS REDUCE			
2			MURDER AND SUICIDE?	
3		A REVIEW OF INTERNATIONAL AND		
4			SOME DOMESTIC EVIDENCE	
5			DON B. KATES* AND GARY MAUSER"	
6	INTRODUCTION			650
7	I.	VIOLENCE: THE DECISIVENESS OF		
8		SOCIAL FACTORS		660
9	II.	AsKiNC THE WRONG QUESTION		662
10	III.	Do	ORDINARY PEOPLE MURDER?	665
11	IV. MORE GUNS, LESS CRIME?			670
12	V.	GEOGRAPHIC, HISTORICAL AND DEMOGRAPHIC		
13		PATTERNS		673
14		A.	Demographic Patterns	676
15		B.	Macro-historical Evidence: From the	
16			Middle Ages to the 20* Century	678
17		C.	Later and More Specific Macro-Historical	
18			Evidence	684
19		D.	Geographic Patterns within Nations	685
20	entirely ours.			
21				
22	650	Harvard Journal of Law & Public Policy	[Vol. 30	
23	E.	Geographic Comparisons: European		
24		Gun Ownership and Murder Rates	687	
25	F.	Geographic Comparisons: Gun-Ownership		
26		and Suicide Rates	690	
27	CONCLUSION		693	
28		INTRODUCTION		
29	of	any kind	of gun is minimal,	
30	higher than Germany in 2002.^			
31		Table 1: European Gun Ownership and		
32		(rates given are per 100,000 people and in		
33	Nation		Murder Rate	
34	Russia		20.54 [2002]	
35	Luxembourg		9.01 [2002]	
36	Hur^gary		2.22 [2003]	
37	Finland		1.98 [2004]	
38	Sweden		1.87 [2001]	
39	Poland		1.79 [2003]	
40	France		1.65 [2003]	
41	Derimark		1.21 [2003]	
42	Greece		1.12 [2003]	
43	Switzerland		0.99 [2003]	
44	Germany		0.93 [2003]	
45	Norway		0.81 [2001]	
46	Austria		0.80 [2002]	
47	Notes: This table covers all the Continental European nations for which			
48	the two data sets given are both available. In every case, we have given			
49	the homicide data for 2003 or the closest year thereto because that is the			
50	year of the publication from which the gun ownership data are taken. Gun			
51	ownership data comes from GRADUATE INSTITUTE OF INTERNATIONAL			
52	STUDIES, SMALL ARMS SURVEY 64 tbl.2.2,65 tbl.2.3 (2003).			
53	The homicide rate data comes from an annually published report,			
54	CANADIAN CENTRE FOR JUSTICE STATISTICS, HOMIGIDE IN CANADA,			

What does PDF Tables cost?

TOWSON UNIVERSITY

- Free test up to 50 pages;
- Free Registration;
- Free 50 pages;
- Buy Pages;
- PDF Pages…



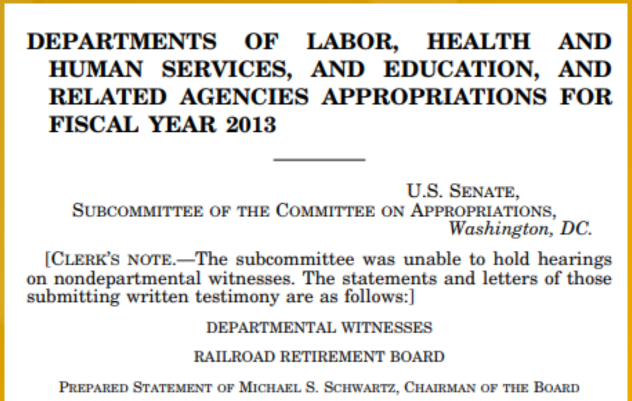
Simple, flexible pricing

All our pricing options come with the same set of essential features:

- ✓ Automatic table recognition
- ✓ Export to XLSX, CSV, XML & HTML
- ✓ Full automation via our API

$30	$100	$200
1,000 page credits	5,000 page credits	10,000 page credits
BUY NOW	BUY NOW	BUY NOW

Or set your own: 20000 page credits for $400.00 BUY NOW



DEPARTMENTS OF LABOR, HEALTH AND HUMAN SERVICES, AND EDUCATION, AND RELATED AGENCIES APPROPRIATIONS FOR FISCAL YEAR 2013

U.S. Senate,
Subcommittee of the Committee on Appropriations,
Washington, DC.

[Clerk's note.—The subcommittee was unable to hold hearings on nondepartmental witnesses. The statements and letters of those submitting written testimony are as follows:]

DEPARTMENTAL WITNESSES

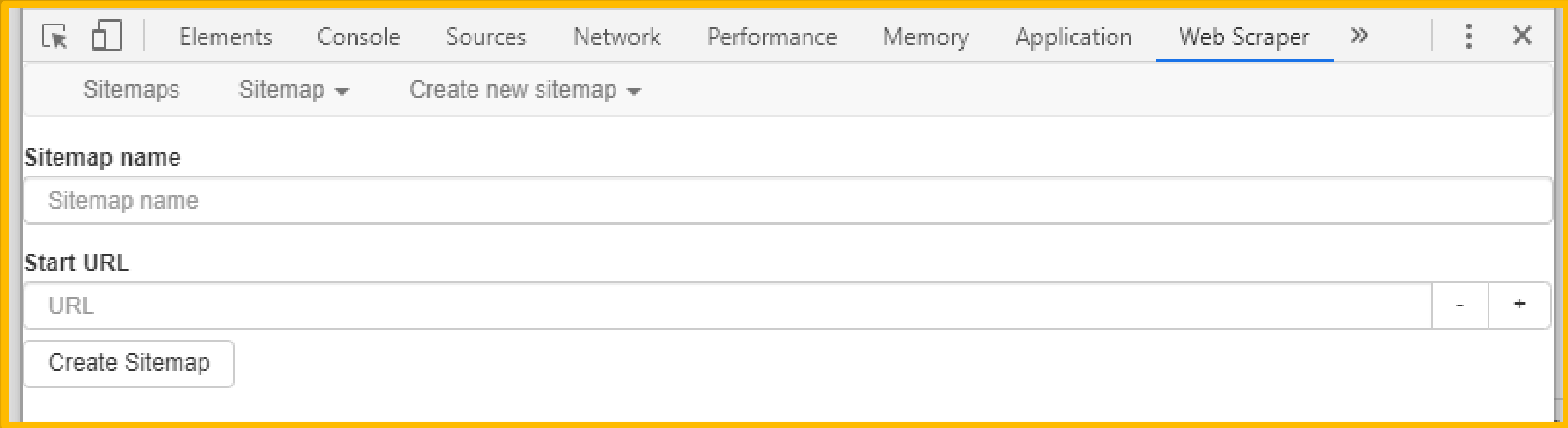
RAILROAD RETIREMENT BOARD

Prepared Statement of Michael S. Schwartz, Chairman of the Board

http://pdftables.com

Scraping for Tech Tigers

- Tiger teams scrape from dynamic pages (Census, Amazon, Monster Jobs, etc.);
- Such programs have a longer learning curve;



https://www.webscraper.io/

($) Scraping Tables ($)

- Best for high-volume harvesting;
- Choose the program :
  - safest; most reliable;
  - shortest learning curve;
  - best fit to workflow.

- **https://www.outwit.com/**
- **https://www.parsehub.com/**
- **http://www.visualscraper.com/**
- **http://scrapinghub.com/**
- **https://www.import.io/**
- **https://www.webhose.io/**
- **https://dexi.io/**
- **http://scrapinghub.com/**
- **http://www.spinn3r.com/**

Free Tools: Tabula Java to Browser

http://tabula.technology

Tabula scrapes PDF;

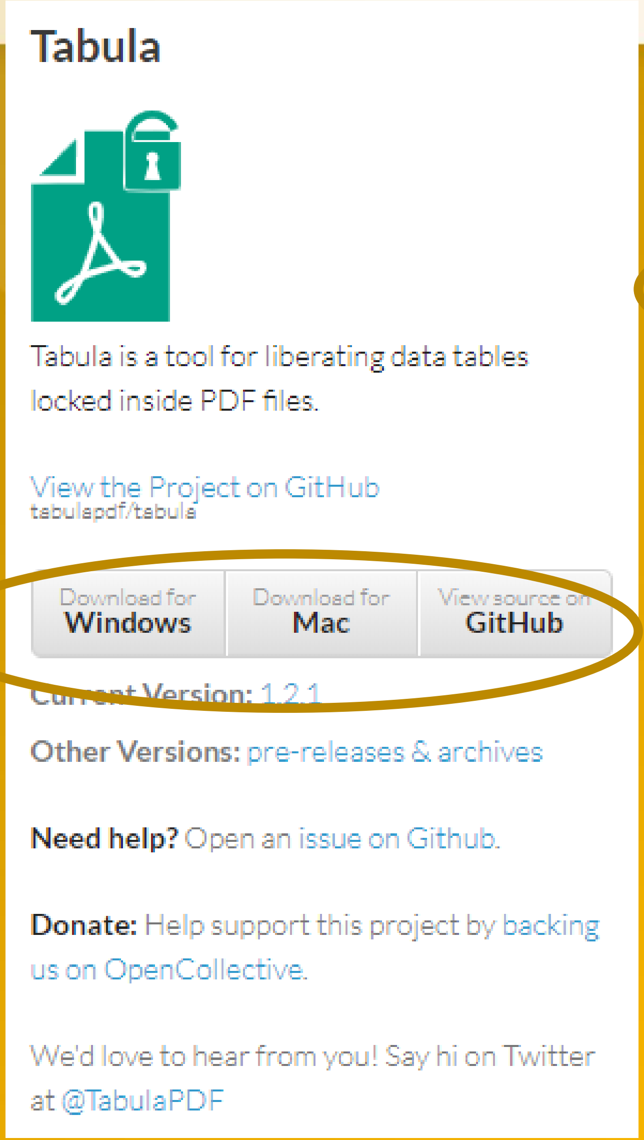
-User download;

- Update Java;

- Download/Install;

- Open tabula.exe

- Troubleshoot…



Tabula

Tabula is a tool for liberating data tables locked inside PDF files.

View the Project on GitHub
tabulapdf/tabula

Download for **Windows** | Download for **Mac** | View source on **GitHub**

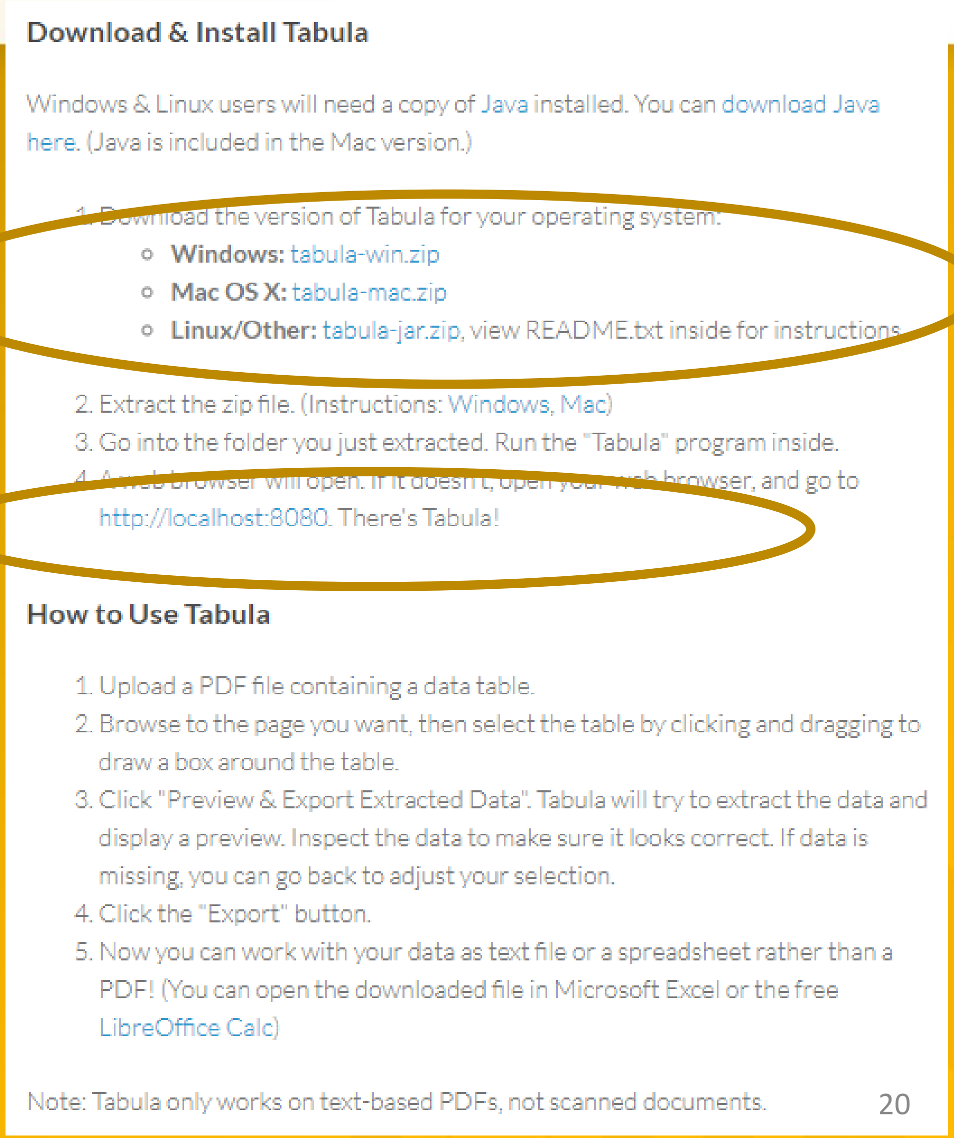
Current Version: 1.2.1

Other Versions: pre-releases & archives

Need help? Open an issue on Github.

Donate: Help support this project by backing us on OpenCollective.

We'd love to hear from you! Say hi on Twitter at @TabulaPDF



Download & Install Tabula

Windows & Linux users will need a copy of Java installed. You can download Java here. (Java is included in the Mac version.)

1. Download the version of Tabula for your operating system:
   - **Windows:** tabula-win.zip
   - **Mac OS X:** tabula-mac.zip
   - **Linux/Other:** tabula-jar.zip, view README.txt inside for instructions
2. Extract the zip file. (Instructions: Windows, Mac)
3. Go into the folder you just extracted. Run the "Tabula" program inside.
4. A web browser will open. If it doesn't, open your web browser, and go to http://localhost:8080. There's Tabula!

How to Use Tabula

1. Upload a PDF file containing a data table.
2. Browse to the page you want, then select the table by clicking and dragging to draw a box around the table.
3. Click "Preview & Export Extracted Data". Tabula will try to extract the data and display a preview. Inspect the data to make sure it looks correct. If data is missing, you can go back to adjust your selection.
4. Click the "Export" button.
5. Now you can work with your data as text file or a spreadsheet rather than a PDF! (You can open the downloaded file in Microsoft Excel or the free LibreOffice Calc)

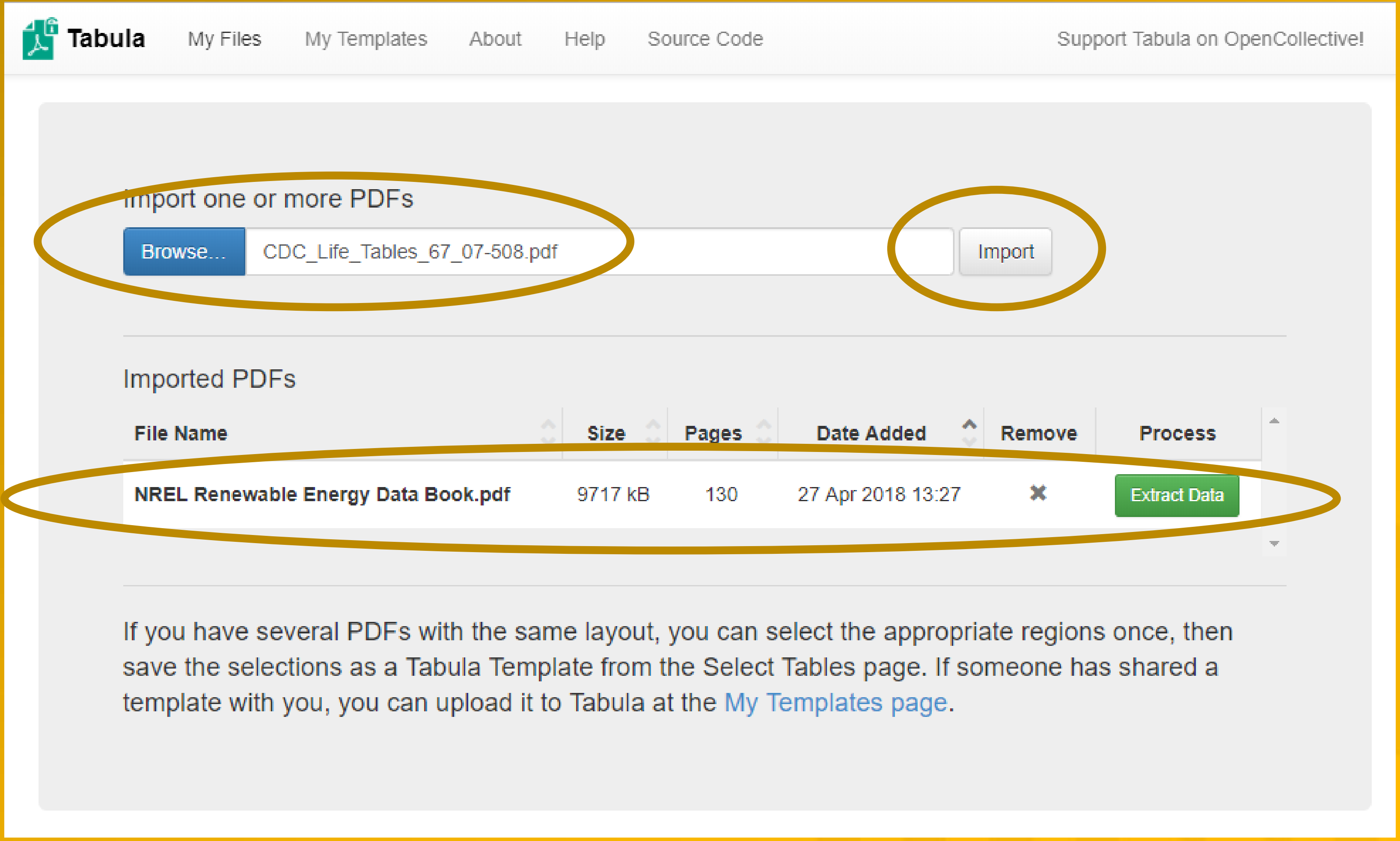
Note: Tabula only works on text-based PDFs, not scanned documents.

Tabula Java to Browser - http://tabula.technology

Save PDF to local drive

Tabula scrapes PDF

- Use for PDF saved to your computer;
- Will keep a tab on what has been imported;



Import one or more PDFs

CDC_Life_Tables_67_07-508.pdf

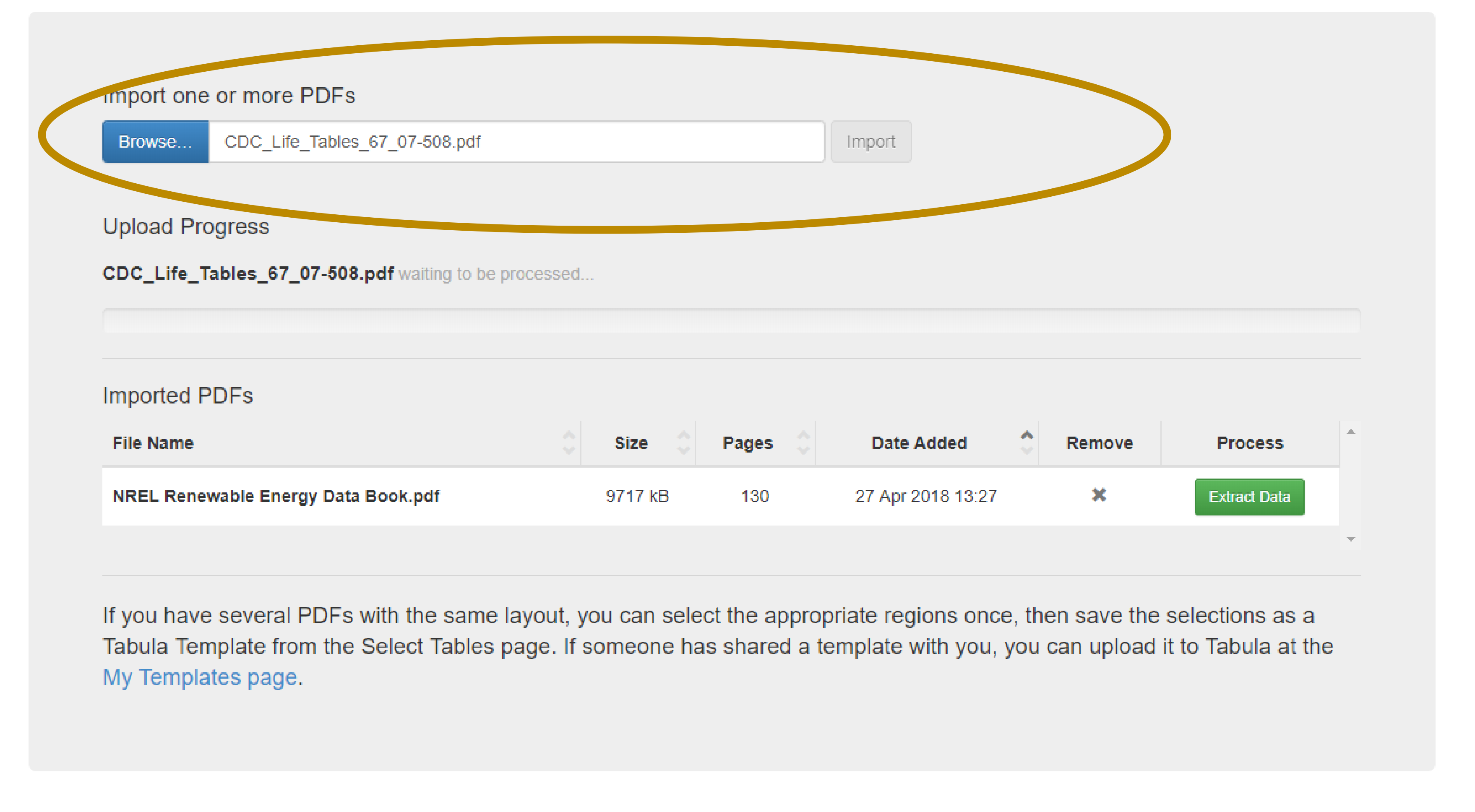
Imported PDFs

File Name	Size	Pages	Date Added	Remove	Process
NREL Renewable Energy Data Book.pdf	9717 kB	130	27 Apr 2018 13:27	✖	Extract Data

If you have several PDFs with the same layout, you can select the appropriate regions once, then save the selections as a Tabula Template from the Select Tables page. If someone has shared a template with you, you can upload it to Tabula at the My Templates page.

Tabula PDF to Excel http://tabula.technology

Once loaded, select Import



Import one or more PDFs

Upload Progress

CDC_Life_Tables_67_07-508.pdf waiting to be processed...

Imported PDFs

File Name	Size	Pages	Date Added	Remove	Process
NREL Renewable Energy Data Book.pdf	9717 kB	130	27 Apr 2018 13:27	✖	Extract Data

If you have several PDFs with the same layout, you can select the appropriate regions once, then save the selections as a Tabula Template from the Select Tables page. If someone has shared a template with you, you can upload it to Tabula at the My Templates page.

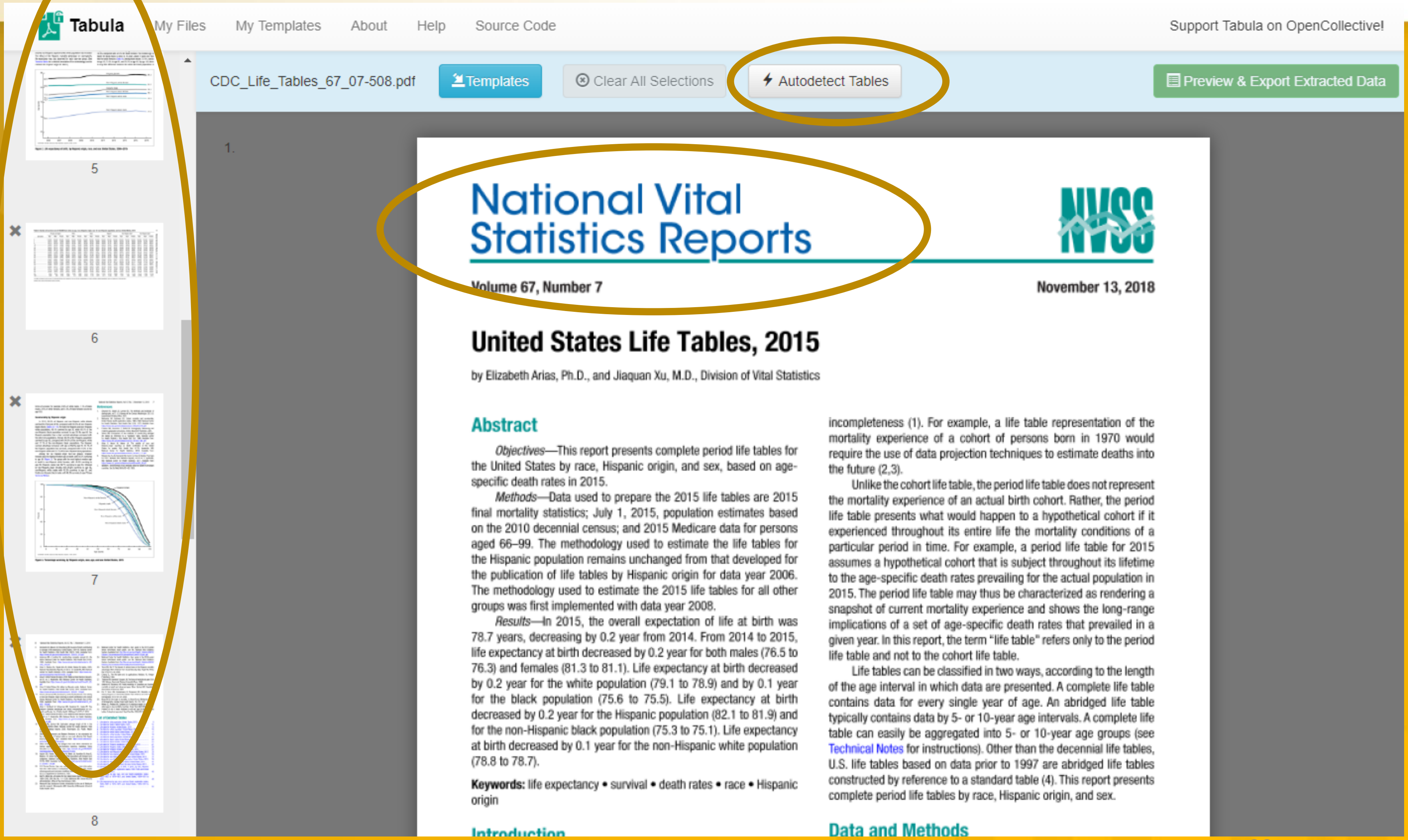
Tabula Java to Browser - http://tabula.technology

Tabula Improvements

- User download; opens in browser;
- Able to autodetect tables in a document;

Main limitation:

- Confused by formatting.



National Vital Statistics Reports

NVSS

Volume 67, Number 7 November 13, 2018

United States Life Tables, 2015

by Elizabeth Arias, Ph.D., and Jiaquan Xu, M.D., Division of Vital Statistics

Abstract

Objectives—This report presents complete period life tables for the United States by race, Hispanic origin, and sex, based on age-specific death rates in 2015.

Methods—Data used to prepare the 2015 life tables are 2015 final mortality statistics; July 1, 2015, population estimates based on the 2010 decennial census; and 2015 Medicare data for persons aged 66–99. The methodology used to estimate the life tables for the Hispanic population remains unchanged from that developed for the publication of life tables by Hispanic origin for data year 2006. The methodology used to estimate the 2015 life tables for all other groups was first implemented with data year 2008.

Results—In 2015, the overall expectation of life at birth was 78.7 years, decreasing by 0.2 year from 2014. From 2014 to 2015, life expectancy at birth decreased by 0.2 year for both males (76.5 to 76.3) and females (81.3 to 81.1). Life expectancy at birth decreased by 0.2 year for the white population (79.1 to 78.9) and by 0.1 year for the black population (75.6 to 75.5). Life expectancy at birth decreased by 0.2 year for the Hispanic population (82.1 to 81.9) and for the non-Hispanic black population (75.3 to 75.1). Life expectancy at birth decreased by 0.1 year for the non-Hispanic white population (78.8 to 78.7).

Keywords: life expectancy • survival • death rates • race • Hispanic origin

Introduction

incompleteness (1). For example, a life table representation of the mortality experience of a cohort of persons born in 1970 would require the use of data projection techniques to estimate deaths into the future (2,3).

Unlike the cohort life table, the period life table does not represent the mortality experience of an actual birth cohort. Rather, the period life table presents what would happen to a hypothetical cohort if it experienced throughout its entire life the mortality conditions of a particular period in time. For example, a period life table for 2015 assumes a hypothetical cohort that is subject throughout its lifetime to the age-specific death rates prevailing for the actual population in 2015. The period life table may thus be characterized as rendering a snapshot of current mortality experience and shows the long-range implications of a set of age-specific death rates that prevailed in a given year. In this report, the term "life table" refers only to the period life table and not to the cohort life table.

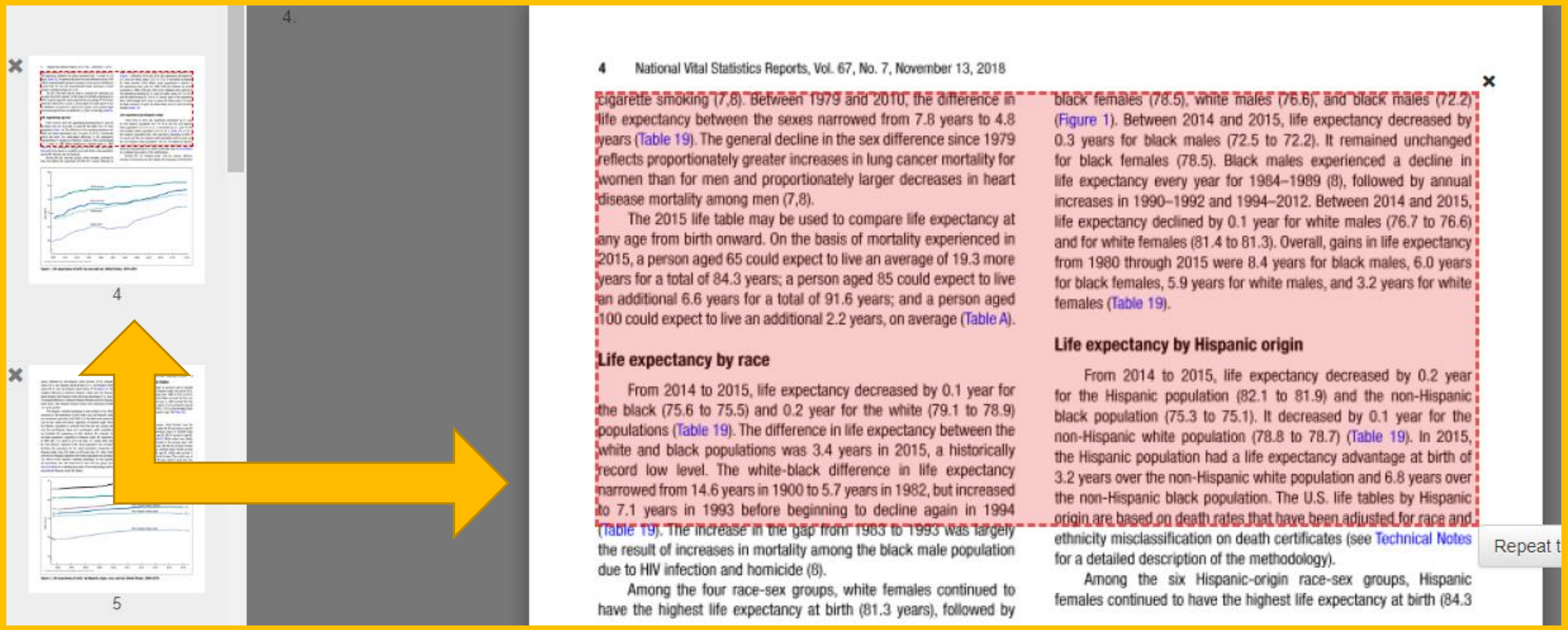
Life tables can be classified in two ways, according to the length of the age interval in which data are presented. A complete life table contains data for every single year of age. An abridged life table typically contains data by 5- or 10-year age intervals. A complete life table can easily be aggregated into 5- or 10-year age groups (see Technical Notes for instructions). Other than the decennial life tables, U.S. life tables based on data prior to 1997 are abridged life tables constructed by reference to a standard table (4). This report presents complete period life tables by race, Hispanic origin, and sex.

Data and Methods

Tabula Autoselect Tables

Autodetect Tables:

- Searches and high-lights what it believes is tabular data;
- Confused by format;
- **X** in upper right removes unwanted elements.



4 National Vital Statistics Reports, Vol. 67, No. 7, November 13, 2018

cigarette smoking (7,8). Between 1979 and 2010, the difference in life expectancy between the sexes narrowed from 7.8 years to 4.8 years (Table 19). The general decline in the sex difference since 1979 reflects proportionately greater increases in lung cancer mortality for women than for men and proportionately larger decreases in heart disease mortality among men (7,8).

The 2015 life table may be used to compare life expectancy at any age from birth onward. On the basis of mortality experienced in 2015, a person aged 65 could expect to live an average of 19.3 more years for a total of 84.3 years; a person aged 85 could expect to live an additional 6.6 years for a total of 91.6 years; and a person aged 100 could expect to live an additional 2.2 years, on average (Table A).

Life expectancy by race

From 2014 to 2015, life expectancy decreased by 0.1 year for the black (75.6 to 75.5) and 0.2 year for the white (79.1 to 78.9) populations (Table 19). The difference in life expectancy between the white and black populations was 3.4 years in 2015, a historically record low level. The white-black difference in life expectancy narrowed from 14.6 years in 1900 to 5.7 years in 1982, but increased to 7.1 years in 1993 before beginning to decline again in 1994 (Table 19). The increase in the gap from 1983 to 1993 was largely the result of increases in mortality among the black male population due to HIV infection and homicide (8).

Among the four race-sex groups, white females continued to have the highest life expectancy at birth (81.3 years), followed by black females (78.5), white males (76.6), and black males (72.2) (Figure 1). Between 2014 and 2015, life expectancy decreased by 0.3 years for black males (72.5 to 72.2). It remained unchanged for black females (78.5). Black males experienced a decline in life expectancy every year for 1984–1989 (8), followed by annual increases in 1990–1992 and 1994–2012. Between 2014 and 2015, life expectancy declined by 0.1 year for white males (76.7 to 76.6) and for white females (81.4 to 81.3). Overall, gains in life expectancy from 1980 through 2015 were 8.4 years for black males, 6.0 years for black females, 5.9 years for white males, and 3.2 years for white females (Table 19).

Life expectancy by Hispanic origin

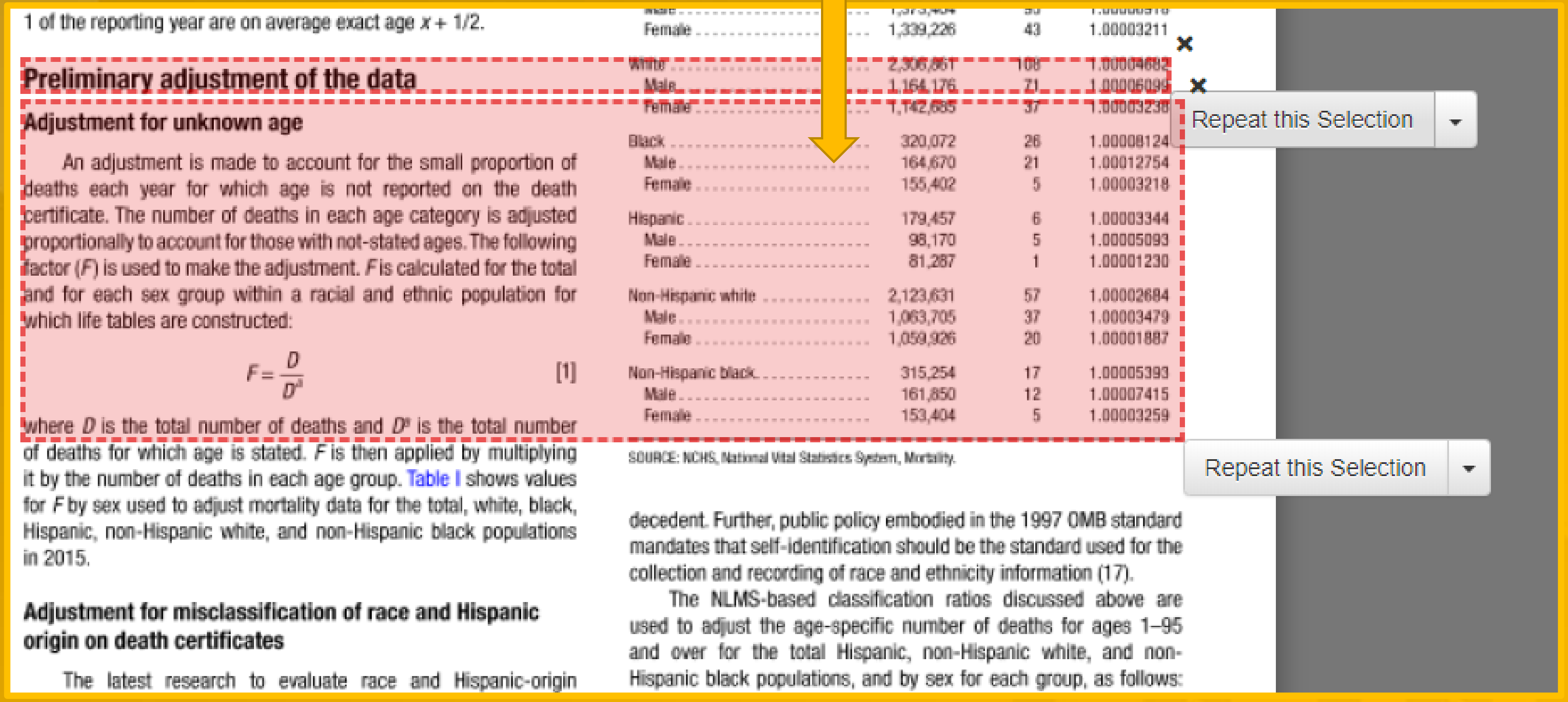
From 2014 to 2015, life expectancy decreased by 0.2 year for the Hispanic population (82.1 to 81.9) and the non-Hispanic black population (75.3 to 75.1). It decreased by 0.1 year for the non-Hispanic white population (78.8 to 78.7) (Table 19). In 2015, the Hispanic population had a life expectancy advantage at birth of 3.2 years over the non-Hispanic white population and 6.8 years over the non-Hispanic black population. The U.S. life tables by Hispanic origin are based on death rates that have been adjusted for race and ethnicity misclassification on death certificates (see Technical Notes for a detailed description of the methodology).

Among the six Hispanic-origin race-sex groups, Hispanic females continued to have the highest life expectancy at birth (84.3

Tabula Autodetect Tables

Autodetect Tables:

- Searches and high-lights what it believes is tabular data;
- OK to shape tables.



1 of the reporting year are on average exact age $x + 1/2$.

Preliminary adjustment of the data

Adjustment for unknown age

An adjustment is made to account for the small proportion of deaths each year for which age is not reported on the death certificate. The number of deaths in each age category is adjusted proportionally to account for those with not-stated ages. The following factor (*F*) is used to make the adjustment. *F* is calculated for the total and for each sex group within a racial and ethnic population for which life tables are constructed:

$$F = \frac{D}{D^a} \qquad [1]$$

where *D* is the total number of deaths and D^a is the total number of deaths for which age is stated. *F* is then applied by multiplying it by the number of deaths in each age group. Table I shows values for *F* by sex used to adjust mortality data for the total, white, black, Hispanic, non-Hispanic white, and non-Hispanic black populations in 2015.

Adjustment for misclassification of race and Hispanic origin on death certificates

The latest research to evaluate race and Hispanic-origin

Male	[illegible]	[illegible]	[illegible]
Female	1,339,226	43	1.00003211
White	2,306,861	108	1.00004682
Male	1,164,176	71	1.00006099
Female	1,142,685	37	1.00003238
Black	320,072	26	1.00008124
Male	164,670	21	1.00012754
Female	155,402	5	1.00003218
Hispanic	179,457	6	1.00003344
Male	98,170	5	1.00005093
Female	81,287	1	1.00001230
Non-Hispanic white	2,123,631	57	1.00002684
Male	1,063,705	37	1.00003479
Female	1,059,926	20	1.00001887
Non-Hispanic black	315,254	17	1.00005393
Male	161,850	12	1.00007415
Female	153,404	5	1.00003259

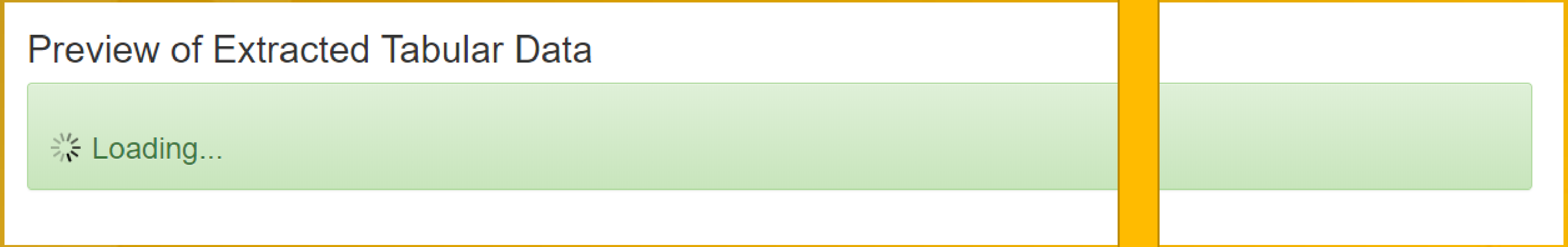
SOURCE: NCHS, National Vital Statistics System, Mortality.

decedent. Further, public policy embodied in the 1997 OMB standard mandates that self-identification should be the standard used for the collection and recording of race and ethnicity information (17).

The NLMS-based classification ratios discussed above are used to adjust the age-specific number of deaths for ages 1–95 and over for the total Hispanic, non-Hispanic white, and non-Hispanic black populations, and by sex for each group, as follows:

Tabula Export to Excel

Preview & Export Extracted Data



Preview of Extracted Tabular Data

Loading...

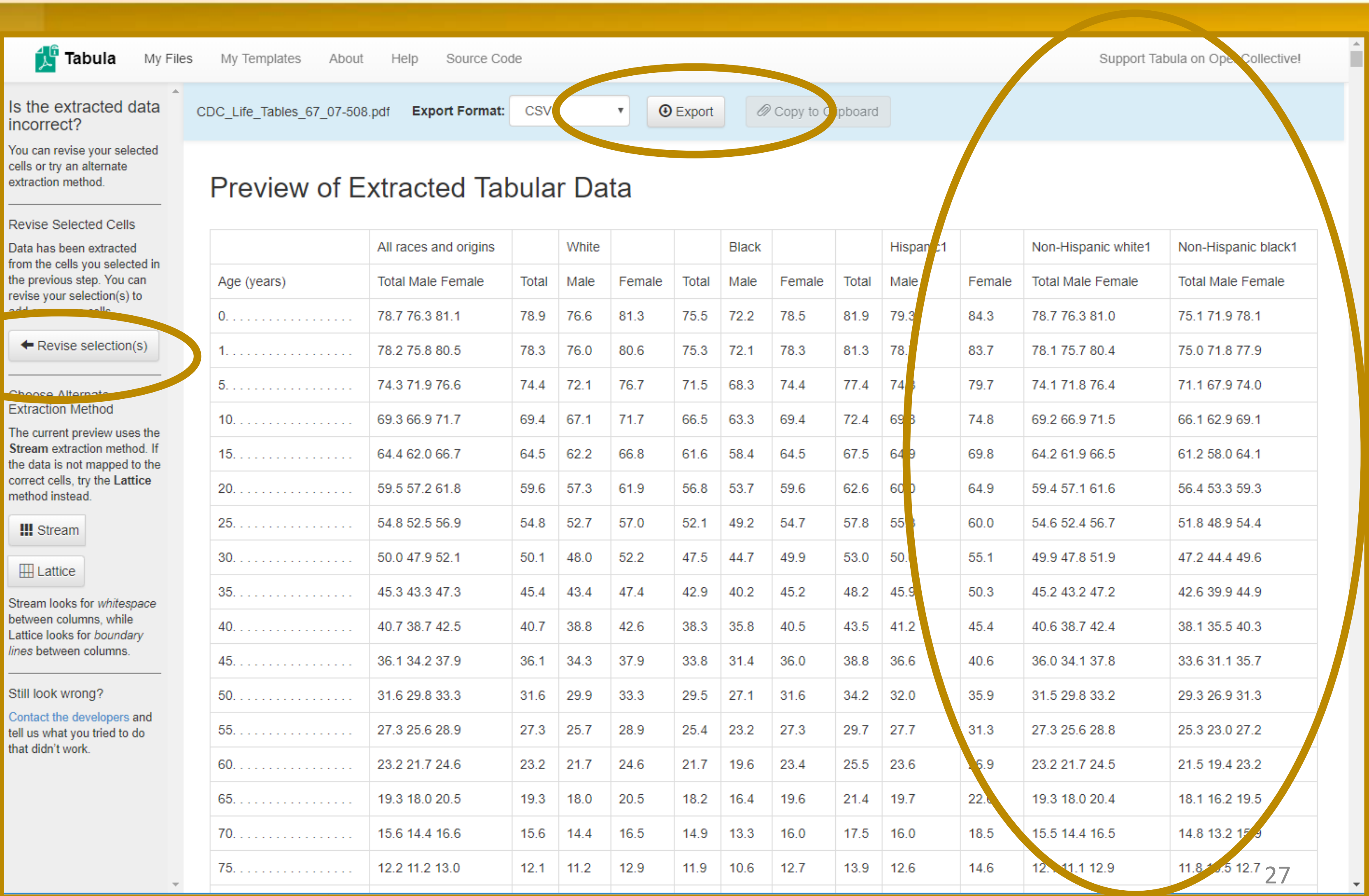
Tabula Java Program

TOWSON UNIVERSITY

Revise Selections to go back and adjust;

Note the fused tables;

Export Extracted Data



Is the extracted data incorrect?

You can revise your selected cells or try an alternate extraction method.

Revise Selected Cells

Data has been extracted from the cells you selected in the previous step. You can revise your selection(s) to add ... cells.

Choose Alternate Extraction Method

The current preview uses the **Stream** extraction method. If the data is not mapped to the correct cells, try the **Lattice** method instead.

Stream looks for *whitespace* between columns, while Lattice looks for *boundary lines* between columns.

Still look wrong?

Contact the developers and tell us what you tried to do that didn't work.

CDC_Life_Tables_67_07-508.pdf Export Format: CSV

Preview of Extracted Tabular Data

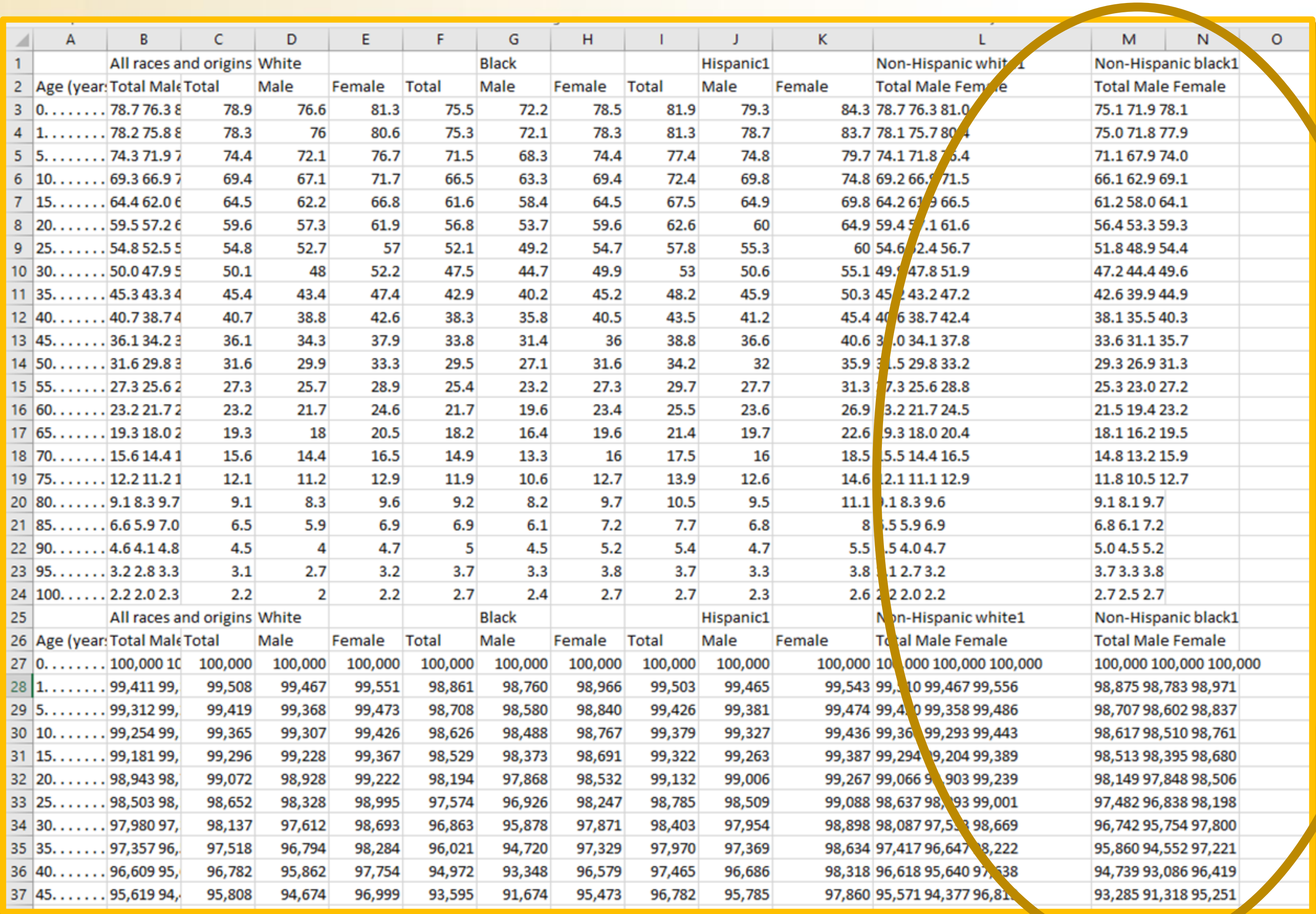
	All races and origins		White			Black			Hispanic1		Non-Hispanic white1	Non-Hispanic black1
Age (years)	Total Male Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total Male Female	Total Male Female
0.	78.7 76.3 81.1	78.9	76.6	81.3	75.5	72.2	78.5	81.9	79.3	84.3	78.7 76.3 81.0	75.1 71.9 78.1
1.	78.2 75.8 80.5	78.3	76.0	80.6	75.3	72.1	78.3	81.3	78.[illegible]	83.7	78.1 75.7 80.4	75.0 71.8 77.9
5.	74.3 71.9 76.6	74.4	72.1	76.7	71.5	68.3	74.4	77.4	74.8	79.7	74.1 71.8 76.4	71.1 67.9 74.0
10.	69.3 66.9 71.7	69.4	67.1	71.7	66.5	63.3	69.4	72.4	69.8	74.8	69.2 66.9 71.5	66.1 62.9 69.1
15.	64.4 62.0 66.7	64.5	62.2	66.8	61.6	58.4	64.5	67.5	64.9	69.8	64.2 61.9 66.5	61.2 58.0 64.1
20.	59.5 57.2 61.8	59.6	57.3	61.9	56.8	53.7	59.6	62.6	60.0	64.9	59.4 57.1 61.6	56.4 53.3 59.3
25.	54.8 52.5 56.9	54.8	52.7	57.0	52.1	49.2	54.7	57.8	55.8	60.0	54.6 52.4 56.7	51.8 48.9 54.4
30.	50.0 47.9 52.1	50.1	48.0	52.2	47.5	44.7	49.9	53.0	50.[illegible]	55.1	49.9 47.8 51.9	47.2 44.4 49.6
35.	45.3 43.3 47.3	45.4	43.4	47.4	42.9	40.2	45.2	48.2	45.9	50.3	45.2 43.2 47.2	42.6 39.9 44.9
40.	40.7 38.7 42.5	40.7	38.8	42.6	38.3	35.8	40.5	43.5	41.2	45.4	40.6 38.7 42.4	38.1 35.5 40.3
45.	36.1 34.2 37.9	36.1	34.3	37.9	33.8	31.4	36.0	38.8	36.6	40.6	36.0 34.1 37.8	33.6 31.1 35.7
50.	31.6 29.8 33.3	31.6	29.9	33.3	29.5	27.1	31.6	34.2	32.0	35.9	31.5 29.8 33.2	29.3 26.9 31.3
55.	27.3 25.6 28.9	27.3	25.7	28.9	25.4	23.2	27.3	29.7	27.7	31.3	27.3 25.6 28.8	25.3 23.0 27.2
60.	23.2 21.7 24.6	23.2	21.7	24.6	21.7	19.6	23.4	25.5	23.6	26.9	23.2 21.7 24.5	21.5 19.4 23.2
65.	19.3 18.0 20.5	19.3	18.0	20.5	18.2	16.4	19.6	21.4	19.7	22.[illegible]	19.3 18.0 20.4	18.1 16.2 19.5
70.	15.6 14.4 16.6	15.6	14.4	16.5	14.9	13.3	16.0	17.5	16.0	18.5	15.5 14.4 16.5	14.8 13.2 15.9
75.	12.2 11.2 13.0	12.1	11.2	12.9	11.9	10.6	12.7	13.9	12.6	14.6	12.[illegible] 11.1 12.9	11.8 10.5 12.7

Tabula Java Program

Tabula confused by merged cells;

All tables are on one sheet in a pile;

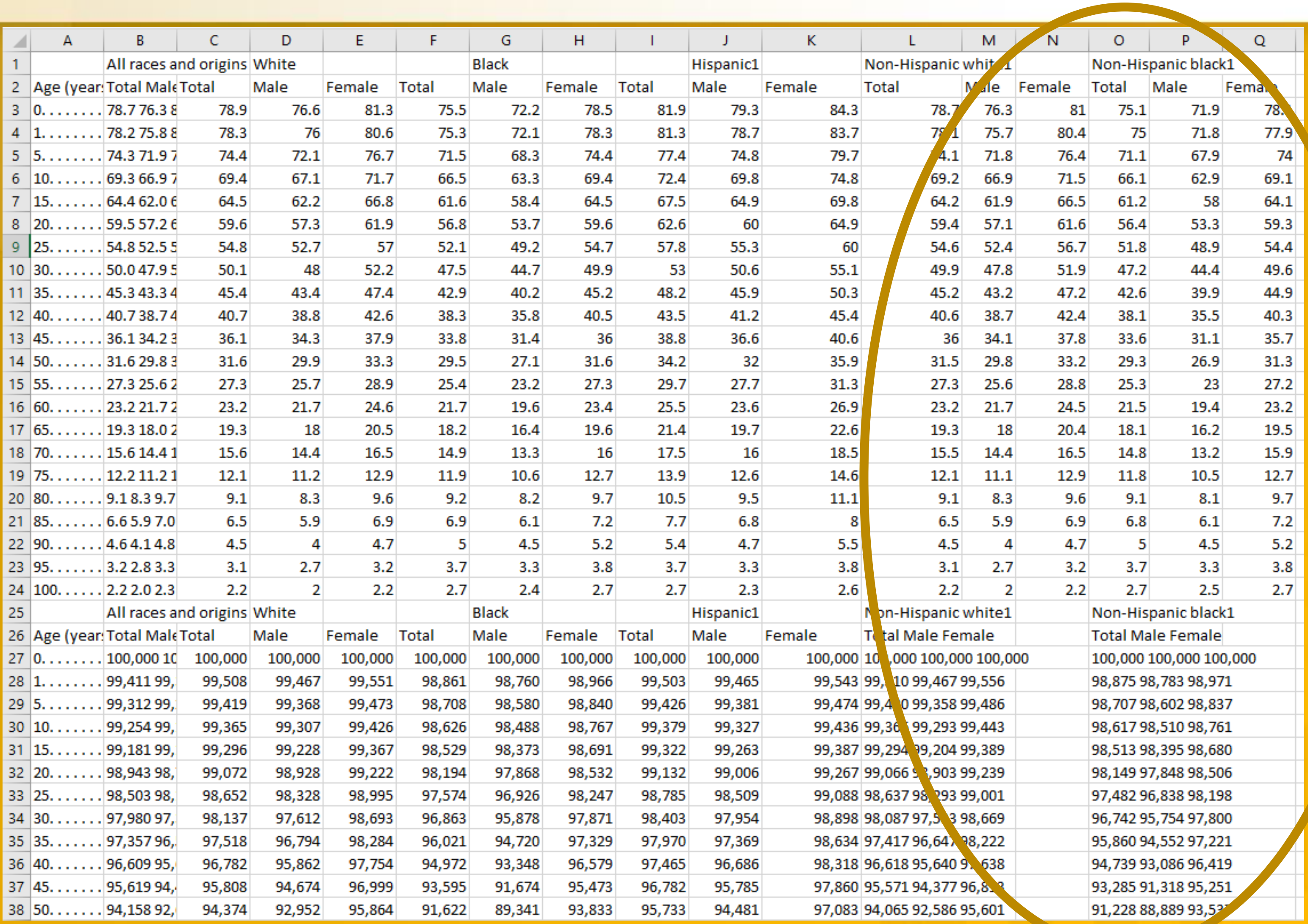
Far columns fused;



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		All races and origins		White			Black			Hispanic1		Non-Hispanic white1	Non-Hispanic black1		
2	Age (year:	Total Male	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total Male Female	Total Male Female		
3	0.	78.7 76.3 8	78.9	76.6	81.3	75.5	72.2	78.5	81.9	79.3	84.3	78.7 76.3 81.0	75.1 71.9 78.1		
4	1.	78.2 75.8 8	78.3	76	80.6	75.3	72.1	78.3	81.3	78.7	83.7	78.1 75.7 80.4	75.0 71.8 77.9		
5	5.	74.3 71.9 7	74.4	72.1	76.7	71.5	68.3	74.4	77.4	74.8	79.7	74.1 71.8 76.4	71.1 67.9 74.0		
6	10.	69.3 66.9 7	69.4	67.1	71.7	66.5	63.3	69.4	72.4	69.8	74.8	69.2 66.9 71.5	66.1 62.9 69.1		
7	15.	64.4 62.0 6	64.5	62.2	66.8	61.6	58.4	64.5	67.5	64.9	69.8	64.2 61.9 66.5	61.2 58.0 64.1		
8	20.	59.5 57.2 6	59.6	57.3	61.9	56.8	53.7	59.6	62.6	60	64.9	59.4 57.1 61.6	56.4 53.3 59.3		
9	25.	54.8 52.5 5	54.8	52.7	57	52.1	49.2	54.7	57.8	55.3	60	54.6 52.4 56.7	51.8 48.9 54.4		
10	30.	50.0 47.9 5	50.1	48	52.2	47.5	44.7	49.9	53	50.6	55.1	49.9 47.8 51.9	47.2 44.4 49.6		
11	35.	45.3 43.3 4	45.4	43.4	47.4	42.9	40.2	45.2	48.2	45.9	50.3	45.2 43.2 47.2	42.6 39.9 44.9		
12	40.	40.7 38.7 4	40.7	38.8	42.6	38.3	35.8	40.5	43.5	41.2	45.4	40.6 38.7 42.4	38.1 35.5 40.3		
13	45.	36.1 34.2 3	36.1	34.3	37.9	33.8	31.4	36	38.8	36.6	40.6	36.0 34.1 37.8	33.6 31.1 35.7		
14	50.	31.6 29.8 3	31.6	29.9	33.3	29.5	27.1	31.6	34.2	32	35.9	31.5 29.8 33.2	29.3 26.9 31.3		
15	55.	27.3 25.6 2	27.3	25.7	28.9	25.4	23.2	27.3	29.7	27.7	31.3	27.3 25.6 28.8	25.3 23.0 27.2		
16	60.	23.2 21.7 2	23.2	21.7	24.6	21.7	19.6	23.4	25.5	23.6	26.9	23.2 21.7 24.5	21.5 19.4 23.2		
17	65.	19.3 18.0 2	19.3	18	20.5	18.2	16.4	19.6	21.4	19.7	22.6	19.3 18.0 20.4	18.1 16.2 19.5		
18	70.	15.6 14.4 1	15.6	14.4	16.5	14.9	13.3	16	17.5	16	18.5	15.5 14.4 16.5	14.8 13.2 15.9		
19	75.	12.2 11.2 1	12.1	11.2	12.9	11.9	10.6	12.7	13.9	12.6	14.6	12.1 11.1 12.9	11.8 10.5 12.7		
20	80.	9.1 8.3 9.7	9.1	8.3	9.6	9.2	8.2	9.7	10.5	9.5	11.1	9.1 8.3 9.6	9.1 8.1 9.7		
21	85.	6.6 5.9 7.0	6.5	5.9	6.9	6.9	6.1	7.2	7.7	6.8	8	6.5 5.9 6.9	6.8 6.1 7.2		
22	90.	4.6 4.1 4.8	4.5	4	4.7	5	4.5	5.2	5.4	4.7	5.5	4.5 4.0 4.7	5.0 4.5 5.2		
23	95.	3.2 2.8 3.3	3.1	2.7	3.2	3.7	3.3	3.8	3.7	3.3	3.8	3.1 2.7 3.2	3.7 3.3 3.8		
24	100.	2.2 2.0 2.3	2.2	2	2.2	2.7	2.4	2.7	2.7	2.3	2.6	2.2 2.0 2.2	2.7 2.5 2.7		
25		All races and origins		White			Black			Hispanic1		Non-Hispanic white1	Non-Hispanic black1		
26	Age (year:	Total Male	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total Male Female	Total Male Female		
27	0.	100,000 10	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000 100,000 100,000	100,000 100,000 100,000		
28	1.	99,411 99,	99,508	99,467	99,551	98,861	98,760	98,966	99,503	99,465	99,543	99,510 99,467 99,556	98,875 98,783 98,971		
29	5.	99,312 99,	99,419	99,368	99,473	98,708	98,580	98,840	99,426	99,381	99,474	99,420 99,358 99,486	98,707 98,602 98,837		
30	10.	99,254 99,	99,365	99,307	99,426	98,626	98,488	98,767	99,379	99,327	99,436	99,368 99,293 99,443	98,617 98,510 98,761		
31	15.	99,181 99,	99,296	99,228	99,367	98,529	98,373	98,691	99,322	99,263	99,387	99,294 99,204 99,389	98,513 98,395 98,680		
32	20.	98,943 98,	99,072	98,928	99,222	98,194	97,868	98,532	99,132	99,006	99,267	99,066 98,903 99,239	98,149 97,848 98,506		
33	25.	98,503 98,	98,652	98,328	98,995	97,574	96,926	98,247	98,785	98,509	99,088	98,637 98,293 99,001	97,482 96,838 98,198		
34	30.	97,980 97,	98,137	97,612	98,693	96,863	95,878	97,871	98,403	97,954	98,898	98,087 97,518 98,669	96,742 95,754 97,800		
35	35.	97,357 96,	97,518	96,794	98,284	96,021	94,720	97,329	97,970	97,369	98,634	97,417 96,647 98,222	95,860 94,552 97,221		
36	40.	96,609 95,	96,782	95,862	97,754	94,972	93,348	96,579	97,465	96,686	98,318	96,618 95,640 97,638	94,739 93,086 96,419		
37	45.	95,619 94,	95,808	94,674	96,999	93,595	91,674	95,473	96,782	95,785	97,860	95,571 94,377 96,815	93,285 91,318 95,251		

Tabula Java Program

1. **Workaround:**
2. **Create two columns;**
3. **Data → Text to Columns;**
4. **Space-Separated;**
5. **Repeat as needed.**

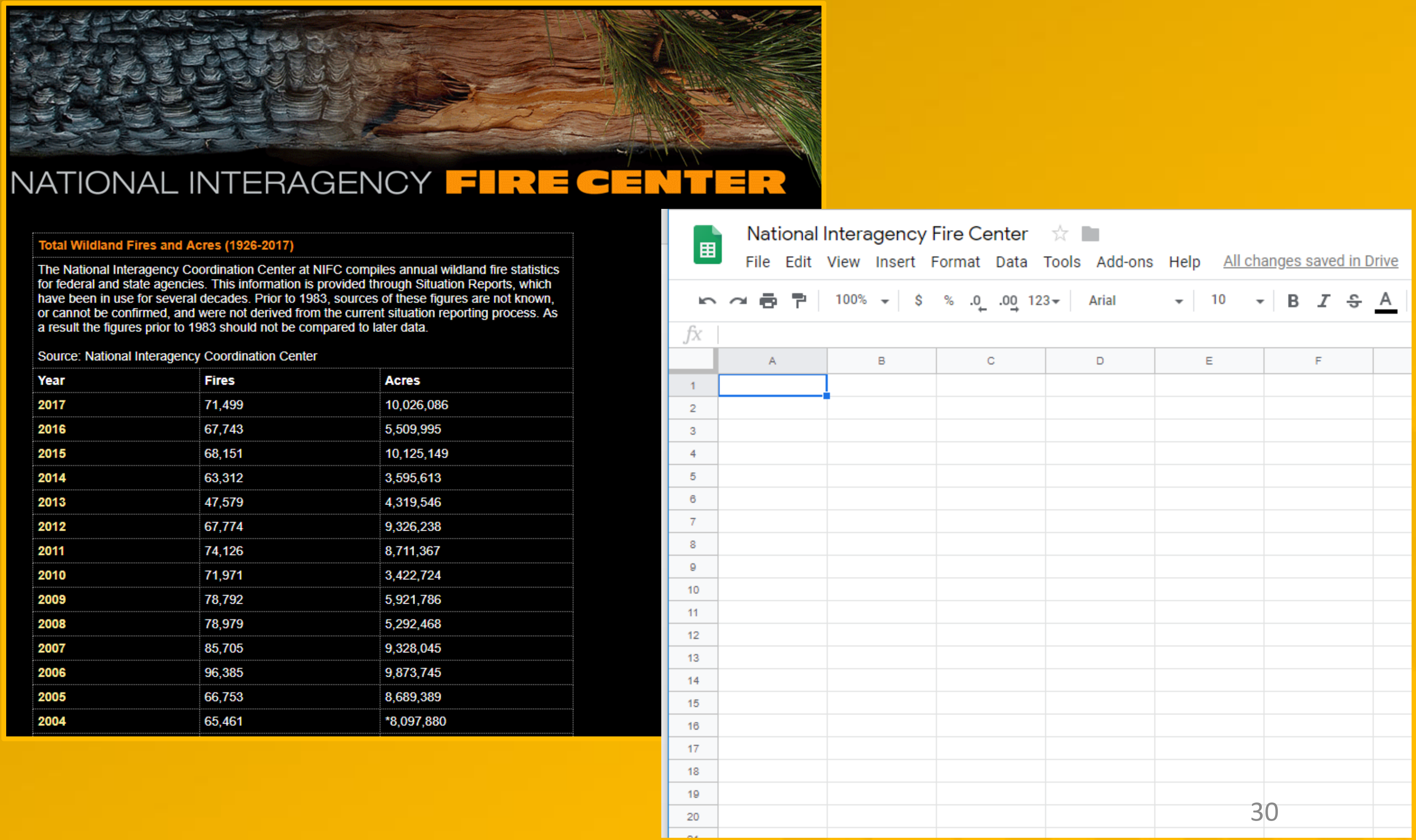


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		All races and origins		White			Black			Hispanic1		Non-Hispanic white1			Non-Hispanic black1		
2	Age (year:	Total Male	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
3	0.	78.7 76.3 8	78.9	76.6	81.3	75.5	72.2	78.5	81.9	79.3	84.3	78.7	76.3	81	75.1	71.9	78.[illegible]
4	1.	78.2 75.8 8	78.3	76	80.6	75.3	72.1	78.3	81.3	78.7	83.7	78.1	75.7	80.4	75	71.8	77.9
5	5.	74.3 71.9 7	74.4	72.1	76.7	71.5	68.3	74.4	77.4	74.8	79.7	74.1	71.8	76.4	71.1	67.9	74
6	10.	69.3 66.9 7	69.4	67.1	71.7	66.5	63.3	69.4	72.4	69.8	74.8	69.2	66.9	71.5	66.1	62.9	69.1
7	15.	64.4 62.0 6	64.5	62.2	66.8	61.6	58.4	64.5	67.5	64.9	69.8	64.2	61.9	66.5	61.2	58	64.1
8	20.	59.5 57.2 6	59.6	57.3	61.9	56.8	53.7	59.6	62.6	60	64.9	59.4	57.1	61.6	56.4	53.3	59.3
9	25.	54.8 52.5 5	54.8	52.7	57	52.1	49.2	54.7	57.8	55.3	60	54.6	52.4	56.7	51.8	48.9	54.4
10	30.	50.0 47.9 5	50.1	48	52.2	47.5	44.7	49.9	53	50.6	55.1	49.9	47.8	51.9	47.2	44.4	49.6
11	35.	45.3 43.3 4	45.4	43.4	47.4	42.9	40.2	45.2	48.2	45.9	50.3	45.2	43.2	47.2	42.6	39.9	44.9
12	40.	40.7 38.7 4	40.7	38.8	42.6	38.3	35.8	40.5	43.5	41.2	45.4	40.6	38.7	42.4	38.1	35.5	40.3
13	45.	36.1 34.2 3	36.1	34.3	37.9	33.8	31.4	36	38.8	36.6	40.6	36	34.1	37.8	33.6	31.1	35.7
14	50.	31.6 29.8 3	31.6	29.9	33.3	29.5	27.1	31.6	34.2	32	35.9	31.5	29.8	33.2	29.3	26.9	31.3
15	55.	27.3 25.6 2	27.3	25.7	28.9	25.4	23.2	27.3	29.7	27.7	31.3	27.3	25.6	28.8	25.3	23	27.2
16	60.	23.2 21.7 2	23.2	21.7	24.6	21.7	19.6	23.4	25.5	23.6	26.9	23.2	21.7	24.5	21.5	19.4	23.2
17	65.	19.3 18.0 2	19.3	18	20.5	18.2	16.4	19.6	21.4	19.7	22.6	19.3	18	20.4	18.1	16.2	19.5
18	70.	15.6 14.4 1	15.6	14.4	16.5	14.9	13.3	16	17.5	16	18.5	15.5	14.4	16.5	14.8	13.2	15.9
19	75.	12.2 11.2 1	12.1	11.2	12.9	11.9	10.6	12.7	13.9	12.6	14.6	12.1	11.1	12.9	11.8	10.5	12.7
20	80.	9.1 8.3 9.7	9.1	8.3	9.6	9.2	8.2	9.7	10.5	9.5	11.1	9.1	8.3	9.6	9.1	8.1	9.7
21	85.	6.6 5.9 7.0	6.5	5.9	6.9	6.9	6.1	7.2	7.7	6.8	8	6.5	5.9	6.9	6.8	6.1	7.2
22	90.	4.6 4.1 4.8	4.5	4	4.7	5	4.5	5.2	5.4	4.7	5.5	4.5	4	4.7	5	4.5	5.2
23	95.	3.2 2.8 3.3	3.1	2.7	3.2	3.7	3.3	3.8	3.7	3.3	3.8	3.1	2.7	3.2	3.7	3.3	3.8
24	100.	2.2 2.0 2.3	2.2	2	2.2	2.7	2.4	2.7	2.7	2.3	2.6	2.2	2	2.2	2.7	2.5	2.7
25		All races and origins		White			Black			Hispanic1		Non-Hispanic white1			Non-Hispanic black1		
26	Age (year:	Total Male	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total Male Female			Total Male Female		
27	0.	100,000 10	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000 100,000 100,000			100,000 100,000 100,000		
28	1.	99,411 99,	99,508	99,467	99,551	98,861	98,760	98,966	99,503	99,465	99,543	99,[illegible]10 99,467 99,556			98,875 98,783 98,971		
29	5.	99,312 99,	99,419	99,368	99,473	98,708	98,580	98,840	99,426	99,381	99,474	99,4[illegible]0 99,358 99,486			98,707 98,602 98,837		
30	10.	99,254 99,	99,365	99,307	99,426	98,626	98,488	98,767	99,379	99,327	99,436	99,36[illegible] 99,293 99,443			98,617 98,510 98,761		
31	15.	99,181 99,	99,296	99,228	99,367	98,529	98,373	98,691	99,322	99,263	99,387	99,294 99,204 99,389			98,513 98,395 98,680		
32	20.	98,943 98,	99,072	98,928	99,222	98,194	97,868	98,532	99,132	99,006	99,267	99,066 98,903 99,239			98,149 97,848 98,506		
33	25.	98,503 98,	98,652	98,328	98,995	97,574	96,926	98,247	98,785	98,509	99,088	98,637 98,[illegible]93 99,001			97,482 96,838 98,198		
34	30.	97,980 97,	98,137	97,612	98,693	96,863	95,878	97,871	98,403	97,954	98,898	98,087 97,5[illegible]3 98,669			96,742 95,754 97,800		
35	35.	97,357 96,	97,518	96,794	98,284	96,021	94,720	97,329	97,970	97,369	98,634	97,417 96,64[illegible] 98,222			95,860 94,552 97,221		
36	40.	96,609 95,	96,782	95,862	97,754	94,972	93,348	96,579	97,465	96,686	98,318	96,618 95,640 97,638			94,739 93,086 96,419		
37	45.	95,619 94,	95,808	94,674	96,999	93,595	91,674	95,473	96,782	95,785	97,860	95,571 94,377 96,8[illegible]			93,285 91,318 95,251		
38	50.	94,158 92,	94,374	92,952	95,864	91,622	89,341	93,833	95,733	94,481	97,083	94,065 92,586 95,601			91,228 88,889 93,53[illegible]		

Star Attraction: Google Sheet Hack

TOWSON UNIVERSITY

- Plenty of data or directories are still in static HTML tables;
- Lena Groeger, ProPublica.org, has a Google Sheet hack;
- Populates a Google sheet with a static HTML table in a single formula:



NATIONAL INTERAGENCY FIRE CENTER

Total Wildland Fires and Acres (1926-2017)

The National Interagency Coordination Center at NIFC compiles annual wildland fire statistics for federal and state agencies. This information is provided through Situation Reports, which have been in use for several decades. Prior to 1983, sources of these figures are not known, or cannot be confirmed, and were not derived from the current situation reporting process. As a result the figures prior to 1983 should not be compared to later data.

Source: National Interagency Coordination Center

Year	Fires	Acres
2017	71,499	10,026,086
2016	67,743	5,509,995
2015	68,151	10,125,149
2014	63,312	3,595,613
2013	47,579	4,319,546
2012	67,774	9,326,238
2011	74,126	8,711,367
2010	71,971	3,422,724
2009	78,792	5,921,786
2008	78,979	5,292,468
2007	85,705	9,328,045
2006	96,385	9,873,745
2005	66,753	8,689,389
2004	65,461	*8,097,880

National Interagency Fire Center

Google Sheet Hack

TOWSON UNIVERSITY

- Step One

- Gather these data elements;

❑ URL;
❑ type of element;
❑ first data element.

→

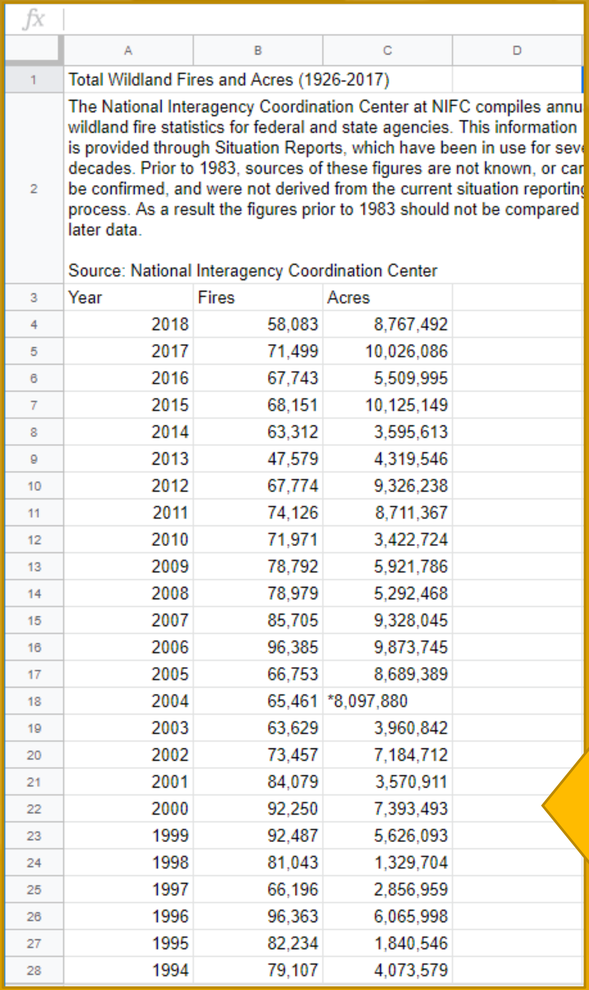
❑ The target page
❑ Table
❑ 0 (starts at the top)

❑ https://www.nifc.gov/fireInfo/fireInfo_stats_totalFires.html

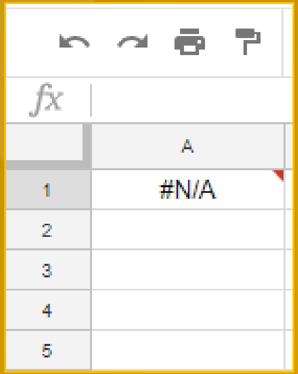
❑ "table"

❑ "0"

=IMPORTHTML("https://www.nifc.gov/fireInfo/fireInfo_stats_totalFires.html","table","0")



	A	B	C	D
1	Total Wildland Fires and Acres (1926-2017)			
2	The National Interagency Coordination Center at NIFC compiles annu wildland fire statistics for federal and state agencies. This information is provided through Situation Reports, which have been in use for sev decades. Prior to 1983, sources of these figures are not known, or car be confirmed, and were not derived from the current situation reportin process. As a result the figures prior to 1983 should not be compared later data. Source: National Interagency Coordination Center			
3	Year	Fires	Acres	
4	2018	58,083	8,767,492	
5	2017	71,499	10,026,086	
6	2016	67,743	5,509,995	
7	2015	68,151	10,125,149	
8	2014	63,312	3,595,613	
9	2013	47,579	4,319,546	
10	2012	67,774	9,326,238	
11	2011	74,126	8,711,367	
12	2010	71,971	3,422,724	
13	2009	78,792	5,921,786	
14	2008	78,979	5,292,468	
15	2007	85,705	9,328,045	
16	2006	96,385	9,873,745	
17	2005	66,753	8,689,389	
18	2004	65,461	*8,097,880	
19	2003	63,629	3,960,842	
20	2002	73,457	7,184,712	
21	2001	84,079	3,570,911	
22	2000	92,250	7,393,493	
23	1999	92,487	5,626,093	
24	1998	81,043	1,329,704	
25	1997	66,196	2,856,959	
26	1996	96,363	6,065,998	
27	1995	82,234	1,840,546	
28	1994	79,107	4,073,579	

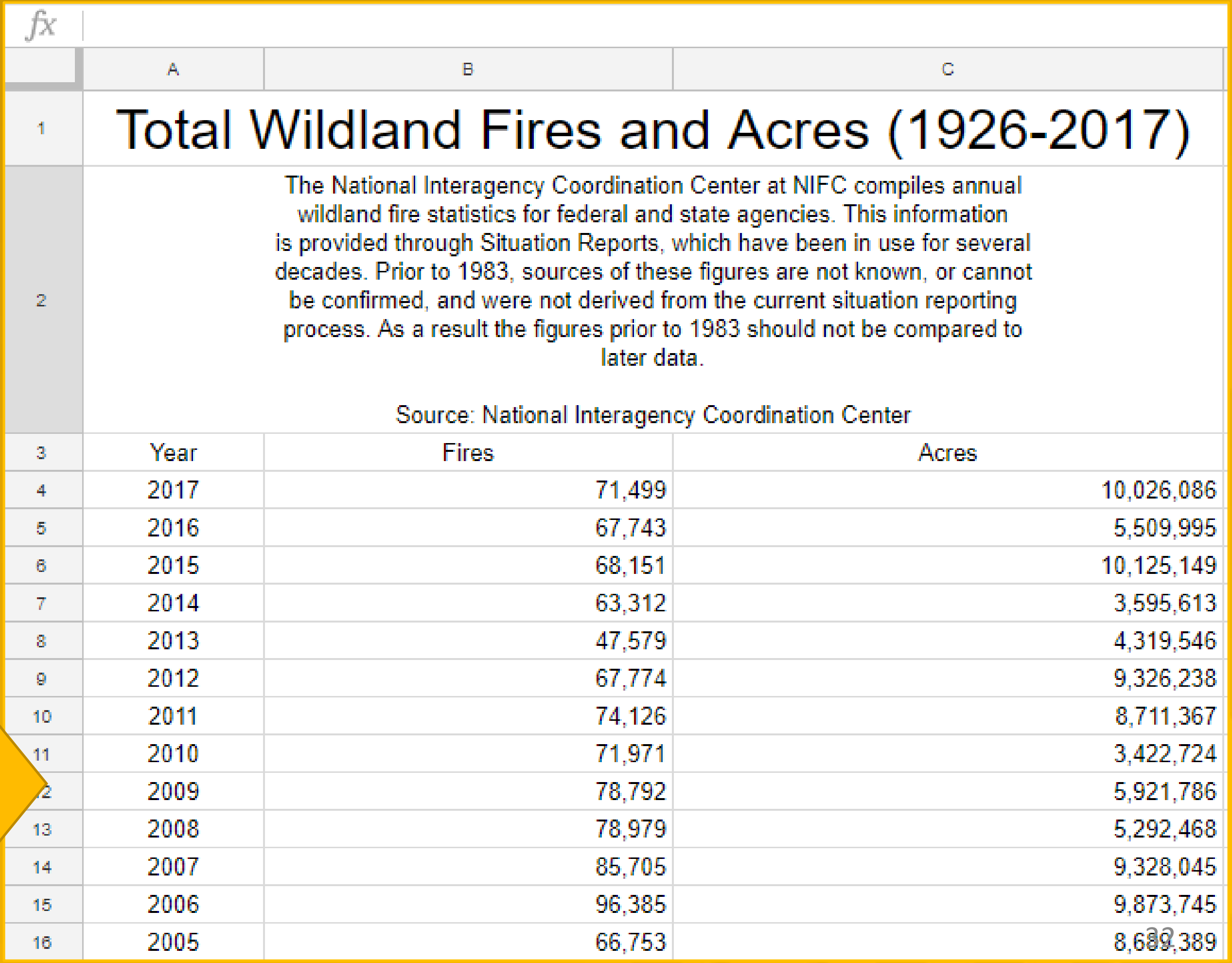


	A
1	#N/A
2	
3	
4	
5	

Error

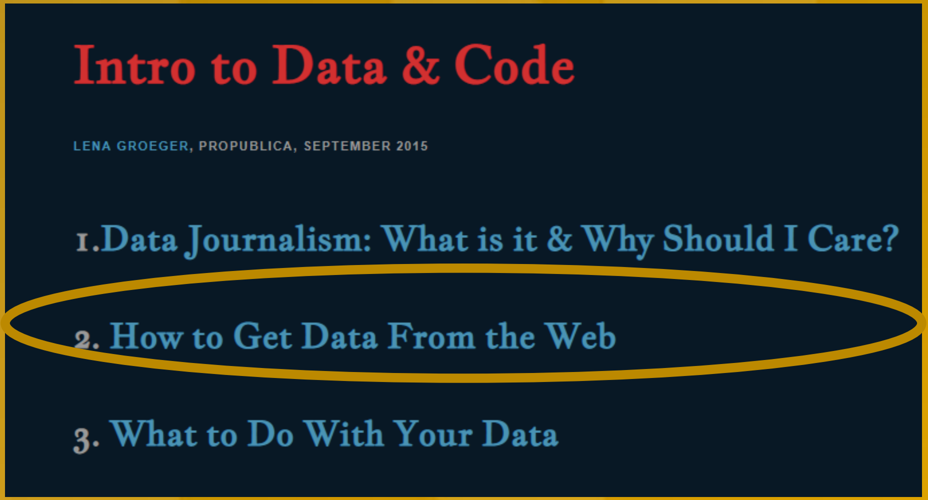
Rough

Pretty



	A	B	C
1	Total Wildland Fires and Acres (1926-2017)		
2	The National Interagency Coordination Center at NIFC compiles annual wildland fire statistics for federal and state agencies. This information is provided through Situation Reports, which have been in use for several decades. Prior to 1983, sources of these figures are not known, or cannot be confirmed, and were not derived from the current situation reporting process. As a result the figures prior to 1983 should not be compared to later data. Source: National Interagency Coordination Center		
3	Year	Fires	Acres
4	2017	71,499	10,026,086
5	2016	67,743	5,509,995
6	2015	68,151	10,125,149
7	2014	63,312	3,595,613
8	2013	47,579	4,319,546
9	2012	67,774	9,326,238
10	2011	74,126	8,711,367
11	2010	71,971	3,422,724
12	2009	78,792	5,921,786
13	2008	78,979	5,292,468
14	2007	85,705	9,328,045
15	2006	96,385	9,873,745
16	2005	66,753	8,689,389

Lena Groeger, ProPublica

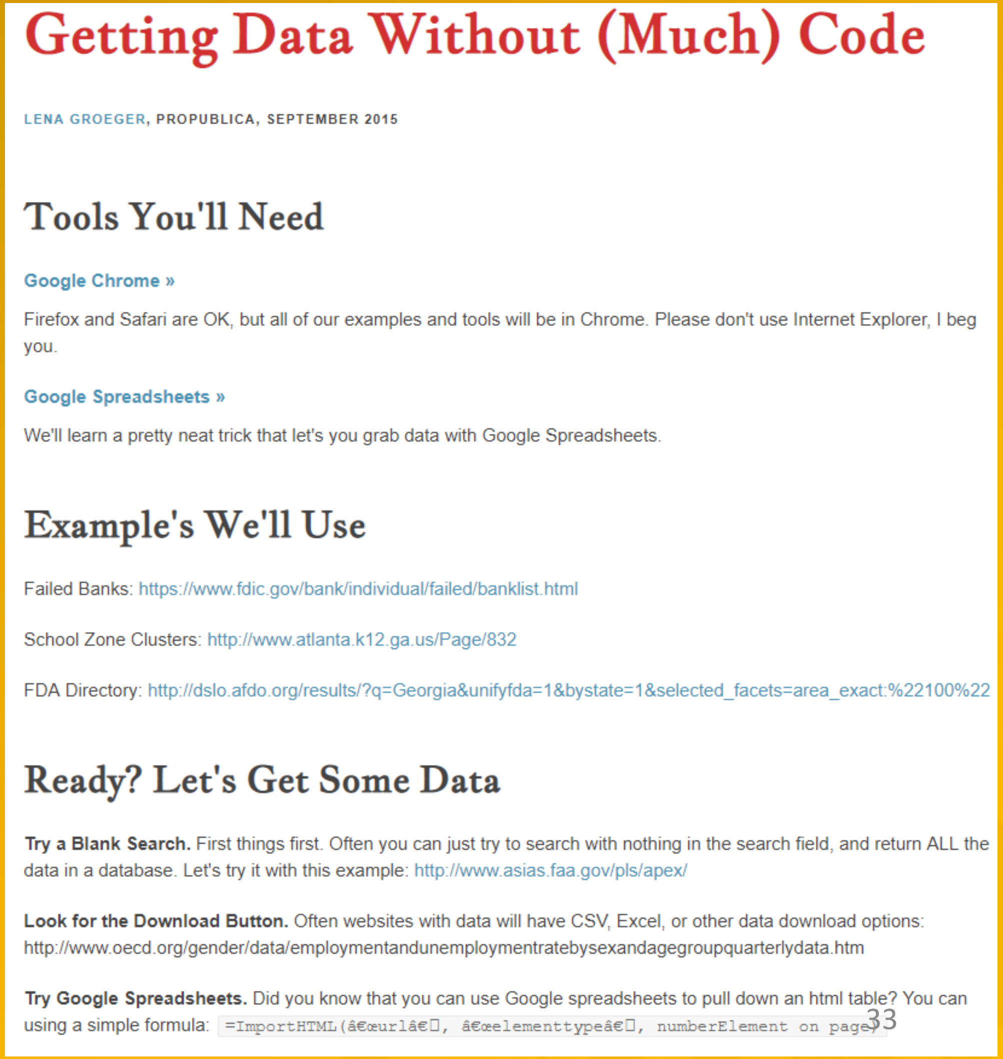


Intro to Data & Code

LENA GROEGER, PROPUBLICA, SEPTEMBER 2015

1. Data Journalism: What is it & Why Should I Care?
2. How to Get Data From the Web
3. What to Do With Your Data

https://bit.ly/1Kn6Eav



Getting Data Without (Much) Code

LENA GROEGER, PROPUBLICA, SEPTEMBER 2015

Tools You'll Need

Google Chrome »

Firefox and Safari are OK, but all of our examples and tools will be in Chrome. Please don't use Internet Explorer, I beg you.

Google Spreadsheets »

We'll learn a pretty neat trick that let's you grab data with Google Spreadsheets.

Example's We'll Use

Failed Banks: https://www.fdic.gov/bank/individual/failed/banklist.html

School Zone Clusters: http://www.atlanta.k12.ga.us/Page/832

FDA Directory: http://dslo.afdo.org/results/?q=Georgia&unifyfda=1&bystate=1&selected_facets=area_exact:%22100%22

Ready? Let's Get Some Data

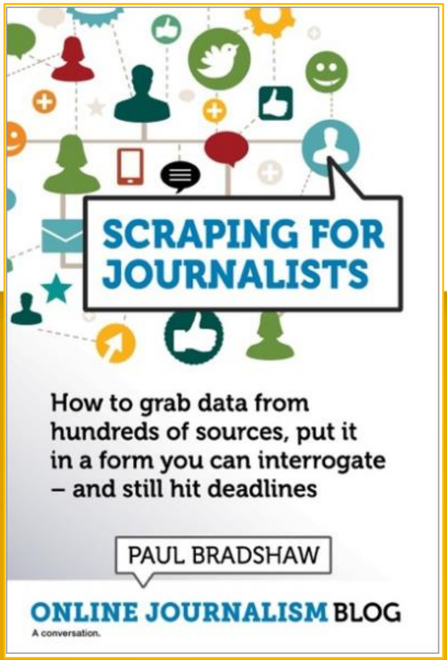
Try a Blank Search. First things first. Often you can just try to search with nothing in the search field, and return ALL the data in a database. Let's try it with this example: http://www.asias.faa.gov/pls/apex/

Look for the Download Button. Often websites with data will have CSV, Excel, or other data download options: http://www.oecd.org/gender/data/employmentandunemploymentratebysexandagegroupquarterlydata.htm

Try Google Spreadsheets. Did you know that you can use Google spreadsheets to pull down an html table? You can using a simple formula: `=ImportHTML(â€œurlâ€, â€œelementtypeâ€, numberElement on page)`

Further Information

- Paul Bradshaw,
- Master's Program, Birmingham City University,
- Online Journalism Blog: https://onlinejournalismblog.com/;
- Ebooks for sale from LeanPub:
    - Scraping for Journalists ($20.01) - https://leanpub.com/scrapingforjournalists
    - Data Journalism Heist ($9.99) - https://leanpub.com/DataJournalismHeist



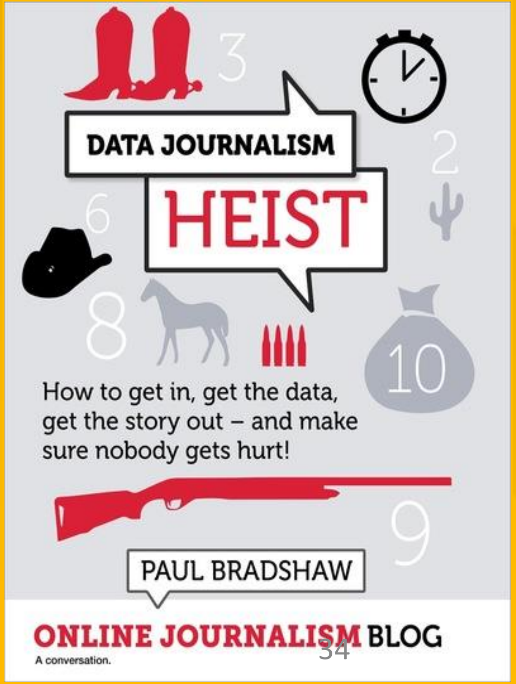
SCRAPING FOR JOURNALISTS

How to grab data from hundreds of sources, put it in a form you can interrogate – and still hit deadlines

PAUL BRADSHAW

ONLINE JOURNALISM BLOG

A conversation.



DATA JOURNALISM HEIST

How to get in, get the data, get the story out – and make sure nobody gets hurt!

PAUL BRADSHAW

ONLINE JOURNALISM BLOG

A conversation.

Data Scraping for the Coding-Challenged

Carl P. Olson

colson@towson.edu

Thank You!

TOWSON UNIVERSITY