

# Obtaining Data

---

**DATA 201: Thinking With Data**  
**Winter 2022**

Jonathan Hudson, Ph.D  
Instructor  
Department of Computer Science  
University of Calgary

**Tuesday, March 22, 2022**

# Why do we want data?

# Ethics — Key Points

---

- Ethics boards
- Informed consent
- Storing data
- Accessing data
- Anonymizing data
- Deleting data

# How to obtain data?



# Collecting Data

---

- Collect
- Generate
- Find / Extract

# Collecting Data

---

- Collect
  - Observation, surveys, logging, sensors, crowdsourcing
- Generate
- Find / Extract

# Collecting Data

---

- Collect
  - Observation, surveys, logging, sensors, crowdsourcing
- Generate
  - Simulations, experiment
- Find / Extract

# Collecting Data

---

- Collect
  - Observation, surveys, logging, sensors, crowdsourcing
- Generate
  - Simulations, experiment
- Find / Extract
  - Open corpuses, data retailers, Application Program Interfaces (APIs), scraping the Web

# Observation

---




# Surveys

Qualtrics: Online Survey Sc x

www.qualtrics.com

1-800-340-9194



1-800-340-9194


## Ask Questions. Get Answers. Act Fast.



Collect



Analyze



Act

Create your free account


Chat w/ Sales

### The Ridiculously Powerful Insight Platform



University Student Satisfac x

https://www.surveymonkey.com/blog/en/university-student-sati...



### 1. How well do the professors teach at this university?

☐ Extremely well

☐ Quite well

☒ Moderately well

☐ Slightly well

☐ Not at all well

### 2. How effective is the teaching outside your major at this university?

☐ Extremely effective


☐ Very effective


☐ Moderately effective


☐ Slightly effective


This survey asks students to assess educational, social, and other aspects of your university. How effective is the teaching? Is their academic advisor helpful? Do students feel safe on campus? When you get to know what your students want, you can improve certain programs and your overall enrollment and retain students. Customize this template or add questions if you want to know how students experienced specific programs or events.

To create a survey using the University Student Satisfaction survey template, just sign up or sign in to SurveyMonkey. You'll be able to choose the template when you begin creating a survey.

 0


 0

 Share

 Pin it

Use Cases x Use Cases x Google Con x How it work x Home | GoC x

www.google.com/insights/consumersurveys/home



Report Custom Insights

What would most influence your decision to buy clothes online instead of in-store?  
Results for respondents with demographics. Weighted by Age, Gender, Region. (1,538 responded) @ Winner statistically significant. @

Sum Compare

Inferred Gender

Male Female

Sum Compare

Inferred Age

18-24 25-34 35-44 45-54 55-64 65+

Demographics

Free shipping 40.5% (+0.5 / -0.5)

Online discounts 24.9% (+0.4 / -0.4)

Ability to return in store 16.4% (+0.4 / -0.4)

Free returns 16.1% (+0.4 / -0.4)

Options

Weighted ON

Percentages ON

Sort by winner ON

Answers

☒ Free shipping



☒ Free returns

☒ Ability to return in store

☒ Online discounts

## To find out what people really think, just ask the Internet.

When you want answers to your business questions, you need to reach everyday people — not just those who choose to participate in research panels.



Introducing

Mobile surveys and language targeting.

Learn more

Y OF  
RY

# General Guidelines

---

- Clear purpose/goal
- Every question is necessary
- Inform about survey length
- Simple and direct questions
- Logical order
- Avoid biased questions
- Use simple language
- Consistent with rating scale
- Test your survey
- Offer incentives

<b>Month</b>	<b>Treatment</b>	<b>Pressure</b>
March	Control	165
March	Placebo	163
March	300 mg	166
March	450 mg	168
April	Control	162
April	Placebo	159
April	300 mg	161
April	450 mg	163
May	Control	164



# Surveys

---

Thank you for being a client of RBC.

This is a friendly reminder regarding an email we sent you recently.

You recently dealt with RBC over the phone.

We've shortened our survey to just 7 minutes and would like your thoughts on how we can better serve you.

Please complete the entire survey no later than **Wednesday, July 5, 2017.**

[Start the Survey](#)

Thank you for your time!

RBC Royal Bank and Ipsos

*This survey is for service quality purposes only, you will not be asked to reveal personal account information. The information and feedback you provide will be shared with RBC to help them improve the service they provide to you and will be held in the strictest confidence by the bank.*

# Surveys

---

Dear Colleague,

You recently received an invitation to participate in a survey to tell us about the teaching culture at the University of Calgary. The purpose of this research is to develop a new survey instrument that can be used to provide feedback to universities on their teaching cultures, and how they can strengthen support for learning and teaching over time.

The research is funded by the Social Sciences and Humanities Research Council (SSHRC) and is being conducted at nine Canadian universities. The University of Calgary Conjoint Faculties Research Ethics Board has approved this study.

The survey will take 20 minutes to complete and is available until February 28, 2018.

Survey participation is confidential and voluntary; your decision regarding participation will not affect your position as a University of Calgary employee. You may decline to respond to any question by leaving it blank. If you have any questions about the research, please email the researchers at [TCPS@ucalgary.ca](mailto:TCPS@ucalgary.ca) or visit the [Institutional Teaching Culture Study](#) website

[Take the Institutional Teaching Culture Perception Survey](#)

Or copy and paste the URL below into your internet browser:

[https://ucalgary.ca1.qualtrics.com/jfe/form/SV\\_57Q](https://ucalgary.ca1.qualtrics.com/jfe/form/SV_57Q)

In order for the questions to display properly, we strongly encourage you to complete this survey on a computer.

Thank you for considering this invitation.

Sincerely,

Lynn Taylor (former vice-provost teaching and learning)

# Surveys

---





# Surveys



# Surveys

---

- Which movie genre do you think you like the most?
- What is your favourite movie genre?

# Surveys

---

- Do you use ASCII to write emails?
- Do you use plain text to write emails?

# Surveys

---

- Wouldn't you like fries as a side dish?
- Do you like fries as a side dish?

# Surveys

---

I am very satisfied with the service.

1

2

3

4

5

Strongly  
Disagree

Strongly  
Agree



# Surveys

---

I am ~~very~~ satisfied with the service.

1

Strongly  
Disagree

2

3

4

5

Strongly  
Agree

# Likert Scale

---

1

Strongly  
Disagree

2

Disagree

3

Neutral

4

Agree

5

Strongly  
Agree

# Likert Scale

Please indicate how much you agree or disagree with each of these statements:					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
-----					
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

# Likert Scale

Likert  
Item

Please indicate how much you agree or disagree with each of these statements:					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
-----	-----	-----	-----	-----	-----
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

# Likert Scale

Likert  
Item

Likert  
Item

Please indicate how much you agree or disagree with each of these statements:					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
-----	-----	-----	-----	-----	-----
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

# Likert Scale

Likert  
Item

Likert  
Item

Likert  
Item

Please indicate how much you agree or disagree with each of these statements:					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
-----					
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

# Likert Scale

Please indicate how much you agree or disagree with each of these statements:					
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
	-----1	-----2	-----3	-----4	-----5
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

Likert Item

Response Level = 5

# Likert Scale

Please indicate how much you agree or disagree with each of these statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
-----	1	2	3	4	5
The president is doing a good job.	1	2	3	4	5
The Congress is doing a good job.	1	2	3	4	5
The Secretary of Defense is doing a good job.	1	2	3	4	5

Likert Scale

Likert Item

Response Level = 5



# Likert-Type Item

---

How often do you eat out?

1

2

3

4

5

Never

Sometimes

Average

Often

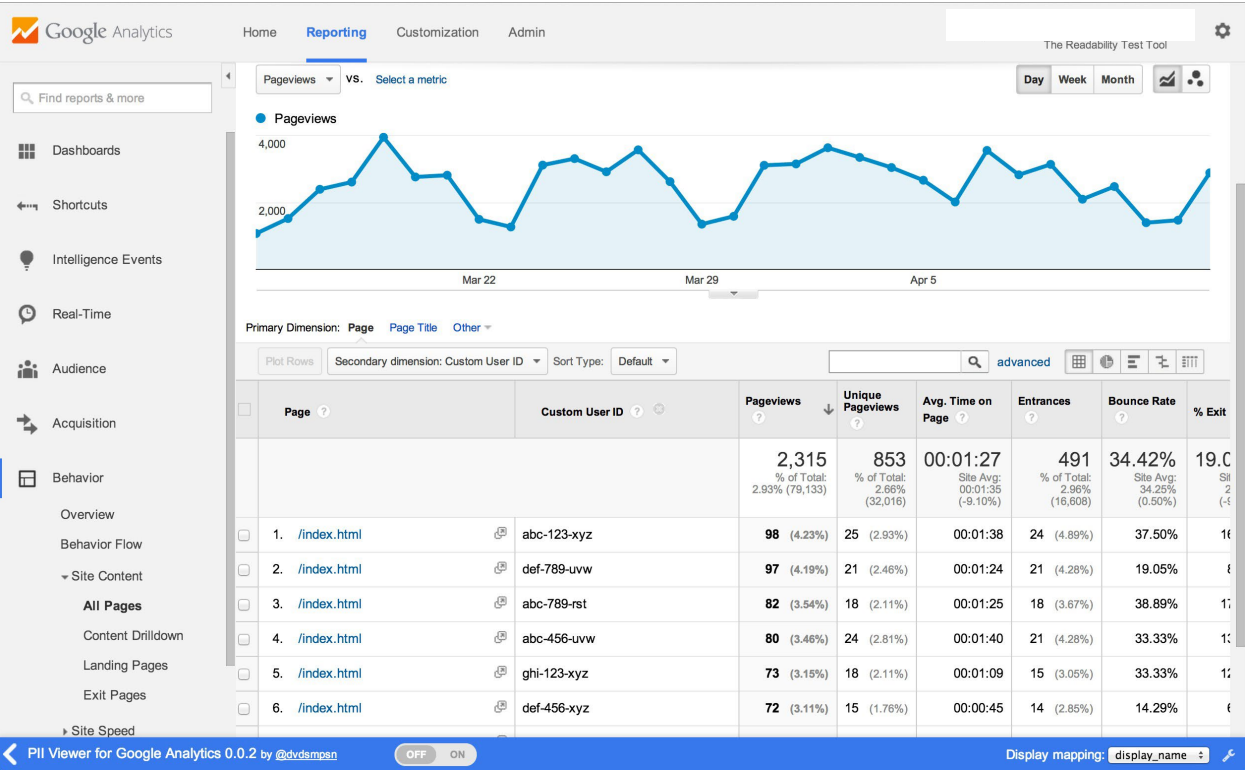
Very often

# Other Types of Questions

---

- Closed-ended questions
- Open-ended questions

# Logging

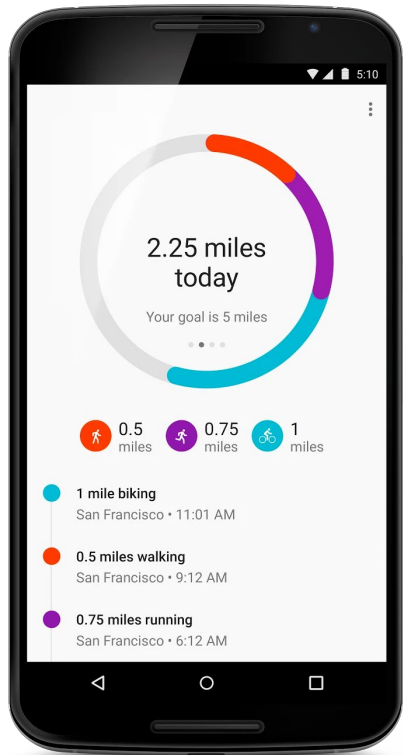


Google Analytics



Open Web Analytics

# Sensors



# Crowdsourcing

Amazon Mechanical Turk - x

← → ↺ <https://www.mturk.com/mturk/welcome> ☆

### HITs containing 'short survey'

11-20 of 49 Results

Sort by: HITs Available (most first) GO! [Show all details](#) | [Hide all details](#) [First](#) << [Previous](#) < 1 2 3 4 5 > [Next](#) >> [Last](#)

<a href="#">Answer a short survey about Work Team Dynamics</a> <a href="#">Request Qualification (Why?)</a>   <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Whitney Ohmer</a>	<b>HIT Expiration Date:</b> Oct 12, 2014 (2 weeks 5 days)	<b>Reward:</b> \$0.25	
	<b>Time Allotted:</b> 60 minutes	<b>HITs Available:</b> 1	
<a href="#">Answer a short survey about Work Team Dynamics</a> <a href="#">Request Qualification (Why?)</a>   <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Whitney Ohmer</a>	<b>HIT Expiration Date:</b> Oct 12, 2014 (2 weeks 5 days)	<b>Reward:</b> \$0.25	
	<b>Time Allotted:</b> 60 minutes	<b>HITs Available:</b> 1	
<a href="#">Short Survey</a> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">David Tannenbaum</a>	<b>HIT Expiration Date:</b> Oct 12, 2014 (2 weeks 5 days)	<b>Reward:</b> \$0.10	
	<b>Time Allotted:</b> 60 seconds	<b>HITs Available:</b> 1	
<a href="#">Short survey about website experience (on average it takes 13 minutes)</a> <a href="#">Not Qualified to work on this HIT (Why?)</a>   <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">peter slattery</a>	<b>HIT Expiration Date:</b> Oct 6, 2014 (1 week 6 days)	<b>Reward:</b> \$1.50	
	<b>Time Allotted:</b> 40 minutes	<b>HITs Available:</b> 1	
<a href="#">Take a short survey</a> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">Jonathan Pirc</a>	<b>HIT Expiration Date:</b> Oct 6, 2014 (1 week 6 days)	<b>Reward:</b> \$1.50	
	<b>Time Allotted:</b> 30 minutes	<b>HITs Available:</b> 1	
<a href="#">Answer a FUN and EASY survey about on consumer decision making!</a> <a href="#">View a HIT in this group</a>			
<b>Requester:</b> <a href="#">CJJ</a>	<b>HIT Expiration Date:</b> Oct 6, 2014 (1 week 6 days)	<b>Reward:</b> \$0.25	
	<b>Time Allotted:</b> 55 minutes	<b>HITs Available:</b> 1	

[FAQ](#) | [Contact Us](#) | [Careers at Mechanical Turk](#) | [Developers](#) | [Press](#) | [Policies](#) | [Blog](#)

©2005-2014 Amazon.com, Inc. or its Affiliates

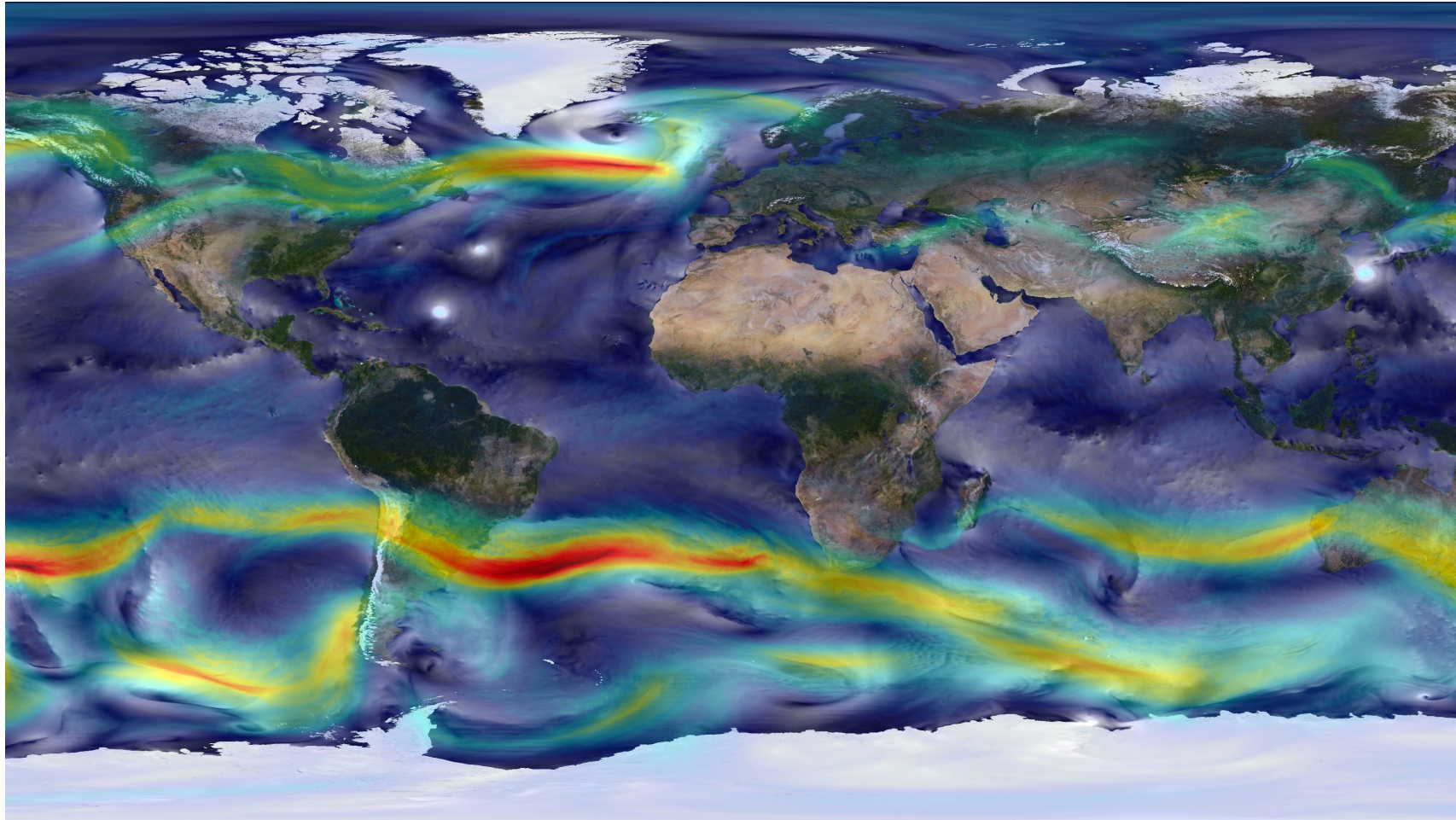
An [amazon.com](#) company

# Generating Data



# Simulations

---



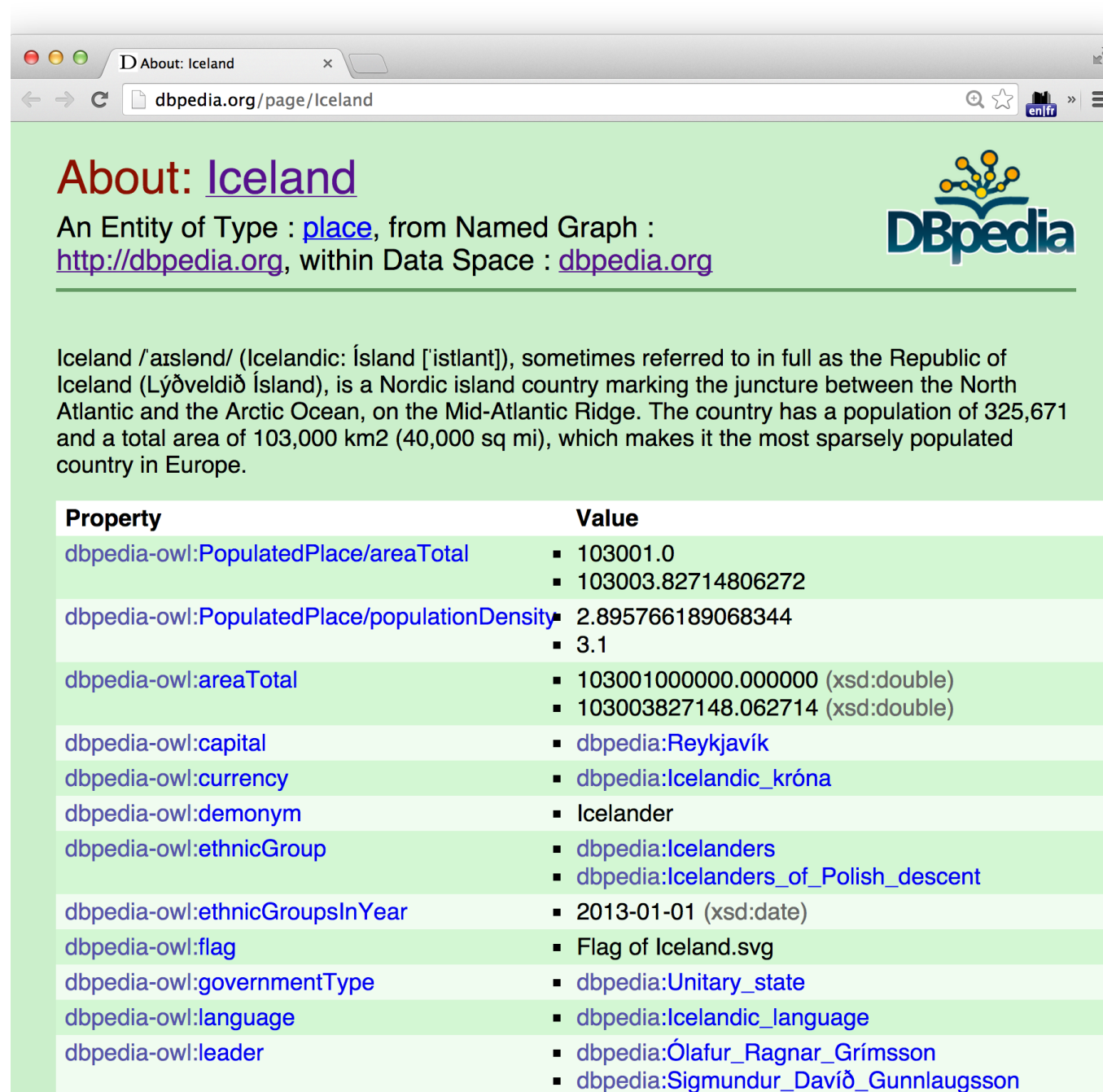
# Experiment

---





# Extracting / Finding Data



**About: Iceland**

An Entity of Type : [place](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](http://dbpedia.org)

Iceland /ˈaɪslənd/ (Icelandic: Ísland [ˈistlant]), sometimes referred to in full as the Republic of Iceland (Lýðveldið Ísland), is a Nordic island country marking the juncture between the North Atlantic and the Arctic Ocean, on the Mid-Atlantic Ridge. The country has a population of 325,671 and a total area of 103,000 km<sup>2</sup> (40,000 sq mi), which makes it the most sparsely populated country in Europe.

Property	Value
dbpedia-owl:PopulatedPlace/areaTotal	<ul style="list-style-type: none"> <li>103001.0</li> <li>103003.82714806272</li> </ul>
dbpedia-owl:PopulatedPlace/populationDensity	<ul style="list-style-type: none"> <li>2.895766189068344</li> <li>3.1</li> </ul>
dbpedia-owl:areaTotal	<ul style="list-style-type: none"> <li>103001000000.000000 (xsd:double)</li> <li>103003827148.062714 (xsd:double)</li> </ul>
dbpedia-owl:capital	<ul style="list-style-type: none"> <li>dbpedia:Reykjavík</li> </ul>
dbpedia-owl:currency	<ul style="list-style-type: none"> <li>dbpedia:Icelandic_króna</li> </ul>
dbpedia-owl:demonym	<ul style="list-style-type: none"> <li>Icelander</li> </ul>
dbpedia-owl:ethnicGroup	<ul style="list-style-type: none"> <li>dbpedia:Icelanders</li> <li>dbpedia:Icelanders_of_Polish_descent</li> </ul>
dbpedia-owl:ethnicGroupsInYear	<ul style="list-style-type: none"> <li>2013-01-01 (xsd:date)</li> </ul>
dbpedia-owl:flag	<ul style="list-style-type: none"> <li>Flag of Iceland.svg</li> </ul>
dbpedia-owl:governmentType	<ul style="list-style-type: none"> <li>dbpedia:Unitary_state</li> </ul>
dbpedia-owl:language	<ul style="list-style-type: none"> <li>dbpedia:Icelandic_language</li> </ul>
dbpedia-owl:leader	<ul style="list-style-type: none"> <li>dbpedia:Ólafur_Ragnar_Grímsson</li> <li>dbpedia:Sigmundur_Davíð_Gunnlaugsson</li> </ul>

Wikidata

https://www.wikidata.org/wiki/Wikidata:Main\_Page

English Not logged in Talk Contributions Create account Log in

Main Page Discussion Read View source View history Search

**WIKIDATA**

Main page  
Community portal  
Project chat  
Create a new item  
Item by title  
Recent changes  
Random item  
Nearby  
Help  
Donate

Print/export  
Create a book  
Download as PDF  
Printable version

Tools  
What links here  
Related changes  
Special pages  
Permanent link  
Page information  
Wikidata items

open

multilingual

free

collaborative

# Welcome to Wikidata

the free knowledge base with 16,194,852 data items that anyone can edit

[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

**Welcome!**

Wikidata is a free, linked database that can be read and edited by both humans and machines.

Wikidata acts as central storage for the structured data of its

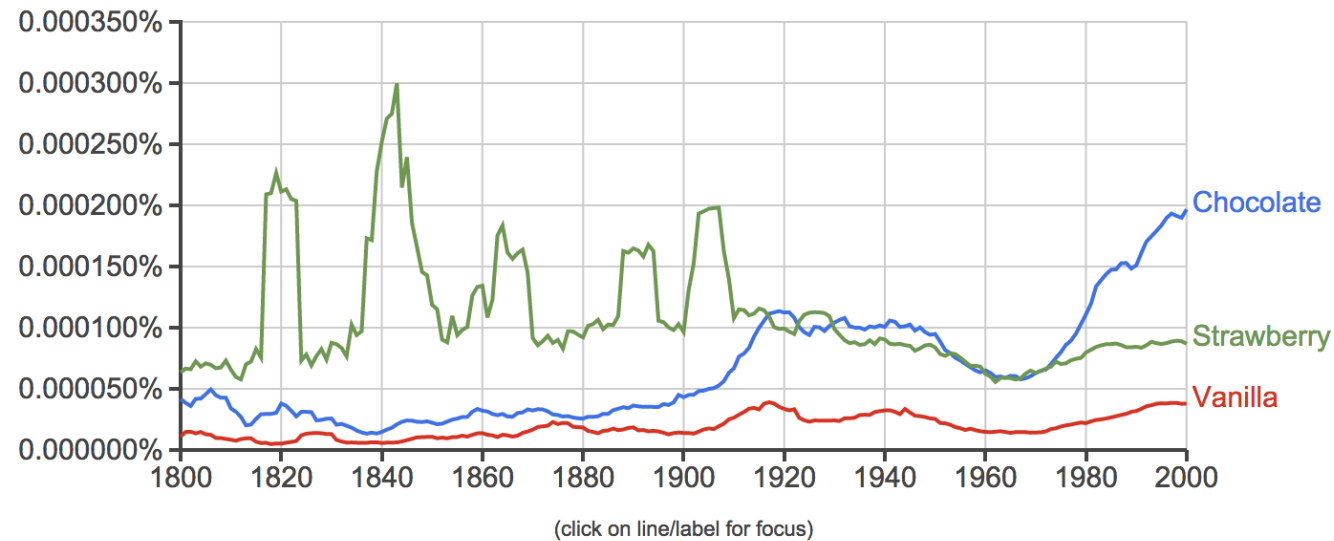
**Learn about data**

New to the wonderful world of data? [Develop and improve your data literacy through content](#) designed to get you up to speed and feeling comfortable with the fundamentals in no time.

## Google books Ngram Viewer

Graph these comma-separated phrases:  ☐ case-insensitive

between  and  from the corpus  with smoothing of  . [Search lots of books](#)



Search in Google Books:

<a href="#">1800 - 1845</a>	<a href="#">1846 - 1974</a>	<a href="#">1975 - 1985</a>	<a href="#">1986 - 1993</a>	<a href="#">1994 - 2000</a>	<a href="#">chocolate</a>
<a href="#">1800 - 1840</a>	<a href="#">1841 - 1913</a>	<a href="#">1914 - 1924</a>	<a href="#">1925 - 1990</a>	<a href="#">1991 - 2000</a>	<a href="#">vanilla</a>
<a href="#">1800 - 1820</a>	<a href="#">1821 - 1840</a>	<a href="#">1841 - 1852</a>	<a href="#">1853 - 1977</a>	<a href="#">1978 - 2000</a>	<a href="#">strawberry</a>

Run your own experiment! Raw data is available for download [here](#).

Sponsored by: **Institute for Reproducible Research** and  **Whatbox** and

Ads!  
Disable your ad-blocker!

Used by:



Academic  
Torrents



UNIVERSITY OF  
CALGARY



## Sports Reference | Sports Stats, fast, easy, and up-to-date

A group of sites providing both basic and sabermetric statistics and resources for sports fans everywhere. Our aim is to be the easiest-to-use, fastest, most complete sources for sports statistics anywhere.

### Baseball-Reference.com (Mobile)

Baseball statistics from 1871 to the present for major league [players](#), [teams](#), and [leagues](#). Complete [postseason](#) and [managerial](#) data is included as well. [Minor league stats](#) back to 1877, and [box scores](#), gamelogs and splits back to 1912. Updated daily.



### Basketball-Reference.com

Professional basketball statistics for every season from [Abdul-Jabbar](#) to [Zaslofsky](#). Find statistics for your favorite [player](#), [team](#), or [league](#). The site also includes sections for [coaches](#), [awards](#), [leaders](#), and the [playoffs](#). A user favorite is the [Play Index](#) custom leaderboards. Updated daily.



### Pro-Football-Reference.com

Football statistics and game results for all of NFL, AFL, AAFC & APFA history. The site includes [player](#), [Coach](#), and [team](#) stats, games scores for each team, Pro Bowl selections for each season, [extensive leaderboards](#), and every [draft pick](#) ever selected. Our [play index](#) allows for custom searches through every NFL box score since 1960, every TD in NFL history and every play since 1994. Additionally, you can search through every play in Super Bowl history. Updated the morning after every NFL game.



### Hockey-Reference.com

Hockey statistics from the creators of *Total Hockey* to our website. Find statistics for your favorite [player](#) from [Aalto](#) to [Zyzun](#), [team](#), or [league](#). The site also includes sections for [coaches](#), [drafts](#), [awards](#), [leaders](#), and the [playoffs](#). Analyze players using [the Play Index](#).



### College Football @ Sports-Reference.com

College Football game results for every major school back to [Princeton vs Rutgers](#) in 1869. [Player Stats](#) for [recent players](#) [schools](#), or [conferences](#) back to 1956. The site also includes sections for [coaches](#), [awards](#), [leaders](#), and [bowl games](#). Use the [Play Index](#) to create your own custom searches.



### College Basketball @ Sports-Reference.com

College Basketball Stats and Game Result. Game results for every game featuring two major schools since 1949. [Player Stats](#) from [Stephen Curry](#) to [Bill Walton](#). Stats for [schools](#) and [conferences](#). The site also includes sections for [coaches](#), [awards](#), [leaders](#), and the [postseason](#). Use the [Play Index](#) to create your own custom searches.



New to this site? [Start Here](#)

[DataBank](#) [Microdata](#) [Data Catalog](#) 

This page in: [English](#) [Español](#) [Français](#) [العربية](#) [中文](#)


# World Bank Open Data

Free and open access to global development data

Search data e.g. GDP, population, Indonesia ▾

Browse by [Country](#) or [Indicator](#)


## MOST RECENT

**New country classifications by income level: 2017-2018** 

World Bank Data Team, Jul 01, 2017

**Q2 2017 Update of World Development Indicators Available** 

World Bank Data Team, Jul 01, 2017

**Chart: Globally, The Number of People Without Access to Electricity is Falling** 

T. Khokhar, Jun 30, 2017

**Chart: Globally, Over 1 Billion**

[View all news](#) 

[View all blogs](#) 

## WHAT YOU CAN LEARN WITH OPEN DATA

**Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)**

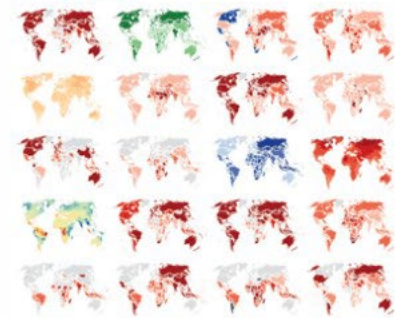


### Extreme Poverty

The proportion of the world's population living in extreme poverty has dropped significantly

## Atlas of Sustainable Development Goals 2017

From World Development Indicators



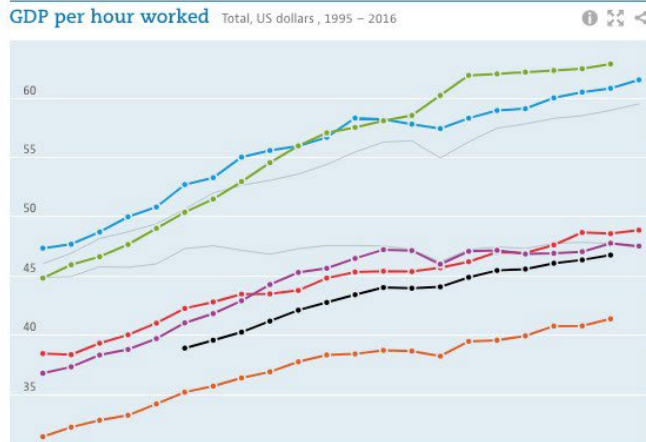
**A Visual Guide to the 17 SDGs** 

Apr 17, 2017

OECD Data

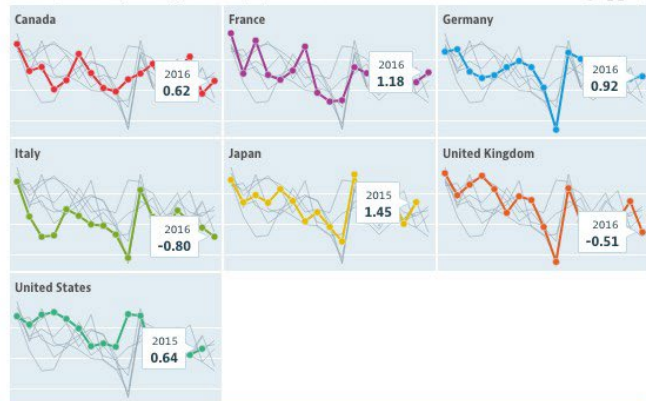
## Featured Charts

Labour productivity differences across OECD countries remain significant. Labour productivity growth continues its long-term decline, including in some G7 economies.



## Labour productivity and utilisation

Labour productivity, Annual growth rate (%), 2000 – 2016



## Latest news

### Statistical news releases

See recent statistical [news releases](#).

### Compare your income

Try OECD's [Compare your income](#) to check your perception of income inequality against reality!

## Statistical resources

### Database access



Make your own queries on large databases in our data warehouse, OECD.Stat.

[Query databases](#)

### Statistical tools

Learn about OECD methodologies and stay informed through statistical newsletters.

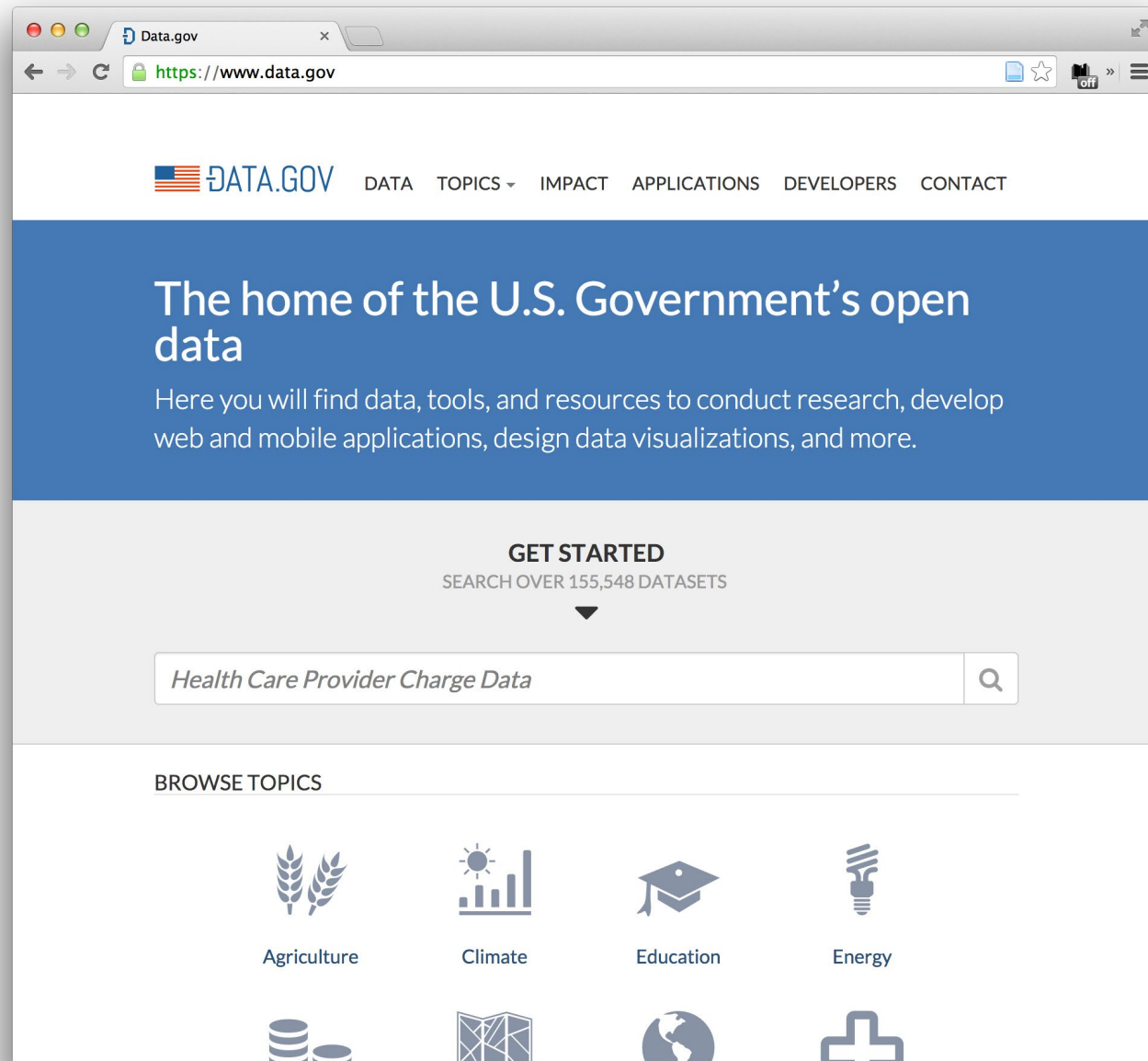
[Go to statistical tools](#)

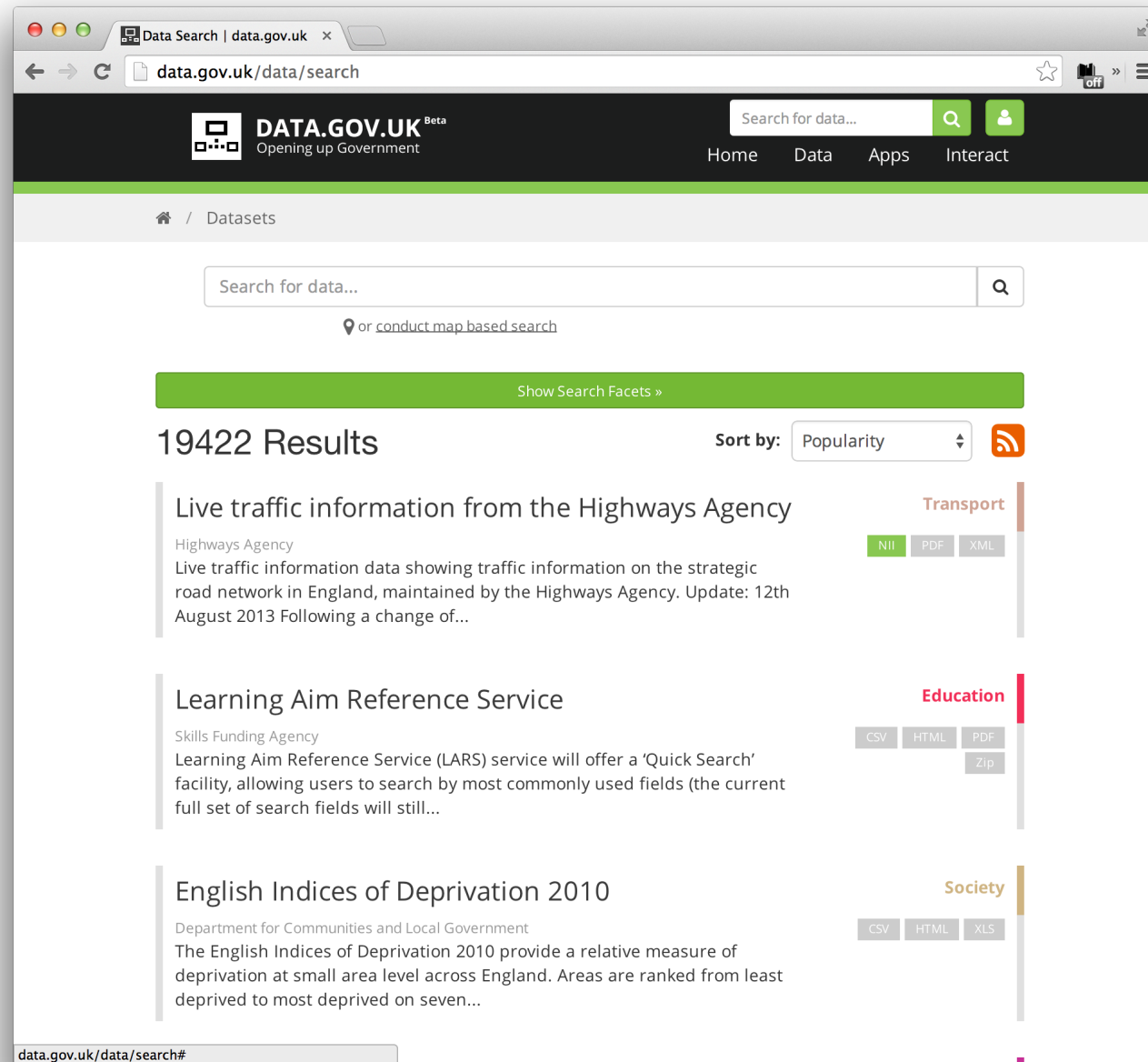
### Developer API

Learn about OECD's application programming interface (API) for accessing data.

[Go to API for developers](#)







data.gov.be

data.gov.be/fr/datasets

nl fr de en

Beta

Data.gov.be

HOME CONDITIONS D'UTILISATION DATA APPS IDÉES FORUM CONTACT LIENS

Liste de sets de données disponibles comme "open data".

Catégorie Type Granularité

- Tout - - Tout - - Tout - Appliquer

Titre	Catégorie	Type	Granularité
<a href="#">Zones de stationnement voirie 2013</a>	<a href="#">Mobilité</a>	<a href="#">Téléchargement</a>	
<a href="#">Usages TIC des ménages wallons</a>	<a href="#">TIC</a>	<a href="#">Téléchargement, Service web</a>	
<a href="#">Usages TIC des citoyens wallons</a>	<a href="#">Population, Economie, TIC</a>	<a href="#">Téléchargement, Service web</a>	
<a href="#">UDP Mars 2013 par commune</a>	<a href="#">Energie, Pouvoirs publics</a>	<a href="#">Téléchargement</a>	
<a href="#">UDP Mai 2013 par commune</a>	<a href="#">Energie, Pouvoirs publics</a>	<a href="#">Téléchargement</a>	
<a href="#">UDP Janvier 2013 par habitant</a>	<a href="#">Population, Energie</a>	<a href="#">Téléchargement</a>	
<a href="#">UDP Janvier 2013 par commune</a>	<a href="#">Energie, Pouvoirs publics</a>	<a href="#">Téléchargement</a>	
<a href="#">UDP Février 2013 par commune</a>	<a href="#">Energie, Pouvoirs publics</a>	<a href="#">Téléchargement</a>	
<a href="#">UDP Evolution Mensuelle</a>	<a href="#">Energie</a>	<a href="#">Téléchargement</a>	

data.gov.be/fr/apps



Government  
of Canada

Gouvernement  
du Canada

[Français](#)

Search Canada.ca



[Jobs](#) ▾

[Immigration](#) ▾

[Travel](#) ▾

[Business](#) ▾

[Benefits](#) ▾

[Health](#) ▾

[Taxes](#) ▾

[More services](#) ▾

[Home](#) → [Open Government](#)



## Open Data

Search open data that is relevant to Canadians, learn how to work with datasets, and see what people have done with open data across the country.

Follow:  

### [Search through our Open data portal](#)

Looking for data about Government of Canada services, financials, national demographic information or high resolution maps? Discover that and more through our open data portal, your one-stop shop for Government of Canada open datasets.

### [Open Data Inventory](#)

Review the open data inventories submitted by Government of Canada departments and agencies.

### [Working with data and application programming interfaces](#)

Learn more about how to use a variety of data formats, including application programming interfaces (APIs).

### [Apps gallery](#)

Browse and download apps that were built both by the Canadian public and the Government of Canada that use open data.

Search through our Open Government Portal.

 Find

### [Open maps](#)

Explore the Government of Canada's geospatial data, services, and applications and create customized maps.

### [Open data 101](#)

Learn what open data is, how to get the most out of it, and explore the principles that are guiding the Government of Canada's open data initiatives.

### [Suggest a dataset](#)

Can't find the dataset that you are looking for? Let us know!

### [Submit your app](#)

Have you made an application using Government of Canada open data or open information? If so, let us know and we will share it with Canadians in our [apps gallery](#).



OPEN GOVERNMENT PROGRAM

## Open Data

Search a listing of Government of Alberta open datasets. Can't find what you are looking for? [Suggest a resource.](#)

Search Open Data

Enter search terms...

Refine Results

Topics

Energy and Natural Resources (1364)

Health and Wellness (357)

Economy and Finance (158)

Population and Demography (137)

Agriculture (128)

Employment and Labour (123)

Business and Industry (105)

Society and Communities (63)

Arts, Culture and History (62)

Government (54)

SHOW MORE

Publisher

Alberta Energy Regulator (1339)

Health (308)

Treasury Board and Finance (243)

Agriculture and Forestry (123)

Economic Development and Trade (80)

2,308 results

Sort Results by

Relevance

Average Weekly Earnings (including overtime), Alberta

OPENDATA

CSV

PDF

LAST MODIFIED

June 29, 2017

DESCRIPTION

This Alberta Official Statistic describes the average weekly earnings of all payroll employees in Alberta. Earnings include overtime and are seasonally adjusted (current dollars).

TAGS

AOS

AWE

ALBERTA OFFICIAL STATISTICS

OFFICIAL STATISTIC

SEPH

Provincial Base Map Series - 1:1 000 000 Scale Maps

OPENDATA

PDF

LAST MODIFIED

June 28, 2017

DESCRIPTION

This cartographic quality series of 1:1 000 000 scale colour maps cover the provincial extent of Alberta. The primary provincial base map displays the Alberta Township System (ATS), major hydrographic features, municipalities, major roads, railways and select geoadministrative features (parks,...

TAGS

CARTOGRAPHY

EARTH SCIENCES

GIS

GEOGRAPHIC DATA

GEOGRAPHY

GEOMATICS

MAP

TOPOGRAPHY

TOPONYMY





Browse Data



Citizen Dashboard



Help



### Solar Potential Images

Explore solar potential images of buildings in Calgary.



### Historic Rainfall Data

Check out our rainfall gauge data since 1998.



### 2015 Normalized Bike Volumes

See 2015 data on two way bike volumes in the downtown core.

# Data Sources for Data Science

Non-GIS and non-statistical data sources, primarily to support the Data Science Certificate

- Home
- Subject-specific data sources
- Directories and search engines
- Licenses and permissions
- Submit your data

## This guide

This guide links to sources of non-GIS and non-statistical data, including research data. It was created primarily to support the Data Science Certificate program, and mostly links to sources for raw data files.

Do you have a resource to suggest for this guide? Email [Jennifer Lee](#), Science Librarian.

For GIS (Geographic Information System), Statistics Canada, and City of Calgary data, see the [Spatial and Numeric Data Services](#) guides.

 Print Page

Search this Guide Search



**Tweepy**

[Read the Docs](#)

[Discuss with Others](#)

[Report a Bug](#)

# Tweepy

**An easy-to-use Python library for accessing the Twitter API.**

 Fork

1,552

 Star

3,630



## The New York Times Developer Network

All the APIs Fit to POST

You already know that NYTimes.com is an unparalleled source of news and information. But now it's a premier source of data, too — why just read the news when you can hack it?

### Getting Started

The Times Developer Network is our API clearinghouse and community. Here's how to get started:

1. Request an [API key](#)
2. Read the API documentation, [FAQ](#) and [Terms of Use](#)
3. Use the API Tool associated with each API to experiment without writing code

### Using The NYT APIs

#### [Terms of Use](#)

Before you can use New York Times APIs, you must agree to the Terms of Use.

#### [Attribution Guidelines and Restrictions](#)

Please review these guidelines before you use our APIs.

#### [Branding Guidelines](#)

Please review the Branding guidelines before you use our APIs, this includes information on the use of our logo, image, and branding components.

#### [API Key Registration](#)

Ready to start coding? Request a key for each API you want to use.

# Spotify Developer

Music, meet code. Powerful APIs, SDKs and widgets for simple and advanced applications



## Spotify Web API

Our Web API lets your applications fetch data from the Spotify music catalog and manage user's playlists and saved music. [Read more](#)



## Web API Console

Quickly test requests to our Web API, get back live JSON data and formatted cURL requests, with full OAuth and scope support.



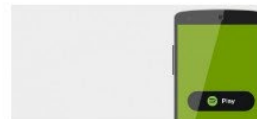
## Spotify Play Button

Add a Spotify Play Button to your blog, website, or social page—all your fans have to do is hit "Play" to enjoy the music. [Read more](#)



## Spotify iOS SDK

(Beta) The Spotify iOS SDK makes it easy to add streaming, user authentication and other Spotify features to iOS apps. [Read more](#)



## Spotify Android SDK

(Beta) The Spotify Android SDK makes it easy to add audio streaming, user authentication and other Spotify features to Android apps. [Read more](#)



## Integrate Spotify Connect into your product

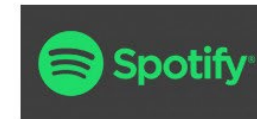
Want to integrate Spotify Connect to give your users a seamless music experience over wifi?

[Read more](#)



## Developer Showcase

Here are just a few of the latest applications that have been built with our APIs, SDKs, and other developer tools. [Read more](#)



## Design Resources

Hello partner. Welcome to our hub for partner guidelines and assets. We want to make it easy for you to integrate Spotify in your app while respecting our brand and legal/licensing restrictions. Note that by using these resources, you accept

[Read More](#)



## Steam Web API Documentation

Valve provides these APIs so website developers can use data from Steam in new and interesting ways. They allow developers to query Steam for information that they can present on their own sites. At the moment the only APIs we offer provide item data for Team Fortress 2, but this list will grow over time.

### Steam Web APIs available

**ISteamNews:** Steam provides methods to fetch news feeds for each Steam game.

**ISteamUserStats:** Steam provides methods to fetch global stat information by game.

**ISteamUser:** Steam provides API calls to provide information about Steam users.

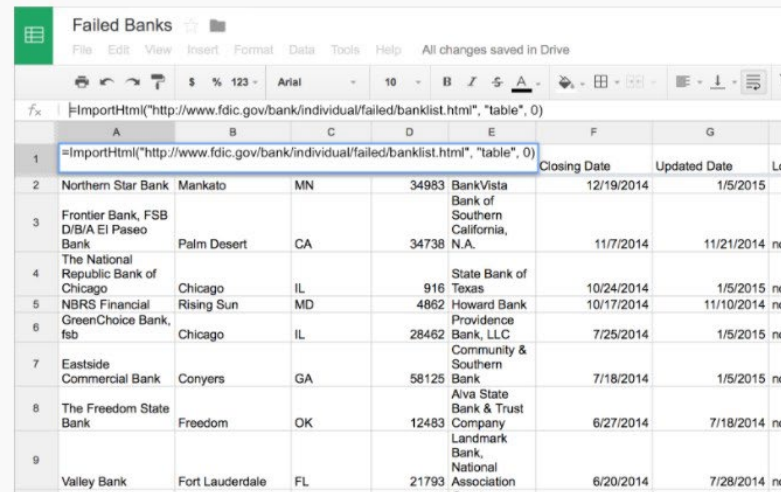
**ITFItems\_440:** Team Fortress 2 provides API calls to use when accessing player item data.

### Obtaining an Steam Web API Key

All use of the Steam Web API requires the use of an API Key. You can acquire one **by filling out this form**. Use of the APIs also requires that you agree to the **Steam API Terms of Use**.

<https://www.fdic.gov/bank/individual/failed/banklist.html>





	A	B	C	D	E	F	G
1	=ImportHtml("http://www.fdic.gov/bank/individual/failed/banklist.html", "table", 0)						
2	Northern Star Bank	Mankato	MN	34983	Bank Vista Bank of Southern California, N.A.	12/19/2014	1/5/2015
3	Frontier Bank, FSB D/B/A El Paseo Bank	Palm Desert	CA	34738		11/7/2014	11/21/2014
4	The National Republic Bank of Chicago	Chicago	IL	916	State Bank of Texas	10/24/2014	1/5/2015
5	NBRS Financial	Rising Sun	MD	4862	Howard Bank	10/17/2014	11/10/2014
6	GreenChoice Bank, fsb	Chicago	IL	28462	Providence Bank, LLC	7/25/2014	1/5/2015
7	Eastside Commercial Bank	Conyers	GA	58125	Community & Southern Bank	7/18/2014	1/5/2015
8	The Freedom State Bank	Freedom	OK	12483	Ava State Bank & Trust Company	6/27/2014	7/18/2014
9	Valley Bank	Fort Lauderdale	FL	21793	Landmark Bank, National Association	6/20/2014	7/28/2014

ROBERT KOSARA / NOVEMBER 15, 2009

## The Simple Way to Scrape an HTML Table: Google Docs

Raw data is the best data, but a lot of public data can still only be found in tables rather than as directly machine-readable files. One example is the FDIC's [List of Failed Banks](#). Here is a simple trick to scrape such data from a website: Use Google Docs.

The table on that page is even relatively nice because it includes some JavaScript to sort it. But a large table with close to 200 entries is still not exactly the best way to analyze that data.

After some digging around – and even considering writing my own throw-away extraction script –, I remembered having read something about [Google Docs](#) being able to [import tables from websites](#). And indeed, it has a very useful function called [ImportHtml](#) that will scrape a table from a page.

### MORE DATA ARTICLES

[Temperature Baseline Differences](#)[A Look At Climate Data](#)[Interactively Explore Climate Data](#)[A Browser for Data.gov](#)[Data Is A Dish Best Served Raw](#)

### RECENTLY POPULAR

[Understanding Pie Charts](#)[The Simple Way to Scrape an HTML](#)[Table: Google Docs](#)[Data: Continuous vs. Categorical](#)[What is Visualization? A Definition](#)[You Only See Colors You Can Name](#)[EuroVis 2017 Conference Report. Part 2](#)[EuroVis 2017 Conference Report. Part 1](#)

### SUBSCRIBE VIA EMAIL

# Data Formats

# Data Formats

---

- Structured Data:
- Unstructured Data:
- Semi-structured Data:

# Data Formats

---

- Structured Data: spreadsheet or database
- Unstructured Data:
- Semi-structured Data:



# Data Formats

---

- Structured Data: spreadsheet or database
- Unstructured Data: raw text, streaming data, images, videos
- Semi-structured Data:

# Data Formats

---

- Structured Data: spreadsheet or database
- Unstructured Data: raw text, streaming data, images, videos
- Semi-structured Data: somewhere in between

# CSV (Comma-Separated Value)

---

```
1 firstName,lastName,age,streetAddress,city
2 John,Smith,25,21 2nd Street,New York,NY,10021
```

firstName	lastName	age	streetAddress	city	state	postalCode	homePhoneNumber	faxPhoneNumber
John	Smith	25	21 2nd Street	New York	NY	10021	212 555-1239	646 555-4567

# JSON (JavaScript Object Notation)

```
{  
  "firstName": "John",  
  "lastName": "Smith",  
  "age": 25,  
  "address": {  
    "streetAddress": "21 2nd Street",  
    "city": "New York",  
    "state": "NY",  
    "postalCode": "10021"  
  },  
  "phoneNumber": [  
    {  
      "type": "home",  
      "number": "212 555-1239"  
    },  
    {  
      "type": "fax",  
      "number": "646 555-4567"  
    }  
  ]  
}
```

firstName	lastName	age	streetAddress	city	state	postalCode	homePhoneNumber	faxPhoneNumber	gender
John	Smith	25	21 2nd Street	New York	NY	10021	212 555-1239	646 555-4567	male

# XML (eXtensible Markup Language)

```
<person>
  <firstName>John</firstName>
  <lastName>Smith</lastName>
  <age>25</age>
  <address>
    <streetAddress>21 2nd Street</streetAddress>
    <city>New York</city>
    <state>NY</state>
    <postalCode>10021</postalCode>
  </address>
  <phoneNumbers>
    <phoneNumber type="home">212 555-1234</phoneNumber>
    <phoneNumber type="fax">646 555-4567</phoneNumber>
  </phoneNumbers>
  <gender>
```

firstName	lastName	age	streetAddress	city	state	postalCode	homePhoneNumber	faxPhoneNumber	gender
John	Smith	25	21 2nd Street	New York	NY	10021	212 555-1239	646 555-4567	male

# Handling Data

# Storing Data

# Storing Data

---

- Backup
- Password protect or encrypt sensitive information



# Provenance

---

- Keep track of where and when data was collected
- Record data collection / processing steps

# IP and Copyright

---

- Know who owns the data
- Do you need to publish or (re)share data?
- Do NOT violate copyright

# Privacy

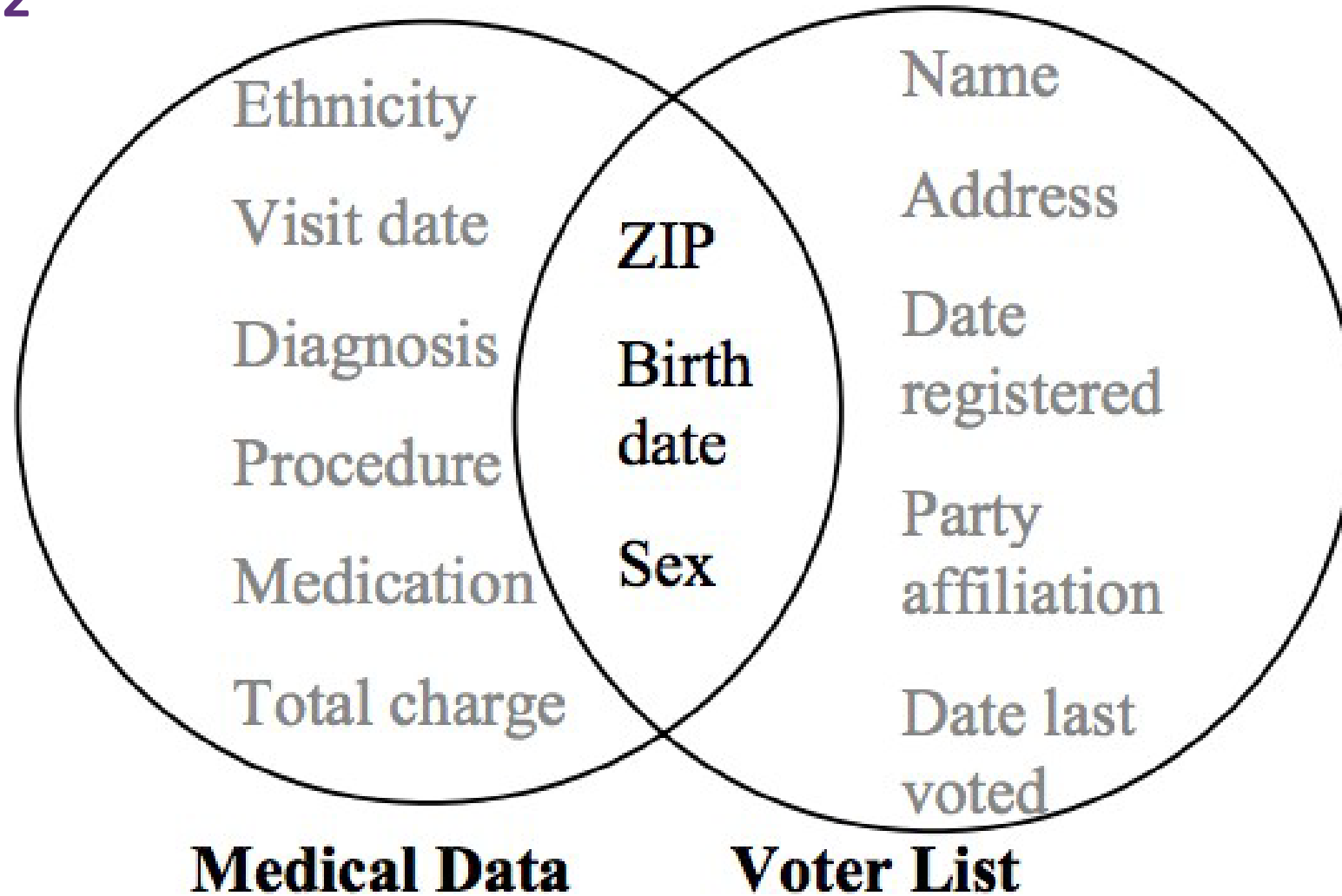
---

- Any information that could be used to identify individuals is sensitive
- May have legal repercussions for releasing it
- Anonymize data before sharing

# How to anonymize data?

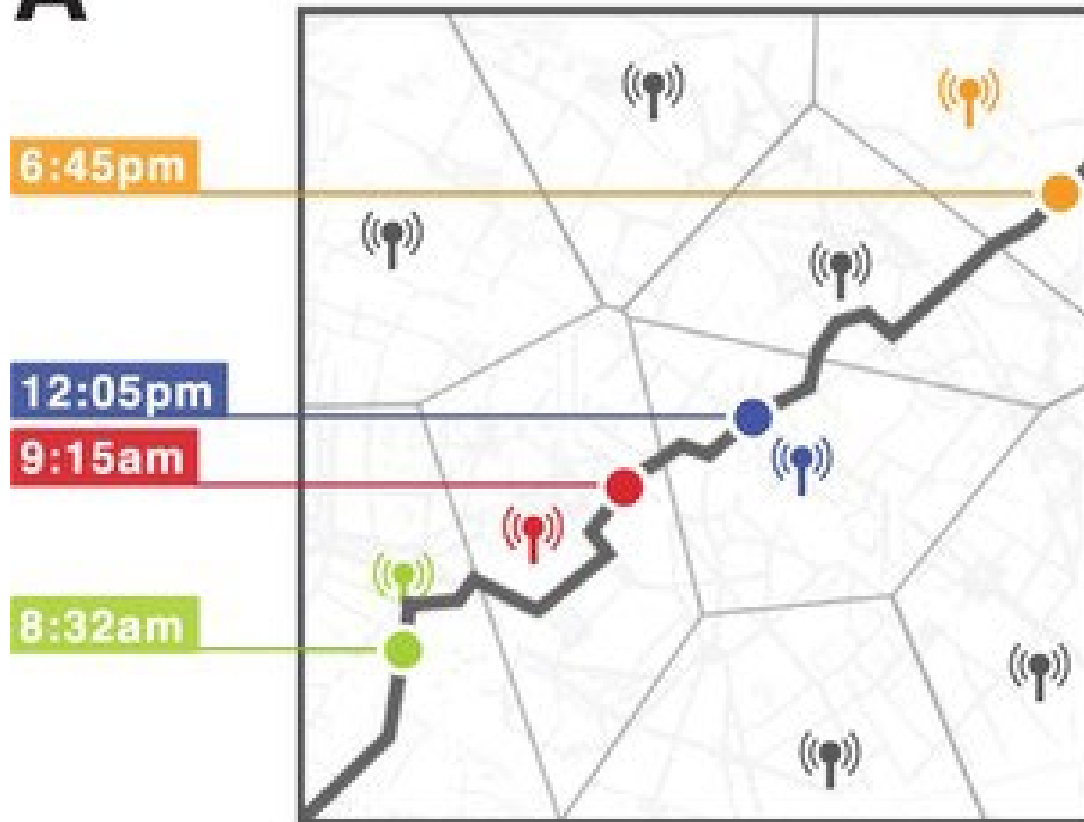
## L. Sweeney. 2002

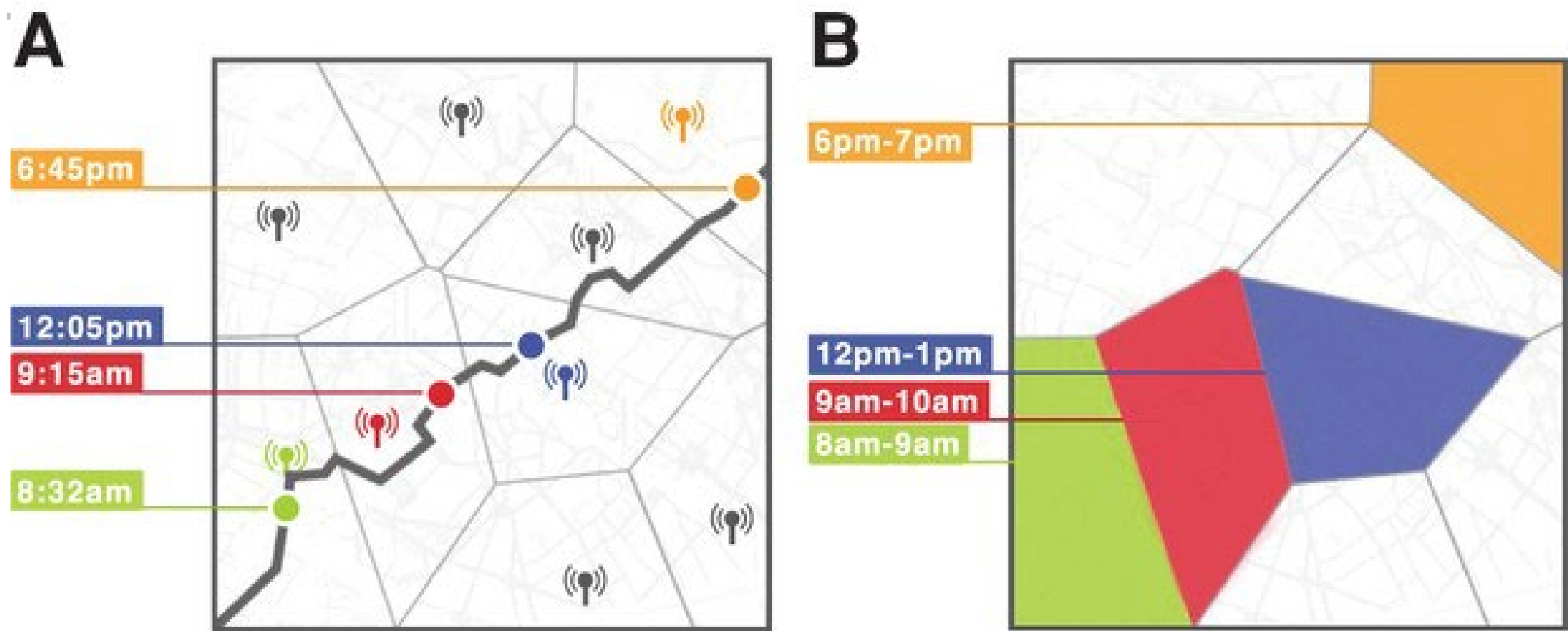
---



de Montjoye et al. 2013

A







# Ethics

WEATHER



NORDSTR



# Onward to ... Data Cleaning

---

Jonathan Hudson  
[jwhudson@ucalgary.ca](mailto:jwhudson@ucalgary.ca)  
<https://pages.cpsc.ucalgary.ca/~jwhudson/>



UNIVERSITY OF  
CALGARY