Actionable Data: the Wolfram Approach

Matthew Day
matthewday@wolfram.com

Manager, Data Repository
Wolfram Partnerships
Wolfram Research

WOLFRAM

Introduction to Wolfram Research

- ***Mathematica*** **is a computation engine.**
- **Wolfram|Alpha is a knowledge engine. Natural language queries + data + computation.**
- **CDF is a technology for creating and sharing user-friendly, interactive documents.**

If a question can be answered on the basis of human knowledge, a computer should be able to provide that answer.

Wolfram|Alpha

A knowledge engine for working with:

1. **Curated facts.**

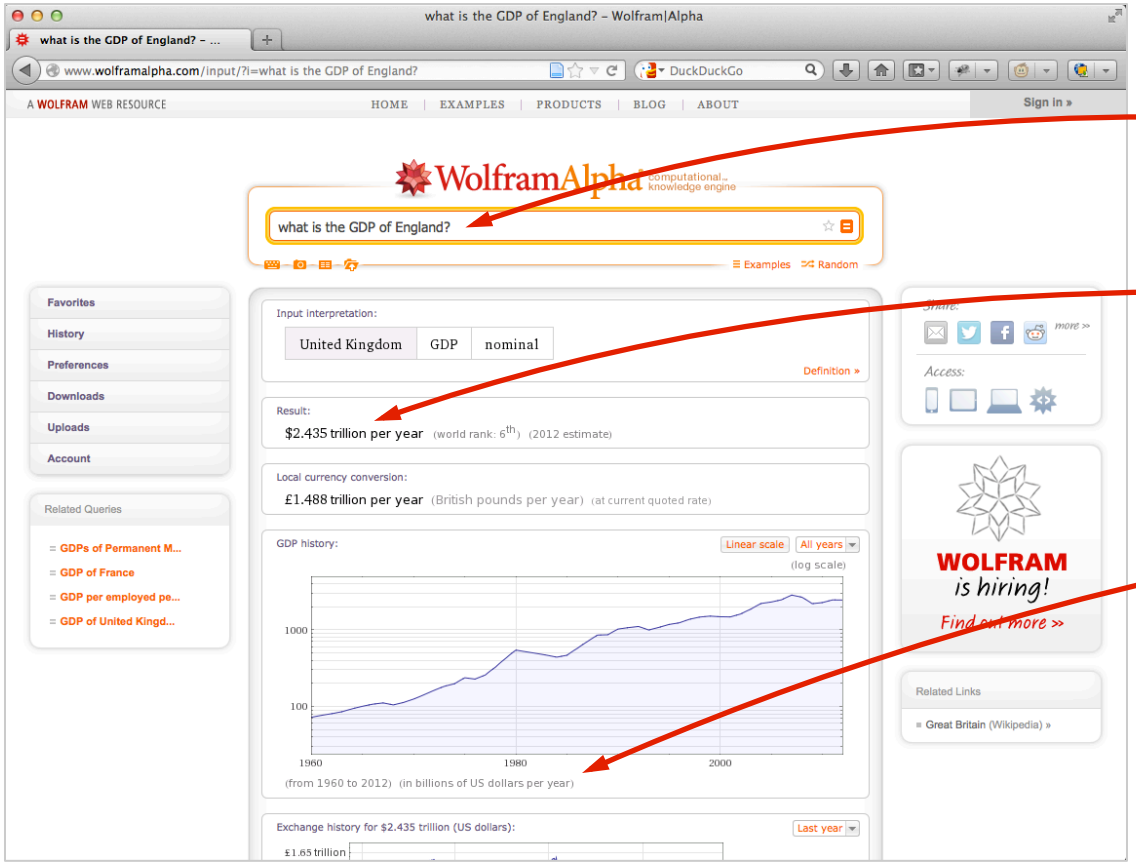
2. **Models.**

3. **Random pieces of data.**

Simple facts

What is the GDP of England?

wolframalpha.com/input/?i=what is the GDP of England?



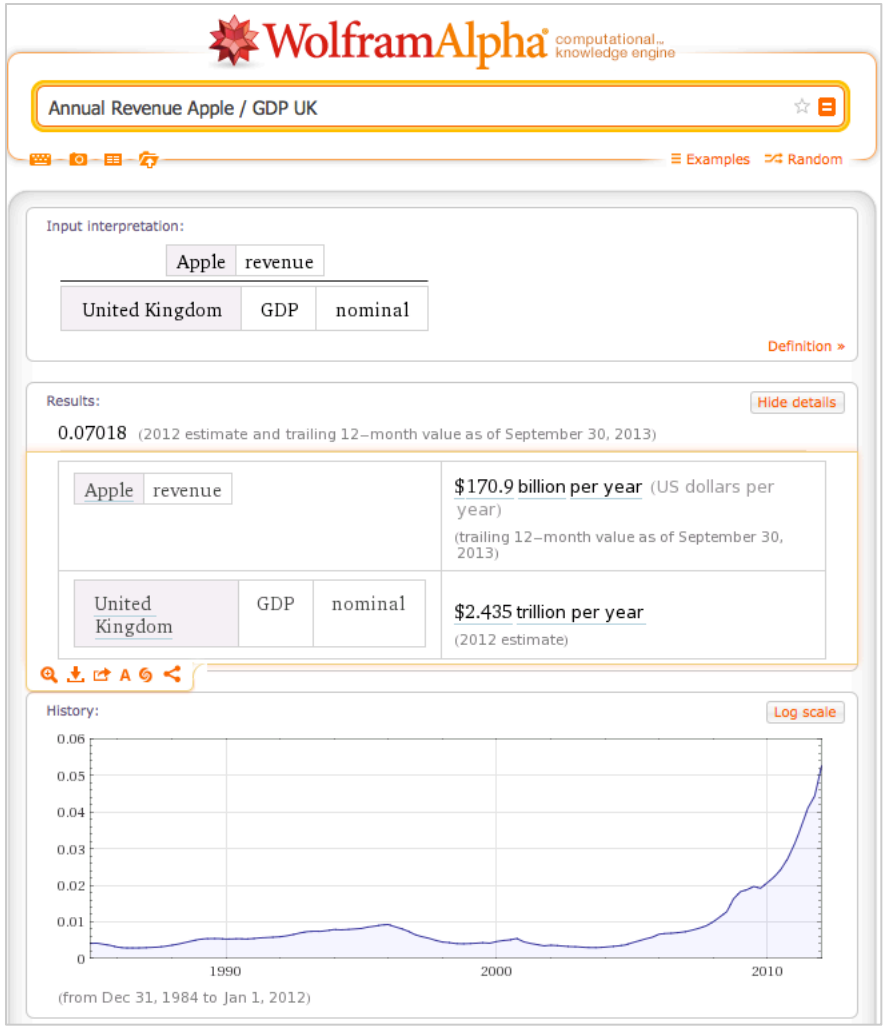
natural language question

the answer

an automated report giving context to the answer

WOLFRAM

Computed facts

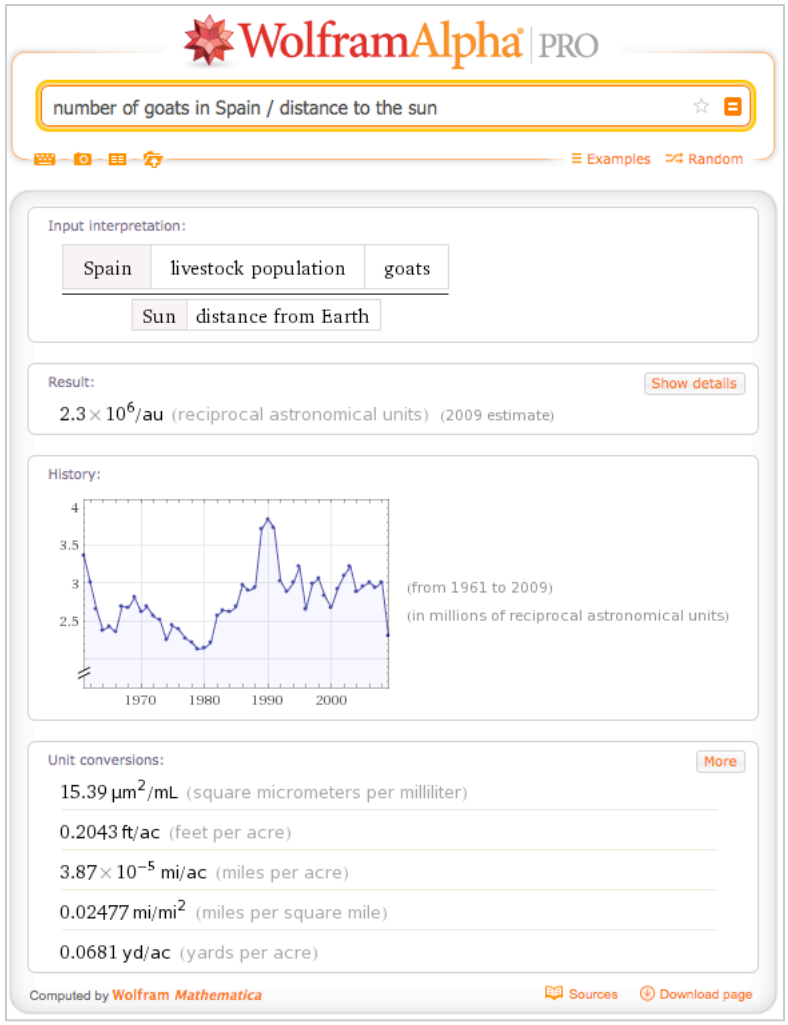


Revenue of Apple / GDP of the UK?

To answer this, Wolfram|Alpha needs to know about currencies and exchange rates.

WOLFRAM

Everything is computed on the fly

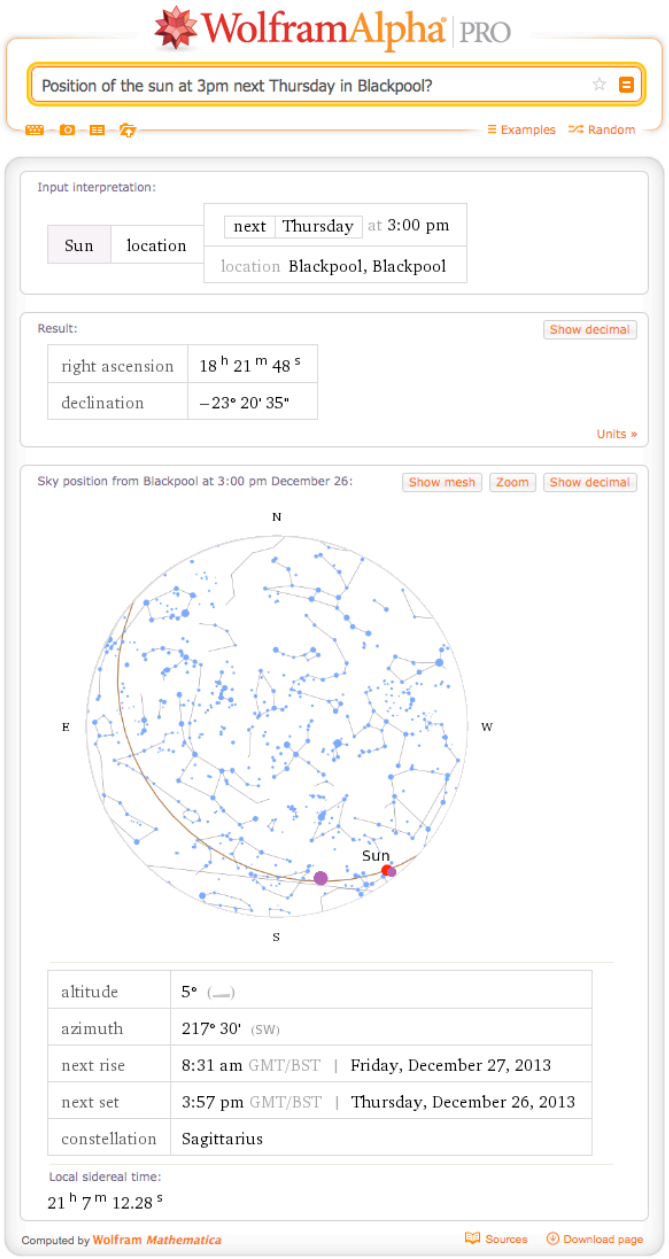


Number of goats in Spain / distance to the sun.

We didn't foresee this question, but it is answerable.

WOLFRAM

Models of the world

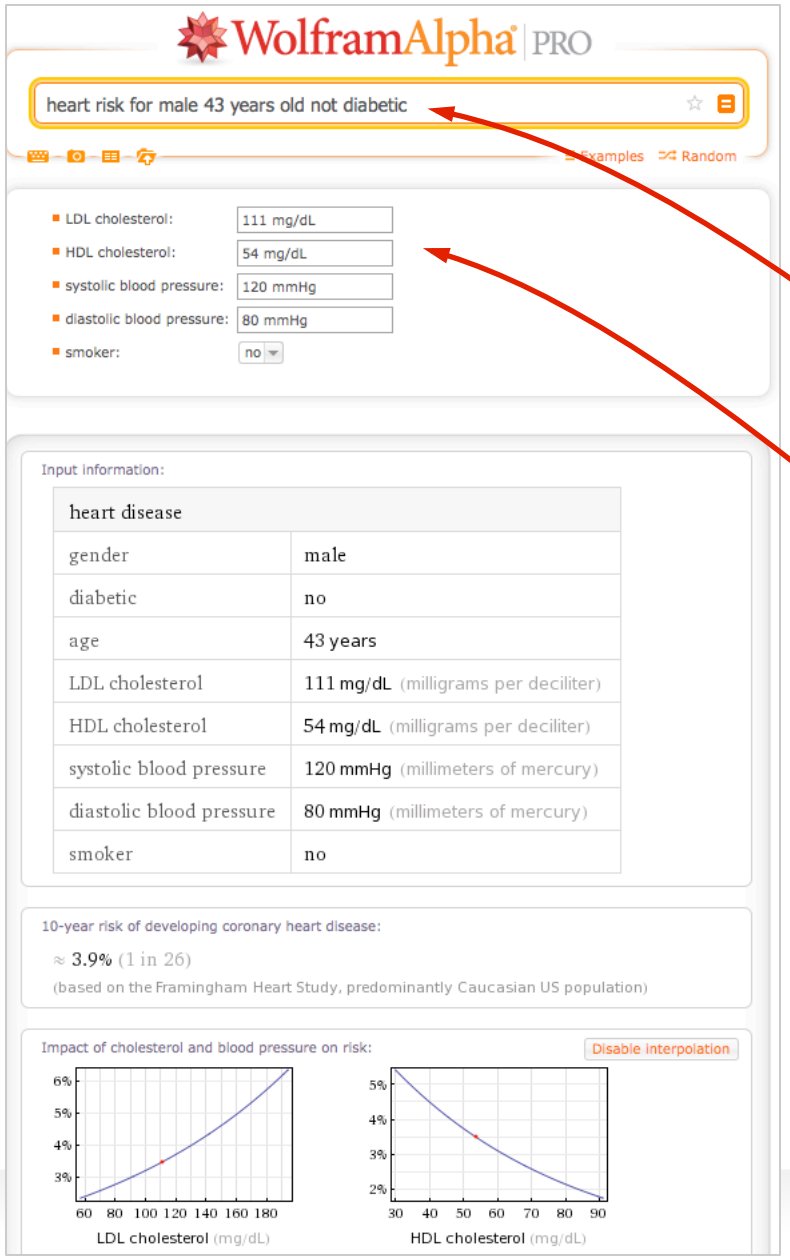


"Position of the sun" ?
"3pm next Thursday" ?
"in Blackpool" ?

Wolfram|Alpha needs to know quite a lot about the world to compute this answer.

WOLFRAM

A biomedical model



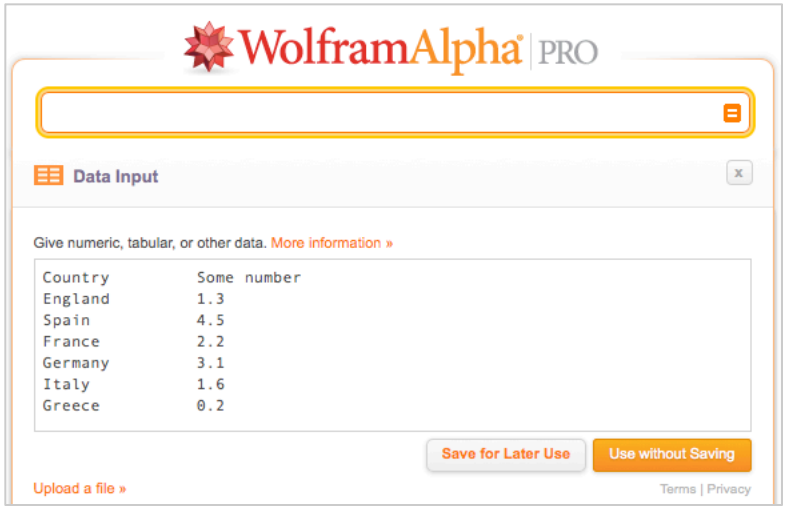
Wolfram|Alpha has been taught a model of heart disease risk in humans (from the Framingham Heart Study).

I gave some of the required information in the question …

and editable assumptions have been provided for the other variables.

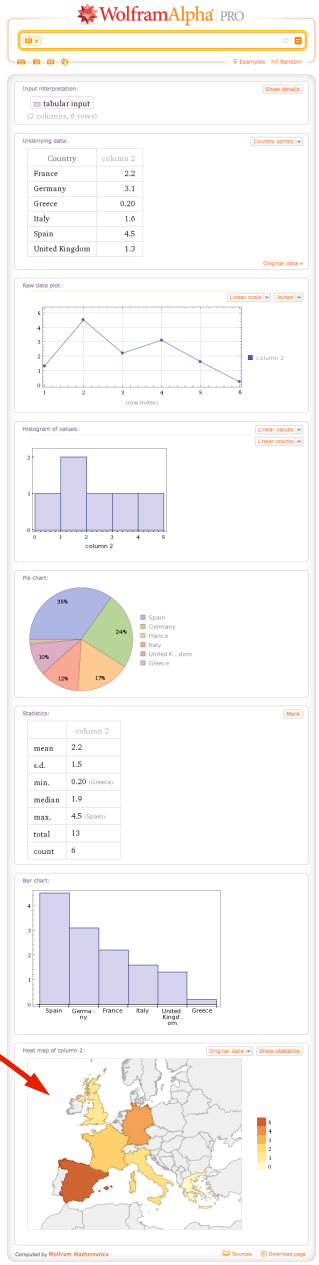
WOLFRAM

Random data



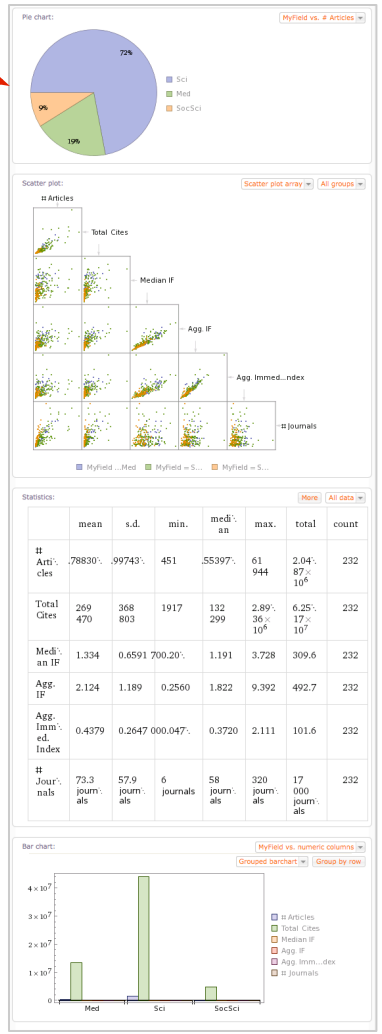
I gave Wolfram|Alpha some random data ...

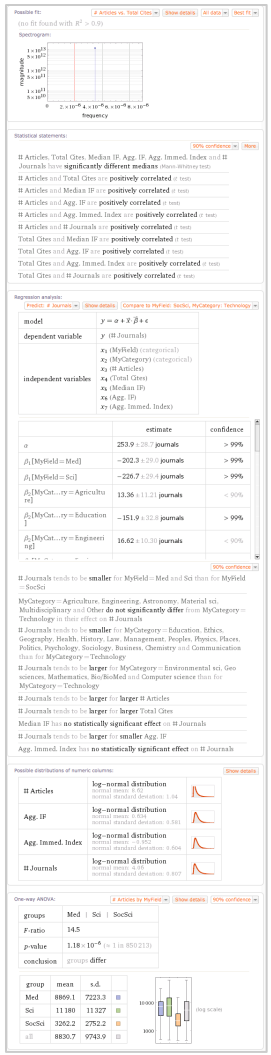
... and I got back something useful, including a nice coloured map that I could use in a publication.



Another data example

	A	B	C	D	E	F	G	H	I	J
1	MyField	MyCategory	ISI Category	# Articles	Total Cites	Median IF	Agg. IF	Agg. Immed. Index	Agg. Cited 1/2 Life	# Journals
35	Med	Clinical	EMERGENCY MEDICINE	3465	67594	1.081	1.76	0.355	6.7	24
36	Med	Clinical	TROPICAL MEDICINE	2979	62811	1.107	2.096	0.372	6.9	20
37	Med	Clinical	MEDICAL LABORATORY TECHNOLOGY	2960	76496	1.319	2.103	0.438	7.4	32
38	Med	Clinical	INTEGRATIVE & COMPLEMENTARY MEDICINE	2628	41091	1.329	2.195	0.291	6.1	22
39	Med	Clinical	MEDICAL INFORMATICS	2083	46210	1.516	1.782	0.438	7.1	23
40	Med	Clinical	AUDIOLOGY & SPEECH-LANGUAGE PATHOLOGY	1955	64894	1.55	1.744	0.274	>10.0	21
41	Med	Clinical	MEDICINE, LEGAL	1660	25546	1.229	1.82	0.28	6.4	15
42	Med	Clinical	SUBSTANCE ABUSE	1649	57124	2.359	2.847	0.598	7.2	14
43	Med	Clinical	PRIMARY HEALTH CARE	1084	19994	1.269	1.462	0.744	6.8	14
44	Med	Clinical	MEDICAL ETHICS	676	8391	0.917	1.264	0.731	5.1	17
45	Med	Clinical	ANDROLOGY	477	9840	1.535	2.222	0.499	6.8	6
46	Sci	Agriculture	FOOD SCIENCE & TECHNOLOGY	18470	421932	1.165	1.898	0.315	7.2	128
47	Sci	Agriculture	AGRONOMY	7485	171793	1	1.452	0.281	8.5	79
48	Sci	Agriculture	AGRICULTURE, MULTIDISCIPLINARY	6068	125325	0.495	1.439	0.25	7.8	57
49	Sci	Agriculture	AGRICULTURE, DAIRY & ANIMAL SCIENCE	5494	121768	0.698	1.296	0.228	8.6	55
50	Sci	Agriculture	FISHERIES	4661	129802	1.177	1.548	0.301	9.1	48
51	Sci	Agriculture	FORESTRY	4068	96892	0.868	1.557	0.324	8.1	59
52	Sci	Agriculture	SOIL SCIENCE	3659	132592	1.53	1.78	0.365	9.8	33
53	Sci	Agriculture	AGRICULTURAL ENGINEERING	3344	63520	1.01	3.193	0.616	4.9	12
54	Sci	Agriculture	HORTICULTURE	3156	71876	0.75	1.28	0.202	9.1	31
55	Sci	Agriculture	LIMNOLOGY	1840	67019	1.207	1.843	0.503	>10.0	19
56	Sci	Agriculture	AGRICULTURAL ECONOMICS & POLICY	536	10575	0.893	1.069	0.17	8.7	14
57	Sci	Astronomy	ASTRONOMY & ASTROPHYSICS	16525	680837	1.683	4.242	1.439	6.8	56
58	Sci	Bio/BioMed	BIOCHEMISTRY & MOLECULAR BIOLOGY	51425	2893583	2.857	4.276	0.873	7.7	289
59	Sci	Bio/BioMed	PHARMACOLOGY & PHARMACY	33464	1085783	2.153	2.943	0.541	6.7	261
60	Sci	Bio/BioMed	NEUROSCIENCES	33223	1666914	2.759	3.951	0.782	7.4	243
61	Sci	Bio/BioMed	CELL BIOLOGY	23181	1581141	3.271	5.779	1.186	6.9	180





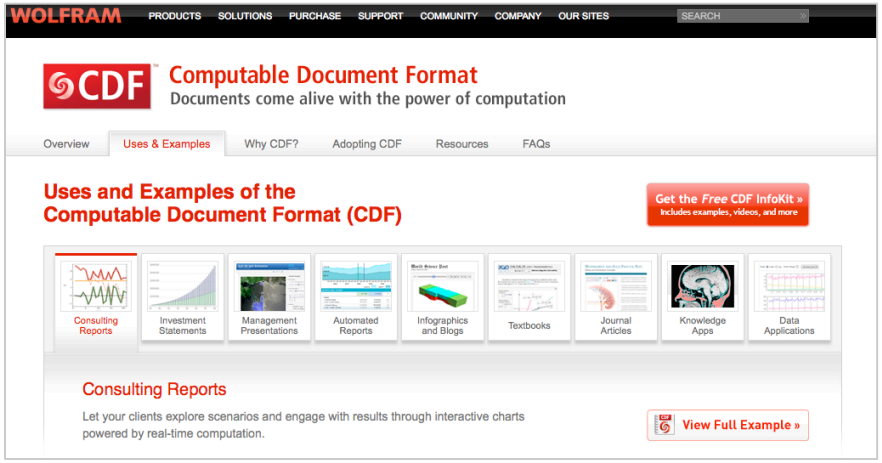
This example is of some more extensive, real-life data on journal impact factors.

We get back a lot of automated analysis, much of which is genuinely useful. It includes some natural language statements that can be drawn from the data.

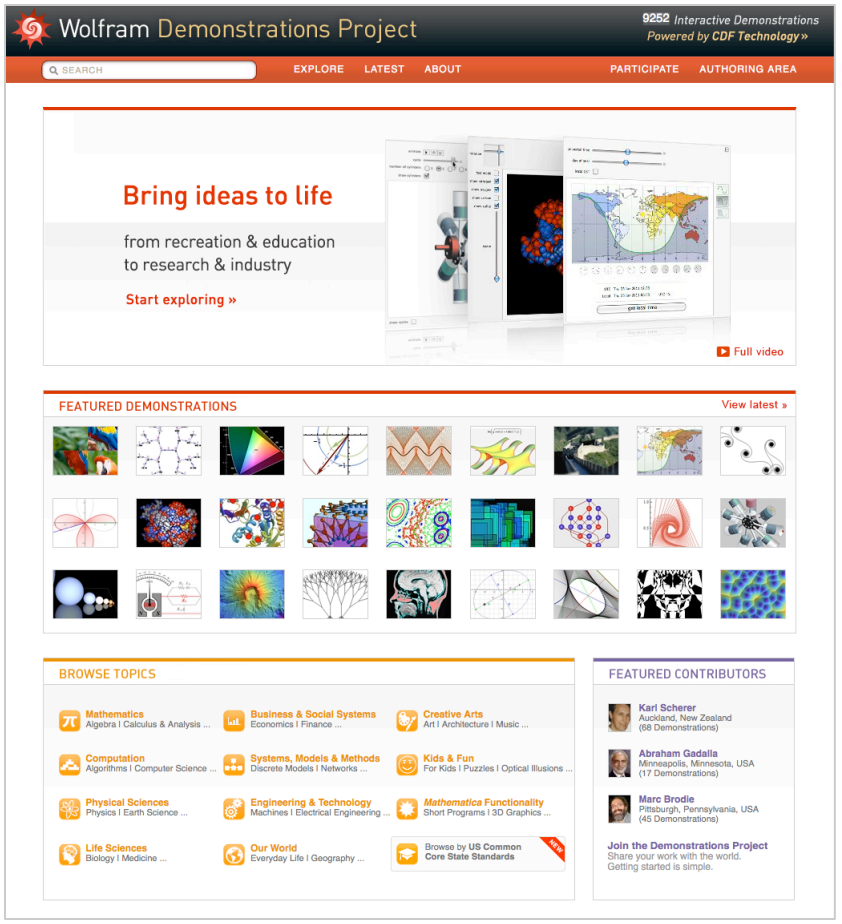
WOLFRAM

Publishing interactive content

CDF (Computable Document Format)
wolfram.com/cdf



CDF Demonstrations
demonstrations.wolfram.com



WOLFRAM

The future …

- More powerful data analysis, yet still easy to use.
- New mechanisms for publishing actionable data.
- Cloud-based, plug-in free.