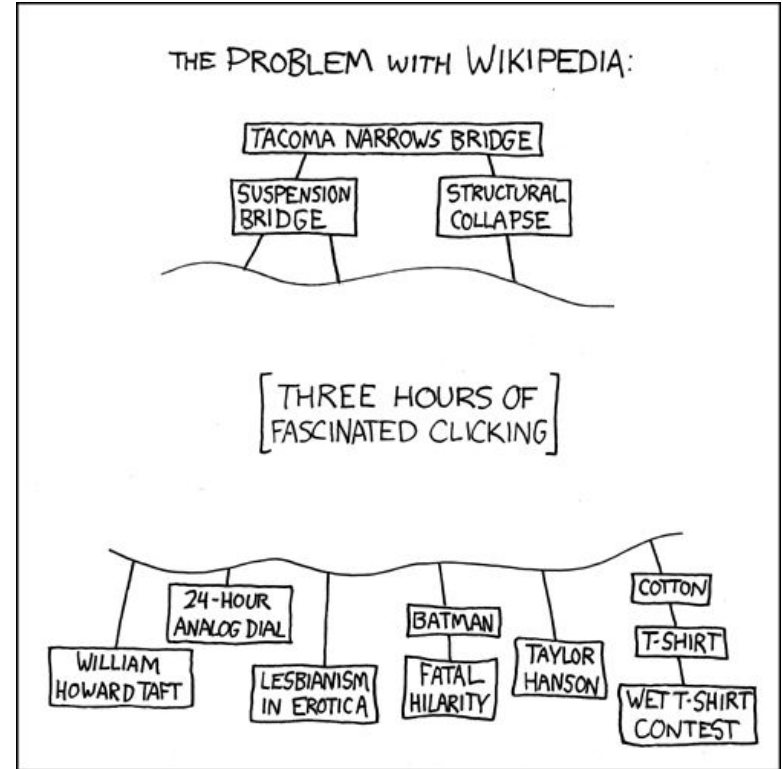


ROOT DIRECTORY

SINGLE-SOURCE TRUTH FOR WIKIPEDIA CATEGORIES

BUSINESS PROBLEM

THE RABBIT HOLE OF DEATH



TEN TOO MANY CATEGORY TOOLS

1. 'Portal' pages
2. 'Outline' pages
3. 'Indices' pages
4. 'Areas of' pages
5. 'Contents' pages
6. 'Category' pages
7. 'Glossaries' pages
8. 'Overviews' pages
9. 'Contents/Lists' page
10. 'Category: Lists' page
11. 'Lists of lists of lists' page

SIMPLE QUESTIONS HARD ANSWERS

How many articles on Wikipedia are
related to mathematics?

SIMPLE QUESTIONS HARD ANSWERS

How many articles on Wikipedia are related to mathematics?

PORTAL ESTIMATE

31,444

Portal:Mathematics

From Wikipedia, the free encyclopedia

Portal topics: [Activities](#) · [Culture](#) · [Geography](#) · [Health](#) · [History](#) · [Mathematics](#) · [Nature](#) · [People](#) · [Technology](#) · [Random portal](#)

The Mathematics Portal

[Shortcut](#)
[P:MAT](#)

Mathematics is the study of numbers, quantity, space, structure, and change. Mathematics is used throughout the world as an essential tool in many fields, including [natural science](#), engineering, medicine, and the social sciences. Applied mathematics, the branch of mathematics concerned with application of mathematical knowledge to other fields, inspires and makes use of new mathematical discoveries and sometimes leads to the development of entirely new mathematical disciplines, such as statistics and [game theory](#). Mathematicians also engage in pure mathematics, or mathematics for its own sake, without having any application in mind. There is no clear line separating pure and applied mathematics, and practical applications for what began as pure mathematics are often discovered.

[More about Mathematics...](#)

[Selected article](#) | [Selected picture](#) | [Did you know...](#) | [Topics in mathematics](#)
[Categories](#) | [WikiProjects](#) | [Things you can do](#) | [Index](#) | [Related portals](#)

There are approximately 31,444 mathematics articles in Wikipedia

[Refresh with new selections below](#) (purge)

SIMPLE QUESTIONS HARD ANSWERS

How many articles on Wikipedia are related to mathematics?

Portal estimate: 31,444

WIKIDATA ESTIMATE 23,051

```
1 SELECT distinct ?item ?article ?sitelink ?linkTo WHERE {
2   { ?item wdt:P361* wd:Q395 .}
3   union
4   { ?item wdt:P361/wdt:P279* wd:Q395 .}
5   union
6   { ?item wdt:P31/wdt:P279* wd:Q1936384 .}
7   union
8   { ?item wdt:P921/wdt:P279* wd:Q395 .}
9   optional
10  { ?sitelink ^schema:name ?article .
11    ?article schema:about ?item ;
12      schema:isPartOf <https://en.wikipedia.org/> .
13  }
14  OPTIONAL { ?item wdt:P361 ?linkTo . }
15  SERVICE wikibase:label { bd:serviceParam wikibase:language "en". }
16 }
```

23051 results in 3590 ms

</> Code Download Link

SIMPLE QUESTIONS HARD ANSWERS

How many articles on Wikipedia are related to mathematics?

Portal estimate: 31,444

Wikidata estimate: 23,051

PetScan ESTIMATE
34,425

Do it!

Link to a pre-filled form for *the query you just ran* [with](#) and [without](#) auto-run. PSID is **6752301**. [Find images for Wikidata.](#)

34425 results

Only the first 1000 results are shown in HTML, so as to not crash your browser; other formats will have complete results.

#	Title	Page ID	Namespace	Size (bytes)	Last change
1	Aristotle	308	(Article)	132294	20181210161637
2	Alain Connes	340	(Article)	7499	20181207081309
3	Arithmetic mean	612	(Article)	12361	20181209191936
4	Abacus	655	(Article)	50675	20181210161637
5	Ambiguity	677	(Article)	27644	20181204200313
6	Albert Einstein	736	(Article)	154667	20181210205058
7	Algorithms (journal)	742	(Article)	3273	20181120014534
8	Algorithm	775	(Article)	101321	20181210161637

DAFUQ

SOLUTION

**SINGLE
SOURCE OF
TRUTH**

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression		0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression		0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression		0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression		0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression		0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

Simple Wiki Article

TITLE	Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression	0.92	0.74	0.2

OBJECTIVE:

TRUTH TABLE

TITLE	Wiki Article	Simple Wiki Article	Probability Score	Complexity Score	Quality Score
Logistic Regression	https://en.wikipedia.org/wiki/Logistic_regression	https://simple.wikipedia.org/wiki/Logistic_Regression	0.92	0.74	0.2

MODELING

TF-IDF

MULTINOMIAL NAIVE BAYES

LOGISTIC REGRESSION

FEATURE IMPORTANCE

VERY MATHY WORDS

MATHEMATICIAN

THEOREM

NUMBER

ALGORITHM

VERY NOT MATHY WORDS

ENGINEER

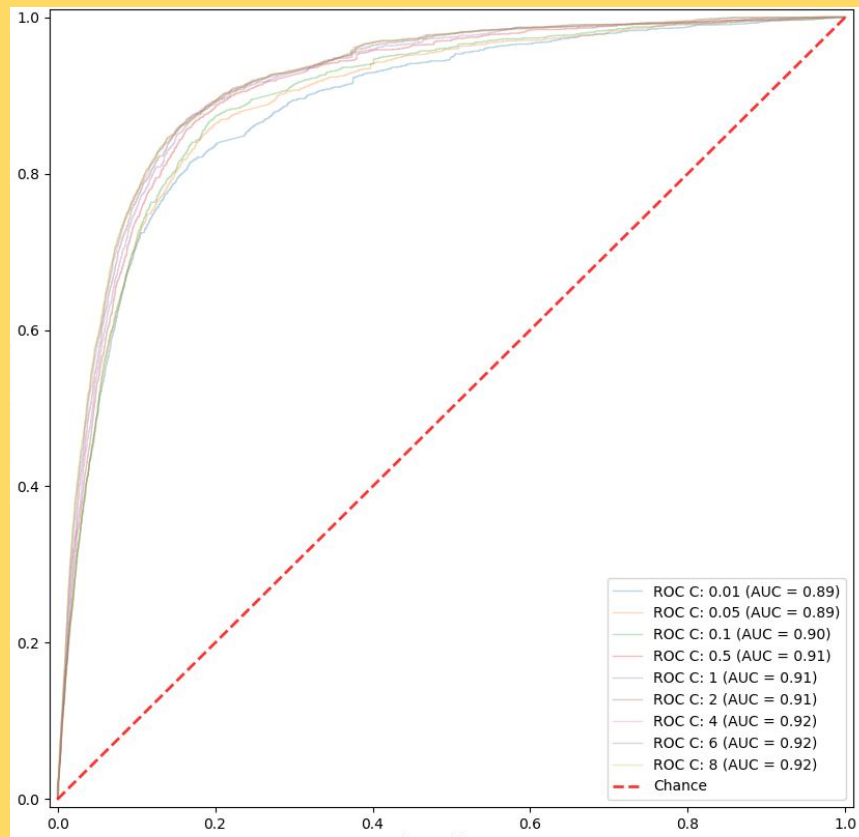
PHYSICS

ART

CHEMISTRY

ROC AUC

TRUE POSITIVE RATE



FALSE POSITIVE RATE

CONFUSION MATRIX @ 0.20

TRUE NEGATIVES

11792

FALSE POSITIVES

2780

FALSE NEGATIVES

277

TRUE POSITIVES

1823

FALSE POSITIVES

	title	predicted	actual	threshold_pred
2760	Criss-cross algorithm	0.952502	0.0	True
2761	Vector logic	0.953342	0.0	True
2762	Category:Biological theorems	0.954540	0.0	True
2763	Chaitin's constant	0.955203	0.0	True
2764	Computational complexity	0.958522	0.0	True
2765	P versus NP problem	0.959869	0.0	True
2766	Non-constructive algorithm existence proofs	0.960831	0.0	True
2767	Generic-case complexity	0.962550	0.0	True
2768	Time complexity	0.962793	0.0	True
2769	Computing the permanent	0.964047	0.0	True
2770	Integer factorization	0.965000	0.0	True
2771	Computable number	0.971522	0.0	True
2772	Theory of computation	0.971899	0.0	True
2773	Discrete logarithm	0.972355	0.0	True
2774	Linear programming	0.972618	0.0	True
2775	Existential theory of the reals	0.978830	0.0	True
2776	Computable function	0.980849	0.0	True
2777	Computational complexity theory	0.982427	0.0	True
2778	Numerical analysis	0.983520	0.0	True

FALSE POSITIVES

		predicted	actual	threshold_pred
	title	0.952502	0.0	
2760	Criss-cross algorithm	0.953342	0.0	True
2761	Vector log	0.954540	0.0	True
2762	Category:Biological theorem	0.955203	0.0	True
2763	Chaitin's constant	0.958522	0.0	True
2764	Computational complexity	0.959869	0.0	True
2765	P versus NP problem	0.960831	0.0	True
2766	Non-constructive algorithm existence proof	0.962550	0.0	True
2767	Generic-case complexity	0.962793	0.0	True
2768	Time complexity	0.964047	0.0	True
2769	Computing the permanent	0.965000	0.0	True
2770	Integer factorization	0.965000	0.0	True
2771	Computable number	0.971522	0.0	True
2772	Theory of computation	0.971899	0.0	True
2773	Discrete logarithm	0.972355	0.0	True
2774	Linear programming	0.972618	0.0	True
2775	Existential theory of the reals	0.978830	0.0	True
2776	Computable function	0.978830	0.0	True
2777	Computational complexity theory	0.980849	0.0	True
2778	Numerical analysis	0.982427	0.0	
		0.983520	0.0	

FALSE

NEGATIVES

	title	predicted	actual	threshold_pred
0	A Guide-Book to Mathematics for Technologists ...	0.008715	1.0	False
1	Nuria Juncosa	0.011367	1.0	False
2	Lin Hsin Hsin	0.011919	1.0	False
3	NinKi: Urgency of Proximate Drawing Photograph	0.013497	1.0	False
4	Hidden Figures	0.015482	1.0	False
5	Category:Cryptography organizations	0.016806	1.0	False
6	Homeokinetics	0.023209	1.0	False
7	Verification and validation	0.024878	1.0	False
8	David Suter	0.025197	1.0	False
9	Dunham expansion	0.028177	1.0	False
10	Category:Cryptography books	0.028557	1.0	False
11	Biconcave disc	0.030608	1.0	False
12	The Road to Reality	0.034167	1.0	False
13	Possible Worlds (play)	0.035817	1.0	False
14	Ratchet effect	0.037957	1.0	False
15	Desmond Paul Henry	0.039584	1.0	False
16	Everything	0.039781	1.0	False
17	Open Energy Modelling Initiative	0.042159	1.0	False
18	Rotational–vibrational coupling	0.043692	1.0	False
19	Systems medicine	0.045823	1.0	False

FALSE NEGATIVES

16

	title	predicted	actual	threshold_pred
0	A Guide-Book to Mathematics for Technologists ...	0.008715	1.0	False
1	Nuria Juncosa	0.011367	1.0	False
2	Lin Hsin Hsin	0.011919	1.0	False
3	NinKi: Urgency of Proximate Drawing Photograph	0.013497	1.0	False
4	Hidden Figures	0.015482	1.0	False
5	Category:Cryptography organizations	0.016806	1.0	False
6	Homeokinetics	0.023209	1.0	False
7	Verification and validation	0.024878	1.0	False
8	David Suter	0.025197	1.0	False
9	Dunham expansion	0.028177	1.0	False
10	Category:Cryptography books	0.028557	1.0	False
11	Biconcave disc	0.030608	1.0	False
12	The Road to Reality	0.034167	1.0	False
13	Possible Worlds (play)	0.035817	1.0	False
14	Ratchet effect	0.037957	1.0	False
15	Desmond Paul Henry	0.039584	1.0	False
	Everything	0.039781	1.0	False
17	Open Energy Modelling Initiative	0.042159	1.0	False
18	Rotational–vibrational coupling	0.043692	1.0	False
19	Systems medicine	0.045823	1.0	False

FALSE

NEGATIVES

	title	predicted	actual	threshold_pred
258	Gentleman's Diary	0.191730	1.0	False
259	American Regions Mathematics League	0.191816	1.0	False
260	Thomas Rawson Birks	0.191820	1.0	False
261	Bruce Irons (engineer)	0.192257	1.0	False
262	Portal:Number theory	0.193504	1.0	False
263	Category:Set families	0.193807	1.0	False
264	Everything and More (book)	0.194760	1.0	False
265	List of things named after Hermann Grassmann	0.195324	1.0	False
266	Category:Operations researchers	0.195498	1.0	False
267	Category:Coordinate systems by dimensions	0.195721	1.0	False
268	Book:Grandi's series	0.195768	1.0	False
269	Portal:Analysis	0.196186	1.0	False
270	The New York Journal of Mathematics	0.196348	1.0	False
271	Hobbes–Wallis controversy	0.197043	1.0	False
272	Template:Grandi's series	0.197455	1.0	False
273	Willem Klein	0.197694	1.0	False
274	Category:Proof without words	0.197800	1.0	False
275	Swim lane	0.198298	1.0	False
276	Echo removal	0.199882	1.0	False
277	NaN	NaN	NaN	NaN

FALSE

NEGATIVES

264

Everything and More (book)

0.194760

1.0

False

	title	predicted	actual	threshold_pred
258	Gentleman's Diary	0.191730	1.0	False
259	American Regions Mathematics League	0.191816	1.0	False
260	Thomas Rawson Birks	0.191820	1.0	False
261	Bruce Irons (engineer)	0.192257	1.0	False
262	Portal:Number theory	0.193504	1.0	False
263	Category:Set families	0.193907	1.0	False
264	Everything and More (book)	0.194760	1.0	False
265	List of things named after Hermann Grassmann	0.195324	1.0	False
266	Category:Operations researchers	0.195498	1.0	False
267	Category:Coordinate systems by dimensions	0.195721	1.0	False
268	Book:Grandi's series	0.195768	1.0	False
269	Portal:Analysis	0.196186	1.0	False
270	The New York Journal of Mathematics	0.196348	1.0	False
271	Hobbes–Wallis controversy	0.197043	1.0	False
272	Template:Grandi's series	0.197455	1.0	False
273	Willem Klein	0.197694	1.0	False
274	Category:Proof without words	0.197800	1.0	False
275	Swim lane	0.198298	1.0	False
276	Echo removal	0.199882	1.0	False
277	NaN	NaN	NaN	NaN

SO, HOW
MANY MATH
ARTICLES ARE ON
WIKIPEDIA?

4.5%

~235,000 mathematics related articles

~5,800,000 total articles

JAKUB SVEC

www.github.com/jakubsvec001

www.linkedin.com/jakubsvec001

jakubsvec001@gmail.com