
Wikidata

**Structured data repository that
anyone can edit**

— Filip Maljković —
PSSOH, October 26, 2019

Two main problems with Wikipedia

1. Outdated data; no triggers for change; boring and repetitive work
 - There's a large number of Wikipedias; some will have updated info, some won't
2. Inflexible ways for fetching specific data
 - Limits of categories
 - Example: all authors whose parents were born in Paris

Two main problems with Wikipedia

One solution for both problems:

Wikidata - central multilingual repository for linked and structured data, on a wiki, that anyone can edit, under a free license



WIKIDATA

Structured data

Statement = Item: property \rightarrow value

- Item (Q) - any topic (person, object, place, concept)
 - Property (P) - specific type of data relevant for an item (e.g. *gender* for people, *capital* for countries, *length* for rivers)
 - Value - either a “literal value” (e.g. length of Danube is 2860 km), or a reference to another item (e.g. capital of Serbia is Belgrade)
-
- Item is described by a set of statements (each statement provides a fact or a data point about an item)

Structured data

- Earth (Q2): highest point (P610) → Mount Everest (Q513)
- Mount Everest (Q513): elevation above sea level (P2044) → 8848 m
- Earth (Q2): lowest point (P1589) → Challenger Deep (Q459173)
- Challenger Deep (Q459173): elevation above sea level (P2044) → -10,994±40 m

Structured data

Or:

- Q2: P610 \rightarrow Q513
- Q513: P2044 \rightarrow 8848 m
- Q2: P1589 \rightarrow Q459173
- Q459173: P2044 \rightarrow -10,994 \pm 1 m

the third planet in the Solar System

Gaia | ⊕ | the world | Blue Planet | Terra Mater | Terra | Planet Earth | Tellus | The Earth | Sol III | Sol 3 | Sol d | ♁

↵ In more languages Configure

Language	Label	Description	Also known as
English	Earth	the third planet in the Solar System	Gaia ⊕ the world Blue Planet Terra Mater Terra Planet Earth Tellus The Earth Sol III Sol 3 Sol d ♁
Croatian	Zemlja	treći unutarnji planet Sunčeva sustava	Svijet Zemlja Majka Zemlja Geja Terra
Serbian	Земља	трећа планета у Сунчевом систему	
српски	No label defined	No description defined	

All entered languages

Statements

instance of	inner planet of the Solar System	0 references
-------------	---	--------------

part of	Solar System	0 references
---------	---------------------------	--------------

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

Tools

- What links here
- Related changes
- Special pages
- Permanent link
- Page information
- Concept URI
- Cite this page

Wikipedia (238 entries) [Collapse]

- ab Адгъыл
- ace Вумоѳ
- ady Чыгу
- af Aarde 🇺🇹
- als Erde
- am መሬት
- ang Eorðe
- an Tierra
- arc ܐܕܡܐ
- ar الأرض 🇺🇹
- arz الارض
- ast Tierra
- as পৃথিৱী
- av Ракъ (планета)
- ay Aka pacha
- azb بیز
- az Yer
- bar Eadn
- bat_smg Žemė
- ba Ер
- bcl Kinaban
- be_x_old Зямля
- be Планета Зямля 🇺🇹
- bg Земя 🇺🇹
- bh धरती
- bjn Bumi
- bn পৃথিবী
- bo ལྗོངས་ལྗོངས་
- br Douar (planedenn)
- bs Zemlja (planeta)
- bxr Дэлхэй
- ca Terra
- cbk_zam Tierra
- cdo Dê-giù

- Localization
- Description
- Aliases
- Interwiki

highest point

Chimborazo	
applies to part	centre
country	Ecuador
1 reference	
reference URL	https://www.theweathernetwork.com/us/news/articles/climate-and-environment/ecuadors-mt-chimborazo-is-officially-highest-spot-on-earth/66219
Mount Everest	
elevation above sea level	8,848±20 metre
applies to part	Earth's surface
country	Nepal
0 references	

deepest point

Challenger Deep	
elevation above sea level	-10,994±40 metre
1 reference	
reference URL	http://abstractsearch.agu.org/meetings/2011/FM/OS13B-1517.html
retrieved	8 January 2017
publication date	2011
Galathea Depth	
end time	1952
elevation above sea level	-10,540±1 metre
1 reference	
imported from	German Wikipedia
Vityas Deep 1	
start time	1957
elevation above sea level	-10,950±50 metre

- hak Thi-khii
- ha Duniya
- haw Honua
- he כדור הארץ
- hif Dunia
- hi पृथ्वी
- hr Zemlja
- hsb Zemja
- ht Latè
- hu Föld
- hy Երկիր
- ia Terra
- id Bumi 🌍
- ie Terra
- ig Àlà
- ilo Daga (planeta)
- io Tero
- is Jörðin 🌍
- it Terra 🌍
- iu ܕܢܘܗܐ
- jam Oert
- ja 地球 🌍
- jbo terdi
- jv Bumi
- kaa Jer
- kab Tagnit
- ka დედამიწა
- kg Ntoto
- kk Жер 🌍
- kl Nunarsuaq
- km ផែនដី
- kn ಭೂಮಿ
- koi Mywap
- ko 지구
- krc Джар
- ksh Ääd (Planeet)
- ku Erd

- Qualifiers
- Temporal data
- Data types
- Error ranges
- References

Data retrieval

- SPARQL
- Wikidata will give you all they *know*, but not more than that
- The more data there is, the more useful Wikidata is
- <https://query.wikidata.org>

Female mayors

```
9
10 ?city wdt:P31/wdt:P279* ?c . # find instances of subclasses of city
11 ?city p:P6 ?statement . # with a P6 (head of government) statement
12 ?statement ps:P6 ?mayor . # ... that has the value ?mayor
13 ?mayor wdt:P21 ?sex . # ... where the ?mayor has P21 (sex or gender) female
14 FILTER NOT EXISTS { ?statement pq:P582 ?x } # ... but the statement has no P582 (end date) qualifier
15
16 # Now select the population value of the ?city
17 # (wdt: properties use only statements of "preferred" rank if any, usually meaning "current population")
18 ?city wdt:P1082 ?population .
19 # Optionally, find English labels for city and mayor:
20 SERVICE wikibase:label {
21 bd:serviceParam wikibase:language "sr", "en" .
22 }
23 }
24 ORDER BY DESC(?population)
25 LIMIT 10
```

Press [CTRL-SPACE] to activate auto completion.

Data updated a few seconds ago

▶ Run Clear 10 Results in 995 ms

👁 Display ⬇ Download 🔗 Link

city	cityLabel	mayor	mayorLabel
Q wd:Q1490	Токио	Q wd:Q261703	Yuriko Koike
Q wd:Q1156	Мумбај	Q wd:Q18218029	Snehal Ambekar
Q wd:Q38283	Јокохама	Q wd:Q529363	Fumiko Hayashi
Q wd:Q1533	Каракас	Q wd:Q21573102	Helen Fernández
Q wd:Q2807	Мадрид	Q wd:Q19592761	Manuela Carmena
Q wd:Q11462	Сурабаја	Q wd:Q12522317	Tri Rismaharini
Q wd:Q220	Рим	Q wd:Q23766020	Вирџинија Рађи
Q wd:Q90	Париз	Q wd:Q2851133	Anne Hidalgo
Q wd:Q16555	Хјустон	Q wd:Q213847	Annise Parker
Q wd:Q3915	Антананариво	Q wd:Q14618311	Lalao Ravalomanana

Billionaires

item	itemLabel	billions
Q wd:Q5284	Бил Гејтс	77.7
Q wd:Q29279	Амансио Ортега	73.5
Q wd:Q47213	Ворен Бафет	60.8
Q wd:Q170419	Карлос Слим	49.9
Q wd:Q1066551	Charles G. Koch	42.9
Q wd:Q4934	Лари Пејџ	34.8
Q wd:Q92764	Сергеј Брин	34
Q wd:Q92764	Сергеј Брин	33.4
Q wd:Q92764	Сергеј Брин	32.6
Q wd:Q2105413	Wang Jianlin	28.7
Q wd:Q12908	Џорџ Сорос	24.9
Q wd:Q181162	Steve Ballmer	24.8
Q wd:Q557305	Phil Knight	23.4
Q wd:Q92764	Сергеј Брин	18.7
Q wd:Q293679	Al-Waleed bin Talal	18.3
Q wd:Q1394965	Aliko Dangote	18
Q wd:Q1225503	Dilip Shanghvi	17.5
Q wd:Q717038	Masayoshi Son	17
Q wd:Q92764	Сергеј Брин	16.6

Eponyms

eponym	eponymLabel	count	sample	sampleLabel
Q wd:Q41688	Хендрик Антон Лоренц	97	Q wd:Q172137	Лоренцова сила
Q wd:Q6722	Карл Фридрих Гаус	80	Q wd:Q2658	Gaussian elimination
Q wd:Q41585	Давид Хилберт	61	Q wd:Q190056	Хилбертов простор
Q wd:Q173746	Пал Ердеш	49	Q wd:Q243972	Ердешев број
Q wd:Q935	Исак Њутн	44	Q wd:Q12438	ЊУТН
Q wd:Q7604	Леонард Ојлер	43	Q wd:Q82435	е
Q wd:Q318362	Johan de Witt	42	Q wd:Q290492	HNLMS Johan de Witt
Q wd:Q42299	Bernhard Riman	36	Q wd:Q152860	4167 Riemann
Q wd:Q8814	Огистен Луј Коши	32	Q wd:Q45286	Cauchy–Binet formula
Q wd:Q9041	Џејмс Ват	29	Q wd:Q25236	ват
Q wd:Q23696	Simon Stevin	27	Q wd:Q1344596	Stevinus
Q wd:Q176659	Андреј Марков	25	Q wd:Q2733	Lempel–Ziv–Markov chain Algorithm
Q wd:Q80222	Жозеф Луј Лагранж	25	Q wd:Q190463	Лагранжова тачка
Q wd:Q153224	Андреј Колмогоров	24	Q wd:Q575766	Kolmogorov–Smirnov test
Q wd:Q77141	Александар Гротендик	24	Q wd:Q1062242	Grothendieck topology
Q wd:Q762	Леонардо да Винчи	23	Q wd:Q19101	Leonardo da Vinci–Fiumicino Airport
Q wd:Q17455	Џон фон Нојман	22	Q wd:Q77887	Von Neumann universe
Q wd:Q7251	Алан Тјуринг	21	Q wd:Q163310	Тјурингова машина
Q wd:Q44481	Пјер Симон Лаплас	20	Q wd:Q44666	Laplace's demon
Q wd:Q81082	Анри Поенкаре	19	Q wd:Q146944	2021 Поенкаре
Q wd:Q619	Никола Коперник	18	Q wd:Q1278	коперницијум

Cats



commons:James bowen (author).jpeg
Bob
 wd:Q27744042



commons:Ayasofya Gll cat.jpg
Gll
 wd:Q28792126



Mrs. Chippy
 commons:Mrschippy2.png
 wd:Q851190



Hodge
 commons:Hodgecat flickr.jpg
 wd:Q1622688



Scarlett
 commons:ScarlettTheCat.jpg
 wd:Q4421603



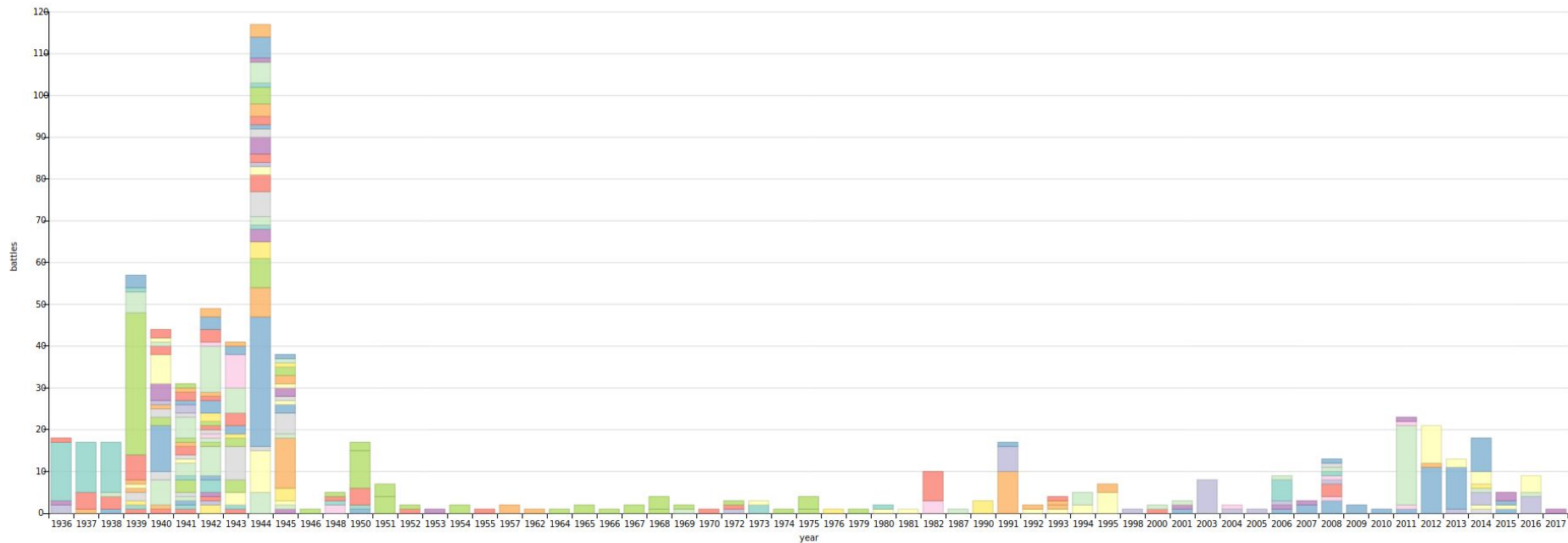
Mr. Nuts
 commons:Mr Nuts Executive Portrait.jpg
 wd:Q6928344



Tsim Tung Brother Cream
 commons:Brother cream in stall.jpg
 wd:Q10813994

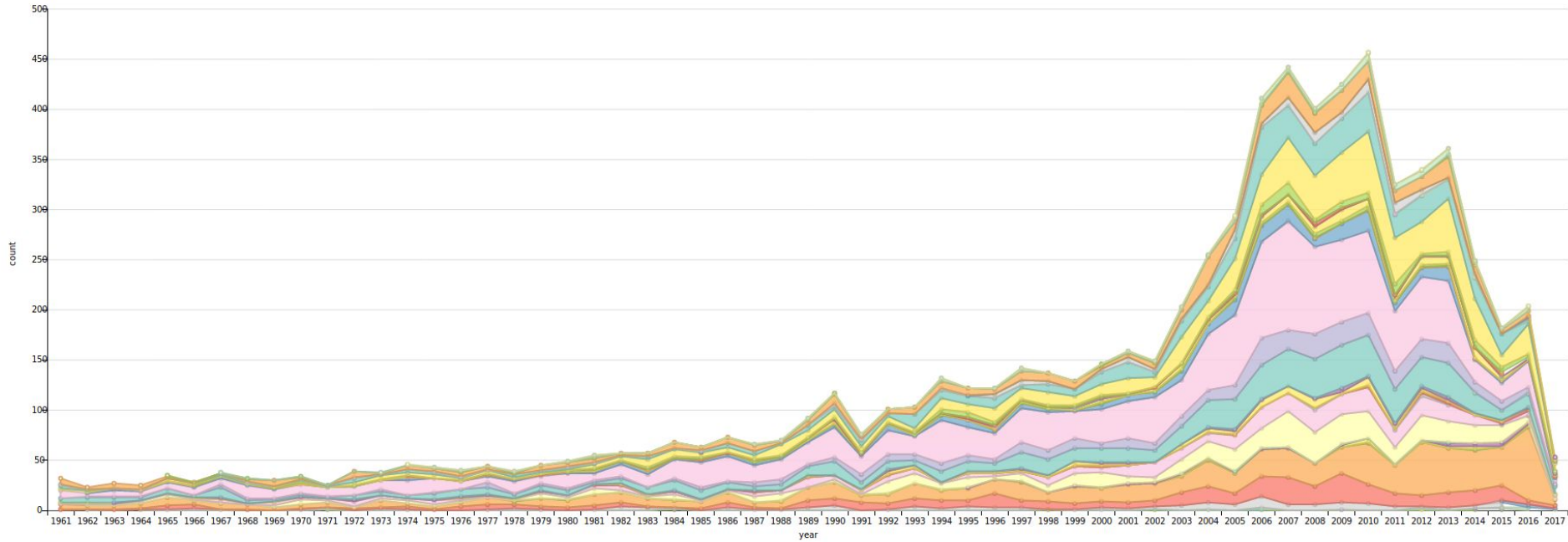


Wars by country by year



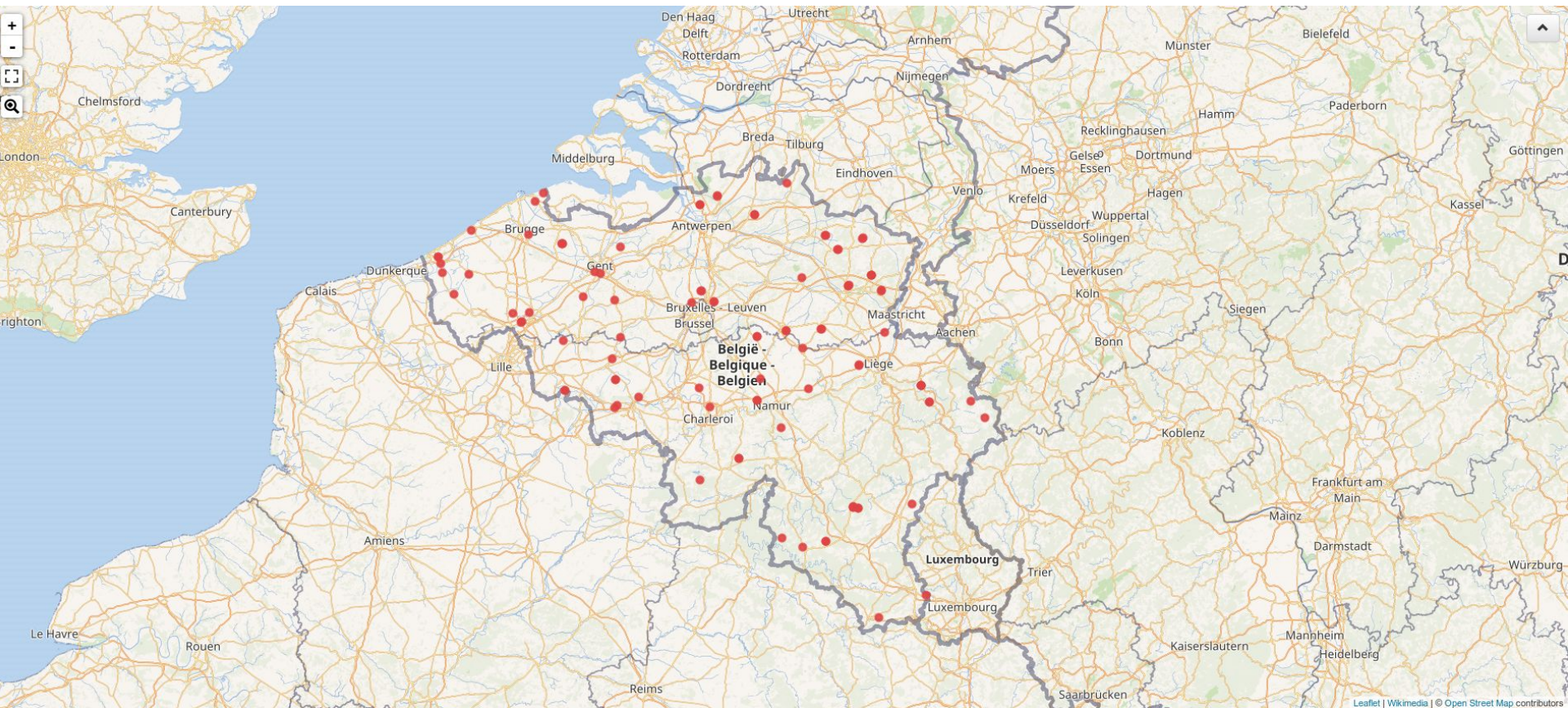
- | | | | | | | | | | |
|---|---|---|--|--|--|---|---|---|---|
| ■ Afghanistan | ■ Albania | ■ Algeria | ■ Angola | ■ Australia | ■ Austria | ■ Belarus | ■ Belgium | ■ Bosnia and Herzegovina | ■ Brunei |
| ■ Bulgaria | ■ Burundi | ■ Cambodia | ■ Canada | ■ Croatia | ■ Cyprus | ■ Czech Republic | ■ East Timor | ■ Egypt | ■ Eritrea |
| ■ Estonia | ■ Ethiopia | ■ Falkland Islands | ■ Finland | ■ France | ■ Georgia | ■ Germany | ■ Greece | ■ Guernsey | ■ Hungary |
| ■ India | ■ Indonesia | ■ Iran | ■ Iraq | ■ Israel | ■ Italy | ■ Japan | ■ Jersey | ■ Kingdom of the Netherlands | ■ Kiribati |
| ■ Kuwait | ■ Laos | ■ Lebanon | ■ Libya | ■ Lithuania | ■ Luxembourg | ■ Madagascar | ■ Malaysia | ■ Mali | ■ Marshall Islands |
| ■ Mongolia | ■ Morocco | ■ Myanmar | ■ Netherlands | ■ Prussia | ■ Northern Ireland | ■ Norway | ■ Oman | ■ Pakistan | ■ Palau |
| ■ Papua New Guinea | ■ People's Republic of China | ■ Philippines | ■ Poland | ■ Somalia | ■ Republic of Macedonia | ■ Romania | ■ Russia | ■ Senegal | ■ Serbia |
| ■ Singapore | ■ Slovakia | ■ Slovenia | ■ Solomon Islands | ■ Ukraine | ■ South Korea | ■ South Sudan | ■ Soviet Union | ■ Spain | ■ Sri Lanka |
| ■ Syria | ■ Thailand | ■ Tunisia | ■ Turkey | | ■ United Kingdom | ■ United States of America | ■ Vietnam | ■ Western Sahara | ■ Yemen |

Incidence of cancer by year



- | | | | | | | | | | |
|--------------------------|--------------------------|-------------------|----------------------|-------------------------|----------------------|-------------------------------|------------------------|------------------------------|--------------------------|
| Acute erythroid leukemia | Burkitt lymphoma | Ewing's sarcoma | Hodgkin's lymphoma | Merkel cell carcinoma | acute leukemia | acute lymphocytic leukemia | acute myeloid leukemia | adenocarcinoma of the lung | anus cancer |
| astrocytoma | bone cancer | brain cancer | breast cancer | cancer | carcinoma | cervical cancer | cholangiocarcinoma | chronic lymphocytic leukemia | chronic myeloid leukemia |
| colorectal cancer | duodenum cancer | esophageal cancer | gallbladder cancer | glioblastoma multiforme | head and neck cancer | hepatocellular carcinoma | intestinal cancer | kidney cancer | laryngeal cancer |
| leukemia | linitis plastica | liposarcoma | leukemia | lung cancer | lymphoid leukemia | lymphoma | lymphoma | melanoma | mesothelioma |
| multiple myeloma | myelodysplastic syndrome | myelofibrosis | myeloid neoplasm | nasopharynx carcinoma | non-Hodgkin lymphoma | non-small-cell lung carcinoma | mantle cell lymphoma | oligodendroglioma | oral cavity cancer |
| osteosarcoma | ovarian cancer | pancreatic cancer | peritoneal carcinoma | prostate cancer | rectum cancer | renal cell carcinoma | ocular melanoma | rhabdomyosarcoma | skin cancer |
| small cell carcinoma | spinal cancer | stomach cancer | testicular cancer | throat cancer | thyroid cancer | renal pelvis carcinoma | urinary bladder cancer | uterine cancer | |

Airports of Belgium



Links

- <http://tools.wmflabs.org/wikidata-game/> - Gamification
- <http://tools.wmflabs.org/fist/wdfist/> - Image search
- <http://tools.wmflabs.org/mix-n-match/> - Linking to external databases
- <https://tools.wmflabs.org/scholia/> - Citations, papers, authors
- <https://tools.wmflabs.org/reasonator/> - Narrative view
- <http://tools.wmflabs.org/magnustools/#wikidata> - More...
- https://www.wikidata.org/wiki/Wikidata:Tools/External_tools - Misc

Questions?

Thank you!

filip@vikimedija.org