Wikidata

and

Querying Wikidata with SPARQL

What are the ten largest cities with a female mayor?

👁▾ ❓   10 results in 52 ms   </> Code   ⬇ Download▾   🔗 Link▾

cityLabel ⇅	mayorLabel ⇅	population ⇅
Tokyo	Yuriko Koike	13784212
Hong Kong	Carrie Lam	7336585
Baghdad	Zekra Alwach	6960000
Surabaya	Tri Rismaharini	4975000
Yokohama	Fumiko Hayashi	3731706
Rome	Virginia Raggi	2873494
Paris	Anne Hidalgo	2206488
Havana	Marta Hernández Romero	2141652
Caracas	Helen Fernández	1943901
Bucharest	Gabriela Firea	1883425

Where are people born who travel to space?

(colour-coded by gender)
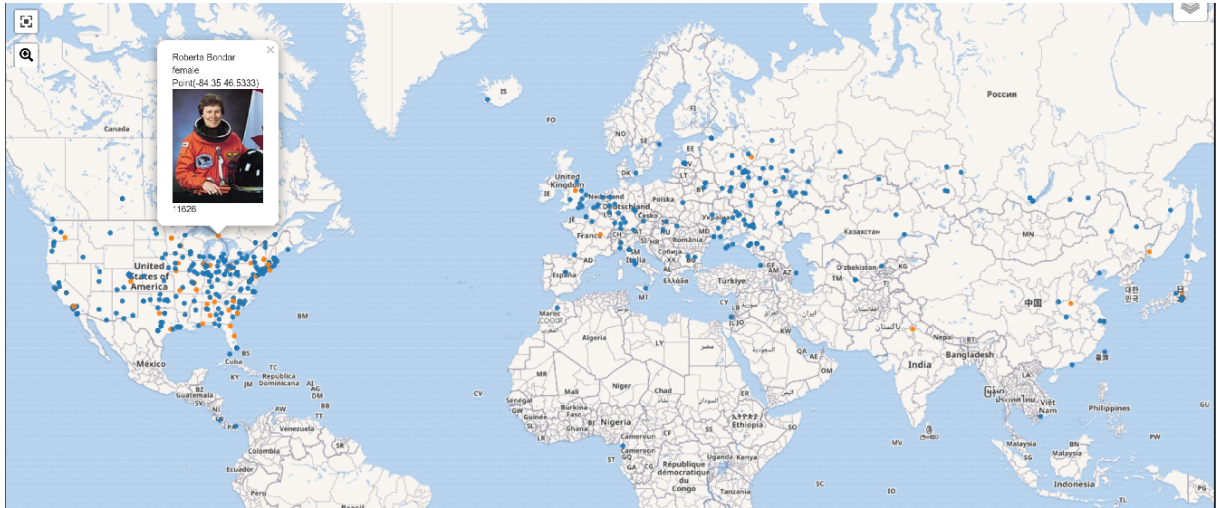
Which days of the week do disasters occur?

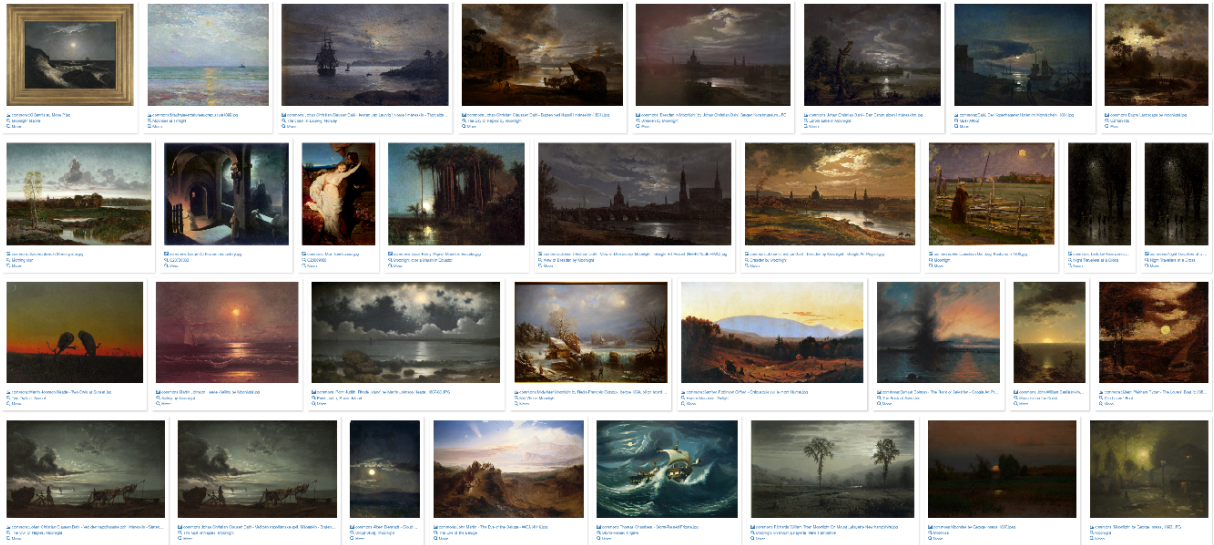| Date	Mon	Tue	Wed	Thu	Fri	Sat	Sun
1	25	33	22	18	26	28	23
2	24	26	23	23	22	32	12
3	24	27	21	31	23	28	38
4	24	25	33	25	26	26	24
5	37	23	32	18	19	17	29
6	25	28	32	20	24	33	22
7	18	22	25	16	22	18	17
8	32	28	19	25	22	23	19
9	20	25	29	29	27	21	23
10	20	20	19	14	25	25	29
11	30	34	28	23	22	20	20
12	41	33	27	30	20	20	23
13	35	26	29	21	25	24	25
14	24	23	27	23	22	28	17

Which 19th century paintings show the moon?

Which films co-star more than one future head of government?

Star in the Dust	1956 film by Charles F. Haas	2	Clint Eastwood, mayor; George Wallace, Governor of Alabama
The Two Who Stole the Moon	1962 Polish film by Jan Batory	2	Jarosław Kaczyński, Prime Minister of Poland; Lech Kaczyński, Mayor of Warsaw
Ragasiya Police 115	1968 film by B. R. Panthulu	2	M. G. Ramachandran, Chief Minister of Tamil Nadu; Jayalalithaa, Chief Minister of Tamil Nadu
Québec : Duplessis et après…	documentary	2	Bernard Landry, Premier of Quebec; René Lévesque, Premier of Quebec
Q3541438	1994 film by Claude Lanzmann	2	Ariel Sharon, Prime Minister of Israel; Ehud Barak, Prime Minister of Israel
Batman & Robin	1997 American superhero film based on the DC Comics character Batman	2	Arnold Schwarzenegger, Mr. Freeze / Governor of California; Jesse Ventura, Governor of Minnesota

A Free Knowledge Graph

WIKIDATA

Wikidata

- Wikipedia's knowledge graph
- Free, community-built database
- Large graph

    (October 2018: >570M statements on >50M entities)
- Large, active community

    (October 2018: >230,000 logged-in human editors)
- Many applications

Freely available, relevant, and active knowledge graph

A short history of Wikidata

- August 2005: Presentation "Wikipedia and the Semantic Web—The Missing Links" at the 1st Wikimedia Conference "Wikimania", Frankfurt, Germany
- 29th October 2012: wikidata.org is launched
- 15th Dec 2012: Item with ID number 1000000 created
- 4th Feb 2013: The first statements can be created
- Early 2013: Most Wikipedia language links relocate to Wikidata
- Late 2013: More than 100,000,000 edits on over 15M items
- Dec 2014: Google announces the closure of Freebase and migration to
  Wikidata
- 2014–2018: A total of >700M edits produce >55M items and >570M
  statements
- May 2018: Wikidata starts storing data about lexemes
  (= expressions in a language)
- Oct 2018: Senses of lexemes become supported

Many applications (1)

As of today, Wikidata content has been used in many ways.

Wikipedia & the Wikimedia community:

- Wikipedia inter-language links (see any Wikipedia page)
- Data displays in pages (auto-generated info boxes, article placeholders,
  result tables, . . .)
- Quality checks & edit-a-thons

External re-uses of data:

- Application-specific data-excerpts (e.g., Eurowings in-flight app)
- Data integration and quality control
  (e.g., Google checks own KG against Wikidata)
- Authority control & identity provider (VIAF, Open Streetmaps, DBLP, . . . link their content to Wikidata)
- Data-driven journalism (individual analyses as well as data-driven
  information portals)

Many applications (2)

As of today, Wikidata content has been used in many ways.

In research:

- Test data for KG-related algorithms
- Training data for machine-learning approaches
- Wikidata as a subject of study (social dynamics, internationality, biases, . . .)

Uses by Wikidata community:

- Software-supported error and vandalism detection
- Feature-based integration with other datasets
- Data-driven statistics as a measure of progress

What is Wikidata?

Wikidata is often described as **"the free knowledge base that anyone can edit"**
or the **"knowledge graph of Wikipedia"**

It is useful to distinguish several of these aspects:

Wikidata is . . .

- *. . . a Wikimedia project* like Wikipedia and Wikimedia Commons; represented and supported by the Wikimedia Foundation (WMF)
- *. . . a dataset* that can be downloaded and freely used and distributed
- *. . . a website* through which the data can be viewed and modified
- *. . . a community* of volunteer editors that creates and controls all content

Two views on the Wikidata knowledge base

The website and its main data services expose Wikidata as a **document-centric knowledge base:**

- Data is grouped by subject entity (one page per entity)
- Documents are structured into different sections
- The order of content is (mostly) preserved and shown

⤳ Useful for display & management

Conceptually and for most applications, Wikidata is a **graph-structured knowledge base:**

- Main content are binary relationships (from entities to entities/data values)
- Properties are first-class objects with a global scope and definition
- Order does not affect the meaning of statements

⤳ Useful for sharing and re-use

We will mostly view Wikidata as a knowledge graph

Wikidata

`https://www.wikidata.org/wiki/Q80`

`https://towardsdatascience.com/a-brief-introduction-to-wikidata-bb4e66395eb1`

Querying Wikidata with SPARQL for Absolute Beginners

`https://www.youtube.com/watch?v=kJph4q0Im98`

`https://www.wikidata.org/wiki/Wikidata:SPARQL_tutorial`

Wikidata Query Service

`https://query.wikidata.org`