

SPARQL for absolute beginners

Tobias Schönberg

What is SPARQL?

- Query language which is similar to spoken language
- Used extensively outside of Wikimedia

Why learn SPARQL?

lv

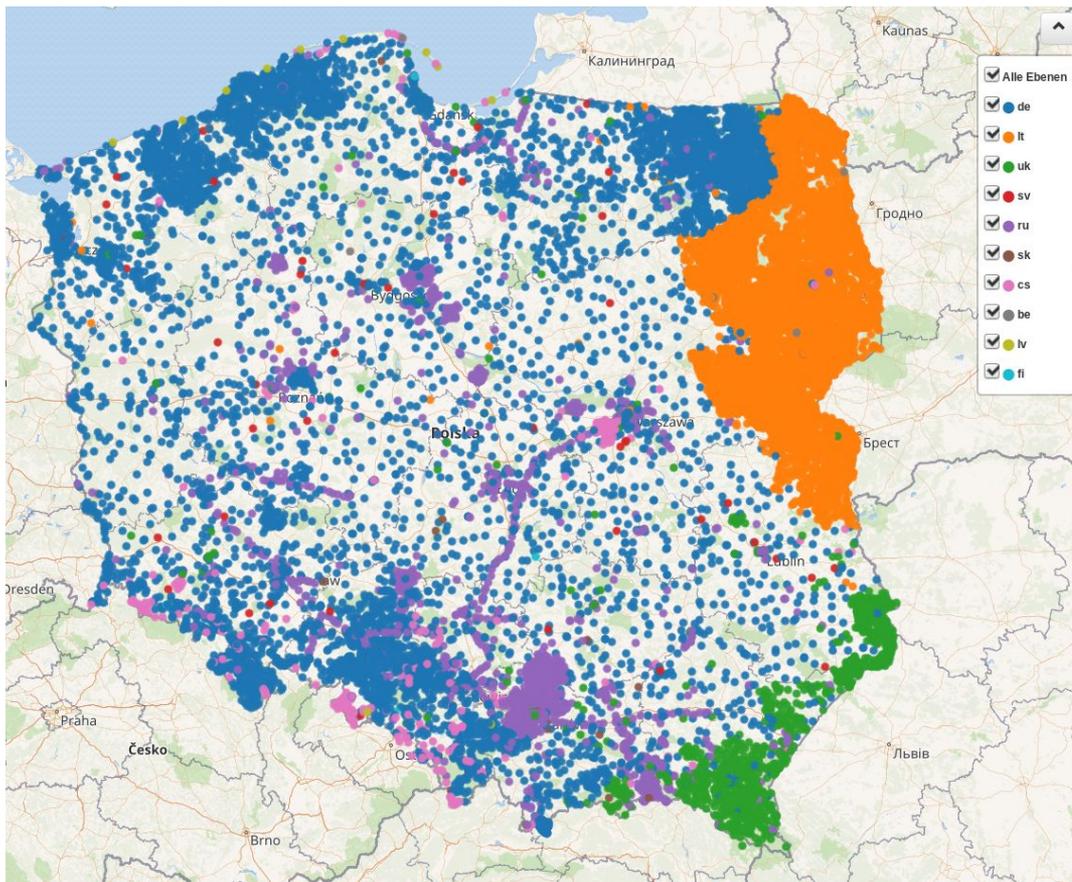
fi

de

cs

Map showing neighboring languages of Polish that have a Wiki-page that no other neighboring language has.

Query in Appendix



it

be

sv

sk

ru

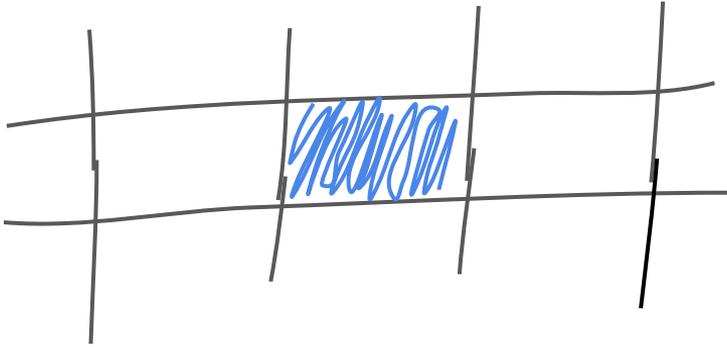
uk

Based on: <https://twitter.com/WikidataFacts/status/908444126999441408>

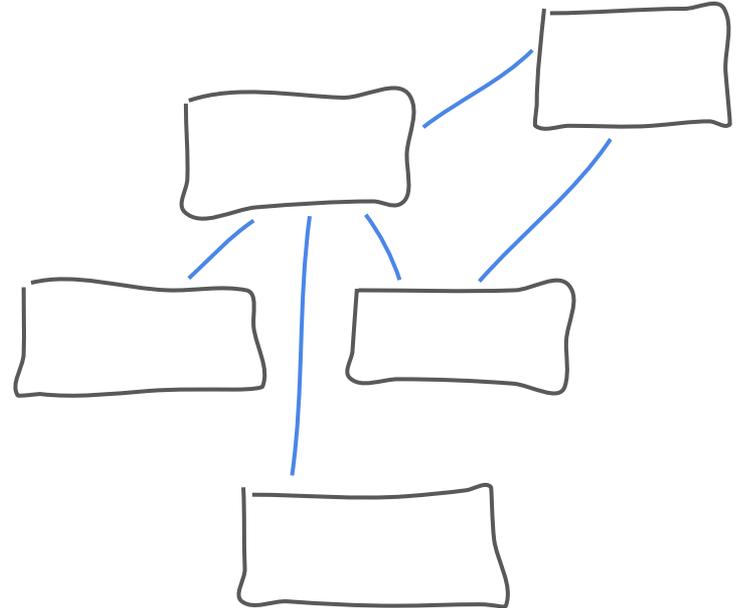
Some basics

How does SPARQL compare to SQL?

SQL is about tables

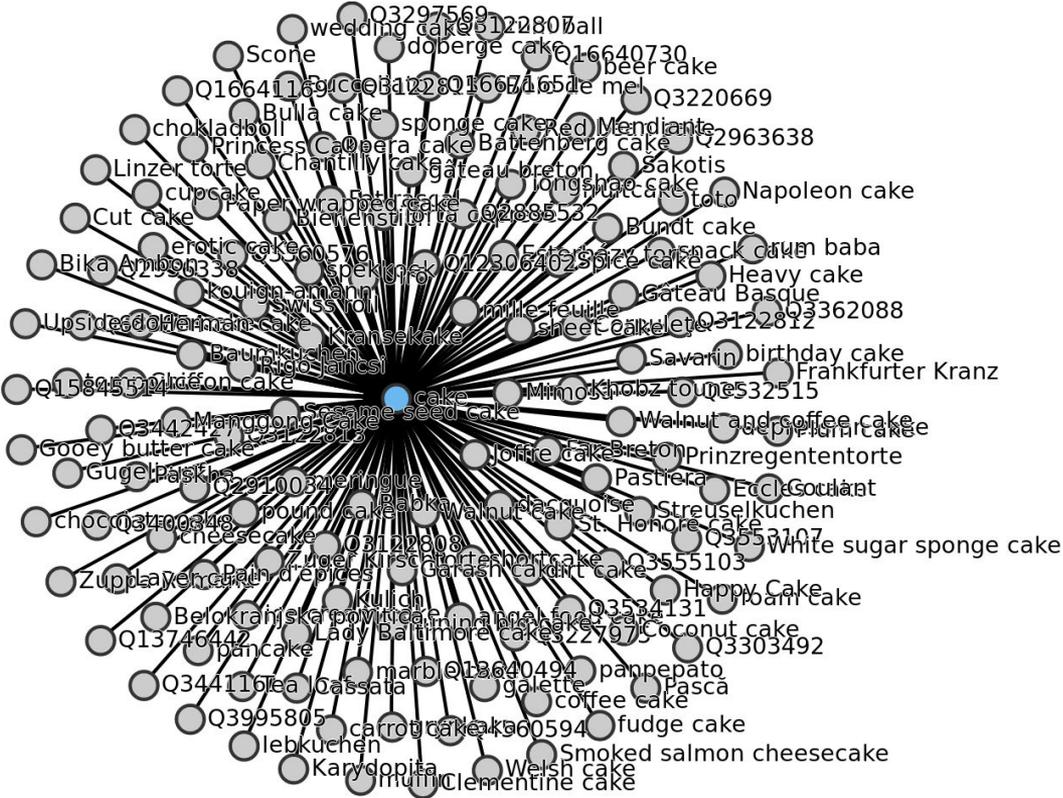


SPARQL is about trees / graphs



SPARQL loves trees and graphs

?items subclass-of cake



Difference SQL and SPARQL syntax

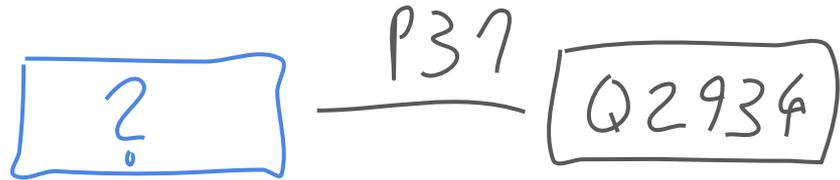
column = value

name = 12



subject predicate object

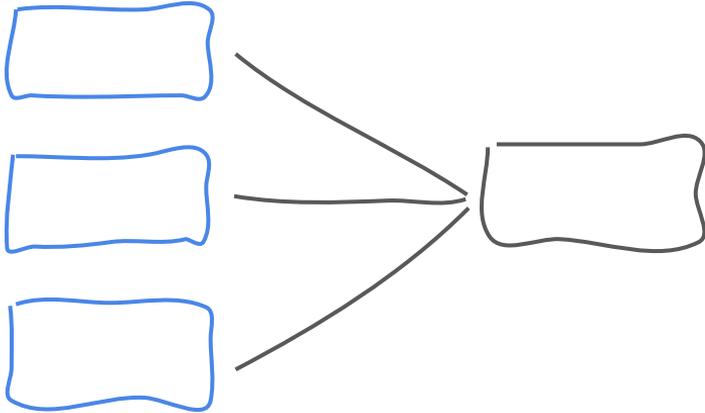
 ? P31 Q2934



How does SPARQL select items?

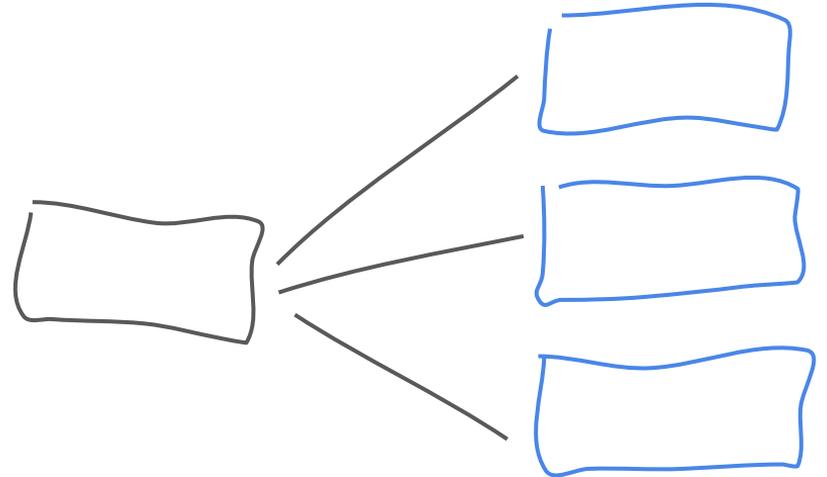
Object is variable

?people born-in Warsaw



Subject is variable

Warsaw is-a ?types



Overview over query.wikidata.org

Examples

Language

The screenshot shows the Wikidata Query interface. At the top left, the 'Wikidata Query' logo is visible. A purple box highlights the 'Beispiele' button. To the right, the language is set to 'Deutsch', also highlighted with a purple box. The main interface is divided into three sections: a 'Helper' section on the left with filter and display options, a 'Query Input' section in the center with a SQL query, and a 'Run query' button at the bottom left. The 'Query Input' section contains the following SQL code:

```
1 #Goats
2 SELECT ?item ?itemLabel
3 WHERE
4 {
5   ?item wdt:P31 wd:Q2934.
6   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
7 }
```

Below the query input, a table of results is displayed. The table has two columns: 'item' and 'itemLabel'. The results are as follows:

item	itemLabel
wd:Q3569037	William Windsor
wd:Q5738485	Goat of the Spanish Legion
wd:Q20908547	Capra hircus aegagrus
wd:Q24287064	Taffy

Query Input

Query Input

Helper

Run query

Results

Example 0: Notable cats

Get the cats query by pressing on “Example”

```
SELECT ?item ?itemLabel
```

```
WHERE
```

```
{
```

```
  ?item wdt:P31 wd:Q146.
```

```
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
```

```
}
```

Example 0: Can you find notable dogs?

Notable dogs

```
SELECT ?item ?itemLabel
```

```
WHERE
```

```
{
```

```
  ?item wdt:P31 wd:Q144 .
```

```
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
```

```
}
```

Can you find people which are
musicians?

Musicians

```
SELECT ?item ?itemLabel
```

```
WHERE
```

```
{
```

```
  ?item wdt:P106 wd:Q639669 .
```

```
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
```

```
}
```

Example 1: Statues of Polish speakers

Get all the instances of statues

```
SELECT ?item ?itemLabel
```

```
WHERE
```

```
{
```

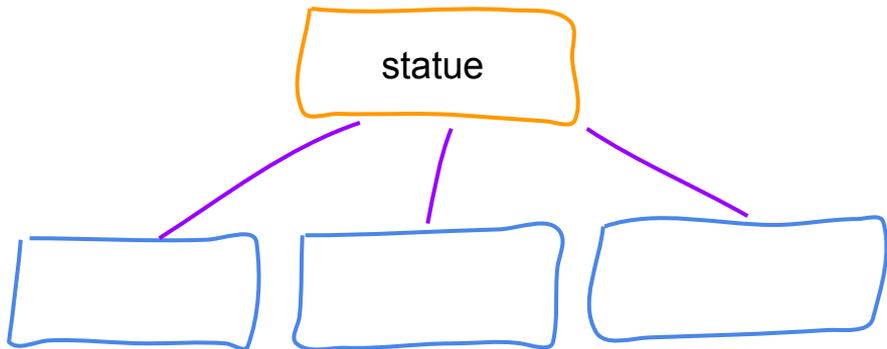
```
  ?item wdt:P31 wd:Q179700 .
```

```
  SERVICE wikibase:label {
```

```
    bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" .
```

```
  }
```

```
}
```



What are we selecting?

All items that have this statement:

Statements

instance of	 statue	 edit
	▼ 0 references	+ add reference
		+ add value

Add the condition that something is depicted

```
SELECT ?item ?itemLabel ?person ?personLabel
```

```
WHERE
```

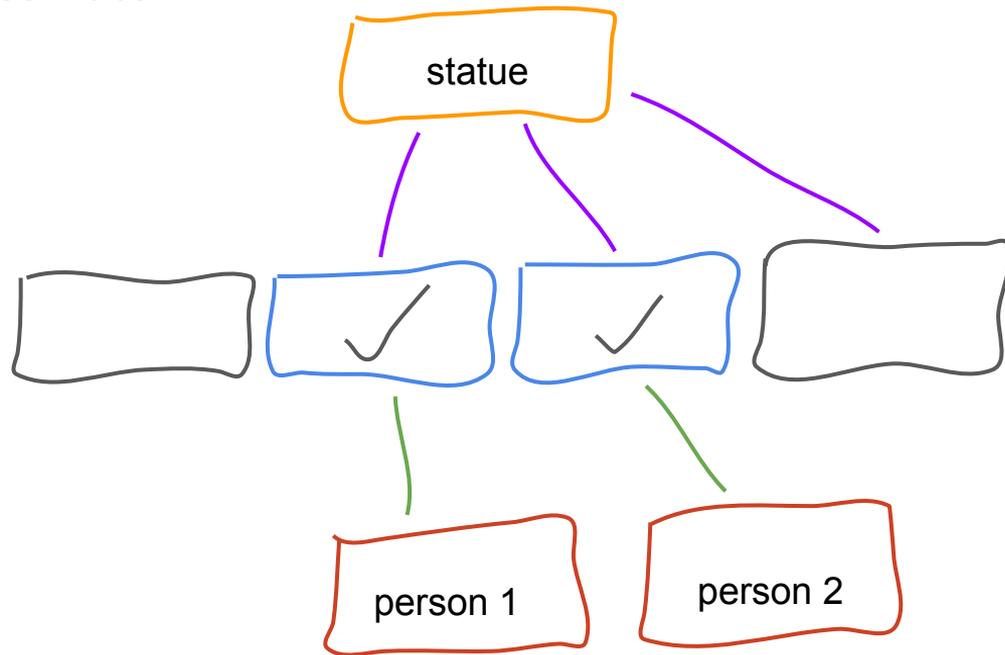
```
{
```

```
?item wdt:P31 wd:Q179700 .
```

```
?item wdt:P180 ?person .
```

```
SERVICE wikibase:label { ... }
```

```
}
```



Note: depicts can also be something different from a person

What are we selecting now?

All items that have these 2 statements:

Statements

<u>instance of</u>	 statue  edit
	▼ 0 references
	+ add reference
<u>depicts</u>	 man  edit
	▼ 0 references
	+ add reference
	 nudity  edit
	▼ 0 references
	+ add reference
	 Contrapposto  edit
	▼ 0 references
	+ add reference
	 Attalus II Philadelphus  edit
	▶ 1 reference

Let's now select something about the person

```
SELECT ?item ?itemLabel ?person ?personLabel
```

```
WHERE
```

```
{
```

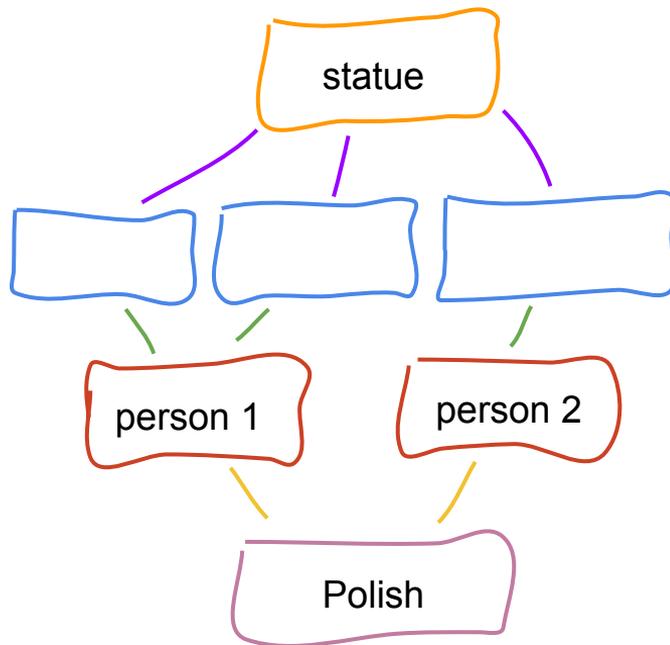
```
  ?item wdt:P31 wd:Q179700 .
```

```
  ?item wdt:P180 ?person .
```

```
  ?person wdt:P1412 wd:Q809 .
```

```
  SERVICE wikibase:label { ... }
```

```
}
```



Exciting results

person	personLabel
 wd:Q153132	Władysław Sikorski
 wd:Q1268	Frédéric Chopin
 wd:Q1268	Frédéric Chopin
 wd:Q1268	Frédéric Chopin
 wd:Q1268	Frédéric Chopin
 wd:Q1268	Frédéric Chopin
 wd:Q1268	Frédéric Chopin
 wd:Q989	Johannes Paul II.
 wd:Q683515	Stefan Rowecki
 wd:Q508834	August Hlond

Let us try to put the statues on a map

```
SELECT ?item ?itemLabel ?person ?personLabel ?coords
```

```
WHERE
```

```
{
```

```
  ?item wdt:P31 wd:Q179700 .
```

```
  ?item wdt:P180 ?person .
```

```
  ?person wdt:P1412 wd:Q809 .
```

```
  ?item wdt:P625 ?coords .
```

```
  SERVICE wikibase:label { ... }
```

```
}
```

Blue things were added on this slide.
query.wikidata.org will allow you to
display a map, if a variable contains
coordinates.

Awesome result, but probably a lot of missing data



Example 2: Text analysis in the CEE region

Example 2: Get All the CEE countries

```
SELECT DISTINCT ?item
```

```
WHERE
```

```
{
```

```
    VALUES ?countries { wd:Q225 wd:Q28 wd:Q403 wd:Q212 wd:Q218  
wd:Q224 wd:Q236 wd:Q221 wd:Q215 wd:Q214 wd:Q36 wd:Q219 wd:Q40  
wd:Q1246 } .
```

```
    ?item wdt:P17 ?countries .
```

```
}
```

Get coordinates and the label

```
SELECT DISTINCT ?item ?itemLabel ?coord
```

```
WHERE
```

```
{
```

```
    VALUES ?countries { wd:Q225 wd:Q28 wd:Q403 wd:Q212 wd:Q218  
wd:Q224 wd:Q236 wd:Q221 wd:Q215 wd:Q214 wd:Q36 wd:Q219 wd:Q40  
wd:Q1246 } .
```

```
    ?item wdt:P17 ?countries .
```

```
    ?item wdt:P625 ?coord .
```

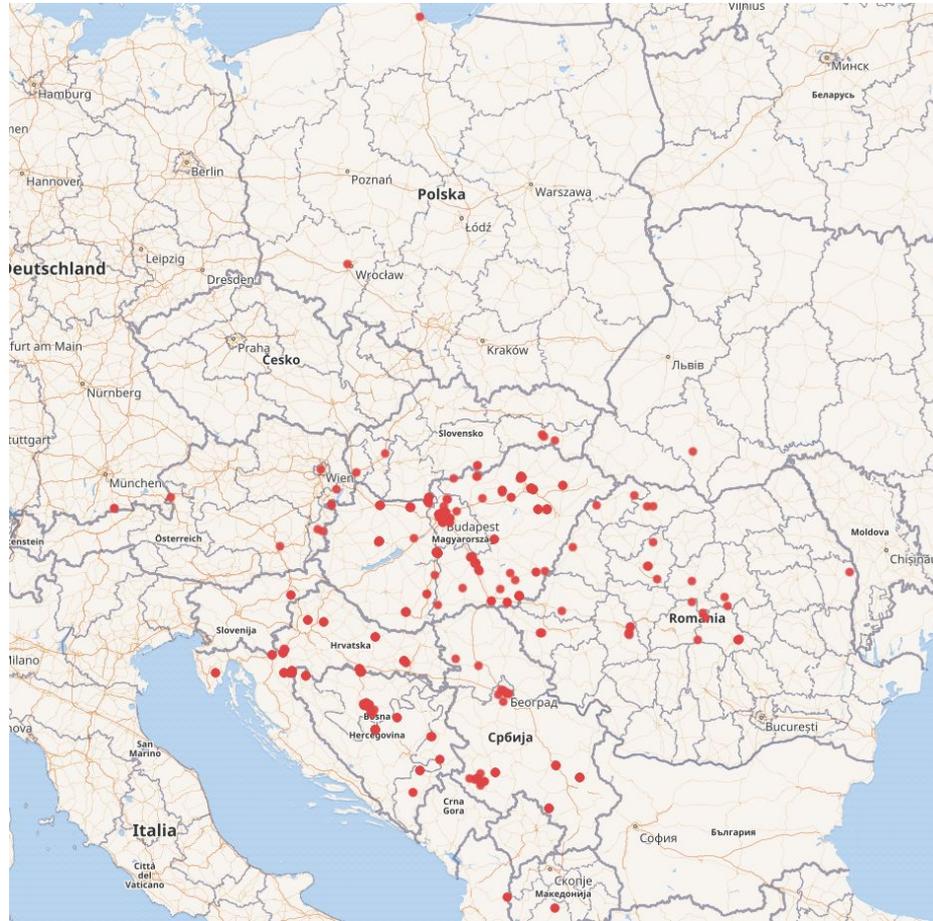
```
    ?item rdfs:label ?itemLabel .
```

```
}
```

Filter for words containing “város” or “varoš”

```
SELECT DISTINCT ?item ?itemLabel ?coord
WHERE
{
    VALUES ?countries { wd:Q225 wd:Q28 wd:Q403 wd:Q212 wd:Q218
wd:Q224 wd:Q236 wd:Q221 wd:Q215 wd:Q214 wd:Q36 wd:Q219 wd:Q40
wd:Q1246 } .
    ?item wdt:P17 ?countries .
    ?item wdt:P625 ?coord .
    ?item rdfs:label ?itemLabel .
    FILTER REGEX(STR(?itemLabel), "v(a|á)ro(s|š)", "i")
}
```

The result: Hungarian cult. influence in CEE



Disclaimer: Some results might be translations of the word “city” that never had a Hungarian name. Or the fragment “varos” in another word.

Appendix 1: Neighboring wikis article in Poland

```
#defaultView:Map
SELECT ?item (SAMPLE(?title) AS ?itemLabel) (SAMPLE(?location) AS ?location) (SAMPLE(?language) AS ?layer) WITH {
  SELECT * WHERE {
    wd:Q36 p:P1332/psv:P1332/wikibase:geoLatitude ?n;
    p:P1333/psv:P1333/wikibase:geoLatitude ?s;
    p:P1334/psv:P1334/wikibase:geoLongitude ?e;
    p:P1335/psv:P1335/wikibase:geoLongitude ?w.
  }
} AS %boundingBox WHERE {
  VALUES ?wiki {
    <https://de.wikipedia.org/>
    <https://ru.wikipedia.org/>
    <https://cs.wikipedia.org/>
    <https://uk.wikipedia.org/>
    <https://be.wikipedia.org/>
    <https://lv.wikipedia.org/>
    <https://lt.wikipedia.org/>
    <https://sv.wikipedia.org/>
    <https://fi.wikipedia.org/>
    <https://sk.wikipedia.org/>
  }
  ?item wdt:P17 wd:Q36;
  wdt:P625 ?location.
  ?article a schema:Article;
  schema:about ?item;
  schema:isPartOf ?wiki;
  schema:inLanguage ?language;
  schema:name ?title.
  INCLUDE %boundingBox.
  BIND(geof:latitude(?location) AS ?lat)
  BIND(geof:longitude(?location) AS ?lon)
  FILTER(?s <= ?lat && ?lat <= ?n &&
    ?w <= ?lon && ?lon <= ?e)
}
GROUP BY ?item
HAVING(COUNT(DISTINCT ?wiki) = 1)
```