



Czech heritage data in Wikidata

Vojtěch Dostál | vojtech.dostal@wikimedia.cz

Wikidata love monuments!

Wikidata:WikiProject Built heritage



Obsah [skrýt]

- 1 [Aim and Scope](#)
- 2 [Background](#)
- 3 [Ways to Contribute](#)
- 4 [Case Reports](#)
- 5 [Related Projects](#)
- 6 [Participants \(en\)](#) – *(Toto prosím přeložte do jazyka čeština.)*
- 7 [Navigation](#)

[Aim and Scope](#) [editovat]

The aim of the present project is to create the world's most complete high quality database of built heritage. It builds upon, consolidates, and extends the work carried out in the context of numerous [Wiki Loves Monuments](#) contests for which official monuments lists have been compiled for many countries.

The project more specifically serves to coordinate the following activities:



Entrance hall of the regional court of Berlin located in Littenstrasse 12-17 in Berlin-Mitte



Cultural monument tables on Czech Wikipedia – created for the 1st WLM contest in 2012

<p>Kašna s Rolandem</p>	 	<p>20135/4-3655 Pam. katalog MonumNet MIS</p>	<p>Cheb</p>	<p>náměstí Krále Jiřího z Poděbrad, pp. 2273/1 50°4'43,41" s. š., 12°22'15,09" v. d.</p>	<p>Kašna s Rolandem (Horní kašna). Polygonální kašna s kamennou nádrží, s hladkými stěnami a vysazenou římsou. Z 19. století, na ní na středním hranolovém pilířku osazena socha Rolanda s vročením 1591. Originál sochy je umístěn v muzeu, na místě stojí faksimile. Roland v podživotní velikosti drží v pravé ruce kopí, levou se opírá o štít se znakem města Cheb.</p> <p>Památkově chráněno od 9. června 1970.</p>
<p>Kašna s Herkulem</p>	 	<p>18537/4-3656 Pam. katalog MonumNet MIS</p>	<p>Cheb</p>	<p>náměstí Krále Jiřího z Poděbrad, pp. 2273/1 50°4'46,33" s. š., 12°22'12,23" v. d.</p>	<p>Kašna s Herkulem (Dolní kašna). Renesanční polygonální kašna s nádrží s hladkými stěnami s vysazeným okrajem. Ze 16. století. Na středním pilíři, přecházejícím v hlavici se stylizovaným dekorem, je doplněna sochou Herkula, jež nese vročení 1738. Herkules je v podživotní velikosti, s kyjem a štítem se znakem města Chebu.</p> <p>Památkově chráněno od 9. června 1970.</p>
<p>Náměstí Krále Jiřího z Poděbrad 43, Špalíček, čp. 494</p>		<p>21787/4-3598 Pam. katalog MonumNet MIS</p>	<p>Cheb</p>	<p>náměstí Krále Jiřího z Poděbrad 494/43 50°4'47,64" s. š., 12°22'11,78" v. d.</p>	<p>Měšťanský dům Špalíček. Dvoupatrový empírový dům v rámci seskupení Špalíčku situovaný v jeho dolním, severozápadním rohu. Architektonický výraz v duchu klasicismu. Pětiosý, dvoupatrový s mansardovou střechou se dvěma vikýři. Obě boční průčelí jsou dvouosá. V prvním patře na nároží k čp. 507 je nika pro sochy, od roku 2007 osazená moderní plastikou, původně zde byla gotická Madona s dítětem která je uložena v</p>

The tables use templates but they are still rather unflexible and difficult to search and edit

```
{{Památky v Česku/začátek|kraj-iso=CZ-KA|okres-iso=CZ-411}}
{{Památky v Česku
|Obrázek=
|Commons=
|Název=Usedlost čp. 40
|Článek=
|Adresa=Hrozňatov 40
|Obec=Hrozňatov
|Zeměpisná_šířka=50.0212886
|Zeměpisná_délka=12.3934175
|Popis=Venkovská usedlost z 18. století skládající se ze dvou obytných stavení a stodoly. Z původně uzavřeného dvora chebského typu z konce 18. století se zachovaly 2 obytné objekty a stodola. Druhé obytné stavení bylo původně kůlnou.
|Poznámka=Parcelní číslo v MonumNetu neuvedeno.
|Id_objektu=32148/4-4524
|Památkou_od=1990-08-03
|Wikidata=Q33242163}}
```

New challenges...

Vítejte na Wikidatech

svobodné databázi vědomostí s 42 289 659 položkami, kterou může každý upravovat.

[Úvod](#) • [Diskuse o projektu](#) • [Komunitní portál](#) • [Nápověda](#)

[Chcete pomoci překládat? Přeložte chybějící zprávy.](#)

2017 developments

- 19 Jan 2017 – first meeting with National Heritage Institute (NHI) on heritage data on Wikidata
- 5 Apr 2017 – default state of the data – 6000 alleged Czech cultural monuments in WD
- 14 Jun 2017 – laborious clean-up of data – some monuments removed, some correctly marked → 6053 items
- 10 Jul 2017 – import of cultural monument tables to Wikidata, we're at 16466 (and at 40570 on 2nd Sept 2017 – close to the total number of monuments).
- 7 Sep 2017 – agreement with NHI signed, access gained to CC-BY-SA data in the NHI Geoportal (19th Sep 2017) – **including coordinates**

2017 developments

- 21 Sep 2017 – Jiri Sedláček added coordinates from Geoportal to lists in Czech (license allows Wikipedia usage) - **974 edits in 961 Wikipedia pages**
- 8 Oct 2017 – export from Wikipedia to Wikidata - **34253 edits in 17217 items - 40598 monuments with coordinates** now available!
- 22 Oct 2017 – all monument identifiers are unique (used in 1 item max)
- 6-9 Nov 2017 – 640 articles on Czech Wikipedia enriched with new infoboxes which data available in Wikidata (and thousands more improved with automatical changes in the infobox source codes)
- 9 Nov 2017 – import of a more modern monument identifier (P4075) by automatical matching with P762 and semiautomatical checks afterwards

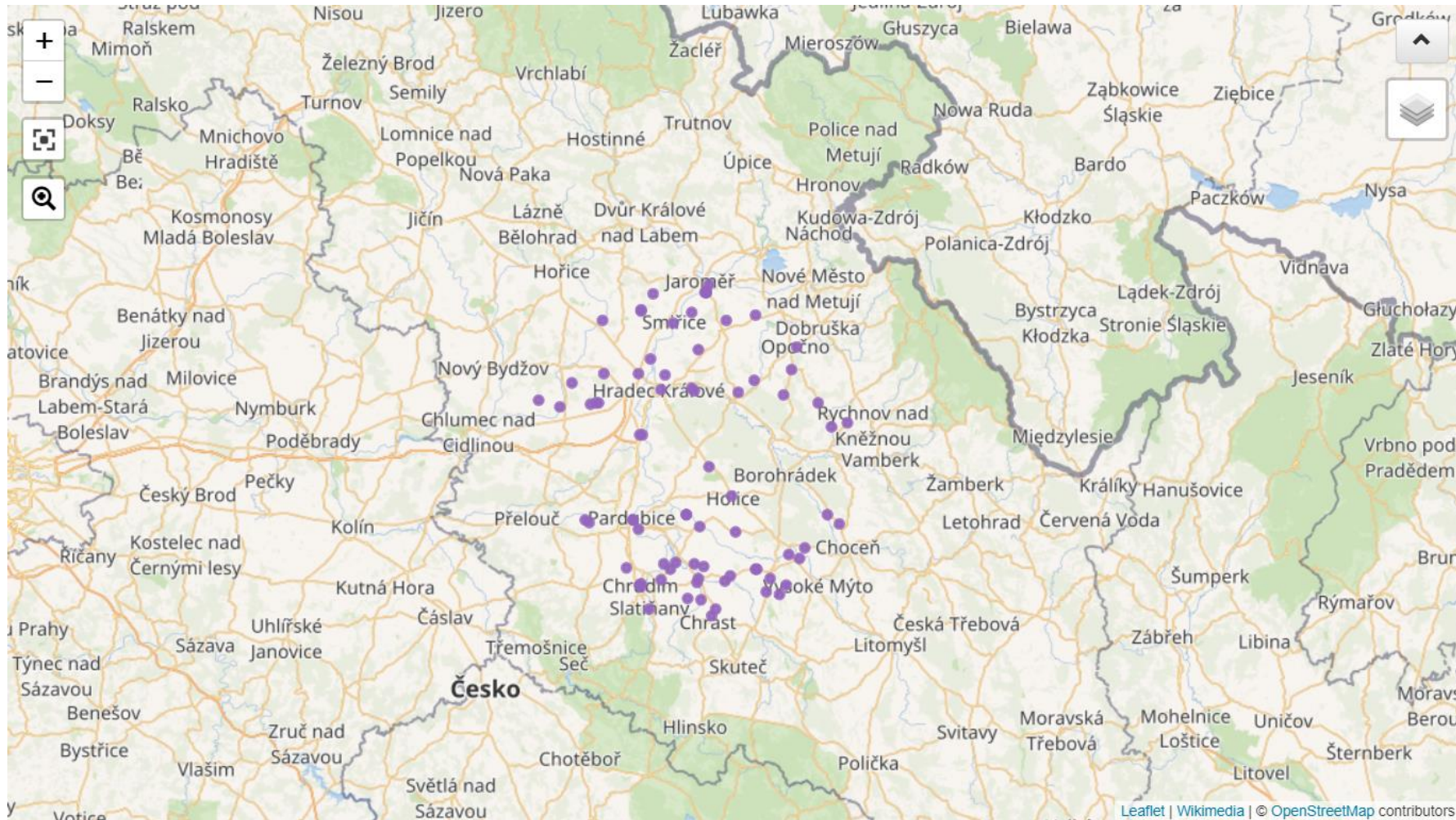
2017 developments

New cultural monuments are being proclaimed, and the current ones are sometimes revoked. Wikidata now reflect the state on 19th Sep 2017 ... updates must continue

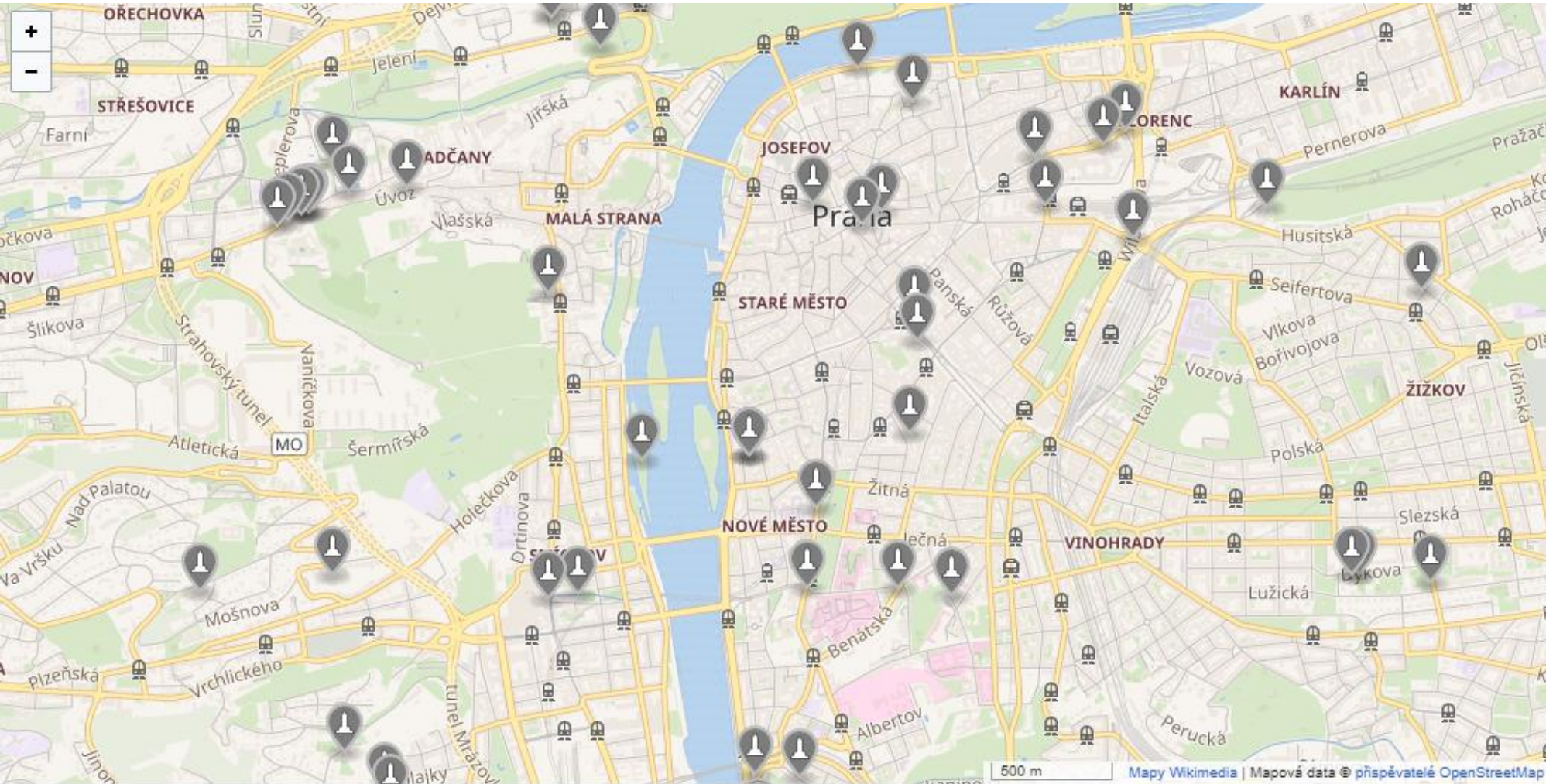
KATALOGOVÉ ČÍSLO	1000024187
NÁZEV	hasičská zbrojnice
KRAJ	Plzeňský kraj
OKRES	Tachov
OBEC	Bezručice
ČÁST OBCE	Bezručice
KATASTRÁLNÍ UZEMÍ	Bezručice
STUPEŇ PAMÁTKOVÉ HODNOTY	hodnotný
VÝZNAM PAMÁTKOVÉ HODNOTY	regionální
TYP	hasičská zbrojnice
DRUH PAMÁTKY	stavba-budova
KATEGORIE	objekt
PAMÁTKOVĚ CHRÁNĚNO OD	4. 10. 2017
EXISTUJÍCÍ	ano

Why are monument coordinates important for Wikimedia Czech Republic?

<https://tools.wmflabs.org/wikinity/>



Prepared competition „Describe a monument“ – missing articles on a map using data from Wikidata



Map details

Coordinates 50.0794, 14.4305

View this location in one of the services listed below. You can filter the list by map type:

Map

External services


Service	
Bing Maps	Map
GeoHack	Map
Google Maps	Map
OpenStreetMap	Map
ACME Mapper	Map
Apple Maps	Map
HERE	Map
MapQuest	Map
Waze	Map
Wikimapia	Map

Structure of Czech heritage data in Wikidata


1) Is/is not a cultural monument

- Property „heritage status“
 - P1435: Q385405 – cultural monument
 - P1435: Q649434 – national cultural monument
 - Total: **41026 items** (<http://tinyurl.com/y846jayn>)
 - From that: 30 movable items (data not available to public due to risk of theft)
 - Possible identifiers
 - **From/until**
 - **Applies to part**
- For items on parts of cultural items
 - P1435: [part of a cultural monument](#)
 - Qualifier: „of“
 - In total: **595 items** (<http://tinyurl.com/yb6e2qnl>)

Uherčice (Q1440293)

památkový status	 Národní kulturní památka České republiky
	od 12. 2. 1973
	▶ 1 reference

Letohrádek královny Anny (Q2424070)

památkový status	 část kulturní památky České republiky
	čeho Pražský hrad
	▼ 0 referencí

1) Is/is not a cultural monument



Katedrála svatého Víta
[commons:Praga 0947 globus belarusi-2.jpg](#)



Letohrádek královny Anny
[commons:Pražský hrad, Letohrádek královny Anny 02.jpg](#)



Zahrada Na Valech
[commons:Praha, Zahrada Na Valech 01.jpg](#)



Matyášova brána
[commons:Prague Castle Mathias Gate.jpg](#)

All parts of the monument „Prague Castle“

<http://tinyurl.com/y8x7kcb5>

1) Is/is not a cultural monument

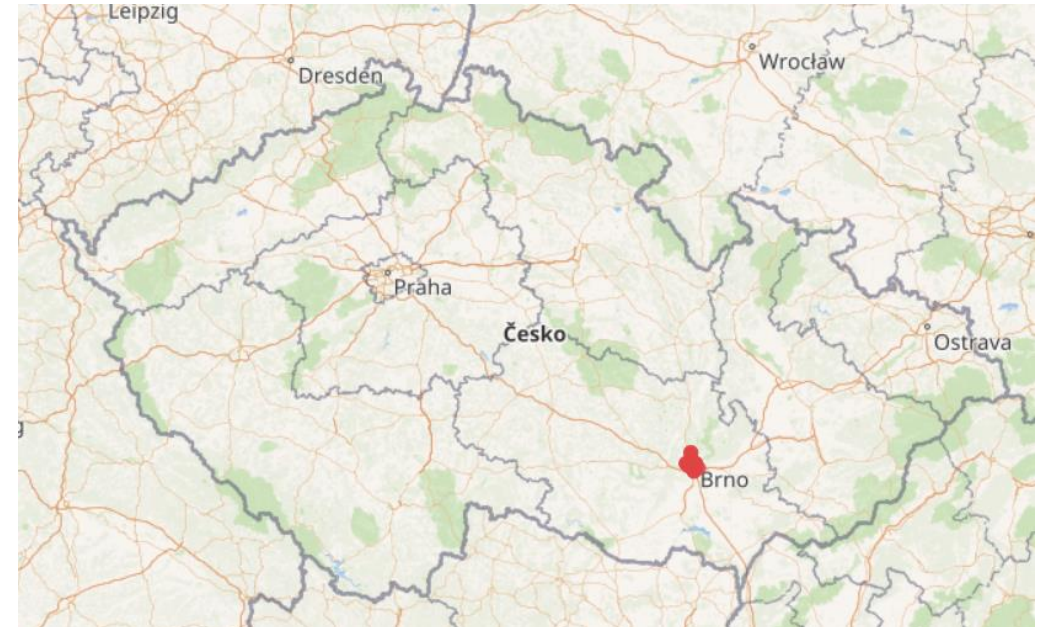
„Applies to part“ – equivalent of the string „applies to“ in available heritage data



Currently used for the front facades of buildings (sometimes protected)

<http://tinyurl.com/h3xbcnj>

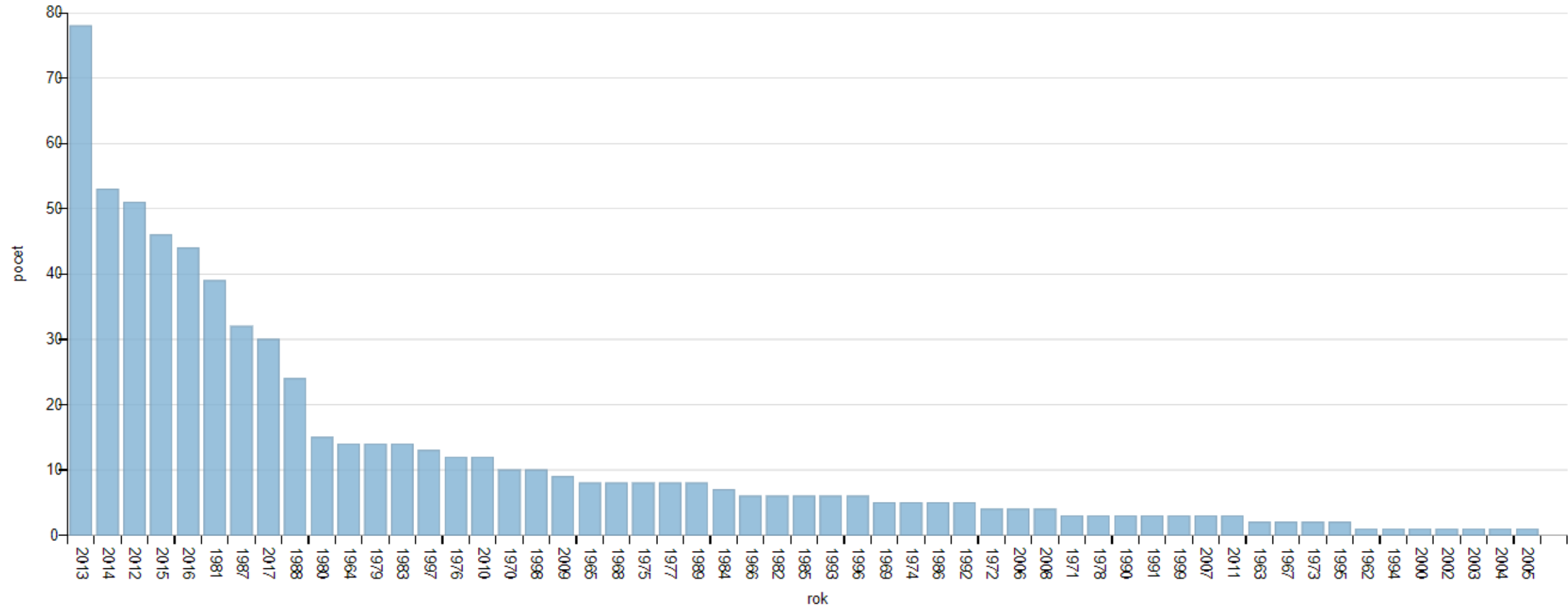
Should an item in Wikidata be the facade or the complete building? Or both?



1) Is/is not a cultural monument

Date of revoked monument protection

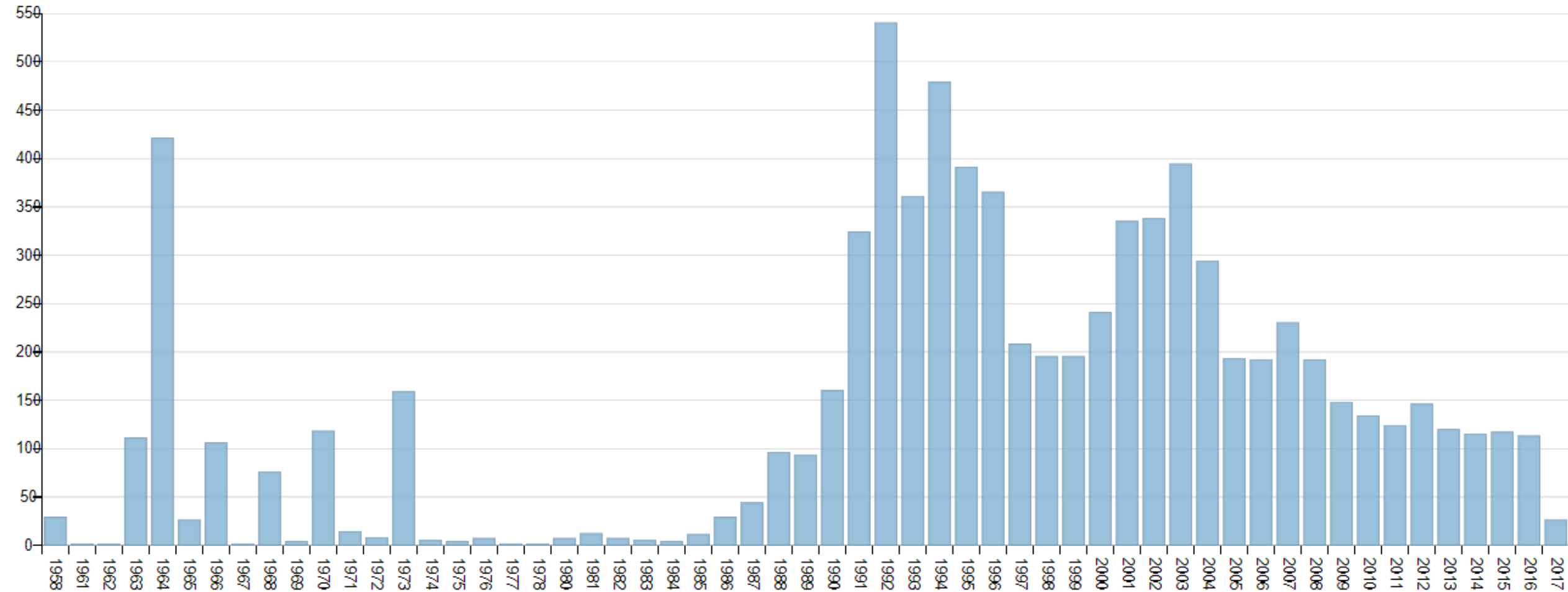
<http://tinyurl.com/y9ljx9dx>



1) Is/is not a cultural monument

Date when monument protection started

<http://tinyurl.com/y8tdowbz>



2) Heritage identifiers in CZ

Rejstříkové číslo ÚSKP (P762)

- 41019 items with this property – **all immovable monuments**
- Format 10141/4-4994-96006/5-397 or "100001"

Katalogové číslo Památkového katalogu (P4075)

- 40471 items with this property
- Format 1000000014 to 1999994429, exception: „N_152327_001“ (error)

2) Heritage identifiers in CZ

Duplicates in the heritage database – two identifiers for the same object

fotografie není veřejně
dostupná

objekt boží muka

katalogové číslo: 1000158334

památková ochrana: KP

číslo ÚSKP: 45864/3-5294

duplicita k rč. 32984/3-5293, kat. č. 144599

In total six items: <http://tinyurl.com/y7xlnts3>

We mark this using qualifier „object in position“: „duplicate“

2) Heritage identifiers in CZ

Missing catalogue numbers (P4075)

- **562** items

Failure to find the appropriate catalogue number marked by:

P762: [value]

P2241: Q42097313

- **Currently 95 items** <http://tinyurl.com/yd4u3tcc>, usually revoked cultural monuments
- **467 – not inspected yet: mostly manual work**
<http://tinyurl.com/yb5ntdyl>

Some objects do not have a database page in NHI catalogue but only a page in „Legal protection“ tab. How to link these? Such as

<https://www.wikidata.org/wiki/Q38488854> -

<http://pamatkovykatalog.cz/?order=relevance%3Adesc&mode=fulltext&keywords=54312%2F5-348&presenter=LegalStatesResults>

2) Heritage identifiers in CZ

Catalogue numbers with underscore

- Probably signify a part of a larger object (in some cases only this part is protected as the cultural monument)
- 419 items with catalogue number with underscore (<http://tinyurl.com/ybdrjnzo>)

2) Heritage identifiers in CZ

Two items with the same catalogue number (duplicates?)

Need to check regularly for duplicates (<http://tinyurl.com/y9vrpfq2>). Currently only 1 pair of items - ambiguous

- Dům dělnické kolonie čp. 935 (49163/5-1002, kat. 1000162286)
<https://www.wikidata.org/wiki/Q37805859>
- Dům dělnické kolonie čp. 936 (49164/5-1002, kat. 1000162286)
<https://www.wikidata.org/wiki/Q37805867>

3) Souřadnice památek

Monuments without coordinates

- **469 items – various reasons**
 - Location unknown (value „unknown“)
 - Objects missing in official databases and we don't know why

Monuments with two coordinates set, but far apart from each other:

<https://www.wikidata.org/wiki/Q38047303>

<http://tinyurl.com/ybajoqyc>

Several monuments with identical coordinates – major problém – usually arised in data from Czech Wikipedia tables:

<http://tinyurl.com/y7o2b23m>

Coordinates of an object can be far from a municipality in which (P131) the object is located

<http://tinyurl.com/y84aylp2>

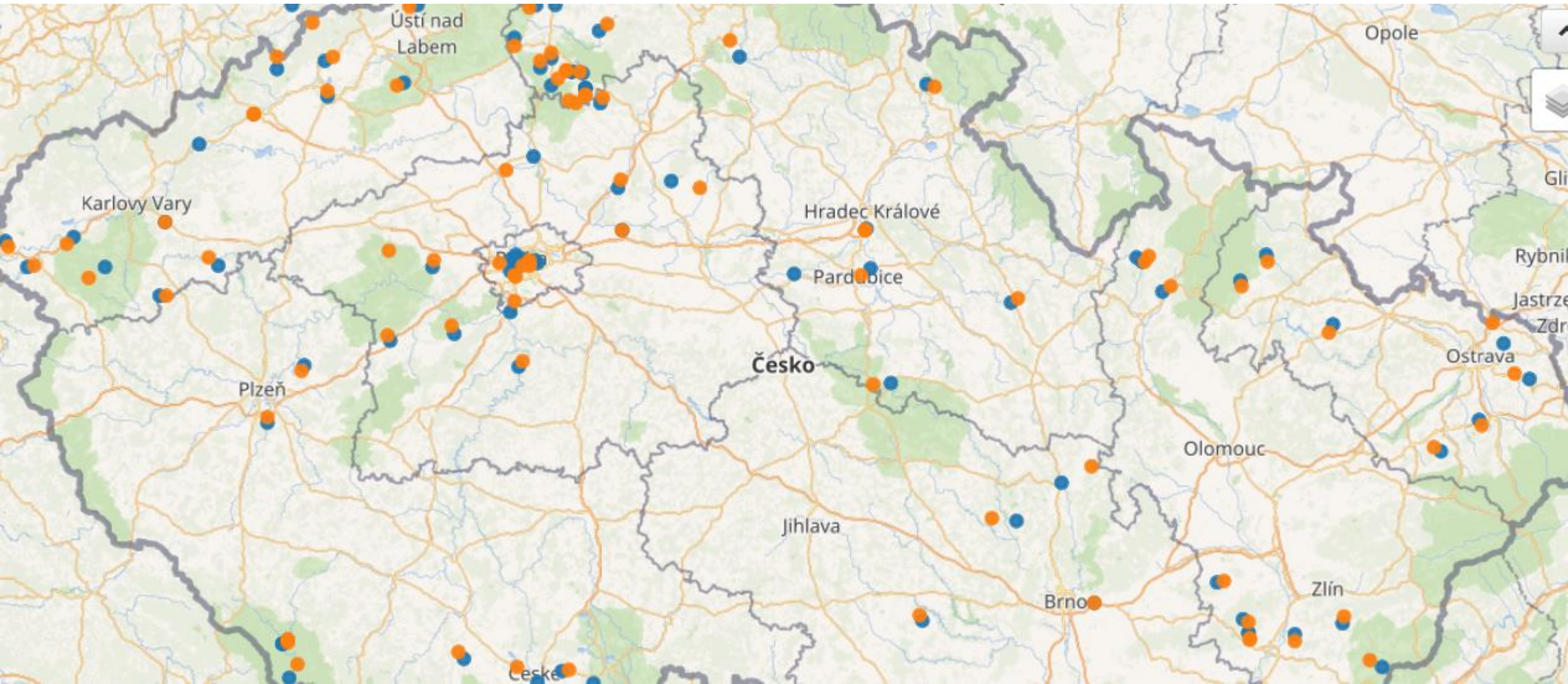
Not counting line constructions:

<http://tinyurl.com/y8o39sal>

3) Coordinates

<http://tinyurl.com/ybd6kmeh>

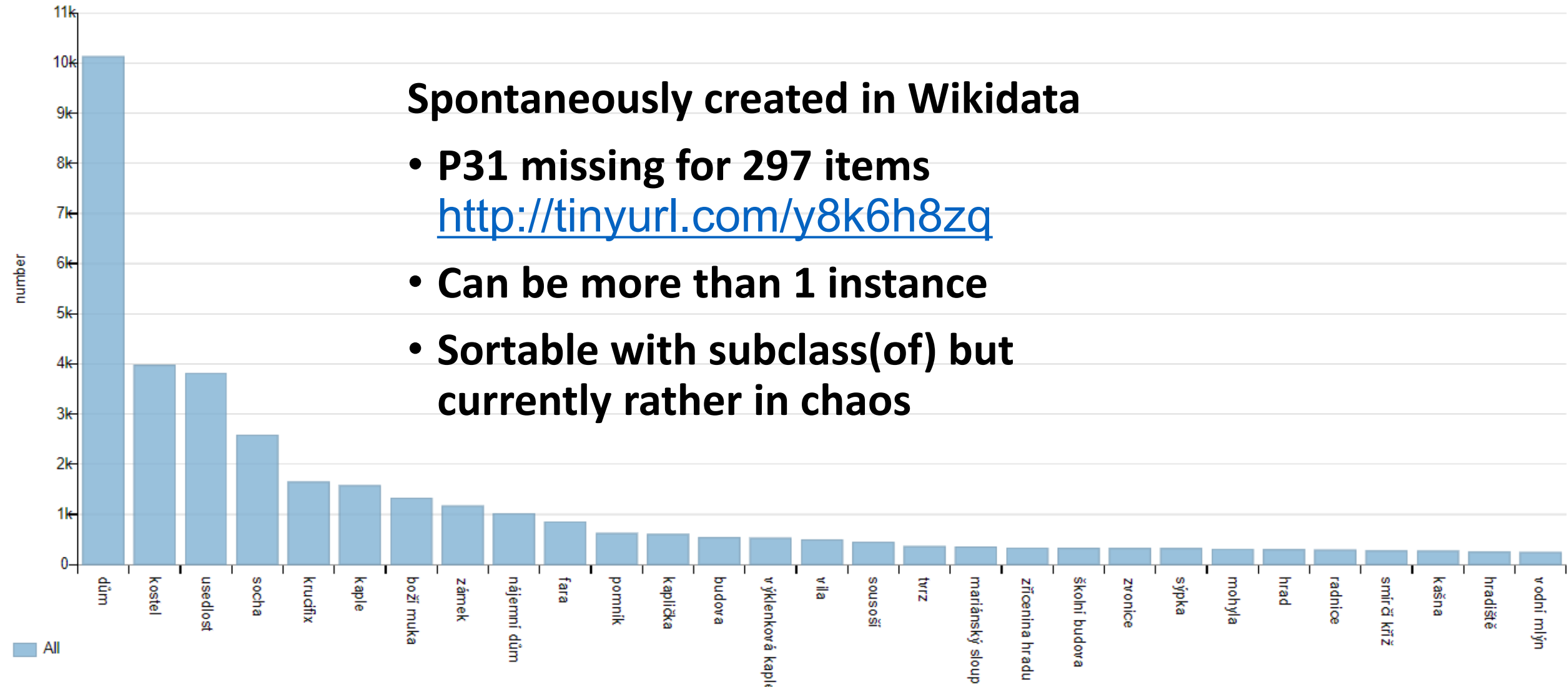
Biggest differences (distance-wise) between WD and NHI data



4) Instances (of) monuments

Spontaneously created in Wikidata

- **P31 missing for 297 items**
<http://tinyurl.com/y8k6h8zq>
- **Can be more than 1 instance**
- **Sortable with subclass(of) but currently rather in chaos**



100 rarest instances for Czech monuments

	zemědělská budova	Q1792401	skleník	lesní cesta
	simultánní kostel	žentour	Q1195705	psychiatrická nemocnice
	Viereckschanze	Q254036	lékařská fakulta	vládní budova
	řeka	úvozová cesta	oltář	fotbalové hřiště
	pastouška	mateřská škola	galerie umění	Q10273077
	Kartouza	ratejna	nivelační bod	Q1497375
	zrušená železniční trať	Q47036812	hospic	tunel
	dopravní společnost	zaniklá osada	observatoř	sluneční hodiny
	klášterní kostel	výklenek	bludiště	propustek
	barbakan	domov pro seniory	Q8719053	kulturní instituce
	roubená stavba	rybník	Svatý obrázek	válcovna
vojenské muzeum	sádka	nejvyšší soud	Q455403	chrám
plavební kanál	crag	vodní nádrž	Q14911880	manufaktura
letecká základna	vesnice	samota	univerzitní budova	Q350514
letišťe	budova banky	Patník	telefonní ústředna	obchodní podnik
zrušené chráněné území	komunitní centrum	mlat	porodnice	
toaleta	kolumbárium	Q677678	internát	
skalní stěna	vysoká pec	Sloup Proměnění Páně	ateliér	
ozubnicová dráha	budova parlamentu	rumpál	hrad bergfritového typu	
podzemí	kolotoč	miska	protipovodňová ochrana	
uhelný důl	fotbalový stadion	Q13424425	nástěnná malba	
vítězný sloup	bazén	Cheder	freska	
solárium	Vrchní soud	poustevena	Goethe-Institut	
chráněné území	konzervatoř	vodní dílo	řetězový most	

Instances grouped with subclass(of)

<http://tinyurl.com/y75tatz5>

class	classLabel	direct	ct
Q811979	stavba	7	36679
Q41176	budova	545	29365
Q3947	dům	10227	18248
Q838948	umělecké dílo	1	8156
Q16970	kostel	3978	6744
Q4989906	pomník	633	4552
Q860861	plastika	8	4419
Q179700	socha	2592	4316
Q5890494	usedlost	3901	3901
Q108325	kaple	1585	2785
Q4817	sloup	13	1719

Monuments which are gold panning sites <http://tinyurl.com/y7kts8qh>



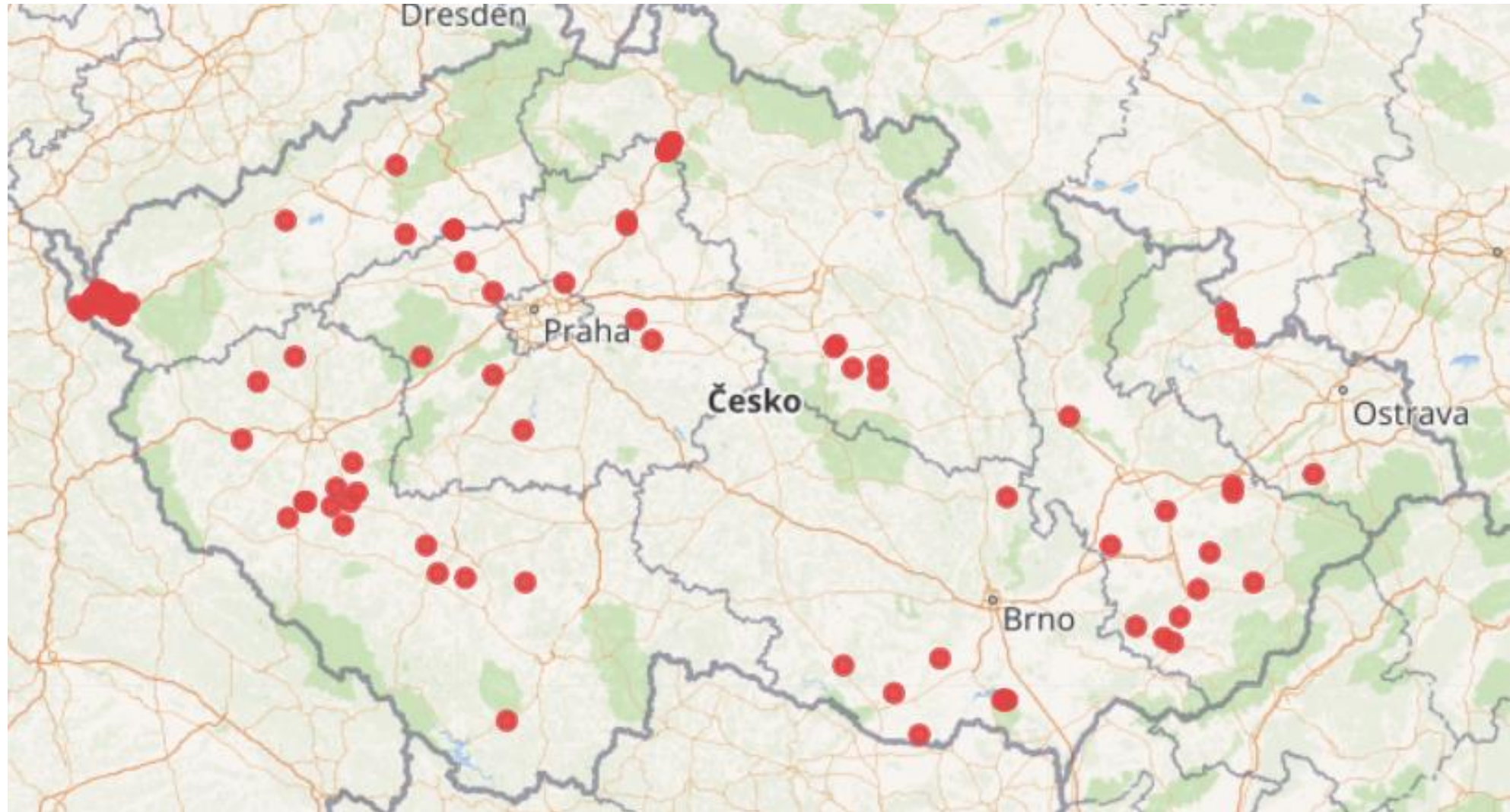
Wine houses

<http://tinyurl.com/yb8zvulg>



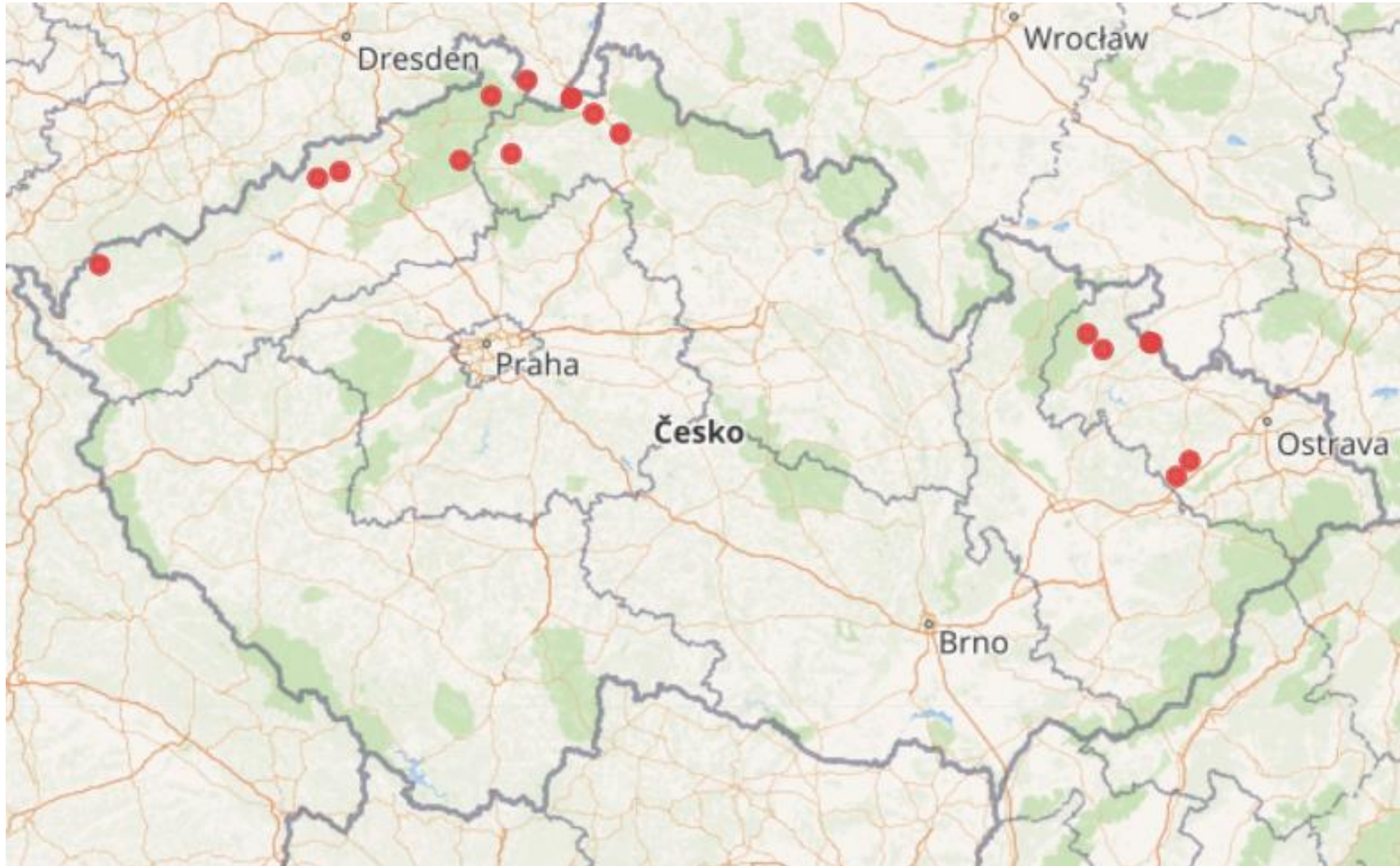
Ancient settlements

<http://tinyurl.com/yb95dsx5>



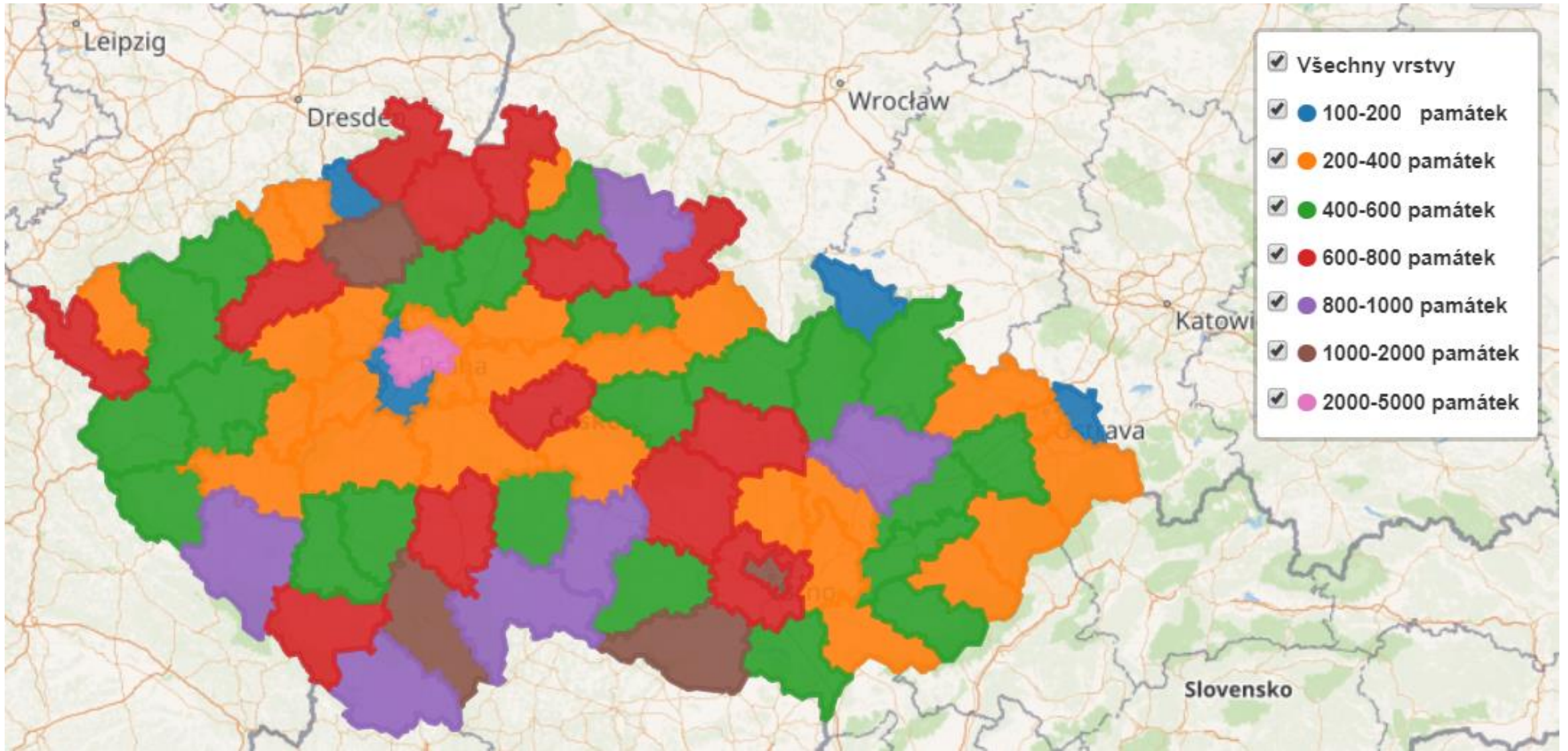
Textile factories

<http://tinyurl.com/ybbrd62y>



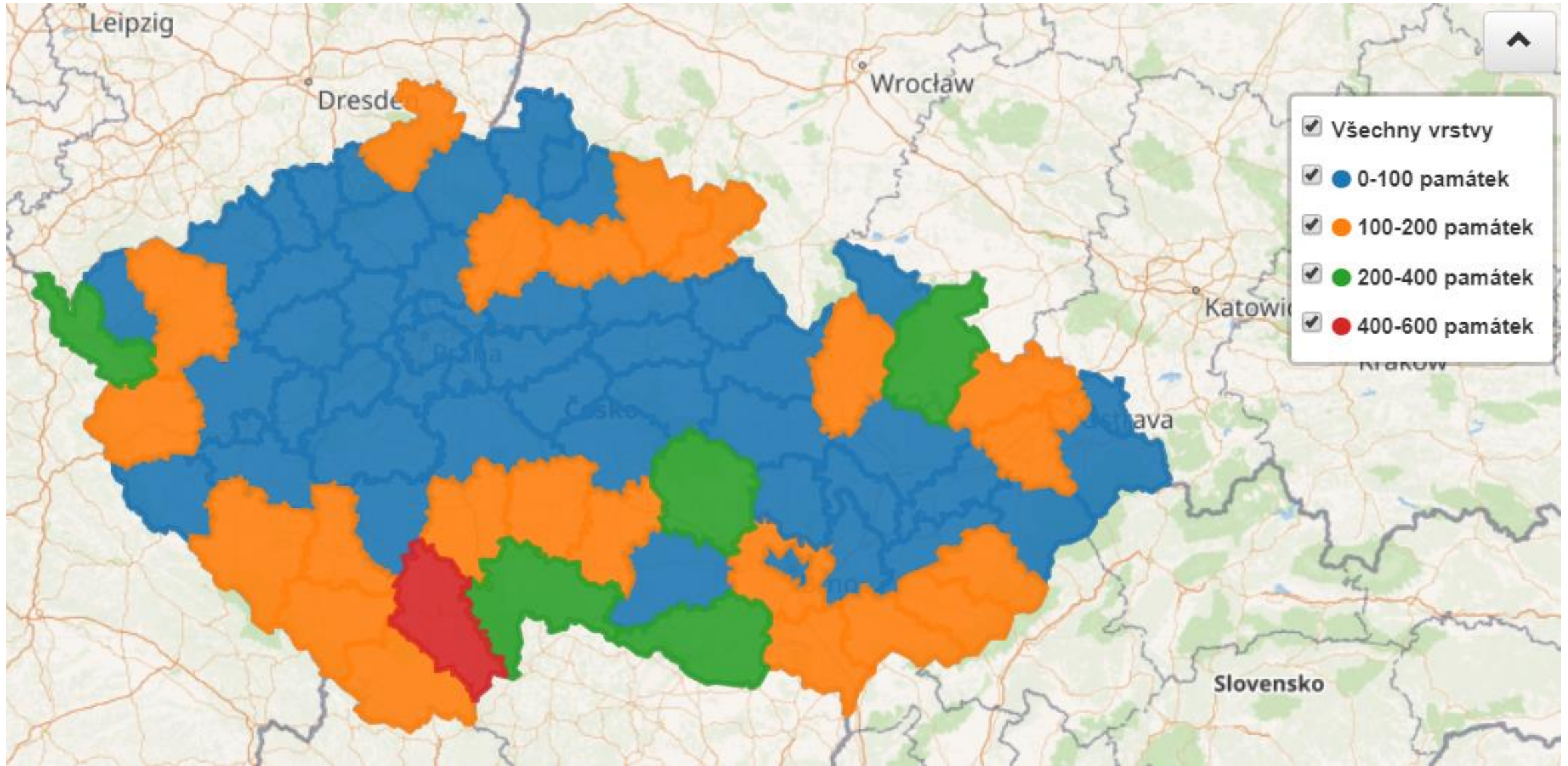
Number of monuments in each district

<http://tinyurl.com/ycdp6v4c>



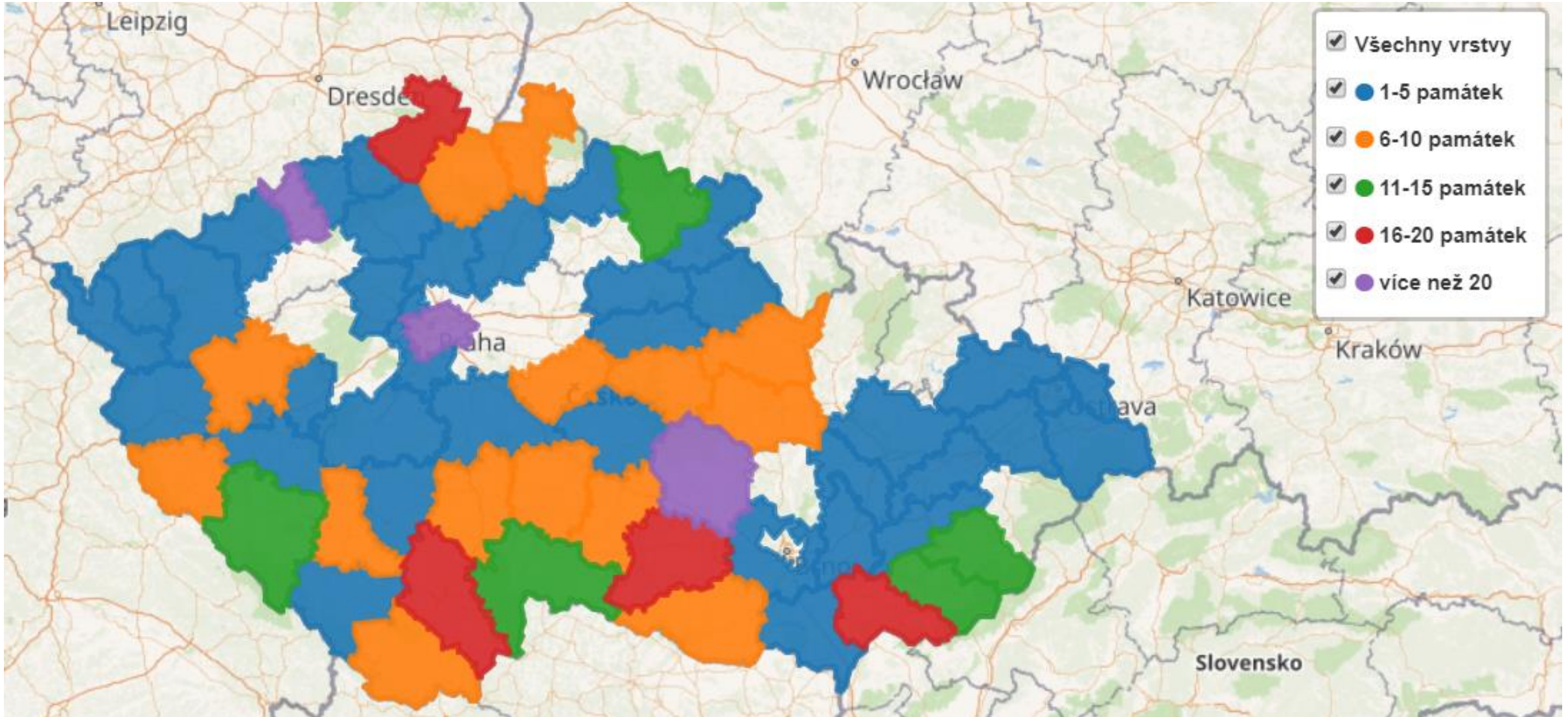
Number of monuments without picture in each district

<http://tinyurl.com/y9unrcfj>



Number of monuments in each district without coordinates

(<http://tinyurl.com/y8p2dy4z>)

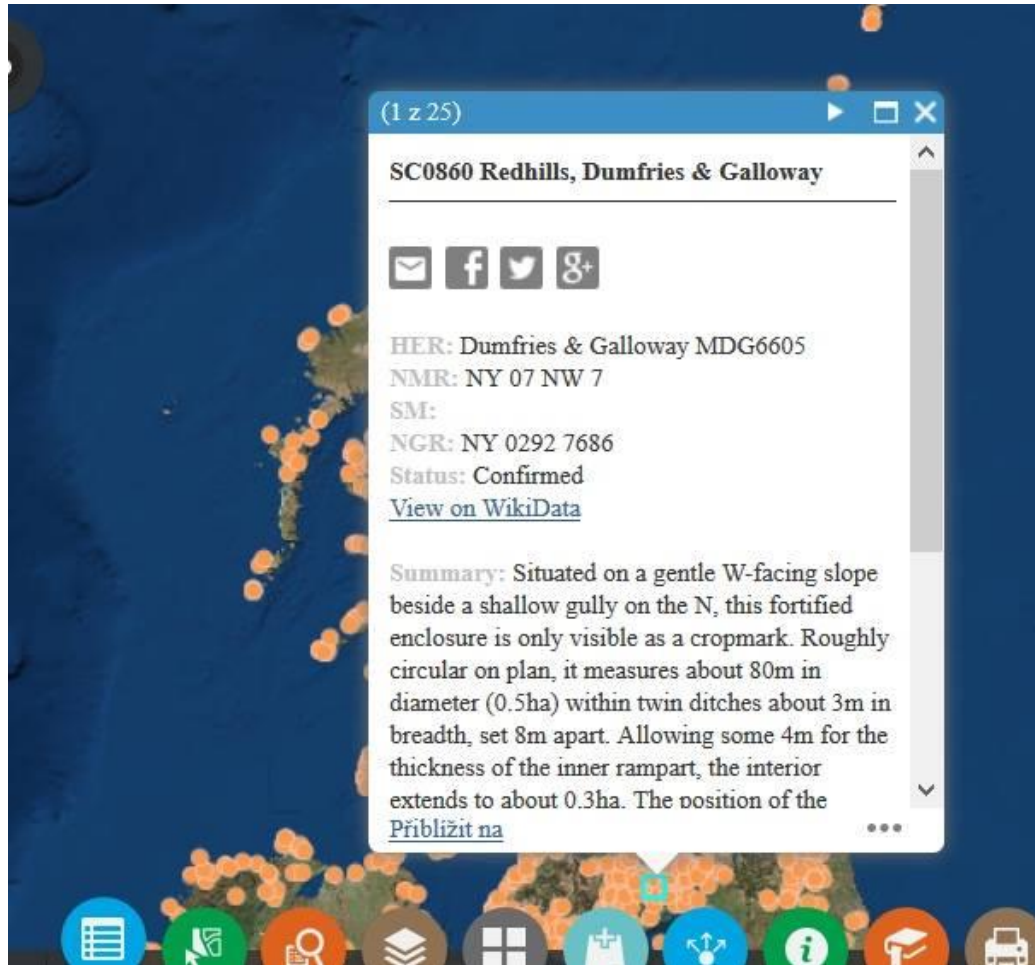


What next?

Separation of buildings from institutions which are housed in these buildings

- Often libraries, theatres, schools... mixing information on institutions and buildings
- We need to separate into distinct items but connect them through appropriate properties – using P466 or P159

Linking to Wikidata from external sites



New tools

<https://tools.wmflabs.org/hub/P4075:1000148122?site=wikidata>

Which other similar databases can we link to?

Drobné památky

Drobné památky Databáze Druhy památek Mapy

Rychlé hledání v databázi

Zadejte název okresu či obce, vsi. Například: Plzeň

 **Prázdné domy** + Přidat dům

Mapa **Databáze** **Autoři databáze** **Návody** **O nás** **Blog**

Změnit mapu

