

Prostorová data

Koncept + vizualizace

Agenda

- 1) Prostorová data – v čem jsou jiná?
- 2) Svět GISu × svět erka
- 3) Implementace v R / `{sf}` & přátelé
- 4) Kde prostorová data hledat a najít?

Prostorová data

Čím jsou jiná než ostatní?

Prostorová data

- Spojují v sobě složky:
 - Prostorovou / kde?
 - Datovou / co?
- Explicitní určení prostoru / souřadnice
 - 50.0842706N, 14.4411622E
- Implicitní určení prostoru / dohledatelný odkaz
 - náměstí Winstona Churchilla 1938/4, 130 00 Praha 3

Problémy s prostorem

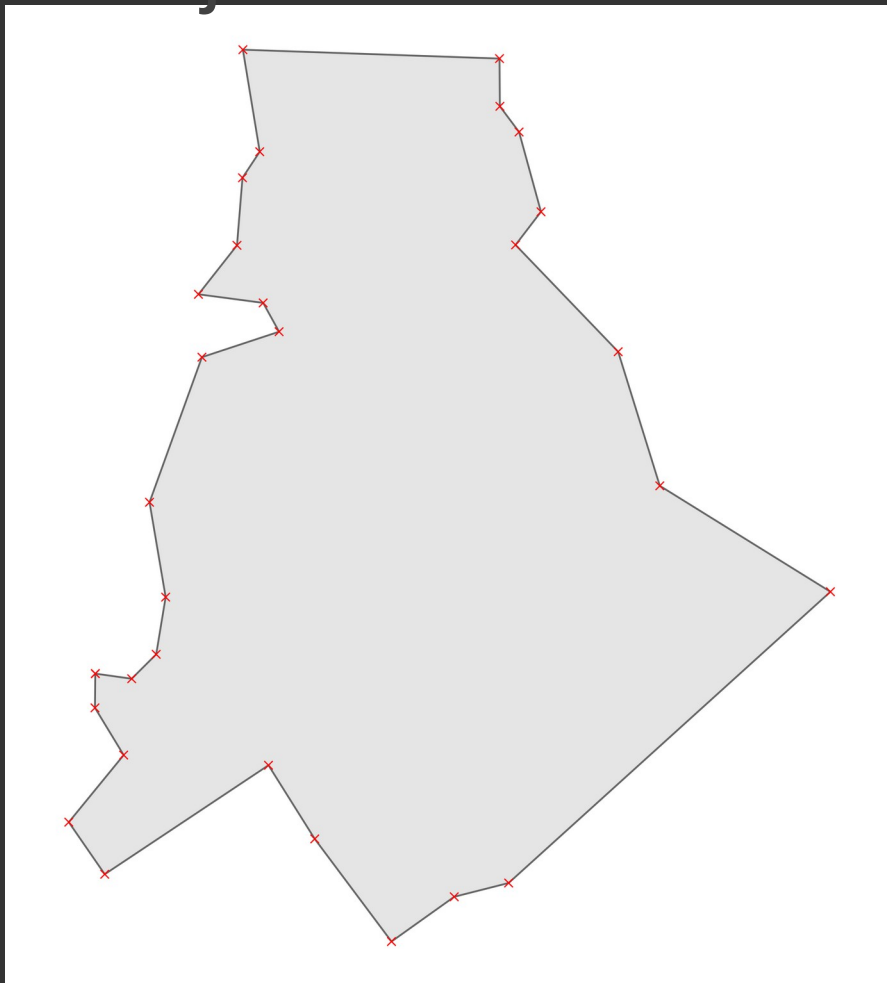
- Všechny mapy jsou rovinné (kromě glóbusu :)
- Zpravidla se vztahují k povrchu koule
- Povrch koule není euklidovský prostor
- Nelze se vyhnout zkreslení
- Neexistuje jedno "nejlepší" řešení

Základní přístupy

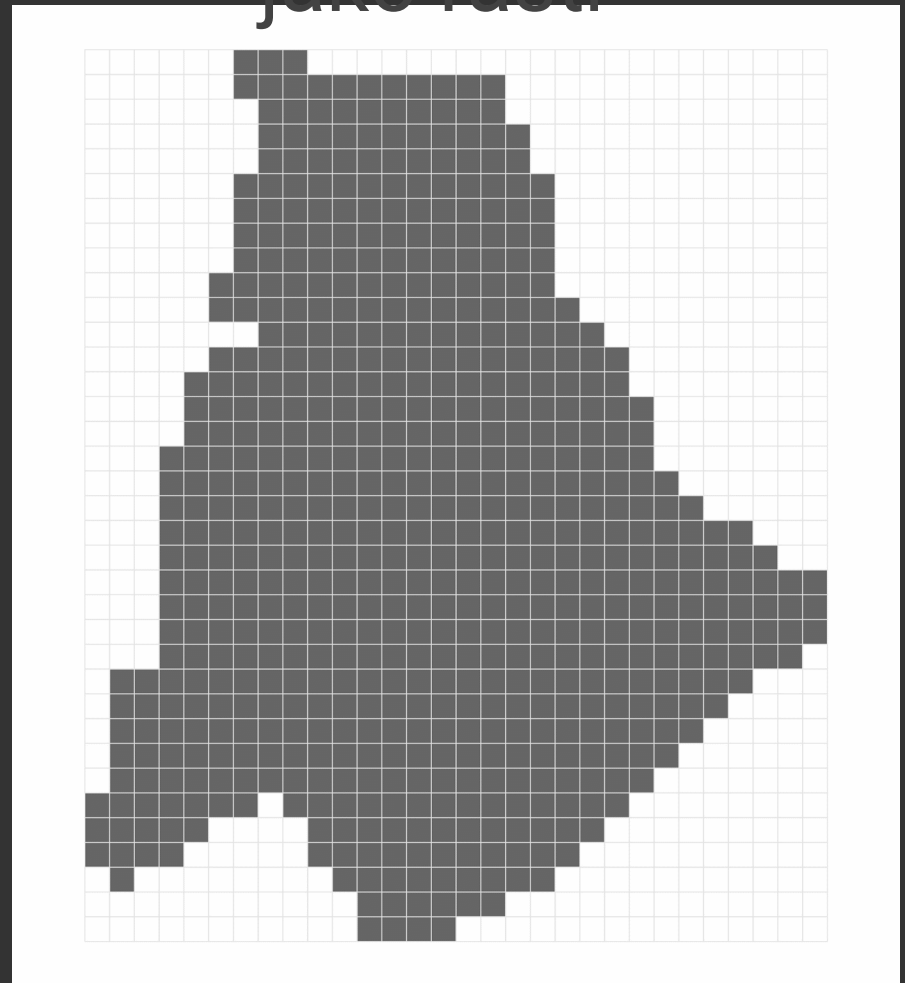
- Vektorová data
 - vycházejí z idealizovaných bodů
 - typická pro *lidmi zaměřené* objekty
- Rasterová data
 - mají jasně definované rozlišení
 - typická pro *remote sensing* (družice)

Rastery × vektory

Mecklenburg, NC
jako vektor



Mecklenburg, NC
jako rastr

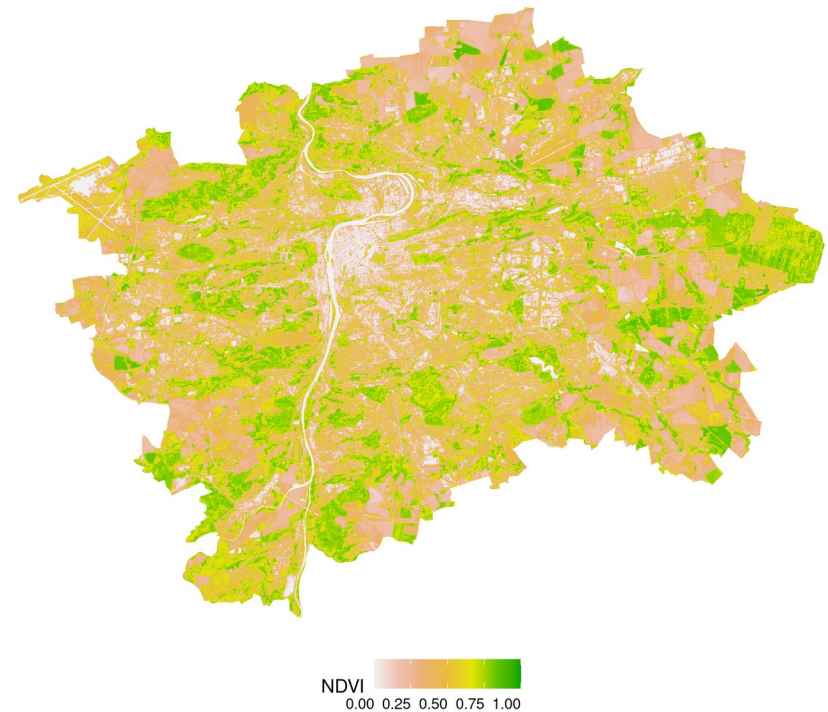


Rastery × vektory

Vektor: město Praha
a řeka Vltava



Raster: index NDVI
z družice Sentinel 2

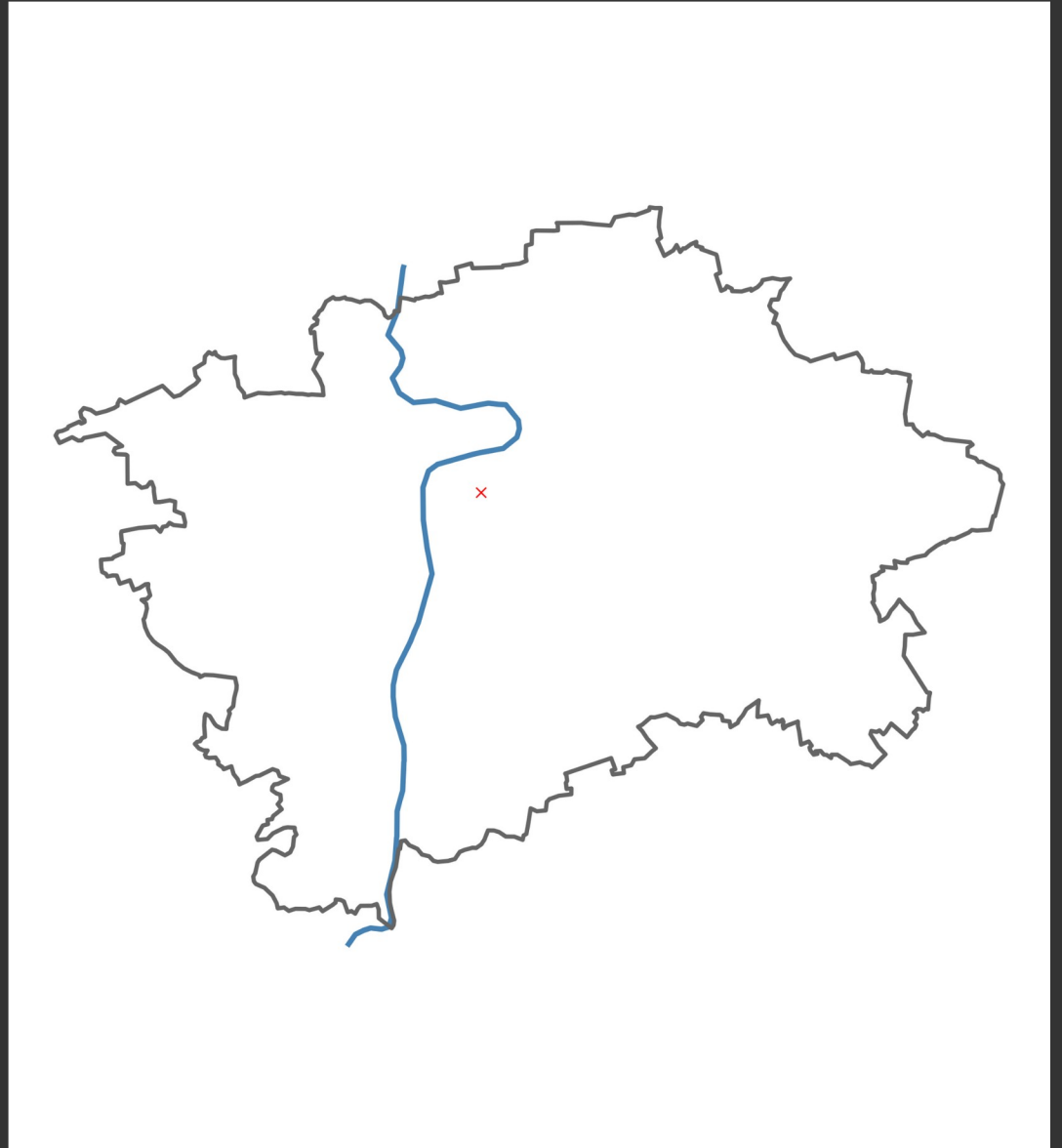


Rastery × vektory - použití

- Typické use cases rasterů:
 - přírodní jevy: počasí, geologie, nadmořská výška
 - výstupy z kamer / přístrojů
 - cena a rychlost jsou víc než přesnost
- Typické use cases vektorů:
 - administrativní uspořádání
 - geodezie / zaměření na místě
 - přesnost na metry, opakované použití

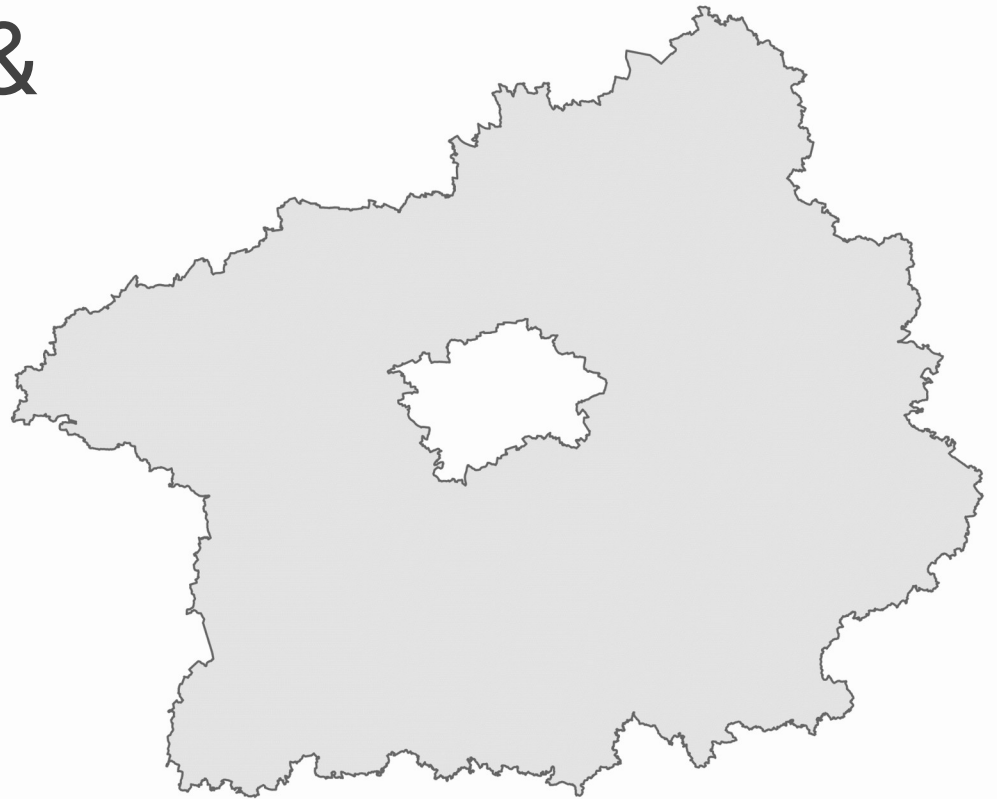
Tři úhelné kameny vektorů

- Bod
 - adresa, místo
- Čára
 - řeka, silnice
- Polygon
 - administrativní jednotky



Díra v polygonu

- Modelový příklad:
Středočeský kraj &
Praha
- Jeden polygon
může obsahovat
obecně mnoho
děr

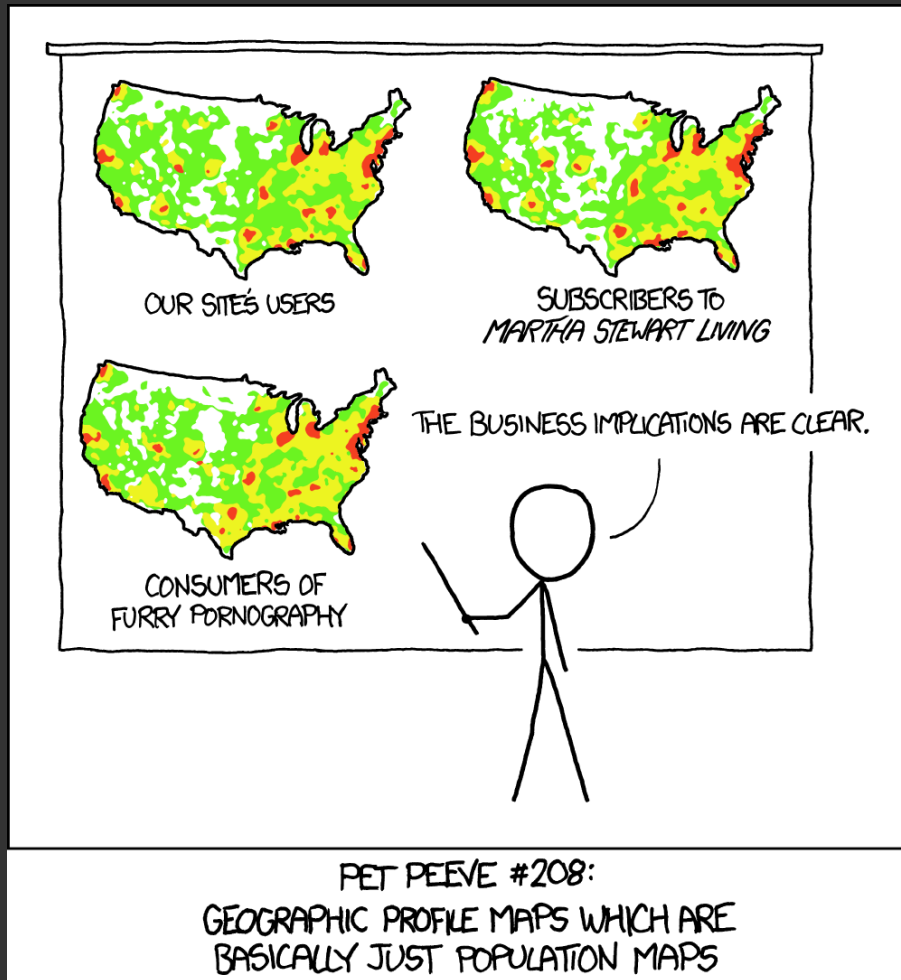


Multipolygon

- Modelový příklad: Japonské císařství
- Jeden státní útvar
- Na 364 ostrovech (polygonech)
- Analogicky fungují multi-čáry a body



Problém EKO dat



- Pracujeme s lidmi
- (takřka) všechny EKO metriky korelují s populací
- Řešení existuje:
 - Znát populaci (a pracovat per capita)
 - Normalizovat polygony na populaci (NUTS2, ZUJ)

Svět GIS

Svět R

Krátké srovnání

Světy GISu



- Komerční leader
 - ArcGIS Pro
 - konference, licence, podpora...



- OS Alternativa
 - QGIS / OSM
 - volnější seskupení oddělených projektů

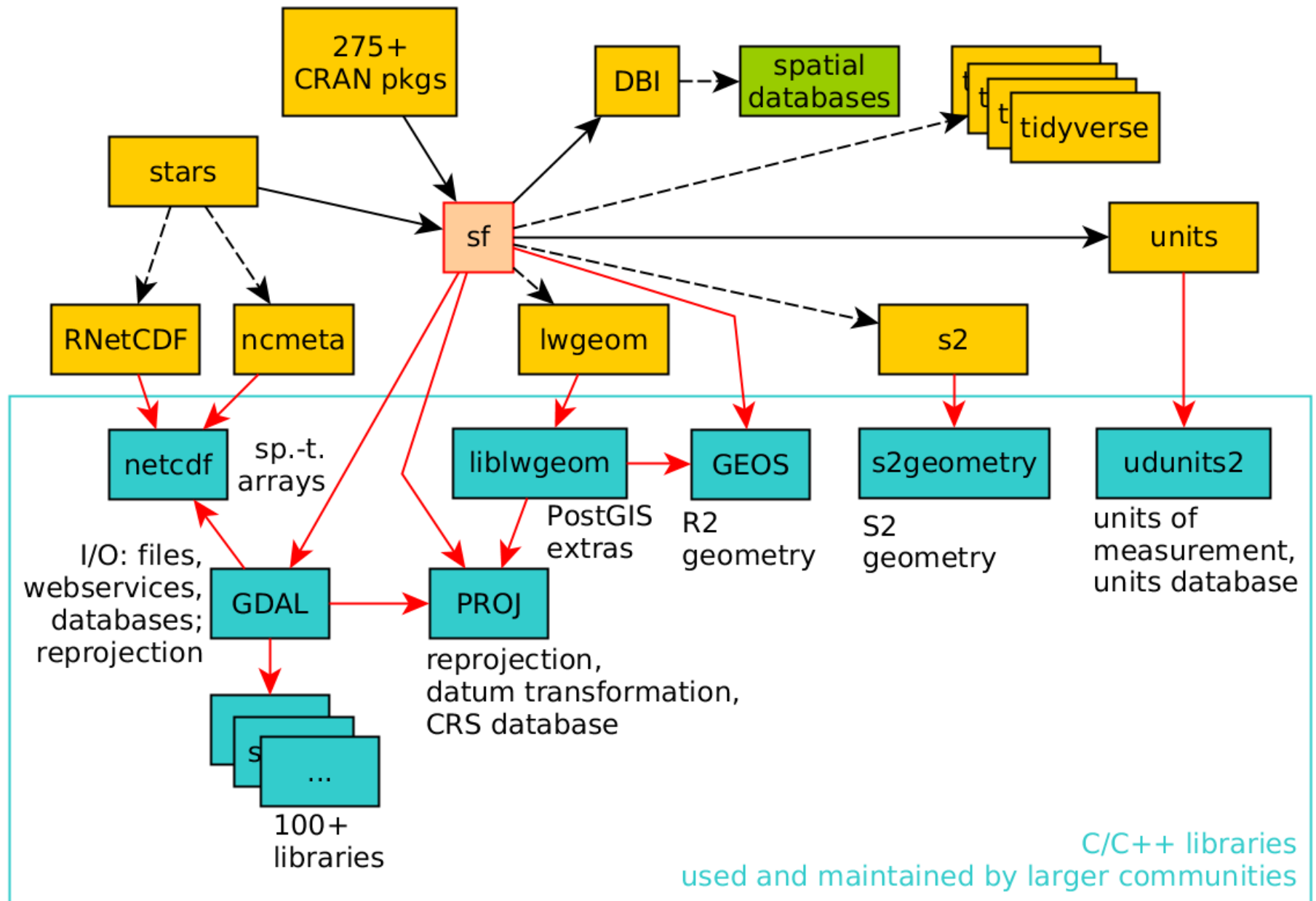
Uplatnění GIS světů

- ESRI:
 - status quo
 - vlády a ropa
- OSGeo
 - otevřenost
 - neziskovky
- Světy samy pro sebe:
 - IT svět: FAANG / Google & Apple maps
 - Legacy navigace (HERE, Garmin)

Společné body GISů a erka

- Erkové balíčky prostorových dat staví na světě OSGeo — nástroje GDAL, GEOS & PROJ
- Vstupní formáty jsou čitelné v obou světech (kompatibilita vstupů)
- Zeměkoule je stejná pro všechny

Závislosti kolem {sf}



Rozdíly GISů a erka

- GISové nástroje jsou typicky klikací
- erko je důsledně skriptovací
- GIS je optimalizován na grafické výstupy (a statistika je něco navíc)
- erko je optimalizované na statistiku (a grafické výstupy jsou něco navíc)

Kdy sahnout po...

- GISu

- když je to "zvykem domu"
- když je prioritou grafický výstup
- když je důležité data vidět

- erku

- když je to "zvykem domu"
- když je prioritou statistika
- když je vhodné opakovat běhy

ukázka QGIS

Intermezzo

Implementace v erku

High Level Overview

Implementace v erku

- Datové formáty: `{sf}`
- Statická prezentace: `{ggplot2}`
- Interaktivní prezentace: `{leaflet}`
- Česká geodata: `{RCzechia}`
- Geocoding + Open Street Map

{sf} datový formát

- Modifikovaný `data.frame`
- 1 zvláštní sloupec s prostorovými informacemi (zpravidla `geometry`)
- Všechny metody pro `data.frame` zůstávají v platnosti
- A k nim pár navíc... :)

{sf} datový formát

- `sf::st_as_sf()` vytvoří prostorový objekt z "obyčejného"
- `dplyr::inner_join()` propojí prostorovou a datovou složku
- `dplyr::filter()` vybere řádky splňující podmínky

Statická prezentace

- `ggplot2::geom_sf()` přímo podporuje data v `{sf}` formátu
- Všechny ostatní `{ggplot2}` techniky fungují analogicky — legendy, popisky, barvy a výplně

Typický {ggplot2} call

```
ggplot() +
```

```
  geom_sf(data = x, aes()))
```

- Estetika – `aes()` – má ggplot2 logiku (color, fill, label...)
- Škály – `*_scale_*` – opět obecně platné

Dynamická prezentace

- Balíček `{leaflet}` zpřístupňuje javascriptovou knihovnu `leaflet.js` do erka
- HTML výstup / interaktivita
- Syntaxe podřízená externí knihovně

Typický {leaflet} call

```
leaflet() %>%
```

```
  addProviderTiles() %>%
```

```
  addMarkers(data = x)
```

- Spojujeme pajpou z {magritte}
- Přidáváme tiles (basemap) + markers, circle markers nebo polygons

{RCzechia} české geo

- Administrativní jednotky – od obcí výše (ORP, okresy, kraje, republika)
- Klíče dle metodiky ČSÚ – snadné doplnění datové složky
- Navíc: řeky, lesy, chráněná území...

Řešený příklad

1-nezaměstnanost.R

{giscoR} EU geodata

- Administrativní jednotky EU
 - NUTS (kraje a výše)
 - LAU (okresy a obce)
- Státy světa
- Klíče dle metodiky Eurostatu

Řešený příklad

1-gisco.R

Geocoding

- Zaměření bodu – dohledání souřadnic
- Standardní problém = standardizovaná řešení (REST API) s erkovou implementací
- `RCzechia :: geocode()` API ČUZK
- `tidygeocoder :: geo_osm()` API OSM

Řešený příklad

1-geocode.R

OpenStreetMap z erka

- OpenStreetMap má Overpass API, které umožňuje dotazovat data
- Overpass API je do erka zpřístupněno přes `{osmdata}`
- Dokumentace na [wiki OSM](#)

Řešený příklad

1-pivo-žižkov.R

Další zdroje dat...

- Geoportál ČR
- Geoportál města Prahy
- Veřejný dálkový přístup ČUZK
- Google is your friend... :)

Formáty prostorových dat

- Shapefile (* .shp)
 - Starý, ale zaběhlý formát (dBase IV)
- Geopackage (* .gpkg)
 - Novější formát (sqlite)
- Geojson (* .json)
 - IT svět (plain text)

Formáty prostorových dat

- Všechny uvedené formáty jsou standardizované = čitelné v GIS nástrojích i erku `sf::st_read()`
- V erku navíc funguje standardní `saveRDS()` a `readRDS()` jako na všechny objekty

Příklad: Migrace

- Zakreslit krajské saldo migrace
- Problém:
 - Získat hodnoty migrace (datová složka)
 - Získat hranice krajů (prostorová složka)
 - Propojit data & prostor
 - Nakreslit obrázek (statický)

Řešený příklad

1-krajská-migrace.R

Příklad: Cenová mapa

- Zjistit nejdražší pražský pozemek
- Problém:
 - Načíst cenovou mapu (datová & prostorová složka v jednom)
 - Setřídít, vybrat nejdražší
 - Podat zprávu (interaktivně)

Řešený příklad

1-cenová-mapa.R

Klíčové body

- Typy prostorových dat
 - bod, čára, polygon + raster
- Jak data získat
 - Balíčkem, ze souboru, geokódováním
- Jak data zobrazit
 - staticky, dynamicky

Klíčové body

- Základní typy dat – vektor, rastr
- Načtení dat `sf::st_read()`
- Propojení dat `dplyr::*_join()`
- Statický obrázek
`ggplot2::geom_sf()`
- Dynamická prezentace `{leaflet}`