

A1 - Einleitung und Motivation

Jan-Philipp Kolb

22 Oktober 2018

Kleine Vorstellungsrunde

- Wo kommt Ihr her?
- Wo arbeitet und studiert Ihr?
- Wie beurteilt Ihr Eure Fähigkeiten mit R?
- Habt Ihr Erfahrungen mit anderen Programmiersprachen / Statistiksoftware? Wenn ja welche?
- Was sind Eure Erwartungen für diesen Kurs?

Disclaimer/ Informationen vorab

Normalerweise gibt es große Unterschiede bei Vorkenntnissen und Fähigkeiten, insbesondere bei diesem Kurs.

- Bitte gebt Bescheid, wenn es zu schnell oder zu langsam geht oder etwas unklar geblieben ist.
- Wenn es Fragen gibt - immer fragen
- Wenn Ihr etwas hinzuzufügen habt - sehr gerne
- In diesem Kurs gibt es viele **Übungen**, denn das Programmieren / die Nutzung von R lernt man am Ende nur allein.
- Ich habe viele **Beispiele** - probiert sie aus
- R macht mehr Spaß zusammen - arbeitet zusammen!

Das Thema Geodatenlandschaft

Georeferenzierung
von Daten

Situation und Zukunft der
Geodatenlandschaft in
Deutschland

Herausgegeben vom Rat für Sozial- und Wirtschaftsdaten

RatSWD.

Rat für Sozial- und
WirtschaftsDaten

GEFÖRDERT VOM

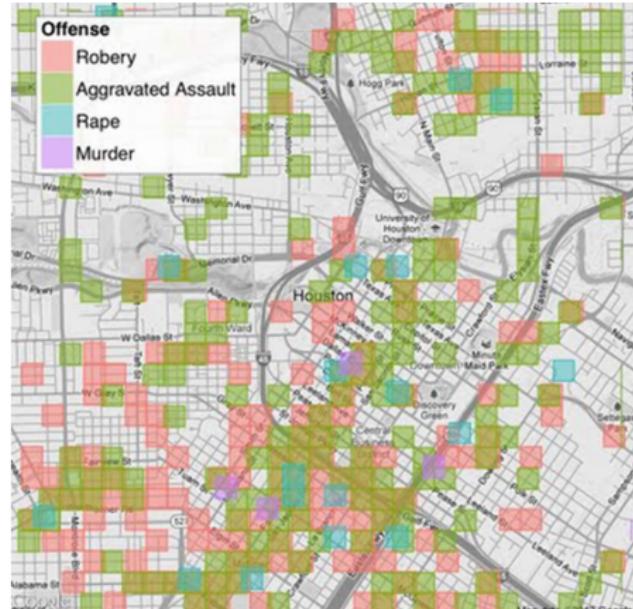
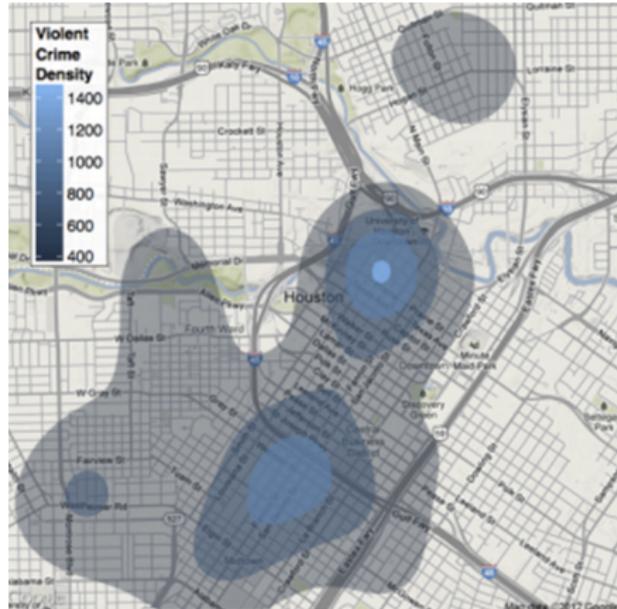


Bundesministerium
für Bildung
und Forschung

R-Pakete - Zum Download von Geo-Information

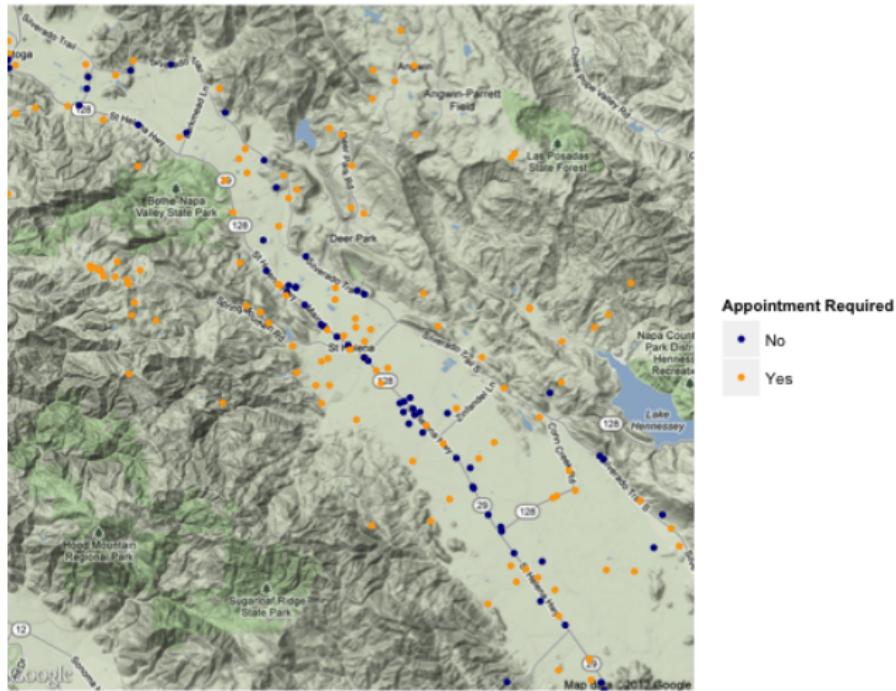
Das Paket `ggmap`

- David Kahle and Hadley Wickham: `ggmap` - Spatial Visualization with `ggplot2`



Worum geht es?

Weine probieren im Napa Valley?



Ziel dieses Kurses

Vorgestellt werden:

- Möglichkeiten für den Download, den Import, die Verarbeitung und die Visualisierung von Geodaten
- Quellen für Geodaten
- Die wichtigsten Programmierschnittstellen (APIs) um die Daten zu bekommen
- R-Pakete um diese Daten zu verarbeiten und zu visualisieren

Motivation

Motivation allgemein

- Raumbezug herstellen/nutzen
- Sekundäranalyse für bestehenden Daten
- Analysepotentiale der Geokodierung vorstellen
- Verbindung von sozial- mit raumwissenschaftlichen Daten

Motivation - Warum die Darstellung in Karten

- Darstellung in Karten ermöglicht besseres Verständnis von sozialwissenschaftlichen Phänomene - Attraktiver Output
- Durch die INSPIRE Richtlinie und *Collaborative Mapping* wächst der verfügbare Bestand an Geodaten.
- Daten sind oft frei verfügbar im Internet (z.B. Nutzung von APIs)
- Die Daten sind oft wenig oder gar nicht strukturiert, heterogen und oft nicht zur räumlichen Visualisierung vorgesehen, beinhalten aber implizit geographische Informationen (Web 2.0)

Laws of Spatial Science

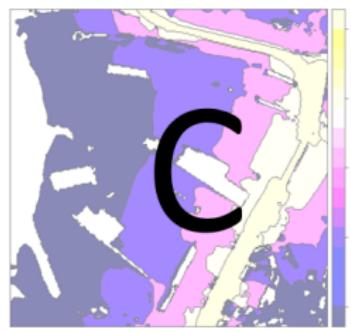
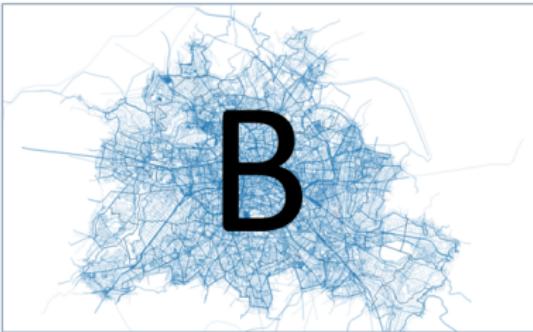
Tobler's law

everything is related to everything else, but near things are more related than distant things.

Spatial Turn

Spatial turn is a term used to describe an intellectual movement that places emphasis on place and space in social science and the humanities.

Datenquellen



Quellen: [Openstreetmap](#) (Graphiken A und B) und [Eionet](#) (Graphik C)

Ergebnisse des Zensus 2011 zum Download



Gemeindeebene

- Bevölkerung nach Geschlecht, Altersgruppe, Familienstatus, Staatsangehörigkeit und Religion

1 km² Raster

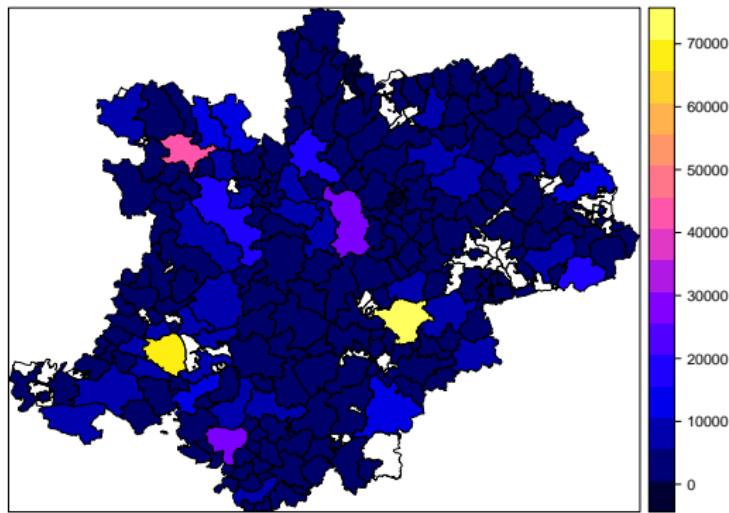
- Bevölkerung, Leerstandsquote, Wohnfläche und Haushaltsgröße

100 m² Raster

Bevölkerung

Zensus Ergebnisse

Beispiel Anteil der Personen aus EU27 Land an Einwohnerzahl pro Gemeinde in Oberfranken



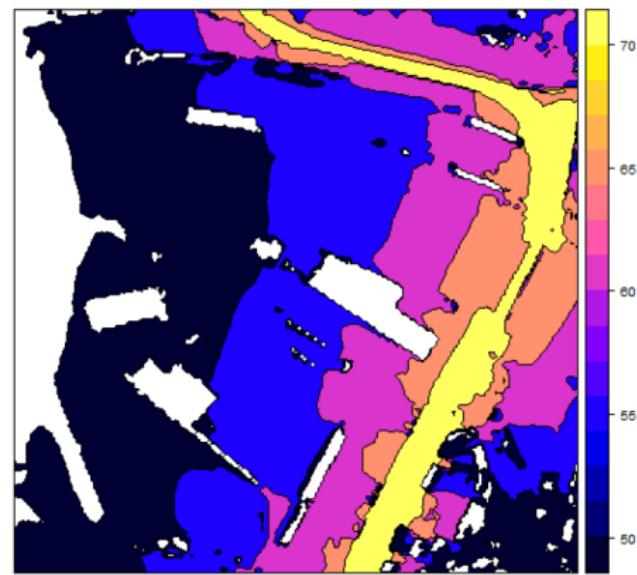
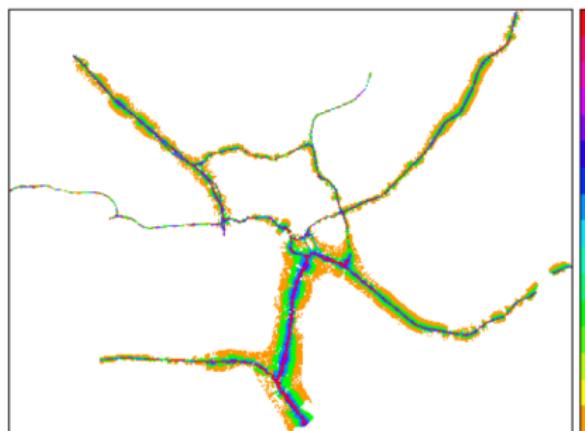
Datenquelle: Eionet

Eionet - Central Data Repository

- Europäisches Umweltinformations- und Umweltbeobachtungsnetz
- Qualitätsgesicherte Daten über den Zustand / Einflussfaktoren auf die Umwelt in Europa
- Strategische Lärmkartierungen

Lärmbelastung durch Schienenlärm

Beispiel: Lärmbelastung durch nächtlichen Schienenlärm in Hamburg



Das Openstreetmap Projekt...

Openstreetmap (OSM)

- Durch kollaboratives Mapping ist eine riesige Datenmenge zugänglich.
- Viele Menschen tragen jeden Tag Informationen bei.
- ... ermöglicht Zugang zu Big Data der Geographie.
- Die wachsende Menge an Geodaten wird von Freiwilligen gesammelt oder über Crowdsourcing gewonnen.

Drei wichtige OSM Objekttypen

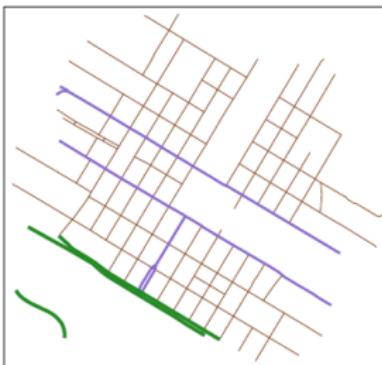
Vektordaten werden für den Betrachter dargestellt durch:

- Polygone sind als (eine Reihe von) verbundenen Punkten mit gleichem Start- und Endpunkt definiert. Die Ausrichtung verläuft gegen den Uhrzeigersinn

Punkte



Linien(abschnitte)



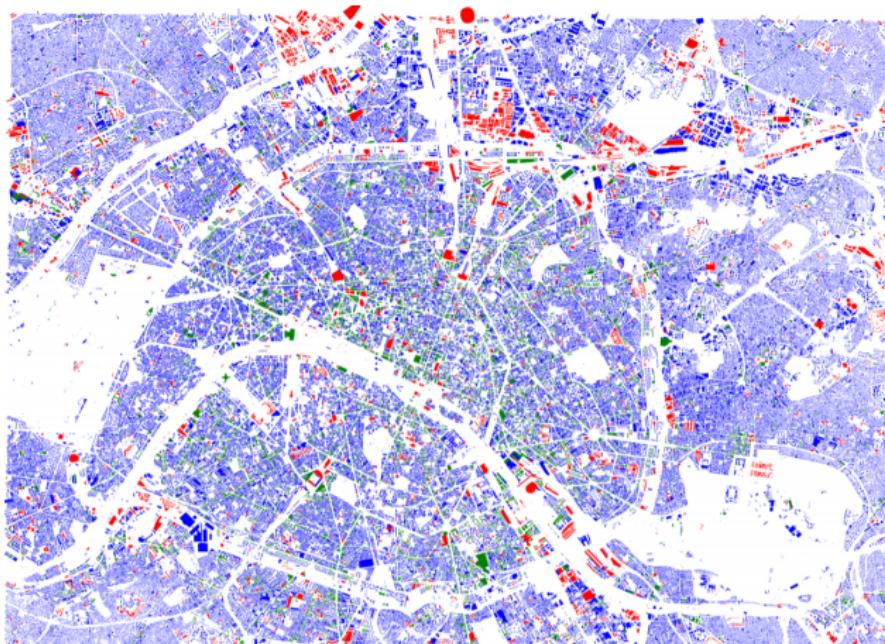
Polygone (Vielecke)



Paper mit OSM Daten

Studie von Gervasoni et al. (2018)

- Convolutional neural networks for disaggregated population mapping using open data



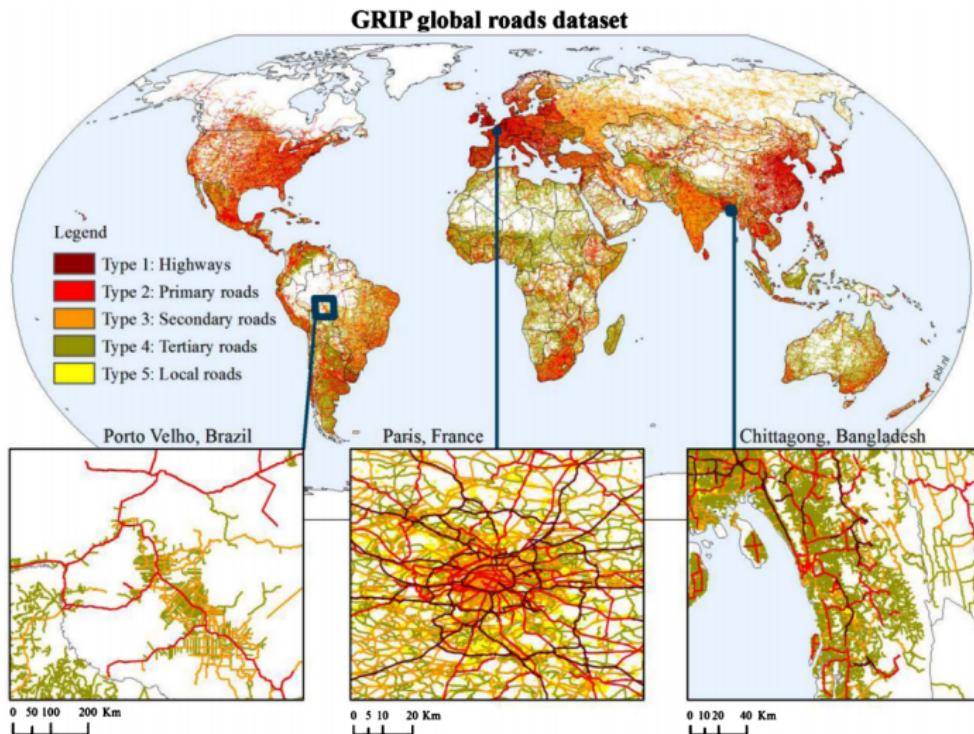
Was ist das Ziel - Straßen in Berlin

Dargestellt werden OpenStreetMap Daten, die mit der Overpass API heruntergeladen wurden.



Globale Muster der Straßeninfrastruktur

Studie von Johan Meijer et al.



Links mit Beispielen

- Shiny App zu **Indikatoren** für Europa
- Räumliche Visualisierung in den USA - **Walmarts in den USA**
- **Race Gap Police USA - Wahl USA**
- Zeit Artikel zum Zustand der **Eisenbahnbrücken**
- **Fahrradunfälle** in Berlin
- **Verteilung Fußballfans**
- **Plastiktüten im Meer**

Datenquellen:

- Datensätze zu **Pegelständen** in Deutschland
- Viele Datensätze auf **driven by data**

Resourcen

- Andreas Plank - **Grafiken und Statistik in R**

A1A Übung - zusätzliche Pakete

Geht auf <https://cran.r-project.org/> und sucht nach Paketen...

- die Umrisse der Länder der Welt enthalten
- mit denen man die Google maps API nutzen kann
- mit denen man Openstreetmap Daten bekommen kann

CRAN Task Views

- Bezuglich mancher Themen gibt es einen Überblick über alle wichtigen Pakete - (**CRAN Task Views**)
- Momentan gibt es 35 Task Views.
- Alle Pakete einer Task-View können mit folgendem Befehl installiert werden: **command:**

```
install.packages("ctv")
library("ctv")
install.views("Spatial")
```