

OpenStreetMap und R

Integration freier Geodaten in die Lingua franca der Statistik

Thomas Schlesinger

Institut für Statistik, LMU München

21.03.2012



Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - `summary`, `find` und `subset`
 - `plot`
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - plot
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

R Fakten

- Programmiersprache und -umgebung für die statistische Datenanalyse
- effektive Datenanalyse und Graphenerstellung
- etwa 250.000 regelmäßige Benutzer
- Open Source (über 3600 packages) von Statistikern für Statistiker
- Vorreiterrolle im Entdecken neuer Statistischer Methoden



Beispielcode (Vektor):

```
> a <- c(1, 5, 9)
> a
[1] 1 5 9
```

Funktionen:

```
> sqrt(a)
[1] 1.000000 2.236068 3.000000
> sqrt( a[3] )
[1] 3
> a * 5
[1] 5 25 45
> sum(a)
[1] 15
```

Beispielcode (Vektor):

```
> a <- c(1, 5, 9)
> a
[1] 1 5 9
```

Funktionen:

```
> sqrt(a)
[1] 1.000000 2.236068 3.000000
> sqrt( a[3] )
[1] 3
> a * 5
[1] 5 25 45
> sum(a)
[1] 15
```

Beispielcode eines data.frame:

```
> cars[1:6,]
```

```
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
```

```
> cars$speed[1:5]
```

```
[1] 4 4 7 7 8
```

```
> summary(cars)
```

```
      speed      dist
Min.   : 4.0    Min.   :  2.00
1st Qu.:12.0    1st Qu.: 26.00
Median :15.0    Median : 36.00
Mean   :15.4    Mean   : 42.98
3rd Qu.:19.0    3rd Qu.: 56.00
Max.   :25.0    Max.   :120.00
```

Beispielcode eines data.frame:

```
> cars[1:6,]
```

```
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
```

```
> cars$speed[1:5]
```

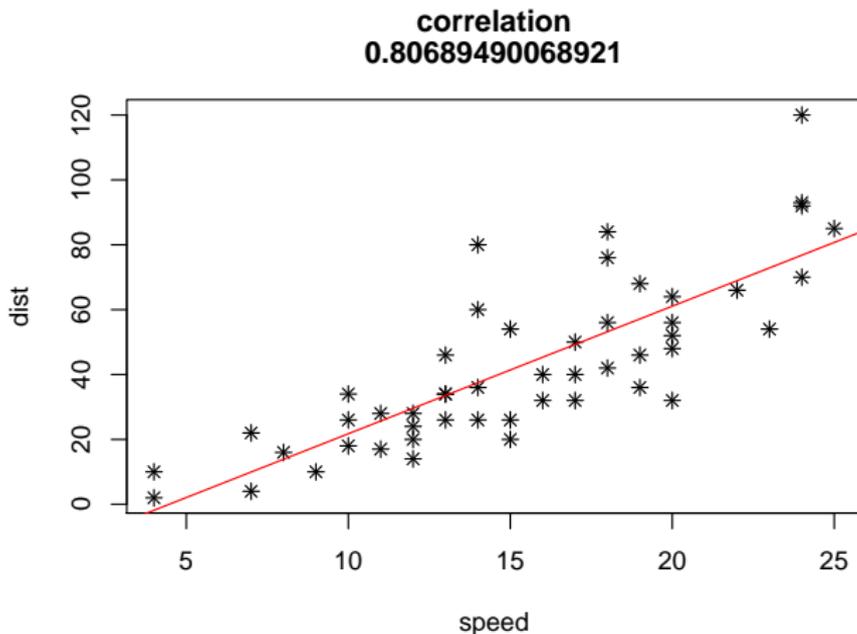
```
[1] 4 4 7 7 8
```

```
> summary(cars)
```

```
      speed      dist
Min.   : 4.0    Min.   : 2.00
1st Qu.:12.0    1st Qu.: 26.00
Median :15.0    Median : 36.00
Mean   :15.4    Mean   : 42.98
3rd Qu.:19.0    3rd Qu.: 56.00
Max.   :25.0    Max.   :120.00
```

Beispiel für einen plot

```
> plot(cars, pch=8)  
> abline( lm(dist~speed, data=cars), col=2)  
> title(c("correlation", cor(cars$dist, cars$speed)))
```



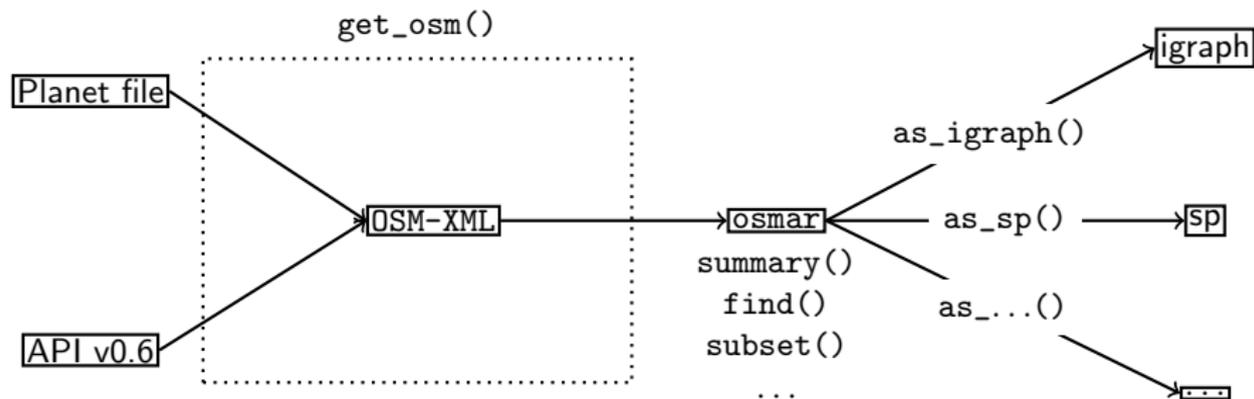
Gliederung

- 1 Einführung in R
- 2 osmar: Einführung**
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - plot
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

osmar-Fakten

- Ziel** Daten der OpenStreetMap-Community dem R-User zur Verfügung stellen
- Paket für R
 - **OpenStreetMap and R** = osmar
 - Im Rahmen der BachelorArbeit „**OpenStreetMap in R**“ entstanden
 - gemeinsame Arbeit zwischen Manuel Eugster und Thomas Schlesinger
 - Version 1.0 am 01.09.2011 auf CRAN hochgeladen (<http://cran.r-project.org/web/packages/osmar/>)

Schema des osmar-Konzepts



Struktur der OpenStreetMap-Datenbank (*kurz*)

- besteht aus Node, Way & Relation
- Alle Objekte enthalten
 - Attribute (ID, user, timestamp, version, ...)
 - Tags der Form key=value (Bspl: highway=traffic_signal)
- Node: geographische Informationen (latitude & longitude)
- Way & Relation: References (Verbindung zu anderen Objekten)

Struktur der OpenStreetMap-Datenbank (*kurz*)

- besteht aus Node, Way & Relation
- Alle Objekte enthalten
 - Attribute (ID, user, timestamp, version, ...)
 - Tags der Form key=value (Bspl: highway=traffic_signal)
- Node: geographische Informationen (latitude & longitude)
- Way & Relation: References (Verbindung zu anderen Objekten)

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen**
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - plot
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen**
 - Datenimport und osmar-Objekt
 - `summary`, `find` und `subset`
 - `plot`
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

Datenimport

- 2 Möglichkeiten: Über API 0.6 oder Planet files

- Bspl mit API 0.6:

```
> osmsource_api()$url
```

```
[1] "http://api.openstreetmap.org/api/0.6/"
```

```
> bb<-center_bbox(center_lon=12.232221, center_lat=51.840577,  
+                width=2000, height=2000)
```

- `.../?bbox=12.217712,51.831589,12.246730,51.849564`

```
> dессau<-get_osm(bb, source= osmsource_api())
```

- `dессau` ist nun ein Objekt der Klasse `osmar`

```
> class(dессau)
```

```
[1] "osmar" "list"
```

Datenimport

- 2 Möglichkeiten: Über API 0.6 oder Planet files

- Bspl mit API 0.6:

```
> osmsource_api()$url
```

```
[1] "http://api.openstreetmap.org/api/0.6/"
```

```
> bb<-center_bbox(center_lon=12.232221, center_lat=51.840577,  
+                width=2000, height=2000)
```

- .../?bbox=12.217712,51.831589,12.246730,51.849564

```
> dessau<-get_osm(bb, source= osmsource_api())
```

- dessau ist nun ein Objekt der Klasse **osmar**

```
> class(dessau)
```

```
[1] "osmar" "list"
```

Inhalt des osmar-Objektes "dessau"

```
List of 3
```

```
$ nodes      :List of 2
```

```
..$ attrs:'data.frame': 10415 obs. of 9 variables:
```

```
..$ tags :'data.frame': 2809 obs. of 3 variables:
```

```
$ ways       :List of 3
```

```
..$ attrs:'data.frame': 2323 obs. of 7 variables:
```

```
..$ tags :'data.frame': 6186 obs. of 3 variables:
```

```
..$ refs :'data.frame': 13805 obs. of 2 variables:
```

```
$ relations:List of 3
```

```
..$ attrs:'data.frame': 67 obs. of 7 variables:
```

```
..$ tags :'data.frame': 257 obs. of 3 variables:
```

```
..$ refs :'data.frame': 5326 obs. of 4 variables:
```

Attribute der nodes (*Ausschnitt*)

```
> dessau$nodes$attrs[1:6,]
```

	id	lat	lon	user	uid	visible
1	23656947	51.83670	12.21799	andy-de	101570	true
2	25529263	51.83101	12.24278	andy-de	101570	true
3	25808609	51.84020	12.23441	geozeisig	66391	true
4	25808610	51.83854	12.23359	BurnyB	83876	true
5	25808612	51.83675	12.23289	BurnyB	83876	true
6	25808613	51.83637	12.23260	andy-de	101570	true
	version	changeset	timestamp			
1	5	8305421	2011-05-31	20:30:33		
2	3	6657952	2010-12-14	12:48:02		
3	5	10079920	2011-12-10	10:56:30		
4	3	8829285	2011-07-25	19:11:37		
5	3	8829285	2011-07-25	19:11:37		
6	4	6576182	2010-12-07	15:45:49		

tags und references der Ways (*Ausschnitt*)

```
> dessau$ways$tags[1:4,]
```

	id	k	v
1	109227254	highway	footway
2	109227254	surface	unpaved
3	87285115	addr:housenumber	70
4	87285115	building	yes

```
> dessau$ways$refs[1:4,]
```

	id	ref
1	109227254	1251020574
2	109227254	1251020524
3	109227254	1251020596
4	109227254	287412505

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - **summary, find und subset**
 - plot
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

summary

Zusammenfassung mittels `summary`

- Ausführung mit

```
> summary(dessau)
```
- Ausgabe der BoundingBox
- Auflistung der häufigsten key=value -Kombinationen

	Key	Value	Freq
1	building	yes	1302
2	highway	footway	214
3	highway	service	210
4	highway	residential	187
5	addr:postcode	06846	149
6	surface	paved	143

find & subset

- Finden der ID's per `find(...)`
 - > bench_ids<-find(dessau, node(tags(k=="amenity" & v=="bench")))
 - > bench_ids[1:5]
 - [1] 417441492 417441496 417441502 417441506 417441520
- Teilmenge von `dessau` mittels `subset(...)` erstellen
 - > bench_dessau<- subset(dessau, node_ids=bench_ids)
 - > bench_dessau
 - osmar object
 - 51 nodes, 0 ways, 0 relations

find & subset

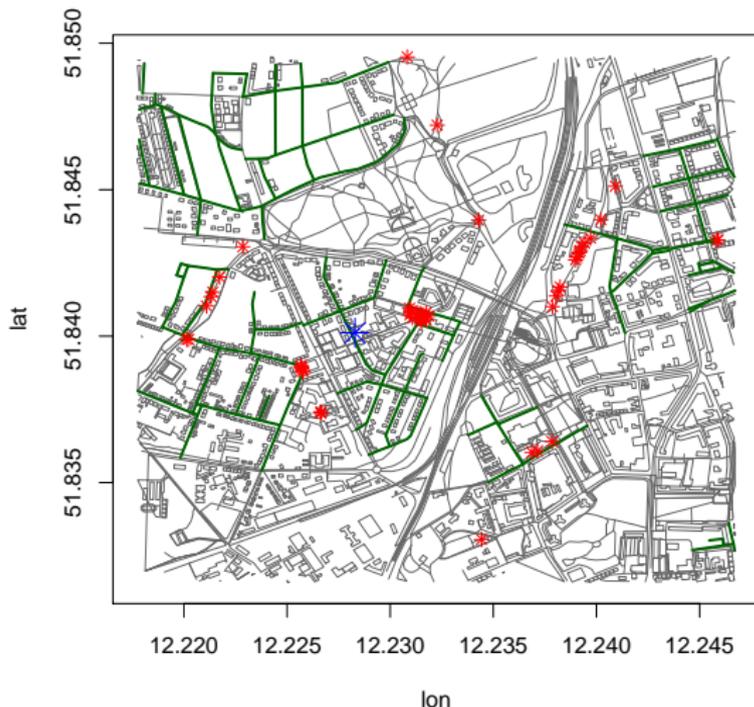
- Finden der ID's per `find(...)`
 - > bench_ids<-find(dessau, node(tags(k=="amenity" & v=="bench")))
 - > bench_ids[1:5]
[1] 417441492 417441496 417441502 417441506 417441520
- Teilmenge von dessau mittels `subset(...)` erstellen
 - > bench_dessau<- subset(dessau, node_ids=bench_ids)
 - > bench_dessau
osmar object
51 nodes, 0 ways, 0 relations

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen**
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - **plot**
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

Graphikerstellung per plot(...)

benches in dessau



- Straßen mit maxspeed=30
- * bench
- * Standort

Darstellung von Rohdaten ohne Transformation

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - plot
- 4 osmar: Anwendungsbeispiel**
- 5 Zusammenfassung

Datensatzerweiterung

```
> bench_coords[1:5,]
```

	id	lon	lat
1	417441492	12.22286	51.84305
2	417441496	12.22174	51.84202
3	417441502	12.22110	51.84103
4	417441506	12.22132	51.84135
5	417441520	12.22014	51.83989

- Ausgangslage:
 - Koordinaten
 - evtl weitere Daten
- Erweiterung mit Straßendaten mithilfe von `find` & `subset`

Vorgehen:

- 50x50 Meter Bounding Box um jeden Punkt herunterladen
- Zählen der Straßenarten und Gebäude
- Hinzufügen zum `data.frame`

```
> bench_data[1:5,]
```

	id	lon	lat	footway	secondary
1	417441492	12.22286	51.84305	2	0
2	417441496	12.22174	51.84202	1	1
3	417441502	12.22110	51.84103	1	0
4	417441506	12.22132	51.84135	1	1
5	417441520	12.22014	51.83989	2	3
	primary	residential	building		
1	0	0	0		
2	0	0	0		
3	0	1	1		
4	0	1	0		
5	0	3	0		

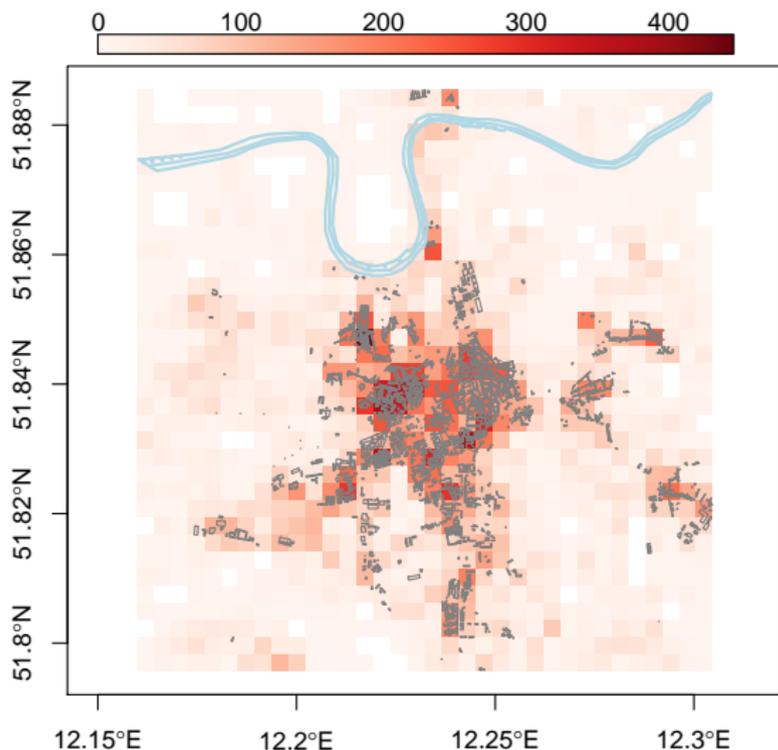
Vorgehen:

- 50x50 Meter Bounding Box um jeden Punkt herunterladen
- Zählen der Straßenarten und Gebäude
- Hinzufügen zum `data.frame`

```
> bench_data[1:5,]
```

	id	lon	lat	footway	secondary	primary	residential	building
1	417441492	12.22286	51.84305	2	0	0	0	0
2	417441496	12.22174	51.84202	1	1	0	0	0
3	417441502	12.22110	51.84103	1	0	0	1	1
4	417441506	12.22132	51.84135	1	1	0	1	0
5	417441520	12.22014	51.83989	2	3	0	3	0

Anzahl der Nodes in Dessau



weitere Pakete & Ausblick

- igraph** Konvertierung in Objekte des `igraph`-Paketes ermöglicht „Benutzung“ als Navigation
- sp** Konvertierung erlaubt Arbeit mit ca. 100 weiteren Paketen
- sp** Möglichkeit der räumlichen Datenanalyse (z.B. räumliche Regression, Kriging, ...)

Ausblick Ziel ist nun das Hochladen von Daten auf den OpenStreetMap-Server zu realisieren.

weitere Pakete & Ausblick

- igraph** Konvertierung in Objekte des `igraph`-Paketes ermöglicht „Benutzung“ als Navigation
- sp** Konvertierung erlaubt Arbeit mit ca. 100 weiteren Paketen
- sp** Möglichkeit der räumlichen Datenanalyse (z.B. räumliche Regression, Kriging, ...)

Ausblick Ziel ist nun das Hochladen von Daten auf den OpenStreetMap-Server zu realisieren.

Gliederung

- 1 Einführung in R
- 2 osmar: Einführung
- 3 osmar: Funktionen
 - Datenimport und osmar-Objekt
 - summary, find und subset
 - plot
- 4 osmar: Anwendungsbeispiel
- 5 Zusammenfassung

Zusammenfassung

- R ist eine Open Source Software für die statistische Datenanalyse
- `osmar` erweitert R um die Datenbanken der OpenStreetMap-Community und ermöglicht statistische Analysen
- Durch das Open-Source-Konzept können weitere User mit dem Paket weiterarbeiten und es erweitern
- weitere Informationen auf <http://osmar.r-forge.r-project.org/>

Quellen

R <http://www.r-project.org/>

sp [http:](http://cran.r-project.org/web/packages/sp/index.html)

[//cran.r-project.org/web/packages/sp/index.html](http://cran.r-project.org/web/packages/sp/index.html)

igraph <http://cran.r-project.org/web/packages/igraph/index.html>

MiKTeX <http://miktex.org/>