

```
library(dplyr)  
  
rladies_global %>%  
  filter(city == 'Munich')
```



#2

Making fRiends with: ggplot2



Hello!

I am Pam

Today I would like to share some tips with you on
how to create *fun* plots using **ggplot2**



Smoothly sliding into ggplot2

1. First
steps

2. Base
Graphics vs
ggplot2

3. ggplot2 &
Grammar of
Graphics

4. (or hands-on
stuff)

Pro tip: Do stop and ask me at any time!



1.

First steps

Take a visit to the GitHub repo

https://github.com/pamelamatias/RLadies_MUC

and either clone or download the files we'll be using ☺



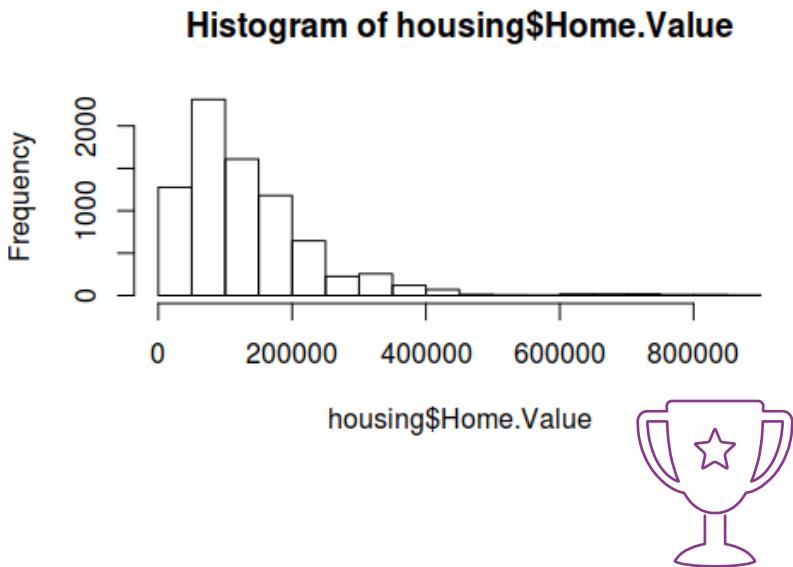
2.

Base Graphics vs ggplot2

or why should I try something new?

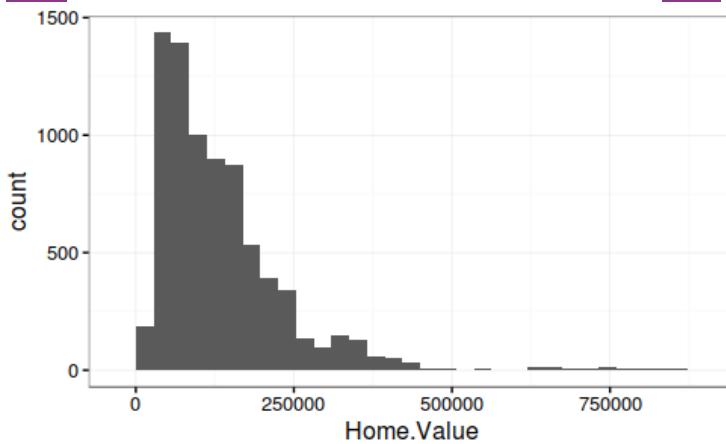
Base graphics

```
hist(housing$Home.Value)
```



ggplot2

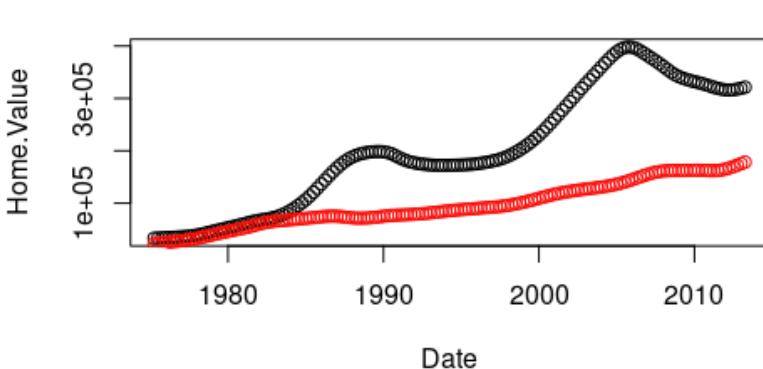
```
library(ggplot2)  
ggplot(housing, aes(x = Home.Value)) +  
  geom_histogram()
```



more verbose for simple plots
):

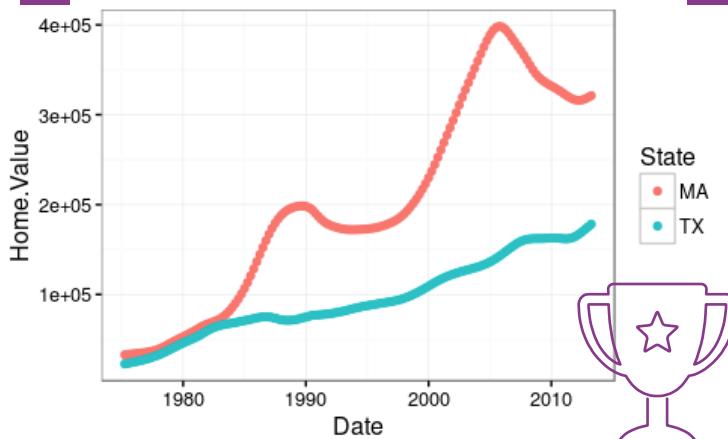
Base graphics

```
plot(Home.Value ~ Date,
      data=subset(housing, State == "MA"))
points(Home.Value ~ Date, col="red",
      data=subset(housing, State == "TX"))
legend(1975, 400000,
      c("MA", "TX"), title="State",
      col=c("black", "red"),
      pch=c(1, 1))
```



ggplot2

```
ggplot(subset(housing, State %in% c("MA", "TX")),
       aes(x=Date,
           y=Home.Value,
           color=State))+
  geom_point()
```



Less verbose for complex graphs!



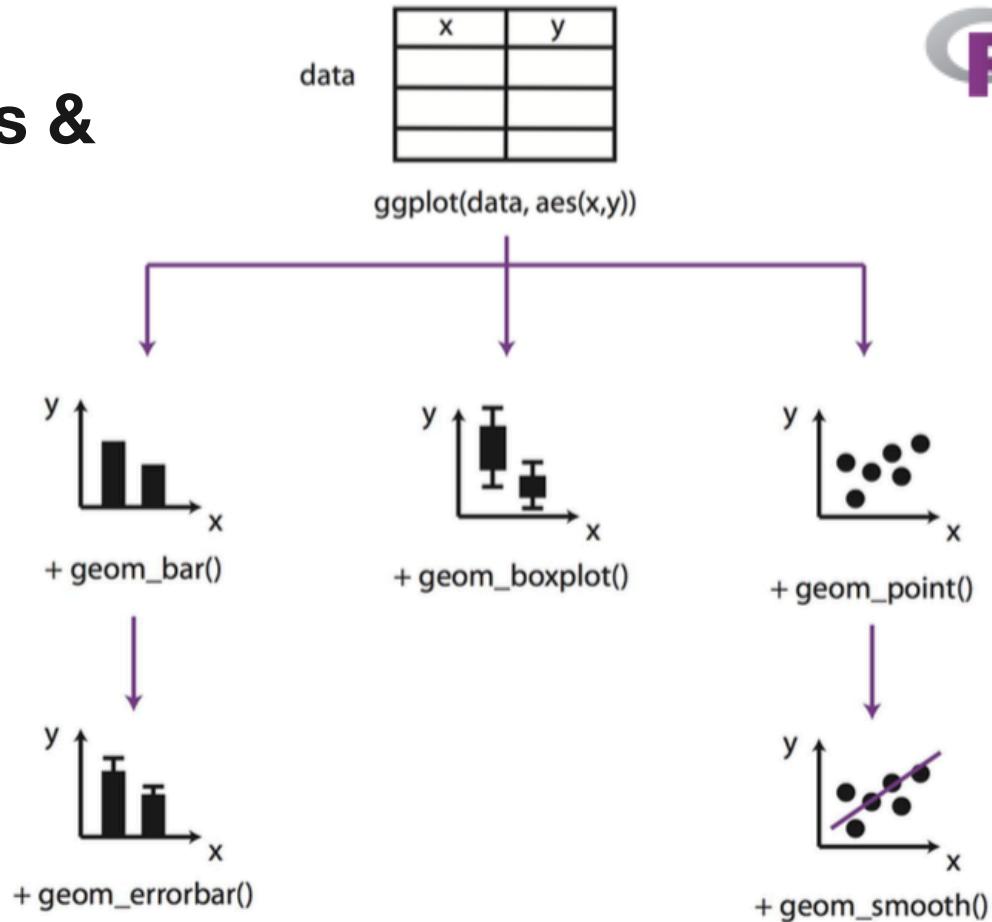
3.

Grammar of Graphics & ggplot2

or “how to make sense out of all of those layers”

Grammar of graphics & ggplot2

- + geom_xy() - geometric objects
- + stat_xy() - statistics
- + scales
- + coord_xy() - coordinate systems
- + facet_xy() - faceting
- + position_xy() - Position adjustments
- + theme_() - Themes (non-data)



```
rladies_global %>%  
  diagram.source(city == 'Berlin')
```



“ *The emphasis in ggplot2 is reducing the amount of thinking time by making it easier to go from the plot in your brain to the plot on the page”*

—Hadley Wickham, 2012

(as cited in Christine Zhang's “Storytelling with Data” Workshop at Boston University)

4.



(or hands-on stuff!)

Dataset

Inside Airbnb

Inside Airbnb
Adding data to the debate

INDEPENDENT, NON-COMMERCIAL,
OPEN SOURCE DATA TOOL

How is Airbnb really
being used in and affecting
your neighborhood?

FILTER by Neighborhood

The data
Airbnb
doesn't want
you to see!

50+
data points
per listing

OUT OF MORE THAN
27,000 LISTINGS:
16K are for the
entire home (58%)
87% highly available
(more than 60 days/year)
29% multi-listings
(where the host has other listings)

Airbnb IN NYC

SEE Airbnb ACTIVITY OVER TIME IN YOUR NEIGHBORHOOD

2012 2013 2014

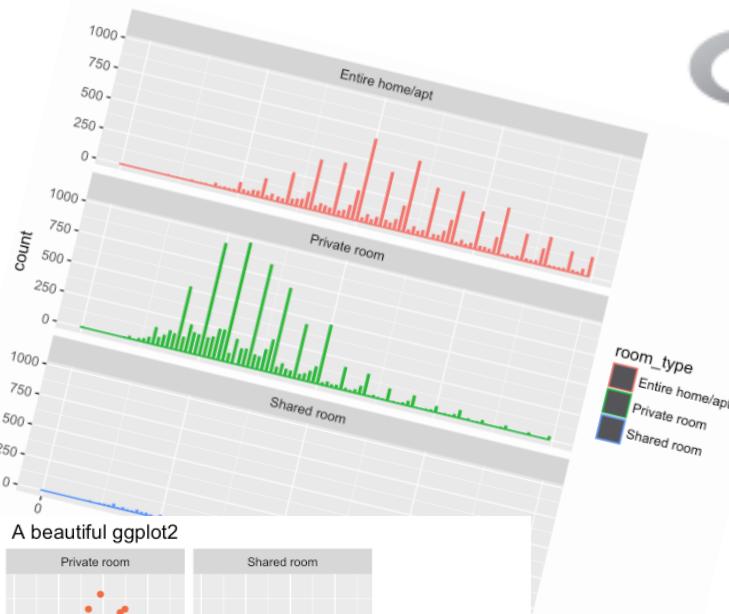
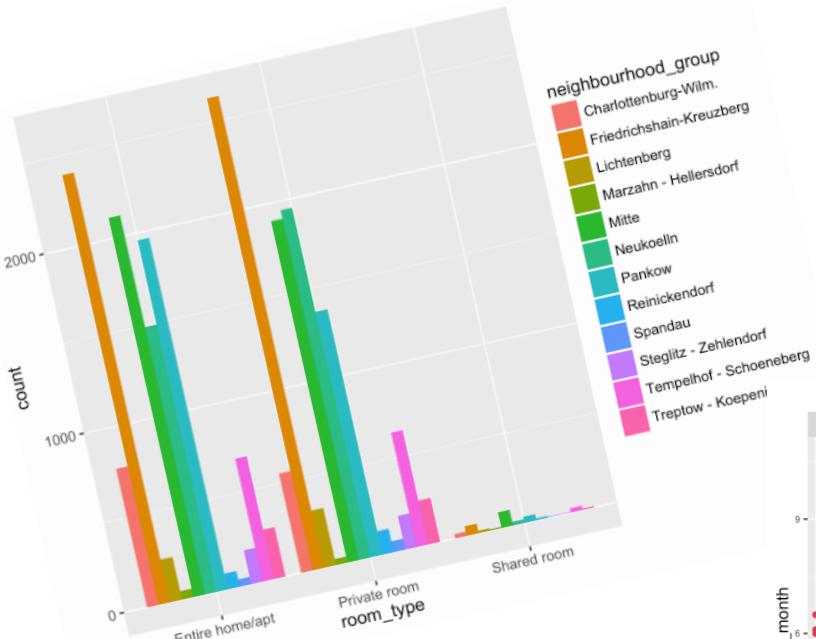
HOST "JOHN D"
17 listings

VIEW TOP HOSTS'
MULTIPLE LISTINGS

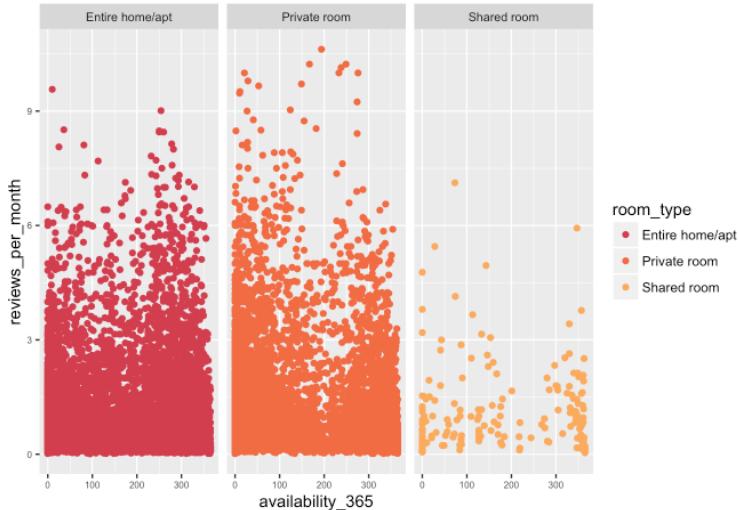
NEXT...

• VISIT insideairbnb.com
• SHARE it widely
 f t #insideairbnb #illegalhotels
#affordablehousing #nyc
• DOWNLOAD the data
(open source; 50+ data points per listing)

Some cool ggplot2s



A beautiful ggplot2





We're Daloha, Dan, Maggie & Pam

We recently joined forces to make **RLadies Munich** happen, but we'd be happy to have you join us on [Meetup](#), [Instagram](#), [Twitter](#), [Facebook](#) and [Slack](#)



Thanks!



**Now let's go for
cookies & beer!**

Not necessarily at the
same time though...

Let's say then
cookies | beer!

