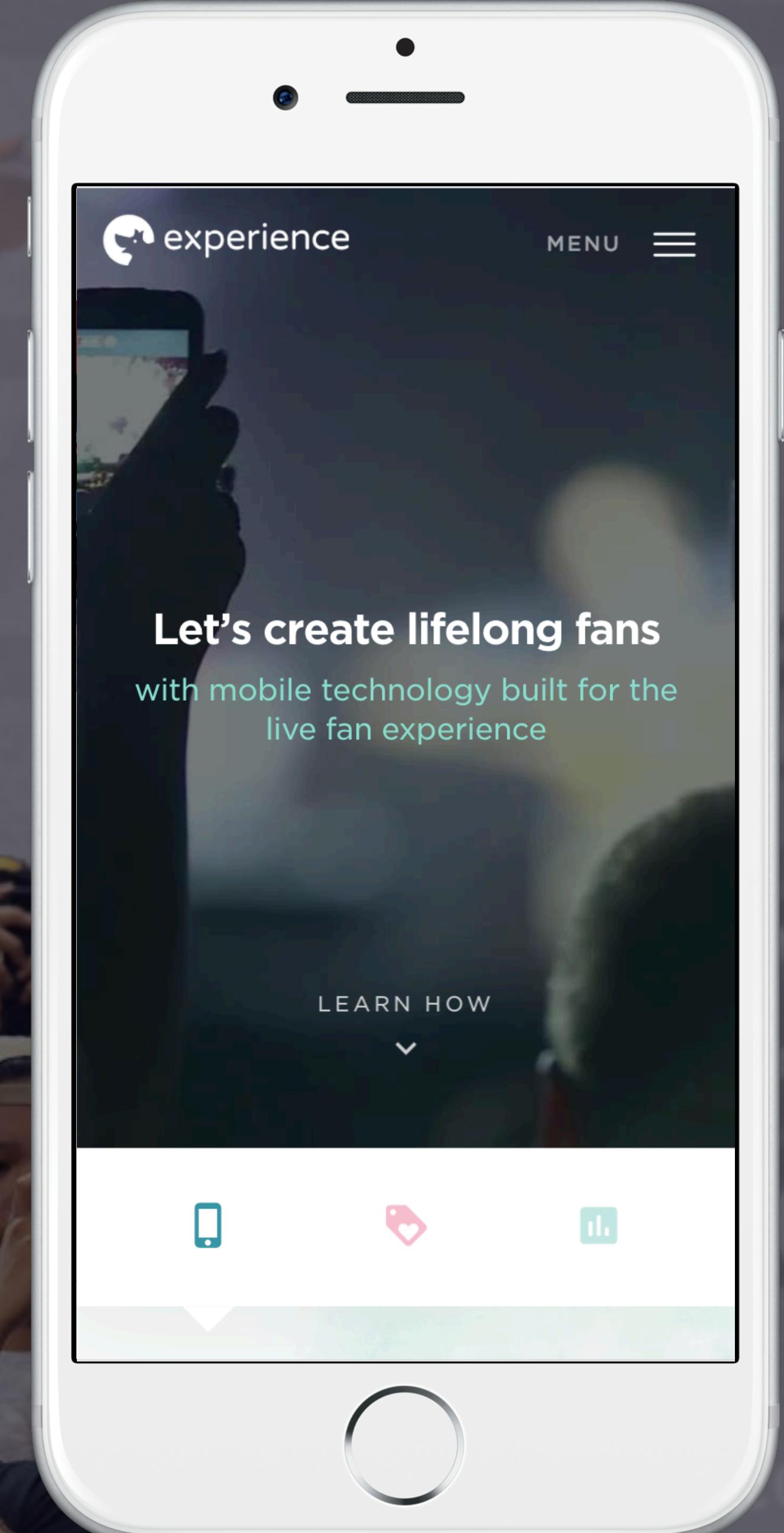


Welcome!

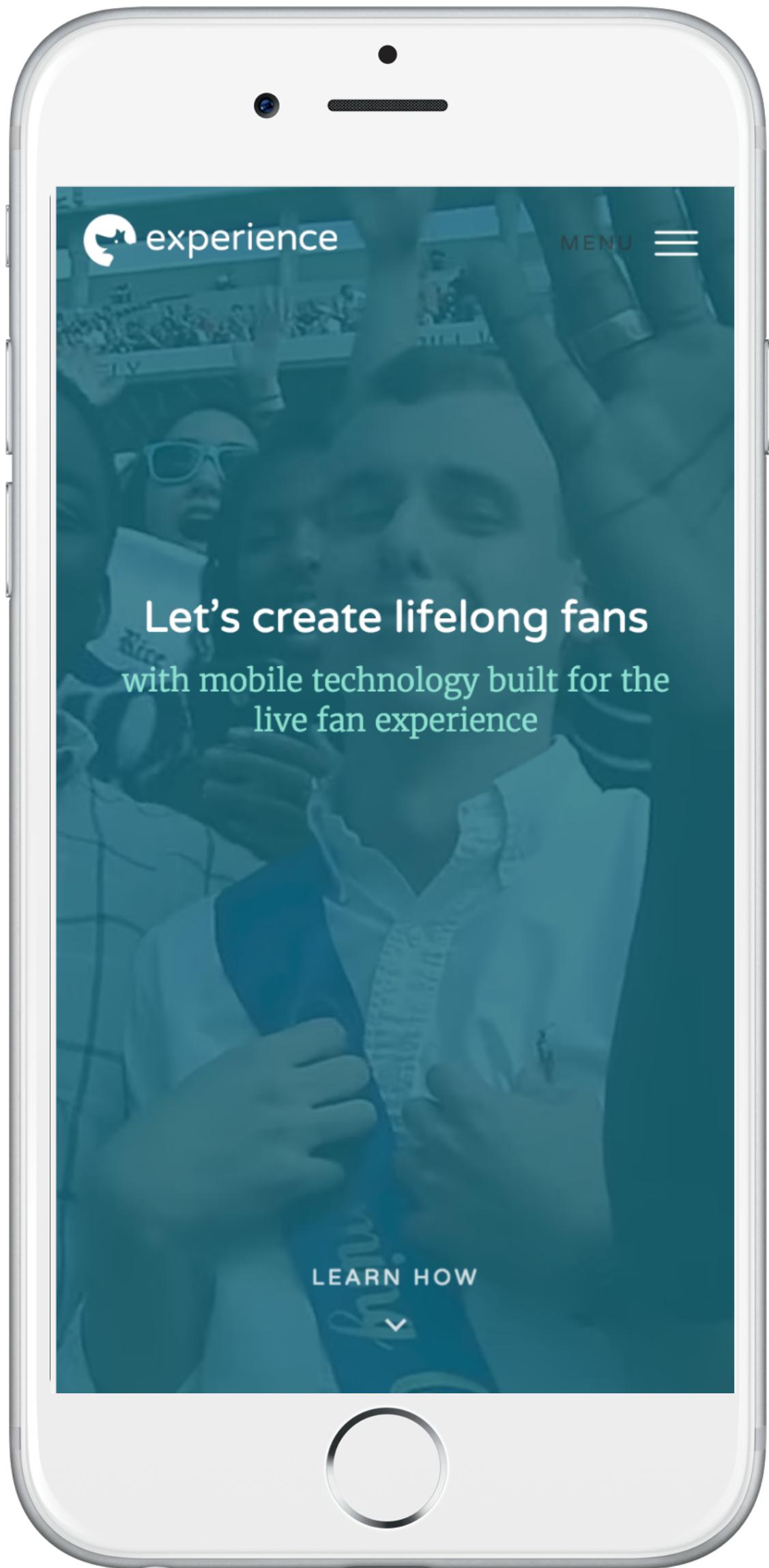
**Wifi:
Network: Experience-Guest
password: domdeluise**



Who am I?

- **Matt Mills**
- **Born and raised in Atlanta**
- **BS in Industrial and Systems Engineering 2014, MS in Analytics 2015**
- **@statmills or www.statmills.com**

What is Experience?



Experience's mobile commerce, ticketing, and data solutions empower sports and entertainment leaders to generate new revenue streams, sell more tickets, and make smarter decisions.

www.expapp.com/solutions

What is Experience?



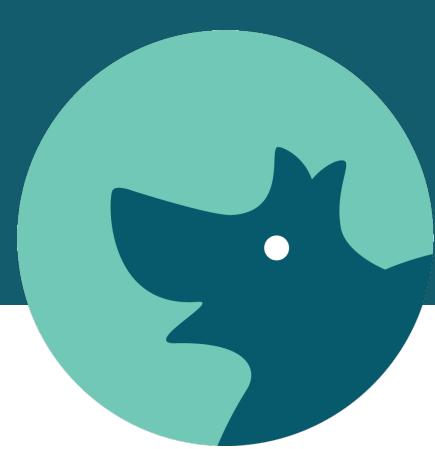
The smartphone screen shows a menu titled "Upgrades" with a back arrow. It lists three experience options:

- Sit Behind Home Plate - \$29 (with an image of a baseball field)
- Meet and Greet Braves Alumni - \$19 (with an image of people at a meet-and-greet)
- Get on the Field for Fireworks - \$19 (with an image of a baseball stadium at night)

The phone has a white case and is shown against a background of five cards representing different experience categories.

- Seat Upgrades**
Find new revenue streams with unsold seat inventory
- VIP Experiences**
Deliver the experiences that make each live event incredible
- Merchandise**
Let your fans take home the action
- Loyalty**
Reward the behavior that matters most - season ticket loyalty

What is Experience?



Single-game tickets

Maximize single-event ticket sales

Subscription access

Deliver “the Netflix” of attending live events

Flexible memberships

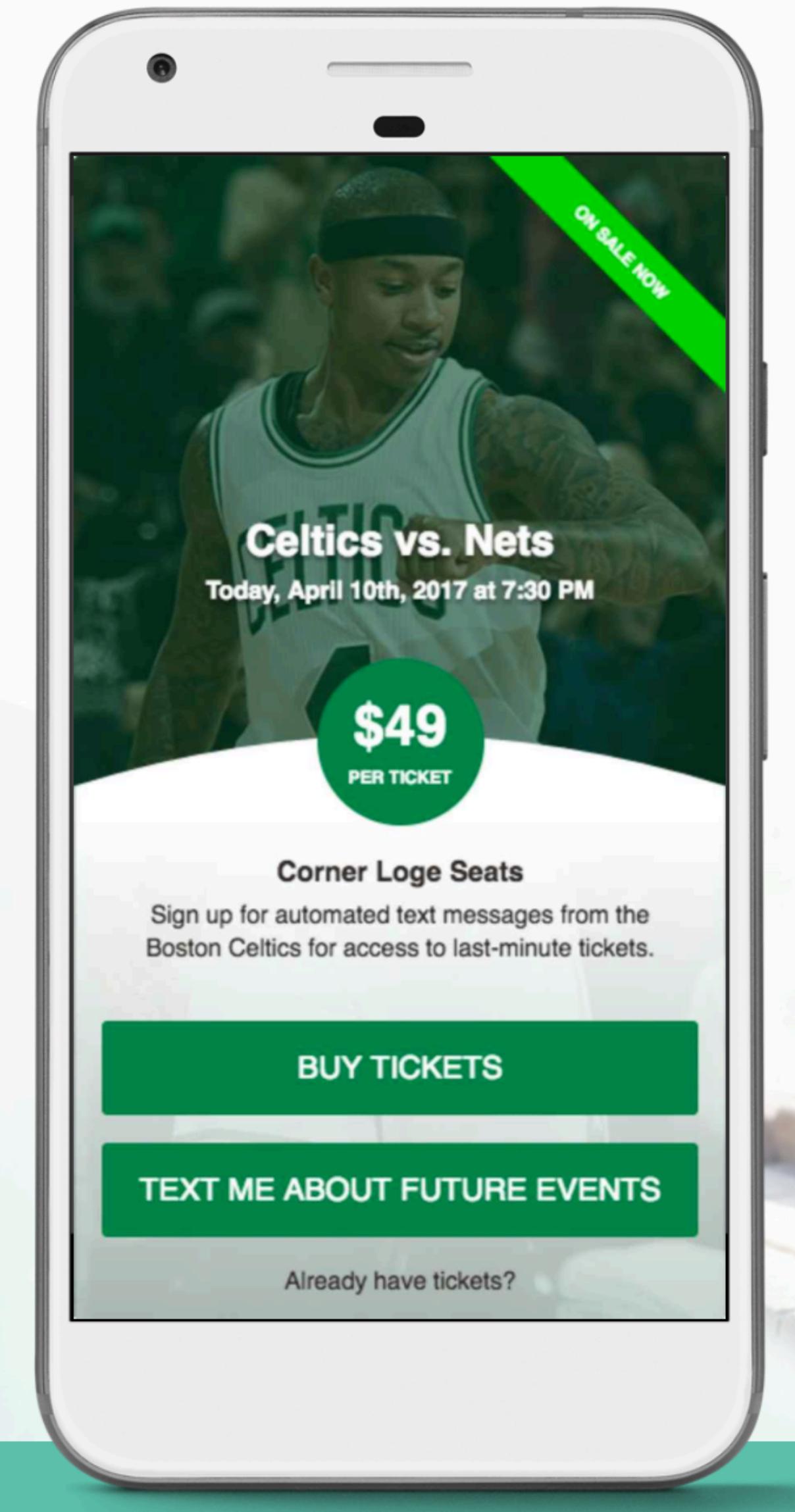
Give avid fans unprecedented flexibility

Fluid Ticket™ technology

Unlock the full value of a mobile season ticket

Student ticketing

Make ticketing easy for students



Atlanta R Users Relevant Links



<https://twitter.com/AtlantaRUsers>



<https://github.com/AtlantaRUsers/Meetups>



<https://www.meetup.com/R-Users-Atlanta/>

Atlanta R Users Group Goals

- **Provide resources to help with learning various aspects of R**
- **Monthly meetups for people in the community to present and share**
- **Annual Atlanta R Conference**

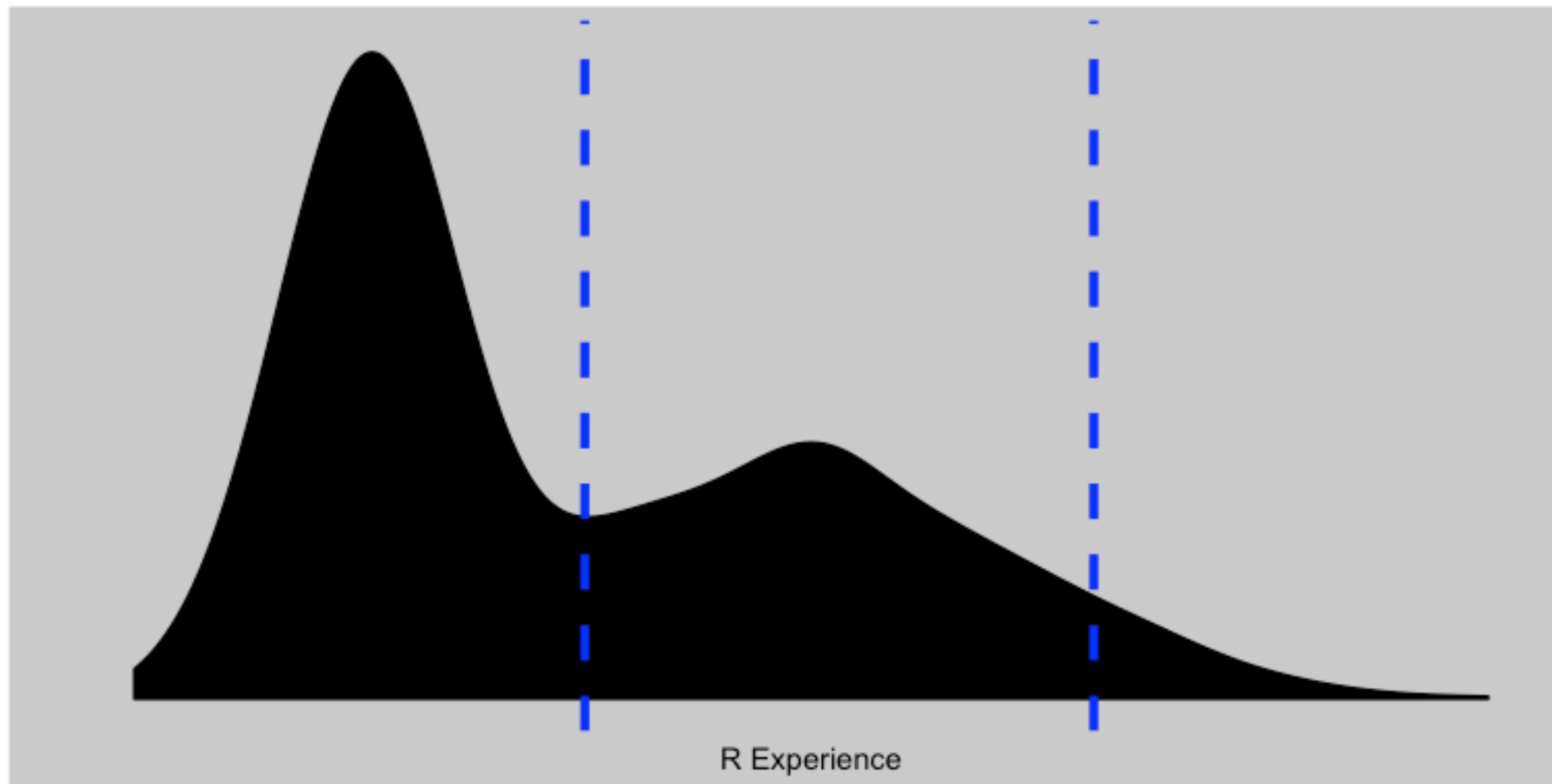
Ways to Help Without Volunteering Time

- We need meeting space
 - Many thoughts on best locations for meetups but we are dependent on space
- Meetup sponsors can help as well!
- Feedback!
- Patience :)

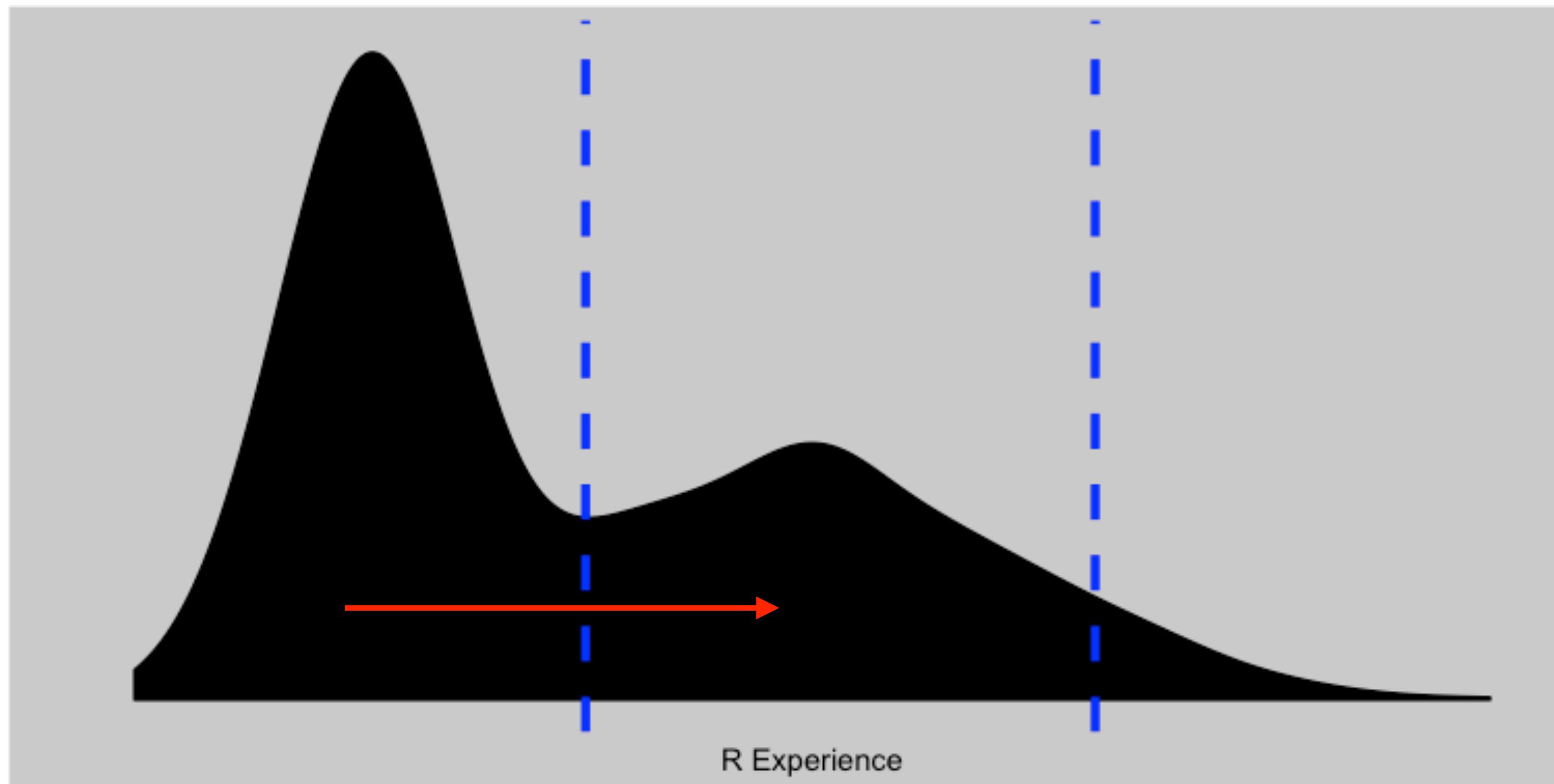
Next Steps

- Need spaces for September/October
- Share the word with people who may be interested
- Call for speakers!

Tonight's Meetup

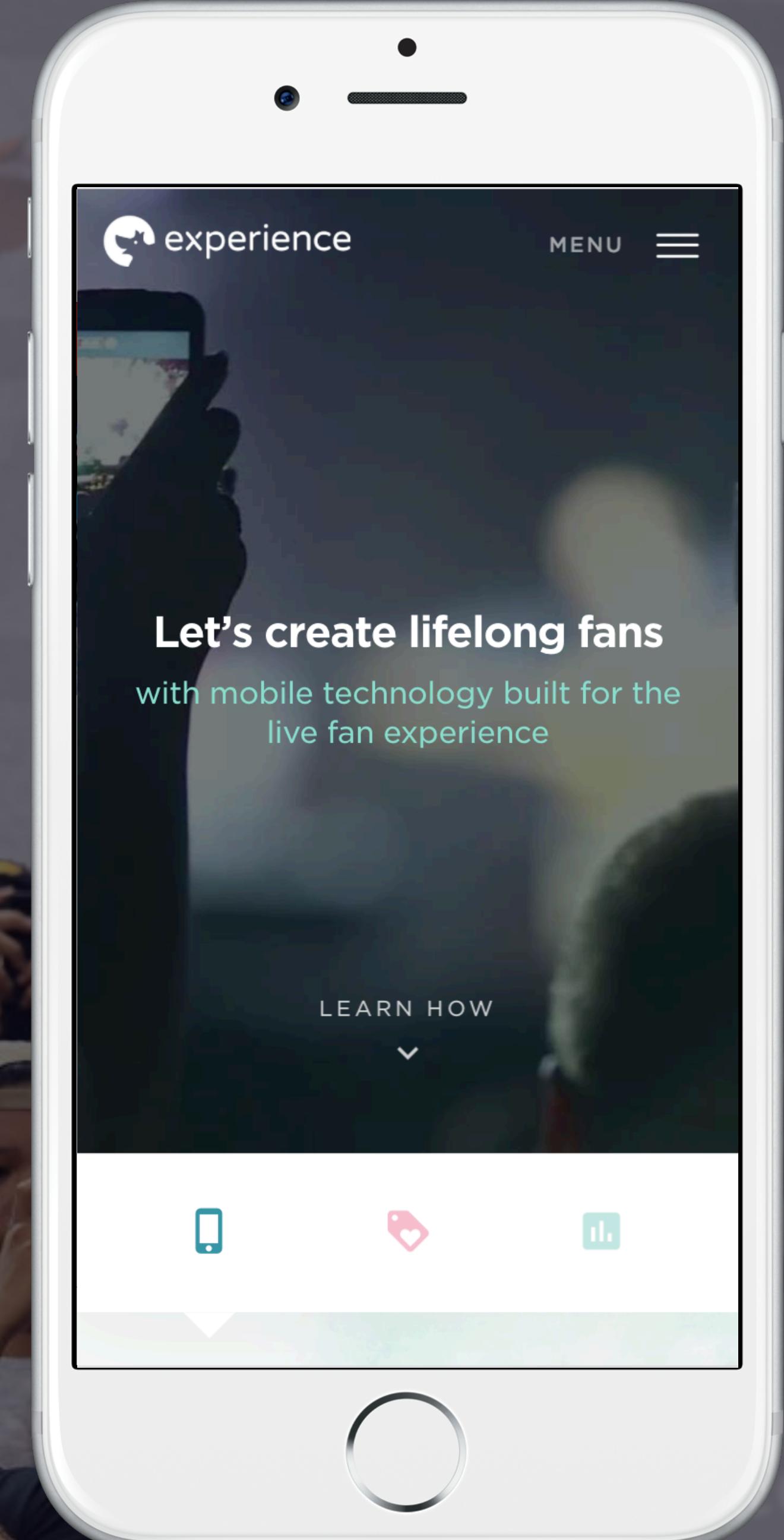


Tonight's Meetup



Replacing Excel with R

Matt Mills - August, 28th



Caveats

- *I don't actually work in Excel, and only do basic work in google sheets when I have to*
- *30 minutes is not nearly enough time to teach you everything*

Agenda

- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- Graphics
- Reporting
- Resources for learning R

Why Switch from Excel to R

- Work in R is
 - ***Reproducible***
 - ***Scaleable***
 - ***Faster***
 - ***More Robust***

What I think you already know how to do

- **Download RStudio**
- **Open R scripts**
- **Run Commands in the Console**
- **Install Packages**

Agenda

- **Differences between R and Excel**
- Reading in Data
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- Graphics
- Reporting
- Resources for learning R

Differences between R and Excel (to me)

- **R is a scripted programming language**

```
97  
98 team_name <- "BOS"  
99  
100 team_data <- readRDS(sprintf("data/team_sales/%s_sales.RDS", team_name))  
101  
102 seat_sales <- transform_sale_records(team_data)  
103  
104 account_info <- generate_account_metrics(seat_sales)  
105  
106 seat_manifest <- create_seat_manifest(seat_sales)  
107  
108 seat_info <- generate_seat_metrics(seat_sales, seat_manifest)  
109  
110 season_info <- combine_seat_and_accounts(seat_sales, seat_manifest, account_info, seat_info)  
111
```

- **Excel is a physical language**

Differences between R and Excel: Data

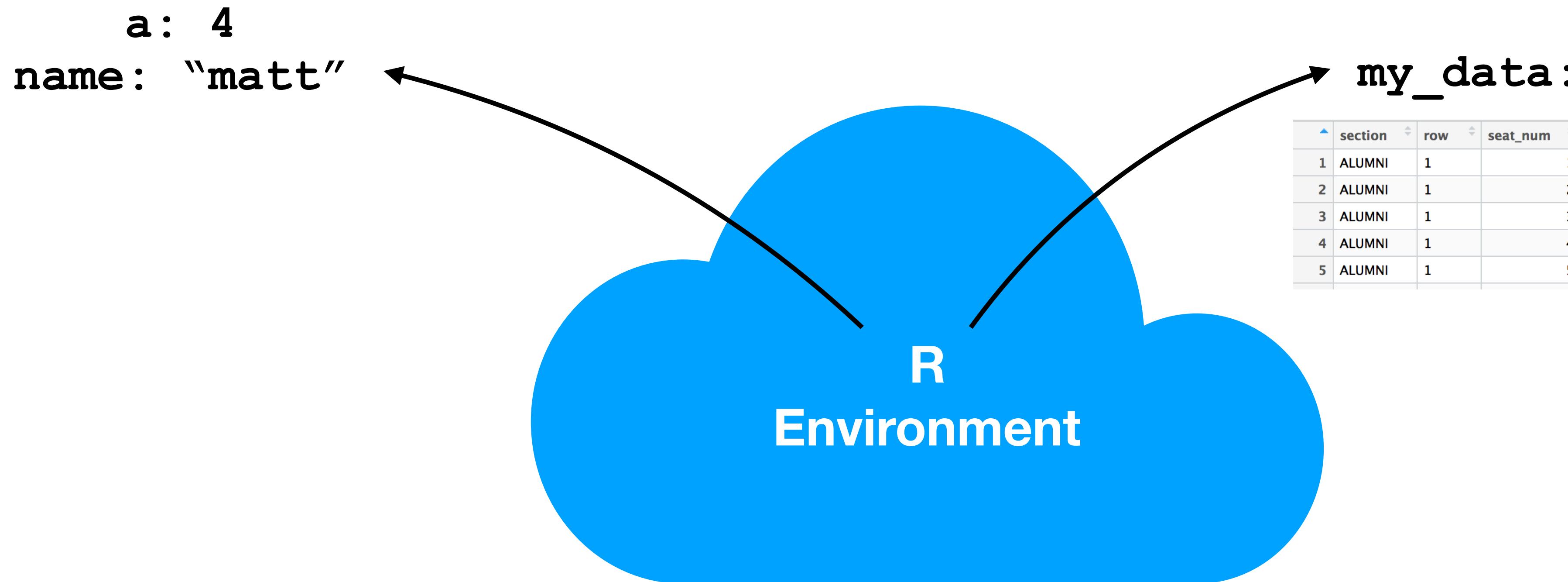
- In Excel *Data* is organized in grids
 - refer to it by it's physical location

	A	B	C
1	Name	3 PA's	24 ×
2	Steph Curry	=B2*3	8

Differences between R and Excel: Data

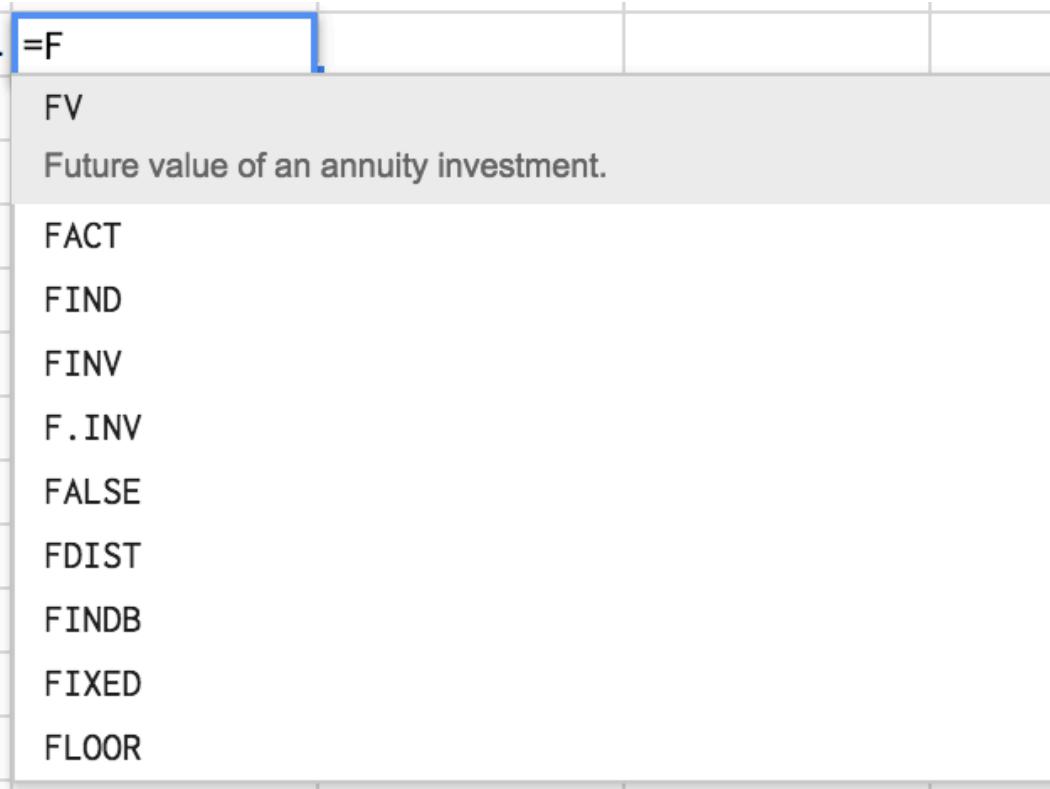
- In Excel **Data** is organized in grids
 - refer to it by it's physical location
- In R **Data is found in your environment**

	A	B	C
1	Name	3 PA's	24
2	Steph Curry		=B2*3



Differences between R and Excel: Manipulation

- In Excel in order to “do” something you use formulas



- In R you use functions

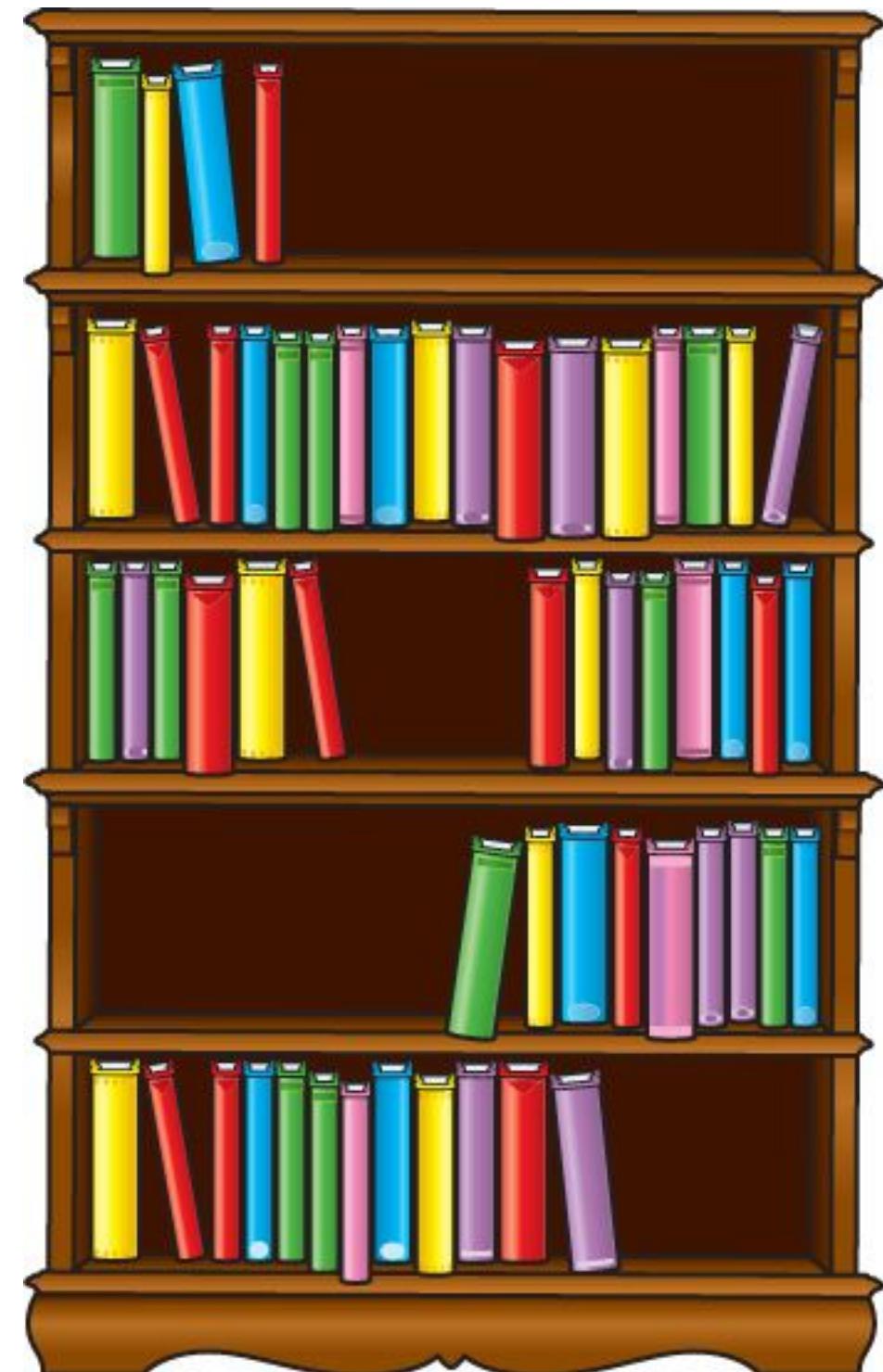
```
> table(sample(c(1, 2, 3), size = 10, replace = T))  
  
1 2 3  
5 3 2
```

Agenda

- Differences between R and Excel
- **Reading in Data**
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- Graphics
- Reporting
- Resources for learning R

Reading in (tabular) Data

- In Excel you open the file...?
- In R you use **packages**
 - packages are collections of functions that share a similar purpose



Reading in (tabular) Data: Packages



- **readxl**
 - can read in specific sheets or cell ranges
 - <https://readxl.tidyverse.org/>
- **readr/data.table**
 - files must be csv, tsv, or txt extensions, not excel.
 - <https://readr.tidyverse.org/> | <https://github.com/Rdatatable/data.table/wiki/Convenience-features-of-fread>
- **googlesheets**
 - Can download files straight from google sheets
 - <https://github.com/jennybc/googlesheets>

Agenda

- Differences between R and Excel
- Reading in Data
- **Basic Summary Statistics**
- Merging Data
- Pivot Tables
- Graphics
- Reporting
- Resources for learning R



Summary Data

- In Excel you use formulas such as **COUNT**, **SUM**, **AVERAGE**
 - Get conditional values with if versions (**COUNTIF**, etc...)
- In R we use basic functions like **sum**, **mean**, **min**, **max** that take in a vector

Summary Data

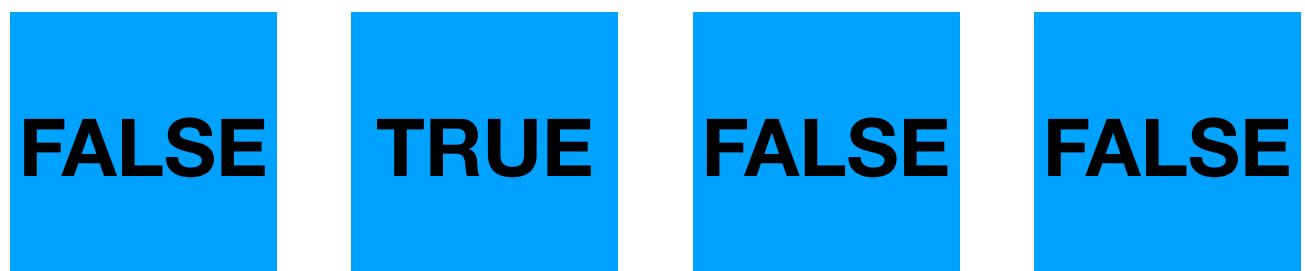
- In Excel you use formulas such as **COUNT**, **SUM**, **AVERAGE**
 - Get conditional values with if versions (**COUNTIF**, etc...)
- In R we use basic functions like **sum**, **mean**, **min**, **max** that take in a vector
 - Get conditional values using logical indexing

Logical Indexing

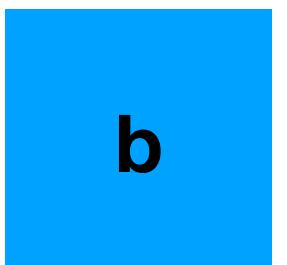
```
> my_vec <- c("a", "b", "c", "d")
> my_vec
[1] "a" "b" "c" "d"
```



my_vec == "b"



my_vec[my_vec == "b"]



Logical Indexing: Continued

Excel

```
SUMIF(range, criterion, sum_range)
```

```
=SUMIF(C:C, "=2007", E:E)
```

R

```
sum(sum_range[range == criterion])
```

```
sum(gapminder$pop[gapminder$year == 2007])
```

Summarizing Data Extended: The Data Frame

- **Summarizing and Manipulating data is easiest* when using Data Frames**
- **Data Frames are**
 - **SQL-like tables**
 - **Vectors of the same length organized into columns**

```
> str(data.frame(gapminder))
'data.frame': 1704 obs. of 6 variables:
 $ country : Factor w/ 142 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 ...
 $ continent: Factor w/ 5 levels "Africa","Americas",...: 3 3 3 3 3 3 3 3 3 ...
 $ year     : int 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...
 $ lifeExp   : num 28.8 30.3 32 34 36.1 ...
 $ pop       : int 8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957
16317921 22227415 ...
 $ gdpPercap: num 779 821 853 836 740 ...
```

Summarizing Data Extended: dplyr / data.table

- **dplyr**
 - provides a grammar for data manipulation
 - SQL-like verbs to manipulate data frames
 - <https://dplyr.tidyverse.org/>
- **data.table**
 - how to *subset, select and compute* and perform aggregations
 - <http://r-database.com>

```
gapminder %>%
  filter(year == 2007) %>%
  group_by(continent) %>%
  summarize(max_gdp = max(gdpPerCap)) %>%
  arrange(desc(max_gdp))
```

Agenda

- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- **Merging Data**
- Pivot Tables
- Graphics
- Reporting
- Resources for learning R

Joining Data in Excel

- In Excel **VLOOKUP**, **INDEX**, and **MATCH** are king*

VLOOKUP(“3”, A:B, 2)

MATCH ("3", A:A)

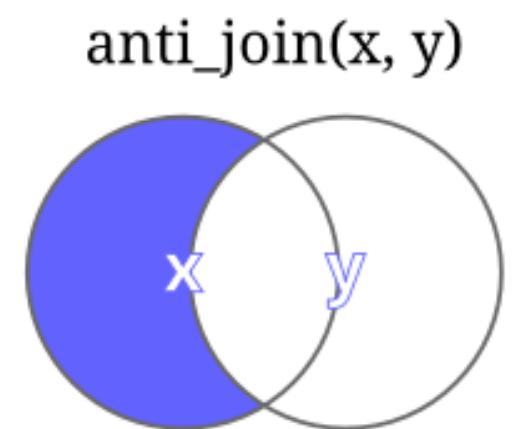
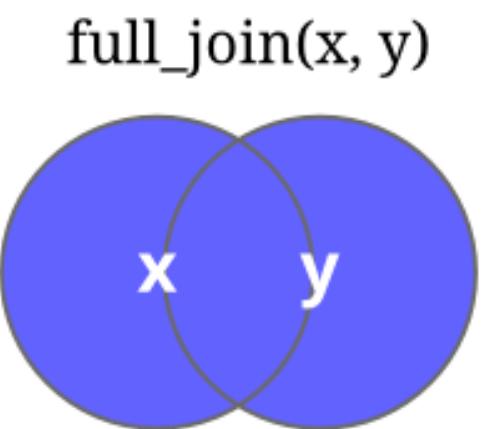
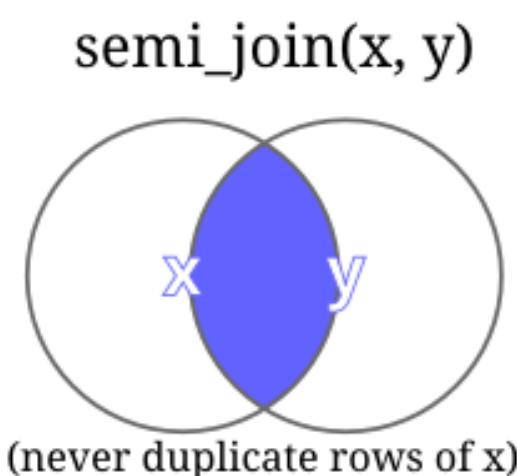
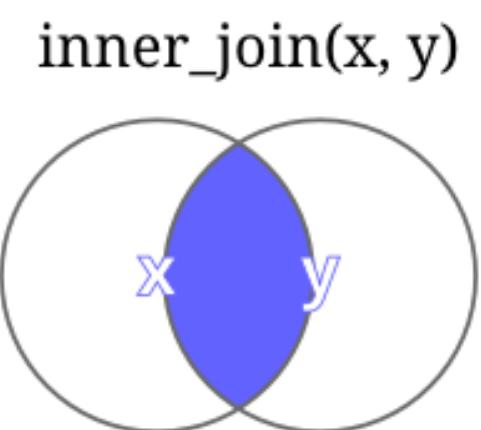
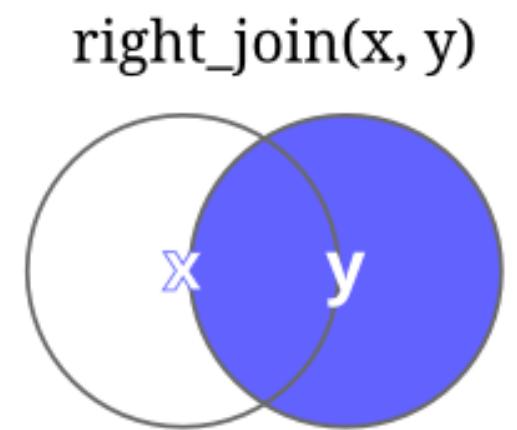
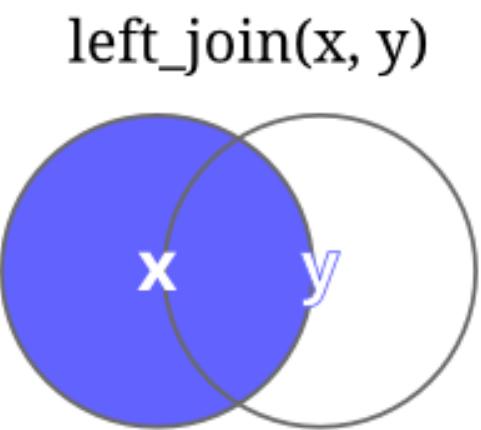
A graph illustrating a piecewise function $f(x)$ plotted against x . The function is zero for $x \leq 1$. For $x > 1$, the function consists of two parts: a horizontal line segment at $y = 1$ and a line segment starting at $(1, 2)$ with a positive slope. The point $(1, 2)$ is marked with a yellow dot.

INDEX(A:B, 3, 2)

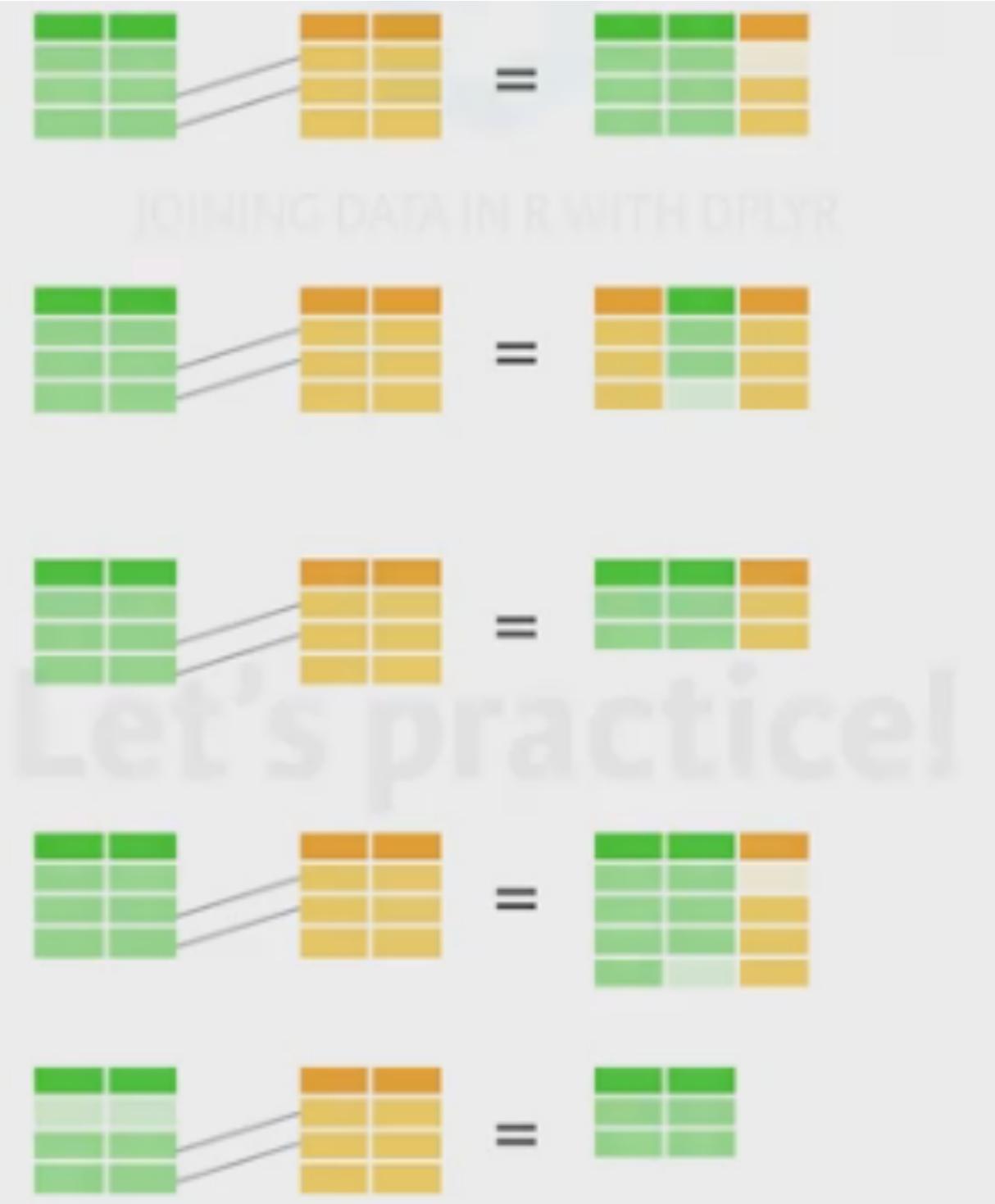
Joining Data in R

- In R we.... well we *join* data frames

dplyr *joins*



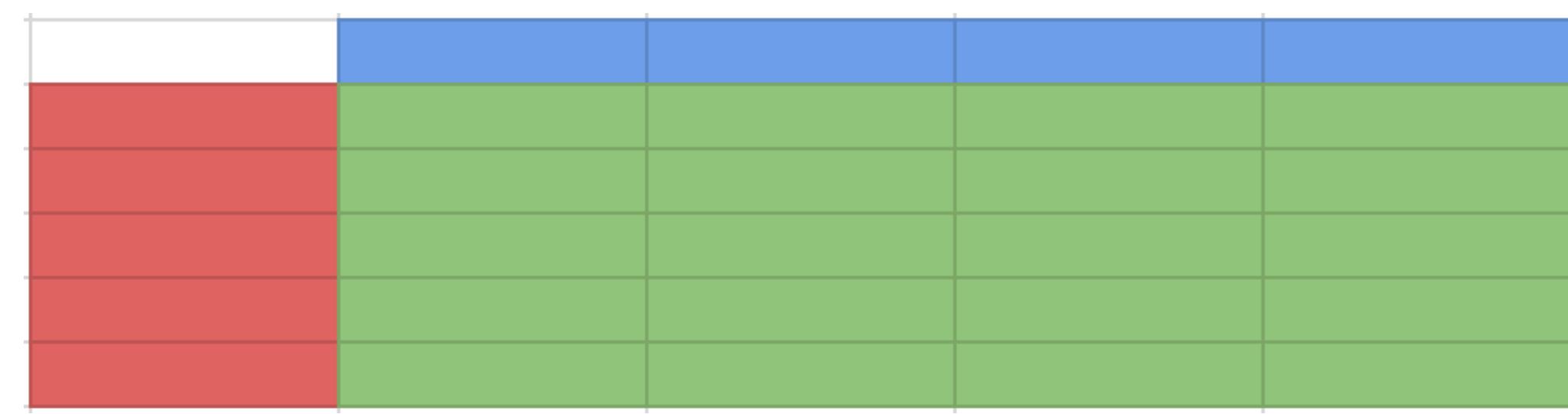
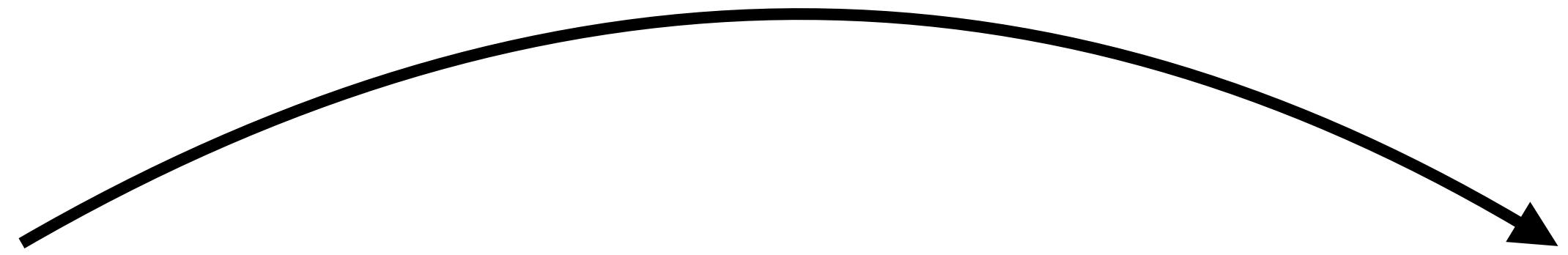
- `left_join()`
- `right_join()`
- `inner_join()`
- `full_join()`
- `semi_join()`



Agenda

- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- Merging Data
- **Pivot Tables**
- Graphics
- Reporting
- Resources for learning R

Pivot Tables

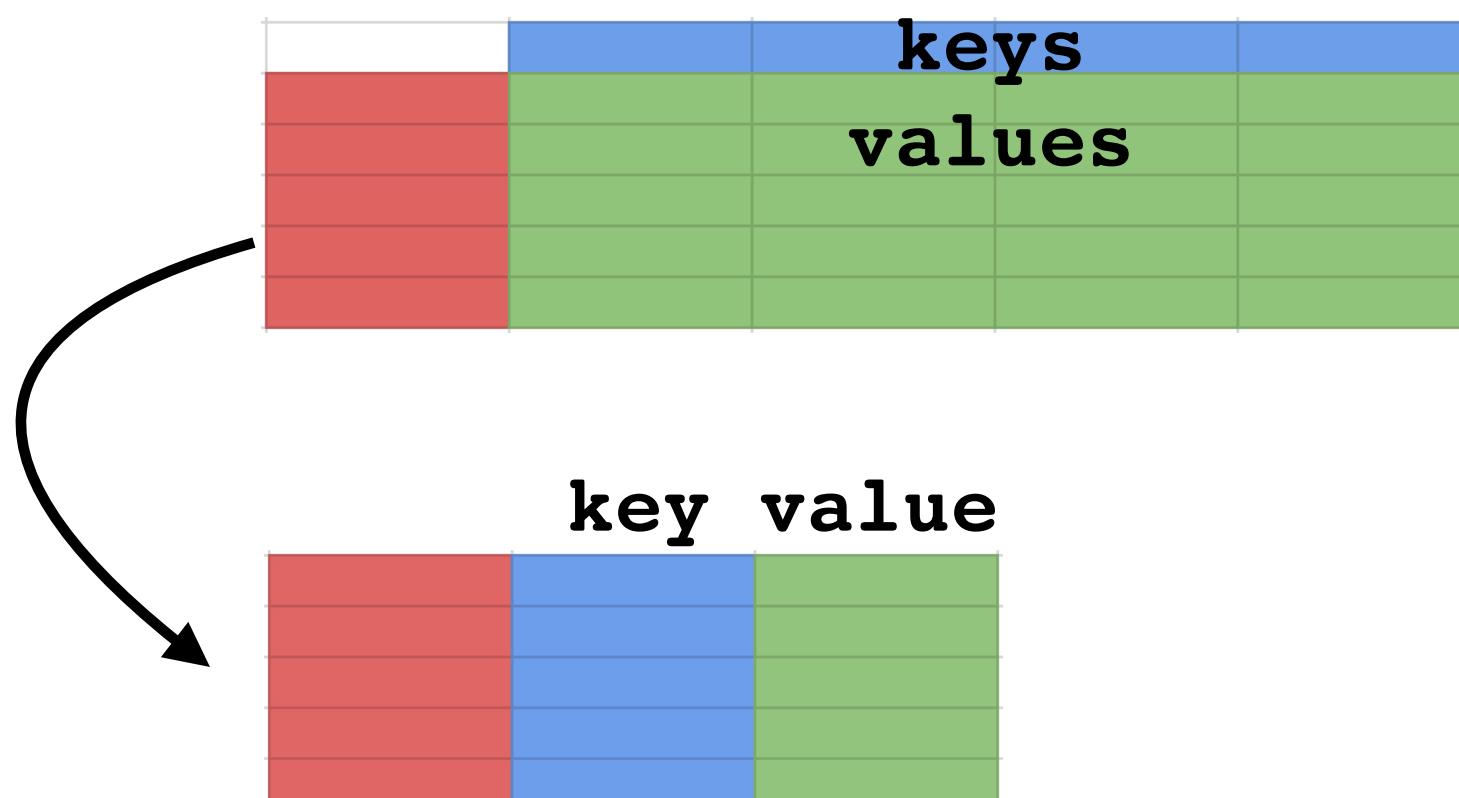


<https://twitter.com/OmniAnalytics/status/1034538215527600128>

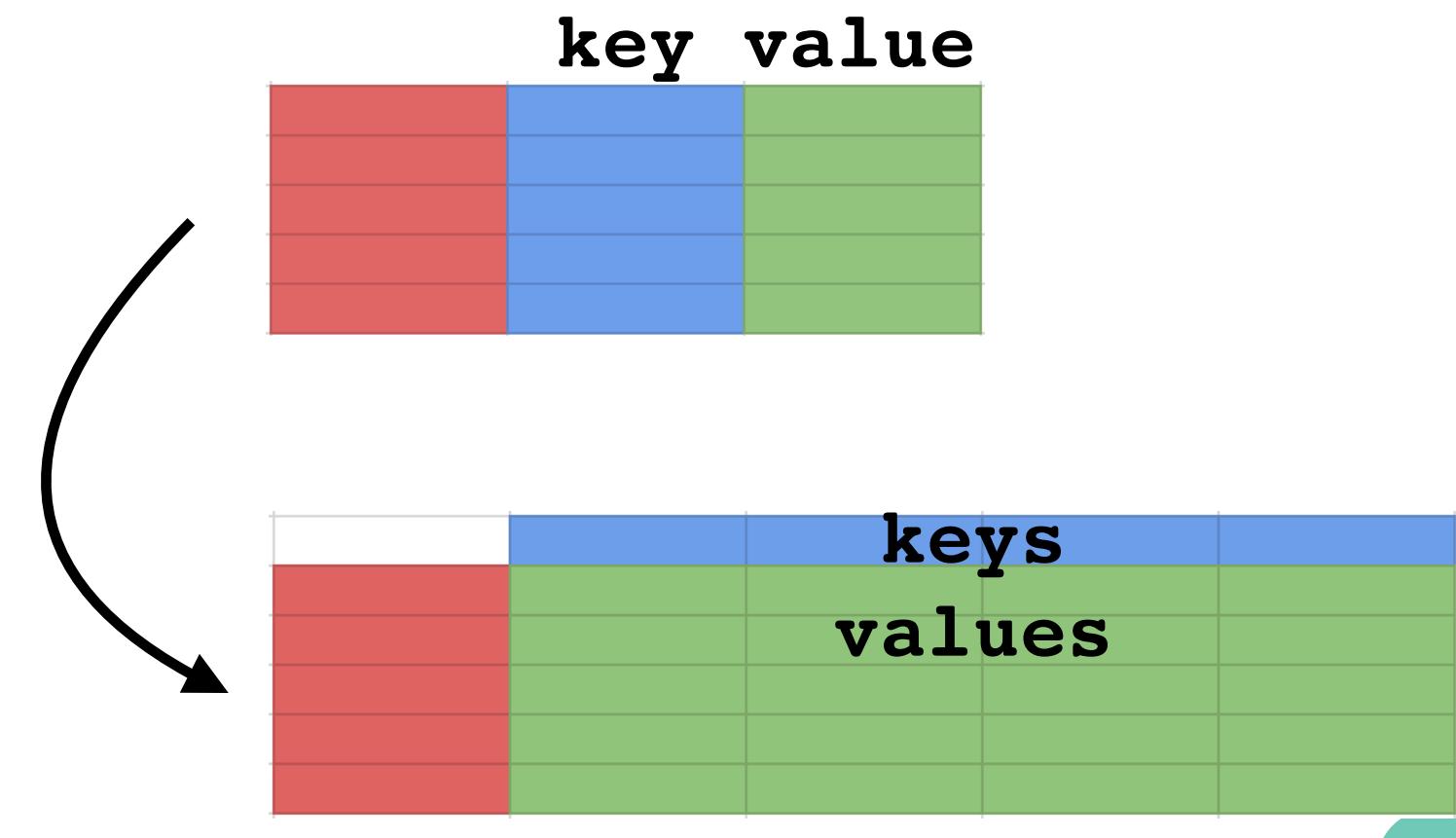
Pivot Tables in R

- **tidyverse**
 - Helps create *Tidy Data*
 - Each variable is in a column.
 - Each observation is a row.
 - Each value is a cell.
 - [https://tidyverse.org/](https://tidyverse.org)

gather(data, key, value, cols)



spread(data, key, value)



Agenda

- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- **Graphics**
- Reporting
- Resources for learning R

Plotting in R

- **ggplot2**
 - **Grammer of Graphics**
 - <https://ggplot2.tidyverse.org/>
 - **maps variables to aesthetics**
 - **x, y, color, size, shape,**
 - **layers on *geometry***

Agenda

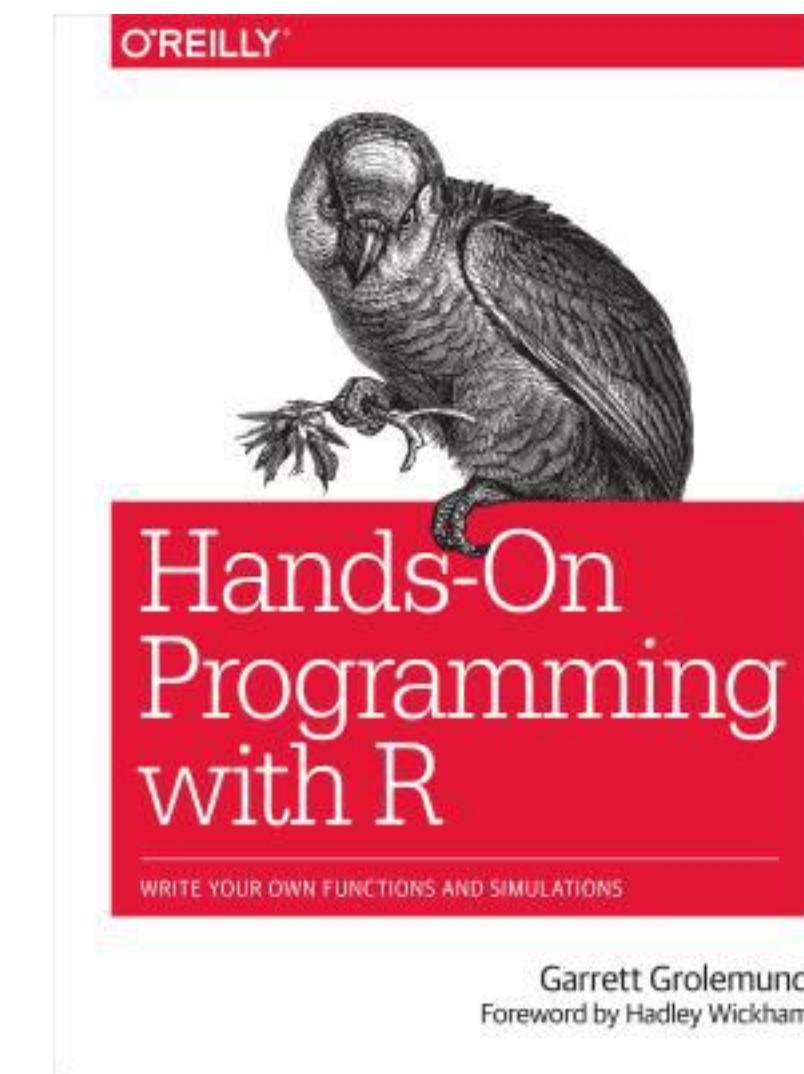
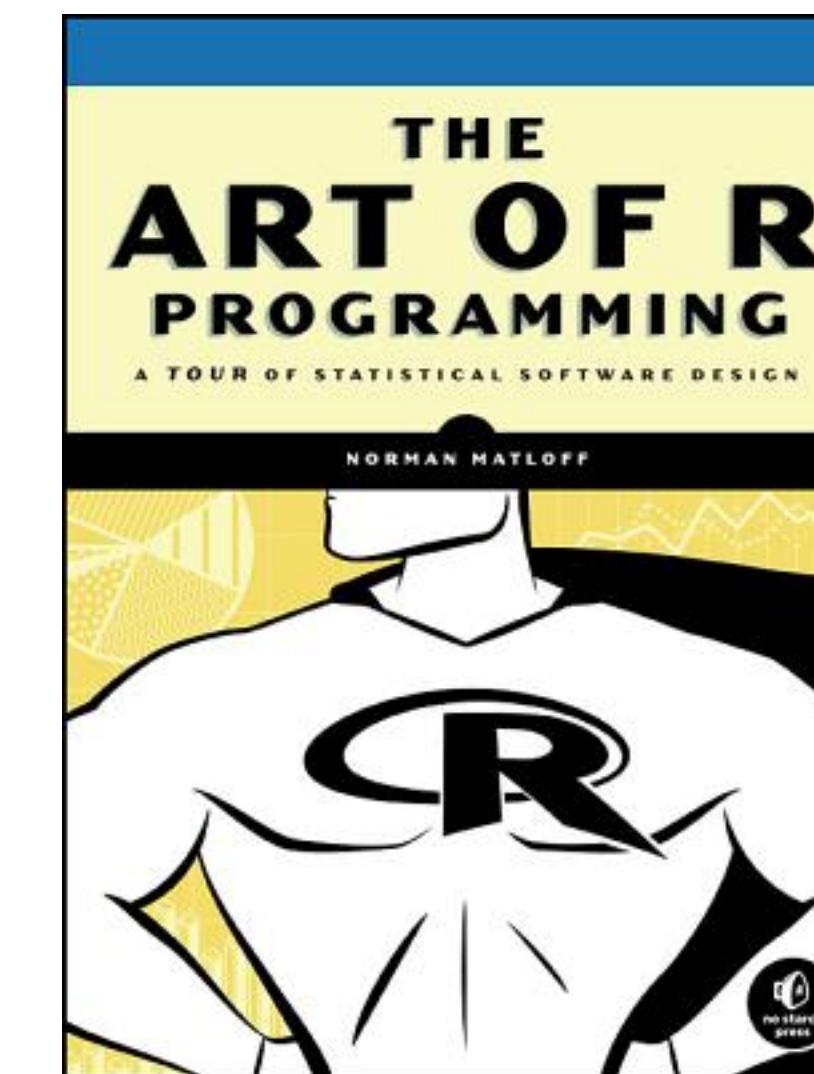
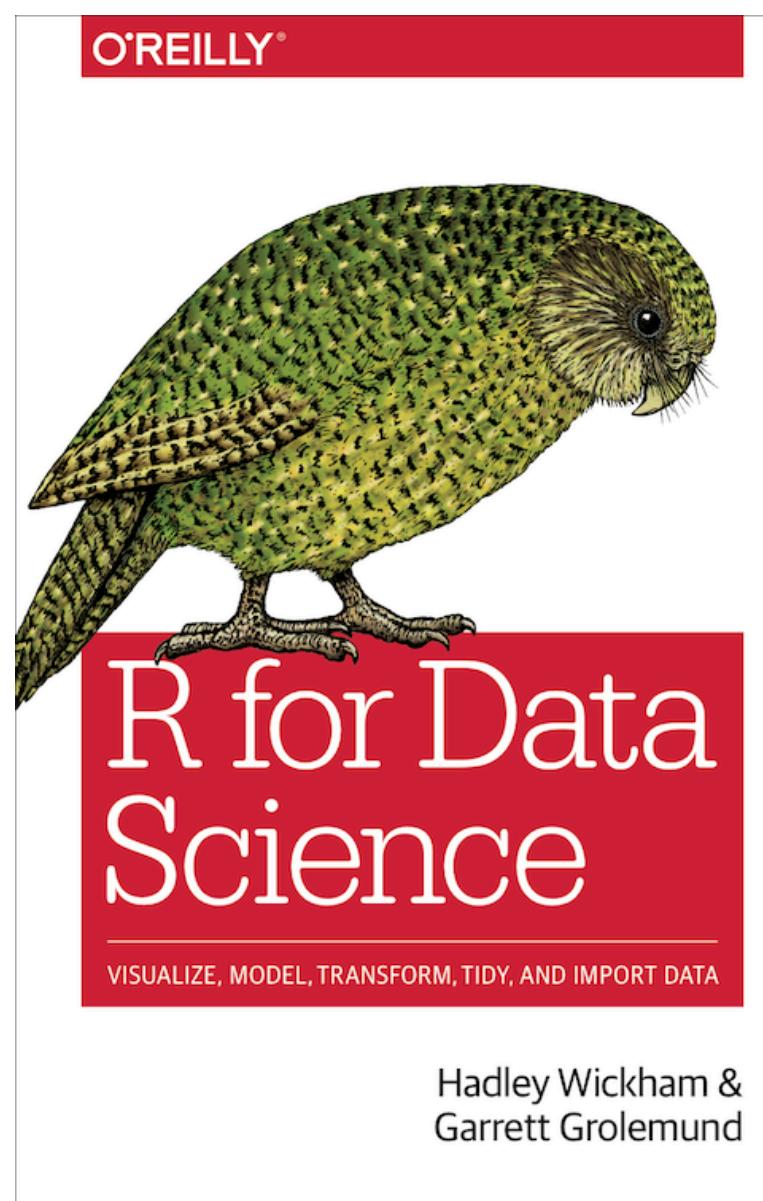
- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- Graphics
- **Reporting**
- Resources for learning R

Agenda

- Differences between R and Excel
- Reading in Data
- Basic Summary Statistics
- Merging Data
- Pivot Tables
- Graphics
- Reporting
- **Resources for learning R**

Books

- **R For Data Science by Hadley Wickham & Garret Grolemund**
 - <http://r4ds.had.co.nz/>
- **The Art of R Programming by Norman Matloff**
- **Hands on Programming with R by Garrett Grolemund**



Websites

- Stack Over Flow - R tag
 - <https://stackoverflow.com/questions/tagged/r>
- Twitter!
 - @hadleywickham, @dataandme, @juliasilge, #rstats
- <https://simplystatistics.org/>
 - more of a general stats blog
- <https://www.r-bloggers.com/>
 - R blog aggregator
- google
 - Pretend you are the Hulk
- ?help

My Journey

- Took some programming courses in college (Matlab & Python)
- Learned R to automate some workflows
- Slowly replaced more and more steps in excel with R
 - Learning was very task oriented, e.g. “how do I join two tables in R”
- Used R to learn statistics and take notes

Questions??



<https://twitter.com/AtlantaRUsers>



<https://github.com/AtlantaRUsers/Meetups>



<https://www.meetup.com/R-Users-Atlanta/>