Fifteen Years of Data Science in NY

Jared P. Lander
Lander Analytics

April 2, 2009

2
APR

Thursday, April 2, 2009

Introductions + Machine Learning with R + Financial Time Series

This is a past event 21 people went



Hosted by
Joshua Reich



Details

This first meeting will be your opportunity to help shape the direction for this group.

With recent exposure in the NY Times, R is become less of an obscure tool confined to academia. The R-SIG-Finance list is full of New Yorkers who apply R in financial settings. Inside Google, Facebook and other data intensive companies, R plays an important role in helping analysts understand ecosystems.

Copy event

Organizer tools

This is a past event



New York Open Statistical Programming Meetup
Public group

Thursday, April 2, 2009
7:00 PM to 10:00 PM
Add to calendar

Union Square Ventures
915 Broadway Suite 1408 · New York, NY



September 30, 2009

Collaborative Filtering with Temporal Dynamics



Hosted By
Paul D.

Details

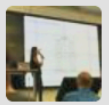
For our first meetup we'll be reviewing Yehuda Koren's paper "Collaborative Filtering with Temporal Dynamics (https://research.yahoo.com/files/kdd-fp074-koren.pdf)". He's on the BellKor Pragmatic Chaos team that is currently in contention for winning the Netflix grand prize.

Collaborative filtering is useful for building recommendation systems. You can apply this to making recommendations for movies, books, music, ads, content, web sites, users, or really anything you can think about wanting to recommend.

I'll lead a discussion of the paper, the Netflix prize in general, and collaborative filtering. Please read the paper before coming so you can participate in the discussion!

We'll start at 6:30 and probably go to 8.

You need to RSVP so security can let you in the door. Please bring an ID with you.



NYC Machine Learning
Public group

Wednesday, September 30, 2009
6:30 PM to 9:30 PM EDT
Add to calendar

kgbWeb
655 Madison Ave 3rd floor · New York, NY

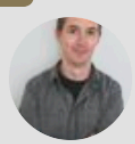
How to find us
call me on my cell at 646.283.6377

Report this event

Attendees (21)

See all

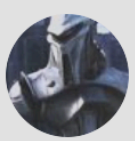
Host



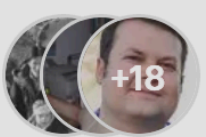
Paul
Organizer



hv
Member



Max
Co-organizer



+18 more
7 on waitlist

December 10, 2009

Recommendation Engine Demystified



Hosted By
Alex L.





NYC Predictive Analytics
Public group

Thursday, December 10, 2009
6:00 PM to 9:00 PM EST
Add to calendar

Bluefly
42 W 39th St. 10F, (between 6th Ave and 5th Ave) · NY, NY

Report this event

Details

Neighborhood based Collaborative Filtering
Presented by Alex Lin, Sr. Architect of Intelligent Mining
Presentation: 6:15 pm to 7:00 pm
Q&A and Group Discussion: 7:00pm - 8:00pm

Recommendation Engines (RE) are frequently used to drive automated predictions about user interests to provide a more relevant user experience. We will discuss the

May 20, 2009


COLUMBIA
UNIVERSITY




SONY
VAIO
Backspace
Enter
Shift
Ctrl
Alt
Fn
Tab
Caps Lock
Home
End
Pg Up
Pg Dn
Web access virtually anywhere in the U.S.
intel
Centrino 2
NVIDIA


NOKIA
3G
9:42 AM
Messages
Calendar
Tuesday
9
Photos
Camera
YouTube
Stocks
Maps
Weather
Voice Memos
Notes
Clock
Calculator
Settings
iTunes
App Store
Compass
Phone
Mail
Safari
iPod

July 6, 2009

This week's New York R Meetup is about Bayes

Posted on July 6, 2009 9:47 PM by Aleks Jakulin

This Thursday at 7pm Jake Hofman and Suresh Velagapundi will present a session at New York R Statistical Programming Meetup at NYU – Silver Center (100 Washington Square East, Room 401). Here's the outline:

Background:

- *Conditional probability & Bayes' Rule*
- *Treating parameters as random variables & putting distributions on them*
- *Bayesian inference: from priors & likelihoods to posteriors*

From Principles to Practice:

- *Simple plan; difficult to execute (normalization)*
- *Resort to approximation methods (variational & MCMC)*
- *Model selection / complexity control a la Bayes*

This entry was posted in Bayesian Statistics, Statistical Computing by Aleks Jakulin. Bookmark the permalink.

July 9, 2009

Attendee list for

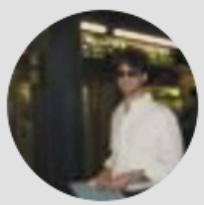
Bayesian Methods + MCMC

July 9, 2009

NYU - Politics Department

Manage attendees

jared



Jared Lander

First event

Went

Answer

"Whatever is most effective for the problem at hand"

Interested in

Open Source | Entrepreneurship | Technology | Startup Businesses | Computer Programming | +3 more

Looking to:

Tell us more about you

Update Profile



January 7, 2010

The R Rosetta Stone



Hosted By
Eric N. and 2 others



Copy event

Organizer tools



New York Open Statistical Programming Meetup
Public group

Thursday, January 7, 2010
7:00 PM to 10:00 PM EST
Add to calendar

Columbia - Statistics Playroom
International Affaird Building, Room 707, Columbia University · New York, NY

How to find us
420 W. 118 St between Amsterdam Ave & Morningside Dr. Enter from 118 St, the elevators are right there, come up to the 7th floor. Rm 707 is right next to the elevator bank; can't miss it

Report this event

Details

The votes are in, and by a slim margin the group has decided that we will focus on how R translates among multiple languages for the January meetup. As such, during this meetup we will discuss how basic programming functions, operations and paradigms translate between R and four other popular programming languages. Given the tight margin in the voting, however, we will also spend slightly more time discussing Python and R.

Here is the tentative schedule:

Python - Drew Conway
Matlab - Harlan Harris
SAS - Georgette Asherman
SQL - Josh Reich
Incanter - Shane Conway

April 8, 2010

What is New York Style?



Charring and air bubbles around the crust



Thin Slices



Just the right bend for folding

Hosts

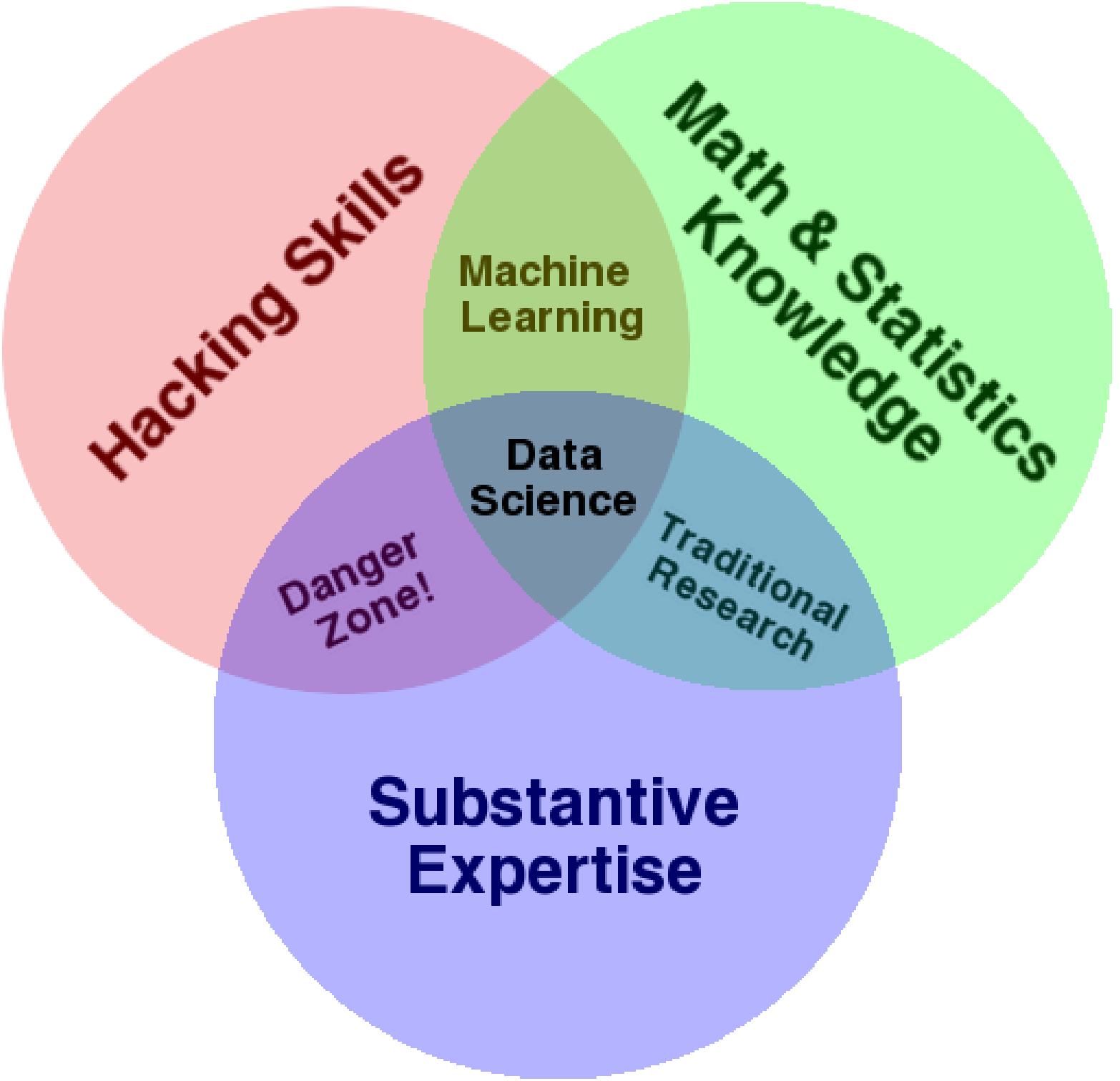
NYC Data Mafia


NYC
Data Mafia
sed
awk
grep
R


NYC Data Mafia
NYC Data Mafia
NY
I R
NY

Tools

- R
- SAS
- SPSS
- Stata
- MySQL
- JAGS/BUGS
- Matlab


Hacking Skills
Math & Statistics Knowledge
Machine Learning
Data Science
Danger Zone!
Traditional Research
Substantive Expertise

August 9, 2011

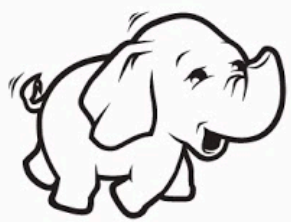
NY R Meetup

NY Open Statistical Programming Meetup

Strata Conference

O'REILLY®

Strata CONFERENCE + HADOOP WORLD



Tools and Techniques That Make Data Work

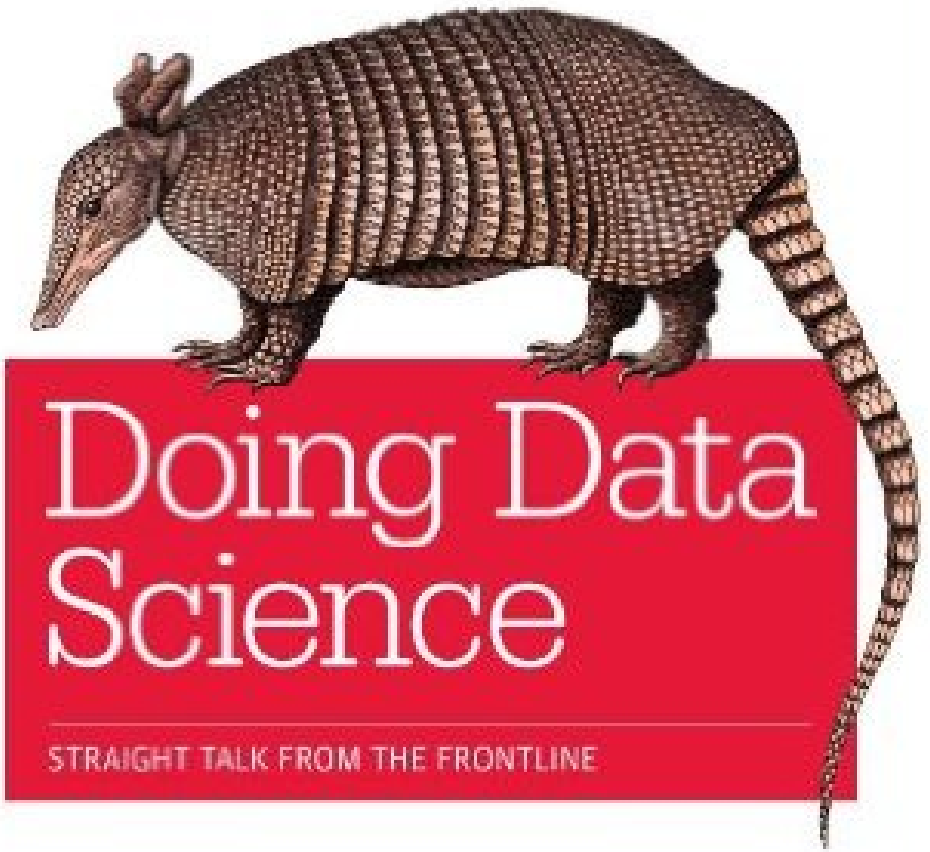
June 5, 2012


I R
NY

Fall 2012


EXIT
data nerd

O'REILLY
Copyrighted Material



Doing Data Science

STRAIGHT TALK FROM THE FRONTLINE

Cathy O'Neil & Rachel Schutt

Copyrighted Material

April 24-25, 2015


NYR
NEW YORK R CONFERENCE


NYR
NEW YORK R CONFERENCE
#rstatsnyc






RStudio
RStudio


I NY
I NY


NYR
NEW YORK R CONFERENCE
#rstatsnyc


NYR

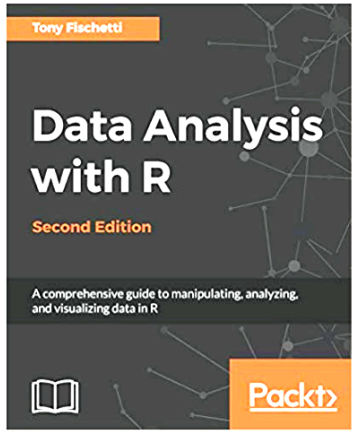

NYR

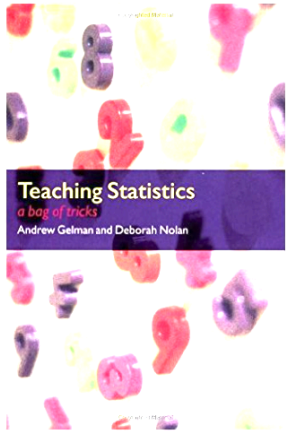

NYR

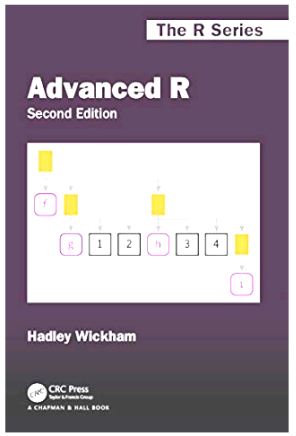
November 21, 2016

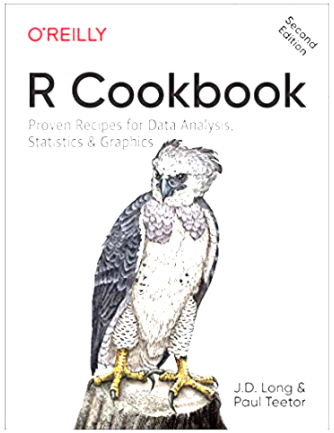

R-Ladies
NYC

Community

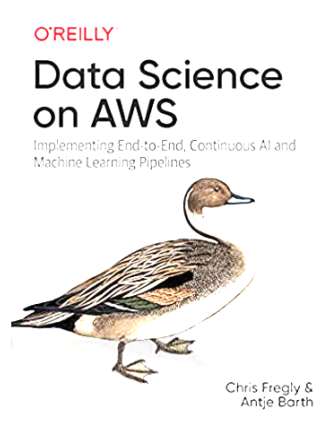

Tony Fischetti
Data Analysis with R
Second Edition
A comprehensive guide to manipulating, analyzing, and visualizing data in R
Packt

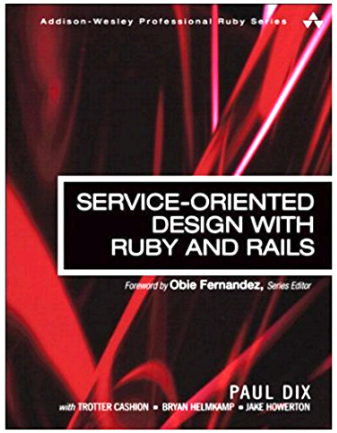

Teaching Statistics
a bag of tricks
Andrew Gelman and Deborah Nolan

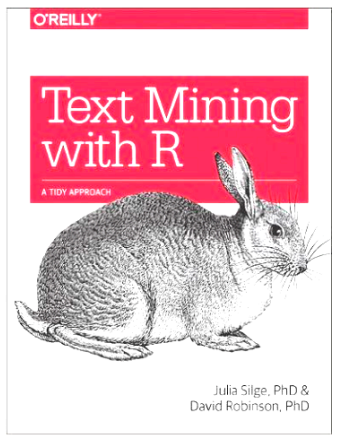

The R Series
Advanced R
Second Edition
Hadley Wickham
CRC Press

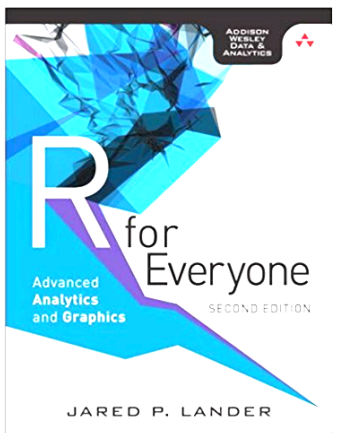

O'REILLY
Second Edition
R Cookbook
Proven Recipes for Data Analysis, Statistics & Graphics
J.D. Long & Paul Teetor

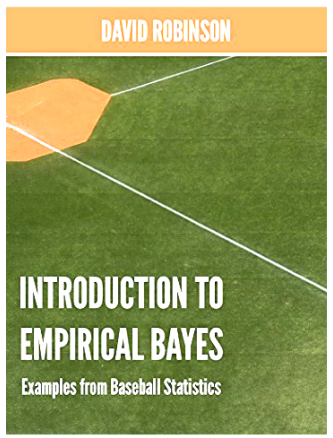

O'REILLY
Doing Data Science
STRAIGHT TALK FROM THE FRONTLINE
Cathy O'Neil & Rachel Schutt

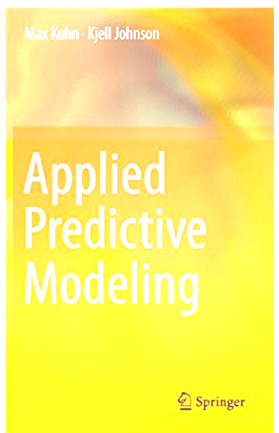

O'REILLY
Data Science on AWS
Implementing End-to-End, Continuous AI and Machine Learning Pipelines
Chris Fregly & Antje Barth

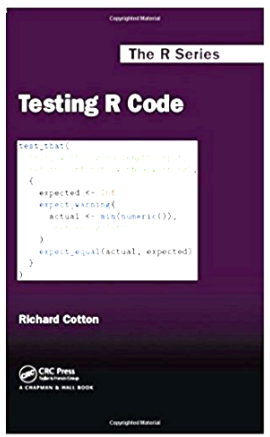

Addison-Wesley Professional Ruby Series
SERVICE-ORIENTED DESIGN WITH RUBY AND RAILS
Foreword by Obie Fernandez, Series Editor
PAUL DIX
with TROTTER CASHION • BRYAN HELMKAMP • JAKE HOWERTON


O'REILLY
Text Mining with R
A TIDY APPROACH
Julia Silge, PhD & David Robinson, PhD

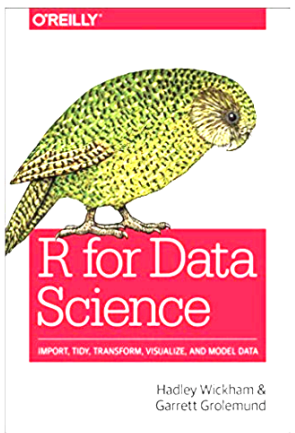

ADDISON WESLEY DATA & ANALYTICS
R for Everyone
Advanced Analytics and Graphics
SECOND EDITION
JARED P. LANDER

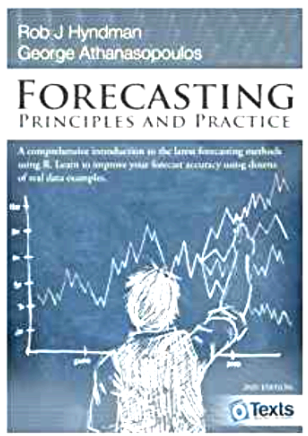

DAVID ROBINSON
INTRODUCTION TO EMPIRICAL BAYES
Examples from Baseball Statistics


Max Kuhn · Kjell Johnson
Applied Predictive Modeling
Springer

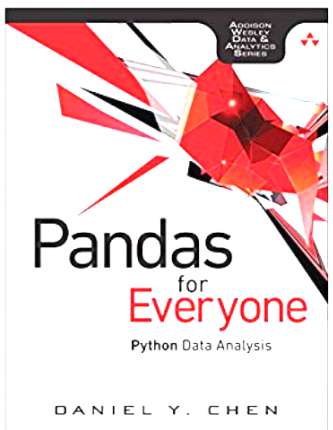

The R Series
Testing R Code
Richard Cotton
CRC Press

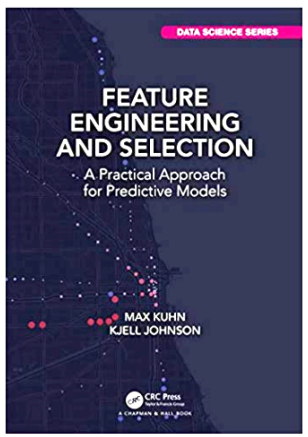

O'REILLY
R Packages
ORGANIZE, TEST, DOCUMENT, AND SHARE YOUR CODE
Hadley Wickham

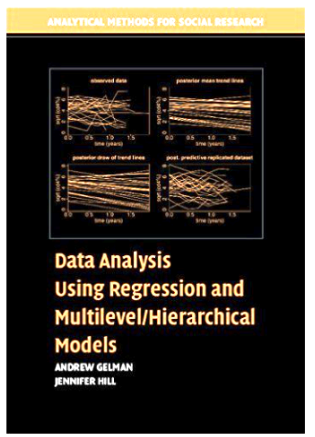

O'REILLY
R for Data Science
IMPORT, TIDY, TRANSFORM, VISUALIZE, AND MODEL DATA
Hadley Wickham & Garrett Grolemund


Rob J Hyndman
George Athanasopoulos
FORECASTING
PRINCIPLES AND PRACTICE
Texts

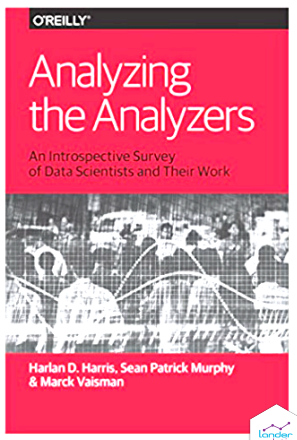

Getting Started with
D3
O'REILLY


ADDISON WESLEY DATA & ANALYTICS SERIES
Pandas for Everyone
Python Data Analysis
DANIEL Y. CHEN


DATA SCIENCE SERIES
FEATURE ENGINEERING AND SELECTION
A Practical Approach for Predictive Models
MAX KUHN
KJELL JOHNSON
CRC Press


ANALYTICAL METHODS FOR SOCIAL RESEARCH
Data Analysis Using Regression and Multilevel/Hierarchical Models
ANDREW GELMAN
JENNIFER HILL


Bad Data
Handbook
O'REILLY


O'REILLY
Analyzing the Analyzers
An Introspective Survey of Data Scientists and Their Work
Harlan D. Harris, Sean Patrick Murphy & Marck Vaisman

nyc.data.bikers <-  +  + 

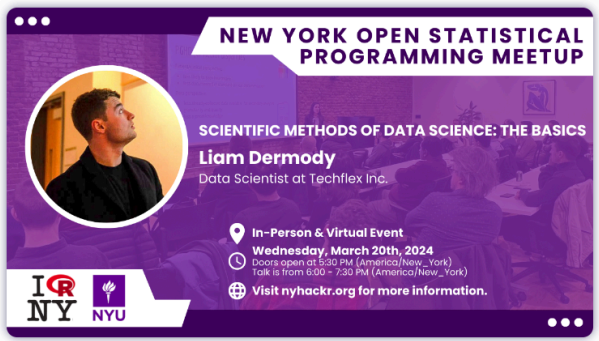

bicycle habitat
4212
3614

DataKind

Past talks







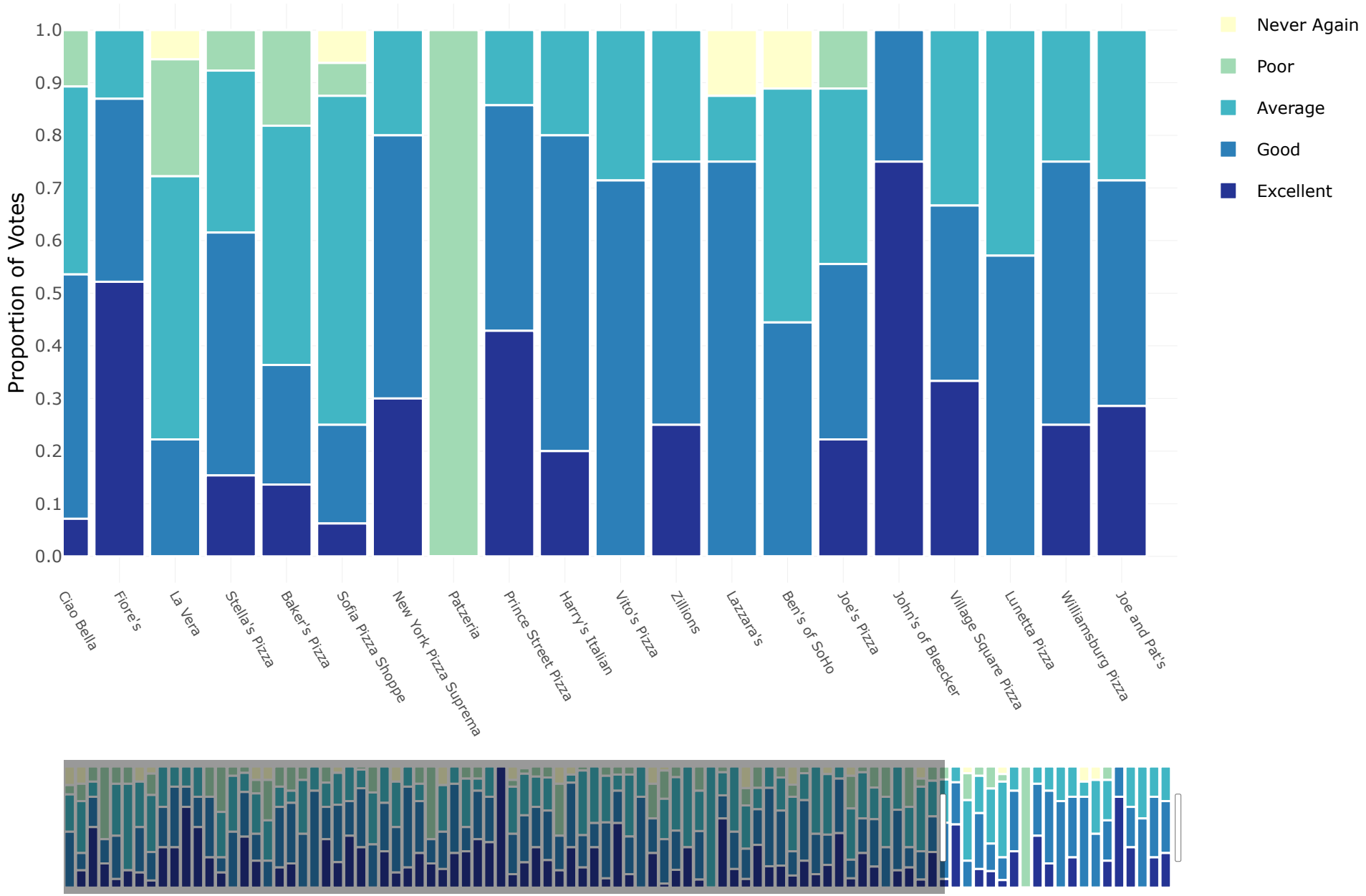


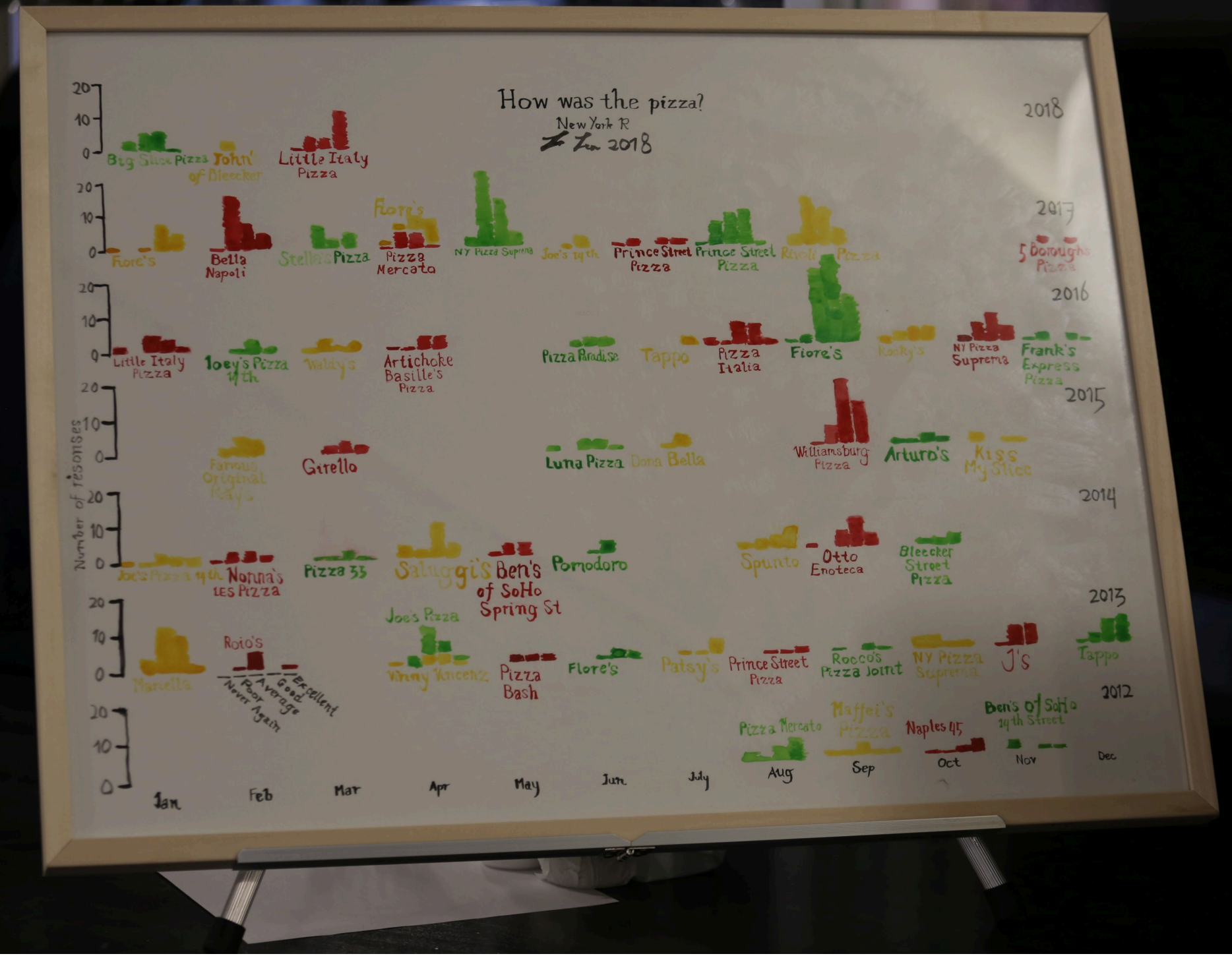
Archive

Show 10 entries

Search:

Detail	Date	Talk	Speaker	Video
	All	All	All	
+	2024-05-07	Turning Polars DataFrames into Pretty Pictures and Great Tables	Jeroen Janssens	-
+	2024-04-25	Put an LLM on it!	Bryan Lewis	▶
+	2024-03-20	Scientific Methods of Data Science: The Basics	Liam Dermody	▶


Proportion of Votes
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Never Again
Poor
Average
Good
Excellent
Ciao Bella
Fiore's
La Vera
Stella's Pizza
Baker's Pizza
Sofia Pizza Shoppe
New York Pizza Suprema
Patzeria
Prince Street Pizza
Harry's Italian
Vito's Pizza
Zillions
Lazzara's
Ben's of SoHo
Joe's Pizza
John's of Bleecker
Village Square Pizza
Lunetta Pizza
Williamsburg Pizza
Joe and Pat's

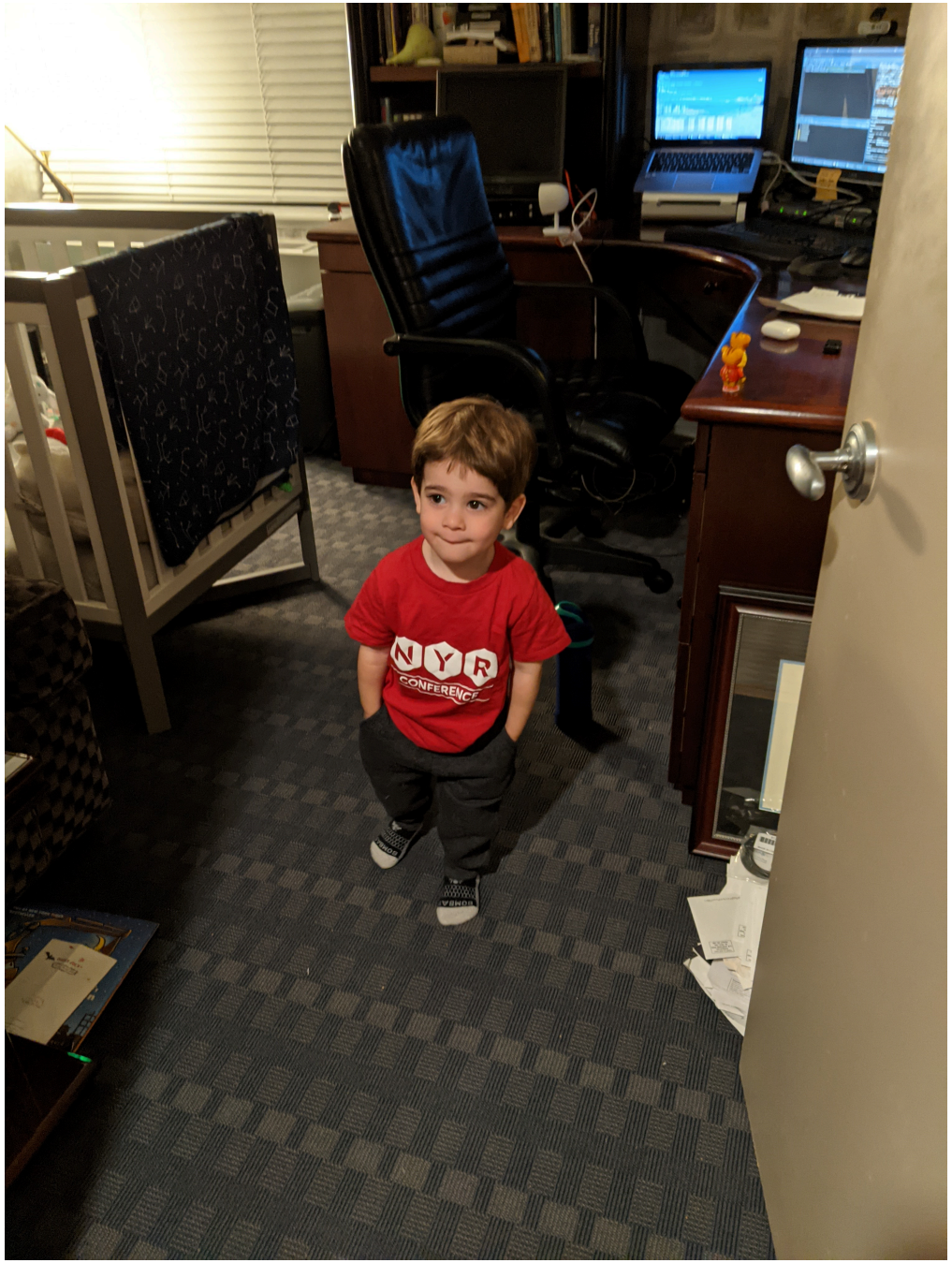

How was the pizza?
New York R
2018
2017
2016
2015
2014
2013
2012
Number of resonses
20
10
0
Big Slice Pizza
John' of Bleecker
Little Italy Pizza
Fiore's
Bella Napoli
Stella's Pizza
Fiore's Pizza Mercato
NY Pizza Suprema
Joe's 14th
Prince Street Pizza
Prince Street Pizza
Rivoli Pizza
5 Boroughs Pizza
Little Italy Pizza
Joey's Pizza 14th
Waldy's
Artichoke Basille's Pizza
Pizza Paradise
Tappo
Pizza Italia
Fiore's
Rocky's
NY Pizza Suprema
Frank's Express Pizza
Famous Original Ray's
Girello
Luna Pizza
Dona Bella
Williamsburg Pizza
Arturo's
Kiss My Slice
Joe's Pizza 14th
Nonna's 1ES Pizza
Pizza 33
Saluggi's
Ben's of SoHo
Spring St
Pomodoro
Joe's Pizza
Spunto
Otto Enoteca
Bleecker Street Pizza
Manella
Roto's
Poor
Average
Good
Excellent
Never Again
Vinny Vincenz
Pizza Bash
Flore's
Patsy's
Prince Street Pizza
Rocco's Pizza Joint
NY Pizza Suprema
J's
Tappo
Pizza Mercato
Maffei's Pizza
Naples 45
Ben's of SoHo 14th Street
Jan
Feb
Mar
Apr
May
Jun
July
Aug
Sep
Oct
Nov
Dec






TARWIN & HENSON
Dewey, Illinois
NOMAD
fresh
local
organic
P·I·Z·Z·A
fresh
local
organic
traveling
wood-fired
brick oven




NYR
CONFERENCE


NYR
NEW YORK R CONFERENCE

Academics



COLUMBIA UNIVERSITY

DATA SCIENCE INSTITUTE



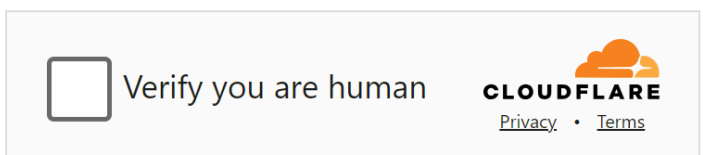
Center for
Data Science

Artificial Intelligence @ Columbia



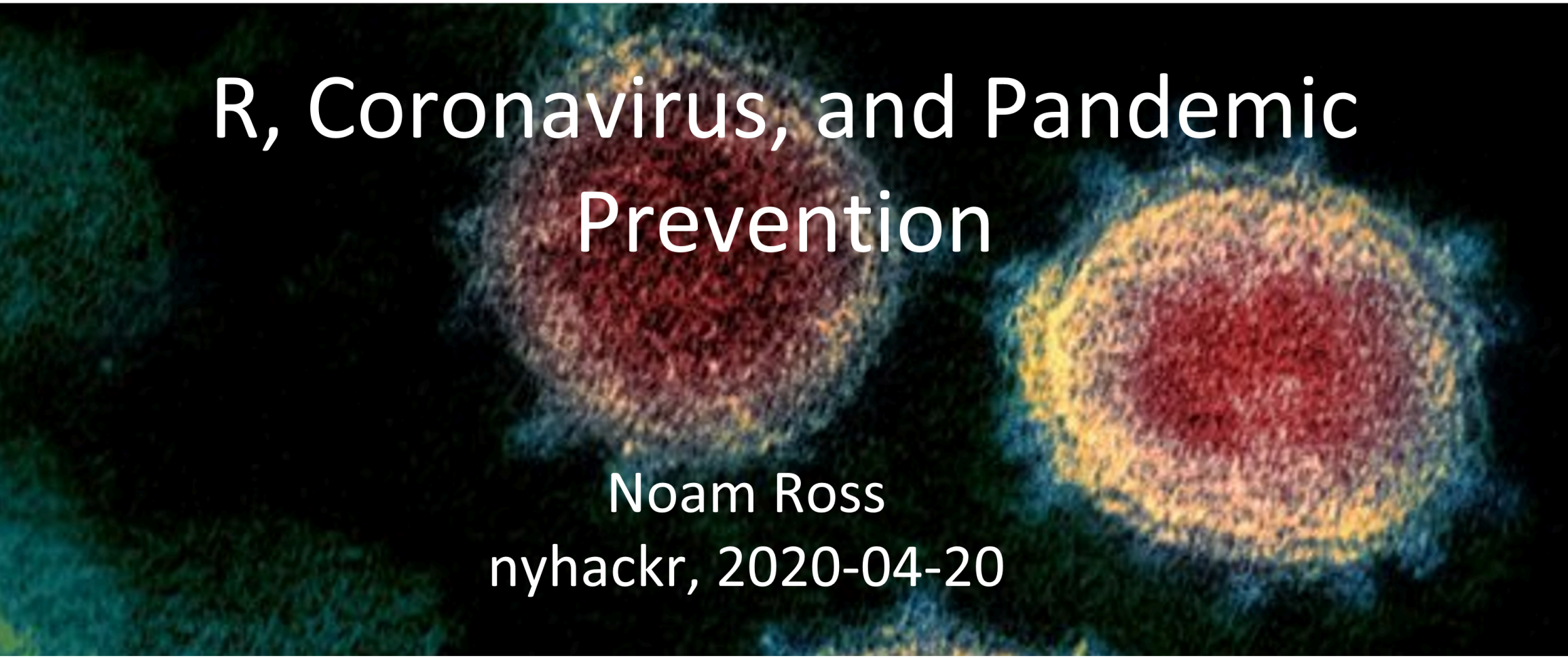
ai.columbia.edu

Verify you are human by completing the action below.



ai.columbia.edu needs to review the security of your connection before proceeding.

Pandemic



R, Coronavirus, and Pandemic Prevention

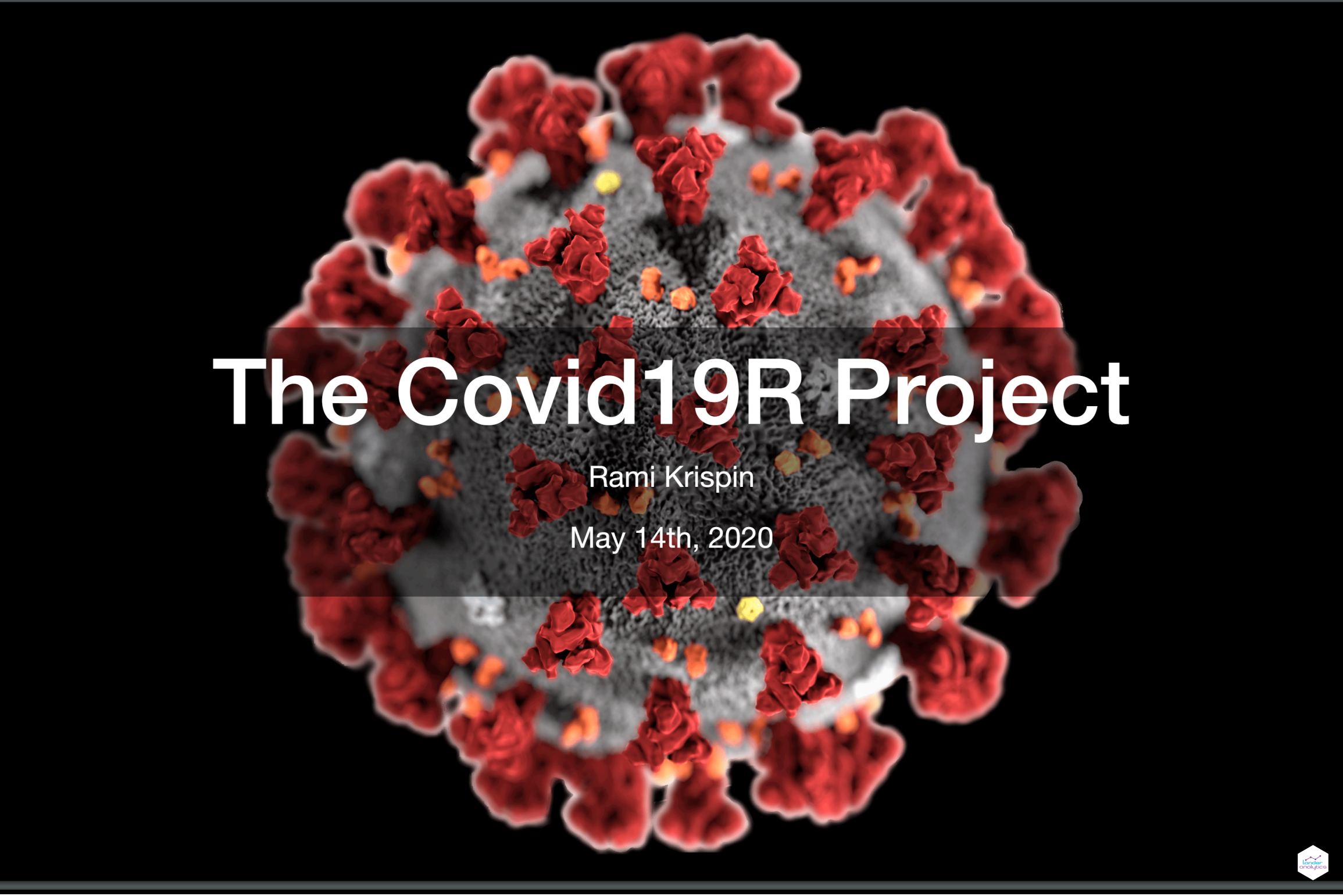
Noam Ross
nyhackr, 2020-04-20

noamross @

EcoHealth Alliance

RopenSci

tarides analytics



The Covid19R Project

Rami Krispin

May 14th, 2020

Learning when, where, and by how much, things change

Gavin Simpson

New York R Meetup • June 22 2020


01 16
Agenda
Job Board
lenovo
ASUS


CONFERENCE
NEW YORK
2020
WHERE R ENTHUSIASTS AND DATA SCIENTISTS
GATHER TO EXPLORE, SHARE AND INSPIRE IDEAS
#rstatsnyc
Presented by landeranalytics
I R DC
RStudio
DCR
WASHINGTON DC R CONFERENCE
NYR
NEW YORK R CONFERENCE
MARK E. FENNER
MACHINE LEARNING WITH PYTHON FOR EVERYONE
lander analytics
PROGRAMMING SKILLS FOR DATA SCIENCE
MICHAEL FREEMAN | JOEL ROSS
R for Everyone
JARED P. LANDER
Pandas for Everyone
Python Data Analysis
DANIEL Y. CHEN
rlang
coefplot
Jon Krohn
with Aglaé Bassens
& Beyleveld Grant
DEEP LEARNING ILLUSTRATED
I R NY
NYC Data MaFia

NEW YORK OPEN STATISTICAL PROGRAMMING MEETUP

SHINY SPEEDRUN: IMPROVISING AN INTERACTIVE DATA VISUALIZATION



David Robinson

Director of Data Science at Heap Analytics

In-Person & Virtual Event

Thursday, June 2nd, 2022

Doors open at 5:45 PM (America/New_York)

Talk starts at 6:30 PM (America/New_York)

Visit nyhackr.org for more info.

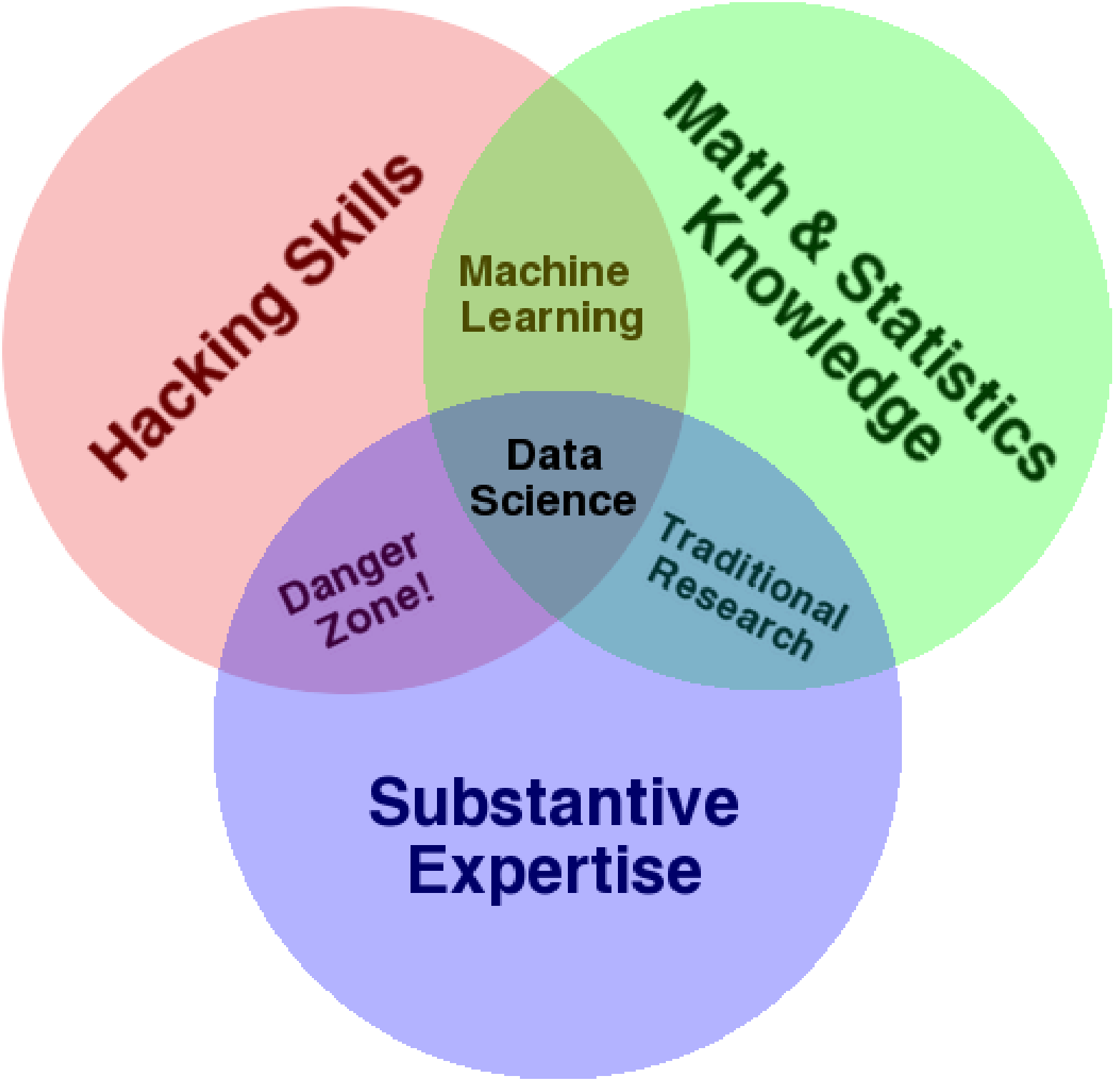
I ♥ R NY

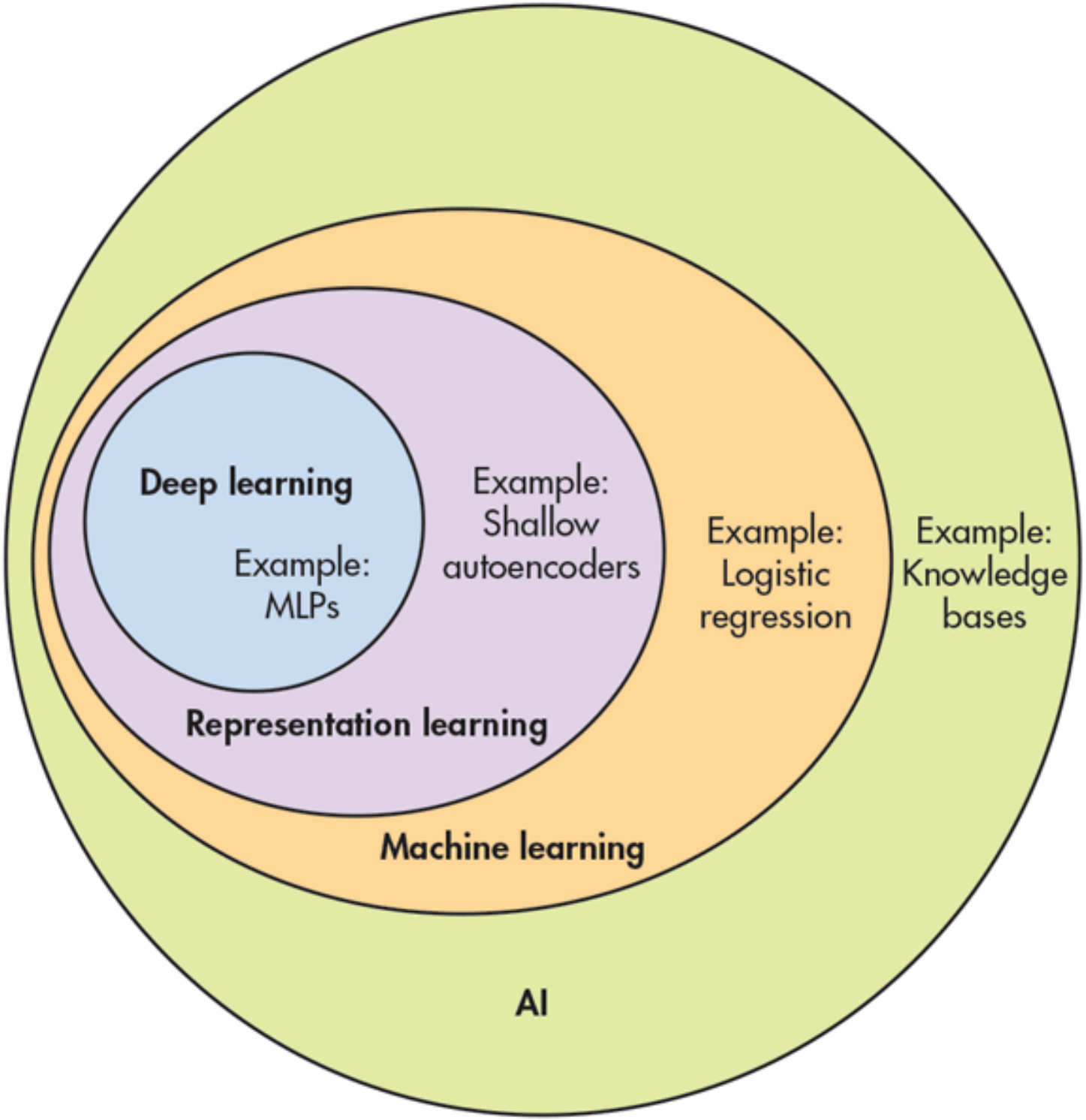
Microsoft


BASEBALL
Coke

New Tools

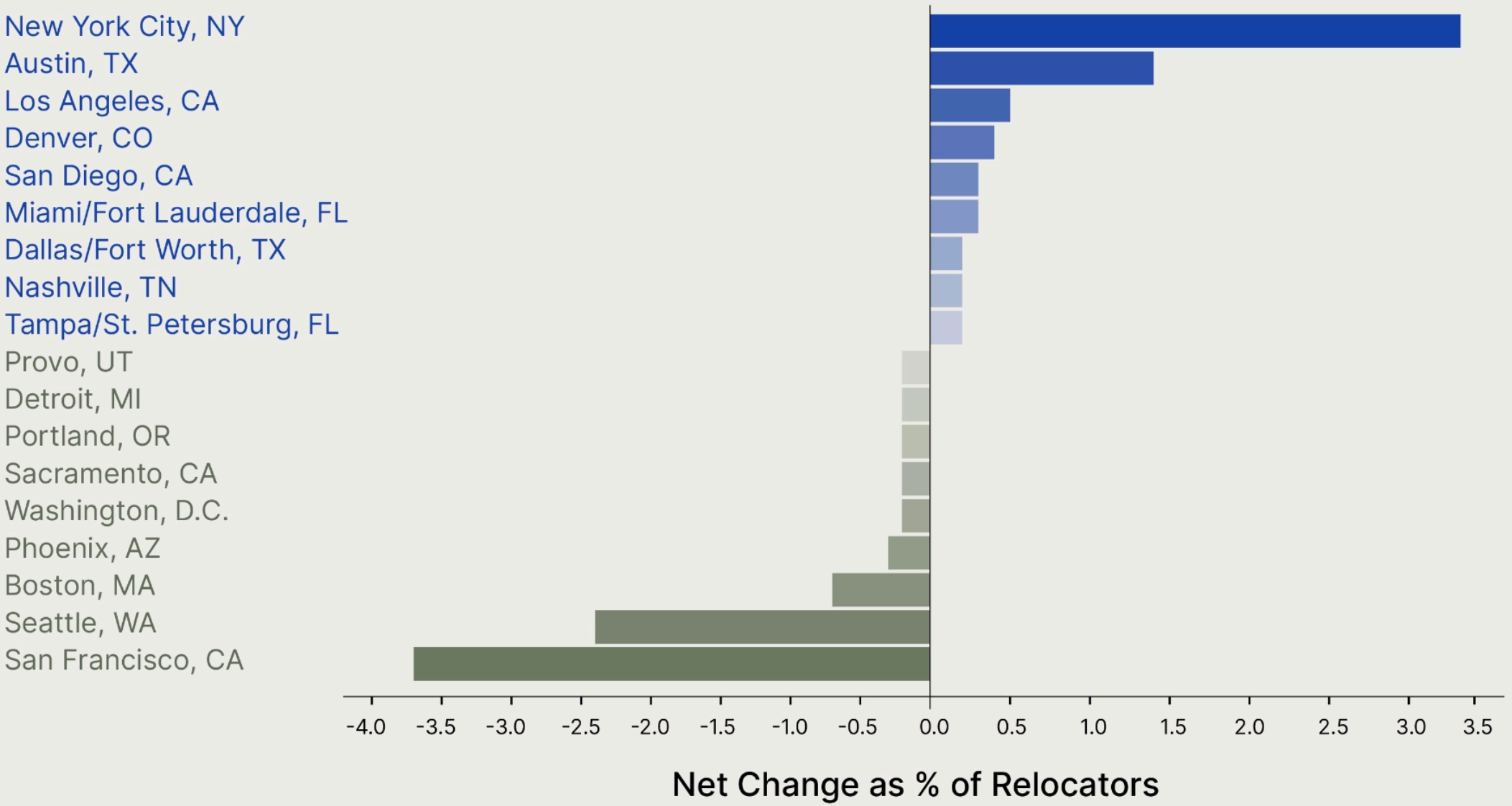
- R
- Python
- SQL
- Snowflake
- Postgres
- DuckDB
- Spark
- Stan
- Torch


Hacking Skills
Math & Statistics Knowledge
Machine Learning
Data Science
Danger Zone!
Traditional Research
Substantive Expertise

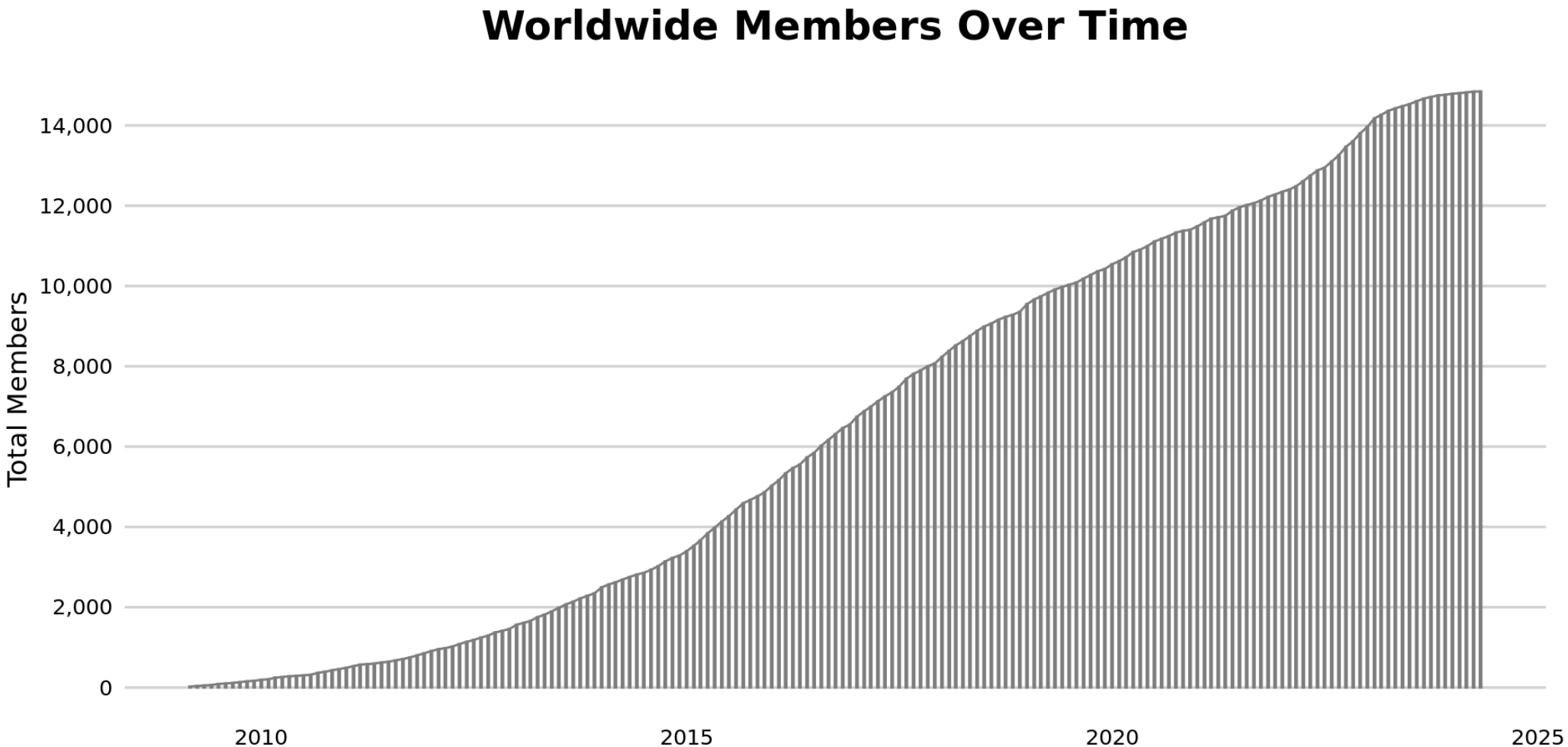

Deep learning
Example: MLPs
Example: Shallow autoencoders
Representation learning
Example: Logistic regression
Machine learning
Example: Knowledge bases
AI

New York Continues to Grow

The Big Apple gets bigger



SignalFire


Worldwide Members Over Time
Total Members
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
2010
2015
2020
2025



AI Startups Are Making Their Home in New York. Can They Turn It Into an AI Powerhouse?

A new class of startups is setting up shop in NYC instead of Silicon Valley, hoping to tap its concentration of businesses ready to buy AI applications

The Big Apple is making a bid to become a premier hub for generative artificial intelligence, according to investors, founders and businesses with ties to the region. MIKE SEGAR/REUTERS

By Belle Lin

May 11, 2024 8:00 am ET | WSJ PRO

Gift unlocked article

Listen (9 min)


AlphaSense
ASAPP
AUGURY
BigID
Bluecore
BlueVoyant
bob
BOWERY
CAPITOLIS
CERTIK
dataiku
Dataminr
Electric
glia
gloat
Hugging Face
(h[s]) HYPERSCIENCE
JELLYSMACK
Justworks
health
LIQUIDITY
NOOM
opentrons
Prove
ramp
runway
SANDBOXAQ
SecurityScorecard
Spring Health
sword
THIRTY MADISON
VAST
verbit
via
vise


EMPIRE AI

NYC

The New York City Artificial Intelligence Action Plan



October 2023

EVEN WHEN THE NEWS IS FREE, JOURNALISM IS NOT.
SUPPORT INDEPENDENT, FACT-BASED JOURNALISM. DONATE

U.S. NEWS

NYC's AI chatbot was caught telling businesses to break the law. The city isn't taking it down



FILE - New York City Mayor Eric Adams speaks during a news conference at City Hall, Dec. 12, 2023, in New York. An artificial intelligence-powered chatbot meant to help small business owners in New York City has come under fire for dispensing bizarre advice that misstates local policies and advises companies to violate the law. But the chatbot remains online, even as Adams acknowledged Tuesday, April 2, 2024, that its answers were "wrong in some areas." (AP Photo/Peter K. Afriyie, File)



BY JAKE OFFENHARTZ
Updated 7:11 PM EDT, April 3, 2024

Share


NYC OpenData


I
R
NY

Thank You


lander
analytics