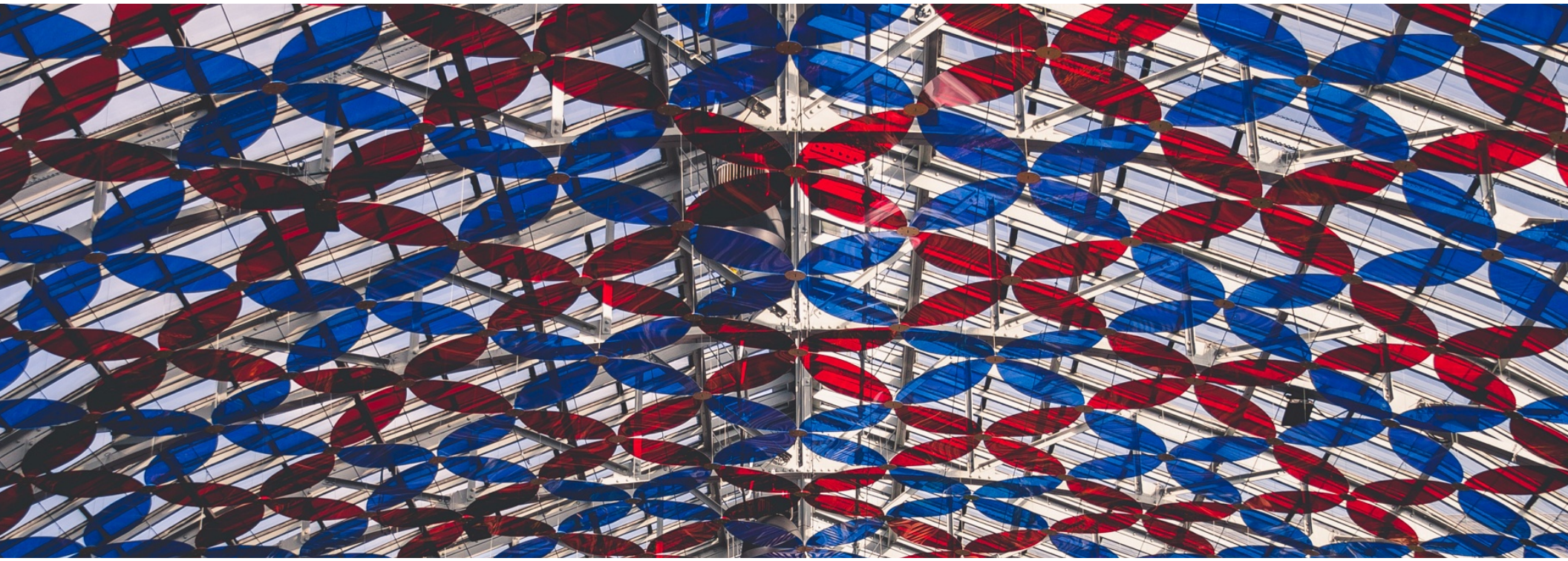


CS 2050 Hackathon

Michael Saelee <saelee@iit.edu>
Department of Computer Science

IIT College of Science
ILLINOIS INSTITUTE OF TECHNOLOGY

Thanks for registering & attending!

"Data Hackathon"

- Traditional hackathons were often driven by language, framework, or API
- Increasing visibility of big data and data science → *data* hackathons
  - Driven by the data being processed!

Data Science

- Glean information from heaps of data
- Leverages statistics, data mining, AI, visualization, and domain-specific knowledge (e.g., natural and social sciences)

Data Sets

- City of Chicago data portal: https://data.cityofchicago.org
- Illinois data portal: https://data.illinois.gov
- Any other open-sourced, Chicago municipal data

Themes

- Technology and Education
- Crime Monitoring and Prevention
- Distribution of Civic Funding

… or propose your own!

Identifying a Topic

1. What sort of issue are we trying to identify or solve?
2. What is the scope of the issue?
3. Who might benefit from the insights we glean?

Three Tracks

1. Analysis
2. Visualization
3. “App”

Track 1: Analysis

- Application of data science techniques to derive useful insights from data
- Based on statistical models and/or machine learning algorithms

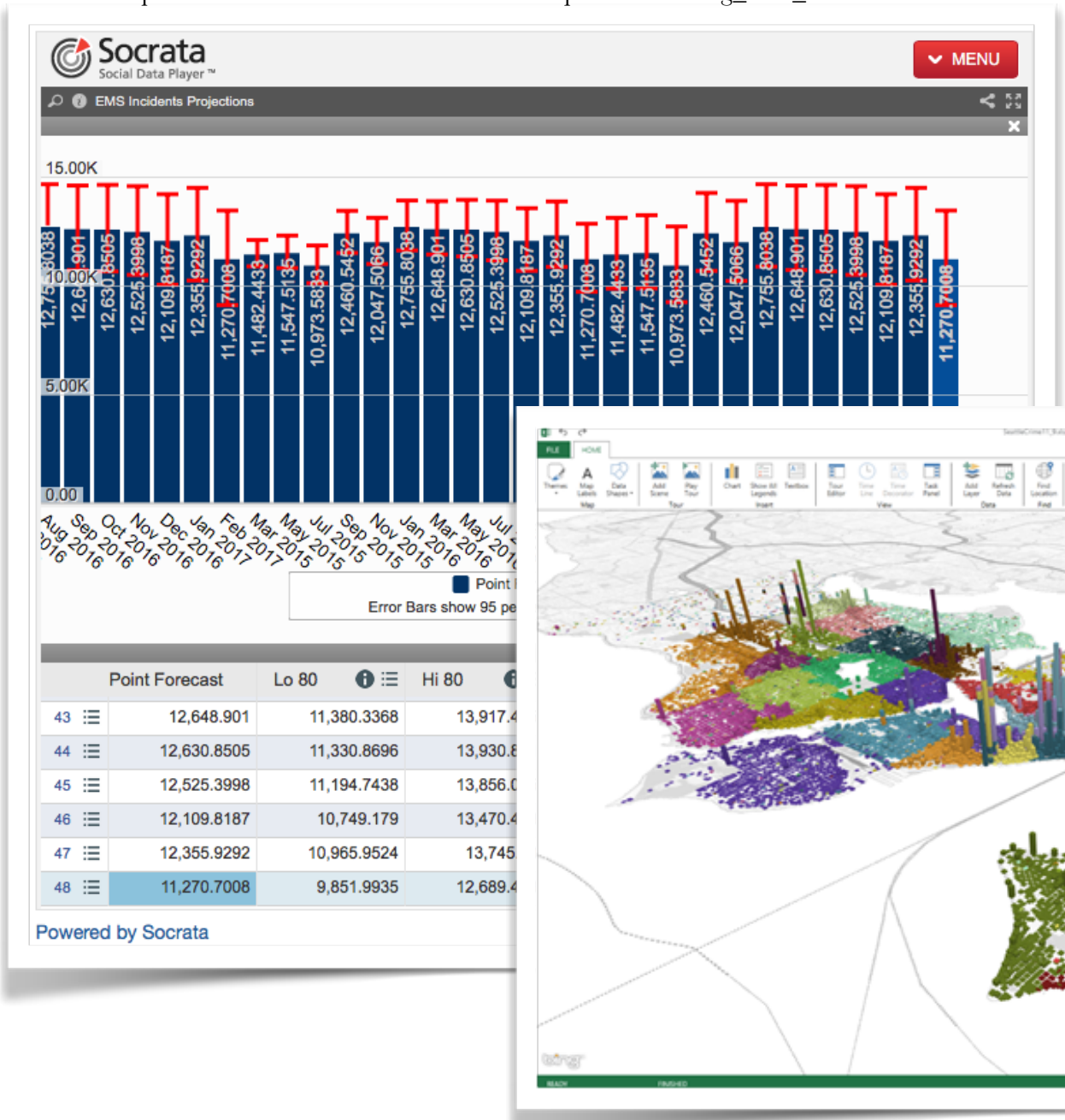
Track 1 judging

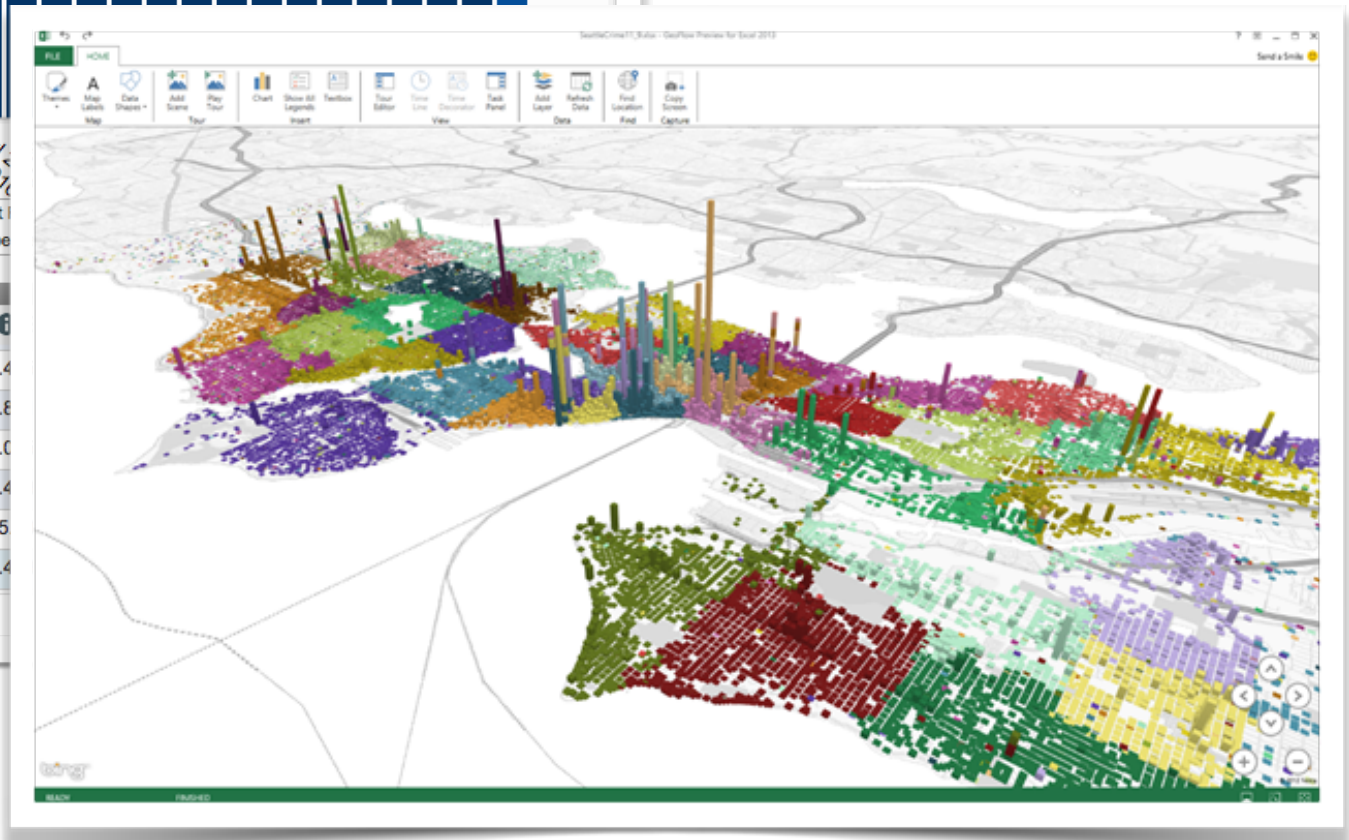
- Criteria:
  - Adherence to good data science practices
  - Technical sophistication
  - Strength and impact of result(s)

Track 2: Visualization

- $\max_{i \in \mathbb{N}}$(A picture is worth i words)
- Graphically render data so as to reveal trends and facilitate intuition

https://dev.socrata.com/consumers/examples/forecasting_with_rsocrata.html





https://blogs.bing.com/search/2013/04/11/geoflow-for-excel-3d-big-data-visualization-built-on-bing-maps/

Track 2 judging

- Criteria:
  - Accurate representation of data
  - Aesthetics and sophistication
  - Added insight / intuition

Useful Tools

- Python: Pandas, Matplotlib
- R (r-project.org)
- Mathematica
- Microsoft Excel!

Track 3: “App”

- An interface that allows users to query and/or otherwise interact with data
- E.g., chat-bot, web service/application, mobile app

App APIs

- Data APIs: Socrata, Plenario
- Frameworks: Express, Sinatra, Flask
- Mapping: Google Maps, Leaflet
- “Bots”: Wit.ai, Api.ai

Track 3 judging

- Criteria:
  - Good usability and interface design
  - Technical sophistication
  - Utility and impact

Collect them all!

- Each team can be entered in 1-3 tracks
  - You don't need to declare a track until presentation/judging

Awards

- Overall winner: $600
- Track winners (3): $400 per team
- Wildcard: $200

Hackathon Rules

- *Teams*: 3-5 members (min is enforced!)
- *Dataset*: must be publicly available, and related to the City of Chicago
- *Code*: must be open sourced — may reuse existing code with appropriate licenses
- *Presentations*: 5 minutes per team after end-of-hacking tomorrow @ 2:30PM

(Official) Timeline

- Thursday:
  - 6:00PM Introductions
  - 6:30PM Team formation & Brainstorming
  - 7:30PM Dinner

(Official) Timeline

- Friday:
  - 7:00AM Breakfast
  - 12:30PM Lunch
  - 2:30PM End of hacking
  - 3:00PM Presentations & Judging
  - 4:30PM Award presentation

Workshops & Mentors

- Ad hoc workshops and mentors to be announced and introduced over the course of the event!