

What Works Cities

Bloomberg
Philanthropies

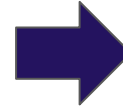
SUMMIT ON
TRANSFORMING
DATA
≡ INTO ≡
ACTION

3

Education



GovEx



How do you know which students are at risk of going off track **BEFORE** they're off track?

- Identify risk factors
- Build a model to predict status using available data
- Apply the model to student cohorts

How does this question apply to cities?

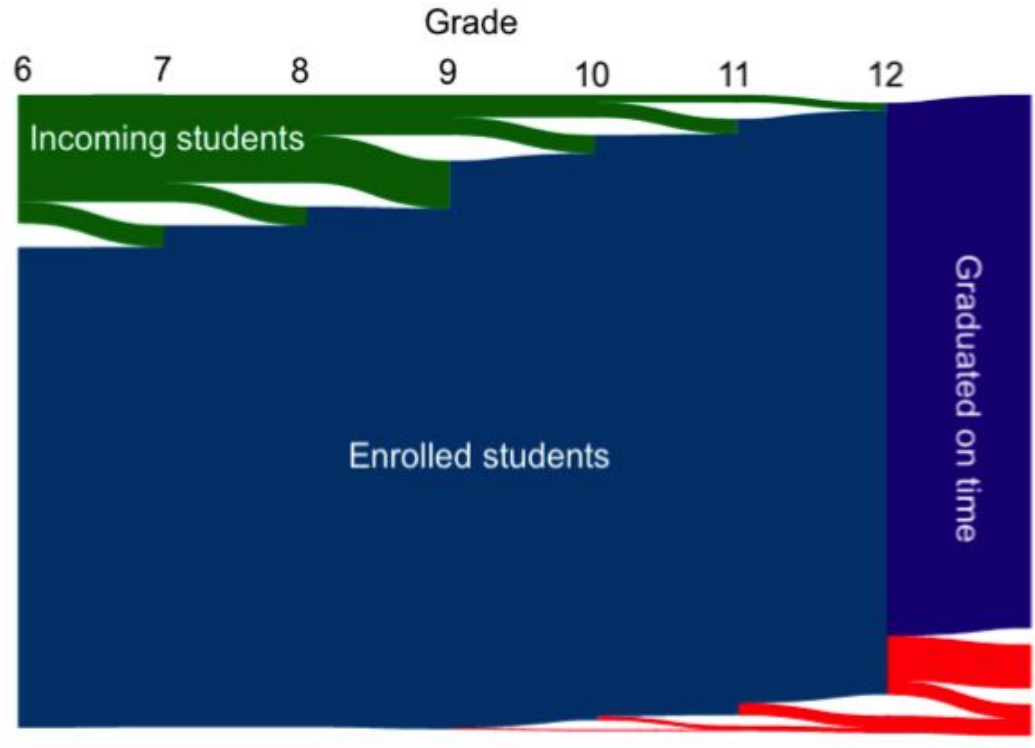
Community interest in education

- Education is key to a city's economic health and intervention helps increase economic potential in the aggregate
- At-risk students may interact with other city services like parks and recreation, social services, or public safety
- Cities can partner in developing intervention strategies through a variety of service areas to benefit mutual interests

Apply a risk model to other service areas

- Use data in other service areas to build a “risk model”
- Combine education data with data from other service areas to target services

The role of data



12% of students
do not graduate on time

- **Individual longitudinal data (over 4-5 years)**
 - Demographics
 - Performance (grades and test scores)
 - Attendance
- **Repeated for each cohort**
- **Data are inherently dirty and need a lot of preparation for analysis**

The role of technology

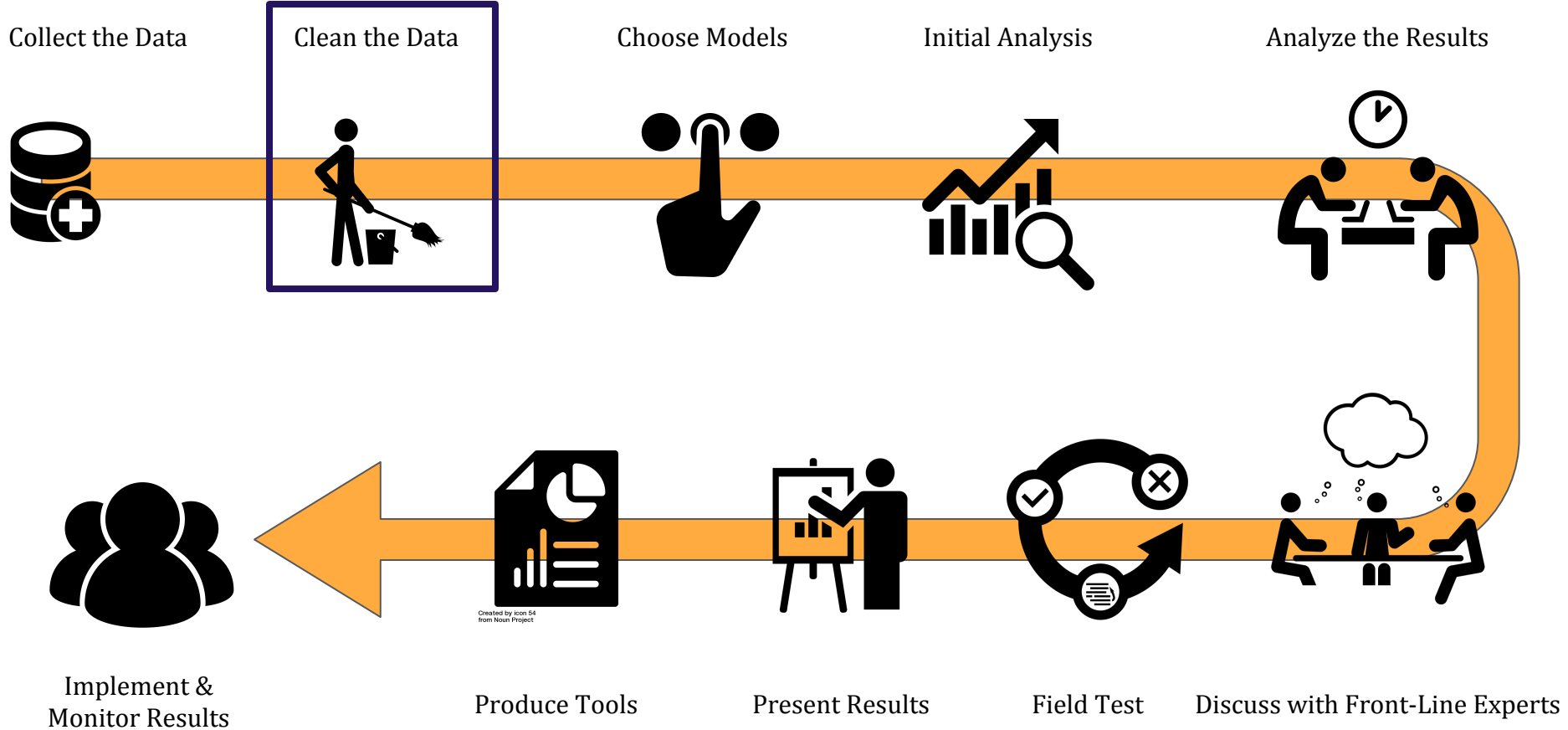
What technology was collecting and storing data?



What was used for analysis and info sharing?



The methodology



The role of people

| | |
|--------------------|--|
| The Students | The subject of the data in this case - the district has a large volume of data about each individual student |
| The Teacher | Has contextual information on students that may not show up in the data |
| The Central Office | Owns the data and tries to combine it with contextual knowledge on a scalable level across the district |
| The Superintendent | Leads efforts from the top on the importance of identifying at-risk students |
| The Partners | Outside partner to help develop the machine learning models, run the analysis, and determine model fit |
| The Policy Drivers | Determine the applicability of a variety of models and determine how it will be put in place in practice |

Summary of Key Insights and Replication Considerations

- **Data readiness is key** - collection, definitions, documentation, and cleaning are all key steps that precede any sort of modeling
- **Cities often don't control, but can influence education** - think about what programs and services that cities influence that interact with education and try to integrate them into school district efforts where possible
- **Many risk-based problems are similar** - applying a machine learning model from education may make sense in another area, so don't ignore efforts
- **Provide space for analysis** - while partners may provide much of the analysis, district and city leaders need reserve time and resources for staff to work on analytical projects

To learn more about this project

Rayid Ghani

rayid@uchicago.edu

University of Chicago - Data Science for Social Good

<http://dssg.uchicago.edu/>



THE UNIVERSITY OF
CHICAGO

**Your next
station is**

4



GovEx

What Works Cities

Bloomberg
Philanthropies



SUMMIT ON
TRANSFORMING
DATA
≡ INTO ≡
ACTION