MDST - A Model for Data Science Education through Public Service

Arya Farahi, Jonathan Stroud

The Michigan Data Science Team (MDST), University of Michigan - Ann Arbor

Abstract. The Michigan Data Science Team (MDST) is a student organization at the University of Michigan where students from all disciplines collaborate on impactful data science projects. Many MDST projects focus on public service, which allows students to exercise their skills to aid local communities. MDST's success in establishing community partnerships serves as a model for using public service projects to advance data science education.
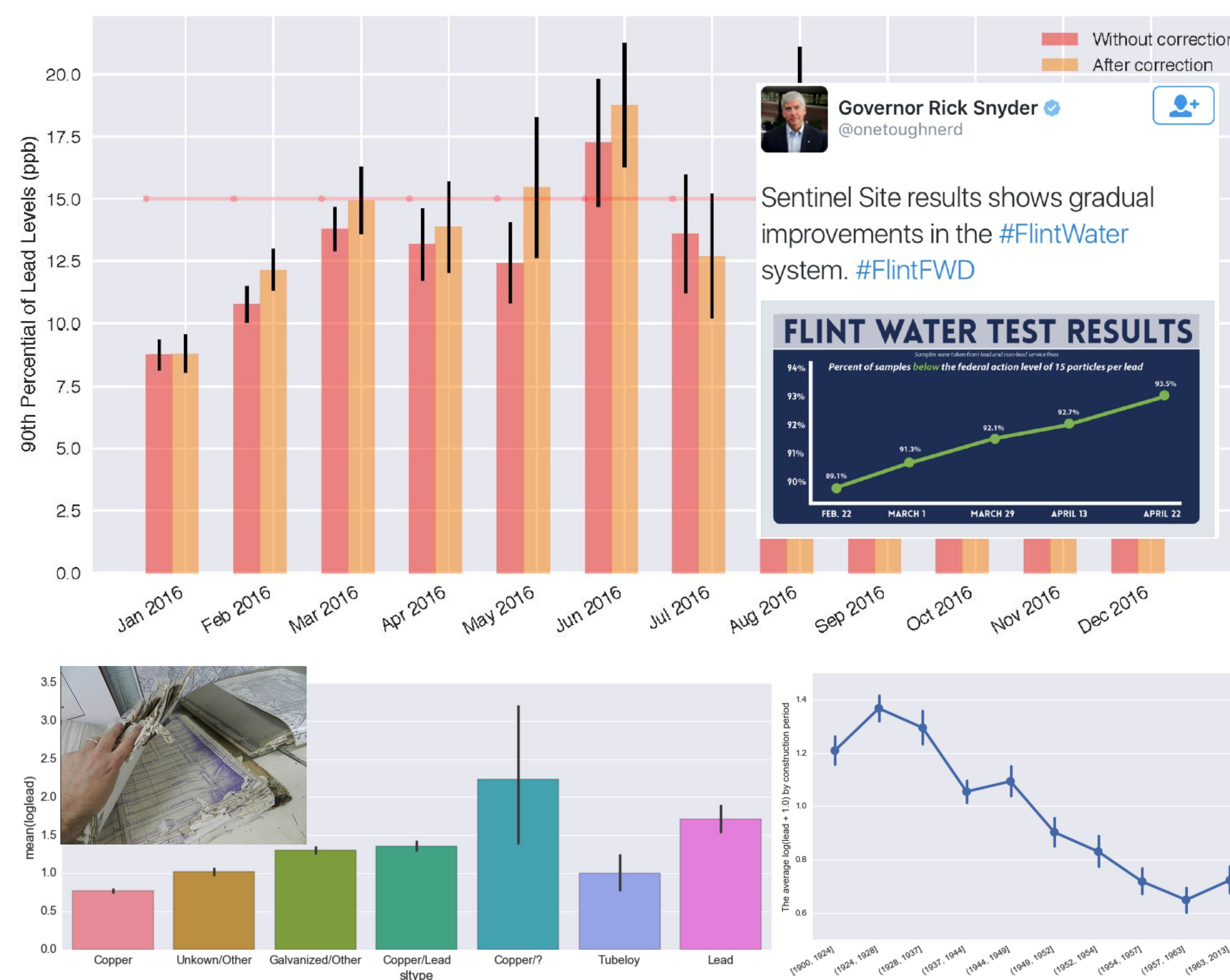
(Left) MDST members presents at the Bloomberg Data For Good Exchange, Sep. '17. (Right) MDST tutorial series, Fall 2016.

Initiatives

Tutorials
- No data science experience required
- Teach practical skills interactively
- Bridges gap between analytics and practice

Competitions
- Intra-team prediction challenges
- Focus on real-world problems
- Encourage collaborative skills

Projects
- Data science consulting for no charge
- Initiate partnerships with community
- Public service and community service focus

MDST visits the City of Flint Water Plant in April 2016,

A Successful Partnership

Flint, Michigan
- Assessed each parcel's risk for lead contamination
- Located lead service lines from historical data
- Quantified selection bias in testing program
- Proposed adaptive sampling strategies
- Crowd-sourced with prediction competition
- Released data via MyWater-Flint app

Some analyses performed by MDST student members. (Top) We quantify selection bias in lead tests performed by the EPA in Flint, Michigan. (Bottom) We find that the strongest predictive factor of lead contamination is the year that parcel was constructed, and not the service line material.

Effects

Student Growth
- Learn reproducible data analytics skills
- Build interpersonal skills
- Provide leadership experience
- Build exposure to real-world problems

Community Partnership
- Provide evidence-based recommendations
- Promote data transparency
- Enhance reuse of data and analytics
- Build relationship with local universities

(Left) MDST members at the NBA analytics Hackathon, Sep. '17. (Right) MDST presents at Neural Information Processing Systems '16.

Results

Membership
- 50+ active members
- 9 colleges at U of M represented
- Freshmen to postdoctoral scholars

Publications
- 5 accepted conference papers
- 10+ accepted posters, 4 poster awards

Partnerships
- Corporate: Quicken Loans, Baltimore Ravens
- City Government: Flint, Detroit