

Student stats analysis: Recall election a go, signatures fall short of ‘million’

by Alex Hummel - Tuesday, February 21, 2012

<http://www.uwosh.edu/today/17130/student-stats-analysis-recall-election-a-go-signatures-fall-short-of-million/>

Math is everywhere. It’s in the traffic of our morning and evening commutes. It’s in the sports we follow. It’s in the vast array of technology we use every day.

And, most assuredly, it’s involved in the day’s top headlines.

That includes Wisconsin’s ongoing recall election saga.

A statistically-valid mathematical analysis – this one, courtesy of Jennifer Szydluk’s University of Wisconsin Oshkosh “Data Exploration” students -- confirms Wisconsinites can plan on a recall election for Governor this year. However, claims that supporters of the recall of Governor Walker collected more than one million signatures are high, according to class’s results.

“Our project is not about politics; It’s about mathematics,” said Szydluk, winner of the 2010 UW System Board of Regents Teaching Excellence Award. “While the (Wisconsin) Government Accountability Board must labor over 152,335 Scott Walker recall petitions, checking every signature, University of Wisconsin Oshkosh students used statistical sampling to do the same job with a lot less effort.”

A day after the release of the submitted recall petitions were made available online to the public, Szydluk and the 30 students in her Data Exploration and Analysis course at UW Oshkosh dug into them, undertaking a mathematical sampling project to answer two questions: “How many total recall signatures were collected?” and “Are there enough signatures to trigger a recall election?”

For student and future teacher Aisha Mahmood, of Appleton, Szydluk’s project represented the kind of relevant, rubber-meets-the-road problem solving many math-averse people overlook.

“I got very excited when I first heard about our project because I believe that our life is surrounded by math,” Mahmood said. “We can find math logic in every activity we do. Therefore, it was nice to know that we would be actually working with real data, which is huge. This data is authentic and has been collected from a large population and is more exact than made up data, which people use to ballpark-figure study.”

Students dug into the recall petition project immediately and took the science of it all seriously.

“We used a random-number-generator first to select 500 petitions and second to select a random name on each sheet,” Szydluk said. “We recorded the total number of signatures on each of the 500 sampled petitions. Then we used the Government Accountability Board’s criteria for determining if each selected signature was valid on inspection, and we attempted to verify addresses using online resources (411.com,

Switchboard.com). Finally, we used statistical methods to address our questions.”

Here’s what the students found and how they arrived at their conclusion:

- With 95 percent confidence, the number of signatures is between 890,000 and 980,000. It is highly unlikely (less than a 1 percent chance) that there are more than one million signatures.
- More than 97 percent of the entries sampled by Szydlik’s students meet the GAB petition criteria.
- Addresses could be verified for approximately 75 percent of signers using online resources.
- Szydlik and students decided to count a signature as invalid in the case where either the signature violated GAB criteria or in the case where they “could not confirm that a person by the signer’s name lived the claimed municipality.” “Our aim was get an upper bound on the signatures that might be declared invalid. We found that fewer than 15 percent of the 500 signatures fall into this category and are, thus, vulnerable to challenge,” Szydlik said.
- The class determined that 756,500 (that’s $890,000 \times 0.85$) is a lower bound on the number of valid signatures.
- Thus, there will most certainly be a recall election.

Szydlik said the only caveat the class’s data analysis has is that their sampling method will not detect the prevalence of duplicate recall petition signatures.

“If this is a *wide-spread* occurrence (i.e. if tens of thousands of people signed more than once) then our conclusions will not be valid,” she said.

Szydlik said the project and the topic was a high-impact, hyper-relevant one for students. In nonpartisan fashion, it brought current headlines into the classroom and clearly helped fuel student learning.

“From this project I learned that we could apply our class knowledge into a real life situation,” Mahmood said. “When sampling, if our data is chosen randomly, then our findings will be accurate. I learned that random data sampling is the most precise way of testing a study because you are working with data which was collected from different sources and cannot be made up or manipulated.”

Szydlik said the value of the project to the class led students’ to set aside any differing political opinions so they could concentrate on their shared goal of arriving at an accurate result.

“In our class, we had both Governor Walker supporters and detractors – but we had no political arguments,” she said. “Our conversation focused on how to assure we had valid data and on why and how analysis of a random sample could provide answers to our questions. In the end, I asked the students to raise their hands if they *believed* our conclusions. All 30 hands went up.”

Read more:

- [UWO teacher-in-residence program yields positive results](#)
- [COEHS professor delivers congressional testimony on STEM program](#)