Well the election is upon us and the only good part of it is that a week after this blog goes live the whole thing will be over. The political posturing and half-baked thinking displayed by most politicians, of course, will still be there and it is possible that the outcome will not be decided for some time to come – but the endless jockeying for position will be over. Sanctimonious speeches will erupt on both sides about how voting integrity was either maintained against the fearsome onslaught of those wishing to taint the election (argument from the winning side) or how the election was actually stolen by bad actors and is now illegitimate (argument from the losing side). Curiously, neither side actually talks about the inadequacies of the election process it self and some of the severe problem and type of election poses to properly ‘reflecting the will of the people’.

In a nutshell, what makes all elections border on illegitimacy is that in all electoral scheme there demonstrably exists situations in which a candidate disfavored by a majority of people participating in the election is elected over other, more desired candidates.

Before talking about the problem in its most general form, let’s look at the case study from *Chapter 15: Election Problems and Engine Failure* from the companion Course Guidebook from Michael Starbird’s Great Courses lecture series entitled *Meaning from Data: Statistics Made Clear*. Some of the structure in the following narrative is Starbird’s but the detailed analysis and presentation is unique.

The case study starts will polling results that show the populations preferences regarding four candidates for the election. Each candidate is scored by what percentage of the population ranked that candidate as their first choice, what percentage had that candidate as their second choice, and so on. The tabular results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Candidate | 1st Place Percentage | 2nd Place Percentage | 3rd Place Percentage | 4th Place Percentage | Total Percentage |
| A | 40 | 14 | 16 | 30 | 100 |
| B | 13 | 46 | 33 | 8 | 100 |
| C | 18 | 18 | 3 | 61 | 100 |
| D | 29 | 22 | 48 | 1 | 100 |
| Total Percentage | 100 | 100 | 100 | 100 |  |

The problem of determining the will of the people boils down to the problem of figuring out how to summarize the data presented in the table to unequivocally determine who is the candidate that best represents the voting population.

Suppose that we hold an election in which all four candidates are on the ballot and voters can only vote for one. In this case, it is clear that Candidate A will be elected since he will have garnered the greatest number of votes. He would then represent the entire population even though 60% of that population would prefer someone else. This system, called [plurality voting](https://en.wikipedia.org/wiki/Plurality_voting), gives a great deal of concern since the result seems to defy the will of a overwhelming majority of the voters.

Dissatisfied with the previous results as the outcome, suppose we decide to allow voters to pick their top two choices – their first place and runner up positions. In this case, the rankings would be: Candidate B with 59% of the voters placing him in either first or second place, followed by Candidate A with 54%, Candidate D with 53%, and Candidate C with 36% of the vote. This vote-for-two method elected Candidate B who was least liked for first place to represent the population. In other words, the one thing the voters had a consensus on was that Candidate B was not the guy for first place.

Now beginning to mutter to ourselves about the stupidity of it all, we finally hit upon a scheme we are sure to work. Let’s let the voters choose their top candidate as in the plurality case but instead of keeping only one candidate we will keep the top two. Now having thinned the field, we re-offer the vote with a clearer choice. This scheme, which is variation of [run-off voting](https://en.wikipedia.org/wiki/Two-round_system), seems to us to have the best of both worlds. Putting this in place, we are happy to see that Candidates A and D are selected to proceed to the next to round. Only now do we realize that we aren’t quite sure of the next step. How will the 31% of the population not committed to either break their support to favor one or the other? Looking beyond the numbers to examine the issues, we realize that the 18% supporting Candidate C will through their support behind Candidate D because of his stand on Issue #1. We further realize that of the 13% supporting B, about 2/3 (say 7%), are going to also support Candidate D for the same reason. Candidate D wins the election with 54% of the vote and we feel good until we realize that something disturbing.

What if Candidate B, totally irrelevant in race for first place, had dropped out before the first round. Again, examining the issues, we now realize the Candidate C would have gained most of the support for B (based on their common viewpoint on Issue #1) bring his percentage up to 30%, enough to edge out Candidate D for the runoff. Now we are deeply unsettled because our system seems to hinge on presence or absence of an irrelevant candidate who splits the vote. Following that scenario further, we realize that of the 29% committed to Candidate D, who is now absent from the second round of the runoff, just over half (say 15%) will split towards Candidate A based on their support for a different issue, Issue #2. This brings Candidate A up to 55% of the vote in the second round making him the clear winner.

The fun doesn’t stop there, having now turned to looking at the issues, we realize that if Candidates D and C had pulled 12% and 11% of the support from Candidates A and B, respectively, over Issue #2 before the first round of the election, that the runoff would now involve Candidate C and D.

|  |  |  |
| --- | --- | --- |
| Candidate | Current Percentage | Hypothetical Change based on Issue #2 |
| A | 40 | 28 |
| B | 13 | 2 |
| C | 18 | 29 |
| D | 29 | 41 |

If the remaining 30% not directly of committed to either candidate, breaks two-to-one in favor of Candidate C, based on Issue #1, then Candidate C would have won even though Candidate D had done better and would have won a plurality vote even more decisively than Candidate A had in the previous scenario.

These results aren’t hypothetical but are based on an actual presidential election in the United States with Candidates A and D being Abraham Lincoln and Frederick Douglas, respectively. Perhaps we can take some comfort that the two most favored candidates were both abolitionists and so that important issue would have been carried forward regardless of who would’ve one but the course of the Civil War may have changed dramatically with Douglas in office. That said, surely we remain disturbed by the impression that it seems that we could select any candidate based on which voting system we used.

In fact, this impression has been fashioned into an actual, mathematically provable results known as [Arrow’s Impossibility Theorem](https://en.wikipedia.org/wiki/Arrow%27s_impossibility_theorem). This theorem says that it is impossible for any ranked system of voting to

* Maintain the consensus (e.g. Candidate B in the vote for two scenario)
* Reward candidates for doing better (e.g. Candidate D in the last runoff scenario lost even though he had improved his standing)
* Remain unswayed by irrelevant candidates (e.g. Candidate B dropping out of the second runoff scenario)

Mathematically, the outcome of an election has much to do with the system used and none are superior to others. Remember that next election day.