



2016/47 Lifestyle

<https://jungle.world/artikel/2016/47/daten-sprechen-nicht-fuer-sich>

Die US-Wahl hat gezeigt, wie groß der menschliche Einfluss auf Wahlprognosen sein kann

Daten sprechen nicht für sich

Von **marcel thiemann**

Die Meinungsforschung habe versagt, lauteten erste Einschätzungen nach dem für viele überraschenden Wahlsieg von Donald Trump. Ein genauere Blick zeigt: Die Umfragen lagen gar nicht so falsch wie zunächst angenommen.

Nate Silver ist so etwas wie der Star unter den Meinungsforschern, doch dieser Tage ist er auch ein Prügelknabe, der sich rechtfertigen muss. Das muss eigentlich die ganze Branche, doch Silver steht stärker im Rampenlicht als andere. Der Macher des Umfrageblogs »FiveThirtyEight« muss enttäuschten Demokraten erklären, wie es dazu kommen konnte, dass Donald Trump die Wahl gewonnen hat, obwohl fast alle Umfragen einen Sieg Hillary Clintons vorhergesagt hatten. Noch wenige Tage vor der Wahl hatte sich Silver auf Twitter eine heftige öffentliche Auseinandersetzung mit Ryan Grim geliefert, dem Umfrageredakteur der Huffington Post, dessen Modell Hillary Clinton eine Siegeswahrscheinlichkeit von 98 Prozent zusprach. Silvers Modell gab ihr lediglich 70 Prozent. Grim warf Silver vor, die Erfolgchance Trumps zu überschätzen. Später zeigte sich, dass Silver weniger falsch lag. Im Laufe der Wahlnacht änderte sich die Vorhersage der New York Times komplett. Als die Stimmen in den ersten Bundesstaaten ausgezählt waren, kreuzten sich die Kurven der Siegchancen: Die von Trump stieg steil an, Clintons fiel rasant. Am Ende der Nacht war für viele klar: Die Umfrageforscher haben versagt. Auch wenn Silver weniger weit daneben lag als andere, musste auch er sich nach der Wahl rechtfertigen. Er und viele andere Forscher verwiesen zunächst auf den Unterschied zwischen Prognosen für Wahlmänner-Stimmen und popular vote. Der Vorsprung Clintons, der von den meisten Medien als sicherer Sieg ausgelegt wurde, war die Schätzung der popular vote, also des Anteils an allen landesweit abgegebenen Stimmen. Ginge es nach diesen, hätte Clinton die Wahl gewonnen. Im Durchschnitt der Umfragen hatte Clinton am 8. November der Meinungsforschungswebsite »RealClearPolitics« zufolge einem Vorsprung von 3,2 Prozentpunkten. »Die Leute sind nicht sehr gut darin, statistische Kennzahlen zu lesen«, und auch Kommentatoren hätten die Tendenz, »Wahrscheinlichkeiten mit Gewissheiten« gleichzusetzen, sagt Silver, der eigentlich gar kein Wahlforscher ist, zumindest war er das ursprünglich nicht. Anfangs sagte er die Ergebnisse von Baseballspielen voraus. Genauer gesagt, er schätzte sie. Und das tat er so erfolgreich, dass die New York Times ihn engagierte, um das Wahlblog der Zeitung zu betreiben. 2012 sagte Silver korrekt den Wahlsieger in allen 50 Bundesstaaten voraus und wurde zum Umfragestar. »Ein bizarrer Zufallstreffer«, sagt Silver mittlerweile. Weil die Meinungsforscher das Endergebnis, eine erneute Präsidentschaft Obamas,

richtig vorhersagten, fiel nicht auf, wie groß die Fehler in ihren Umfragen waren, denn sie hatten die Unterstützung für Obama deutlich unterschätzt. Der gewann mit 3,9 Prozentpunkten Vorsprung. Die Wahlforscher hatten im Durchschnitt einen knappen Sieg mit nur 1,2 Punkten Vorsprung vorhergesagt. Noch vier Tage vor der Wahl 2016 hatte Nate Silvers Kollege Harry Enten prophezeit: »Donald Trump ist nur einen durchschnittlichen Umfragefehler von der Präsidentschaft entfernt.« Er berief sich auf die Fehlergeschichte der modernen Umfrageforschung seit 1968, die zeigt, dass die Demoskopien im Schnitt 2,0 Prozent danebenlagen. Nicht die Meinungsforschung habe bei diesen Wahlen versagt, sondern die politischen Beobachter hätten falsch gelegen, konterten auch die Macher von »RealClearPolitics«. Trotzdem, auch wenn die Prognosen nicht so schlecht waren wie allgemein angenommen, warum haben sie das Endergebnis, die letztlich entscheidenden Wahlmännerstimmen aus den einzelnen Bundesstaaten, falsch vorhergesagt? Die Wahlforschung müsse ihre Annahmen und Modelle überprüfen, sagt Silver. Es werde noch Wochen und Monate dauern, bis klar sei, was schiefgelaufen ist, schließlich laufe die Auszählung der Stimmen in Staaten wie Kalifornien und Washington immer noch. Doch erste Hinweise gibt es. 15 Prozent derer, die angaben, wählen zu wollen, waren noch drei Tage vor den Wahlen unentschlossen, 2012 waren es nur fünf Prozent. Von ihnen entschieden sich offenbar viele für Trump. Doch die meisten Umfragen wurden zwei bis drei Tage vor der Wahl abgeschlossen und bildeten deswegen womöglich die Entscheidung der lange Unentschlossenen nicht mehr ab. Bekannt ist bislang, dass die Wahlbeteiligung insgesamt zumindest nicht deutlich geringer war als 2012. Damals lag sie bei 58,6 Prozent der Wahlberechtigten, diesmal lag sie Angaben des US Elections Project der University of Florida zufolge bei 58,2 Prozent. Doch diese Zahl dürfte sich in den kommenden Wochen noch erhöhen, weil immer noch in mehreren Bundesstaaten Stimmen ausgezählt werden. Entscheidend war offenbar, dass Clinton den Wahltagsbefragungen zufolge weniger Unterstützung in wichtigen demographischen Gruppen hatte, als Umfragen vor der Wahl geschätzt hatten. Und zwar durchgehend bei Afroamerikanern, Latinos, Millenials, Frauen, Gewerkschaftsmitgliedern, weißen Arbeitern, Armen. Ein anderer Grund könnte sein, dass Wähler, die für Trump stimmten, das vorher nicht angaben. Social desirability, soziale Erwünschtheit, heißt das in der Umfrageforschung häufig beobachtete Phänomen, dass Menschen unpopuläre Meinungen nur zögerlich oder nicht äußern. Eine Studie der Beraterfirma Morning Consult vom Oktober ergab, dass zwei Prozent der Wähler nicht zugaben, für Trump stimmen zu wollen. Auch nach Angaben von David Wilkinson vom Unternehmen Cambridge Analytica, das die Trump-Kampagne beriet, war die Unterstützung für Trump in Umfragen, die online und per Roboterstimme ausgeführt wurden, drei Prozentpunkte höher als bei Telefonumfragen. Doch kurz vor der Wahl waren es nur noch ein bis zwei Punkte. Silver vermutet dagegen, dass social desirability keinen großen Einfluss hatte und die Fehler der Institute eher herkömmlicher Natur waren: die Über- beziehungsweise Unterschätzung der Wahlbeteiligung verschiedener Gruppen und der Einfluss von Last-Minute-Unentschlossenen. Andere Umfrageforscher halten dem entgegen, nach dem »Brexit«-Votum und der Wahl Trumps sei klar, dass mit historischen Daten arbeitende Wahlforschung vor allem Rechtspopulisten unterschätze. Die Daten sprächen eben nicht für sich, bemerkte das Wirtschaftsmagazin Fortune nach der Wahl. In der Umfrageforschung werden die Rohdaten gewichtet, um vermutete Verzerrungen in der Stichprobe auszugleichen. Vor allem aber zeigt sich daran der Einfluss der Forscher auf die Daten. Auf diesen weist womöglich auch ein weiteres Phänomen hin. Eine Besonderheit dieser Wahlen ist, dass viele Umfrageinstitute mit der gleichen Tendenz irrten. Um die Fehler einzelner Umfragen auszugleichen, berechnen Umfrageaggregatoren wie Silvers »FiveThirtyEight« und »RealClearPolitics« den Durchschnitt aller Umfragen. Doch wenn viele

Umfragen in die gleiche Richtung irren, funktioniert dieses System kaum. Der Effekt des herding ist aus der Gruppensoziologie bekannt und beschreibt die Tendenz von Anlegern und Journalisten, einem vorherrschenden Trend zu folgen, wenn er sich etabliert hat. Waren es Unentschlossene, Unabhängige, Drittparteiwähler und auf beiden Seiten die pikierten Parteianhänger der beiden Kandidaten mit historisch niedrigen Zustimmungswerten die sich im letzten Moment um den Kandidaten ihrer Partei scharten und mit zugekniffener Nase angaben, diesen wählen zu wollen, oder gab es auch bei den Umfrageforschern »Herdenverhalten«? Jedenfalls näherten sich fast alle Umfragen am Ende an und sagten einen geringen oder deutlicheren Vorsprung für Clinton und eine Mehrheit der Wahlmänner-Stimmen für sie voraus. Journalisten polemisierten immer gerne gegen Daten, weil sie diese als eine Bedrohung ihres Handwerks wahrnahmen. Modelle und Umfragen müssten auch die Unsicherheit der Daten abbilden beziehungsweise diese deutlich machen, auch wenn der Zeitungskonsument Sicherheit verlangt, so Silver. Aber selbstverständlich sei Unsicherheit nicht sehr verkaufsfördernd für Big-Data-Anwendungen, weder in der Umfrageforschung noch in der Wirtschaft, kommentierte ein Fortune-Journalist. Die Zahlen und Modelle würden eben eine unsichere Realität abbilden. Auf den Vorwurf, in seinem Modell zu wenig Unsicherheit zuzulassen, antwortete Silver, statistische Modelle seien vor allem journalistische Werkzeuge, um »zu verstehen, wie Wahlen funktionieren«. Es gehe nicht nur um das Ergebnis. Die Umfrageforschung sei, wie die Demokratie, »das am wenigsten schlechte System«.