OPINION

THE VIEWS EXPRESSED BY CONTRIBUTORS ARE THEIR OWN AND NOT THE VIEW OF THE MESSENGER

How Special Is Your Special Day?

Published 01/16/24 10:00 AM ET | Updated 01/16/24 11:17 AM ET

Sheldon H. Jacobson, Ph.D.

As we begin yet another year, filled with 366 unique days (2024 is a leap year), each of us has a birthday upon which we can celebrate with family and friends, marking the day that we get one year older — and perhaps even wiser.

Yet, with the population of the United States now at more than 335 million people, that translates into just over 900,000 people on average who were born on each day of the year. To put this into perspective, that is the same number as the populations of Columbus, Ohio, or Charlotte, N.C., and more than the populations of each of Wyoming, Vermont, Alaska, North Dakota and the District of Columbia.

Not all days have the same number of people who celebrate their birthdays, however. The month of September has disproportionately more birthdays than any other month. This may be due to the beginning of cooler weather in December that drives more people indoors … enough said.

Not surprisingly, February has the fewest number of people who celebrate their birthday, given that it has either 28 or 29 days, compared to 30 or 31 for the other 11 months.

Finding people with the same birthday is quite easy. Data scientists have studied this problem, aptly named the birthday problem.

How does it work?

Any given person has one of 365 possible days upon which they were born (excluding leap years). If a second person does not match the first person's birthday, they have 364 out of 365 possible days upon which they could be born. If a third person does not match the first and second persons' birthdays, they have 363 out of 365 possible days upon which they could be born.

Continuing in this manner, the likelihood that one of the people shares a common birthday grows quickly as these ratios get combined.

How quickly?

In a group of 12 randomly selected people, there is around a 16% chance that at least two of these people share a common birthday.

In a group of 23 randomly selected people, that chance grows to 50%. In a group of 32 randomly selected people, it is better than 75%. In a group of 41 randomly selected people, it is better than 90%.

When we reach a group of 57 randomly selected people, there is a better than 99% chance that at least two of these people share a common birthday.

What makes this problem thought-provoking (perhaps even head-scratching) is that in the last case, it is almost certain that a group of 57 people contains at least two people with the same birthday, even though 57 is less than one-sixth of the days in a year.

This means that anytime you attend a wedding, some social event (like a large Super Bowl party), or a business gathering with 32 or more people in attendance, there is a good chance that two or more people share the same birthday. When the size of the event reaches 57 or more people, that chance is near certain.

Of course, knowing whether you are part of such a match is difficult to predict. The likelihood that you match a particular birthday in a group of 57 people is around 1-in-7. Nonetheless, somewhere in this group are at least two people who celebrate the same special day.

What the birthday problem illustrates is that properties associated with large numbers can be deceiving, and in many cases, counterintuitive. What may seem like a rare event is just a consequence of large numbers expressing themselves.

In a world filled with large data sets, seemingly rare events occur. It is why professional sports teams go on winning (and losing streaks) over a large enough number of seasons (just ask the NBA's Detroit Pistons).

So, the next time you board an airplane, get onto a bus, attend your child’s soccer or basketball game, or enter a restaurant, know that it is highly likely that two or more people amongst you share the same birthday. You may not be part of such a match, but someone in attendance certainly is. Data dictate such a phenomenon.

Sheldon H. Jacobson, Ph.D., is a professor of computer science at the University of Illinois Urbana-Champaign. A data scientist, he applies his expertise in data-driven risk-based decision-making to evaluate and inform public policy.