

Quality Improvement Basics: Data Analysis Example

Slide 1 Describing Data: Summary Stats

Based on what was covered in the QI Basics Data Analysis module, let's review an example scenario using our fictional health care system, Solar Health.

The QI team at Solar Health wants to look at descriptive data to understand cancer rates in their patient population. They pull together a table that shows their cancer prevalence rate for each month in the most recent calendar year. Their summary statistics are:

- Mean: 10% was the average cancer prevalence rate
- Median*: 9% was the middle value for cancer prevalence rate
- Mode: 9% was the most frequently reported cancer prevalence rate
- Range: Rates ranged from the lowest at 7% up to the highest at 13%

Please note that the median will usually approximate the mean unless you have significant outliers or an abnormal distribution.

Slide 2 Diagnostic Analysis: Identifying Trends and Anomalies

To better understand how their data compares to national rates, the QI team plots their data against a national benchmark. A benchmark is “a standard or point of reference against which things may be compared or assessed.” A good benchmark will be composed of data from a population that is relatively comparable to your population regarding demographics, health status, and social characteristics. From this graph, we can see that Solar Health System's rates are below the national average.

The QI team also wanted to look at rate changes over a period of four years. This comparison shows that the average rate increased slightly from year one to year two and from year two to year three. The average rate then jumps between years three and four.

Slide 3 Stratification

As we've discussed, stratification is arranging or classifying data into smaller groups, or “strata” to identify interactions and relationships within the data.

10% of Solar Health's 70,000 patients across two hospitals and three clinics have been diagnosed with breast cancer. This is a bit lower than the national average and is within a standard range, so it doesn't cause much alarm to leadership or community members.

However,

Slide 4 Diagnostic Analysis: Stratification

- Stratifying by facility, we can see that Moon Hospital and Clinic Venus have much higher cancer prevalence rates masked by the other facilities' lower rates.
- What interactions could be happening that could cause these stark differences in rates?
- When we stratify the data by different groups or categories, we may see different trends in the data that tell a different story. Moon Hospital and Clinic Venus have alarmingly high cancer

prevalence rates masked by the other organizations' lower rates. It turns out that those two facilities serve patients who live in different zip codes than the other clinics and hospital. It may be appropriate to focus QI activities and initiatives on these two first.

- Sometimes, we may need to stratify even deeper. What are the cancer prevalence rates among the different providers at those two facilities? What about across different races, ethnicities, or age groups?

Slide 5 Diagnostic Analysis: Identify Relationships

- When stratifying by the patient income distribution at each facility, we can see that Clinic Venus and Moon Hospital have a much larger low-income patient population than the other sites.
- By stratifying even further, we can see that the cancer rates at Moon Hospital and Clinic Venus are much higher for low-income patients than those with an annual income over \$70k.
- There seems to be a correlated relationship between patient income and cancer prevalence rates.

Slide 6 In Summary

- Summary statistics help us understand the dataset's characteristics, size, and variables. The mean, median, mode, and range give a high-level overview of the data.
- Plot data over time to spot trends, patterns, and anomalies.
- Purposeful stratification compares outcomes across strata to uncover disparities or trends.