

10 QUESTIONS TO ASK ABOUT DATA

There are some basic questions anyone can ask about data. When looking through the Mason County Connecting for Kids data series, you may want to think about these questions to help you add meaning to the data.

- 1. How many people are we talking about?**
 - What is the number of people this happened to?
 - The smaller the number the higher likelihood that the associated % or rate will vary noticeably from year to year.
 - Referred to as cases, counts, or events.
 - Example the number of people arrested or the # of people graduating from high school.
- 2. How many people could we be talking about?**
 - What is the number of people this could happen to?
 - This represents the size of the population of interest.
 - Example all county residents any age, the number of people age 18+, or the number of students enrolled in a public school.
- 3. Do we have a percent, rank or rate?**
 - Number's alone do not help us compare groups of people or places, so this type of statistic is needed to help us do that.
 - Example the motor vehicle death rate or the % of population that was foreign born.
- 4. Do we have data for more than one year?**
 - Having the data for more than one year allows you to compare over time what is happening with your issue of interest.
 - Comparing across time can help you see if the issue has a trend and if there is a trend whether it up, flat, or down.
 - Example in 1990 about 5% of children lived with a relative and in 2000 about 8% did.
- 5. What happens when we compare "our" percent, rank or rate to... the state, other counties, or different age groups?**
 - Comparing two pieces of data helps you to get a feel for whether things are changing or staying the same.
 - It allows you to see where you stand compared to something – are you higher, lower, or the same.
 - This help you see if the issue has an up or down trend.
 - Example in 1999 the rate of child abuse referrals for the county was 68.9 but the state rate was 37.5.
- 6. Is this good news or bad news?**
- 7. Do we know if this was worse or better a few years ago?**
 - Is the change over time (question 4) or the difference in comparison (question 5) – positive, negative, or neither?
 - Example the county child abuse rate is higher than the state and that is negative but it is better than 5 years ago.
- 8. Having looked at the data, what do you think about it?**
- 9. If you showed the data to someone what might they say is the "explanation" for the findings?**
- 10. What statements do you feel confident making based solely on the data?**

Data serve its most valuable purpose when used as a guide, not to make decisions.

To use data and understand statistics, you do not have to be a data geek. You actually use data everyday to make sense of the world around you. For example, asking the weight of a newborn baby to tell how healthy they are, the miles per gallon a car gets before buying it, and a child's grades to determine their progress.

What is data?

Data consist of a series of facts or statements that may have been collected, stored, processed and/or manipulated but have not been organized or placed into context.

How do you use data?

When data are organized, they become information. Information can be processed and used to draw generalized conclusions and, at times, make predictions.

Project funded by
the Family Policy Council.
Derived from materials
developed by
Mary Ann O'Garro,
Thurston County Public Health
& Social Services Department.

STATISTICS USED IN THIS SERIES

Statistics are used to make sense of numbers, that is to analyze them. When looking through the Mason County Connecting for Kids data series, you may want to refer to explanations of the statistics to help you understand what the data is revealing.

(Number)

Can be a person, case, count, or event

This is important because:

- It is a key piece of information used in a percent or rate.
- It tells you if you are dealing with “small” numbers or relatively few events.

Per Year (Number Per Year)

The total of all events divided by number of years counted

Equation is... # + # + # + # + # / Total Years

In the data sheets, you will not see each year of data used to create the # per year statistic.

This is important because:

- It helps stabilize numbers that may vary from year to year.
- It helps you see the average of a set of numbers.

Rate

Describes the number of times something happened in relation to the size of the population

**The # (number) of times something happened
Divided by the total population of interest**

X

**Standard
(1,000 or 100,000)
and Adjustment**

Rates are usually developed using advanced statistics. Often the rate is multiplied by a standard like 100,000 to get the number from a decimal to a whole number (0.000155 to 15.5). Most people are more comfortable with whole numbers. Adjustments are used to even out the effects of age structure of a population and to look at only certain age groups. The effects of age structure include having a younger population (higher volume of kids) which, without adjustments, might make it appear you have more child related health issues than a neighboring county which is much older over all.

This is important because:

- It allows you to compare one group to another.
- It can help you see trends between groups, places, or years.
- It can be used to rank all counties against each other.

These data are old!

You will often notice when you look through data reports that the numbers provided are from a couple of years ago. Understanding how data is collected may be helpful.

- 1) Most data are collected at the local level and then forwarded to the state or agency that processes the data. To release data for use by the public, the state must: wait for locals to report it (which could be months after a year ends), error check it, go back to locals if problems are found, and then put it together in one or many formats to release it.
- 2) After it is released, people who use the data then have to analyze it and write up reports about it. In a really good system this could take a year, but it usually takes longer for all this to happen.

Why is some data unavailable?

Local, state and federal laws and requirements usually guide the collection of data. If the state does not ask for certain pieces of data they will not get it. Some laws restrict what can be released to the public. Some data is available but not publicized. Due to limited resources the state may not be able to provide it all to the public.

% of Total (Percent of Total)

Number of times something occurs out of 100 (per cent)

Equation is... # / Total x 100
(number divided by total multiplied by 100)

For example, to get the percent of total that were female using the table below you would use the following equation = # (female 5) / Total (20) x 100 which produces 20%. You would say 20% of the total population was female.

BREAKDOWN OF TOTAL	# YEAR	% OF TOTAL YEAR
Female	5	25%
Male	15	75%
Total	20	100%

The total in the equation is the whole picture of whatever you are looking at for a specified time period.

This is important because:

- It gives you a proportion to compare others populations to.
- It can be used for comparisons.
- It breaks down the whole, that is total population.

% Change (Percent Change)

The difference between two points in time

Equation is... # (Year 2) - # (Year 1) / Year 1 x 100
(number from current year minus number from previous year divided by previous year multiplied by 100)

% Change is a way to look at growth, that is increases or decreases in a component of the whole (total) population. Percent change gives you a feel for how fast or slow the #'s are changing across a time period.

	YEAR 1	# YEAR 2	% CHANGE YEAR 1-YEAR 2
Female	5	10	↑ 100%
Male	15	20	↑ 33%

This is important because:

- It can be used to show trends over time.
- It gives you a feel for change among components or sub-groups of a population.

Why is the data in the Mason County series different from other reports?

There are many reasons why data from one source, report, or agency may be different from others. Rarely does it mean that the statistics are wrong.

1) Different data sources may actually analyze the data differently and therefore look different. Be careful not to compare two different data sources, you'll probably be comparing apples and oranges.

2) Different people preparing the data may actually be using different definitions of the topic of interest and therefore the data looks different. For example, pregnancy rates... some people include miscarriages in the total and some do not.

3) Different people may actually be looking at different populations. One example is teen pregnancy... some people look at teens 17 or younger and some 19 or younger.

And finally, 4) some people use different statistically calculation methods. There are a few different ways to get rates and none are wrong. Data may vary (slightly) simply because the software the analyst uses calculates it a bit differently.

