Using and Understanding Data

The data user's guide:
Using data to make better decisions

Information in this document is based on *The Data Difference*. Developed by the Data Users Task Group, Oregon Department of Human Services.

Adapted for Hawaii by

Healthy Hawaii Initiative Evaluation Team
Department of Public Health Sciences
University of Hawaii

Using and Understanding Data

Table of Contents

For more information on anything contained in this document, please contact Dr. Jay Maddock at jmaddock@hawaii.edu.

Information in this document is based on *The Data Difference*. Developed by the Data Users Task Group. Oregon Department of Human Services.

Introduction

The following stories take us to a place called Kalana County. There, we'll put ourselves in the shoes of a county staff member. And we'll look in on a Start.Living.Healthy coalition.

We'll gather and analyze data, and share our work with stakeholders and the public.

Of course, there is no Kalana County. We've created the place as a means of presenting our guidance about data in useable terms.

The settings could just as easily be your own community. The principles we present here about sound collection and use of data apply wherever people seek better information to support better decisions.

A data resource directory for Hawaii

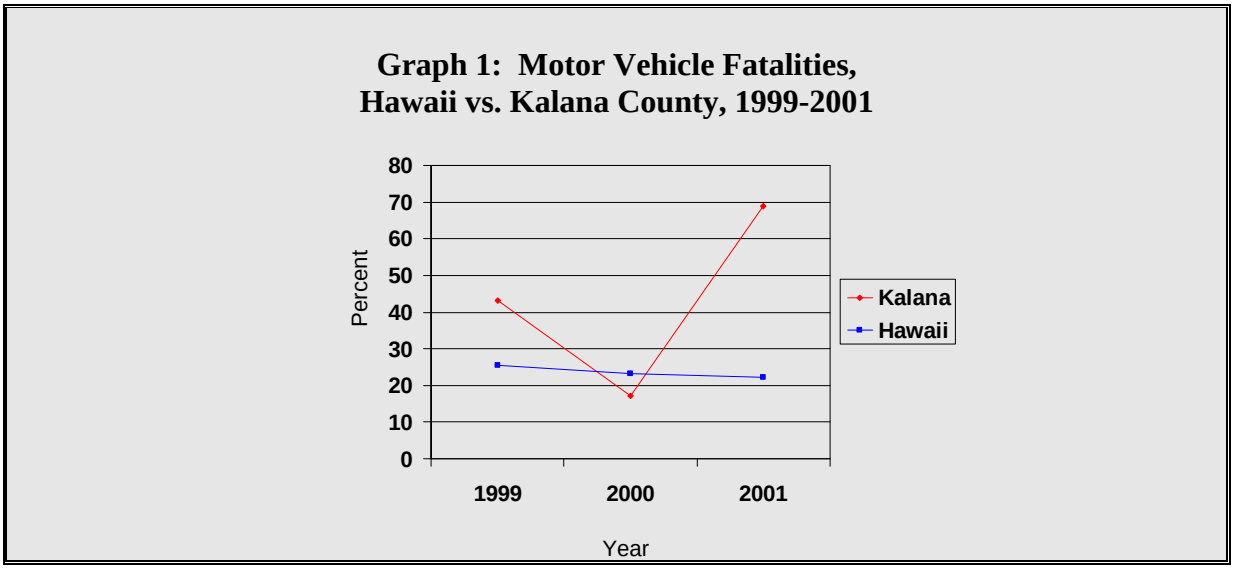
http://www.hhdw.org

The Hawaii Health Data Warehouse (HHDW) is developing online community profiles. HHDW encourages those with an interest in data to use the resource, and to provide us with feedback.

The HHDW website has several resources to aid in understanding the health data reports. They can be viewed and printed online at http://www.hhdw.org/cms/index.php?page=how-to-use-data

Motor Vehicle Fatalities in Kalana County

Calculating rates, aggregating data, and keeping track of that decimal point!



Source: *Kalana County Health Department*

Analyze small numbers by calculating RATES and AGGREGATING data – but watch out for faulty comparisons and misplaced decimal points!

Introduction

Kalana County Mayor Ho's daughter died in a car accident. The lead story in *Kalana News Bulletin* laments "motor vehicle fatalities on the rise in Kalana County this year". The newspaper suggests that the mayor use this latest family tragedy as a "wake-up call about the risks facing the county".

Accompanying the article is the above graph, showing that an astounding 69 percent of Kalana's residents died in motor vehicle accidents in 2001.

Caution! A word about graphs

The numerical scale of graphs is important. When you're looking at several charts, make sure the scales are comparable.

Start at the source

As staff to the distraught mayor, you feel the need to investigate the driving situation and set the record straight. You figure that if 69 percent of the county's residents are dying in motor vehicle accidents, it would have made national news, not just the *Bulletin.*

You call the Kalana County Health Department, which was listed as the source of the data. To find out what proportion of Kalana County residents died in motor vehicle accidents, you need

both the number of motor vehicle fatalities and the total number of residents in the county. A Health Department staffer faxes you these statistics:

Table 1: Kalana County Motor Vehicle Fatalities, 1999-2001

1999	5
2000	2
2001	8

Table 2: Kalana County Population Estimate, 1999-2001

1999	11,579
2000	11,627
2001	11,613

Source: *Kalana County Health Department*

TIP

The Health Department staffer informs you that population data like those in Table 2 come from the U.S. Census Bureau (www.census.gov).

Calculating the rate

It's time to begin analyzing your data so you can present a summary to Mayor Ho that compares the numbers with the media portrayal.

As you have learned from the Health Department staffer, motor vehicle fatalities, being very rare events, are shown as a **rate** per 100,000 residents.

To find the motor vehicle fatality (MVF) rates for Kalana County, you will need data from both Tables 1 and 2 above. A rate is calculated by taking the **frequency** (or number of MVFs) in the population you are interested in, during a specific time period, and dividing it by the total population you are interested in, within the *same* time period. The answer is then multiplied by a standard population number such as 1,000 or 100,000.

Example:

$$\frac{\text{Number of MVF in Kalana County in 2001}}{\text{Estimated county population in 2001}} \times 100{,}000$$

Or, $\frac{8}{11{,}613} \times 100{,}000$

So, 0.0006888 x 100,000 = 68.9 per 100,000

Thus, the MVF rate for Kalana County in 2001 was 68.9 per 100,000 population. In other words, if there were 100,000 people in Kalana County, 69 of them would have died in motor vehicle accidents in 2001.

Rates can be great….

Rates help equalize populations of varying sizes, allowing comparisons among different years, genders, or geographic areas. Just make sure to calculate all rates from the same population base; for example, if you are looking only at the number of MVFs for people aged 65 and older, the denominator should be the total number of people 65 and older, rather than the entire county population. In addition, make sure to multiply by the same number. For example, even though Kalana County doesn't have 100,000 people, in order to compare its MVF rate to the state's rate, both rates are multiplied by 100,000.

… but watch out for that decimal!

A rate of 68.9 MVFs per 100,000 is a far cry from the **69 percent** (or 69 per 100) that the *Bulletin* reported. How could such a mistake happen? Earlier, on the phone, the Health Department staffer hadn't been surprised.

She points out that a rate of 69 per 100,000 is **.069 percent**. Such mistakes, and the resulting mislabeling of charts or tables, are common among people not used to working with specific data.

Check to make sure your comparisons are valid

The graph in the *News Bulletin* was labeled "Motor Vehicle Fatalities". It's important to check with the source of the data to clarify the definition. In this case, the entire population was included, but in other cases only specific age and gender groups should be used. When comparing rates, you must use information from the same population base.

Caution! When numbers are very small

Counties with small populations will likely have small numbers of events such as suicides or motor vehicle fatalities. You might also come up with small numbers when you divide data into age groups or other categories.

Rates calculated from small numbers can vary considerably from year to year. Do these variations reflect meaningful change? It's a good idea to check with the original source of the data on this question.

Another problem involves confidentiality: Small numbers may allow someone reviewing your data to recognize individual cases – the 15-year-old male who died in a motor vehicle accident in a certain county in a certain neighborhood, for example.

Here's a solution

To avoid these problems, the Health Department staffer strongly suggests that you combine, or **aggregate**, the data in Table 1 above.

Example

Add the years together, for events and for population. Then, use the resulting proportion to calculate the rate.

Using Data from Tables 1 and 2 for 1999-2001:

$$\frac{\text{Sum of MVF}}{\text{Sum of population estimates}} \times 100{,}000$$

Or $$\frac{5 + 2 + 8}{11{,}579 + 11{,}627 + 11{,}613} \times 100{,}000$$

Or $$\frac{15}{34{,}819} \times 100{,}000$$

So .000431 x 100,000 = 43.1 per 100,000

Thus, the aggregate rate in this county for the three years covered in the *Bulletin* analysis is 43.1 MVF per 100,000.

The *News Bulletin* story implied that motor vehicle fatalities are a much larger problem this year than in previous years. You figure that the reporter didn't heed the Health Department's warning about small numbers and failed to aggregate the data. Such an error would appear to show a spike in the MVF rate in the county, even though the actual change was from two to eight.

(See the next section, *Heart Disease in Kalana County – Part 1*, for information on calculating percent change)

Some things to keep in mind

When you asked for the most current data, the Health Department staffer gave you 2001 data with the caution that it is "preliminary". This means that some numbers may change. Therefore, caution should be used in interpreting the data.

A data snapshot is just one way to look at what is going on in a particular county. There are many other important perspectives to consider in piecing together an accurate composite picture. For instance, it may be useful to note that although residents often are involved in MVF, tourists can also contribute to the numbers.

TIP

To learn about other data sources, visit the Hawaii Department of Health's Web site:

http://hawaii.gov/health

Conclusion:

Pleased with your finding, the Mayor asks you to draft a response for her to send to the *Bulletin.* The response includes some key points about the newspaper's presentation and use of data. Among them:

- How to calculate and when to use **rates**, which help to standardize data so that comparisons can be made
- When the numbers are very small, the importance of **aggregating** multiple years of data to stabilize rates and protect the confidentiality of individual "events".
- When using graphics, the importance of clear, accurate, **labels.** Data that are misused or inaccurate can lead the reader to the wrong conclusion.

Heart Disease in Kalana County – Part 1

Proper use of percentages can help you sort out the facts

Do anecdotal experiences – our own impressions about the way things are – paint an accurate picture? Proper use of numbers, percentages and rates can help you sort out the facts, and steer you away from faulty conclusions.

Introduction:

Kalana County Mayor Ho has asked you to investigate heart disease in the county.
Last week, while in Honolulu to testify before the Legislature, she heard Kalana's Senator Kim give a floor speech on the issue.

Senator Kim said that Kalana County experienced a 300 percent increase in heart disease last year.

Those statistics restate Mayor Ho's personal experience. During her five years teaching math in the local middle school, she saw 13 heart attacks among teachers – one the first year, two each in the second and third years, three the fourth year, and five during her final year.

In the cases she saw during the last few years, poor lifestyle was a major factor. Also in these cases, the victim was highly stressed. This suggests to her that heart disease is on the rise in Kalana because of unhealthy lifestyles and increasing stress.

But as a math teacher, she knows that her 13 cases may not be representative. She needs some quantitative data. To begin, she's asked you to answer these questions:

- How many myocardial infarctions (MI; heart attacks) has Kalana had in recent years?
- How many of these resulted in death (MI fatalities)?
- How do these figures compare statewide?
- What has been the percent change each year for number of MIs and MI fatalities, both statewide and in Kalana?
- Has the county really seen a 300 percent increase in the number of MIs in the last year?

Some information from the Department of Health starts you on your way.

TIP

Make sure you begin with appropriate, comparable data. In particular, when working with percent change over time, it's essential that you use the same time period for all calculations.

How to calculate percent change

$$\frac{\text{(Most Recent Number – Previous Number)}}{\text{Previous Number}} \times 100$$

To find the percent change in statewide MI reports from 1995, when there were 6753, through 2001, when there were 6937:

$$\frac{6937 - 6753}{6753} \times 100$$

So, $\frac{184}{6753} \times 100 = 2.7\%$

Thus, you can say that the number of heart attacks statewide increased 2.7 percent from 1995-2001.

The following table on the number of myocardial infarctions in Kalana County will help you calculate a similar percent change locally.

Table 1: Myocardial Infarctions – Kalana County

Year	# of MIs	Yrly % Change of MIs	# of MI fatalities	Yrly % Change of MI fatalities	Rate: Fatalities per 100 Cases
1995	625		210		33.6
1996	688	10.1%	297	41.4%	43.2
1997	677	-1.6%	257	-13.5%	38.0
1998	633	-6.5%	235	-8.6%	37.1
1999	645	1.9%	244	3.8%	37.8
2000	599	-7.1%	205	-16.0%	34.2
2001	632	5.5%	206	0.5%	32.6

Source: *Kalana County Health Department*

Plugging the 1995 and 2001 figures into our formula, you calculate that the number of MIs in Kalana County increased 1.1 percent during the period (from 625 to 632), a better track record

than the state's. The county's figures on myocardial infarction fatalities show a decrease. From 1995-2001, the number of MI fatalities in Kalana *dropped* 1.9 percent (from 210 to 206).

So far, you haven't encountered the 300 percent figure that Senator Kim cited. But it's there, as you see when you look at the numbers on heart transplant fatalities.

Caution! Percent change figures aren't always useful

You find that statewide, deaths from heart transplants dropped 50 percent this year, after increasing 100 percent the year before. And fatalities from heart transplants in Kalana County did rise an incredible 300 percent! But are these useful statistics?

Table 2: Heart Transplant Fatalities

	Hawaii		Kalana	
Year	**Fatalities**	**% Change**	**Fatalities**	**% Change**
1999	10		1	
2000	20	100.00%	1	0.00%
2001	10	-50.00%	4	300.00%

Source: *Kalana County Health Department*

Small numbers like these sometimes experience large, but meaningless, fluctuations. That's exactly the case with the 300 percent figure. See the motor vehicle fatalities in Kalana County scenario for more about this problem and its solutions.

Conclusion

You've produced some good information about heart attacks in Kalana County, and about how the county compares statewide. You've also put the 300 percent increase in heart disease into perspective for Mayor Ho. Now, you're ready to take on her questions about lifestyle factors.

Heart Disease in Kalana County – Part 2

Using percents to break down totals into categories

Now that we've calculated percent changes in heart attack data, we'll use percentages another way: to break down totals into categories. That will help us understand heart attacks in our county. But in the end, we'll also learn something about data's limitations.

Sorting heart attack data by age of victim

Mayor Ho also used to volunteer at a senior center and saw many victims of heart disease. All were over age 65. How does this compare to the actual age distribution of victims statewide and in Kalana? As shown in Table 1, adults over age 65 constitute well under half of all victims – about 33 percent statewide and 35 percent in Kalana.

Table 1: Age of Victims – Statewide and Kalana, 2001

	Hawaii		Kalana	
Age	**# of Victims**	**% of Victims**	**# of Victims**	**% of Victims**
< 45	910	26.0%	49	23.8%
45-64	1428	40.8%	85	41.3%
65+	1162	33.2%	72	34.9%

Source: *Kalana County Health Department*

Who is having heart attacks? Calculating proportions

Mayor Ho's impression from her own experience is that Filipinos are the most likely to have heart attacks. The data, however, show something else.

Table 2 shows the numbers of MIs by ethnic group. Unfortunately, the tables you received from the Department of Health (DOH) State Office don't report any additional ethnic groups, such as Samoan, African-American or mixed ethnicity. You make a note of this in your report to Mayor Ho.

Table 2: Myocardial Infarction by Ethnic Group, Kalana County

Ethnic Group	1997	1998	1999	2000	2001
Caucasian	182	176	180	162	175
Filipino	115	112	114	108	118
Japanese	80	67	65	60	73
Chinese	64	65	70	66	60
Hawaiian/Part Hawaiian	155	150	144	148	144
Other	81	63	72	55	62
Total	**677**	**633**	**645**	**599**	**632**

Source: *Kalana County Health Department*

The numbers show that Caucasians and Native Hawaiians have the highest occurrence of heart attacks. Proportions can help you analyze the data, by showing how much of a total is attributable to one category.

Example:

To calculate proportions, use this formula:

$$\frac{\text{The count in one category}}{\text{Total count in all categories}} \times 100 = \text{the one category's proportion of the total}$$

So, using Table 2 to calculate the percentage of instances in which the Caucasians were the victim in 2001:

$$\frac{175}{632} \times 100 \qquad \text{Or,} \quad 0.277 \times 100 = 27.7\%$$

Thus, you can say that Caucasians accounted for about 28 percent of the county's heart disease cases in 2001. Using the same technique, you then calculate the proportion for Filipinos: about 19 percent. The figures are similar statewide.

> *You can add the raw numbers or percents of several categories to come up with new categories. For example, adding the proportions for Caucasians and Filipinos above, you can tell Mayor Ho that these groups accounted for 47 percent of heart disease in Kalana County in 2001.*

Food for thought

You show your data to Mayor Ho and she replies, "so Caucasians are the most likely to get heart disease." Well, not really. In this case, Caucasians accounted for the largest percentage in the county simply because there are more of them!

You need to determine the **incidence rate** *within each ethnic group* to know which group was the most likely to have heart disease. Incidence rate is typically used to describe the number of new cases that develop in a year in a specified population. You can use the following formula to figure it out:

$$\frac{\text{Number of new cases in the specific population}}{\text{Total number of people in the specific population}} = \%\text{ of new cases within the specific population}$$

So, for Caucasians:

$$\frac{\text{175 MIs in Caucasians}}{\text{3,418 total Caucasians}} = 0.0511 \times 100 = 5.1\%$$

But for the Chinese:

$$\frac{\text{60 MIs in Chinese}}{\text{723 total Chinese}} = 0.0829 \times 100 = 8.3\%$$

Be careful with the bottom number, or the **denominator**, of this equation. The specific population number in your table may represent only women or only men. Make sure that the population represented in the top number, or **numerator**, is the same population as the denominator. Again, remember you must use information from the same population base.

Calculations of the MI incidence rate for each ethnic group are shown in Table 3.

Table 3: Ethnic Groups and Heart Disease in Kalana County, 2001

Ethnic Group	Specific Population	MI	% of population
Caucasian	3,418	175	5.1%
Filipino	2,064	118	5.7%
Japanese	1,391	73	5.2%
Chinese	723	60	8.3%
Hawaiian/Part Hawaiian	3,245	144	4.4%
Other	772	62	8.0%
Total	**11,613**	**632**	**5.4%**

When you calculate the percent of each ethnic group's population that had an MI (i.e. the incidence rate of MI for each ethnic group), the numbers change dramatically. These data show that Chinese and people in the "other" category are much more likely to have a myocardial infarction. Caucasians and Hawaiians are less likely to have an MI, but they make up a greater percentage of the MI cases because they constitute the largest segments of the population.

What happens when you add a new category?

The following table gives data on incidence of MIs in Kalana County.

Table 4: Incidence of Myocardial Infarctions, Kalana County

Ethnic Group	1997	1998	1999	2000	2001
Caucasian	182	176	125	122	125
Filipino	115	112	77	73	88
Japanese	80	67	62	55	61
Chinese	64	65	70	60	60
Hawaiian/Part Hawaiian	155	150	144	148	144
Other	81	63	42	28	32
Subtotal	**677**	**633**	**520**	**486**	**510**
Mixed non-Hawaiian	-	-	125	113	122
Total	**677**	**633**	**645**	**599**	**632**

Source: *Kalana County Health Department*

Mayor Ho asks you to report the top three ethnic groups who experienced MIs in each of the past five years. While reviewing the data, you notice you are missing numbers in the 1997 and 1998 columns for "mixed non-Hawaiian".

You contact the DOH analyst you had been working with to get the missing pieces and he tells you that the data do not exist; DOH only began reporting that data in 1999. Working with numbers you have, you find that in Kalana, the top three ethnic groups to experience heart attacks in 1997 and 1998 are Caucasian, Hawaiian and Filipino. However, since "mixed, non-Hawaiian" was added in 1999, it has remained one of the top ethnic groups experiencing MIs in Kalana.

Caution! Understanding the effect of a new category

When a new category is added, it's always important to understand how it affects the bigger picture. In this case, a drop in the number of MIs in the Caucasian and Filipino categories is seen as soon as the category of "mixed, non-Hawaiian" is introduced. This is probably not caused by decreased rates of MI in Caucasians and Filipinos, but instead by people changing their self-identified ethnicity from Caucasian or Filipino to "mixed, non-Hawaiian" after that category became available.

What about the role of lifestyle factors?

Your interview with the DOH analyst reveals that DOH's computerized data do not answer this question. Available data do not identify specific causes for heart attack.

But the numbers do show, statewide, that "obesity" was the condition most often reported along with MI in 2001. Thirty-eight percent of the cases had noted this secondary diagnosis. Fortunately, however, the percent of overweight people dropped 3 percent from its high of 35 percent.

Conclusion – and a word about limits

You present your findings to the Mayor Ho. Generally, she's pleased. Your research serves as a reminder that stories and personal experiences, while they can add depth and richness to data, also can be misleading.

However, she's troubled by the lack of information about other lifestyle factors and stress.

There's an important lesson here: Properly used, data can help you paint an accurate picture. But the data may not fill in every detail. It's important to recognize this, and to proceed carefully where data are incomplete.

Are there blanks in your data? Other information sources you should explore? Means of generating new data to meet your needs? People with statistical training can help you sort out these questions.

Lifestyle Factors in Kalana County
Comparing and presenting data to describe local conditions

Once you've gathered data from a variety of sources, create graphs to tell your story. Show trends with line and bar graphs.

Introduction

The Kalana Start.Living.Healthy Coalition, an advocacy group focusing on positive outcomes for chronic disease, seeks to improve health conditions for residents of Kalana County.

Recent headlines about the "alarming amount of heart attacks" got the Coalition's attention. Coalition members also know that newly elected Mayor Ho has initiated an investigation of heart disease in the county.

The Coalition believes that tobacco use, physical activity and nutrition account for the majority of heart disease cases. The Coalition plans to issue a position paper showing that such a relationship exists and hopes that using quantitative data will be more compelling. They plan to publish the paper on their Web site so community partners can easily access the information.

TIP

Know who your audience is and your purpose.

Showing a trend

First, the Coalition needs to get an overall picture of health behaviors – to show how Kalana County compares to Hawaii and the nation. They also want to determine whether the situation is improving or getting worse over time.

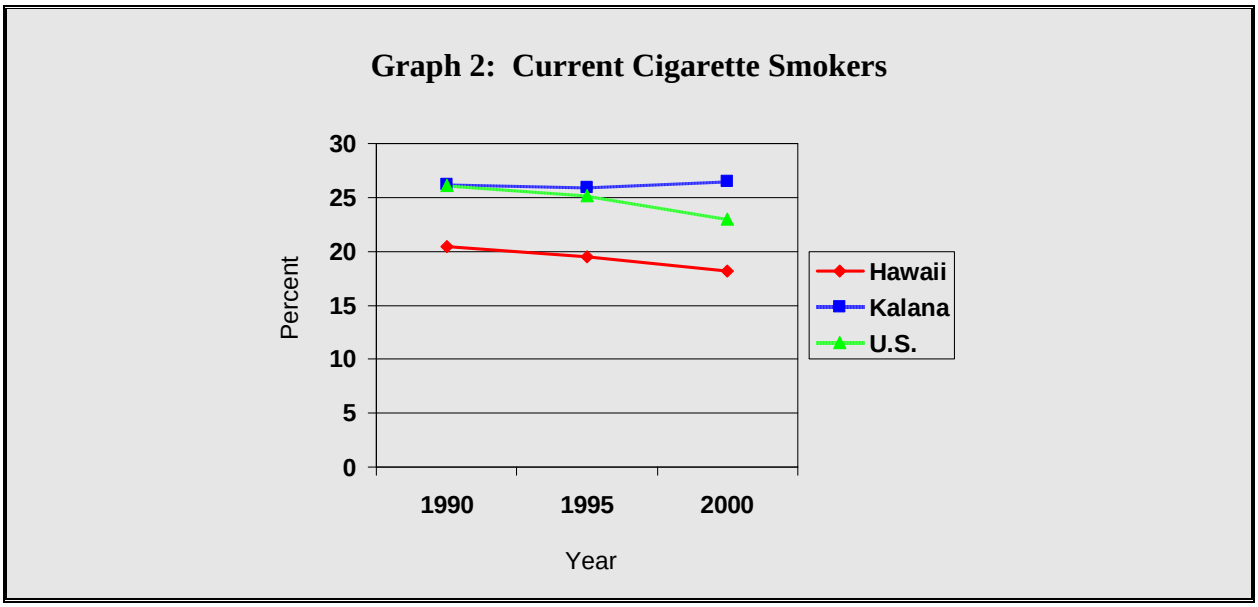
TIP

It's usually the best to compare a local community to the whole state, instead of to another community.

Hawaii has been seen as a healthy state since 1990, consistently ranking in the top ten states for life expectancy. While rural areas such as Kalana have grown, their growth has not kept pace with the state as a whole.

Several factors contribute to a healthy lifestyle, including policies, environmental influences, social norms, and individual skills and beliefs. Understanding the current health behaviors in Kalana County will help the Coalition develop its strategic plan for the coming years.

Cigarette smoking is the most preventable cause of death in the nation. It accounts for over 430,000 deaths annually. Although smoking is decreasing in the nation, the same is not true for all communities.

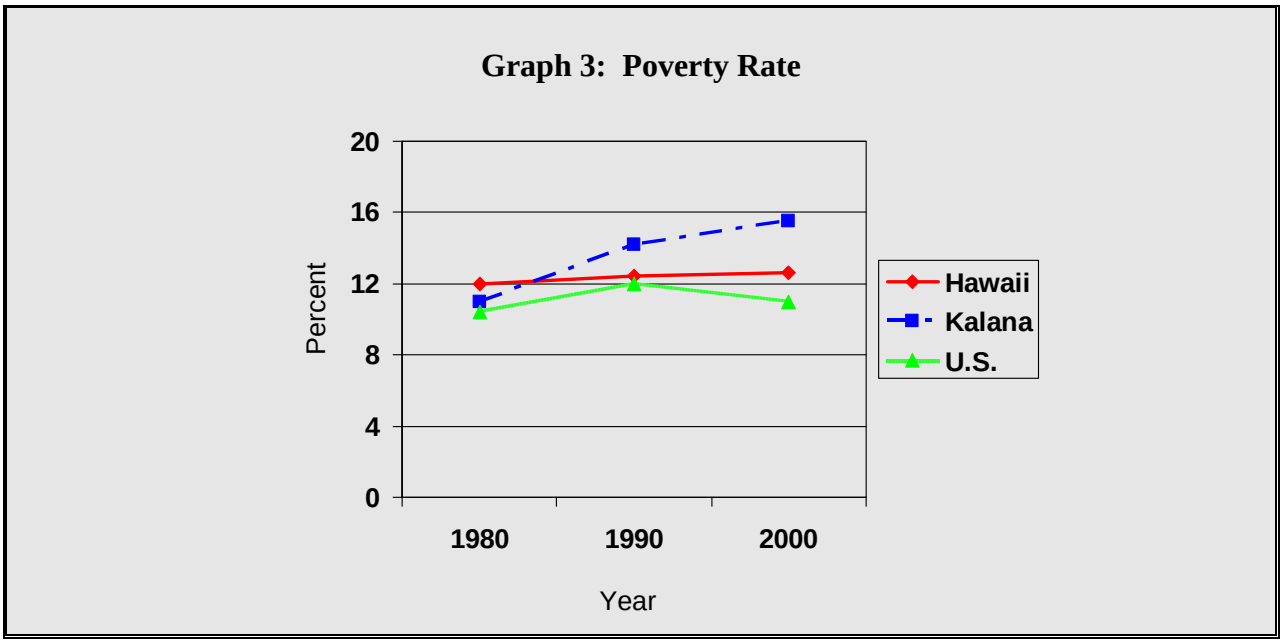


Source: *Kalana Department of Health*

Kalana is a rural community with a non-diversified economic base that has been primarily dependent on the sugar industry. The closing sugar plants, a poor economy, and a recent influx of immigrants from Southeast Asian countries, where smoking rates are high, have all contributed to a smoking rate that is not declining with the rest of the state.

In many ways, Kalana's health behavior indicators have been faring poorly when compared to the state and the nation. Some key measures:

- Kalana's 2000 tobacco use rate was 46 percent higher than the state's (26.5% vs. 18.2%).
- Kalana's 2000 obesity rate was 39 percent higher than the state's (22.8% vs. 16.4%).
- In 2000, 58.2 percent of Kalana's residents got no leisure time physical activity, compared to 49.4 percent for Hawaii as a whole.
- In 1980, Kalana's poverty rate was only 3.7 percent higher than the state rate, but by 2000 it was 43 percent higher.



Source: *Hawaii Department of Business, Economic Development and Tourism*

TIP

Note the sources of your data

A great source of economic data in Hawaii is the Department of Business, Economic Development and Tourism. Each year, they publish a data book with facts and figures about Hawaii. This data book is available to the public for free at http://hawaii.gov/dbedt/.

The Youth Risk Behavior Survey has lots of information on the behavioral risk factors of youth in Hawaii. It can be found on the web at: http://www.cdc.gov/yrbs/

Tapping other data resources

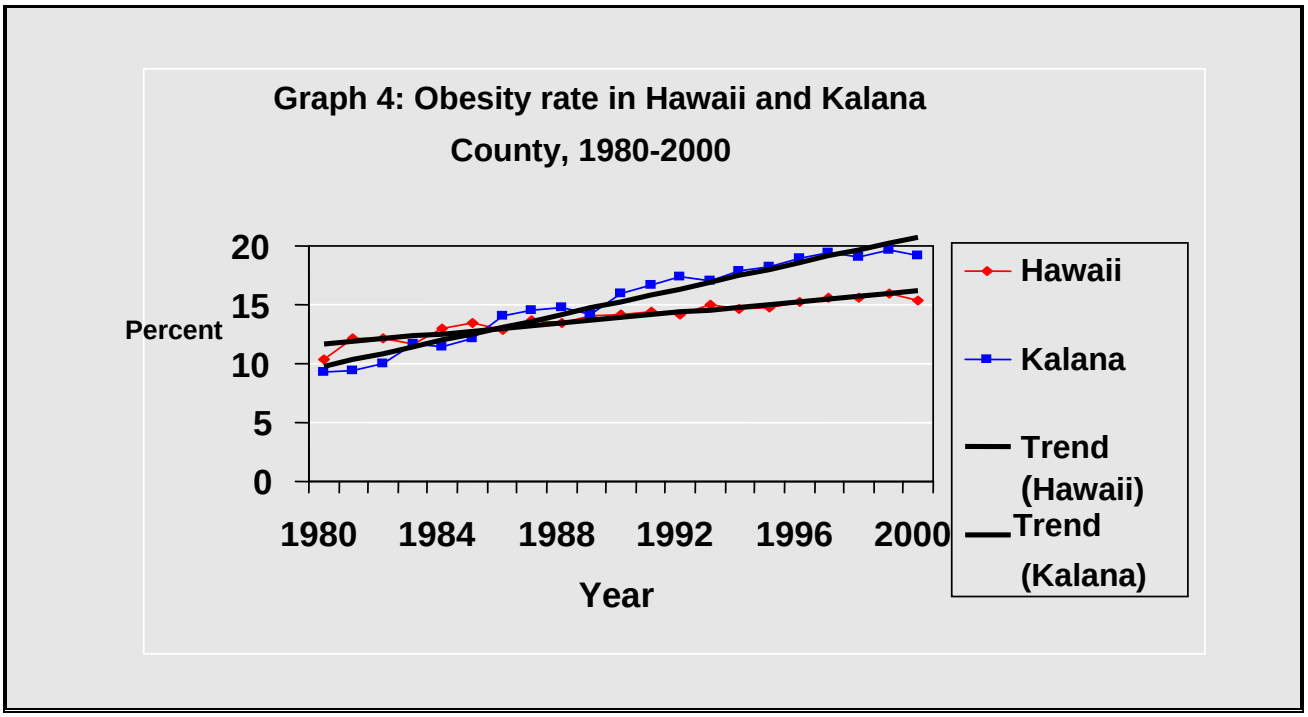
These tobacco and economic data clearly show a widening gap between Kalana and the rest of Hawaii. The Coalition wants to find other social demographic data that demonstrate the impact of the poor economy on the county.

The Coalition goes to the University of Hawaii's Center on the Family Web site at: http://uhfamily.hawaii.edu/index.asp to find more about childhood obesity for Kalana County and Hawaii.

Next, the Coalition goes to the Department of Health's website to get data on physical activity. The DOH takes twenty years worth of data and puts it into a line graph. This graph will allow the Coalition to see any trends for Kalana County and how those trends relate to the state as a whole.

TIP

Bar and line graphs are effective ways to show trends.



Source: *Hawaii Department of Health*

TIP

Use graphic images whenever possible

Kalana County obesity rate is increasing at a faster rate than the rest of the state.

TIP

Find other data that are relevant and meaningful to the issue.

The Coalition knows the value of finding various sources of data. They are aware of a web site that shows an inventory of data resources that are available. The URL is: **http://www.hhdw.org**

The Coalition learns about the Community Health Profiles published by the Hawaii Department of Health and the Hawaii Health Data Warehouse (HHDW). In 2002 the HHDW culled data from dozens of different data sources across the state. From this survey, a "profile" report is produced for each county.

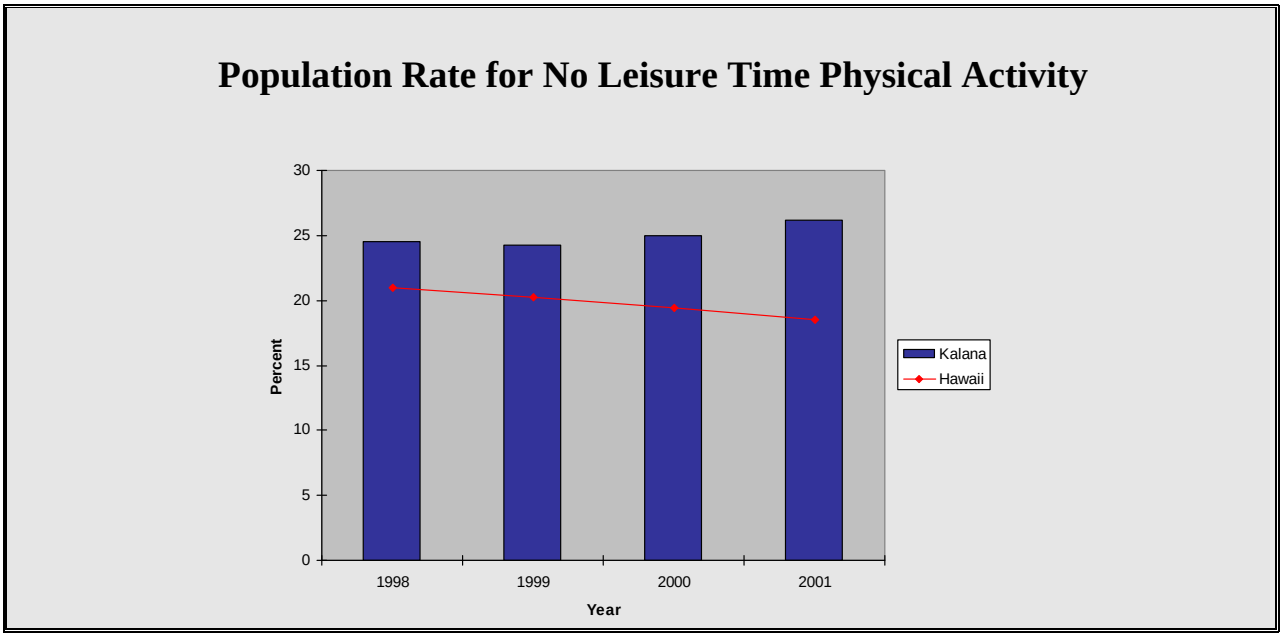
The profiles help community leaders and prevention planners with community planning and collaboration and help to give a more accurate picture of the factors influencing the lives of their

residents. They aid policymakers in directing resources to programs and strategies most likely to promote healthy lifestyles.

Coalition members receive a profile for their county that supports the data they have collected; Kalana County is at a higher risk for obesity and tobacco use than the state overall, possible explanations are:

- Poverty Rate
- Sedentary population

Data from the Health Department show that while the percentage of state residents who do not exercise has decreased, the percentage of Kalana residents who are inactive has stayed the same, slightly increasing in the last year.



***Source**: Kalana County Health Department*

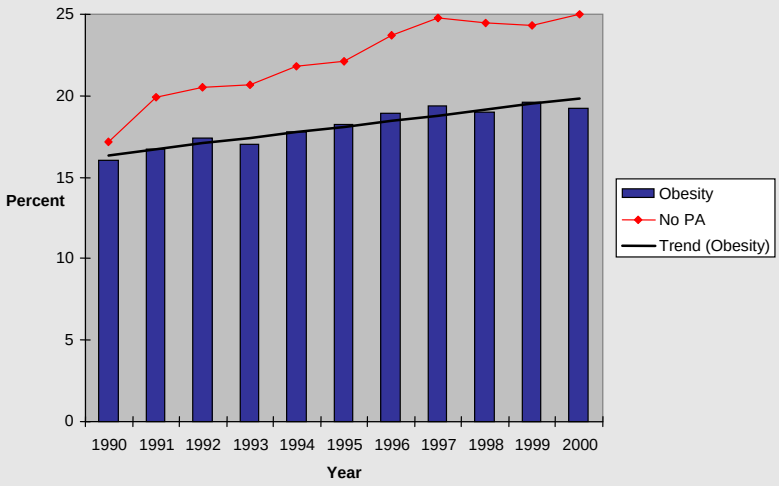
A look at social factors

Kalana's obesity rate has consistently been higher than Hawaii's since 1988. Could an increasingly sedentary population be contributing to this trend? Four data points aren't enough to show any kind of trend, so it's difficult for the Coalition to come to any definite conclusions.

However, they do see that Kalana residents seem to be at greater risk of inactivity than the state overall.

The Coalition compares the past 10 years of no leisure time physical activity to the obesity rate, to assess if the two are related.

Obesity and no leisure time physical activity rates, Kalana County, 1990-2000



Source: *HHDW Community Health Profiles*

Conclusion:

Over the past 10 years, the percent of the population who get no leisure time physical activity is rising with the percent of people who are obese. Although this does not mean that physical inactivity caused the increase in obesity, it is a possible contributing factor.

The Coalition believes that they have examined and used relevant data to tell the whole story. Coalition members feel confident that their position paper will help decision makers to develop services and policies to help the county's residents.

The Coalition will argue in its position paper that the county needs to invest in increasing physical activity, decreasing tobacco use and bringing jobs to the county to improve health. A range of social programs could reduce the risk of poor lifestyle behaviors, and help to improve Kalana residents' wellness.

Using Data – a checklist

How to select data

- Use the most current and complete data you can find.
- Go to the original source for the correct definition of the data.
- Are the data relevant and meaningful to the issue?
- Select data at the most appropriate level (city, ZIP, county, state, etc.).
- Do we have enough information to tell the whole story?

How to do basic calculations

- Proportion (percentage, ratio, fractions, etc.)
- Rate
- Percentage change over time
- Mean, median, mode, frequency distribution

How to compare data

- It's usually best to compare local data to the state.
- Use tables and graphs to show differences.
- Be cautious when comparing data. Are the definitions the same? Were the data collected the same way?
- Follow the same data point through time.

How to analyze and interpret data

- Create a framework that ties all the data together.
- Be careful when taking numbers out of context.
- Have definitions, data collection methods or time frames changed?
- Have we examined and used all relevant data available to tell the full story?
- Be careful when using mean, median, and mode.
- Don't read more into your data than is actually there.

How to address small numbers

- You may need to look at a larger geographic area.
- Combine data from multiple time frames (aggregate).
- You may be able to find related data that have larger numbers.
- Be aware of confidentiality issues.

How to package and present data

- Avoid relying on a single piece of data. Gather data elements and present data in a "bundle" of related information.
- Know your audience and your purpose.
- Use graphics whenever possible.
- Note the sources of data.
- Explain the framework that you used to tie the data together.
- Identify a contact person for additional information/questions.

A Data Glossary

Benchmark – a measure that is tracked against a goal.

Demographics – General characteristics of a population (i.e. age, gender, race, occupation, income, address).

Extrapolate – To infer by projecting or extending known information.

Frequency – The number of occurrences or events during a specified time period for a given population number.

Generalize – The ability to say something accurate about the target population from information collected from a sample or subset of the population.

Interpolate – To infer missing information within the bounds of known information.

Mean (statistical) – the sum of a set of quantities, divided by the number of those quantities. Commonly called "average". For example, the mean of the following set of quantities is 15.8.

$2 + 2 + 7 + 13 + 18 + 20 + 20 + 20 + 40 = 142$

$142 \div 9 = 15.8$

Median (statistical) – The value right in the middle of a series of ranked quantities. For example, 18 is the median in the following series of quantities, because it falls in the middle, with four values above it, and four values below.

2, 2, 7, 13, 18, 20, 20, 20, 40

Metadata – A description of data. Data about data.

Mode (statistical) – The most common value in a series of quantities. For example, 20 is the mode in the following series of quantities.

2, 2, 7, 13, 18, 20, 20, 20, 40

Outcomes – The measurable changes in people, organizations, or community conditions.

Outputs – The quantity of work, activities, services, or other countable things or events that are produced by individual efforts, programs, or service systems.

Percentage – A number that represents the relation of one part to the whole.

Percent Change –

$$\frac{\text{(Most recent number – previous number)}}{\text{Previous Number}} \times 100 = \text{percentage change}$$

Example: To find the percentage change of Hawaii's population from 1990-1998

1990 population = 1,173,321
1998 population = 1,267,550

$$\frac{(1{,}267{,}550 - 1{,}173{,}321)}{1{,}173{,}321} \times 100 = 8\%$$

Proportion – The relation of one part to the whole, usually expressed as a percentage, ratio, or fraction.

Qualitative – Of, relating to, or expressed in relative or subjective terms impossible to precisely quantify.

Quantitative – Of, relating to, or expressed in terms of quantity.

Rate – Measures the frequency of an occurrence of an event in a population in a specified period of time for a standardized population number.

$$\frac{\text{Occurrence} \times 1{,}000 \text{ or } 100{,}000}{\text{Population}} = \text{rate}$$

Scale – A standard by which something can be measured or compared.

Standard Deviation – A statistic that shows the spread or dispersion of scores in a distribution of scores (i.e., a measure of dispersion). The more widely the scores are spread out, the greater the standard deviation. It quantifies how much the values vary from each other.

Statistical Significance – The likelihood that the difference found between groups could have occurred by chance alone. Normally a result is statistically significant if the difference between groups could have occurred by chance alone in less than 1 time in 20. This is expressed as a p value $< .05$.

Survey – A method of gathering information about a specific target population or group of people by sampling a subset of that population.

Trend – A general tendency or direction of movement.