The Data Difference

The data user's guide: Using data for better decisions

Prepared by

The Data Users Task Group
Oregon Department of Human Services

DHS
oregon department
of human services

The Data Difference

Table of Contents

Introduction

The following stories take us to a place called *Cascadia County.* There, we'll put ourselves in the shoes of a county staff member. And we'll look in on a local human services coalition.

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

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

The setting 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 Oregon

http://dhsdirector.hr.state.or.us/data/data_resources.htm

The Data Users Task Group is developing an online directory of data resources in human services and other areas. The group encourages those with an interest in data to use the list, and to offer additions.

The Data Difference: Using Data for Better Decisions can be viewed and printed online at **http://dhsdirector.hr.state.or.us/data/dataguide.pdf**. For more information, or to request paper copies, call **(503) 945-6012**.

January 2001

Dear colleague:

We can't fully understand our world, or make sound policy decisions, without good data. But gathering, interpreting and presenting data accurately can be difficult. This booklet, produced by the Data Users Task Group, can help.

The task group, consisting of state and local data experts and community leaders, promotes better uses of data in making human services decisions.

In preparing this booklet, we engaged scores of data enthusiasts in sorting out the key issues. We then created our own Oregon county, Cascadia, to present those issues in readable fashion.

Whatever your role — whether you're a policy maker, planner or citizen advocate, for example — we hope you'll find our advice valuable.

Our focus here is on the mechanics. But we recognize that what's really at stake is sound use of data for continuous improvement, for accurately measuring results, and for determining how we can best make a difference.

The Data Users Task Group appreciates your interest in using data for better decision-making. We look forward to your feedback about this booklet, and to hearing about your data successes.
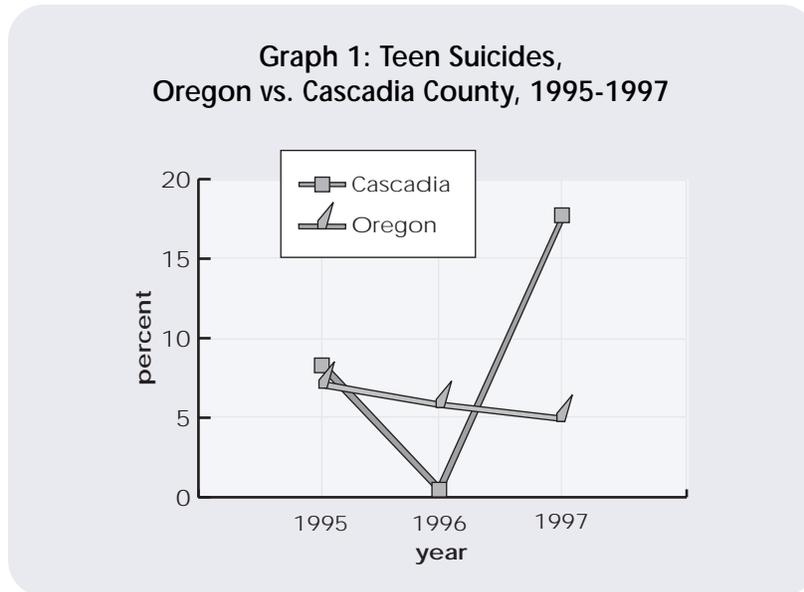
Sincerely,

Lennie Bjornsen, Administrator

Community Partnership Team
Oregon Department of Human Services
Chair, Data Users Task Group

Suicides in Cascadia County
Calculating rates, aggregating data, and keeping track of that decimal point!

**Graph 1: Teen Suicides,
Oregon vs. Cascadia County, 1995-1997**

Source: Cascadia County Health Department

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

Introduction

Cascadia County Commissioner Snow's niece has committed suicide. The lead story in the *Cascadia News Bulletin* laments the "alarming number of teen suicides plaguing Cascadia County this year." The newspaper suggests that the commissioners use this latest tragedy as a "wake-up call about the risks facing the county's teens."

Accompanying the article is the above graph, showing that an astounding 17 percent of Cascadia's teens committed suicide in 1997.

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 commissioner, you feel the need to investigate the teen suicide situation and set the record straight. You figure that if 17 percent of the county's teens are committing suicide, it would have made national news, not just the *Bulletin.*

You call the Cascadia County Health Department, which was listed as the source of the data. To find out what proportion of Cascadia County teens died by their own hand, you need both the number of teen suicides and the total number of teens in the county. A Health Department staffer faxes you these statistics:

Table 1: Cascadia County Teen Suicides, Ages 10-17, 1995-1997

1995	1
1996	0
1997	2

Table 2: Cascadia County Population Estimate, Ages 10-17, 1995-1997

1995	11,579
1996	11,627
1997	11,613

Source: Cascadia County Health Department

TIP

The Health Department staffer informs you that population data like those in Table 2 come from the Center for Population Research and Census, at Portland State University (http://www.upa.pdx.edu/CPRC). The "Pop Center" produces the official population estimates for Oregon's counties and incorporated cities for non-Census years. The U.S. Bureau of the Census counts the population every decade.

Calculating a rate

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

As you have learned from the Health Department staffer, teen suicides, being such rare events, are shown as a **rate** per 100,000 teens.

To find the teen suicide rate for Cascadia County, you will need data from both Tables 1 and 2 above. A rate is calculated by taking the **frequency** (or number of suicides) 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:

$$\frac{\text{Number of suicides (ages 10-17) in Cascadia County in 1997}}{\text{Estimated county population (ages 10-17) in 1997}} \times 100,000$$

or, $\dfrac{2}{11,613} \times 100,000$

so, $0.0001722 \times 100,000 = 17.2$ per $100,000$

Thus, the suicide rate in this age group for Cascadia County in 1997 was 17.2 per 100,000 population. In other words, if there were 100,000 people in Cascadia County in this age range, 17 of them would have taken their own lives in 1997.

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 age base, and to multiply by the same number. Therefore, even though Cascadia County doesn't *have* 100,000 people, in order to compare its suicide rate to the state's rate, for example, both rates are multiplied by 100,000.

. . . but watch out for that decimal!

A rate of 17 suicides per 100,000 is a far cry from the **17 percent** (or 17 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 17 per 100,000 is **.017 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 "Teen Suicides." It's important to check with the source of the data to clarify the definition of "teen." In this case, "teen" means 10 to 17 years old. When comparing rates, you must use information from the same population base.

Caution! When the numbers are very small

Counties with small populations will likely have small numbers of events such as suicides or teen pregnancies. You may 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 committed suicide in a certain county in a certain month, for example.

Here's a solution

To avoid these problems, the Health Department staffer strongly suggests that you combine, or **aggregate**, the age group 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 1995-1997:

$$\frac{\text{Sum of suicides (ages 10-17)}}{\text{Sum of population estimates (ages 10-17)}} \times 100{,}000$$

$$\text{or} \quad \frac{1 + 0 + 2}{11{,}579 + 11{,}627 + 11{,}613} \times 100{,}000$$

$$\text{or} \quad \frac{3}{34{,}819} \times 100{,}000$$

so   .000086 x 100,000 = 8.6 per 100,000

Thus, the aggregate rate in this age group for the three years covered in the *Bulletin* analysis is 8.6 suicides per 100,000.

The *News Bulletin* story implied that teen suicides 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 teen suicide death rate in the county, even though the actual change was from zero to two.

(See *Child Abuse in Cascadia 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 1997 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 young Oregonians are most likely to have suicidal thoughts, the elderly are most likely to actually commit suicide.

TIP

To learn about other data sources, visit the Oregon Department of Human Services Data Resource Directory web site:

http://dhsdirector.hr.state.or.us/data/data_resources.htm

For more information about suicide in Oregon, consult the Oregon Plan for Youth Suicide Prevention:

http://www.ohd.hr.state.or.us/ipe/2000plan/

Conclusion:

Pleased with your findings, the commissioner 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.** In this case, for example, that would include the age range for "teens." Data that are misused or inaccurate can lead the reader to the wrong conclusion.

Child abuse in Cascadia 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:

Cascadia County Commissioner Snow has asked you to investigate child abuse in the county. Last week, while in Salem to testify before the Legislature, she heard Cascadia's Senator Stretch give a floor speech on the issue.

Senator Stretch said that Cascadia County experienced a 300 percent increase in the number of abuse and neglect cases last year.

Those statistics restate Commissioner Snow's personal experience. During her five years teaching math in the local middle school, she saw 13 cases — one the first year, two each 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, methamphetamine was a major factor. Also in these cases, the mother's boyfriend was the perpetrator. This suggests to her that child abuse is on the rise in Cascadia because of meth, that teens are typically the victim, and that boyfriends are the most common offenders.

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 child protective service (CPS) reports has Cascadia had in recent years?

- How many of these reports were founded?

- How do these figures compare statewide?

- What has been the percent change each year for reports and founded cases both statewide and in Cascadia?

- Has the county really seen a 300 percent increase in the number of abuse and neglect cases in the last year?

Some information from the Department of Human Services' State Office for Services to Children and Families 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

Use the following formula:

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

To find the percent change in statewide child abuse reports from 1992, when there were 21,822 reports, through 1998, when there were 31,456 reports:

$$\frac{31,456 - 21,822}{21,822} \times 100$$

so, $\frac{9,634}{21,822} \times 100 = 44.15\%$

Thus, you can say that child abuse reports statewide increased 44.15 percent from 1992-1998.

The following table on the number of CPS reports in Cascadia County will help you calculate a similar percent change locally.

Table 1: Abuse Reports - Cascadia

Year	# of CPS Reports	Yrly % Change of CPS Reports	# of Founded Reports	Yrly % Change of Founded Reports	Rate: Founded Reports per 1000 Reports
1992	1002		224		223.6
1993	935	-6.69%	174	-22.32%	186.1
1994	1309	40.00%	257	47.70%	196.3
1995	1224	-6.49%	235	-8.56%	192.0
1996	1199	-2.04%	244	3.83%	203.5
1997	1158	-3.42%	230	-5.74%	198.6
1998	1105	-4.58%	217	-5.65%	196.4

Source: Cascadia County Health Department

Plugging the 1992 and 1998 figures into our formula, you calculate that child abuse reports in Cascadia County increased 10.28 percent during the period (from 1,002 to 1,105), a better track record than the state's. The county's figures on founded reports also compare favorably. From 1992-1998, the number of founded reports in Cascadia dropped 3.13 percent, while the statewide figure rose 2.7 percent.

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

Caution! Percent change figures aren't always useful

You find that statewide, deaths from child abuse or neglect dropped 50 percent this year, after increasing 100 percent the year before. And child abuse **fatalities** in Cascadia County did rise an incredible 300 percent! But are these useful statistics?

Table 2: Abuse and Neglect Fatalities

Year	Oregon		Cascadia	
	Fatalities	% Change	Fatalities	% Change
1996	17		1	
1997	34	100.00%	1	0.00%
1998	17	-50.00%	4	300.00%

Source: Cascadia 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 Teen Suicides in Cascadia County scenario for more about this problem and its solutions.

Conclusion

You've produced some good information about child abuse in Cascadia County, and about how the county compares statewide. You've also put the 300 percent increase in fatalities into perspective for Commissioner Snow. Now, you're ready to take on her questions about perpetrators, types of abuse and the role of methamphetamine.

Child abuse in Cascadia County – Part 2
Using percents to break down totals into categories

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

Sorting child abuse data by age of victim

As a middle school teacher, Commissioner Snow has encountered child abuse victims in her work. All were over age 11. How does this compare to the actual age distribution of victims statewide and in Cascadia? As shown in Table 1, children over age 11 constitute well under half of all victims – about 17 percent statewide and 21 percent in Cascadia.

Table 1: Age of Victims – Statewide and Cascadia

Age	Oregon		Cascadia	
	# of Victims	% of Victims	# of Victims	% of Victims
0-5	4,818	47.5%	137	40.4%
6-11	3,593	35.4%	130	38.4%
12-17	1,736	17.1%	72	21.2%

Source: Cascadia County Health Department

Who is abusing Cascadia's children? Calculating proportions

Commissioner Snow's impression from her own experience is that boyfriends are the most common offenders. The data, however, show something else.

Table 2 shows the numbers of alleged perpetrators of child abuse and neglect by category. Unfortunately, the tables you received from the Department of Human Services' (DHS) State Office for Services to Children and Families (SCF) don't report any additional categories, such as significant other, boyfriend, or girlfriend. You make a note of this in your report to Commissioner Snow.

Table 2: Alleged Perpetrator of Founded Abuse, Cascadia County

Alleged Perpetrator	1994	1995	1996	1997	1998
Mother	108	101	102	98	91
Father	77	66	66	60	55
Sibling	3	2	2	2	2
Other relative	44	35	39	37	37
Friend	8	9	17	12	15
Neighbor	3	7	5	2	2
Caregiver	10	8	5	7	4
Other	5	6	7	12	10
Total	**258**	**234**	**243**	**230**	**216**

Source: Cascadia County Health Department

Even without data on boyfriends, the numbers show that parents have been the two most prevalent alleged perpetrators of child abuse. 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 mother was the alleged perpetrator in 1998:

$$\frac{91}{217} \times 100$$

or, 0.419 x 100 = 41.9%

Thus, you can say that mothers were the alleged abusers in about 42 percent of the county's child abuse cases in 1998. Using the same technique, you then calculate the proportion for fathers: about 25 percent. The figures are similar statewide.

Food for thought

The data indicate that mothers are much more likely to be perpetrators of child abuse and neglect than fathers. Is this finding surprising? Is it meaningful? Who provides most of the child care? Who are usually the parents in single-parent households?

Types of maltreatment: What happens when you add a new category?

The following table gives data on neglect and on various forms of child abuse.

Table 3: Type of Maltreatment, Cascadia County

Type of Maltreatment	1994	1995	1996	1997	1998
Physical abuse	82	74	74	68	60
Neglect	72	63	60	57	54
Mental injury	26	20	9	9	8
Fatality	5	3	1	1	4
Abandonment	2	4	2	2	2
Sexual abuse/ exploitation	77	75	50	47	44
Subtotal	**264**	**239**	**196**	**184**	**172**
Threat of harm	--	--	113	106	100
Total	**264**	**239**	**309**	**290**	**272**

Source: Cascadia County Health Department

Reviewing the data so you can report to Commissioner Snow on the top three types of abuse in each of the past five years, you notice you are missing data in the 1994 and 1995 columns for "threat of harm."

You contact the SCF analyst you had been working with to get the missing pieces and he tells you that the data do not exist; SCF only began reporting that data in 1996. Working with the numbers you have, you find that in Cascadia, the top three types of abuse in 1994 and 1995 are physical abuse, neglect and sexual abuse. However, since "threat of harm" was added in 1996, it has remained the most common type of abuse in Cascadia.

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, the "threat of harm" category appears to be "pulling" numbers from the other categories; the totals in the other categories are basically on a downward trend as the new category is increasing.

What about the role of methamphetamine?

Your interview with the SCF analyst reveals that SCF's computerized data can't answer this question. Available data don't identify specific drugs.

But the numbers do show, statewide, that "suspected drug/alcohol abuse" was the most prevalent family stress indicator in 1998. Thirty-eight percent of the founded referrals had noted this indicator. Fortunately, however, the percent of founded referrals with alcohol and drug issues dropped 3 percent from its high of 41 percent in 1997.

Conclusion – and a word about limits

You present your findings to Commissioner Snow. 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 boyfriends as perpetrators, and about the influence of methamphetamine on these troubled families.

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, including the contributors listed at the beginning of this booklet, can help you sort out these questions.

Economic conditions in Cascadia 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 charts.

Introduction

The Cascadia Coalition for Healthy Families, an advocacy group focusing on positive outcomes for families and children, seeks to improve economic conditions for residents of Cascadia County.

Recent headlines about the "alarming number of teen suicides" got the Coalition's attention. Coalition members also know that newly elected Commissioner Snow has initiated an investigation of child abuse in the county.

The Coalition believes that the poor economy contributes to suicide, child abuse and other social problems. 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 the economic situation — to show how Cascadia County compares to Oregon and the nation. They also want to determine whether the situation is improving or getting worse over time.

TIP

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

Oregon's economy has been strong since 1990, consistently ranking in the top ten states for job growth. While rural areas such as Cascadia have grown, their growth has not kept pace with the state as a whole

Several factors drive economic growth, including population growth, personal income growth, employment growth, and employment composition. Cascadia's population has grown at approximately one-third the state's rate — a 4.2 percent increase (or percent

change) from 1990 to 1996 as compared to 11.9 percent for all of Oregon. Although personal income and total payrolls have increased, they have not kept pace with state and national trends.

The continuing decline in economic wellness in Cascadia can be seen in the following graphs. From 1986 to 1996, Cascadia's population fell behind in the race for economic wellbeing.

Two standard measures of an area's economic strength — Per Capita Personal Income (PCPI) and Annual Average Wage (AAW) — demonstrate this trend. While Oregon enjoyed increases in income and wages between 1986 and 1996 of 65.8 percent and 47.6 percent — slightly higher than the national increases of 60 percent and 45 percent, Cascadia's increases were only 47 percent and 31 percent respectively.

The gap between Cascadia and the rest of Oregon widened from 1986 to 1996 for both personal income and average wages.

Graph 1: Annual Average Wage (AAW)
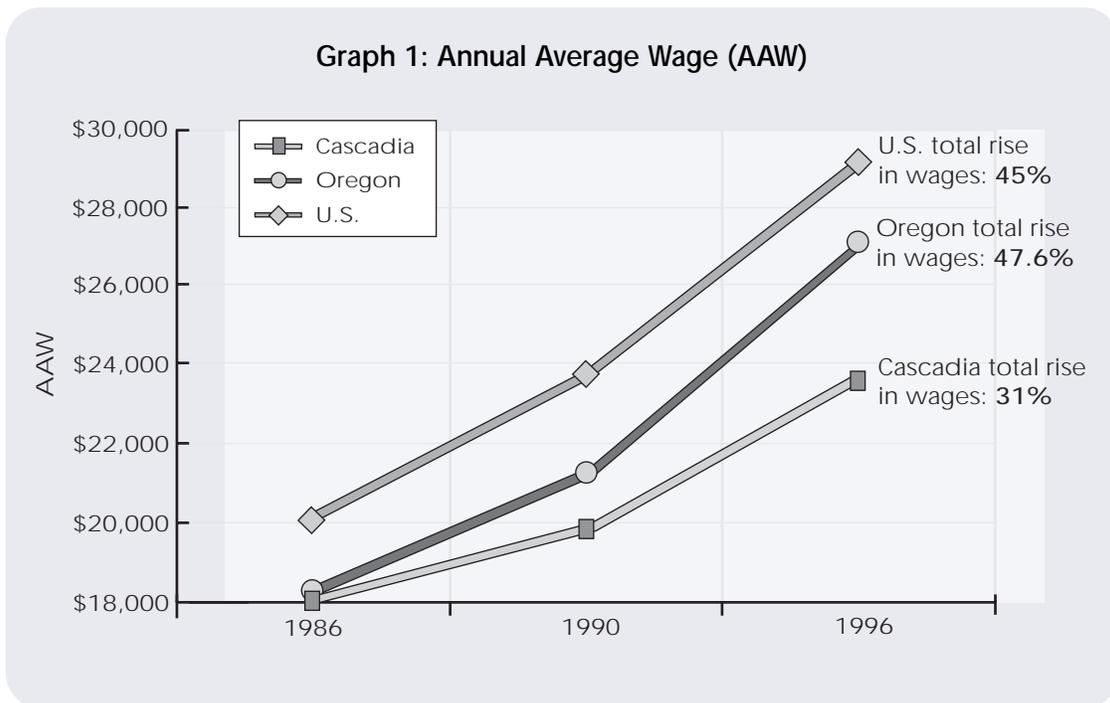
Source: Oregon Economic Department's Regional Economic Profile

Cascadia is a rural community with a non-diversified economic base that has been primarily dependent on the wood products industry. Falling timber sales, automation and environmental factors contribute to the county's poor performance. New jobs created are mostly service jobs at or slightly above minimum wage.

In many different ways, Cascadia's economy has been faring poorly when compared to the state and the nation. Some key measures:

- Cascadia's 1998 unemployment rate was 75 percent higher than the state (9.8% vs. 5.6%).

- Cascadia's 1996 household average adjusted gross income was 75 percent of the state average ($26,100 vs. $34,800).

- In 1996, 28.2 percent of Cascadia's households fell into the very low income category, compared to 19.4 percent for Oregon as a whole.

- In 1979, Cascadia's poverty rate was only 3.7 percent more than the state rate, but by 1996 it was 43 percent higher.

Graph 2: Poverty Rate

	Cascadia
	Oregon
	U.S.
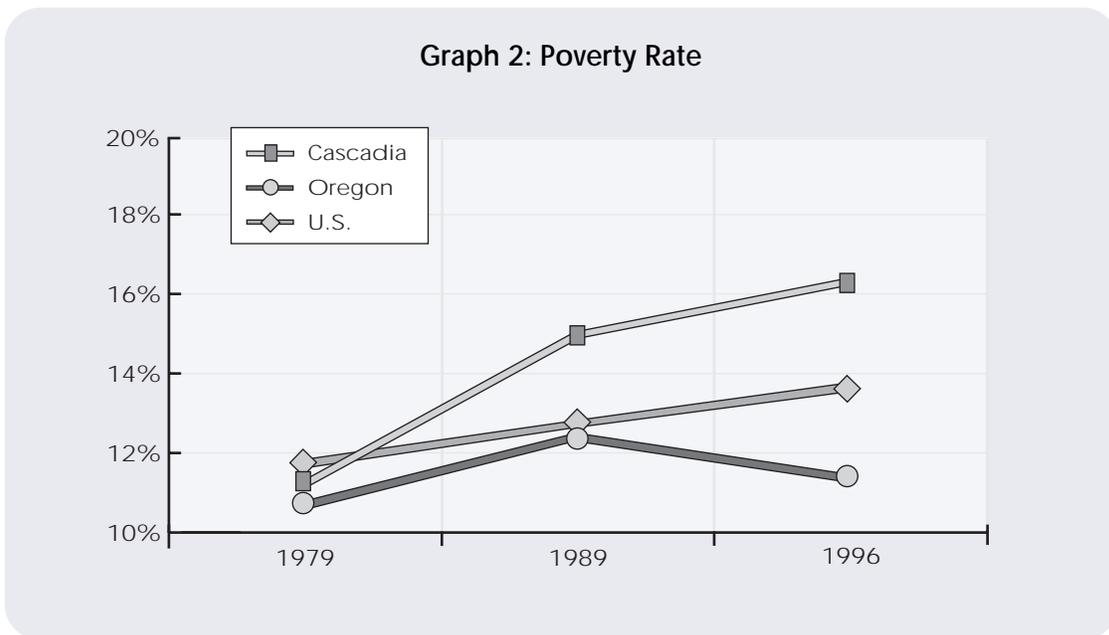
Sources: Resident Oregon Labor Force and Unemployment by Area — Oregon Employment Department; Oregon Personal Income Tax Statistics — Tax Year 1996, Oregon Department of Revenue; Historical Poverty Tables, U.S. Census Bureau

TIP

Note the source(s) of your data.

Tapping other data resources

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

The Coalition goes to the Adult and Family Services (AFS) Web site at: http://www.afs.hr.state.or.us/reports.html to find Temporary Assistance for Needy Families (TANF) and food stamp rates for Cascadia County and Oregon. They take two years worth of data and put it into a line graph. This graph will allow the Coalition to see any trends for Cascadia County and how those trends relate to the state as a whole.

TIP

Bar and line graphs are effective ways to show trends.

In April 1999, the TANF rate was at its highest point in over six months. Since then, the rate has been on a gradual downward trend, fairly consistent with the statewide trend, although at a higher rate than the state overall.
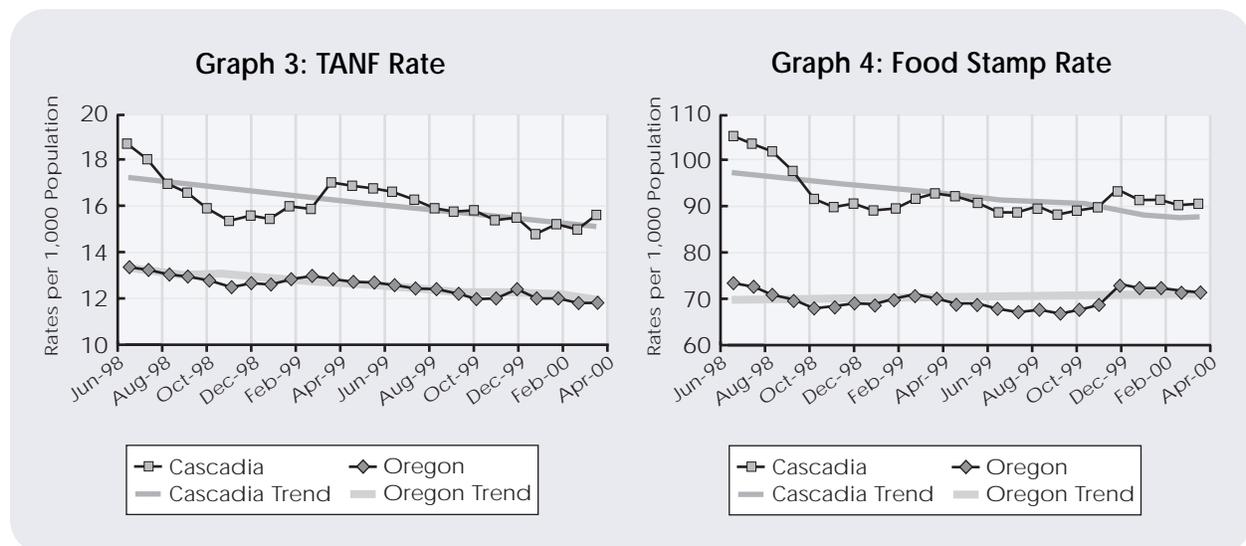
Source: May 2000 AFS Branch and District Data Book online at: http://www.afs.hr.state.or.us/reports.html.

TIP

Use graphic images whenever possible.

Cascadia County's food stamp trend has stayed fairly consistent with Oregon's. While the county's rate is higher, the downward trend is more pronounced than Oregon overall.

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://dhsdirector.hr.state.or.us/data/data_resources.htm*

While looking through the list of resources, the Coalition learns about the County Profiles published by the Office of Alcohol and Drug Abuse Programs (OADAP). Every other year, OADAP conducts the Public School Drug Use Survey. From this survey, a "profile" 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 young people. They aid policymakers in directing resources to programs and strategies most likely to promote healthy development of children and families.

Coalition members receive a profile for their county that supports the economic data they have collected; Cascadia County is at a higher risk of economic and social deprivation than the state overall, being higher in all these areas:

- Unemployment rates

- Temporary Assistance for Needy Families (TANF) caseload

- Families eligible for food stamps

Graph 5: Suicide Rate
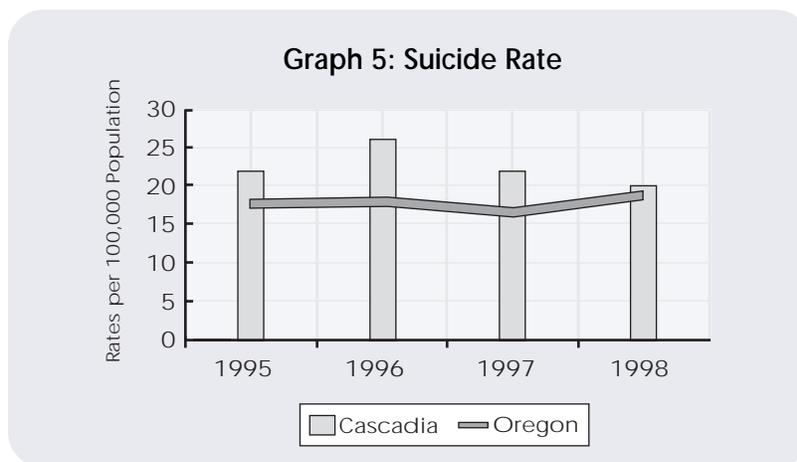
Source: Cascadia County Health Department

A look at social factors

Cascadia's suicide rate has consistently been higher than Oregon's in the past, especially in 1996. 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 Cascadia residents seem to be at greater risk than the state overall.

The Coalition compiles child abuse/neglect rates for the past 10 years and sees that with a few exceptions, Cascadia's rates are higher than the state overall.

Graph 6: Child Abuse/Neglect Rate
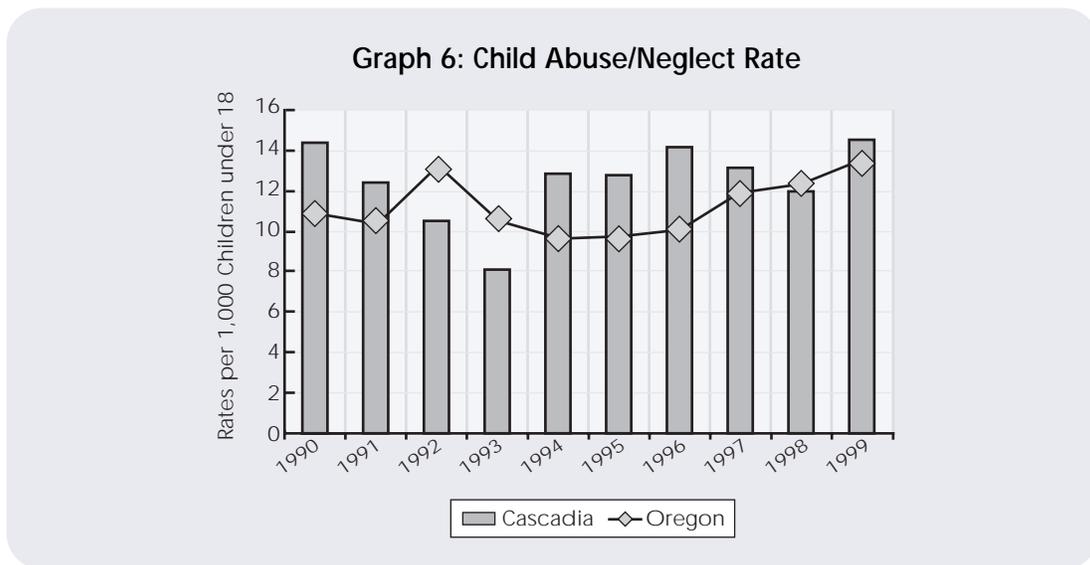
Source: State Office for Services to Children and Families

Conclusion:

Despite Cascadia County's TANF and food stamp rate drop, dependence on public assistance is considerably higher than the state overall. Cascadians are more disadvantaged than Oregonians in general. Per capita income is lower; unemployment and poverty rates are higher. Suicide rates are higher, and with just a few exceptions, Cascadia has generally had higher rates of child abuse/neglect over the past 10 years.

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 to help the county's residents.

The Coalition will argue in its position paper that the county needs to invest in training, education, and other services that can build the skills of county residents who lost jobs in the changing economy. A range of social programs (mental health and alcohol and other drug treatment) could reduce the risk of child abuse and suicide, and help to improve Cascadia 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?

- Are we confident we have examined and used all relevant data that's available to tell the full story?

- Be careful when using mean, median, 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 (Oregon Benchmark) - a measure that is tracked against a goal set out in the state strategic plan known as "Oregon Shines."

Demographics - General characteristics of a population (ie. age, gender, race, occupation, income).

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 quanitites. 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.

Percentage Change -

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

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

1990 population = 2,842,321

1998 population = 3,267,550

$$\frac{(3,267,550 - 2,842,321)}{2,842,321} \times 100 = 15\%$$

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.

Credits and contacts

Produced by:

The Data Users Task Group

The following data experts served as this project's core group. Feel free to contact them with questions about presentation and use of data.

From the Oregon Department of Human Services:

Cathy Iles, Senior and Disabled Services Division
(503) 945-5855, cathy.f.iles@state.or.us

Jennifer Woodward, Health Division
(503) 731-4109, jennifer.a.woodward@state.or.us

Jim White, State Office for Services to Children and Families
(503) 945-5667, jim.m.white@state.or.us

Janet Bubl, Office of Alcohol and Drug Abuse Programs
(503) 378-6197 janet.bubl@state.or.us

From Oregon Housing and Community Services:

Richard Bjelland
(503) 986-0983, richard.bjelland@hcs.state.or.us

Others have contributed over time to the development of this booklet and to other projects of the Data Users Task Group. They include Bob MacKay, Kati Neville, Barbara Groves, Heather Mowry and Pam McVay of the Department of Human Services; Tom Golson, Michael Rindfleisch and Karmen Fore of the Oregon Commission on Children and Families; Robin Peterson of the Oregon Progress Board, and Jono Hildner.

We want to especially thank the leadership of task group member **Cathy Iles** in coordinating the completion of this guide.

Editing and design:

Department of Human Services Communication Group

January 2001