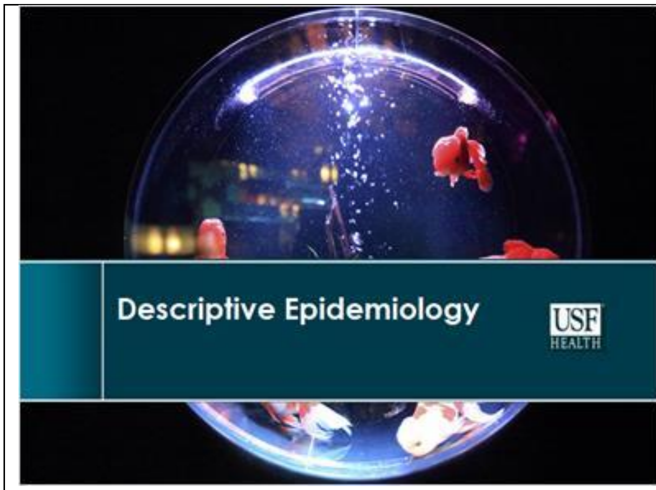
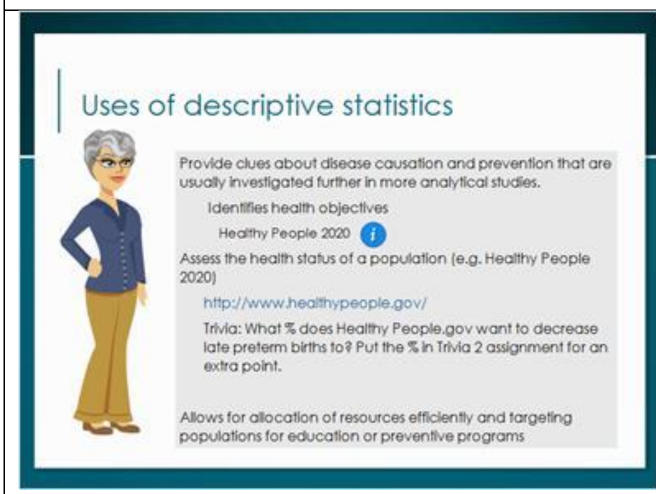
Descriptive Epidemiology

<table>
<tr>
<td>


</td>
<td>Descriptive epidemiology is the type of epidemiological research that provides information on disease patterns by considering various characteristics of person, place and time, using descriptive statistics. The purpose of descriptive epidemiology is to describe the health situation rather than attempting to identify a causative factor. While descriptive studies can be the first step in identifying the causes of disease, they function mostly to find possible associations worth greater study in an analytical study. As I said in the prior lecture descriptive studies be tend to case series and cross-sectional studies. Descriptive studies are said to be hypothesis generating as opposed to hypothesis testing. Sometimes cross-sectional studies are used to test hypotheses but they are not considered as strong a study design as the more analytical studies.</td>
</tr>
<tr>
<td>


</td>
<td>Descriptive statistics have many practical uses. They help us understand more about potential risk factors for diseases and prevention activities that we might want to look at in more detailed studies.

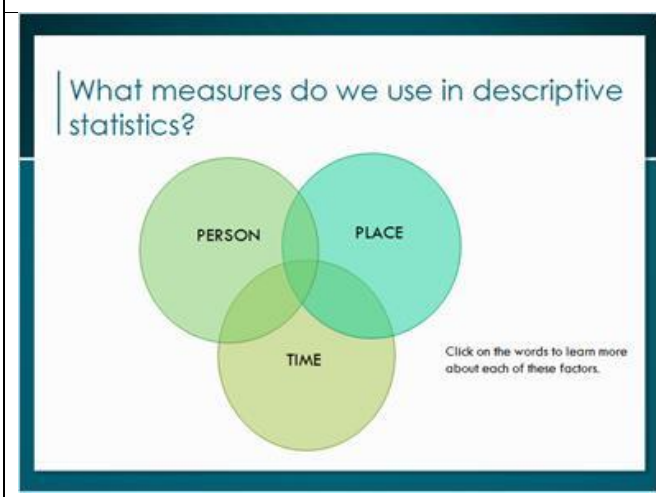
They also help us assess the health status of a population and set health goals. I have included a link to a video on a very important project, Health People 2020, which sets goals for many health measures in the US. I also put in a link to their website. Take some time to look around on the web site to get an idea of the different measures.

Based on these statistics, it is possible to identify needed resources and efficiently allocate resources. For example, if we know the number of people with asthma in a community, we would be able to identify the number of physicians, respiratory therapists, and supplies such as nebulizers that might be needed. If we identify an area in which people have higher rates of disease, we could set up public health programs.

Target 8.1%</td>
</tr>
</table>

We can obtain descriptive statistics through relying on existing data or collecting our own data. People often use data from vital records in descriptive studies. The challenge with using pre-existing data is that the questions were asked for other purposes and they may not collect exactly what you need to know. As you saw on the lecture on sources of data, we can often find descriptive date on the internet. While data we collect ourselves may be more in line with what we need, it is expensive to collect and we are limited in the number of participants.

An example of this issue is we might want to do a study looking at the impact of cigarette smoking on birthweight. If we use birth certificate data we have large number of people in the study, but the data on smoking in the birth certificates is limited to very few questions and it has a lot of missing data. If we collect our own data we can get very accurate data but we would not have such a general population and we would be limited in the number of people in our study unless we had a high level of grant funding. Plus it would take much more time to collect our own data as opposed to using data that was already available. Researchers often struggle with this issue.



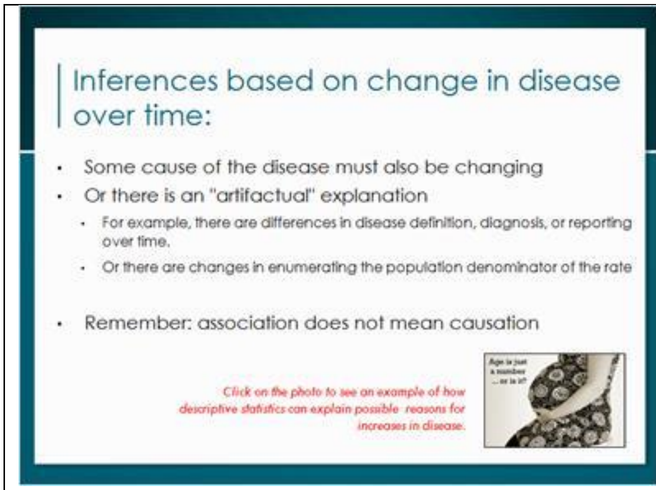
We talked before about the importance of person, place, and time. The TED lecture on heart attack and the analysis you did of environmental risk looked at the role of place.

Time is also an interesting factor to evaluate. It can be as small as hours per day or as large as decades. For example, we know mosquitos carrying different diseases bite at different hours in the day. Or we could look at the variation in disease from season to season. Winter is the season in which we are most likely to acquire influenza. At other times we look at decades or even centuries.

Person characteristics are often evaluated in understanding disease. We look at age, race, SES, immune function, gender, and many others. We will be spending more time on age this semester.

Take some time to click on each measure to look at more information about the different factors. Remember, you are responsible for additional materials so do not skip these in the lectures.

	Time: Malaria is a disease that shows clear seasonal variation as seen in this slide from Bataan. As in other countries, malaria seasonality is related to the viability of vector breeding sites, i.e., the rise in vectors numbers are one of the main causes of the rise in the number of cases. Because the most important vector in Bataan is Anopheles flavirostris, and the ideal breeding sites are “clear, shaded, slow-flowing streams or irrigation ditches, the increase in the number of cases in the majority of the malaria endemic provinces occurs during the time of the year when these characteristic streams abound, be it the rainy or dry season”[1]. In addition, there is great annual variation in Malaria incidence, largely the result of changes in the climate. This slide from the Solomon Islands shows how great the annual variation can be. **Place:** This slide shows the variation in melanoma mortality rates (skin cancer) across the United States. While rates are generally high in the South, the highest rates appear to be in the West. Please note that the dark green states are those for which there is insufficient data to estimate the rates. The assignment you completed in the last module illustrated the importance of place of residence for heart disease. **Person:** While we know Alzheimer’s disease increases with age, this fairly sobering graph shows how high the rates are for our very elderly population, with 44% of people age 80 and over having this disease. Given the increased numbers of more elderly people in the United States, this is a serious upcoming epidemic unless this disease can be prevented or treated.



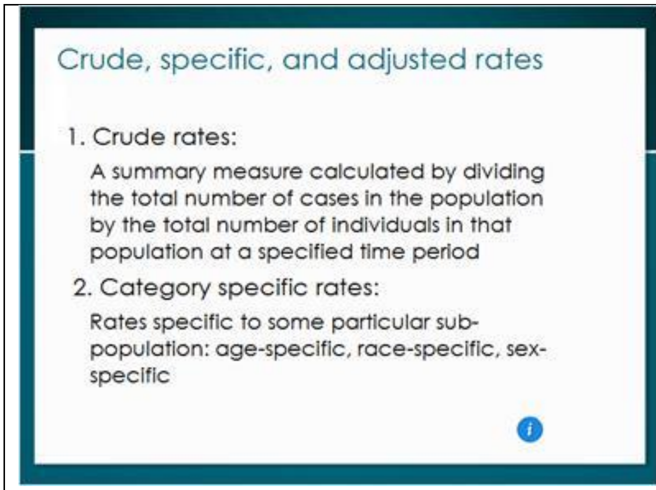
Let's just look at time again as an example of how we think about disease. If a disease changes over time, there are a number of possibilities. The first is that the cause of the disease is changing. We saw that in the example of malaria which changes throughout the year but also from year to year. Changes in malaris incidence can be due to climate changes but also changes in conditions that impact on the reproduction of the vectors. We can also decrease malaria by providing bed nets which has been a very effective public health intervention. Or we can conduct immunization programs to decrease the number of susceptible people for some other diseases.

But there can be other reasons why rates of disease change. We will learn about surveillance systems later in the semester but reporting can change for several reasons. There might be a change in the case definition of a disease, so that new conditions are counted that were not counted before. People may be very aware of a disease because of media reports and physicians might be more likely to test for it and diagnose it. And there could be changes in people going to a medical facility to get a diagnosis, especially if something is in the news. There can be under-reporting of certain conditions, like HIV or AIDS as it is seen as a stigma. This was well documented early in the epidemic.

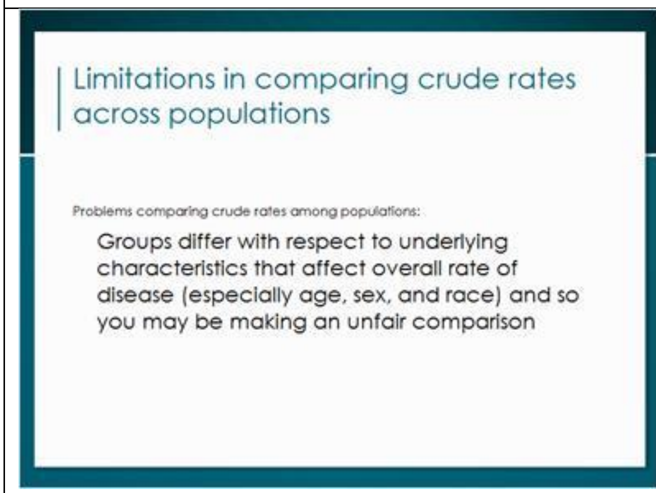
And finally there can be changes in the population. As we saw Alzheimer's increases with age. As the overall population becomes more elderly, there would be an increase in Alzheimer's due to age alone. Click on the photo to see more information on the relationship of maternal age with autism.

Autism: This slide shows the increase in the number of children diagnosed with Autism from 1992 to 2006. Since this slides shows numbers rather than rates, some of the increase may be population growth but we do know rates have increased as well. We don't fully understand the risk but believe it is due to a number of factors, including those that are true and others that are artificial. For example research has found that children with similar symptoms are more likely to be diagnosed with autism now than in the past. This would be an artificial increase rather than a true

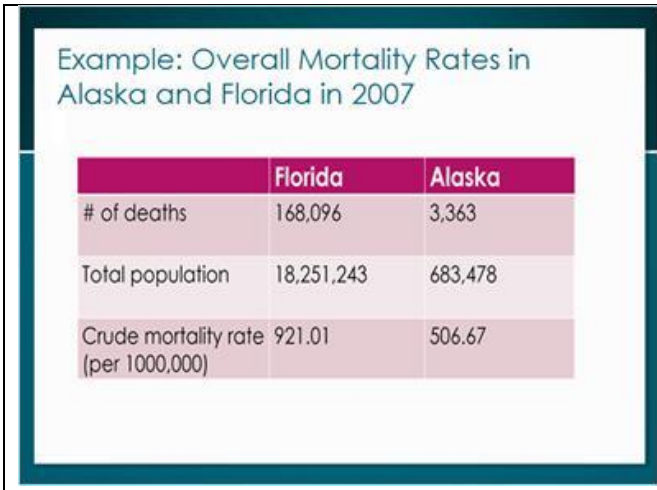
	increase. However, we are exploring other possible reasons for this increase besides the commonly postulated risk of vaccination. One possible factor of interest is the risk from an increase in maternal age. Click more on the slide to see the next graph. **Age at birth.** This slide shows the increase in the number of births to "older" mothers, those in their late 30s and 40s. These rates are for first time mothers showing a real demographic transition in delayed childbearing that is occurring in the United States. Click more on the slide to see the next graph. **Autism by maternal age**. This slide shows the relationship between maternal age and autism risk. There are still more questions than answers about autism but it seems clear that increasing maternal age is part of the picture. This final graph shows the rate of increase in autism cases per 10,000 births as maternal age increases. I have also attached a link to a very interesting article about this finding. This slide has two parts. The first is a link to a short article on maternal age and autism and the second is a blog about being an older mother and thinking about risk.
	As I have said age is fairly consistently shown to be an important predictor of disease. There are diseases that are more common in a variety of age groups. Some are more likely to impact the young, such as otitis media (commonly known as an ear infection), others are present in young adults such as increased rates of sexually transmitted diseases, and there are a number of diseases that are strongly associated with aging. Because age is such an important predictor of health, with a general trend towards increased illness and mortality at older ages, it is very important that if we compare diseases between different populations that we take age into account.



So how do we handle age and other factors when we compare populations? Let's take a look at a few terms. When we speak of crude rates, we are talking about a basic rate of disease that does not take into account any other factors. We simply divide the number of cases by the total population and calculate a rate. If a certain factor is considered important, we can obtain category specific rates. In this case, we divide the population into different categories and look at the rates in each group. This was done in the study looking at maternal age and autism rates. Click on the marker to revisit the autism graph we saw before. You can see that maternal age is subdivided in this graph.



So why do we need to worry about comparing disease between different group? Basically, it is the fact that these underlying characteristics, age but also gender, race, and other factors may lead to biased comparisons. These days, many statistical analyses take into account these differences by doing what is known as a multivariate analysis in which you create a statistical model in which you put different factors into the model and control for their presence. This is beyond the scope of this class but you will learn more about these techniques in future classes. What we have done, since age is generally the most important factor, is calculate age adjusted rates by hand. There are two ways in which this is done. One is known as direct age adjustment and the other indirect age adjustment which uses something called a standardized mortality ratio or SMR. I am less concerned with you learning how to calculate these rates as I am with you understanding how it is done and how to interpret the rates. But I believe to understand this better, you need to perform the calculations by hand. So let's look at an example comparing age between Florida and Alaska.



	Florida	Alaska
# of deaths	168,096	3,363
Total population	18,251,243	683,478
Crude mortality rate (per 1000,000)	921.01	506.67

The difference in the age structure between the populations of Florida and Alaska make a comparison of crude rates an unfair comparison. On the surface Florida looks like a much more dangerous place to live than Alaska with people almost twice as likely to die in Florida as Alaska. But the reason for this is pretty simple.

People in Florida tend to be older than people in Alaska
Older people are more likely to die than younger people

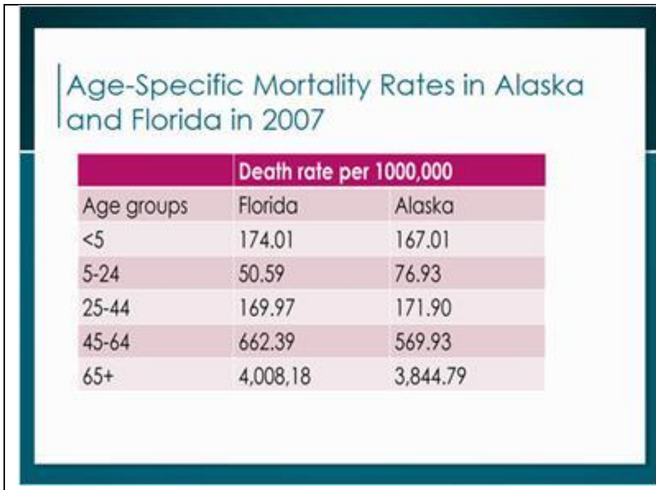
Based on the prior discussion of calculating crude rates, we know that the crude rate of mortality for Florida is calculated by dividing the total number of deaths by the populations and then multiplying it by the number we want in the denominator (in this case 100,000)

Total number of deaths / total population
168,096 / 18,251,243 = 921.01 / 100,000
I want to record this in a brief video showing me doing it

Now you do the same thing for Alaska. Insert quiz on this. Present your results using two decimal places. Do we need a marker saying what a decimal place is?
3,363/683,478 times 1000,000= 492.04 per 1000,000 Answer sheet should say per 100,000 and they put in the number 492.04

	Death rate per 1000,000	
Age groups	Florida	Alaska
<5	174.01	167.01
5-24	50.59	76.93
25-44	169.97	171.90
45-64	662.39	569.93
65+	4,008,18	3,844.79

Fill in the Black Interaction



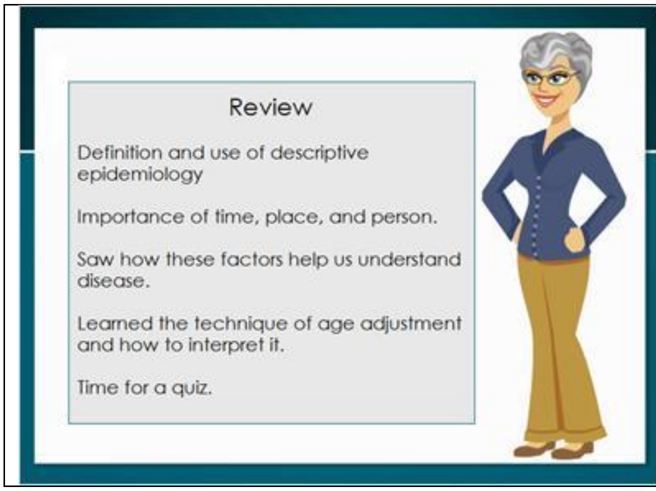
	Death rate per 1000,000	
Age groups	Florida	Alaska
<5	174.01	167.01
5-24	50.59	76.93
25-44	169.97	171.90
45-64	662.39	569.93
65+	4,008.18	3,844.79

You can see when you divide the states into smaller age groups, you get different rates per age. So Florida has higher mortality rates for the young (<5), and older groups (45-64 and 65+) but lower mortality rates for people ages 5-24 and 25-44. Keep in mind that there may be very different distributions of people in the oldest age category as Florida may have more people in their 80 and 90s than Alaska.

To fully compare the rates, you need to perform an age adjustment. As I said previously, there are two types of age adjustment: direct and indirect. For this analysis, we are going to perform direct age adjustment.

To do this, you need to know the death rate per age group for each population, as shown above. You apply a standard population distribution to artificially determine how many deaths there would be in each group if they had the same age distribution. This is known as the weight, although many students find it confusing to think of the terms rates and weights. We often use a population distribution from the US census data as these are available but it does not matter which distribution you use, as you are obtaining numbers for the purpose of comparison and not real numbers that actually exist. This can be done by hand and it is worth doing so to understand how the calculations work. I am going to show you a video By Dr. Pathwari from Youtube which provides a very clear explanation of this procedure.

This vide shows you how to do age adjustment.

	In this lecture we talked about descriptive epidemiology. We specifically looked at the importance of time, place, and person in understanding disease. We saw how these factors can better help us understand disease. And we focused a bit on age as it is such an important predictor of disease. You watched a video on age adjustment and we talked about how to interpret it. Now it is time for you to complete a quiz the quiz on age adjustment to be sure you understand this concept.