Injury Prevention
Indian Health Service
Portland Area

Statistics : The Basics

What is Statistics?

Data is factual information such as number of crashes, most popular car on the road, number of children in a car seat, the population in the United States (Census). Data is special information organized for analysis so that it can be used to reason, design program, and/or make decisions. After the data is collected and organized, **statistics** is the math that is used to analyze and interpret the data. Depending on your project, the statistics can be very basic and easy to use; but can also get very complicated for larger projects. Descriptive statistics is easier to use than inferential statistics. **Biostatistics** are statistics that are applied specifically to biological and health sciences including injury prevention.

Two Types of Statistics

1. **Descriptive**. The methods used here summarize and/or briefly describe the information contained in your data. You are most familiar with this method during the census that is taken every 10 years. Every household is surveyed for the number of people residing there along with many other questions. The Census Bureau then describes this data and lets us know how many people live in this country, how many are males, how many are females, what the age groups are, etc.

2. **Inferential.** The methods used here are for projects that used a sample of the population, and will infer or generalize sample results to the population at large. You are most familiar with this process during election time. Sampling polls are set up to find out how people voted in an election and then the results are inferred or generalized to the entire population. For example, if 60% of the people sampled voted yes on Proposition A, one could infer that 60% of the entire voting population also voted yes on Proposition A. (if the sample size was large enough).

    Inferential statistics is more complex and requires more statistical training than is offered here in this document.

Types of Data

*Quali*tative Data (Also known as Categorical Data).   This data cannot be added. The data can come in **nominal** form (unordered) which is data like colors of hair (brown, black, red), did you like the movie (yes/no), gender (male/female), race (American Indian, Hispanic, Asian). The data can also come in an **ordinal** form (ordered). This data has a sequence (mild, moderate, severe) or how did you rate the movie (bad, fair, good, excellent). **Binary** can be nominal or ordinal and just means that there are just two categories (yes/no, green/blue). **Non-binary** means more than two categories.

*Quan*titative Data (Also known as Numerical Data). Data in the form of numbers gives us information we need to identify injury problems and see certain patterns. The data can come in **discrete** numbers (numbers of crashes, number of injuries) or **continuous** (height, age, blood pressure).

Data Presentation

How you present your data will depend on which type of data you are using. Here are some examples. Consult a statistics book for more kinds of presentation.

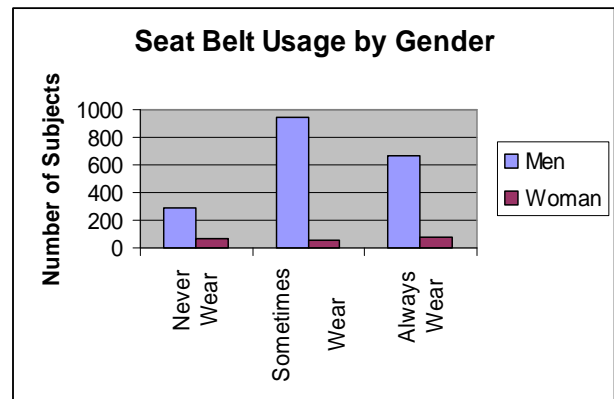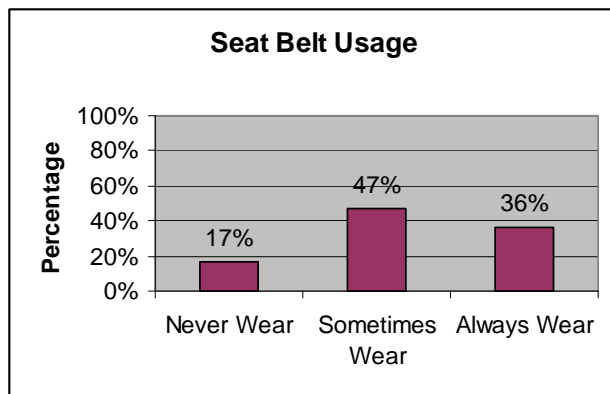| Categorical Data	Numerical Data
Tables and Charts Percentages (row and column) Bar chart Clustered Bar Chart Pie Chart	Tables and Charts Averages (Mean, Median, Mode) Variability (Range and Std Deviation) Histogram

You will often see the letter "N" or "n" in tables, graphs, and charts. An "N" or "n" stands for "number" and denotes the number of people in the group.

Tables and Charts

Seat Belt Usage

	N	Percent
Never Wear	358	17%
Sometimes Wear	994	47%
Always Wear	747	36%
Total	**2099**	

Seat Belt Usage

	Never Wear	Sometimes Wear	Always Wear	Total
Men	293	941	666	**1900**
Woman	65	53	81	**199**
Total	358	994	747	**2099**

Bar Chart

Seat Belt Usage

Seat Belt Usage by Gender

Pie Chart

Seat Belt Usage by Percent

Never Weat Seat Belt