Chapter 1
Overview and Descriptive Statistics

Brittany Cuchta
STAT 3115, Spring 2015

Example 1

The following data set contains measurements on a selection of 1978-79 model automobiles.

Car	Country	MPG	Weight	Horsepower
Buick Estate Wagon	U.S.	16.9	4.360	155
Ford Country Squire Wagon	U.S.	15.5	4.054	142
Chevy Malibu Wagon	U.S.	19.2	3.605	125
Chrysler LeBaron Wagon	U.S.	18.5	3.940	150
Toyota Corona	Japan	27.5	2.560	95
Datsun 510	Japan	27.2	2.300	97
Audi 5000	Germany	20.3	2.830	103
Volvo 240 GL	Sweden	17.0	3.140	125
Saab 99 GLE	Sweden	21.6	2.795	115
Peugeot 694 SL	France	16.2	3.410	133

What is the population? What are the variables and their type? Is the data set multivarate or univariate? What is the sample size?

Example 2

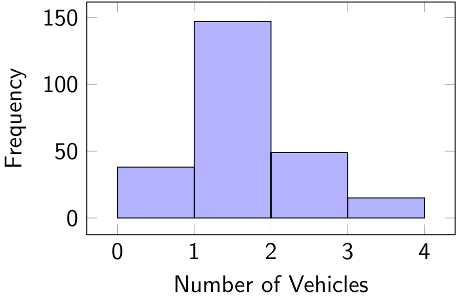
Suppose you want to determine the average amount of time a student at S&T spends studying

1. What is the population of interest?
2. What variables are you interested in and what kind are they?
3. What are the pros and cons of each sampling method?

Example 3

An observer at a major intersection counted the number of cars running red lights each time.

Number of vehicles	Count	Relative Frequency
0	38	0.151
1	147	0.586
2	49	0.195
3	15	0.060
4	2	0.008
Total	251	1



Example 4

We are given the following ordered data set:

1 1 1 1 2 2 2 2 2 3 3 3 4 4 4 4 5 5 5 7 8

1. What is the sample size?
2. Calculate the mean.
3. Find the median.

Example 4

We are given the following ordered data set:

$$1\ 1\ 1\ 1\ 2\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 4\ 4\ 4\ 4\ 5\ 5\ 5\ 7\ 8$$

1. What is the sample size?
2. Calculate the mean.
3. Find the median.

☆ Note that if we changed one of our observations from 8 to 13 we would have the same median but we would now have a mean of 3.5. The mean is very sensitive!

Example 5

Continuing with the same data as Example (4),

1 1 1 1 2 2 2 2 2 3 3 3 4 4 4 4 5 5 5 7 8

1. Find the lower quartile.
2. Find the upper quartile.

Example 6

Recall the Example (4) data set:

1 1 1 1 2 2 2 2 2 3 3 3 4 4 4 4 5 5 5 7 8

What are the measures of spread (range and standard deviation)?

Example 7

Are there any outliers in our Example (4) data set?

1 1 1 1 2 2 2 2 2 3 3 3 4 4 4 4 5 5 5 7 8

Example 8