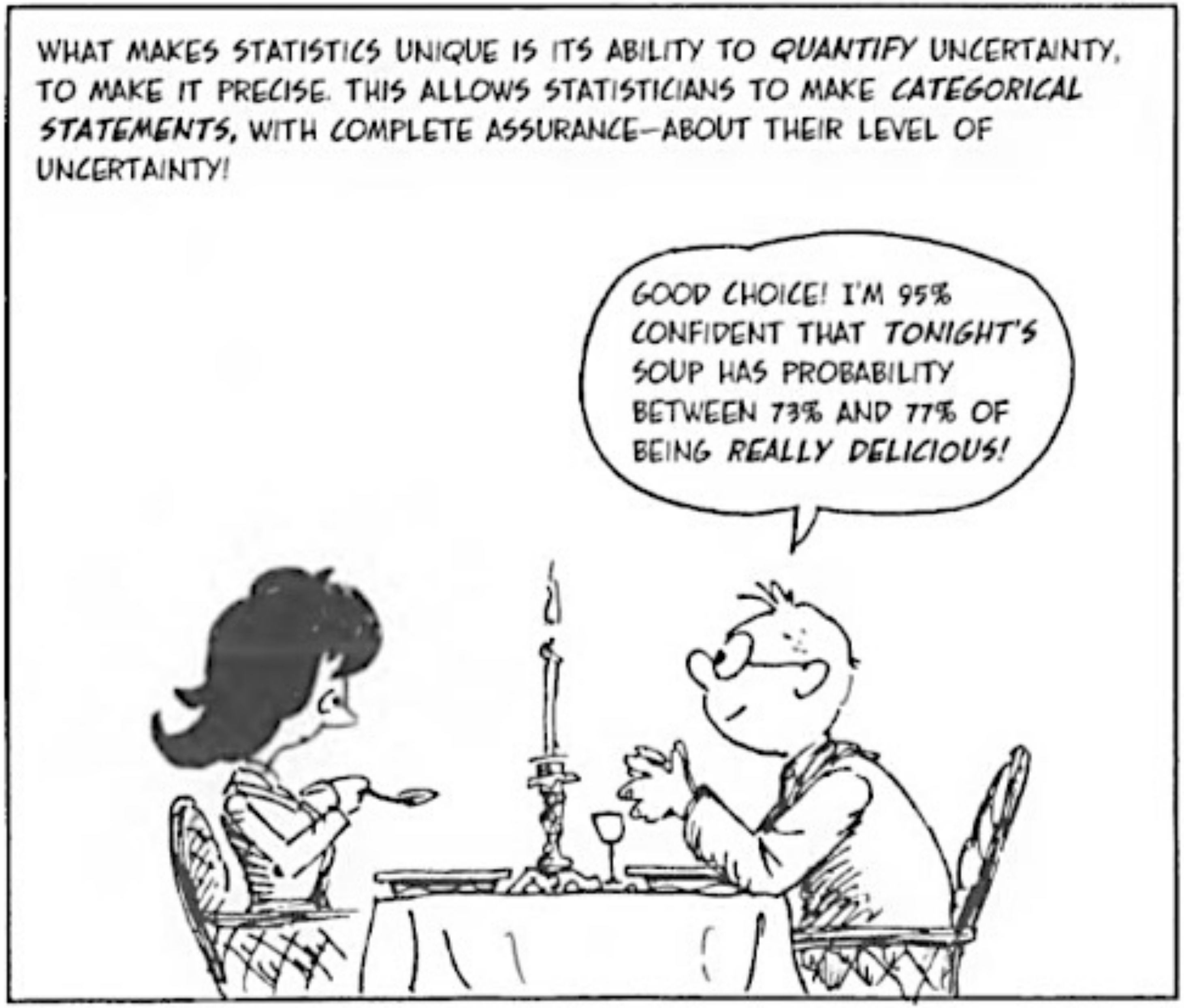
WHAT IS STATISTICS?

WHAT IS STATISTICS?

SHOULD I HAVE THE SOUP? EVERYTHING ELSE IS SO *EXPENSIVE*, AND I DON'T KNOW WHO'S PAYING... ARE STATISTICIANS *STINGY?* I'VE NEVER GONE OUT WITH ONE BEFORE... THOUGH I ONCE KNEW A *VERY* GENEROUS ACCOUNTANT...
SHOULD I HAVE THE SOUP? *27* OUT OF THE *36* TIMES I'VE HAD IT, IT WAS PRETTY GOOD... BUT IS MONDAY THE REGULAR CHEF'S NIGHT OFF? AND WHAT IF ALL THE *AIR MOLECULES* IN THE ROOM SUDDENLY FLY UP TO THE CEILING?

From: Gonick & Smith The cartoon guide to Statistics

WHAT IS STATISTICS?


WHAT MAKES STATISTICS UNIQUE IS ITS ABILITY TO QUANTIFY UNCERTAINTY, TO MAKE IT PRECISE. THIS ALLOWS STATISTICIANS TO MAKE CATEGORICAL STATEMENTS, WITH COMPLETE ASSURANCE—ABOUT THEIR LEVEL OF UNCERTAINTY!
GOOD CHOICE! I'M 95% CONFIDENT THAT TONIGHT'S SOUP HAS PROBABILITY BETWEEN 73% AND 77% OF BEING REALLY DELICIOUS!

From: Gonick & Smith The cartoon guide to Statistics

WHAT IS STATISTICS?

The discipline of statistics also provides methods for organizing and summarizing data and for drawing conclusions based on information contained in the data.

WHY DO YOU NEED TO STUDY STATISTICS?

Examples of everyday life questions that Statistics can help to address

- How many tree are on the planet?
- What's the cancer risk from bacon sandwiches?
- Do busier hospitals have higher survival rates?
- Does going to the university increase the risk of getting a brain tumor?
- Do the daily number of homicides in the UK follow a Poisson distribution?
- Are mothers' heights associated with their sons' heights ?

These examples (and you can find many others) have been taken from the highly recommended general reading book of Sir David Spiegelhalter ”The art of Statistics Learning from Data”

The Scope of Modern Statistics

These days statistical methodology is employed by investigators in virtually all disciplines, including such areas as

- **Marketing** (developing market surveys and strategies for marketing new products).

- **Public health** (identifying sources of diseases and ways to treat them)

- **Physics and engineering** (from assessing the effects of stress of structural elements and the impacts of traffic flows on communities)

- **Molecular biology** (analysis of genomic data)

- **Ecology** (for example by describing quantitatively how individuals in various animal and plant populations are spatially distributed)

- **Materials engineering** (studying properties of various treatments to retard corrosion)

The Scope of Modern Statistics : some examples

ENGINEERING

FOR EXAMPLE, IN 1986, THE SPACE SHUTTLE *CHALLENGER* EXPLODED, KILLING SEVEN ASTRONAUTS. THE DECISION TO LAUNCH IN 29-DEGREE WEATHER HAD BEEN MADE WITHOUT DOING A SIMPLE ANALYSIS OF PERFORMANCE DATA AT LOW TEMPERATURE.

OH...THAT PART OF THE CURVE!!!

MEDICINE

A MORE POSITIVE EXAMPLE IS THE *SALK POLIO VACCINE*. IN 1954, VACCINE TRIALS WERE PERFORMED ON SOME 400,000 CHILDREN, WITH STRICT CONTROLS TO ELIMINATE BIASED RESULTS. GOOD STATISTICAL ANALYSIS OF THE RESULTS FIRMLY ESTABLISHED THE VACCINE'S EFFECTIVENESS, AND TODAY POLIO IS ALMOST UNKNOWN.

From: Gonick & Smith The cartoon guide to Statistics

The Scope of Modern Statistics/Probability theory: some examples

PURE MATHEMATICS

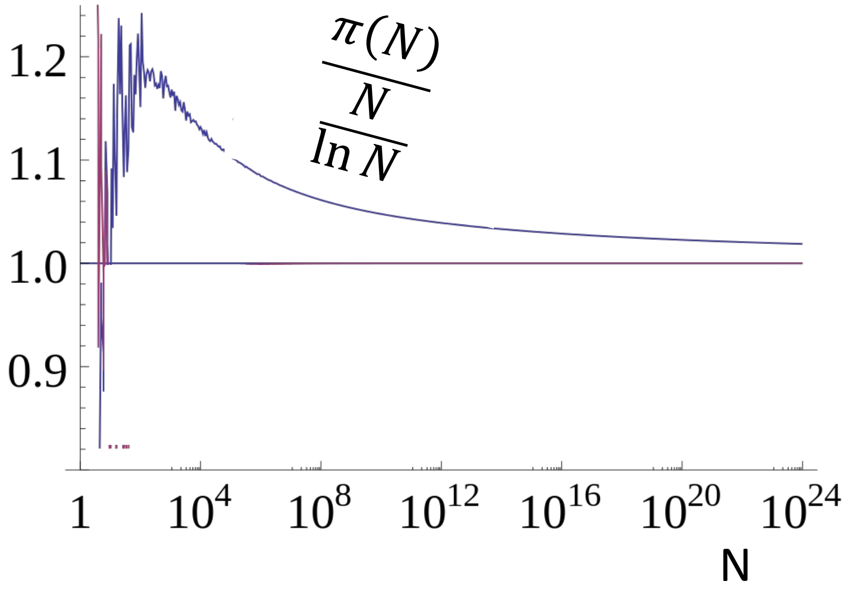
The *prime number theorem*

It gives the asymptotic **distribution of prime numbers** amongst positive integers.

The theorem states that if any random positive integer is selected in the range of zero to a large number N, the probability that the selected integer is a prime is the inverse of the natural logarithm of N.

Therefore, if $\pi(N)$ denotes the number of primes less than or equal to N that are prime. The theorem states that p(N)= $\frac{\boldsymbol{\pi(N)}}{\boldsymbol{N}}$ is asymptotic to $\frac{1}{\ln N}$ which means that

$$\lim_{N\to\infty}\frac{\frac{\pi(N)}{N}}{\frac{1}{\ln N}} = 1$$



As corollary, the theorem also states that the primes become less common as they become larger.

The Scope of Modern Statistics/Probability theory: some examples

CHEMISTRY   The Born Interpretation of the wave function

The square of wave function ψ is a measure for the probability P(r) to find the particle in an infinitesimal volume element dτ around r.

Distribution of a quantum
Particle in a box

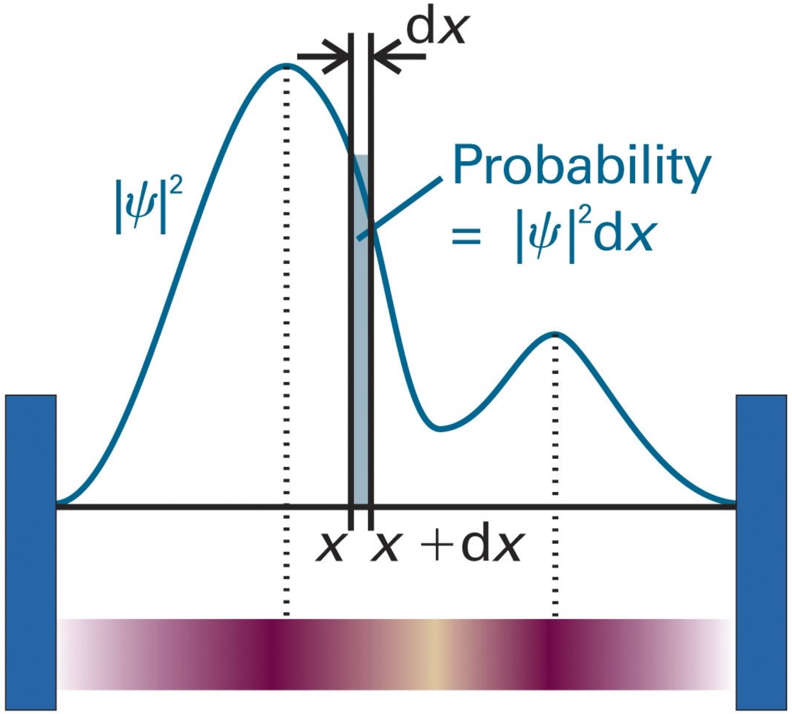
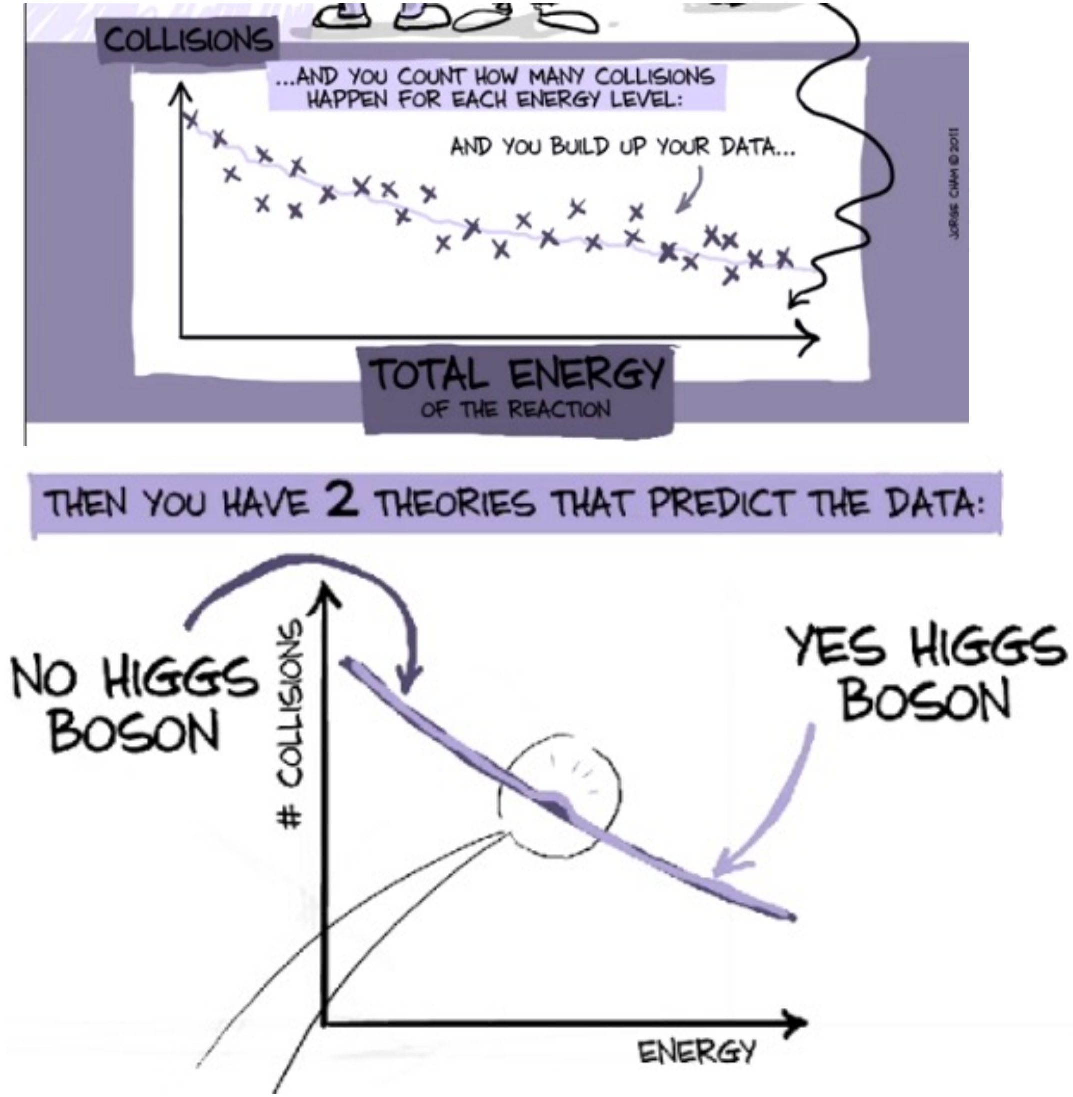


Figure 8-19
Atkins Physical Chemistry, Eighth Edition
© 2006 Peter Atkins and Julio de Paula

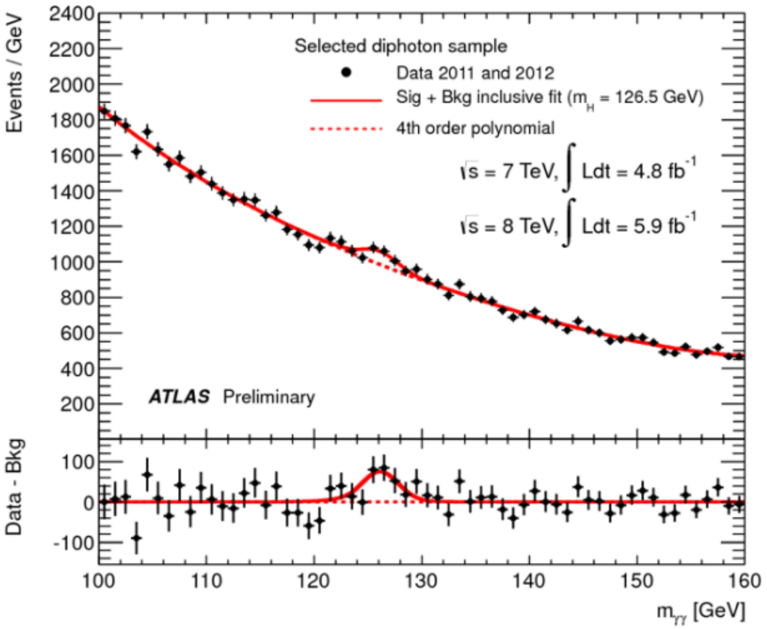
The Scope of Modern Statistics/Probability theory: some examples

PHYSICS

The Discovery of the Higgs Boson



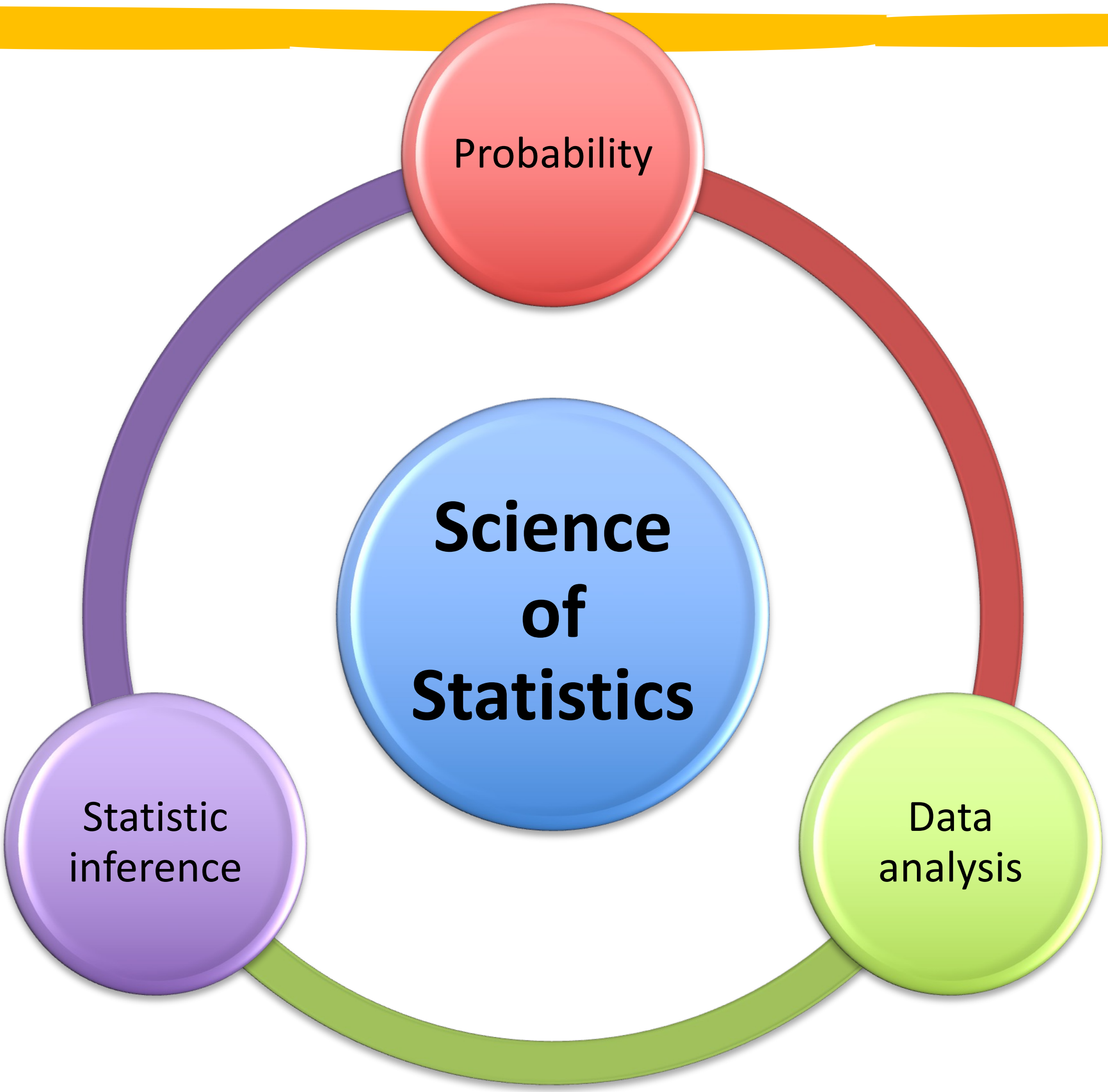
COLLISIONS
...AND YOU COUNT HOW MANY COLLISIONS HAPPEN FOR EACH ENERGY LEVEL:
AND YOU BUILD UP YOUR DATA...
TOTAL ENERGY
OF THE REACTION
JORGE CHAM © 2011
THEN YOU HAVE 2 THEORIES THAT PREDICT THE DATA:
NO HIGGS BOSON
# COLLISIONS
YES HIGGS BOSON
ENERGY



https://atlas.cern/updates/feature/higgs-boson

https://blog.revolutionanalytics.com/2012/07/discovering-the-higgs-boson-with-statistics.html

THE BRANCHES OF STATISTICS



DATA ANALYSIS : The gathering, display, and summary of data.

PROBABILITY : The laws of chance in and out the casino!

STATISTICAL INFERENCE: The science of drawing statistical conclusions from specific data using a knowledge of probability.

THE BRANCHES OF STATISTICS

- An investigator who has collected data may wish simply to summarize and describe important features of the data. This entails using methods from **descriptive statistics.**

- Some of these methods are graphical in nature; the construction of histograms, boxplots, and scatter plots are primary examples.

- Other descriptive methods involve calculation of numerical summary measures, such as means, standard deviations, and correlation coefficients. The wide availability of statistical computer software packages has made these tasks much easier to carry out than they used to be.

THE BRANCHES OF STATISTICS

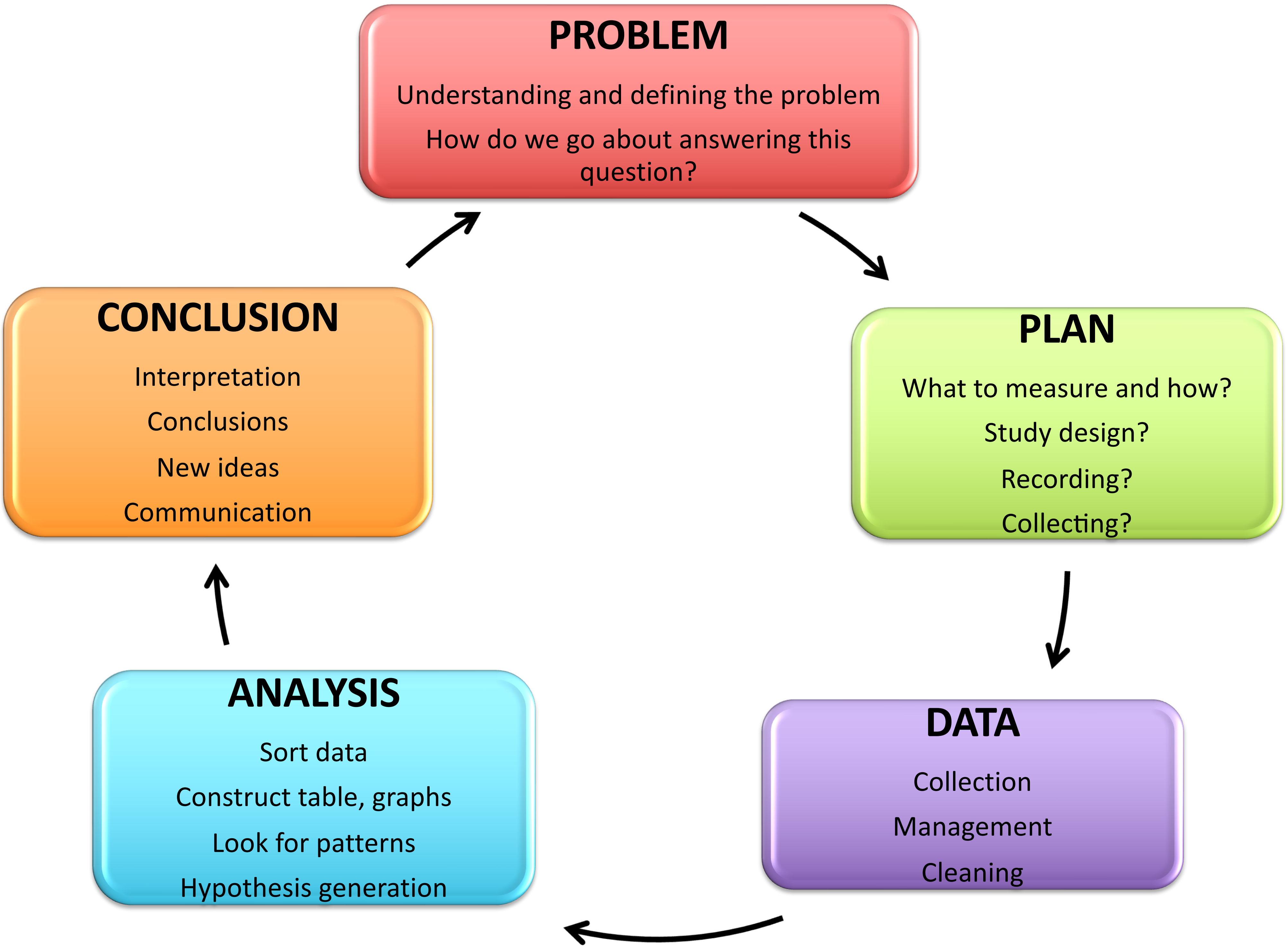
- Computers play an important role in the modern statistics since they allow to easily manipulate large quantity of data.

- *In this module, we are going to learn how to use programming language Python to solve problems in statistics and probability theory that cannot be solved using pen and paper*.

The Art of Statistics

Data are the statistician’s raw material for their analysis we are going to learn methods to represent and represent to extract the relevant information.

All statistical problems can be tackled using the so-called **PPDAC cycle**:



PROBLEM

Understanding and defining the problem

How do we go about answering this question?

PLAN

What to measure and how?

Study design?

Recording?

Collecting?

DATA

Collection

Management

Cleaning

ANALYSIS

Sort data

Construct table, graphs

Look for patterns

Hypothesis generation

CONCLUSION

Interpretation

Conclusions

New ideas

Communication

Populations, Samples, and Processes

Scientists are constantly exposed to collections of facts, or **data,** in their professional capacities and everyday activities.

The discipline of statistics provides methods for organizing and summarizing data and drawing conclusions based on information contained in the data.

Populations, Samples, and Processes

- An investigation will typically focus on a well-defined collection of objects constituting a **population** of interest. In one study, the population might consist of all gelatin capsules of a particular type produced during a specified period.

- Another investigation might involve the population consisting of all individuals who received a B.S. in engineering during the most recent academic year.

Populations, Samples, and Processes

- When desired information is available for all objects in the population, we have what is called a **census.**

- Constraints on time, money, and other scarce resources usually make a census impractical or infeasible. Instead, a subset of the population—**a sample**—is selected in some prescribed manner.

Populations, Samples, and Processes

- Thus, we might obtain a sample of bearings from a particular production run as a basis for investigating whether bearings are conforming to manufacturing specifications, or we might select a sample of last year's engineering graduates to obtain feedback about the quality of the engineering curricula.

- We are usually interested only in certain characteristics of the objects in a population:

  the number of flaws on the surface of each casing, the thickness of each capsule wall, the gender of an engineering graduate, the age at which the individual graduated, and so on.

Populations, Samples, and Processes

- A characteristic may be categorical, such as gender or type of malfunction, or it may be numerical in nature.

- In the former case, the *value* of the characteristic is a category (e.g., female or insufficient solder), whereas in the latter case, the value is a number (e.g., age = 23 or diameter = .502 cm).

Populations, Samples, and Processes

A **variable** is any characteristic whose value may change from one object to another in the population. We shall initially denote variables by lowercase letters from the end of our alphabet.

Examples include

x = brand of calculator owned by a student

y = number of visits to a particular Web site during a specified period

z = braking distance of an automobile under specified conditions

Populations, Samples, and Processes

Data results from making observations either on a single variable or simultaneously on two or more variables.
A **univariate** data set consists of observations on a single variable.

- For example, we might determine the type of transmission, automatic (A) or manual (M), on each of ten automobiles recently purchased at a certain dealership, resulting in the categorical data set

  M A A A M A A M A A

- The following sample of pulse rates (beats per minute) for patients recently admitted to an adult intensive care unit is a numerical univariate data set:

  88 80 71 103 154 132 67 110 60 105

Populations, Samples, and Processes

We have **bivariate** data when observations are made on each of two variables. Our data set might consist of a (height, weight) pair for each basketball player on a team, with the first observation as (72, 168), the second as (75, 212), and so on.

- If an engineer determines the value of both *x* = component lifetime and *y* = reason for component failure, the resulting data set is bivariate with one variable numerical and the other categorical.

Populations, Samples, and Processes

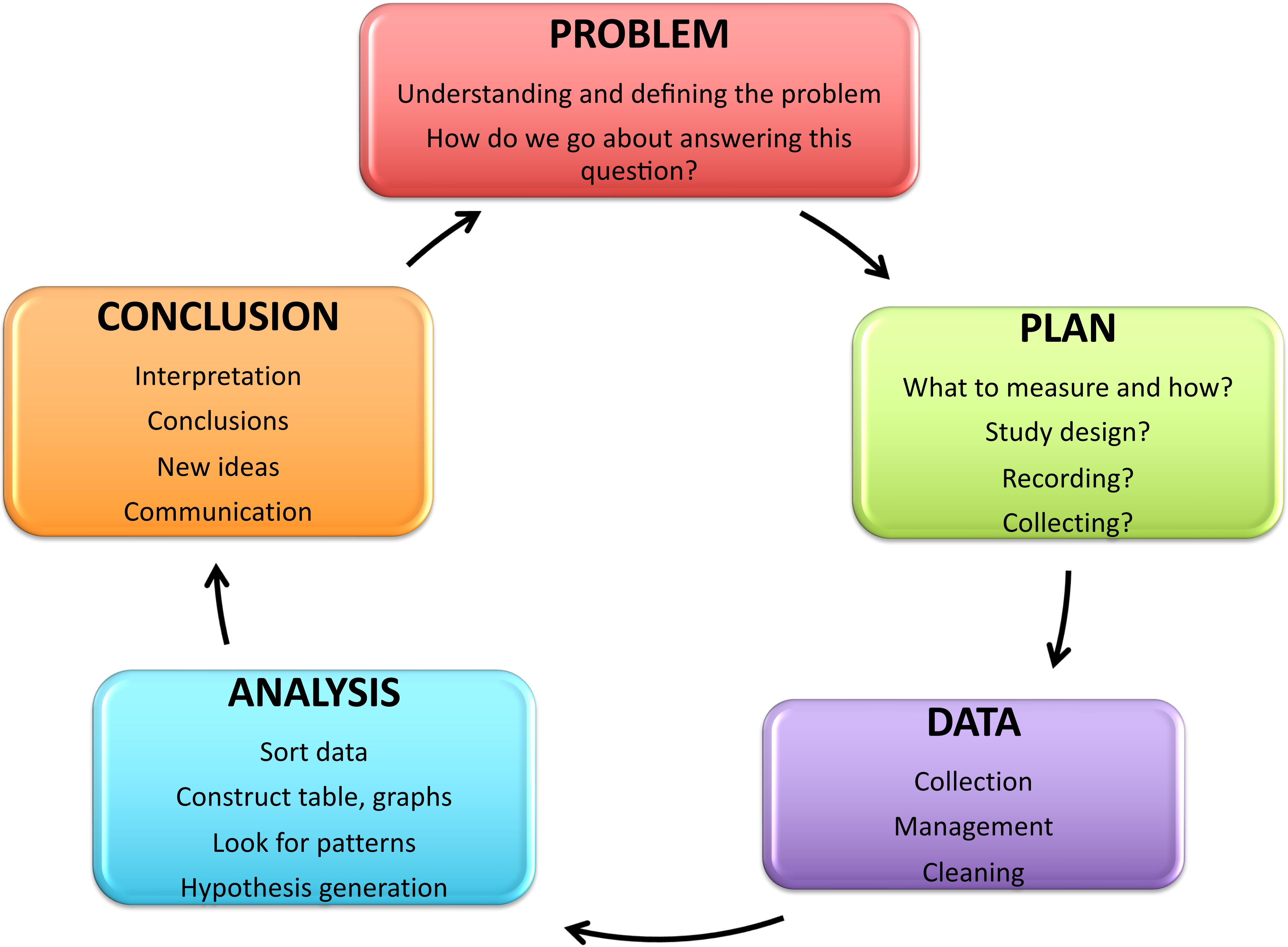
Multivariate data arises when observations are made on more than one variable (so bivariate is a special case of multivariate).

- For example, a research physician might determine the systolic blood pressure, diastolic blood pressure, and serum cholesterol level for each patient participating in a study.

Populations, Samples, and Processes

– Each observation would be a triple of numbers, such as (120, 80, 146). In many multivariate data sets, some variables are numerical, and others are categorical.

Thus, the annual automobile issue of *Consumer Reports* gives values of such variables as type of vehicle (small, sporty, compact, mid-size, large), city fuel efficiency (mpg), highway fuel efficiency (mpg), drivetrain type (rear wheel, front wheel, four wheel), and so on.



PROBLEM

Understanding and defining the problem

How do we go about answering this question?

PLAN

What to measure and how?

Study design?

Recording?

Collecting?

DATA

Collection

Management

Cleaning

ANALYSIS

Sort data

Construct table, graphs

Look for patterns

Hypothesis generation

CONCLUSION

Interpretation

Conclusions

New ideas

Communication

Descriptive statistics: Data collection

- Statistics deals not only with the organization and analysis of data once it has been collected but also with the **development of techniques for collecting the data**.
  - If data is not properly collected, an investigator may not be able to answer the questions under consideration with a reasonable degree of confidence.

- One common problem is that the **target population** *(the one about which conclusions are to be drawn*) may be different from the population actually sampled.

Descriptive statistics: Data collection

For example, advertisers would like various kinds of information about the television viewing habits of potential customers.

The most systematic information comes from placing monitoring devices in a few homes nationwide.

It has been conjectured that placement of such devices in and of itself alters viewing behavior, so that the sample's characteristics may differ from those of the target population.

Descriptive statistics: Data collection

- When data collection entails selecting individuals or objects from a frame, the simplest method for ensuring a representative selection is to take a *simple random sample.*

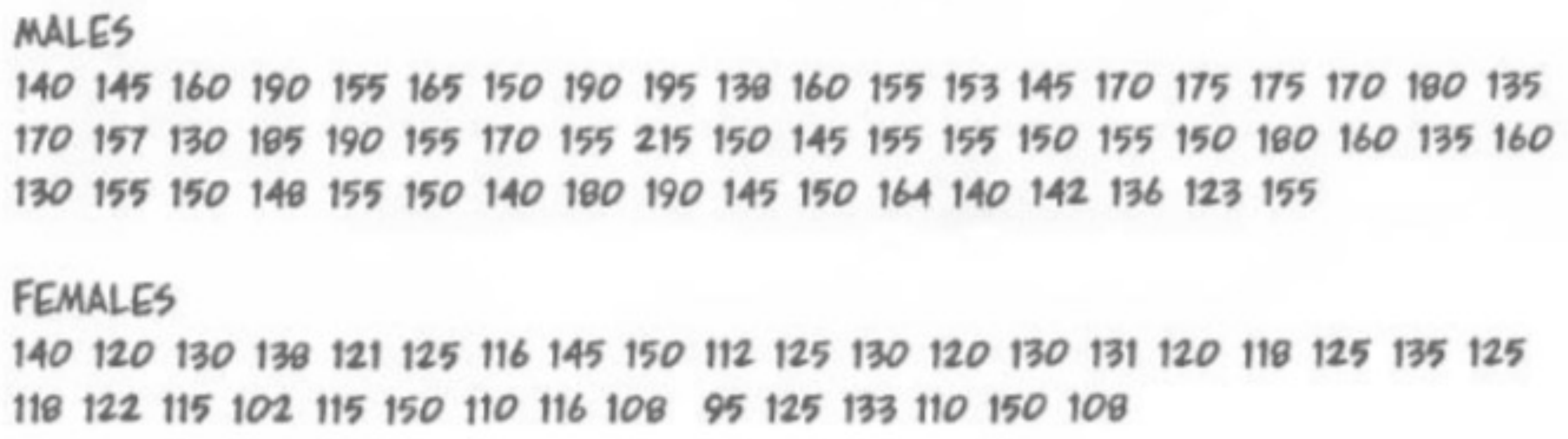
  This is one for which any particular subset of the specified size (e.g., a sample of size 100) has the same chance of being selected.

- Scientists often collect data by carrying out some sort of designed experiment. This may involve deciding how to allocate several treatments (such as fertilizers or coatings for corrosion protection) to the various experimental units (plots of land or pipe).

- Alternatively, an investigator may systematically vary the levels or categories of certain factors (e.g., pressure or type of insulating material) and observe the effect on some response variable (such as yield from a production process).

Descriptive statistics: Organizing data

This set of data represent the weight of 92 students.



MALES

140 145 160 190 155 165 150 190 195 138 160 155 153 145 170 175 175 170 180 135
170 157 130 185 190 155 170 155 215 150 145 155 155 150 155 150 180 160 135 160
130 155 150 148 155 150 140 180 190 145 150 164 140 142 136 123 155

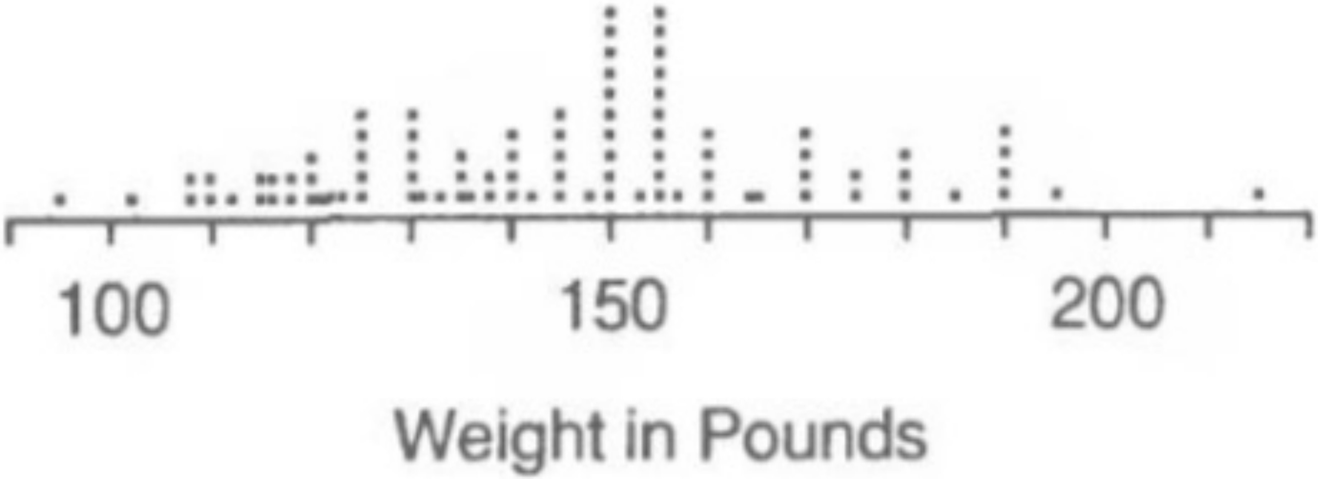
FEMALES

140 120 130 138 121 125 116 145 150 112 125 130 120 130 131 120 118 125 135 125
118 122 115 102 115 150 110 116 108 95 125 133 110 150 108

Example from: Gonick & Smith The cartoon guide to Statistics

Descriptive statistics: Organizing data

The data can be represented in a graphical form as a **dot-plot**

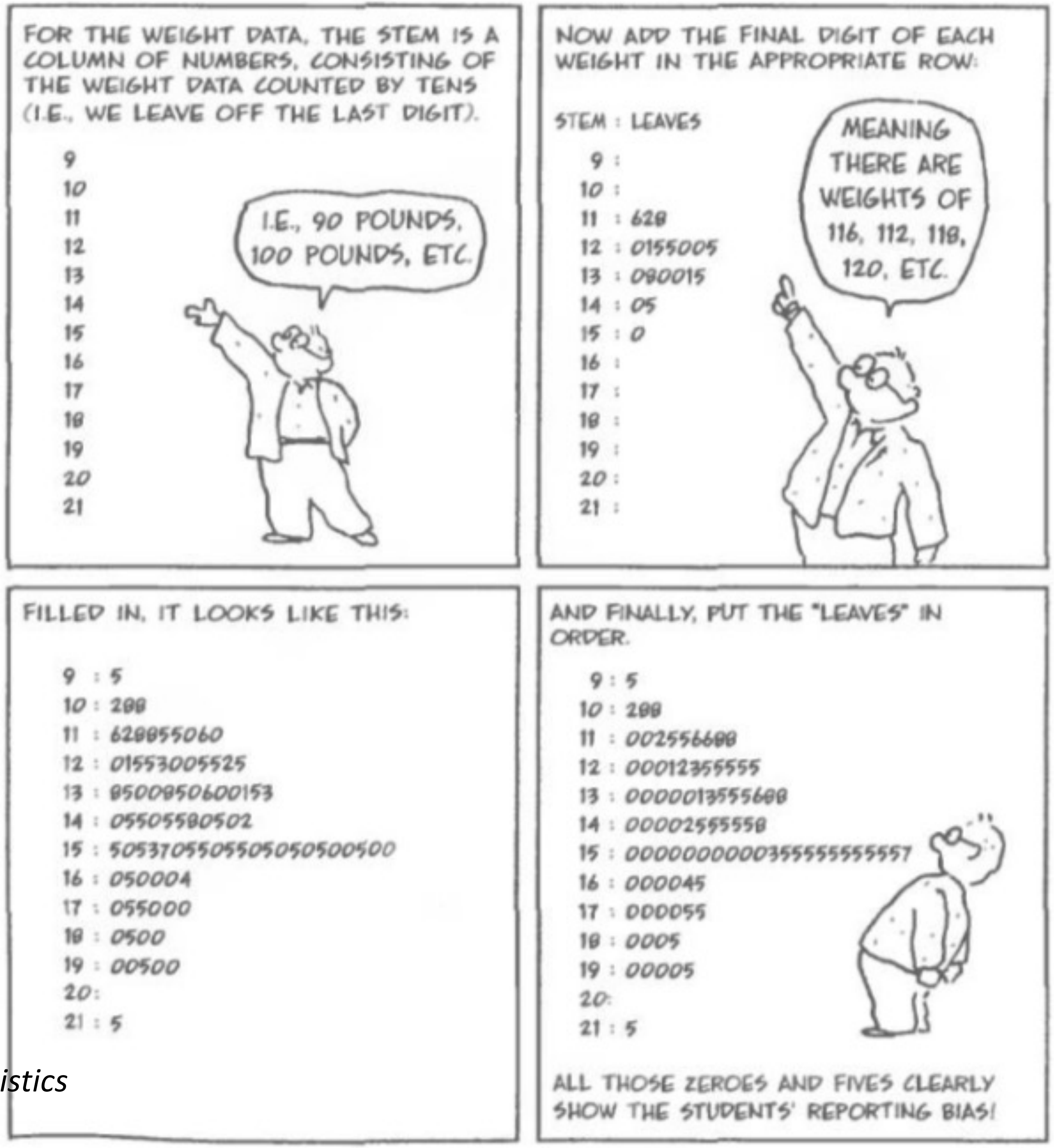


From the two main peaks at 150 and 155 this already show that there is a trend in the data as the students tended to report their weight in multiple of 5 pounds.

From: Gonick & Smith The cartoon guide to Statistics

Descriptive statistics: Organizing data

Another way to analyse
the data is the use
Of the so called stem-and-leaf
diagram invented by the
American mathematician
John Tukey



FOR THE WEIGHT DATA, THE STEM IS A COLUMN OF NUMBERS, CONSISTING OF THE WEIGHT DATA COUNTED BY TENS (I.E., WE LEAVE OFF THE LAST DIGIT).

```
9
10
11
12
13
14
15
16
17
18
19
20
21
```

I.E., 90 POUNDS, 100 POUNDS, ETC.

NOW ADD THE FINAL DIGIT OF EACH WEIGHT IN THE APPROPRIATE ROW:

```
STEM : LEAVES
  9 :
 10 :
 11 : 628
 12 : 0155005
 13 : 080015
 14 : 05
 15 : 0
 16 :
 17 :
 18 :
 19 :
 20 :
 21 :
```

MEANING THERE ARE WEIGHTS OF 116, 112, 118, 120, ETC.

FILLED IN, IT LOOKS LIKE THIS:

```
 9 : 5
10 : 288
11 : 628855060
12 : 01553005525
13 : 8500850600153
14 : 05505580502
15 : 5053705505505050500500
16 : 050004
17 : 055000
18 : 0500
19 : 00500
20:
21 : 5
```

AND FINALLY, PUT THE "LEAVES" IN ORDER.

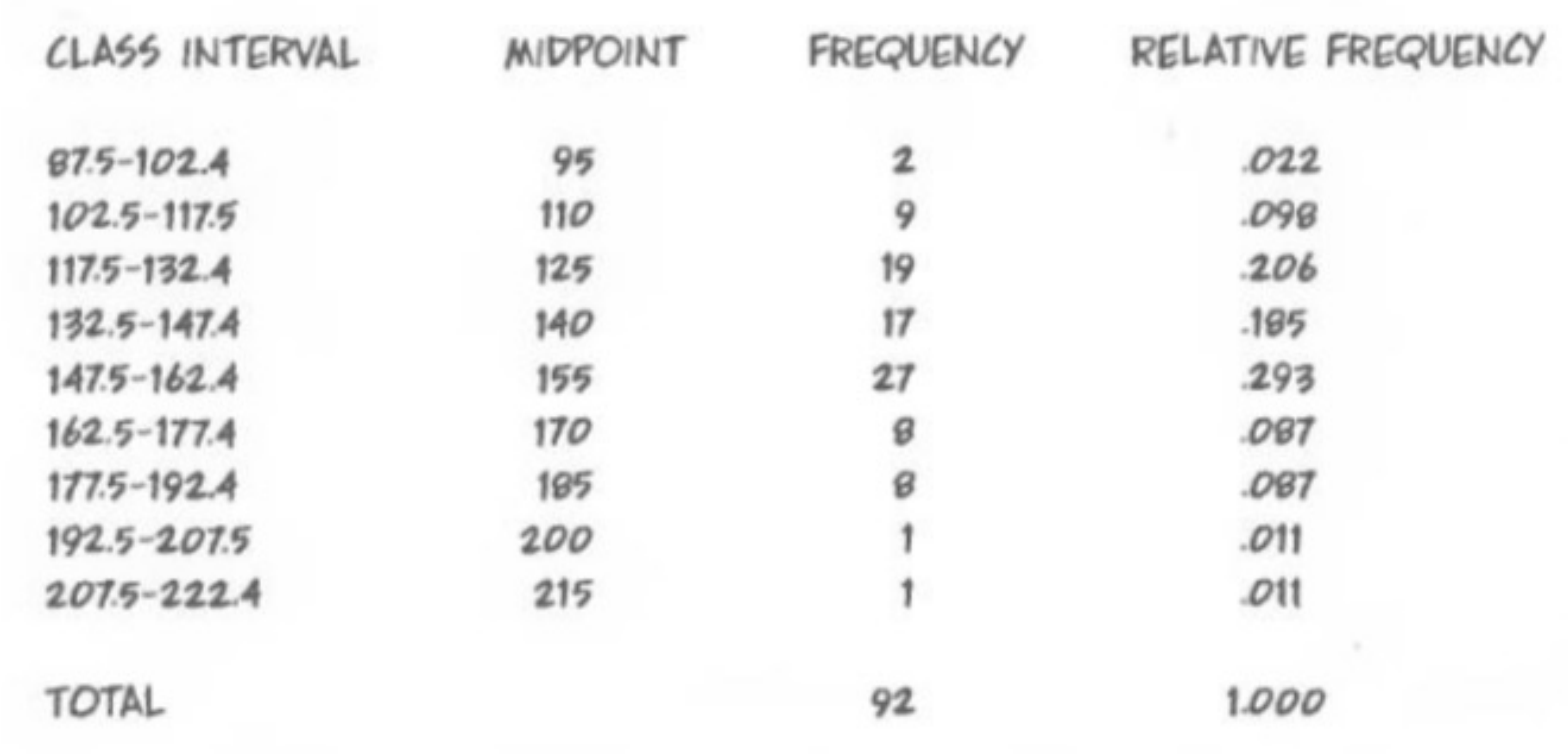
```
 9 : 5
10 : 288
11 : 002556688
12 : 00012355555
13 : 0000013555688
14 : 00002555558
15 : 00000000000355555555557
16 : 000045
17 : 000055
18 : 0005
19 : 00005
20:
21 : 5
```

ALL THOSE ZEROES AND FIVES CLEARLY SHOW THE STUDENTS' REPORTING BIAS!

From: Gonick & Smith The cartoon guide to Statistics and Wikipedia.

Descriptive statistics: data analysis

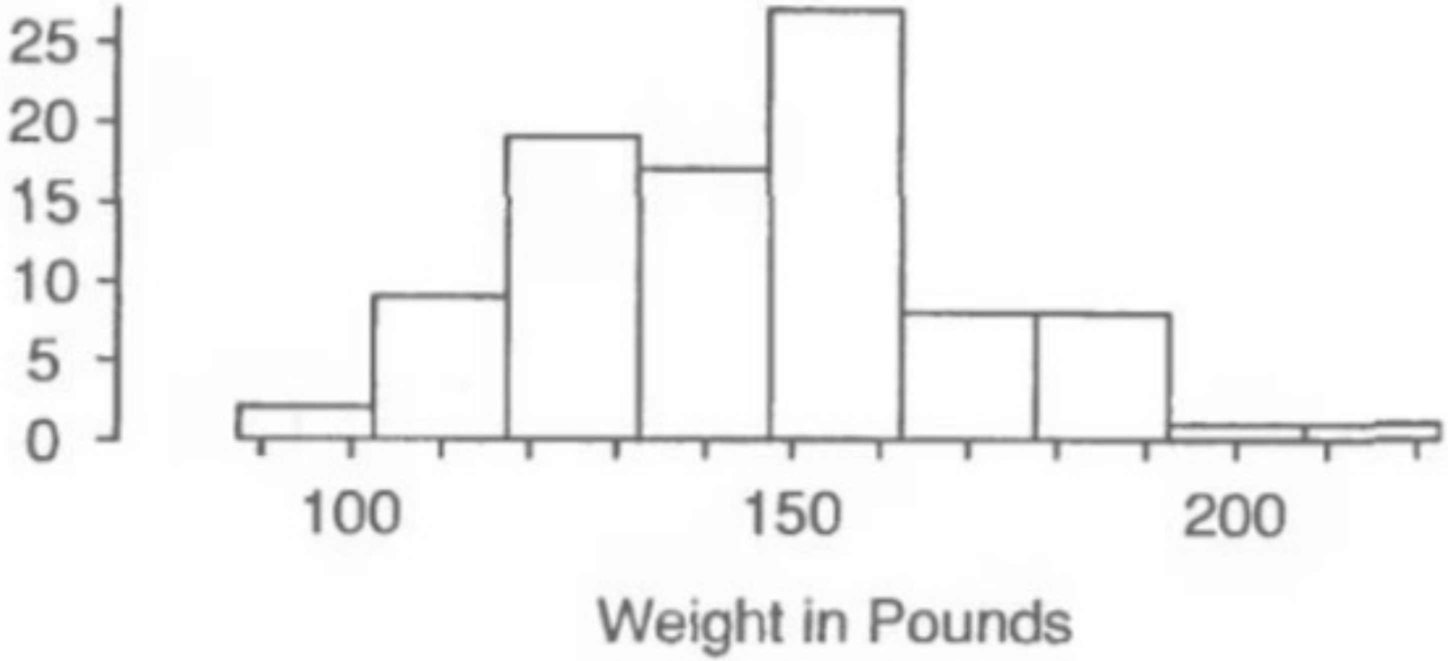
A way to better describe the data set and also remove the bias of the students is to summarize the data in term of frequency and relative frequency:



CLASS INTERVAL	MIDPOINT	FREQUENCY	RELATIVE FREQUENCY
87.5-102.4	95	2	.022
102.5-117.5	110	9	.098
117.5-132.4	125	19	.206
132.5-147.4	140	17	.185
147.5-162.4	155	27	.293
162.5-177.4	170	8	.087
177.5-192.4	185	8	.087
192.5-207.5	200	1	.011
207.5-222.4	215	1	.011
TOTAL		92	1.000

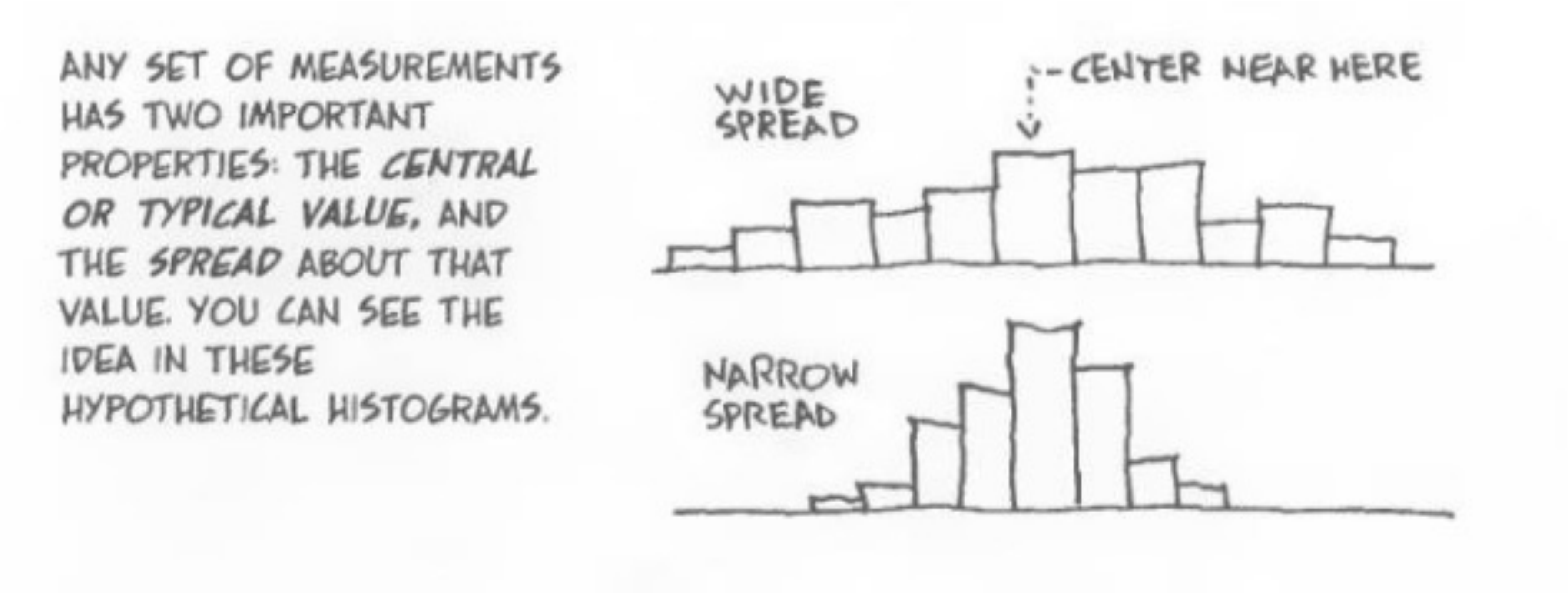
Descriptive statistics: data analysis

and plot the resulting data in form of histogram



The histogram give has a distribution of weights that for example can allow us to learn about the phenomena and eventually help to formulate a model to predict outcomes.

Descriptive statistics: data analysis



ANY SET OF MEASUREMENTS HAS TWO IMPORTANT PROPERTIES: THE *CENTRAL OR TYPICAL VALUE*, AND THE *SPREAD* ABOUT THAT VALUE. YOU CAN SEE THE IDEA IN THESE HYPOTHETICAL HISTOGRAMS.

We are going to lean how to measure these properties and how to predict that the data follow a certain trend in the distribution.

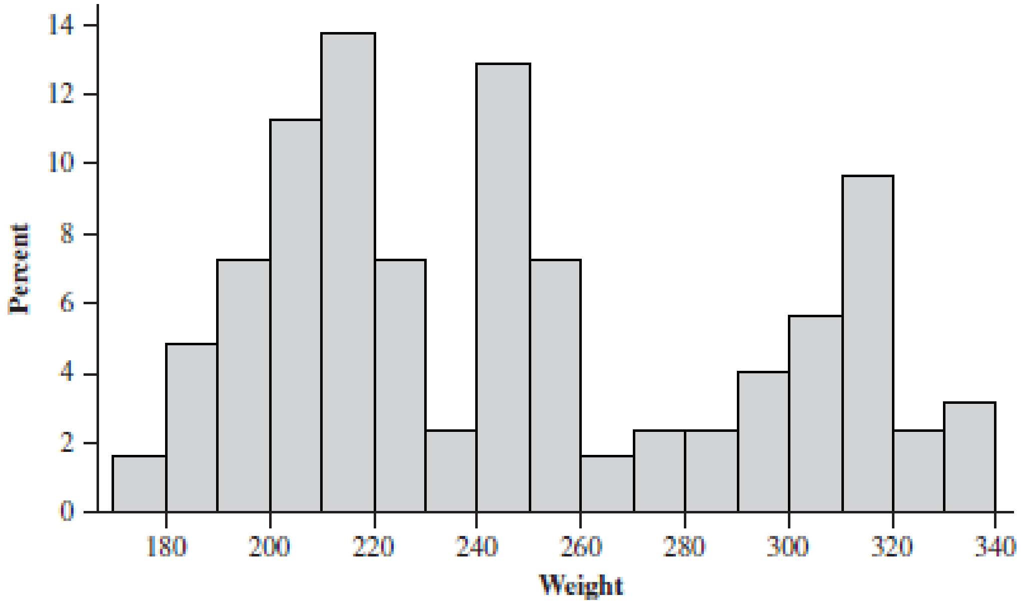
Histogram Shapes

A histogram with more than two peaks is said to be **multimodal.**

However, the number of peaks may well depend on the choice of class intervals, particularly with a small number of observations. The larger the number of classes, the more likely it is that bimodality or multimodality will manifest itself.

Histogram Shapes: Example 1.12

The following histogram shows a histogram of the weights (lb) of the 124 players listed on the rosters of the San Francisco 49ers and the New England Patriots (teams the author would like to see meet in the Super Bowl) as of Nov. 20, 2009.

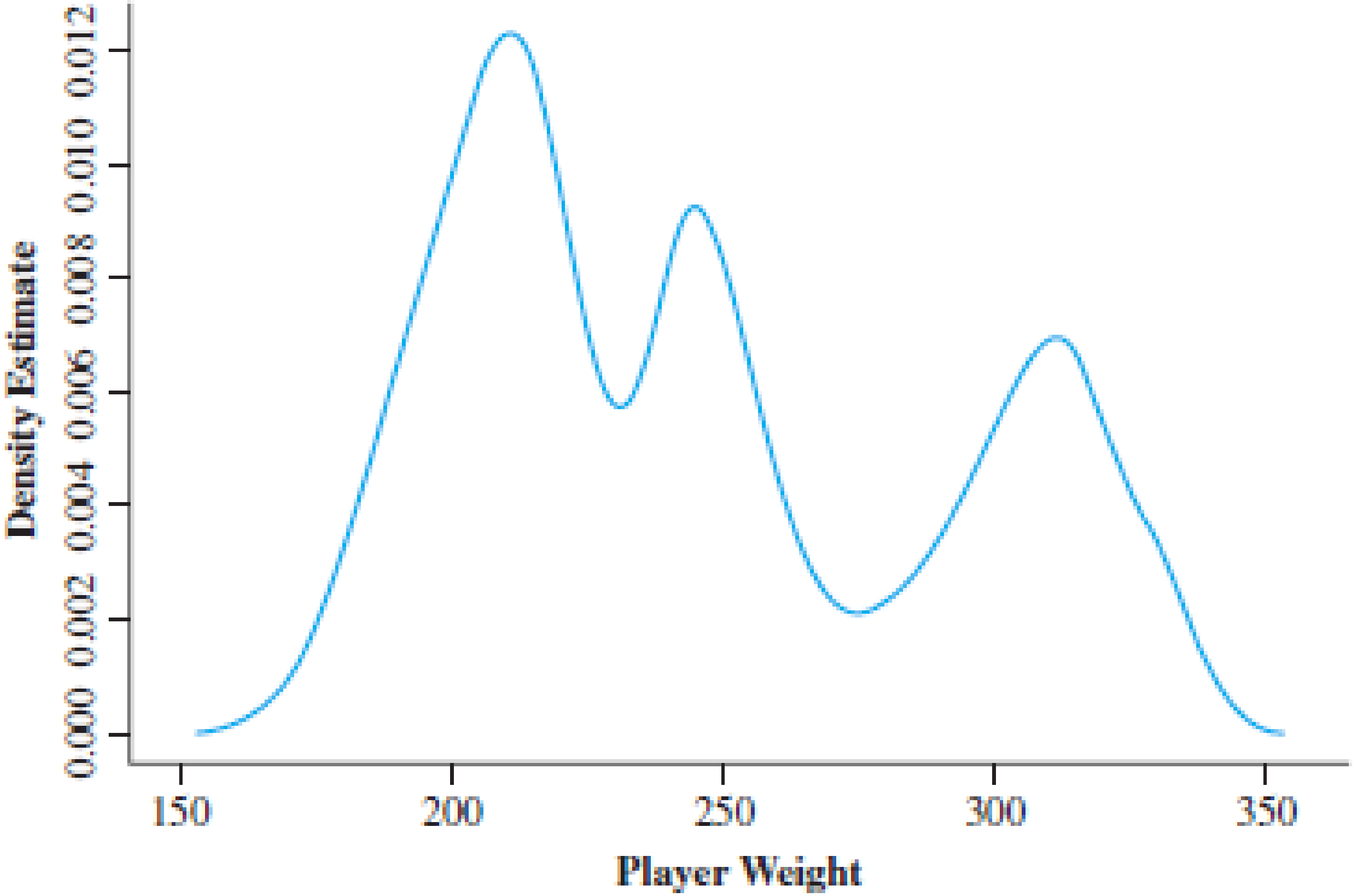


NFL player weights Histogram

Figure 1.11(a)

Histogram Shapes: Example 1.12

The following graph is a smoothed histogram (actually what is called a *density estimate*) of the data.



Histogram Shapes: Example 1.12

- Both the histogram and the smoothed histogram show three distinct peaks; the one on the right is for linemen, the middle peak corresponds to linebacker weights, and the peak on the left is for all other players (wide receivers,
quarterbacks, etc.).

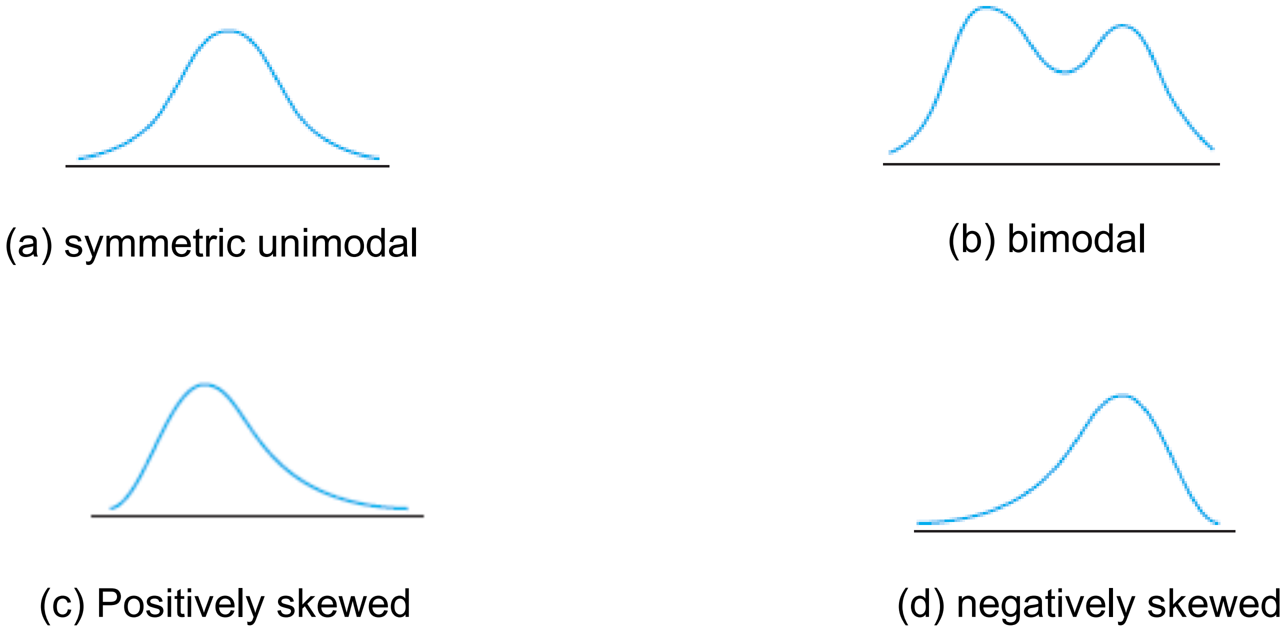
- A histogram is **symmetric** if the left half is a mirror image of the right half. A unimodal histogram is **positively skewed** if the right or upper tail is stretched out compared with the left or lower tail and **negatively skewed** if the stretching is to the left.

Histogram Shapes: Example 1.12

- A histogram is **symmetric** if the left half is a mirror image of the right half.
- A unimodal histogram is **positively skewed** if the right or upper tail is stretched out compared with the left or lower tail and
- **negatively skewed** if the stretching is to the left.

Histogram Shapes: Example 1.12

Figure 1.12 shows “smoothed” histograms, obtained by superimposing a smooth curve on the rectangles, that illustrate the various possibilities.



(a) symmetric unimodal

(b) bimodal

(c) Positively skewed

(d) negatively skewed

Smoothed histograms

Figure 1.12

Inferential statistics/data modelling

- Having obtained a sample from a population, an investigator would frequently like to use sample information to draw some type of conclusion (make an inference of some sort) about the population.

- Techniques for generalizing from a sample to a population are gathered within the branch of our discipline called **inferential statistics.**

Inferential statistics/data modelling

The techniques of **inferential statistics** bases their foundation **on the theory of probability.**

We need to build up a model of the data that allow us to make predictions or quantify a statistical property for the population that is represented by its analyzed sample.

Therefore, we need to learn the foundation of the probability theory to introduce these tools.

AN INTRODUCTION TO THE PROBABILITY THEORY

NOTHING IN LIFE IS CERTAIN. IN EVERYTHING WE DO, WE GAUGE THE CHANCES OF SUCCESSFUL OUTCOMES, FROM BUSINESS TO MEDICINE TO THE WEATHER. BUT FOR MOST OF HUMAN HISTORY, ***PROBABILITY***, THE FORMAL STUDY OF THE LAWS OF CHANCE, WAS USED FOR ONLY ONE THING: ***GAMBLING***.

From: Gonick & Smith The cartoon guide to Statistics

AN INTRODUCTION TO THE PROBABILITY THEORY

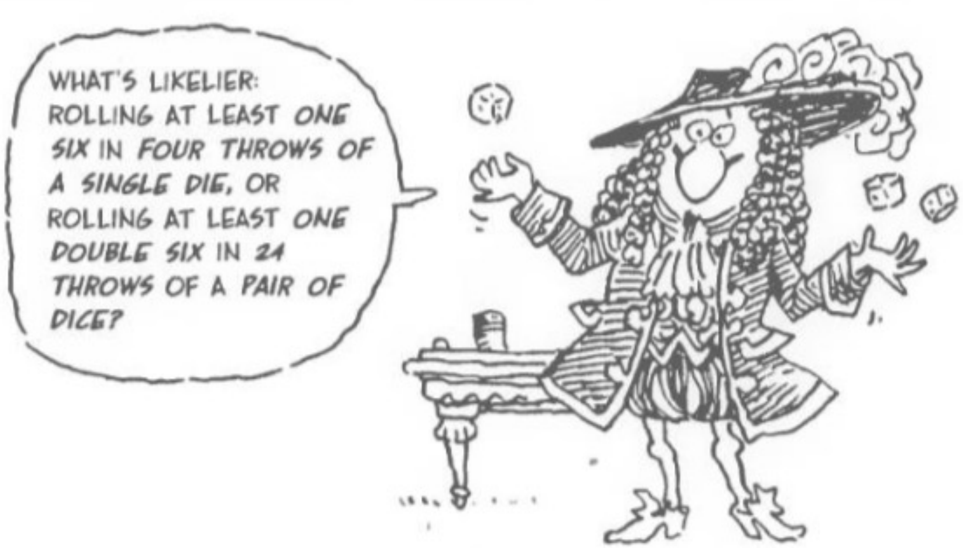
In this part we are going to learn the axioms of probability and connects them to counting problems –including combinatorial and permutations.

Learning outcomes:

- Sample spaces
- Outcomes
- Events
- Frequentist paradigm of probability
- The axioms of probability
- Calculate the probabilities of events occurring in simple sample spaces with equally likely outcomes.

BASIC DEFINITIONS

The theory of probability was developed with the aim to explain a gambling game governed by the law of chance.


WHAT'S LIKELIER: ROLLING AT LEAST ONE SIX IN FOUR THROWS OF A SINGLE DIE, OR ROLLING AT LEAST ONE DOUBLE SIX IN 24 THROWS OF A PAIR OF DICE?

From: Gonick & Smith The cartoon guide to Statistics

In the 18 century, Antoine Gombaud, Chevalier de Méré writer, gambler and friend of two famous Mathematician:

Blaise Pascal and Pierre Fermat

whose work laid the foundations of the Modern Probability theory

PROBABILITY THEORY: A BIT OF HISTORICAL BACKGROUND

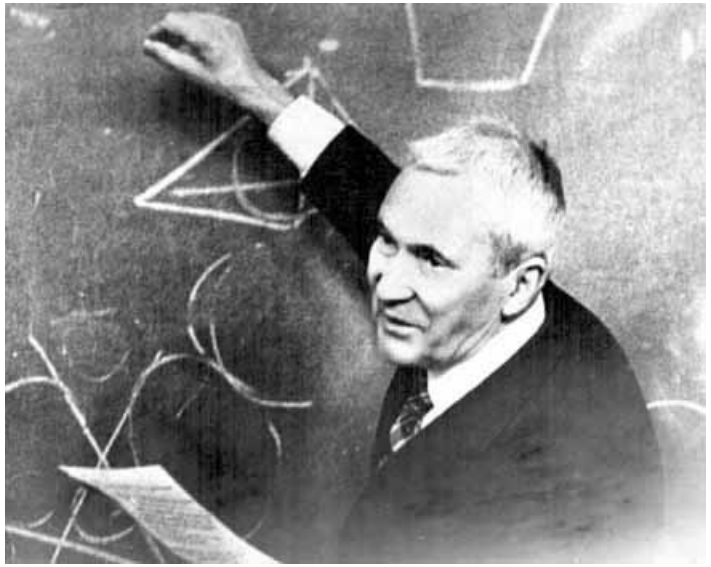
Christiaan Huygens
(1629-1695)

Rev. Thomas Bayes
(1701-1761)

Pierre-Simon Laplace
(1749-1827)

Andrey Markov
(1856-1922)

Modern probability theory is based on an axiomatic description of the properties of probability formulated by the Russian mathematician A. N. Kolmogorov (1903-1987).



http://enacademic.com/dic.nsf/enwiki/54484

What is a probability?

Two possible definitions

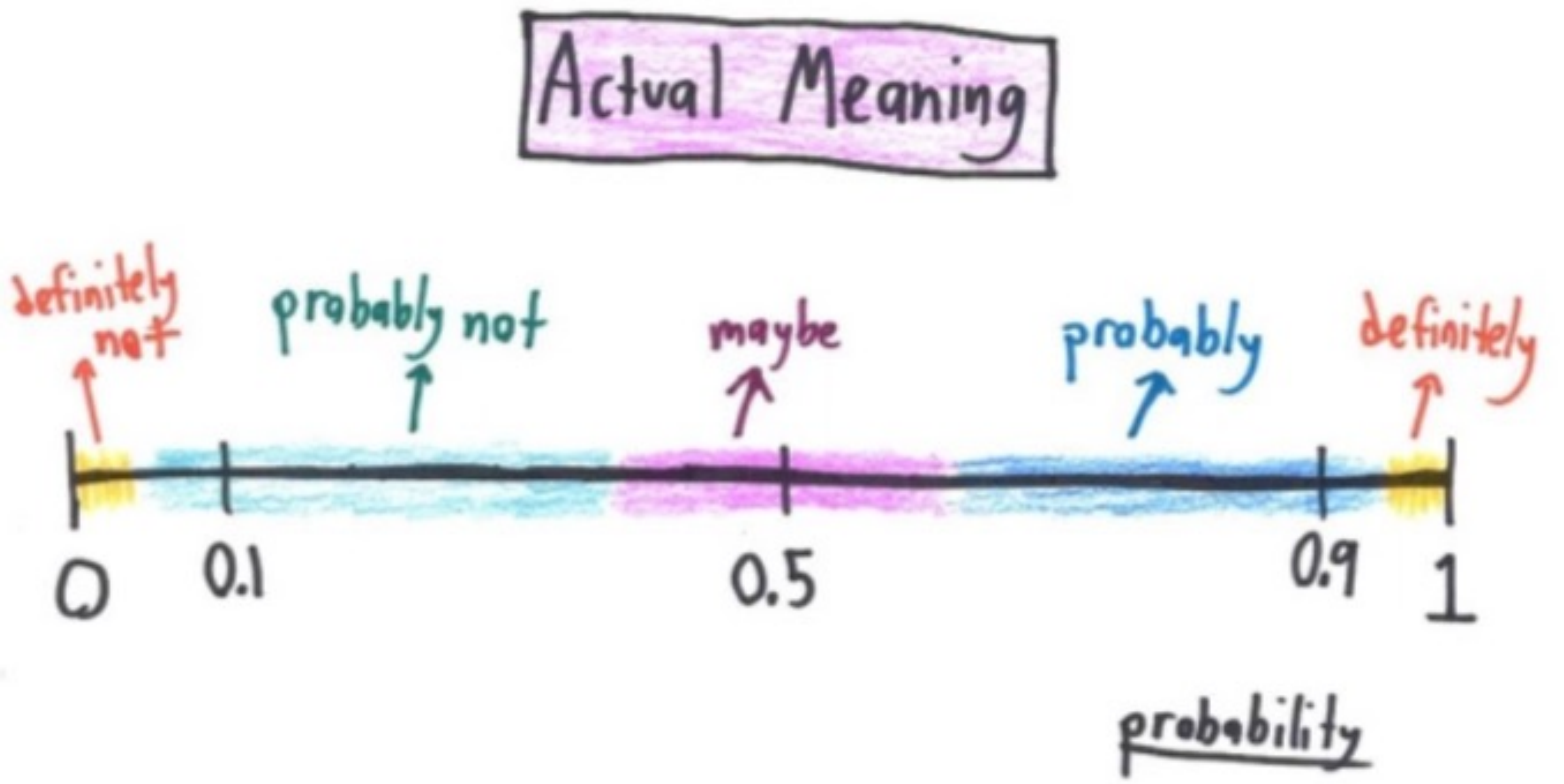
Objective: *the probability is obtained by repeating random experiments and counting the number of times something occurs:*

$$P(A) = p = \lim_{N \to \infty} \frac{N_A}{N}$$

Subjective (Bayseian): Many things can be described by a probability, but it is not possible to repeat the process many times and sample how often the event occurs. In this case the probability is interpreted as reasonable expectation, representing a state of knowledge, or as quantification of a personal belief.

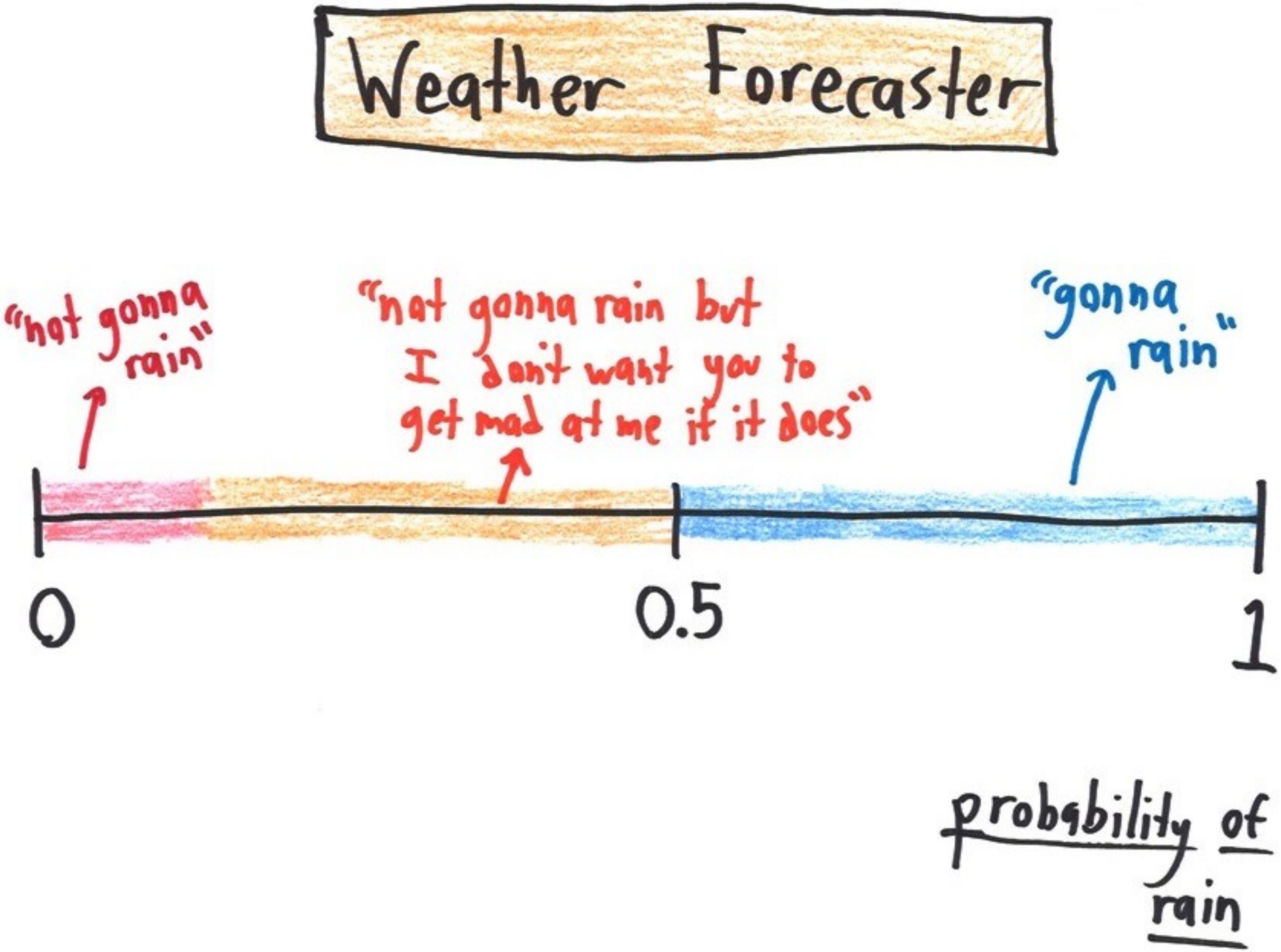
Bayesian methods are widely accepted and used in artificial intelligence in the field of machine learning, computer vision.

What is a probability?



https://mathwithbaddrawings.com/2015/09/23/what-does-probability-mean-in-your-profession/

What is a probability?



https://mathwithbaddrawings.com/2015/09/23/what-does-probability-mean-in-your-profession/

BASIC DEFINITIONS

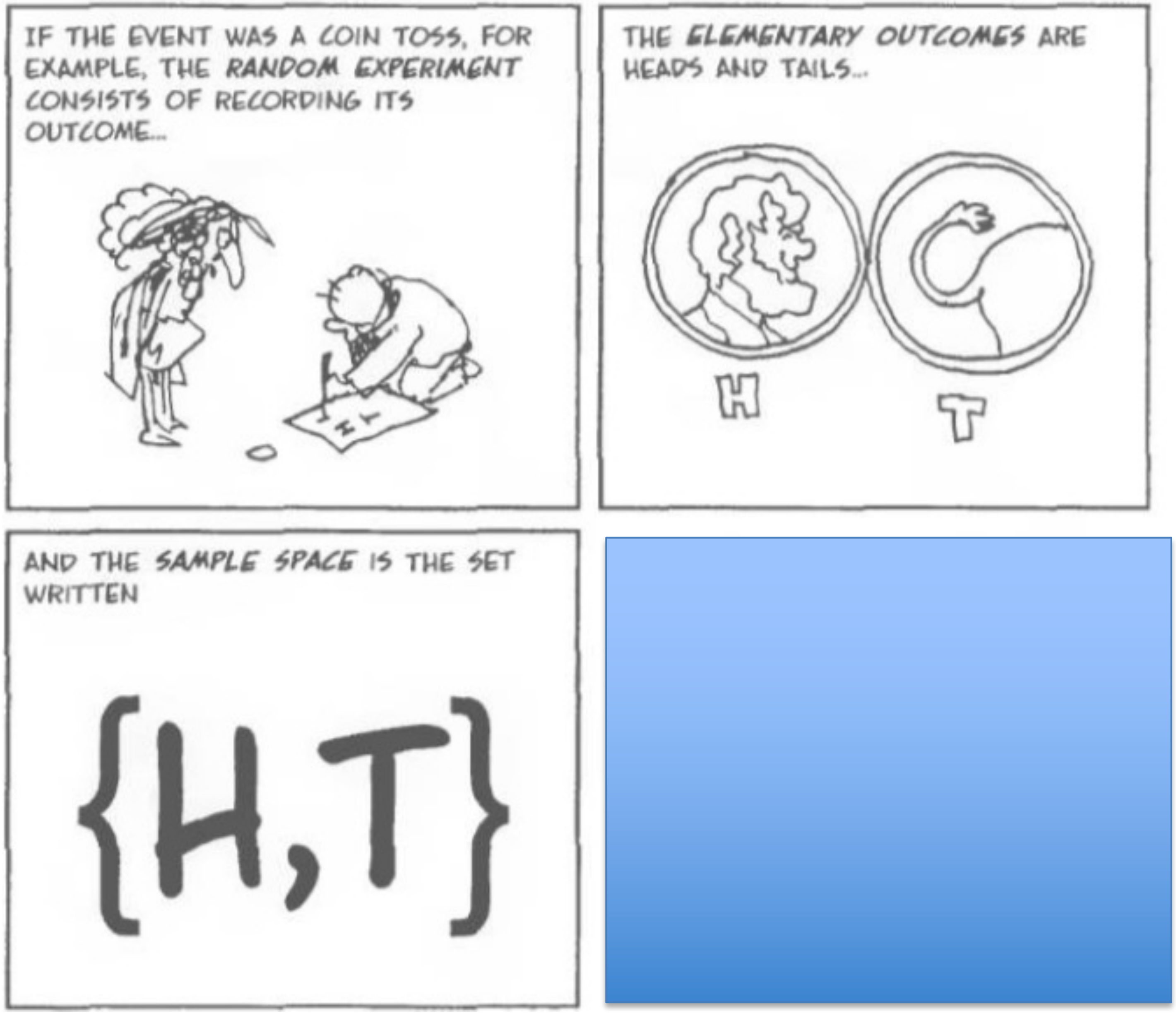
As scientist, we call the game of the gambler a **random experiment**: in general is a process of observing the outcome of a chance event (ex. the rolling of a dice, the radioactive decay, the fluctuation of the stock market).

The **elementary outcomes** are all possible results of the random experiment.

The **sample space** is the set or collection of all the elementary outcome.

SAMPLE SPACES

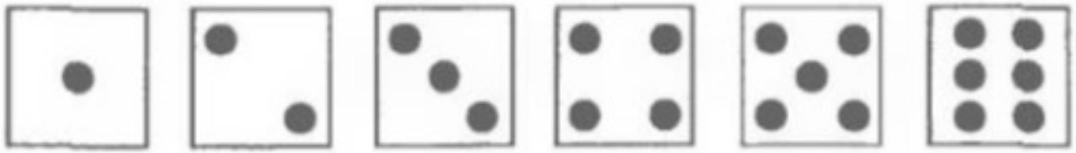
Example *What is the sample space of tossing a coin?*


IF THE EVENT WAS A COIN TOSS, FOR EXAMPLE, THE RANDOM EXPERIMENT CONSISTS OF RECORDING ITS OUTCOME...
THE ELEMENTARY OUTCOMES ARE HEADS AND TAILS...
H
T
AND THE SAMPLE SPACE IS THE SET WRITTEN
{H,T}

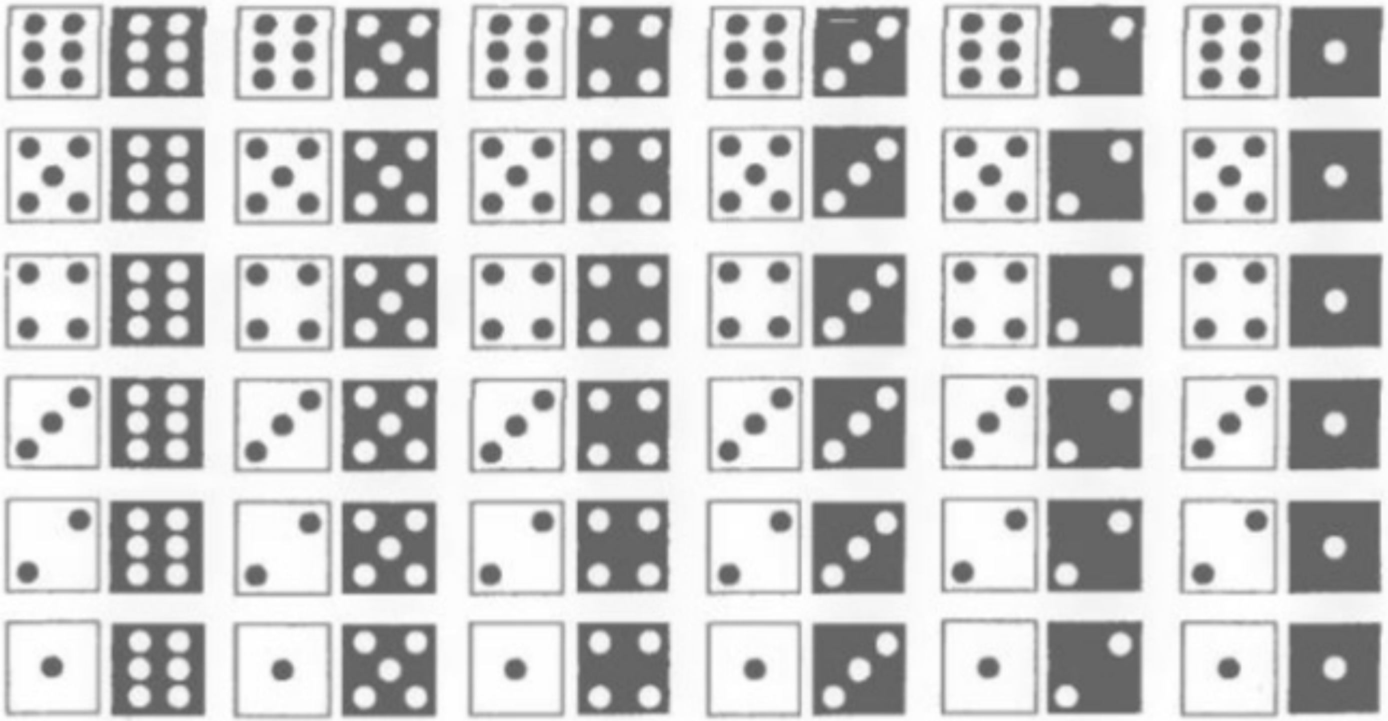
From: Gonick & Smith The cartoon guide to Statistics

BASIC DEFINITIONS

THE SAMPLE SPACE OF THE THROW OF A *SINGLE DIE* IS A LITTLE BIGGER.



AND FOR A *PAIR* OF DICE, THE SAMPLE SPACE LOOKS LIKE THIS (WE MAKE ONE DIE WHITE AND ONE BLACK TO TELL THEM APART):



SAMPLE SPACES

Note that sample spaces are not unique, sometimes there will be more than one way to break down and list the possible outcomes.

Example: An 'experiment' consists of flipping two coins? Give two possible sample spaces.

Define a single coin flip

$$F= \{H, T\}.$$

Then we have

$$S = \{F \times F\} = \{(H, H), (H, T), (T, H), (T, T)\}$$

SAMPLE SPACES

Alternatively, we could have three outcomes - 0 heads come up between the two flips, 1 head or 2 heads

S = {0, 1, 2}

Now our element {0} is identical to {T , T } previously.
But

{1}= {(H, T), (T , H)}

{2}= {(H, H)}

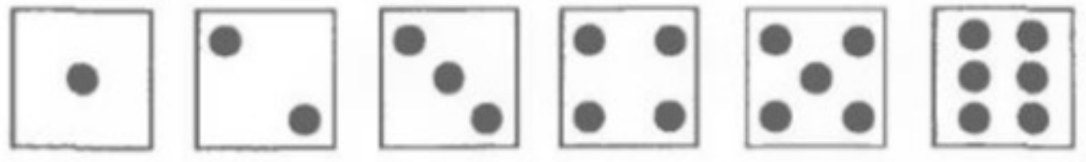
OUTCOMES

Definition: *Outcomes are the individual elementary elements of the sample space.*
(If the sample space is finite, or countably in finite, it will correspond to a discrete random variable.)

Example

A die is rolled once. We let X be the outcome of the experiment. The sample space for the experiment is the 6-element set

S = {1, 2, 3, 4, 5, 6},



and each outcome i , for *i= 1…6*, corresponds to the number of dots on the face that faces up.

EVENTS

Definition: an event is any subset, E , of the sample space, S. $E \subset A$.

(Recall that a set A is a subset of set B , A $\subseteq B$, if all of the members of A are also contained in B)

Example what is the event, E , corresponding to a die coming up with an even number?

E = {2, 4, 6}

EVENTS

Example: *Consider a road race with the runners labelled 1 to 4. What is the event A that runner 1 won the race?*

$S = \{\text{all orderings of } (1{,}2{,}3{,}4)\} = \{(x_1, x_2, x_3, x_4 : x_i = 1, 2, 3, 4, x_i \neq x_j)\}$

$A = 1234, 1243, 1324, 1342, 1423, 1432 = \text{all ordering of } (1234) \text{ beginning with } 1$
$= \{(x_1, x_2, x_3, x_4 : x_1 = 1, x_i = 2, 3, 4, x_i \neq x_j)\}$

note that $A \subset S$

What is the event that runner 2 lost the race?

$B = 1342, 1432, 3142, 3412, 4132, 4312 = \text{all ordering of } (1234) \text{ ending with } 2$
$= \{(x_1, x_2, x_3, x_4 : x_4 = 2, x_i = 2, 3, 4, x_i \neq x_j)\}$

note again that $B \subset S$