

## What do Scientists do with Data?

These are the steps a scientist would go through to analyze and display their data.

**1) Gather Data** (Field trip!): You may have a question in mind when you go out to gather data, or you may just be gathering data that many scientists can use to answer lots of different questions. Make sure you write the date and location on every data sheet, and fill it out as completely and as clearly as possible, being as specific as you can. Your class collected data on phytoplankton. We will be analyzing and graphing this data with you.

**2) Copy the data from the data sheets to a spreadsheet such as Excel, Numbers, or GeoGebra.** You will then have tables of data like you see in the examples on the next page. When you are done, protect the original data sheets. Keep them filed and never let them leave the lab!

**3) Analyze the data.** This information is now organized in a way that you can start analyzing it: you can find totals, averages, means, minimums, and maximums using the spreadsheet software, and enter these values in columns or rows in your spreadsheet. These calculations will be useful in answering many types of questions.

*We will be going over how to do steps two and three in class.*

**4) Ask a question.** Here are some questions you might ask about the data sets on page 3:

- a.) Is the air temperature related to height?
- b.) Are wind speed, height and air pressure related?
- c.) How did the wind speed of Hurricane Katrina compare to the wind speeds of other hurricanes in 2005?
- d.) Have the winning times of the Boston Marathon been improving over time?
- e.) What percentage of Boston marathon winners between 1985 and 2000 were from Kenya?

Can you think of any other questions you can answer with this data?

---

---

There are ways to display the data that can show information much more clearly than just a list of numbers can. Graphs, plots, and charts can show information about relationships in the data that you can see right away instead of comparing numbers in your head. But in order to know what type of chart or graph or plot to use, you have to know what type of question you are asking.

**5) Figure out whether your question is about comparison, relationship, distribution, or composition.**

**Relationship:** How strongly are two variables are related to each other?

Over time: If one of the variables is time, this question is about how the other variable **changes through time.**

**Comparison:** A question that **compares a summary value** (i.e. the mean, median, or sum) of two or more groups

**Time series:** Comparing how values for multiple categories change over time.

**Distribution:** Asks how often each value or outcome shows up for one variable, Or **compares the variability** of groups (i.e. the range, shape of the distribution, and measure of center).

**Composition:** How much of the total does each sub-group make up?

For example, in step 4, question a is are relationship questions, questions b and c are comparison questions, question d is a relationship over time question, and question e is about composition. What types of questions are the ones you added?

Weather Balloon Ascent Data:

Time (s)	Height (m)	Pressure (millibars)	Air Temp (DegC)	Water density (g/cm)	Wind speed (m/s)
-33	227	985.1	11.2	9.31	4.6
18.4	271	980	11.8	9.01	7.7
42.9	400	965	11.4	9.23	12
59.4	487	955	11.4	9.47	20.4
112.8	754	925	11.9	9.73	26
149.5	938	905	12	8.8	28.9
208.7	1220	875	11.7	6.14	25.3
269.5	1512	845	11.8	3.09	18.8
309.3	1712	825	10.1	2.98	16.5
444	2283	770	9.1	1.26	13.8
496.2	2500	750	9	0.9	16.8
546.2	2723	730	8.2	0.84	14.9
691.7	3305	680	4.9	0.82	11.1
790.6	3671	650	2.8	0.61	13.3
985.5	4448	590	-1.7	0.92	16.2
1084.8	4859	560	-5.7	0.68	16.2
1284.2	5660	505	-11	0.11	12
1402	6128	475	-13.4	0.04	11.7
1539.3	6706	440	-17.6	0.39	14.6
1716.5	7412	400	-23.1	0.11	17.2
2514.3	10679	250	-47.1	0.01	24

2005 Hurricane Max Wind Speed:

Year	Name	Max wind speed (knots)
2005	Irene	90
2005	Katrina	150
2005	Cindy	65
2005	Dennis	130
2005	Emily	140
2005	Epsilon	75
2005	Stan	70
2005	Vince	65
2005	Wilma	160
2005	Beta	100
2005	Maria	100
2005	Nate	80
2005	Ophelia	75
2005	Philippe	70
2005	Rita	155

(another weather balloon data set from NASA: [http://mynasadata.larc.nasa.gov/docs/sonde\\_20070731.txt](http://mynasadata.larc.nasa.gov/docs/sonde_20070731.txt))

Boston Marathon Winning Times:

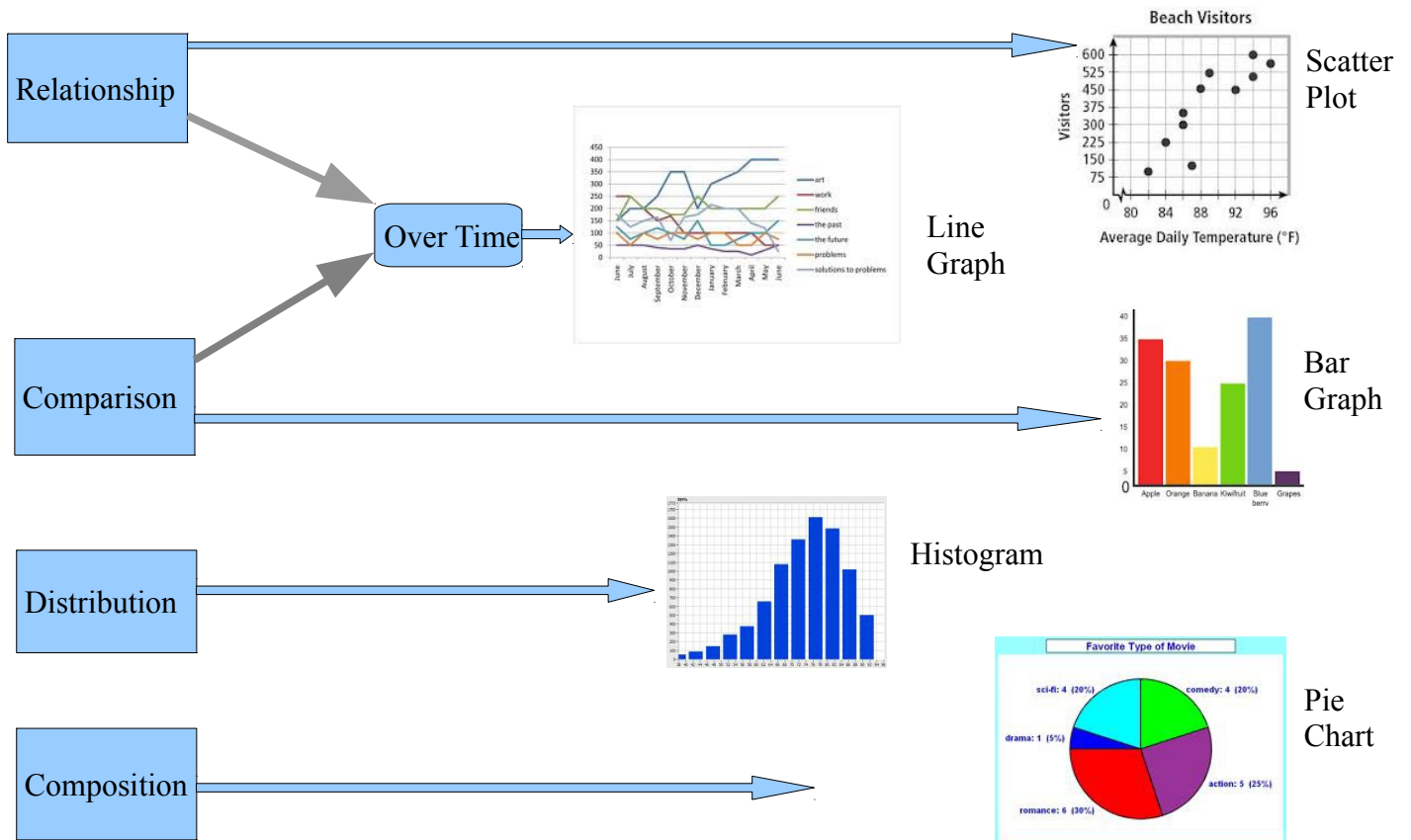
year	name	home	time
1985	Geoff Smith	Great Britain	2:14:05
1986	Robert de Castella	Australia	2:07:51
1987	Toshihiko Seko	Japan	2:11:50
1988	Ibrahim Hussein	Kenya	2:08:43
1989	Abebe Mekonnen	Ethiopia	2:09:06
1990	Gelindo Bordin	Italy	2:08:19
1991	Ibrahim Hussein	Kenya	2:11:06
1992	Ibrahim Hussein	Kenya	2:08:14
1993	Cosmas Ndeti	Kenya	2:09:33
1994	Cosmas Ndeti	Kenya	2:07:15
1995	Cosmas Ndeti	Kenya	2:09:22
1996	Moses Tanui	Kenya	2:09:15
1997	Lameck Aguta	Kenya	2:10:34
1998	Moses Tanui	Kenya	2:07:34
1999	Joseph Chebet	Kenya	2:09:52
2000	Elijah Lagat	Kenya	2:09:47

**Practice: What types of questions are these?** Enter the letter of the type of question in the spaces:

- a) comparison b) comparison: time series c) relationship d) relationship over time e) distribution  
f) not a statistical question

- Which car manufacturer makes the most fuel-efficient vehicles, Chevrolet or Toyota? \_\_\_\_\_
- Has Frenchman Bay warmed over the last 10 years? \_\_\_\_\_
- Which beam shape supports the most weight? \_\_\_\_\_
- What percentage of your family's expenses goes toward food? \_\_\_\_\_
- How tall is this tree? \_\_\_\_\_
- What is the most common height of tree in our yard? \_\_\_\_\_

**6) Choose the type of graph to use.** Once you know the type of question, it is easy to choose a graph:



**Which kind of graph would you use to display data for each of the questions listed under step 4?**

- Is the air temperature related to height? \_\_\_\_\_
- Are wind speed, height and air pressure related? \_\_\_\_\_
- How did the wind speed of Hurricane Katrina compare to the wind speeds of other hurricanes in 2005?  
\_\_\_\_\_
- Have the winning times of the Boston Marathon been improving over time? \_\_\_\_\_
- What percentage of Boston marathon winners between 1985 and 2000 were from Kenya? \_\_\_\_\_

**BONUS: Use graph paper to make a graph for one of these questions!** (We will be doing this in class with the phytoplankton data you collected)