Gathering Data

Discussion

Unstructured data is very hard to work with and many times impractical. Imagine a bunch of post-it notes stuck to your desk each telling you a different assignment and due date for school assignments. It would be nice to order and structure this data so that you could prioritize each assignment. One way to structure data is to store it in a table using rows and columns.

Class Exercise

Answer the following questions based on the table shown.

first_name	last_name	phone	email	street	city	state	zip_code
Debra	Burks	NaN	debra.burks@yahoo.com	9273 Thorne Ave.	Orchard Park	NY	14127
Kasha	Todd	NaN	kasha.todd@yahoo.com	910 Vine Strret	Campbell	CA	95008
Tameka	Fisher	NaN	tameka.fisher@aol.com	758C Honey Creek St.	Redondo Beach	CA	90278
Daryl	Spence	NaN	daryl.spence@aol.com	988 Pearl Lane	Uniondale	NY	11553
Charolette	Rice	(555) 920-2829	charolette.rice@msn.com	107 River Dr.	Sacramento	CA	95820
Lyndsey	Bean	NaN	lyndsey.bean@hotmail.com	786 West Road	Fairport	NY	14450
Latasha	Hays	(555) 539-5920	latasha.hays@hotmail.com	7012 Manor Station Rd.	Buffalo	NY	14215
Jacquline	Duncan	NaN	jacqueline.duncan@yahoo.com	15 Brown St.	Jackson Heights	NY	11372
Genoveva	Baldwin	NaN	genoveva.baldwin@msn.com	8550 Spruce Drive	Port Washington	NY	11050
Pamelia	Newman	NaN	pamelia.newman@gmail.com	475 Chestnut Ave.	Monroe	NY	10950

1. How many rows of data are there? How many columns?

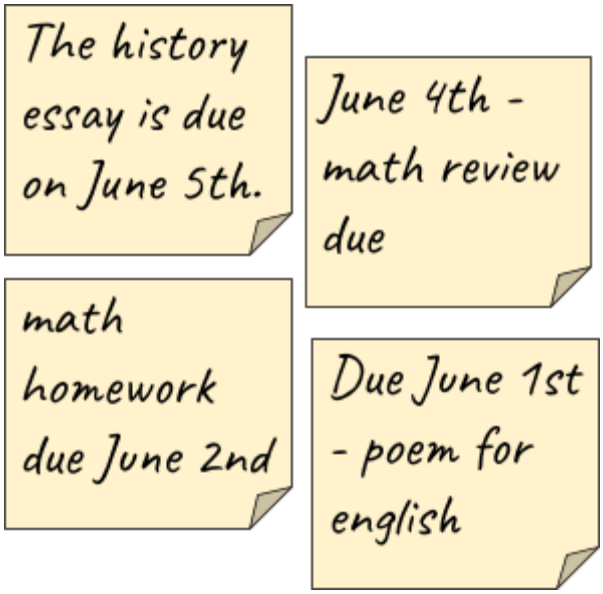
2. What might be a good name/title for this table?

3. What is listed in the 4th row, 3rd column?

4. What is listed in the 9th row, 2nd column?

5. What do you suppose NaN might mean?

Create a table for the following unstructured data.



6. Why did you choose the columns (variables) that you did?

7. How could you sort the data to help you prioritize your tasks?

Research or find the nutrition facts labels for 3 foods or drinks that you regularly consume. Use the information to create a table of usable rows and columns.

8. Which columns in your table include quantitative data (if any)?

9. How could you collect or include more qualitative data?

10. Formulate at least TWO statistical questions that you could use this data to answer.