



Lecture 5

Building Tables

Slides created by John DeNero (denero@berkeley.edu) and Ani Adhikari (adhikari@berkeley.edu)

Announcements

Review: Arrays

An array contains a sequence of values

- All elements of an array should have the same type
 - Arithmetic is applied to each element individually
 - When two arrays are added, they must have the same size; corresponding elements are added in the result
 - A column of a table is an array
-

Ranges

Ranges

A range is an array of consecutive numbers

- `np.arange(end)`:
An array of increasing integers from 0 up to `end`
- `np.arange(start, end)`:
An array of increasing integers from `start` up to `end`
- `np.arange(start, end, step)`:
A range with `step` between consecutive values

The range always includes `start` but excludes `end`

Tables

Ways to create a table

- `Table.read_table(filename)` - reads a table from a spreadsheet
 - `Table()` - an empty table
 - and... `select`, `where`, `sort` and so on all create new tables
-

Example

Charles Joseph Minard, 1781-1870



- French civil engineer who created one of the greatest graphs of all time
 - Visualized Napoleon's 1812 invasion of Russia, including
 - the number of soldiers
 - the direction of the march
 - the latitude and longitude of each city
 - the temperature on the return journey
 - Dates in November and December
-

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont été en Russie, le noir ceux qui en sont sortis. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Fézensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow et qui avaient rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

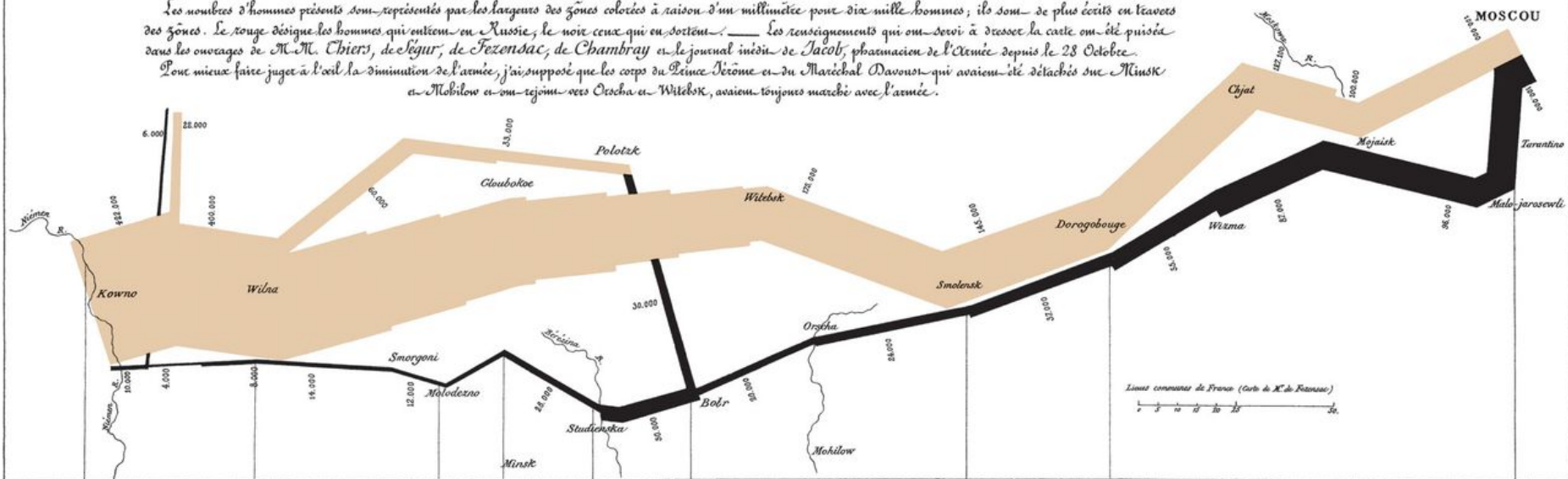
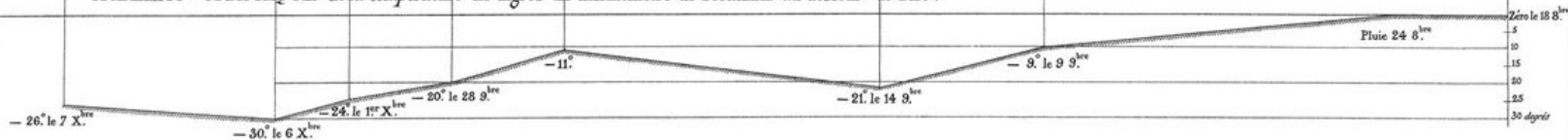


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Some of Minard's Data

Longitude	Latitude	City	Direction	Survivors
32	54.8	Smolensk	Advance	145000
33.2	54.9	Dorogobouge	Advance	140000
34.4	55.5	Chjat	Advance	127100
37.6	55.8	Moscou	Advance	100000
34.3	55.2	Wixma	Retreat	55000
32	54.6	Smolensk	Retreat	24000
30.4	54.4	Orscha	Retreat	20000
26.8	54.3	Moiodexno	Retreat	12000

(Demo)

Table Methods

- Creating and extending tables:
 - `Table().with_column` and `Table.read_table`
- Finding the size: `num_rows` and `num_columns`
- Referring to columns: labels, relabeling, and indices
 - `labels` and `reabeled`; column indices start at 0
- Accessing data in a column
 - `column` takes a label or index and returns an array
- Using array methods to work with data in columns
 - `item`, `sum`, `min`, `max`, and so on
- Creating new tables containing some of the original columns:
 - `select`, `drop`

(Demo)

Manipulating Rows

- `t.sort(column)` sorts the rows in increasing order
 - `t.take(row_numbers)` keeps the numbered rows
 - Each `row` has an index, starting at 0
 - `t.where(column, are.condition)` keeps all rows for which a column's value satisfies a condition
 - `t.where(column, value)` keeps all rows containing a certain value in a column
-