

Data Processing

College 4 - 2018-11-22

Infographics should be....



Functional as a hammer

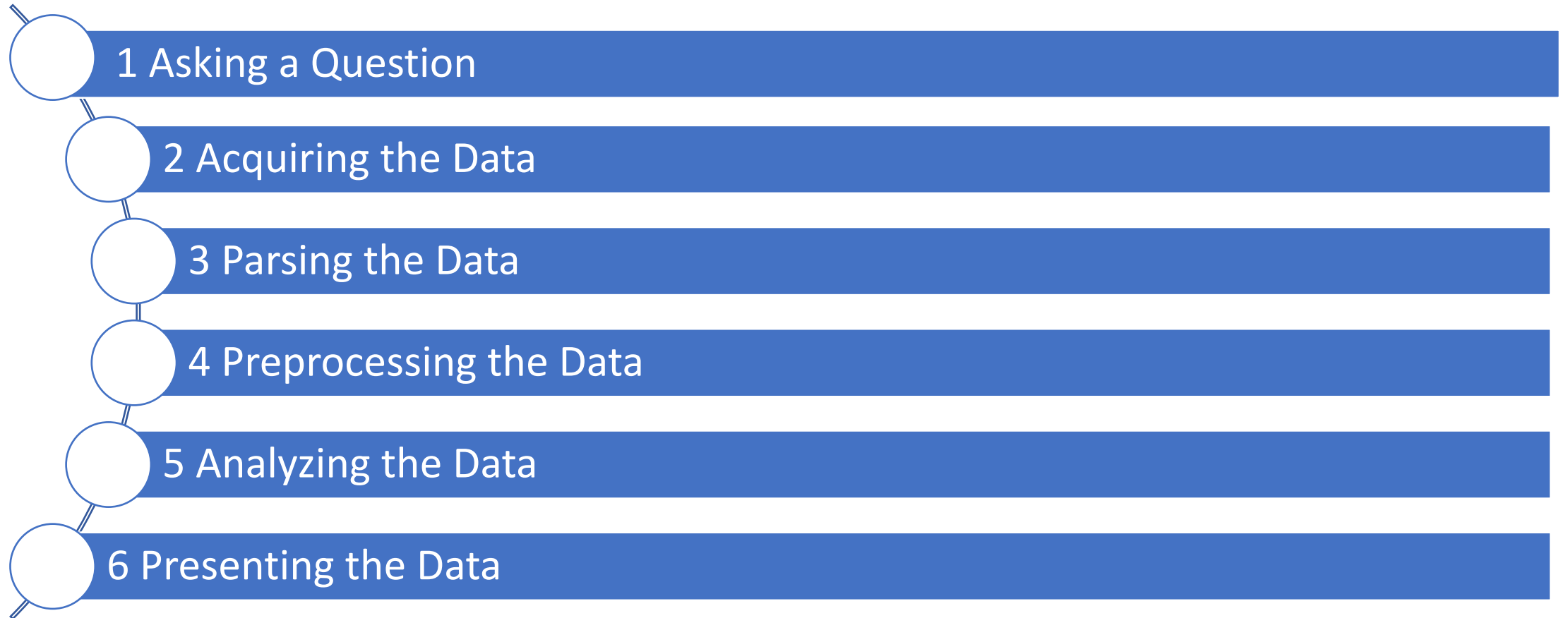


Multilayered as an onion














$$E = mc^2$$

Beautiful as equations

Data Processing Pipeline



Opdrachten

no	Assignment	Programming Languages	deadline
1	Scraping IMDB		Wed 7-Nov
2	Exploratory Data Analysis		Wed 14-Nov
3	JavaScript line		Wed 21-Nov
4	D3 bar chart	 	Tue 27-Nov
5	D3 scatterplot	 	Mon 3-Dec
6	Linked views-interim	  	Mon 10-Dec
7	Linked views	  	Mon 17-Dec

→ 17:00 (!)

Late day wildcard

Dit houdt in dat je bij **één** van de opdrachten van Data Processing **eenmalig** de opdracht een dag later in mag leveren.

Vermeldt in je code of in de **commit message** dat je de late day wildcard gebruikt, zodat de nakijker weet dat het nagekeken moet worden ondanks dat het een dag te laat is.

Aandachtspunten

- Opdrachten bouwen op --> kijk wat je kunt gebruiken uit vorige opdracht
- Begin op tijd (maakt hergebruik ook makkelijker)

Proposal

#Problem statement

#Solution

#Prerequisites

Deadline

Vandaag: donderdag (22 november, 2018) 23:59

<https://data.mprog.nl/project/requirements>

<https://data.mprog.nl/project/proposal>

Proposal

[_write in Markdown_]

#Problem statement

- a) There is a clearly defined problem that a reasonably large group of people have, which an app or visualization can solve particularly well.
- b) There is a widespread lack of knowledge or understanding that an interactive visualization is particularly suited to remedy.

Brief introduction to Markdown



2004 Markdown project
John Gruber
<https://daringfireball.net>

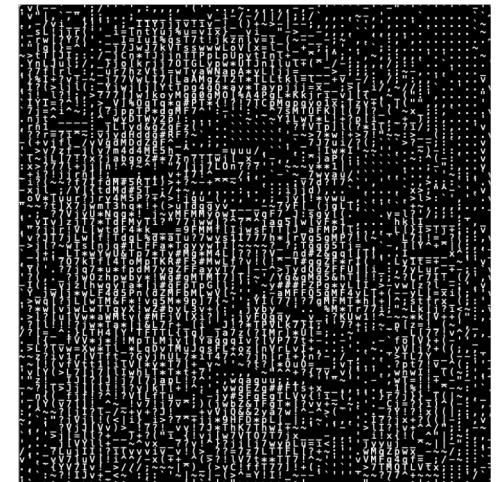
Markdown allows you to write using an easy-to-read, easy-to-write plain text format, then convert it to structurally valid XHTML (or HTML).

Specification:

<https://daringfireball.net/projects/markdown/syntax>

Further standardized as [CommonMark](#)

by John MacFarlane



Small demo

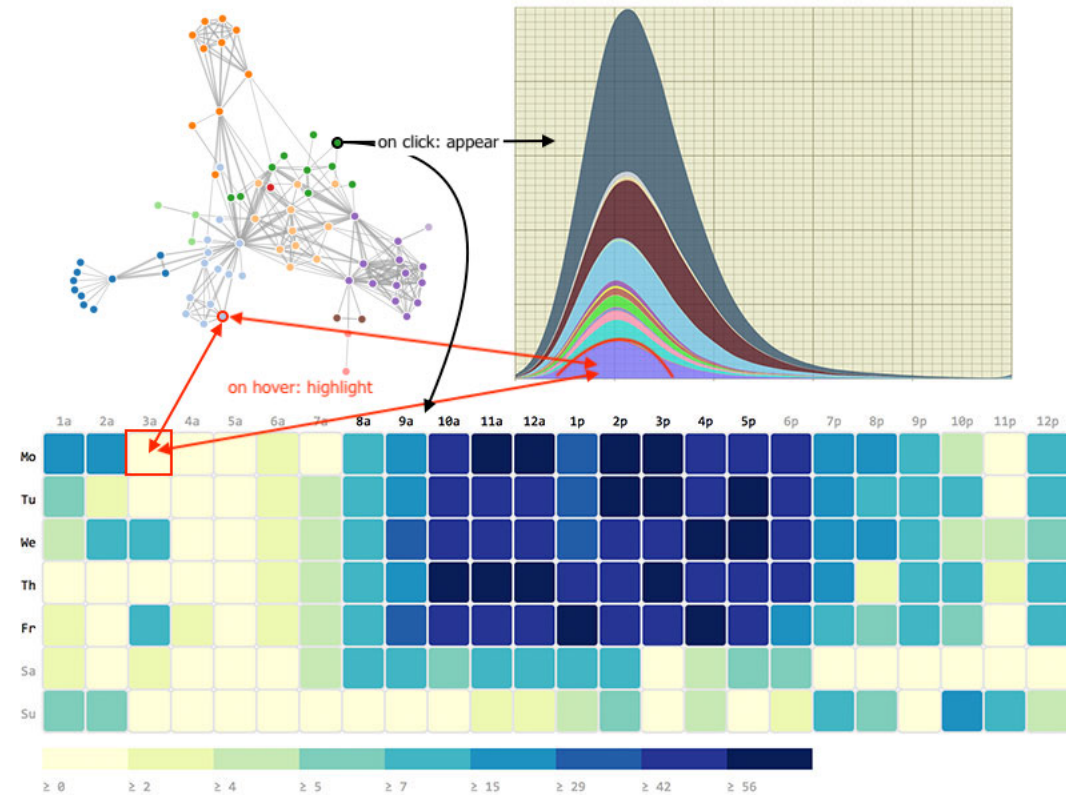
Quick Markdown resources

- Brief tutorial: <https://www.markdowntutorial.com>
- Examples: <https://markdown-it.github.io>

Proposal

#Solution

- Summarize your idea in a single sentence, connecting it to the “gap” that you describe.
- Include a ***visual sketch***
- List ***main features***
- Split the features into 1) the *_minimum viable product_* (MVP) and 2) parts that are optional to implement.



Proposal

#Prerequisites

- List the ***data sources***, *_links_*, and required *_preprocessing_*.
- List the ***external components*** (libraries like d3-tip or SQLite) that you need to implement certain features. Include the names, and if the component is not standard, include a link to its website.
- Include a review ***related visualizations***, in terms of features and technical aspects: what do they do? how have they implemented it? Can you do it in the same way?
- Identify the ***hardest parts*** of implementing your application: think of technical problems or limitations that could arise during development and what possibilities you have to overcome these.

Proposal

#Problem statement

#Solution

#Prerequisites

Deadline

Vandaag: donderdag (22 november, 2018) 23:59

<https://data.mprog.nl/project/requirements>

<https://data.mprog.nl/project/proposal>

Google launches new search engine to help scientists find the datasets they need

Dataset Search could be a scientist's best friend

By **James Vincent** | @jjvincent | Sep 5, 2018, 12:00pm EDT

<https://www.theverge.com/2018/9/5/17822562/google-dataset-search-service-scholar-scientific-journal-open-data-access>

<https://toolbox.google.com/datasetsearch>