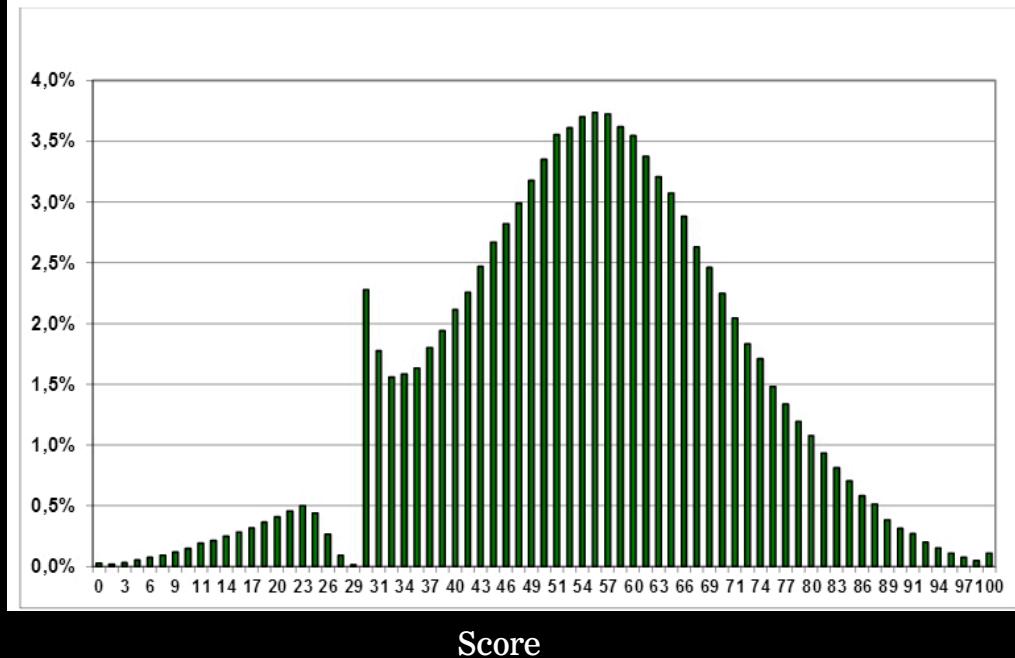


Stuff I learned the hard way in being a
data scientist.

“Human learning” before machine learning

Distribution of results of the Matura (high school exit exam) in Poland

2.1. Poziom podstawowy



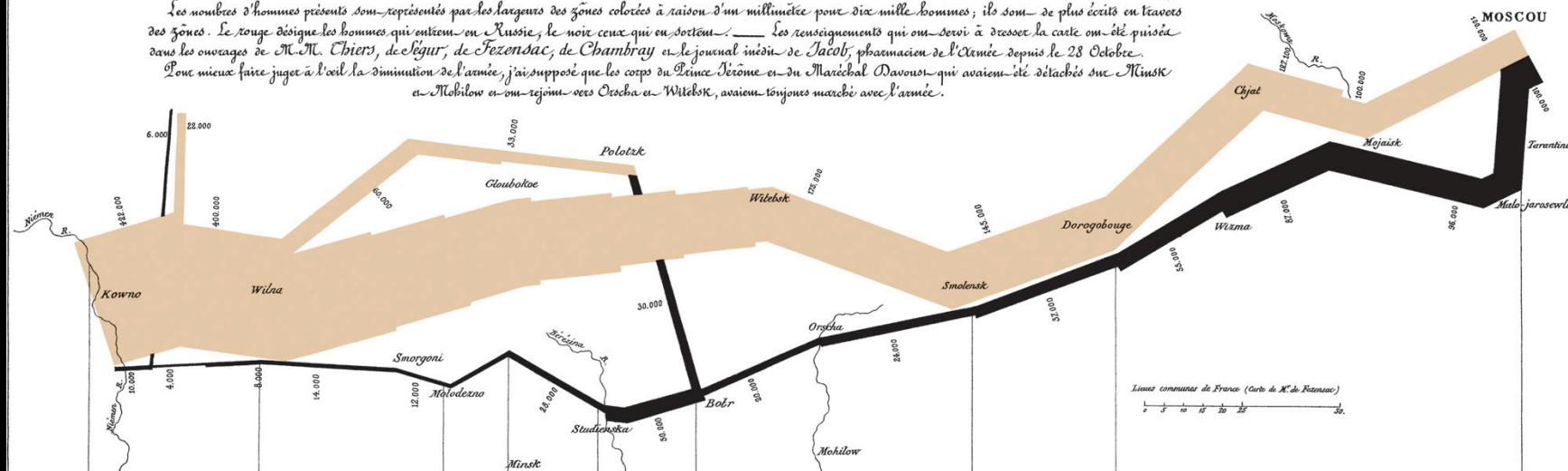
GA GENERAL ASSEMBLY

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 à 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 travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. Chier, de Séjourné, de Fezensac, de Chambray et le journal médical 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 Napoléon et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow et se rejoignirent vers Orsha et Wileïsk, avaient toujours marché avec l'armée.

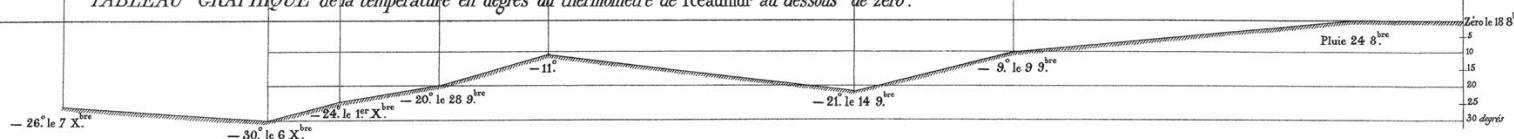


Lieux communs de France (Carte de M. de Fezensac)

1 5 10 15 20 25 30

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

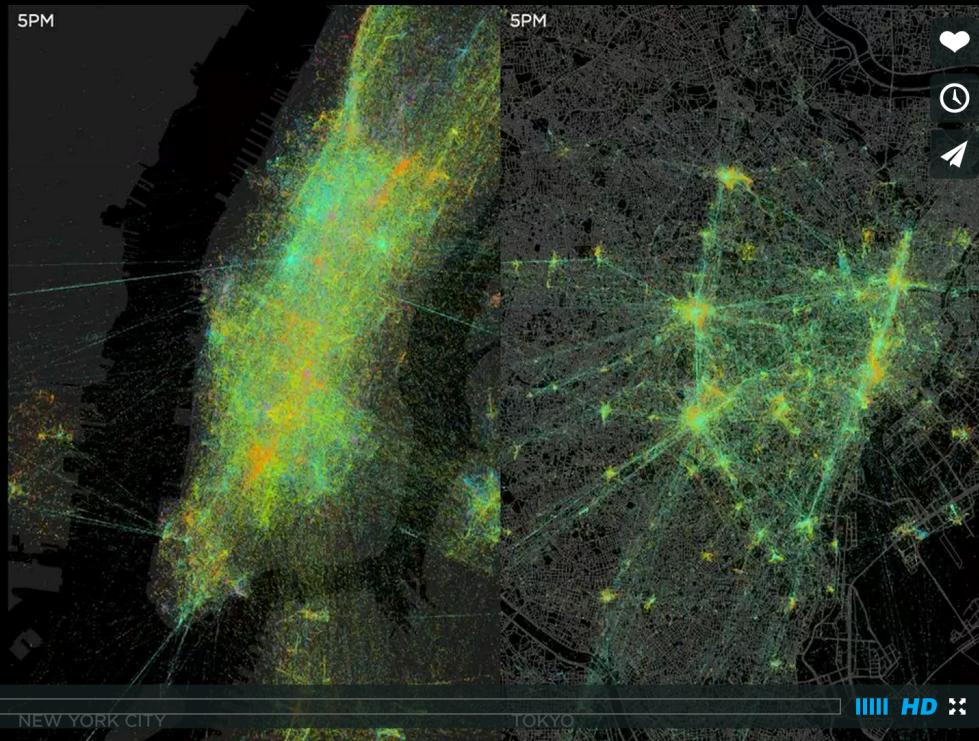
Les cosaques passent au galop
le Niemen gelé.



GA GENERAL ASSEMBLY

foursquare

- RESIDENCE
- FOOD
- ARTS & ENTERTAINMENT
- COLLEGE & UNIVERSITY
- NIGHTLIFE SPOT
- GREAT OUTDOORS
- SHOP & SERVICE
- PROFESSIONAL & OTHER
- TRAVEL & TRANSPORT



<https://vimeo.com/62289901>

Understand the needs of various domains that exist within your organization. You'll need to interact with many of them.

Establish consistent and coherent vocabulary, definitions, and ontologies — both verbally and in code.

There's a difference between building
a model and deploying a model to a
production environment.

Keep it simple - always

“Don’t repeat yourself” also means don’t do the same thing twice

The breadth and depth of material is overwhelming to all of us, *don't sweat the small stuff.*

Data scientists come from a variety of backgrounds and skillsets. *The only requirement is a genuine intellectual curiosity and thirst for solving interesting problems.*

Community:

- Kaggle.com
- Data Tau (<http://www.datatau.com/>)
- reddit.com/r/dataisbeautiful

Learning:

- Data Science Open Source Master's <http://datasciencemasters.org/>
- Harvard CS 109 course: <http://cs109.github.io/2014/>
- Data Science Trello board: <https://trello.com/b/rbpEfMld/data-science>
- Data Science Glossary github.com/rasbt/pattern_classification/blob/master/resources/data_glossary.md
- dataschool.io
- [Coursera.com](https://www.coursera.com)
- [udacity.com](https://www.udacity.com)