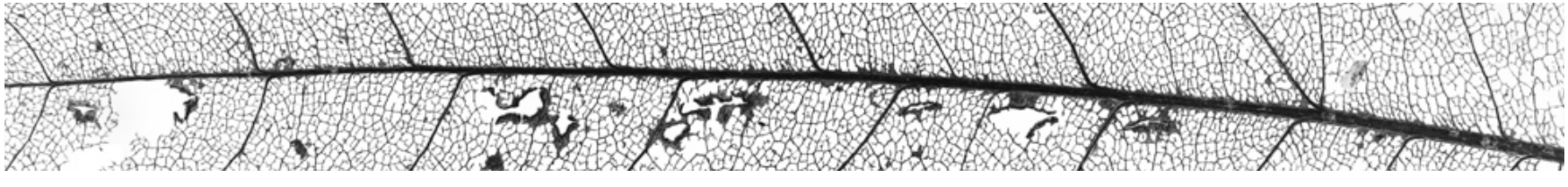
Des archives du commerce à des données quantifiables

une longue chaîne de transformation des données

Paul Girard

Sciences Po, médialab

Collecter et produire des données pour la recherche en SHS
Axe 2 : Collecter des données pour les exploiter : comment les préparer en amont ?
Fréjus, le 16 novembre 2016



SciencesPo
MÉDIALAB

- Conception de méthodes numériques
- Hybridant les approches qualitatives et quantitatives
- Développant des outils-logiciels
- En Sciences Humaines et Sociales

Des archives du commerce à des données quantifiables

une longue chaîne de transformation des données

- Retour d'expérience des projets:
  - RICardo
  - TOFLIT18
- Des archives aux données ?
- Quelles bases de données ?
- L'exploration visuelle au service des données

Latour, Bruno. 1993. *'Le Topofil de Boa-Vista. La Référence Scientifique: Montage Photophilosophique'*. Raisons Pratiques 4: 187–216.

RICardo · XIXème siècle

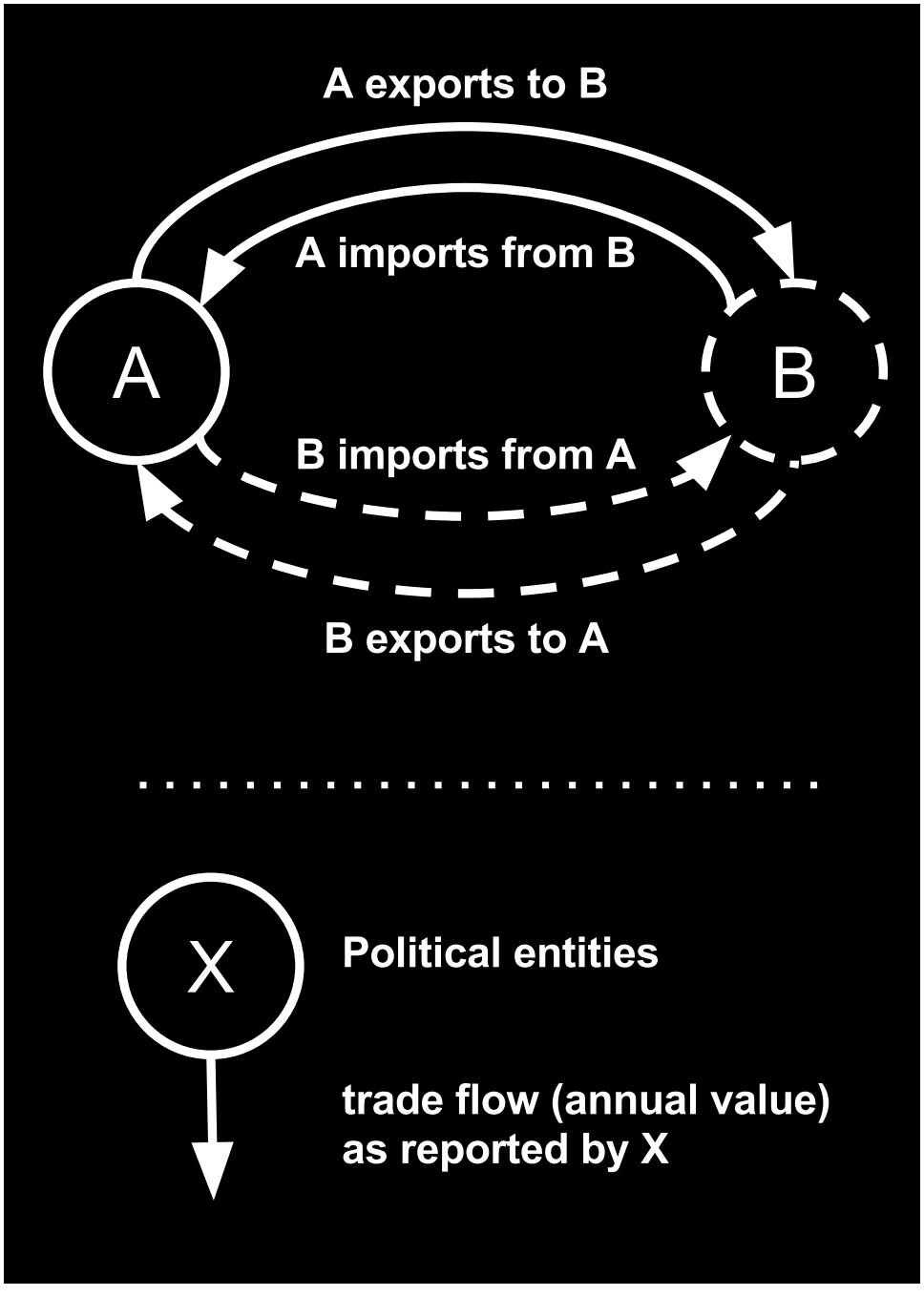
données annuelles

1787 · · · · · · · · · · · · · · · 1938

RICardo · commerce bilateral

Flux de commerce entre ~~pays~~ entités

A <··> B


A exports to B
A imports from B
A
B
B imports from A
B exports to A
X
Political entities
trade flow (annual value)
as reported by X

RICardo · commerce total

Flux de commerce entre les entités A et le monde

A <··> M

ANGLETERRE.

La navigation de Liverpool a donné lieu, en 1833, au mouvement ci-après :

Liverpool.

Mouvement commercial en 1833.

PROVENANCES ET DESTINATIONS.			ENTRÉE.		SORTIE.	
			Navires.	Tonneaux.	Navires.	Tonneaux.
Pays étrangers		États-Unis	562	224,220	408	162,000
		Brésil	93	23,550	138	34,500
		Hollande	249	27,930	261	29,430
		Belgique	14	2,896	15	2,700
		Prusse	7	3,500	9	2,700
		Danemarck	78	15,600	125	24,600
		Russie	68	21,200	84	13,275
		Villes anséatiques	60	11,668	59	9,350
		Deux-Siciles	83	12,190	38	6,655
		Suède et Norwège	42	10,500	24	4,200
		Espagne	57	6,840	36	6,480
		Portugal	60	7,200	34	5,940
		Indes-Occidentales	20	5,248	20	6,100
		France	45	6,109	39	4,831
		Plata (États-Unis de la)	20	3,000	33	7,095
		Mexique	24	3,600	33	5,775
		Sardaigne	18	2,121	50	6,275
		Colombie	18	2,700	12	4,800
		Toscane	6	1,200	32	5,760
		Autriche	20	3,000	24	3,600
		Chili	12	3,000	12	3,000
		Turquie	25	2,750	24	3,000
		Autres contrées	43	10,020	54	12,075
Possessions	Amérique	continentale	345	120,750	321	112,350
		Antilles	230	46,000	201	55,275
	Afrique		48	14,409	51	17,650
	Indes-Orientales		51	12,160	48	11,000

TOFLIT18: commerce de la France

Flux de commerce entre la France et ses partenaires commerciaux

France <··> A,B,C

rapportés par l'état français

TOFLIT18 : les sources

Cabas de palme	Espagne		1515 .
Cabarets	Hollande		38 .
Cabillaud	Danemark	93408 . 8	391531 . 6
	Hollande	297966 . 18	
	Nord	156 .	
Cables	Hollande		800 .
Cacao	Espagne	22979 . 16	839625 . 19
	Guinée	1730 .	
	Hollande	500337 . 13	
	Isles	28687 . 10	
	Italie	92757 .	
	Nord	18751 .	
	Portugal	166183 .	
	Savoye	5200 .	
Cachou	Hollande		757 . 10

TOFLIT18: les produits

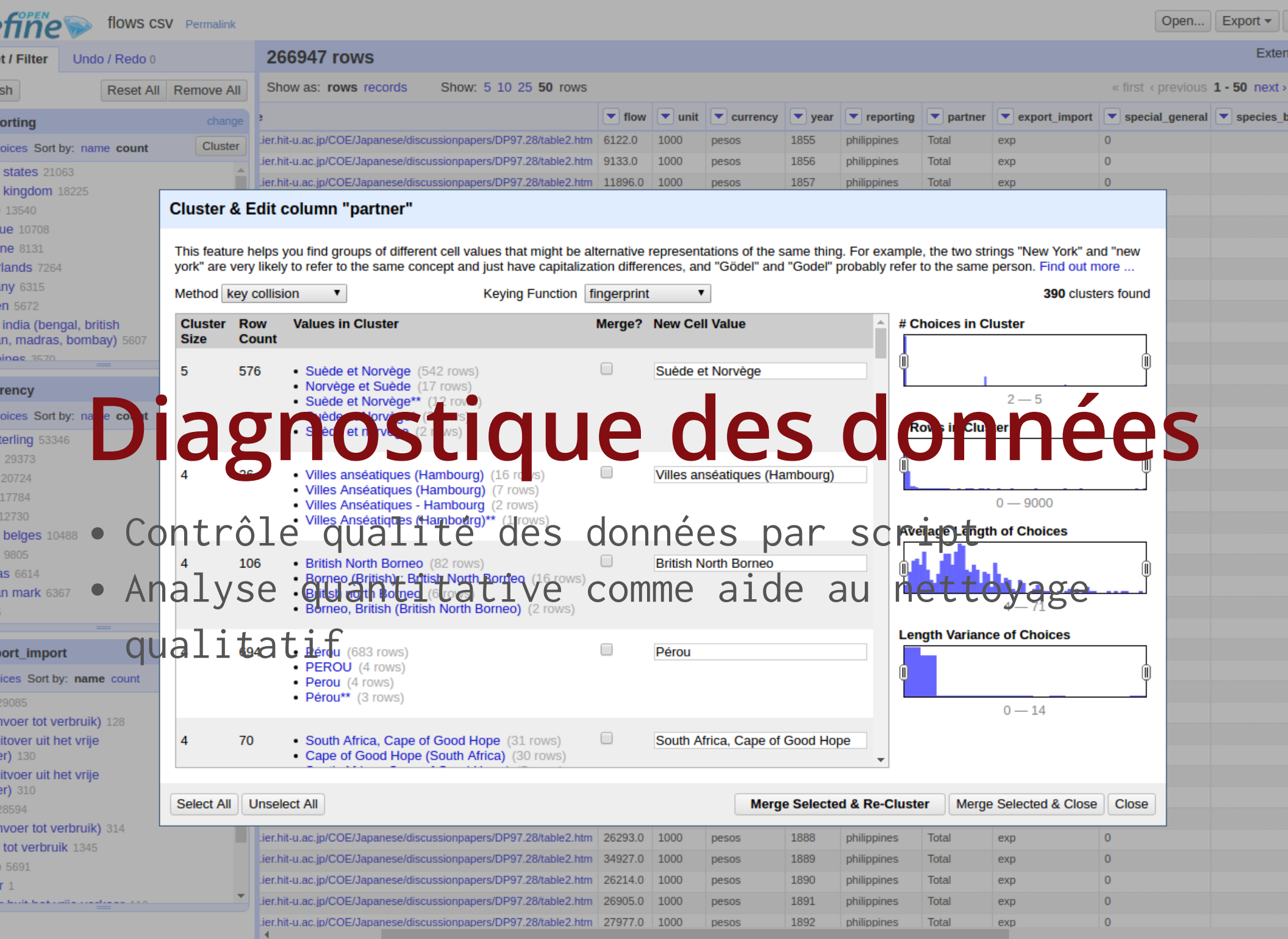
Top 50 des produits

Articles réunis · Indigo · Mercerie · Eau de vie · Librairie · Vinaigre · Cacao · Suif · Beurre · Alun · Liqueurs · Sel · Fromage · Rocou · Confitures · Acier · Fer ; en barres · Thé · Huile d'olive · Porcelaine · Poivre · Farine · Miel · Chandelle · Huile ; d'olive · Fayance · Savon · Amidon · Ris · Verdet · Bierre · Cochenille · Légumes · Bijouterie · Garance · Horlogerie · Chocolat · Meubles · Quinquina · Amandes · Crin · Papier ; blanc · Planches ; de sapin · Jambons · Lard · Drogues réunies · Argenterie · Bougie · Gaudron · Cuivre...

Des sources aux données

Volumes d'archives > images > **?** > chercheurs

Transcription manuelle



Diagnostique des données

- Contrôle qualité des données par script
- Analyse quantitative comme aide au nettoyage qualitatif

Rapport de test-données de Ricardo

```
# RICnames TEST
0 missing reporting in RICnames
0 missing partners in RICnames
missings written in out_data
# CURRENCY TEST
total number of currencies in flow 12366
check number before/after set currency : 12366/12366
check number before/after set modified_currency : 16382/16382
in currency not in flow 4016
in flow not in currency 0
in flow and in currency 12366
in flow in currency not in rate 1
total known currencies in flow 12365
missign rates exported in out_data
# EXP IMP TEST
missing expimp spe/gen in standards :0
EXP IMP TEST : OK
# FLOWS DUPLICATES TEST
## Spe/gen Dups
```

Année 1793 et 1794

MOIS de l'Ere Vulgaire	MOIS Germinal	MOIS de l'Ere Vulgaire	MOIS Floréal	MOIS de l'Ere Vulgaire	MOIS Prairial	MOIS de l'Ere Vulgaire	MOIS Messidor	MOIS de l'Ere Vulgaire	MOIS Thermidor	MOIS de l'Ere Vulgaire	MOIS Fructidor
Mars 1794	VII.e MOIS	Avril 1794	VIII.e MOIS	Mai 1794	IX.e MOIS	Juin 1794	X.e MOIS	Juill. 1794	XI.e MOIS	Août 1794	XII.e MOIS
v. 21	1 Primidi	D. 20	1 Primidi	D.Q. m. 20	1 Primidi	D.Q. j. 19	1 Primidi	D.Q. s. 19	1 Primidi	l. 18	1 Primidi
s. 22	2 Duodi	D.Q. l. 21	2 Duodi	le 21 m. 21	2 Duodi	le 20 v. 20	2 Duodi	le 19 D. 20	2 Duodi	m. 19	2 Duodi
D.Q. D. 23	3 Tridi	le 22 m. 22	3 Tridi	j. 22	3 Tridi	s. 21	3 Tridi	l. 21	3 Tridi	m. 20	3 Tridi
le 24 l. 24	4 Quartidi	m. 23	4 Quartidi	v. 23	4 Quartidi	D. 22	4 Quartidi	m. 22	4 Quartidi	j. 21	4 Quartidi
m. 25	5 Quintidi	j. 24	5 Quintidi	s. 24	5 Quintidi	l. 23	5 Quintidi	m. 23	5 Quintidi	v. 22	5 Quintidi
m. 26	6 Sextidi	v. 25	6 Sextidi	D. 25	6 Sextidi	m. 24	6 Sextidi	j. 24	6 Sextidi	s. 23	6 Sextidi
j. 27	7 Septidi	s. 26	7 Septidi	l. 26	7 Septidi	N.L. m. 25	7 Septidi	N.L. v. 25	7 Septidi	N.L. D. 24	7 Septidi
v. 28	8 Octidi	D. 27	8 Octidi	N.L. m. 27	8 Octidi	le 26 j. 26	8 Octidi	le 26 s. 26	8 Octidi	le 23 l. 25	8 Octidi
N.L. s. 29	9 Nonodi	N.L. l. 28	9 Nonodi	le 28 m. 28	9 Nonodi	v. 27	9 Nonodi	D. 27	9 Nonodi	m. 26	9 Nonodi
le 30 D. 30	10 Décadi	le 29 m. 29	10 Décadi	j. 29	10 Décadi	s. 28	10 Décadi	l. 28	10 Décadi	m. 27	10 Décadi
l. 31	11 Primidi	m. 30	11 Primidi	v. 30	11 Primidi	D. 29	11 Primidi	m. 29	11 Primidi	j. 28	11 Primidi
Avril m. 1	12 Duodi	Mai j. 1	12 Duodi	s. 31	12 Duodi	l. 30	12 Duodi	m. 30	12 Duodi	v. 29	12 Duodi
m. 2	13 Tridi	v. 2	13 Tridi	Juin D. 1	13 Tridi	Juillet m. 1	13 Tridi	j. 31	13 Tridi	s. 30	13 Tridi
j. 3	14 Quartidi	s. 3	14 Quartidi	l. 2	14 Quartidi	m. 2	14 Quartidi	Août v. 1	14 Quartidi	D. 31	14 Quartidi
v. 4	15 Quintidi	D. 4	15 Quintidi	m. 3	15 Quintidi	j. 3	15 Quintidi	s. 2	15 Quintidi	Septem. l. 1	15 Quintidi
Pr.Q. s. 5	16 Sextidi	Pr.Q. l. 5	16 Sextidi	Pr.Q. m. 4	16 Sextidi	P.Q. v. 4	16 Sextidi	P.Q. D. 3	16 Sextidi	P.Q. m. 2	16 Sextidi
le 6 D. 6	17 Septidi	le 6 m. 6	17 Septidi	le 5 j. 5	17 Septidi	le 4 s. 5	17 Septidi	le 3 l. 4	17 Septidi	le 2 m. 3	17 Septidi
l. 7	18 Octidi	m. 7	18 Octidi	v. 6	18 Octidi	D. 6	18 Octidi	m. 5	18 Octidi	j. 4	18 Octidi
m. 8	19 Nonodi	j. 8	19 Nonodi	s. 7	19 Nonodi	l. 7	19 Nonodi	m. 6	19 Nonodi	v. 5	19 Nonodi
m. 9	20 Décadi	v. 9	20 Décadi	D. 8	20 Décadi	m. 8	20 Décadi	j. 7	20 Décadi	s. 6	20 Décadi
j. 10	21 Primidi	s. 10	21 Primidi	l. 9	21 Primidi	m. 9	21 Primidi	v. 8	21 Primidi	D. 7	21 Primidi
v. 11	22 Duodi	D. 11	22 Duodi	m. 10	22 Duodi	j. 10	22 Duodi	P.L. s. 9	22 Duodi	P.L. l. 8	22 Duodi
P.L. s. 12	23 Tridi	l. 12	23 Tridi	m. 11	23 Tridi	v. 11	23 Tridi	le 10 D. 10	23 Tridi	le 9 m. 9	23 Tridi
le 13 D. 13	24 Quartidi	P.L. m. 13	24 Quartidi	P.L. j. 12	24 Quartidi	P.L. s. 12	24 Quartidi	l. 11	24 Quartidi	m. 10	24 Quartidi
l. 14	25 Quintidi	le 14 m. 14	25 Quintidi	le 13 v. 13	25 Quintidi	le 12 D. 13	25 Quintidi	m. 12	25 Quintidi	j. 11	25 Quintidi
m. 15	26 Sextidi	j. 15	26 Sextidi	s. 14	26 Sextidi	l. 14	26 Sextidi	m. 13	26 Sextidi	v. 12	26 Sextidi
m. 16	27 Septidi	v. 16	27 Septidi	D. 15	27 Septidi	m. 15	27 Septidi	j. 14	27 Septidi	s. 13	27 Septidi
j. 17	28 Octidi	s. 17	28 Octidi	l. 16	28 Octidi	m. 16	28 Octidi	v. 15	28 Octidi	D.Q. D. 14	28 Octidi
v. 18	29 Nonodi	D. 18	29 Nonodi	m. 17	29 Nonodi	j. 17	29 Nonodi	D.Q. s. 16	29 Nonodi	le 15 l. 15	29 Nonodi
s. 19	30 Décadi	l. 19	30 Décadi	m. 18	30 Décadi	v. 18	30 Décadi	le 17 D. 17	30 Décadi	m. 16	30 Décadi

Fêtes sans-Culotides.

m. 17	1 de la Vertu
j. 18	2 du Génie
v. 19	3 du Travail
s. 20	4 de l'Opinion
D. 21	5 des Récompenses

est appelée la Franciade.

Le jour intercalaire qui doit terminer cette période est appelé le jour de la Révolution. Ce jour est placé après les cinq complémentaires.

XI. *Le jour, de minuit à minuit, est divisé en dix parties; chaque partie en dix autres, ainsi de suite jusqu'à la plus petite portion commensurable de la durée. Cet article ne sera de rigueur pour les actes publics qu'à compter du premier du premier mois de la troisième année de la république.*

XII. *Le comité d'instruction publique est chargé de faire imprimer en différens formats le nouveau calendrier, avec une instruction simple pour en expliquer les principes et les usages les plus familiers.*

XIII. *Le nouveau calendrier, ainsi que l'instruction seront envoyés aux corps administratifs, aux municipalités, aux tribunaux, aux juges de paix et à tous les officiers publics, aux armées, aux sociétés populaires et à tous les collèges et écoles. Le conseil exécutif provisoire le fera passer aux ministres, consuls et autres agens de France dans les pays étrangers.*

XIV. *Tous les actes publics seront datés suivant la nouvelle organisation de l'année.*

XV. *Les professeurs, instituteurs et institutrices, les pères et mères de famille, et tous ceux qui dirigent l'éducation des enfans de la république, s'empresseront de leur expliquer le nouveau calendrier, conformément à l'instruction qui y est annexée.*

XVI. *Tous les quatre ans, ou toutes les Franciades, au jour de la Révolution, il sera célébré des jeux républicains, en mémoire de la révolution Françoise.*

Air de la Carmagnole.

Sans-Culottes, rassurez-vous, bis
La victoire est pour nous; bis
On [illegible]
Toujours l'on chantera
Chantons la Carmagnole, &c.

A Paris chez [illegible] rue St. Séverin N.° [illegible] et chez Fillion et Valen[illegible] rue St. Jacques N.° 8 près St. Séverin

La calendrier Républicain

Converti au format calendrier grégorien.

```
const AN_REGEX = /An (\d+)/i;

export function normalizeYear(year) {
  const m = year.match(AN_REGEX);

  if (!m)
    return +year;

  const nb = m[1];

  if (nb < 2 || nb > 14)
    throw Error(
      `toflit18.republican_calendar.normalizeYear: invalid year ${year}.`
    );

  return 1792 + (+nb);
}
```

Contrôle de version des données

- git - contrôle de version pour les codes sources
- Appliqué à la gestion de corpus de données
- Implique des fichiers texte brut

fichier texte brut

- Les fichiers texte brut facilitent les traitements informatiques.
- CSV, JSON, XML sont des formats de fichier texte brut.
- Non, XLS, XLSX et ODT n'en sont pas.

No. 19—*continued.*

TRADE WITH PRINCIPAL COUNTRIES.

PORTUGAL—IMPORTS.

TOTAL VALUE of IMPORTS for HOME CONSUMPTION (MERCHANDISE only) into PORTUGAL, distinguishing PRINCIPAL COUNTRIES.

COUNTRIES.	1886.	1887.	1888.	1889.	1890.
	Milreis.	Milreis.	Milreis.	Milreis.	Milreis.
Russia	471,000	500,000	887,000	1,085,000	448,000
Norway and Sweden	741,000	876,000	983,000	1,351,000	1,236,000
Germany	4,681,000	4,508,000	4,706,000	5,356,000	6,303,000
Holland	423,000	330,000	369,000	348,000	411,000
Belgium	1,537,000	1,522,000	1,445,000	1,526,000	2,246,000
United Kingdom	12,125,000	12,250,000	12,309,000	13,914,000	13,263,000
France	5,124,000	4,955,000	4,980,000	6,010,000	6,862,000
Switzerland	—	133,000	195,000	267,000	275,000
Spain	2,593,000	2,265,000	2,551,000	3,240,000	2,931,000
Italy	799,000	437,000	535,000	675,000	796,000
Austria	8,000	140,000	371,000	567,000	488,000
Egypt	—	325,000	308,000	216,000	89,000
Morocco	424,000	316,000	239,000	251,000	444,000
China	—	109,000	243,000	263,000	266,000
United States	4,978,000	5,307,000	4,484,000	3,665,000	5,148,000
Brazil	2,008,000	1,874,000	2,148,000	1,803,000	1,946,000

1891.	1892.	1893.	1894.	COUNTRIES.
Milreis.	Milreis.	Milreis.	Milreis.	
757,000	267,000	408,000	229,000	Russia.
1,103,000	817,000	1,465,000	1,375,000	Norway and Sweden.
5,163,000	2,785,000	4,440,000	4,258,000	Germany.
409,000	328,000	362,000	267,000	Holland.
1,292,000	888,000	1,031,000	1,183,000	Belgium.
11,780,000	9,303,000	10,872,000	9,809,000	United Kingdom.
5,321,000	3,434,000	3,826,000	3,756,000	France.
241,000	221,000	178,000	220,000	Switzerland.
2,492,000	1,701,000	2,857,000	2,960,000	Spain.
754,000	788,000	365,000	551,000	Italy.
469,000	266,000	233,000	226,000	Austria.
311,000	439,000	458,000	430,000	Egypt.
447,000	106,000	213,000	181,000	Morocco.
229,000	228,000	288,000	260,000	China.
5,252,000	6,038,000	7,291,000	5,761,000	United States.
2,054,000	1,911,000	2,428,000	2,484,000	Brazil.

PORTUGAL—EXPORTS.

TOTAL VALUE of EXPORTS of DOMESTIC PRODUCE (MERCHANDISE only) from PORTUGAL, distinguishing PRINCIPAL COUNTRIES.

COUNTRIES.	1886.	1887.	1888.	1889.	1890.
	Milreis.	Milreis.	Milreis.	Milreis.	Milreis.
Russia	296,000	318,000	306,000	336,000	323,000
Norway and Sweden	213,000	223,000	299,000	287,000	187,000
Denmark	206,000	160,000	134,000	208,000	231,000
Germany	1,518,000	1,620,000	1,903,000	1,987,000	2,066,000
Holland	206,000	231,000	257,000	278,000	282,000
Belgium	349,000	495,000	376,000	469,000	564,000
United Kingdom	6,707,000	6,764,000	7,828,000	8,528,000	7,993,000
France	9,491,000	4,818,000	5,207,000	3,768,000	1,522,000
Spain	1,156,000	1,210,000	939,000	1,105,000	879,000
Italy	164,000	188,000	168,000	197,000	221,000
United States	647,000	646,000	554,000	509,000	730,000
Brazil	4,575,000	3,686,000	4,195,000	4,260,000	5,181,000
Portuguese Possessions:					
In Africa	540,000	601,000	879,000	1,007,000	1,111,000
In Asia	17,000	19,000	28,000	22,000	27,000
Other Countries	223,000	260,000	370,000	382,000	281,000
TOTAL Milreis	26,108,000	21,239,000	23,443,000	23,343,000	21,538,000
TOTAL £	5,874,000	4,779,000	5,275,000	5,252,000	4,846,000

1891.	1892.	1893.	1894.	COUNTRIES.
Milreis.	Milreis.	Milreis.	Milreis.	
284,000	406,000	621,000	735,000	Russia.
362,000	173,000	219,000	303,000	Norway and Sweden.
302,000	280,000	269,000	304,000	Denmark.
2,308,000	2,221,000	1,961,000	2,053,000	Germany.
301,000	330,000	396,000	390,000	Holland.
442,000	440,000	632,000	1,048,000	Belgium.
7,483,000	8,719,000	6,553,000	6,692,000	United Kingdom.
1,289,000	1,015,000	906,000	757,000	France.
888,000	1,409,000	1,371,000	2,513,000	Spain.
179,000	105,000	187,000	221,000	Italy.
806,000	900,000	841,000	539,000	United States.
5,274,000	6,782,000	7,155,000	5,994,000	Brazil.
				Portuguese Possessions:
1,280,000	1,624,000	2,061,000	2,187,000	In Africa.
7,000	26,000	34,000	24,000	In Asia.
174,000	201,000	202,000	164,000	Other Countries.
21,379,000	24,631,000	23,408,000	23,924,000	Milreis TOTAL.
4,810,000	5,542,000	5,267,000	5,383,000	£ TOTAL.

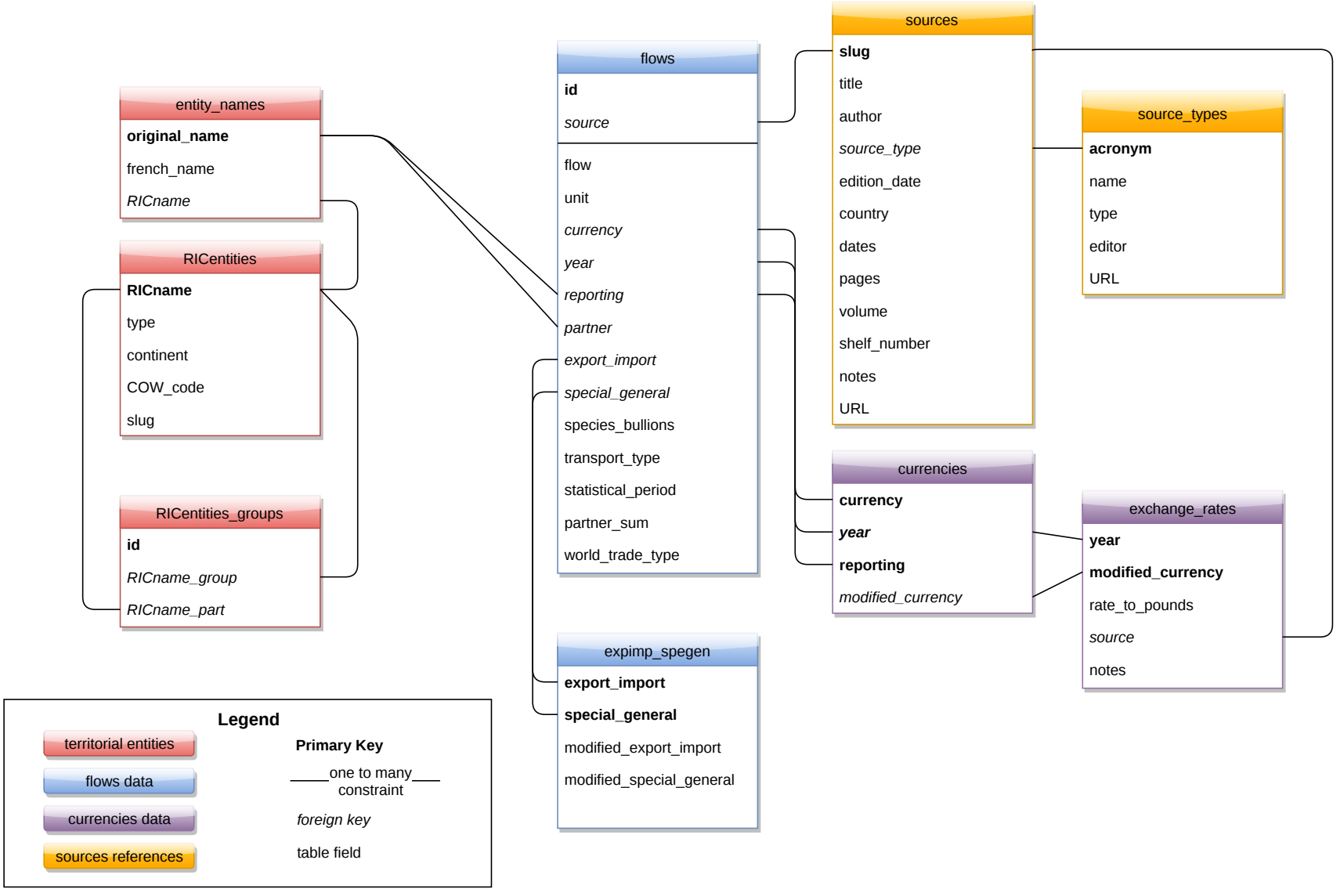
Prendre soin des données

- Contrôles qualité des données
- Hybridant des approches qualitatives et quantitatives

Bases de données

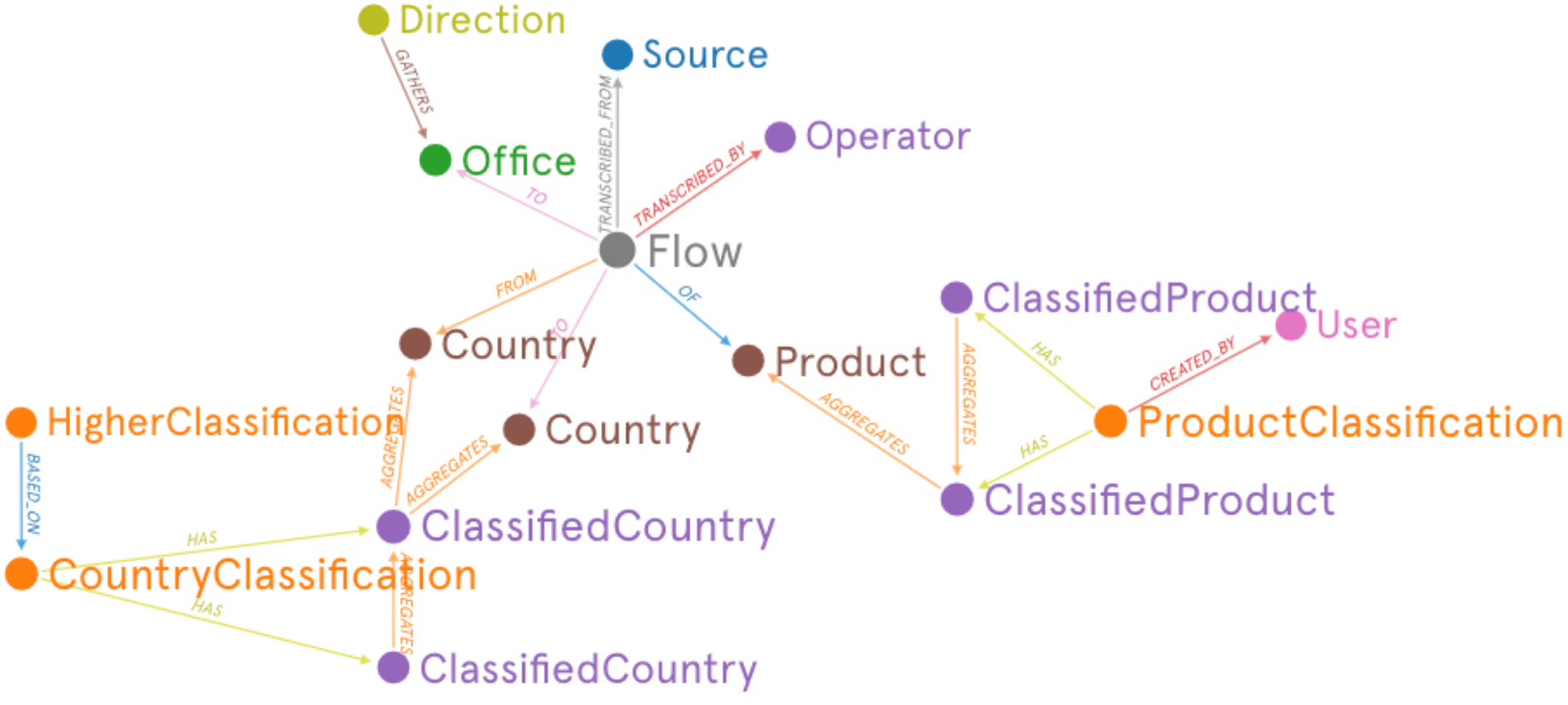
- Les base de données facilitent accès et manipulation
- Le choix de la technologie dépends des usages

RICardo relational database schema



collaboration avec Karine ONFROY G.R.E.T.h.A (UMR CNRS 5113) Bordeaux

Toflit18 NEO4J schema



Classifications

Un des objectifs principaux de TOFLIT18 :
créer des classifications dédiées aux questions de recherche.

base de données en graphe

Le modèle a été conçu pour proposer des **classifications** :

- **hiérarchiques** : aggrégation progressive
- **concurrentes** : aggrégation dédiée à une question
- **dynamiques** : toute analyse commence par un choix de classification

la base de données TOFLIT18

- 419729 flux
- 47732 produits
- 843 pays
- 51 bureaux de commerce français
- 120 années
- 807 volumes d'archive

la base de données RICardo

- 294138 flux
- 1492 RICentities
- 152 années
- 120 monnaies
- 7206 taux de change vers le £
- 73 types de sources (919 volumes)

L'exploration visuelle au service des données

Datascape:

- Des visualisations de données intéractives
- Proposant divers points de vue sur les données
- Aide à s'approprier la complexité par les dynamiques d'exploration

Leclercq, C. and Girard, P. (2013). *The Experiments in Art and Technology Datascape*. Collections Électroniques de l'INHA. Actes de Colloques et Livres En Ligne de l'Institut National D'histoire de L'art. INHA http://inha.revues.org/4926 (accessed 27 October 2015).

Concevoir un datascape

Ateliers appelés *«data sprints»* avec:

- historiens
- économistes
- développeurs
- designers

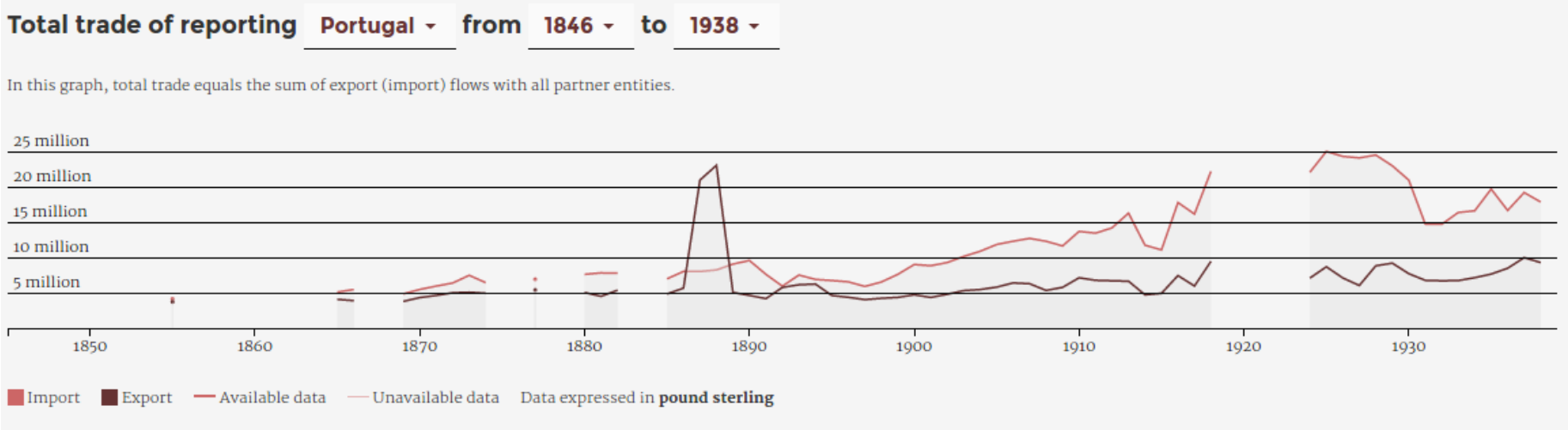
Traitant les enjeux de **contenu**, **implémentation** et de **design**
en même temps et au même endroit.

Analyse Exploratoire de Données

« The greatest value of a picture is when it forces us to notice what we never expected to see. »
Tukey, J. W.

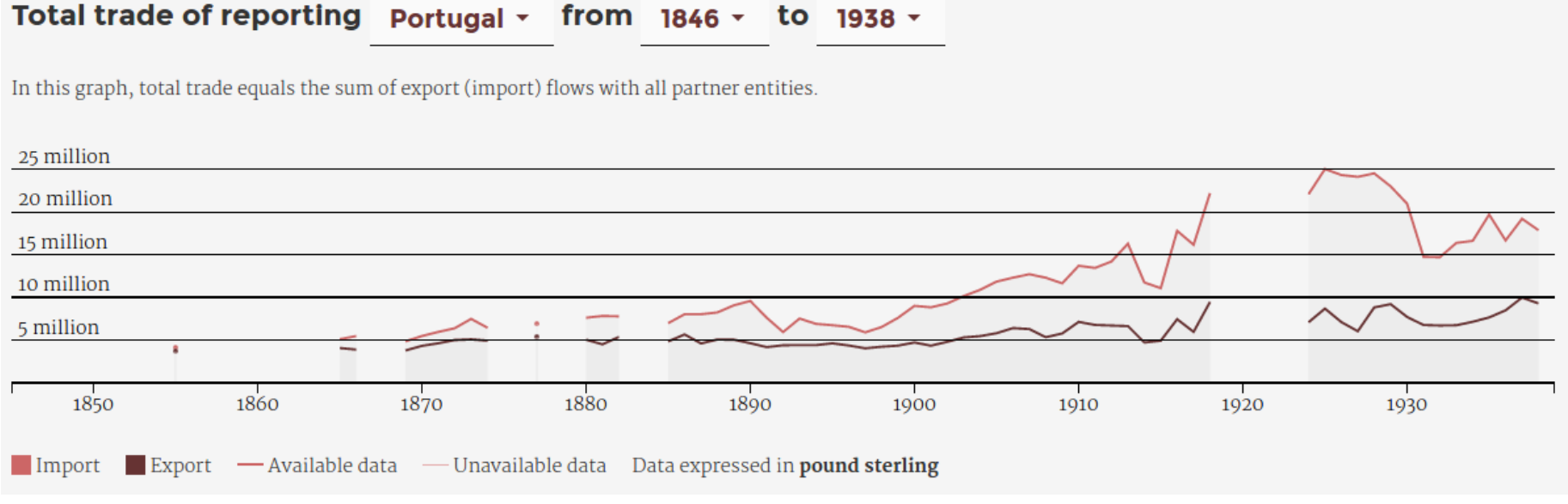
Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Company.

Visualisation de données



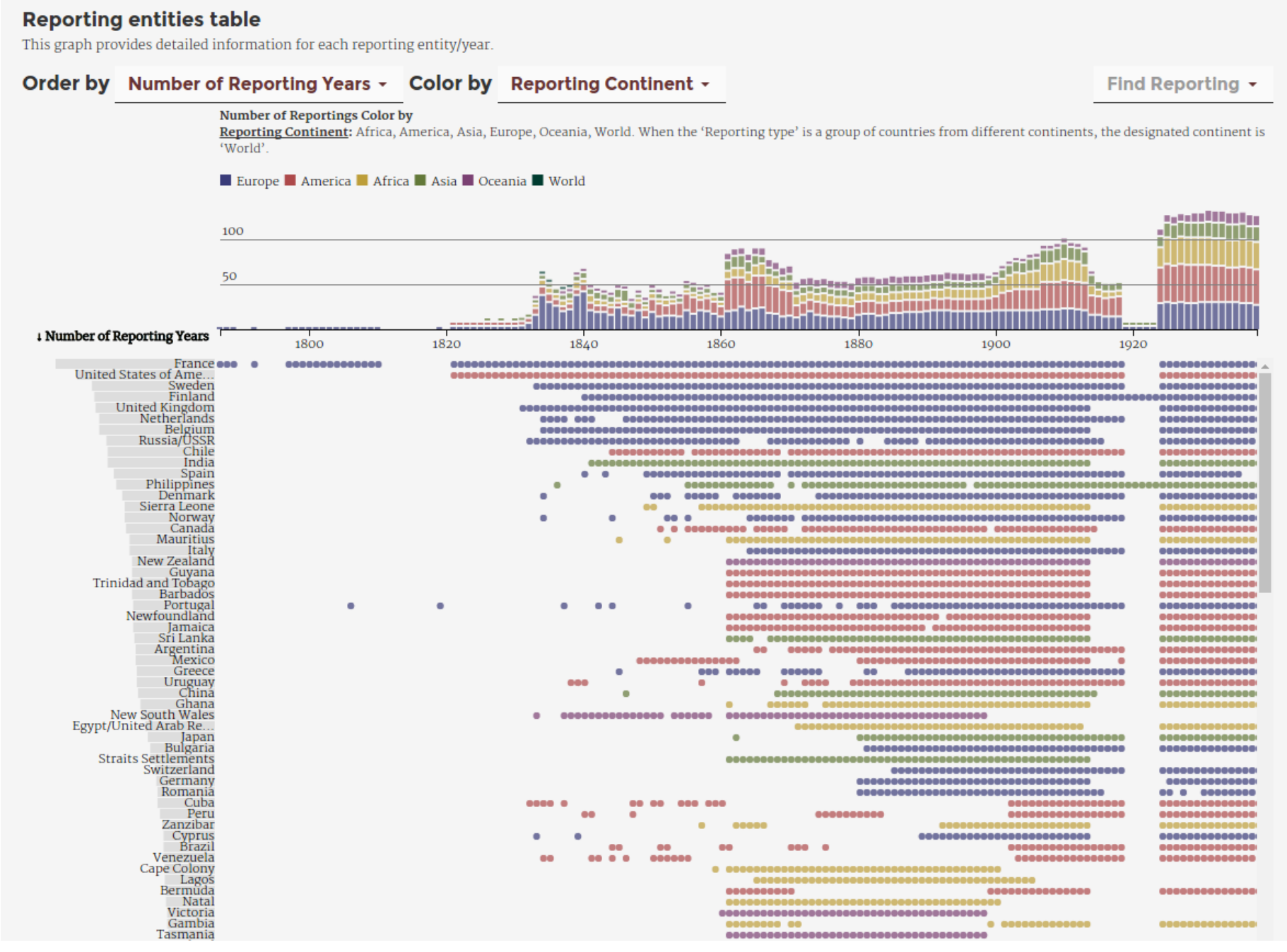
⬆ oups !

Vérifier et corriger les données

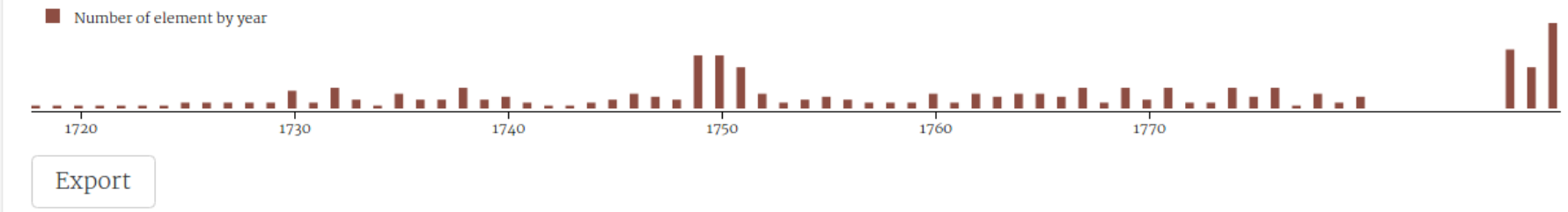


- corrigé comme indiqué slide 20 -

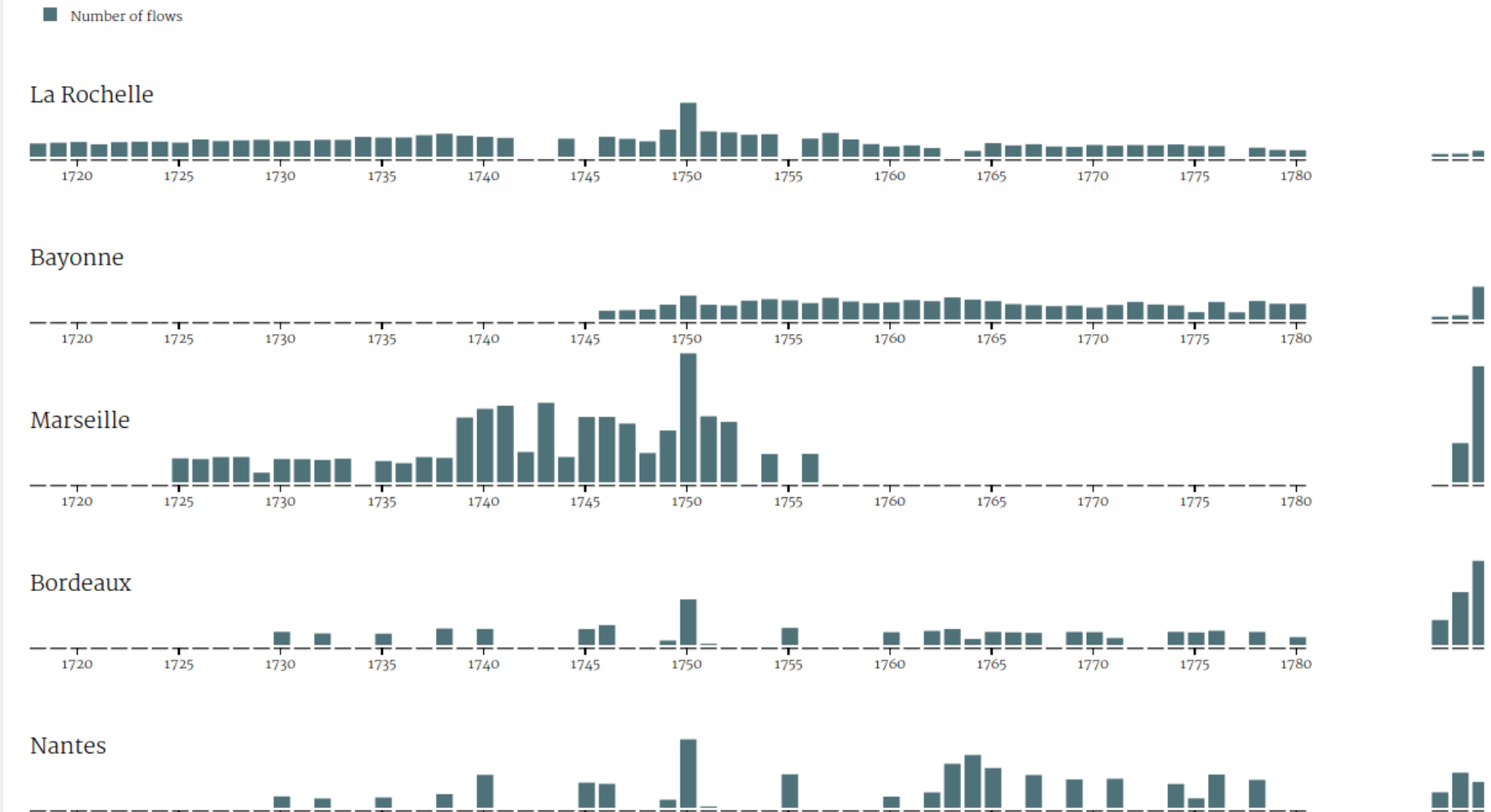
La vue métadonnée de Ricardo



La vue métadonnée de TOFLIT18



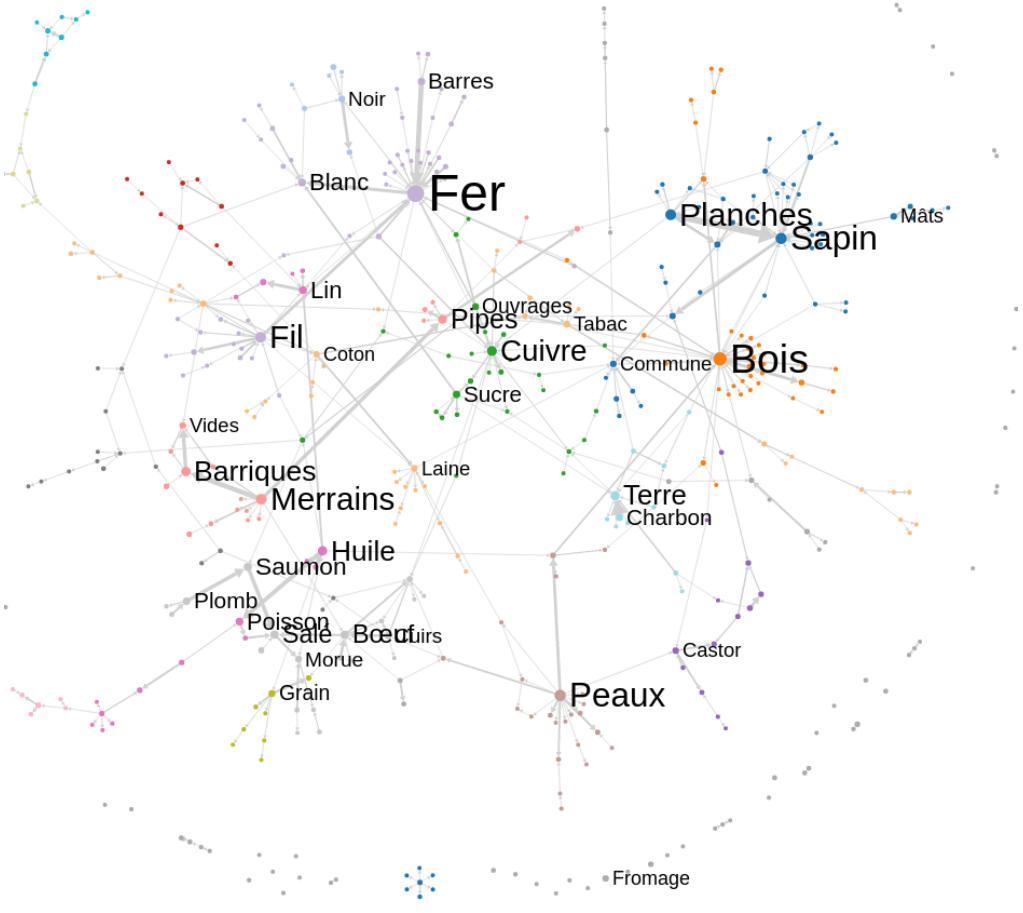
Export



La chaîne de transformations

volumes d'archive > images > excel > git(csv) > base de données > visualisation de données **> csv**

Un avant goût de l'atelier TOFLIT18/Ricardo



Co-occurrences de termes dans les noms de produits dans les exports de "La Rochelle" entre 1720 et 1729

Rendez vous ce soir à 18h30.

Les enjeux liés à la visualisation de données historiques

medialab.github.io/ricardo

Girard, P., Dedinger, B., Ricci, D., Ooghe-Tabanou, B., Jacomy, M., Plique, G. and Tible, G. (2016). RICardo Project : Exploring XIX Century International Trade. Kraków, Poland http://ricardo.medialab.sciences-po.fr/Girardetal_RICardo_dh2016_en.pdf.

Science ouverte

- télécharger les données des visualizations en csv
- corpus de données seront ouverts en 2017
  CC BY-SA
- RICardo : http://ricardo.medialab.sciences-po.fr
- sources RICardo : github.com/medialab/ricardo
- sources TOFLIT18 : github.com/medialab/toflit18
- cette présentation:
  http://medialab.github.io/toflit18/ANFmateSHS

Bibliographie & Liens (1/2)

Dedinger, Béatrice, and Paul Girard. 2016. *'Exploring Trade Globalization in the Long Run : The RICardo Project'*. Historical Methods.
http://ricardo.medialab.sciences-po.fr.

Girard, Paul, Béatrice Dedinger, Donato Ricci, Benjamin Ooghe-Tabanou, Mathieu Jacomy, Guillaume Plique, and Grégory Tible. 2016. *'RICardo Project : Exploring XIX Century International Trade'. In . Kraków, Poland.*
http://ricardo.medialab.sciences-po.fr/Girardetal_RICardo_dh2016_en.pdf.

Bibliographie & Liens (2/2)

Latour, Bruno. 1993. *‘Le Topofil de Boa-Vista. La Référence Scientifique: Montage Photophilosophique’*. Raisons Pratiques 4: 187–216.

Latour, Bruno, Pablo Jensen, Tommaso Venturini, Sébastian Grauwin, and Dominique Boullier. 2012. *‘“The Whole Is Always Smaller than Its Parts” - a Digital Test of Gabriel Tardes’ Monads’*. The British Journal of Sociology 63 (4): 590–615. doi:10.1111/j.1468-4446.2012.01428.x.

Tukey, John Wilder. 1977. *Exploratory Data Analysis*. Addison-Wesley Publishing Company.