

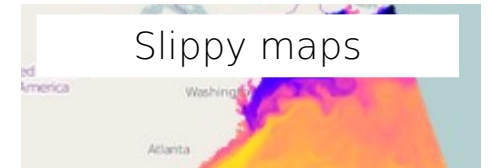
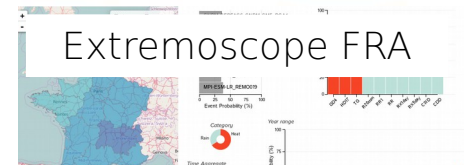
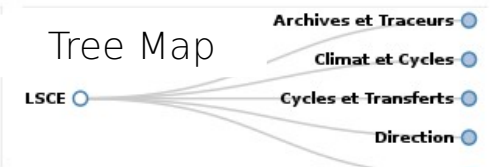
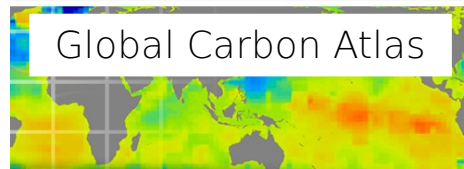
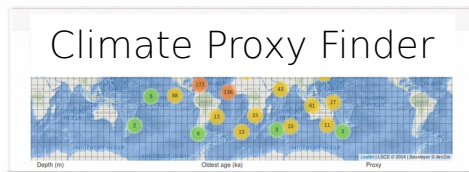
Exploration interactive de données multidimensionnelles climatiques

Cathy Nangini, PhD

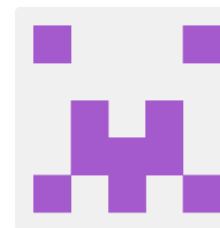
LSCE

Laboratoire des Sciences du Climat et de l'Environnement

LSCE Data Viz Group



Patrick Brockmann
PBrockmann



KatiRG

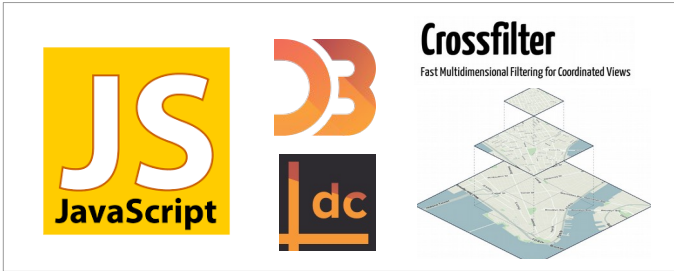
Paris
cnangini@gmail.com
Joined on 11 Dec 2011



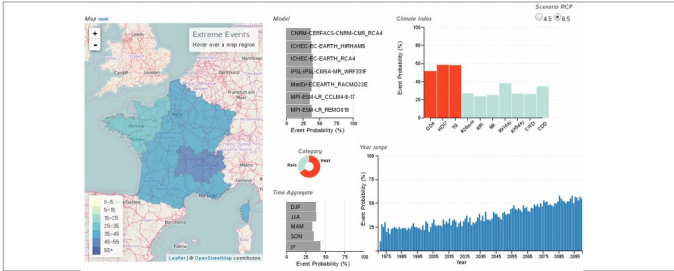
Cette présentation va se focaliser sur les visualisations interactives que le navigateur rend possible aujourd'hui



La montée en puissance du navigateur
14500 av. J.-C – aujourd'hui



Outils disponibles pour accéder au navigateur
développeurs - - - > scientifiques

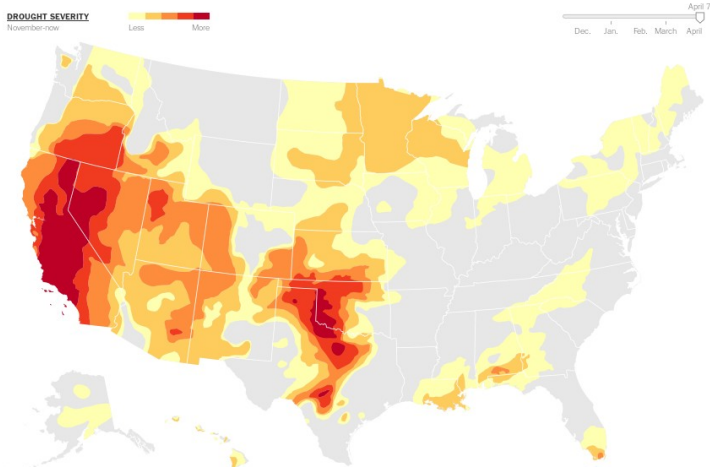


Extremoscope
exemple climatique multi-dimensionnel

La data viz est pour mettre en forme des résultats et les rendre interactifs

parallel coordinates

*interactive
timeline*

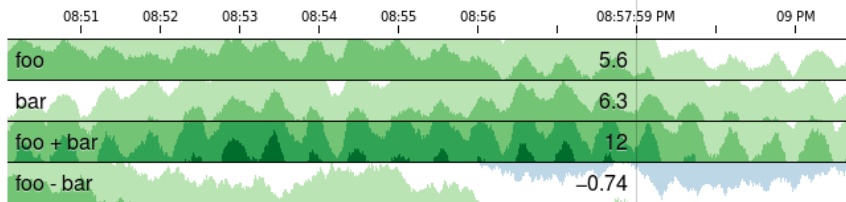


cartes

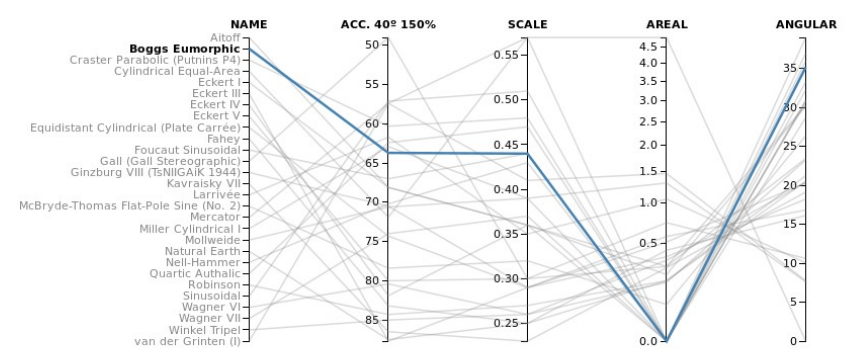
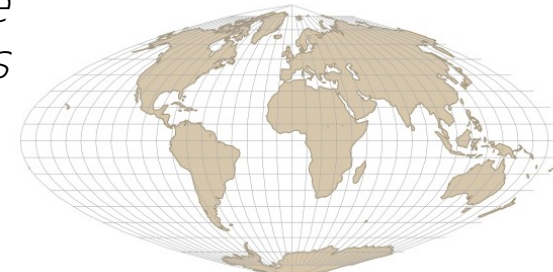
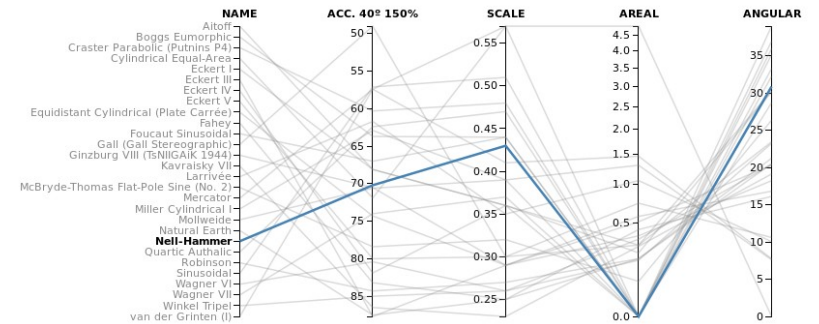
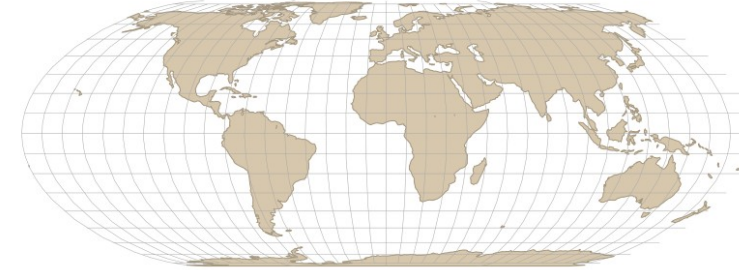
*interactive
cursor*

*interactive
lines*

Time Series Visualization

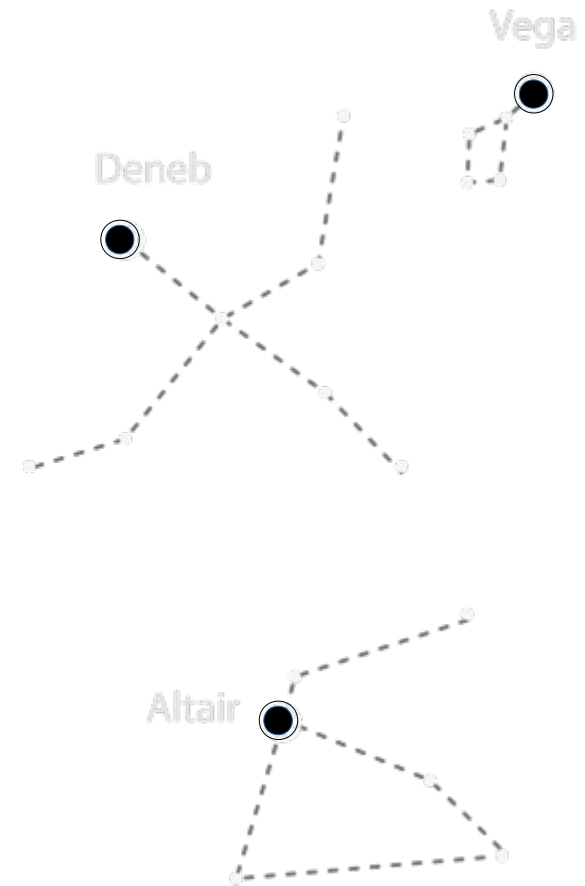


séries temporelles



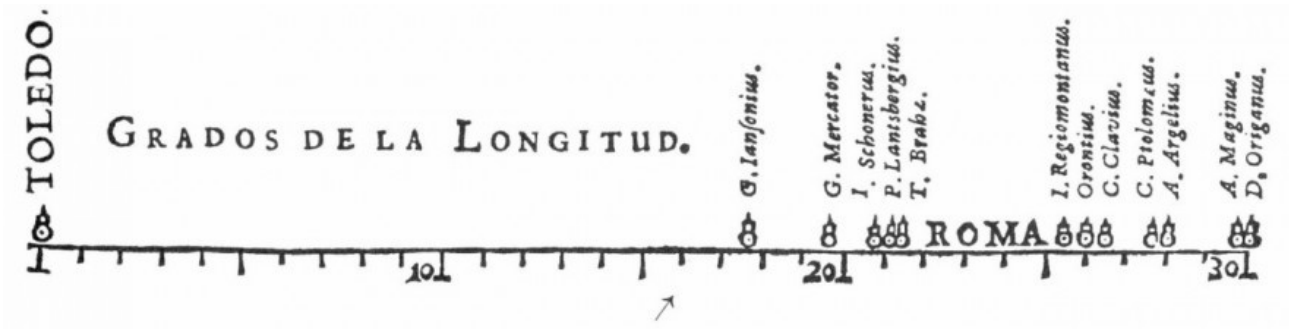
Le besoin de cartographier, de tracer, et finalement de comprendre l'univers est aussi vieux que la société

14500 BC : Grotte de Lascaux
- montre les étoiles dans Le Triangle d'été



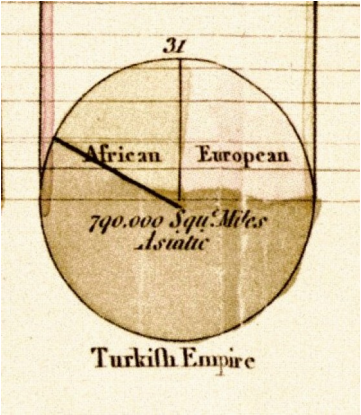
Le graphe statistique, fondation de la data viz, a été inventé pendant La Renaissance.

1644 : astronome Michael Florent Van Langren
- première représentation des données statistiques



mesures longitudinales entre Rome et Toledo

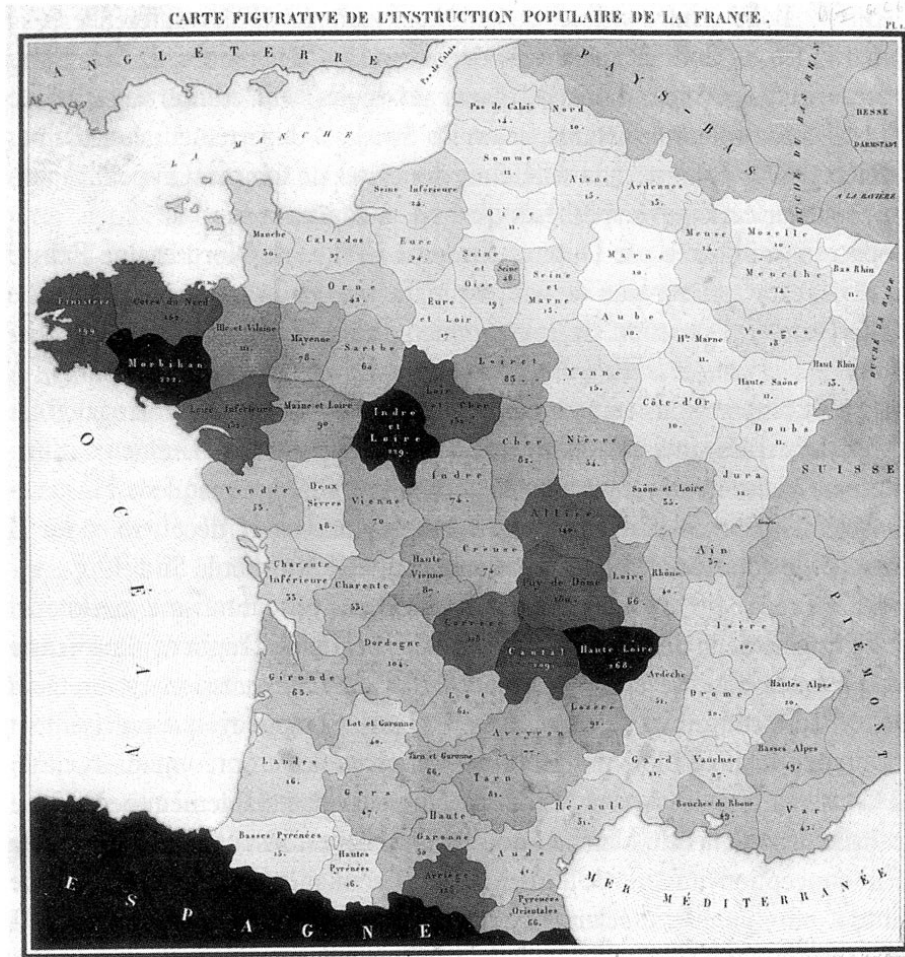
c. 1780 : William Playfair
- l'invention du bar chart, line graph et pie chart (camembert) (assiégé)



proportion de terres détenues par l'Empire turc

c. 1850-1900 : "Age d'or" des graphiques statistiques

1826, Pierre Charles François Dupin
- première carte "choroplèthe"



du grec :

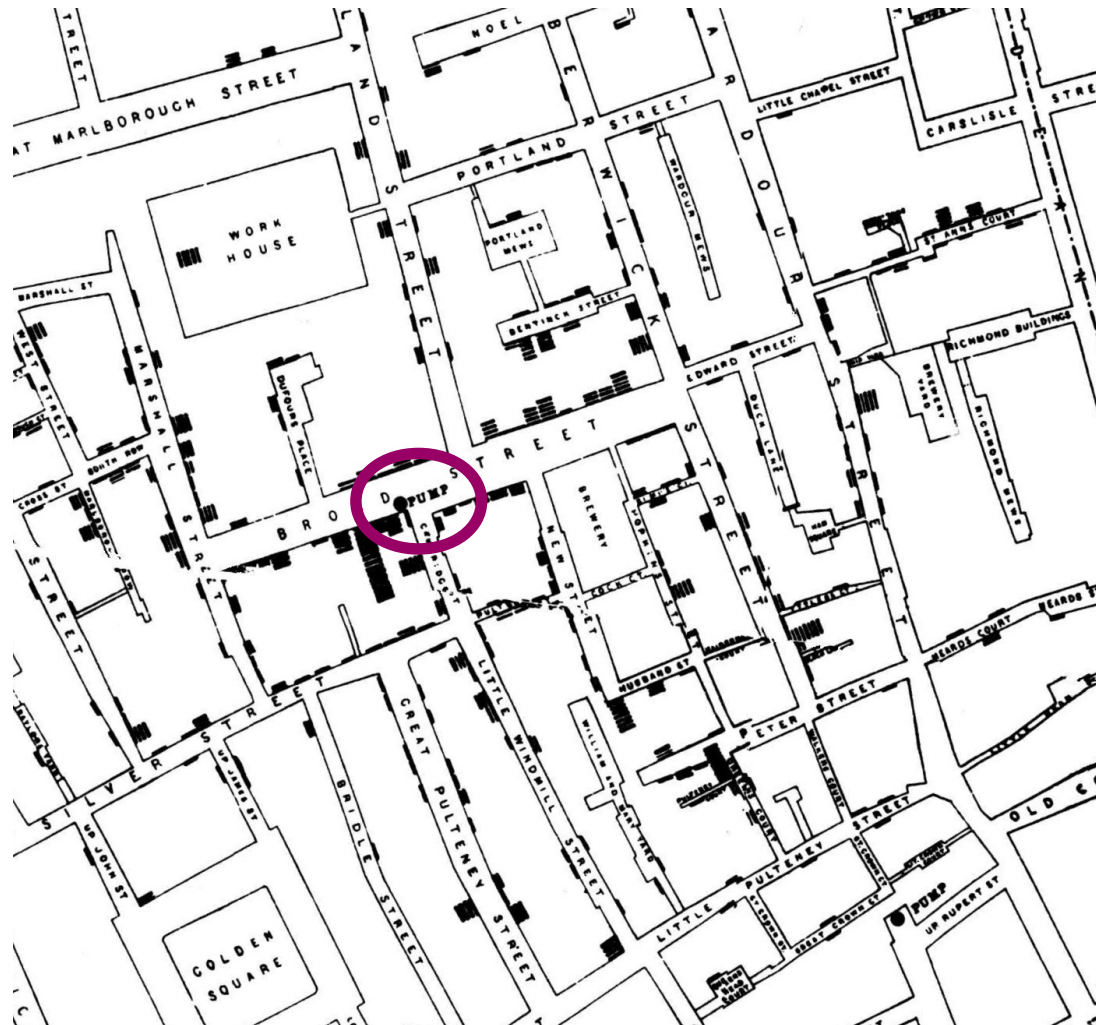
χώρος : « zone/région »
πληθαίν : « multiple »

La distribution de l'analphabétisme □ haut ■ bas

c. 1850-1900 : "Age d'or" des graphiques statistiques

1849, John Snow

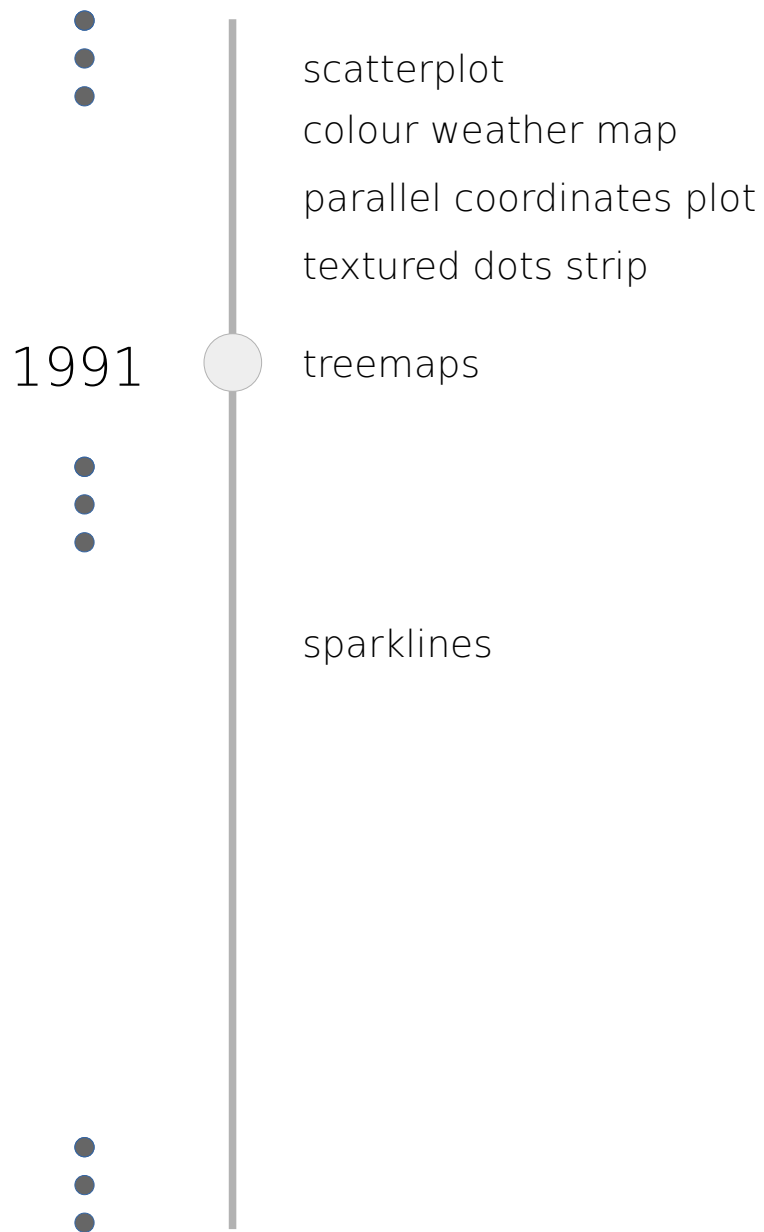
- cartographie les morts du cholera avec les pompes à eau du quartier
- discrédite l'idée fausse que le choléra est causé par des `miasmes`



Une victoire pour la data viz :
la carte a révélé que les morts se trouvaient au tour d'une pompe à eau qui dessert la zone

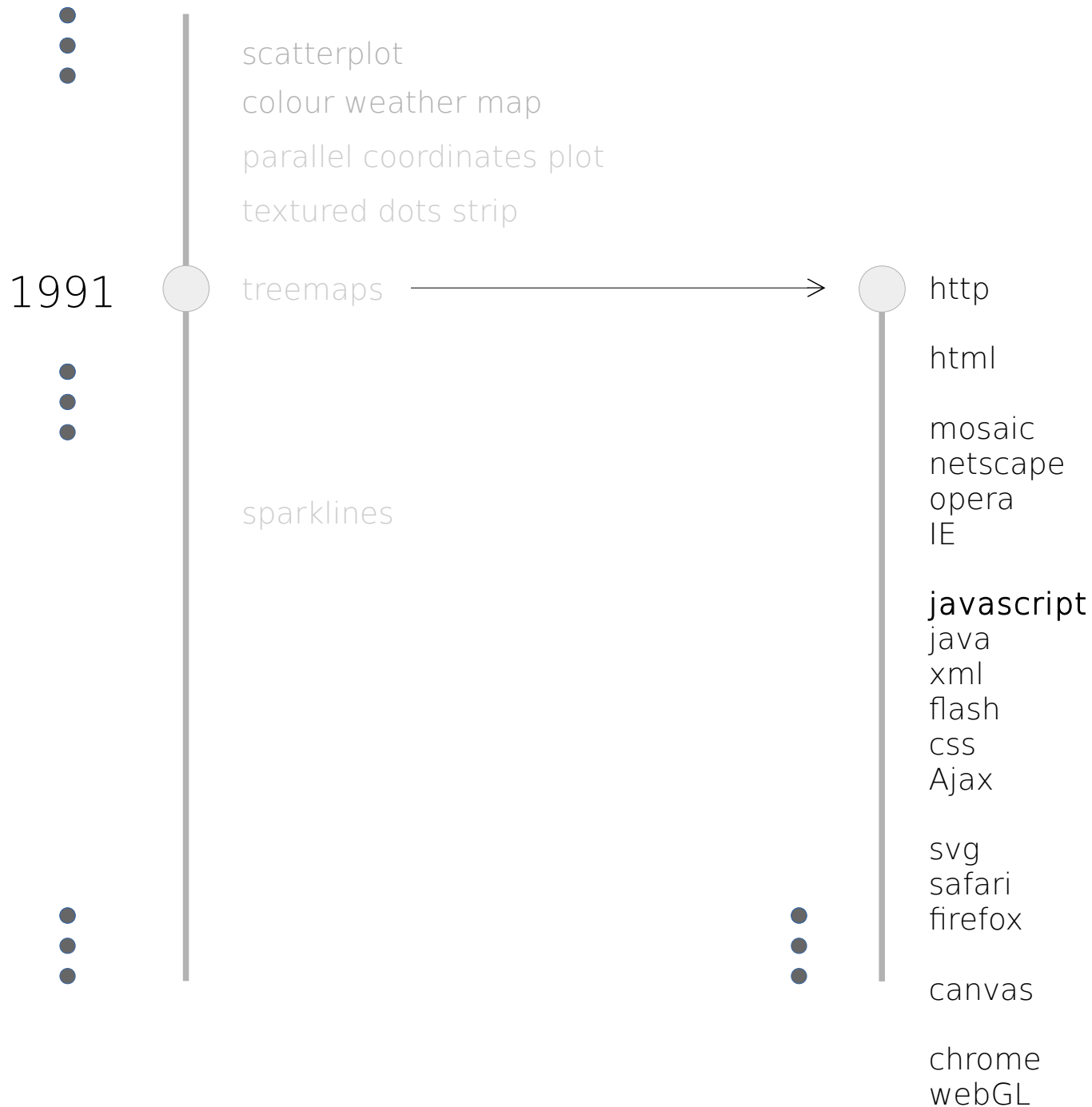
■ maisons avec des morts

○ la pompe qui était la source du problème



Le développement des graphiques statistiques se poursuit...

...mais en 1991 un univers parallèle émerge...



Technologies qui apporte la visualisation dans le navigateur



Permet l'interaction avec les éléments sur la page

B **R** **OW** **S** **ER** **POWER**



Le navigateur est devenu très puissant et ils comportent des moteurs javascript très efficace.

Un nombre croissant de visualisations de données interactive est aujourd'hui visible dans le navigateur.

Aujourd'hui il y a une grande variété d'outils



Data-Visualization Tools & Books

339 Visualization Tools

Highlights

★★★	9
★★	28
★	49

Features

Charting	154
Web-based	150
Design focused	55
Spreadsheet	37
Color	29
Infographics	24
Statistical	23
Scientific	16

Data Types

Multivariate ...	89
Time	78
Geographical	77
Network	27
Text	17

Data Processing

Programming	7
Formula	6
Convert	6
Extract	5
Aggregate	5
Advanced	4
Yes	3
Merge	3
Transform	2
Search	2
Scrape	2
Query	2
Indexing	2
Graph processing	2
Filter	2
Clean	2
Text processing	1
Survey	1
Stream	1

Cost

Free	167
Paid	55
Free & Paid	12

Search: Title

alteryx
Charting + Statistical + Graphical Interface + R

RStudio
Charting + Statistical + Web-based + R

Leaflet.js
Charting + Web-based + JavaScript

NodeXL
Charting + Graphical Interface

Google Charts
Charting + Web-based + JavaScript

ColorBrewer 2.0
Color + Web-based + Graphical Interface

Tableau Public
Charting + Statistical + Infographics + Web-based + Scripting + Graphical Interface

D3.JS
Charting + Web-based + JavaScript

Keshif
Charting + Web-based + JavaScript + Graphical Interface

Scrapy
Web-based + Python

Sketch
Infographics + Graphical Interface

Gimp
Color + Infographics + Graphical Interface

EventFlow
Charting + Graphical Interface

Shiny by RStudio
Charting + Statistical + Web-based + R + Graphical Interface

Raw
Charting + Web-based + JavaScript + Graphical Interface

Cubism
Charting + Web-based + JavaScript

Main

Categorized	236
Active	197
Open-Source	119

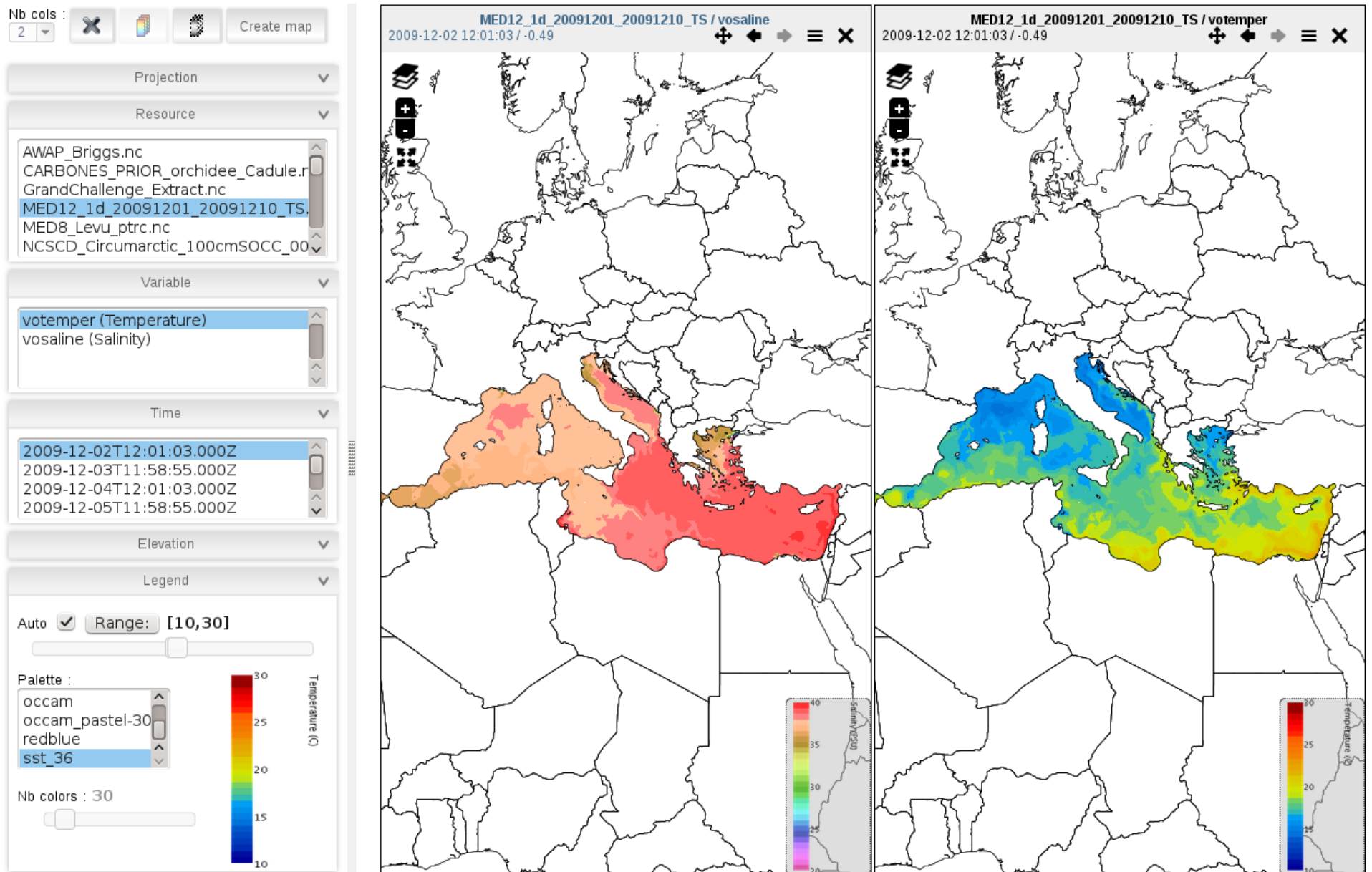
Tags

Platform	11
d3	11
Excel	7
Storytelling	5
PDF	5
VIS2015	4
Technology	4
jQuery	4
Mockups	3
Faceted Browsing	3
3D	3
Treemap	2
SVG	2
Share	2
Server	2
Journalism	2
HTML Canvas	2
Framework	2

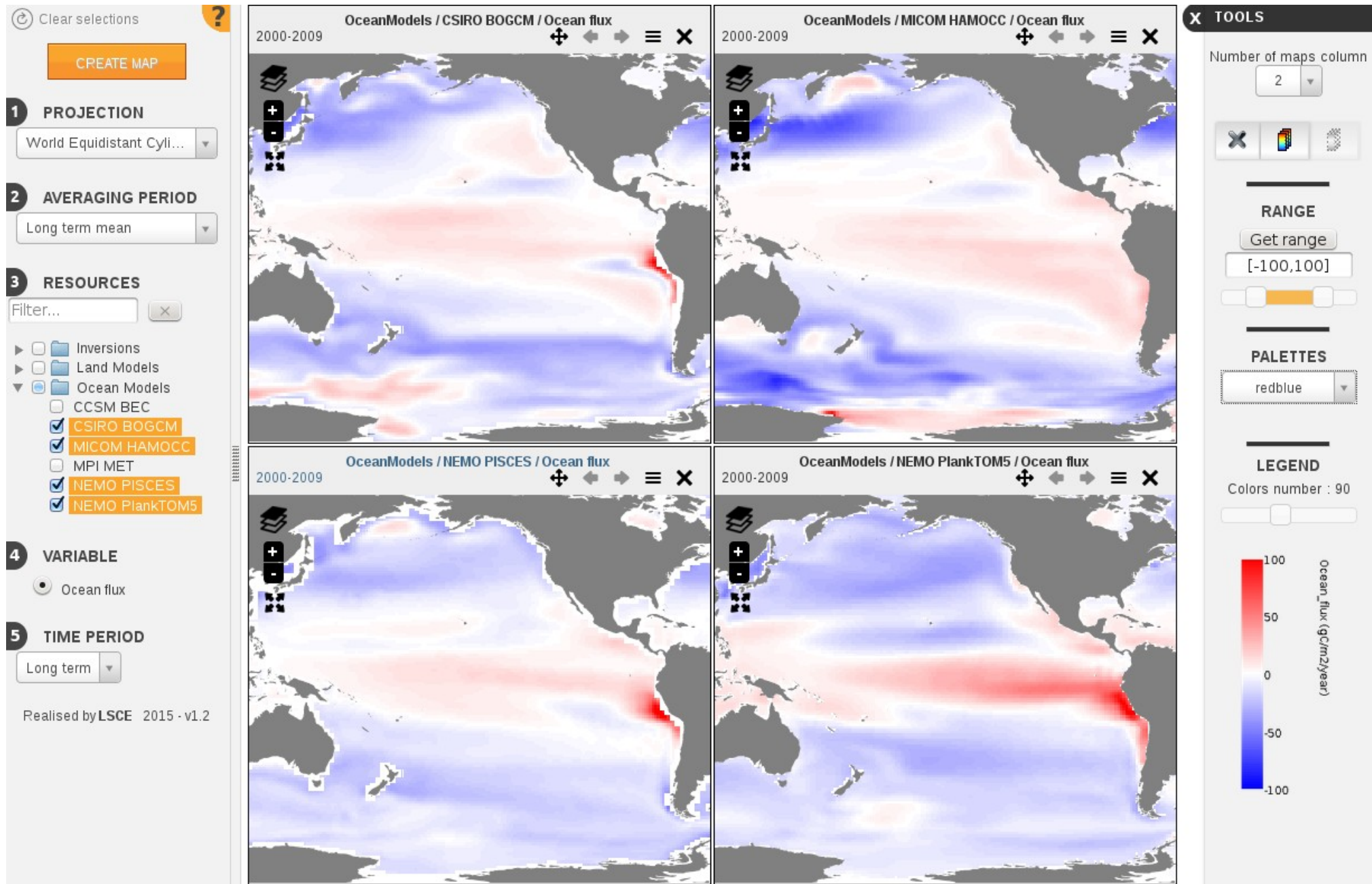
Programming

JavaScript	63
Graphical Interface	56
R	13
Scripting	12
Python	11
HTML	9
Java	7
RESTful API	4
JSON	4
Ruby	3
Processing	2
PHP	2
Other	2
Matlab	2
Command Line	2
API	2
ActionScript	2
XML	1

La recherche sur le climat peut aussi profiter de graphiques interactifs dans le navigateur



Global Carbon Atlas



Un grand joueur dans la domaine de data viz: d3.js



2011



The New York Times

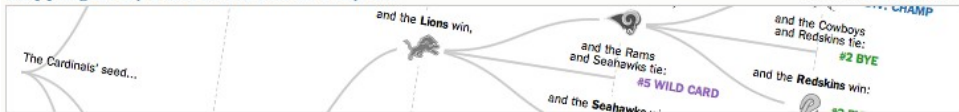
theguardian

Des centaines de réalisations publiées

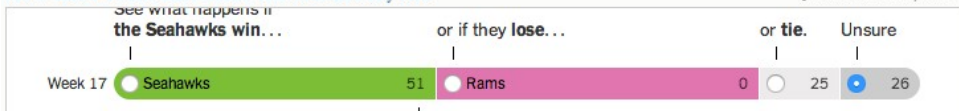
Wind Speeds Along Hurricane Sandy's Path October 30, 2012



Mapping Every Path to the N.F.L. Playoffs December 27, 2014



How Each Team Can Make the N.F.L. Playoffs December 20, 2014



The Most Detailed Maps You'll See From the Midterm Elections November 5, 2014



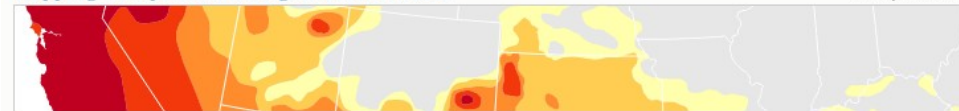
How To Scroll November 3, 2014

In typical usage, scrolling is a continuous gesture. In contrast, a reader can slide the viewport to any position between the top and bottom. In contrast, a reader can slide the viewport to any position between the top and bottom. In contrast, a reader can slide the viewport to any position between the top and bottom. In contrast, a reader can slide the viewport to any position between the top and bottom.

'Stop-and-Frisk' Is All but Gone From New York September 19, 2014



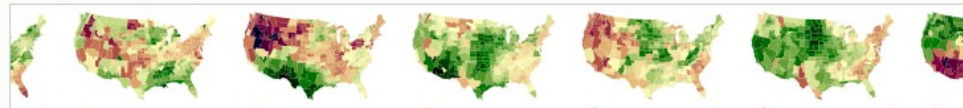
Mapping the Spread of Drought Across the U.S. July 26, 2014



Visualizing Algorithms June 26, 2014



Drought during Month August 24, 2012



State-by-State Premiums Under the Health Care Law September 23, 2013

State	premium for subsidies	and getting subsidies	\$10,000 a year	and getting subsidies
Average, 36 States	\$774	\$282	\$214	\$145
Tennessee	584	282	161	145
Arizona	600	282	166	145
Kansas	610	282	174	145

Constellations of Directors and Their Stars September 6, 2013



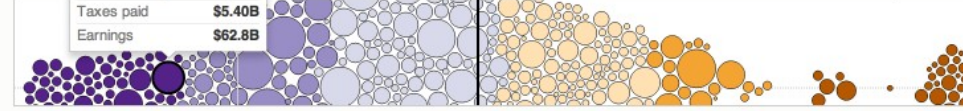
How To Infer Topology September 2, 2013



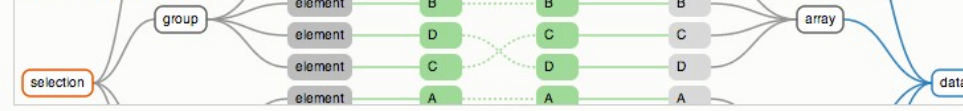
For Yankees, Little Bang for Their Buck August 2, 2013



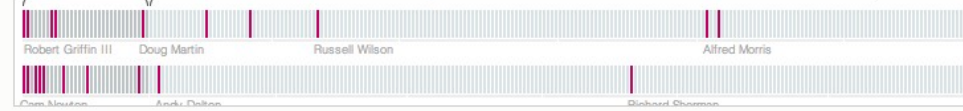
Across U.S. Companies, Tax Rates Vary Greatly May 25, 2013



How Selections Work April 26, 2013



N.F.L. Draft: How Good Are Teams at Picking the Best? April 25, 2013



Il y a beaucoup de bibliothèques qui sont basées sur d3.js



DC pour Dimensional Charting qui permet de réaliser des filtrages sur des données multi-dimensionnelles.

Crossfilter

Fast Multidimensional Filtering for Coordinated Views

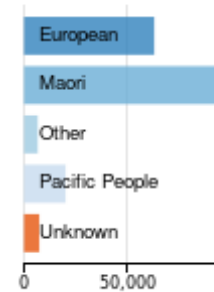
-
- Epoch – general purpose charting library
 - D3GH – 3d modelling
 - Insights.js – for graphs with nodes
 - Dashku – realtime dashboards
 - Rcharts – js visualizations from R
 - JSNetworkX – network processing (python)
 - Dimple – for business charts
 - Visual Sedimentation – streaming data
 - Vega – visualizing via css
 - Dynamic Graphs – realtime visualization
 - Datamaps – collection of interactive maps
 - Cubism.js – time series charts

New Zealand Youth Court Data

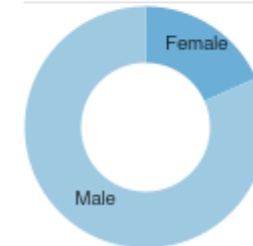
Offence



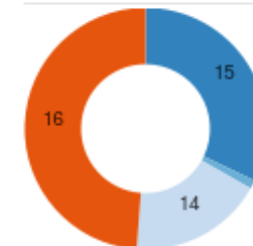
Ethnicity



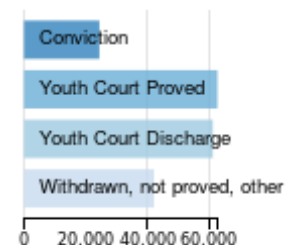
Gender



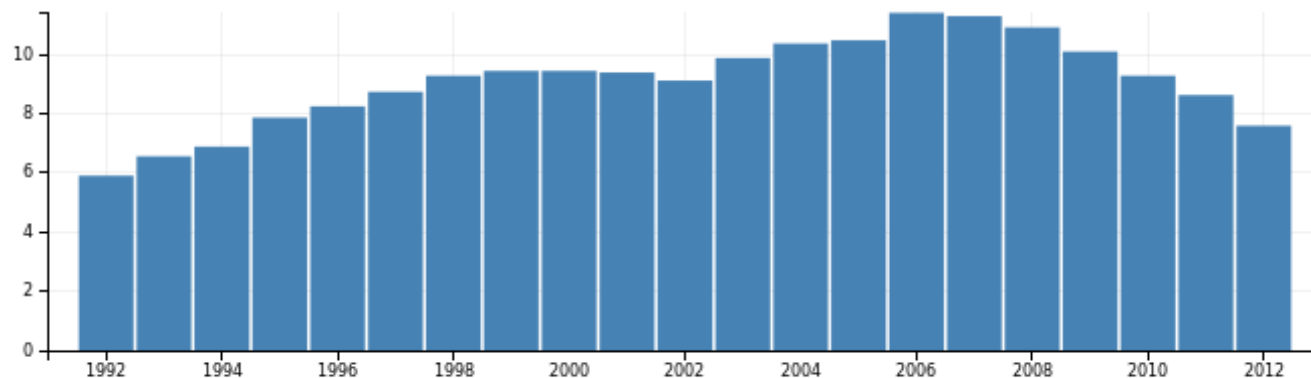
Age



Outcome



Year (in 1000's)



En plus, on peut intégrer des cartes à la volée dans ces bibliothèques



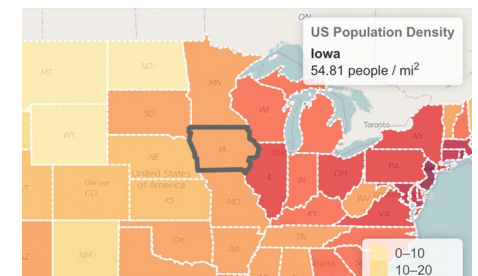
Permet les graphiques d'être filtrés aux mêmes temps, utilisant aussi une autre bibliothèque js:

Crossfilter

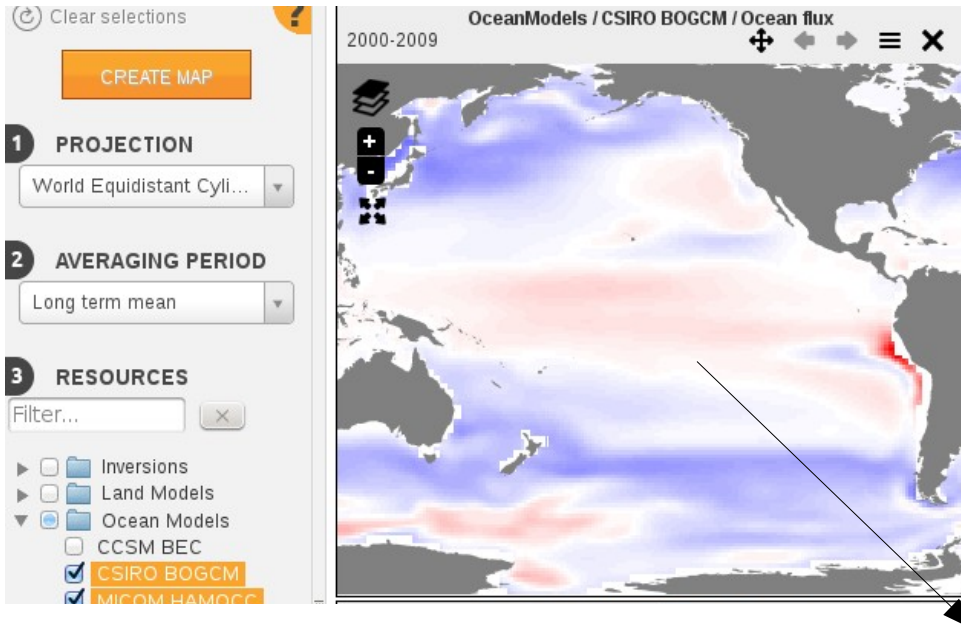
Fast Multidimensional Filtering for Coordinated Views

- Epoch – general purpose charting library
- D3GH – 3d modelling
- Insights.js – for graphs with nodes
- Dashku – realtime dashboards
- Rcharts – js visualizations from R
- JSNetworkX – network processing (python)
- Dimple – for business charts
- Visual Sedimentation – streaming data
- Vega – visualizing via css
- Dynamic Graphs – realtime visualization
- Datamaps – collection of interactive maps
- Cubism.js – time series charts

+ mapping tools



Les cartes sont générées à la volée par un serveur qui produit les tuiles aux différentes échelles



“Slippy maps”



Exploration interactive de données multi-dimensionnelles climatiques : exemple Extremoscope

Nous choisissons :

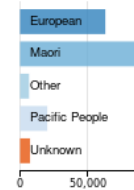
dc.js (inclus d3.js + crossfilter.js)
pour un `dashboard` interactive et inter-connecté



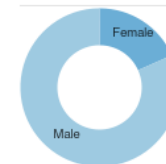
Offence



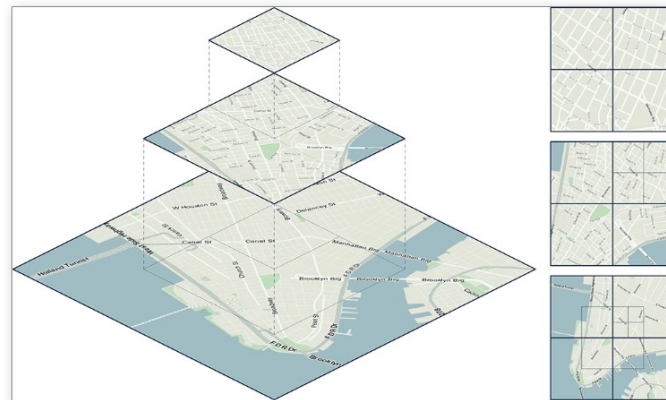
Ethnicity



Gender



dc-leaflet.js
pour les cartes



La visualisation Extremoscope affiche les événements extrêmes pour plusieurs dimensions

11 indices climatiques dans 2 catégories:
Heat (3) et Precipitation (8)

e.g.

Heat

TG : mean daily
temperature (°C)

Precipitation

RX1day: highest 1-day precipitation
amount (mm)

6 modèles climatiques : calculé par plusieurs groupes au tour de monde pour **2 scénarios**: RCP 4.5 et 8.5

échelle spatiale : 13 régions en France

échelles temporelles :

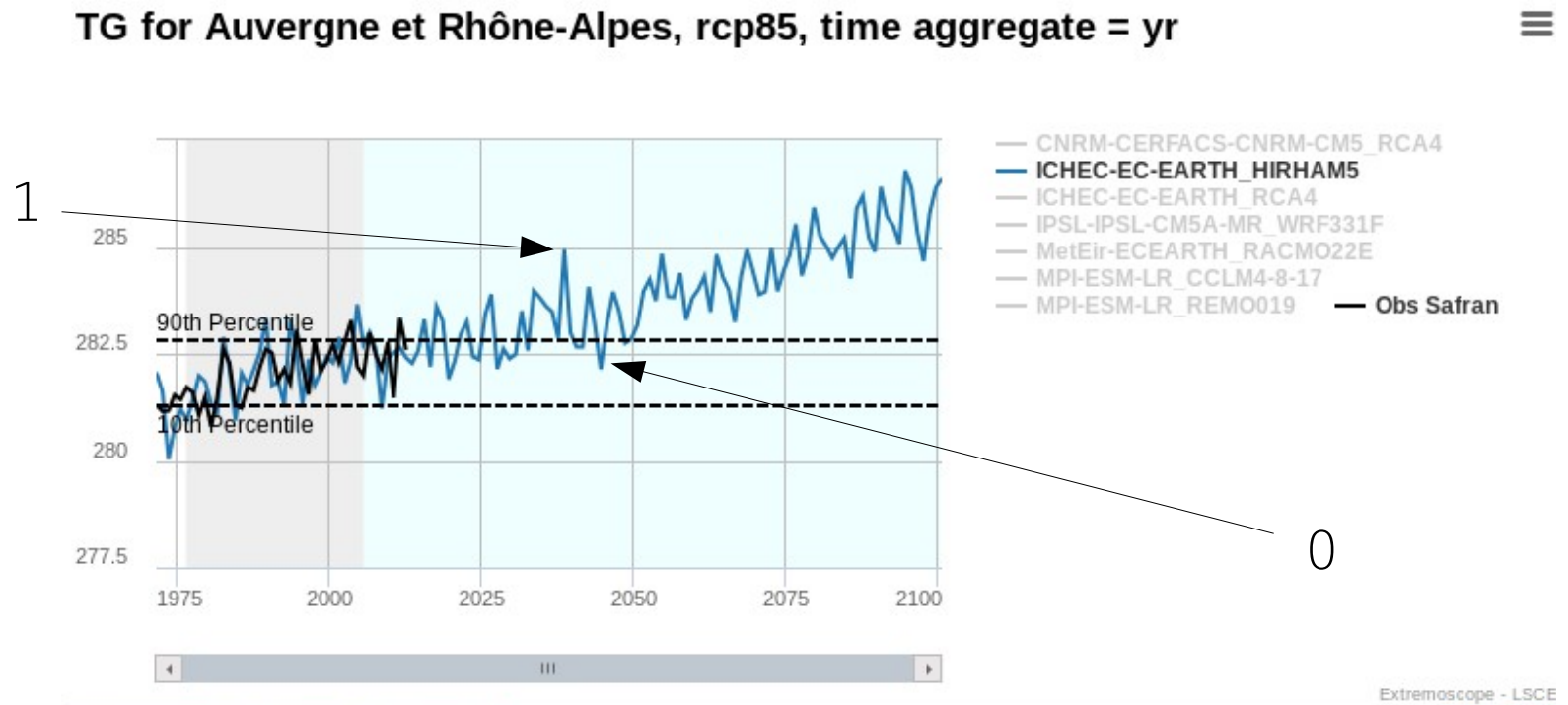
gamme des données: 1972 – 2100

mésures : par jour

aggregation par 4 saisons et toute l'année

Un point est extrême s'il dépasse le seuil (<10ème ou > 90ème centile)

dépasse le seuil = 1
ne dépasse pas le seuil = 0



on fait la moyenne des 0s et 1s sur une période du temps, et aussi les régions, les indices, etc. => on obtient la probabilité (%)

Les points qui dépassent le seuil sont sauvguardés dans un fichier avec les autres paramètres

Jeu de données (csv file) :

Toutes les indices qui passent une seuile (<10ème & > 90ème centile) sont sauvgardées dans un fichier.

Index	Model	Percentile	Region	Scenario	TimeAggregate	Year
1,	1,	90,	1,	1,	1,	1972
1,	1,	90,	1,	1,	1,	1973
1,	1,	90,	1,	1,	1,	1975
...						
1,	1,	10,	1,	1,	1,	1978
1,	1,	10,	1,	1,	1,	1979
1,	1,	10,	1,	1,	1,	1983

=> Plusieurs dimensions, plusieurs échelles

Comment les organise ?

Traditionnellement...



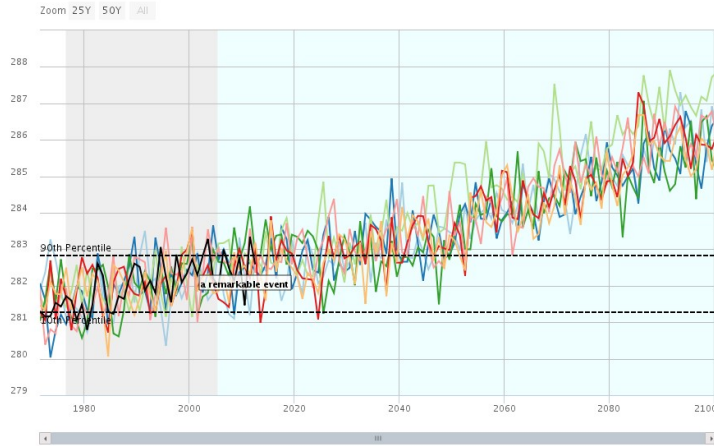
Auvergne

From 1971 To 2100

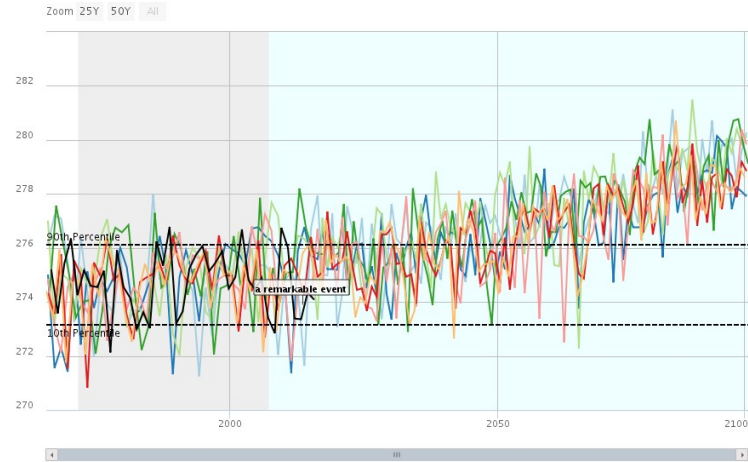
- CNRM-CERFACS-CNRM-CM5_RCA4
- ICHEC-EC-EARTH_HIRHAM5
- ICHEC-EC-EARTH_RCA4
- IPSL-IPSL-CM5A-MR_WRF331F
- MetEir-ECEARTH_RACMO22E
- MPI-ESM-LR_CCLM4-8-17
- MPI-ESM-LR_REMO019
- Obs Safran

pick TG (mean daily temp)

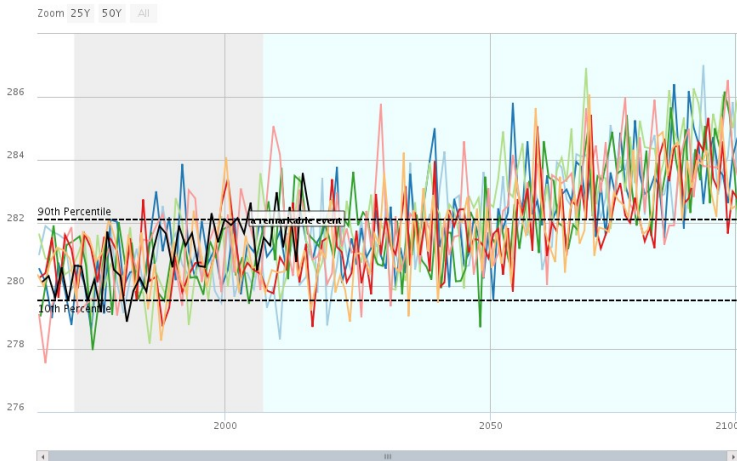
TG for Auvergne et Rhône-Alpes, rcp85, time aggregate = yr



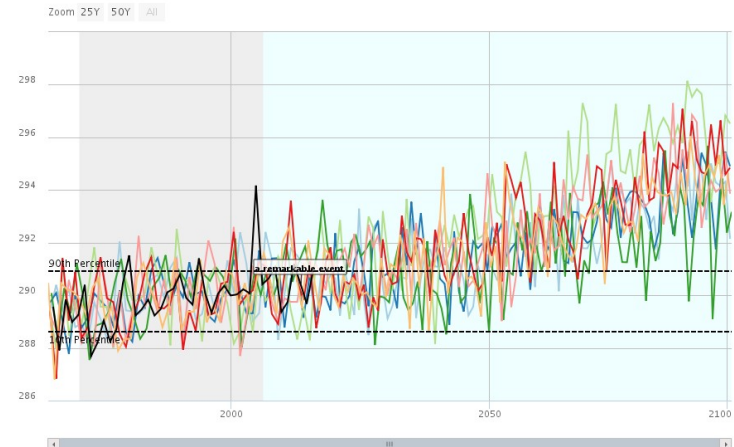
TG for Auvergne et Rhône-Alpes, rcp85, time aggregate = DJF



TG for Auvergne et Rhône-Alpes, rcp85, time aggregate = MAM



TG for Auvergne et Rhône-Alpes, rcp85, time aggregate = JJA



Voici notre "dashboard" interactif fait avec dc.js qui contient tout l'information

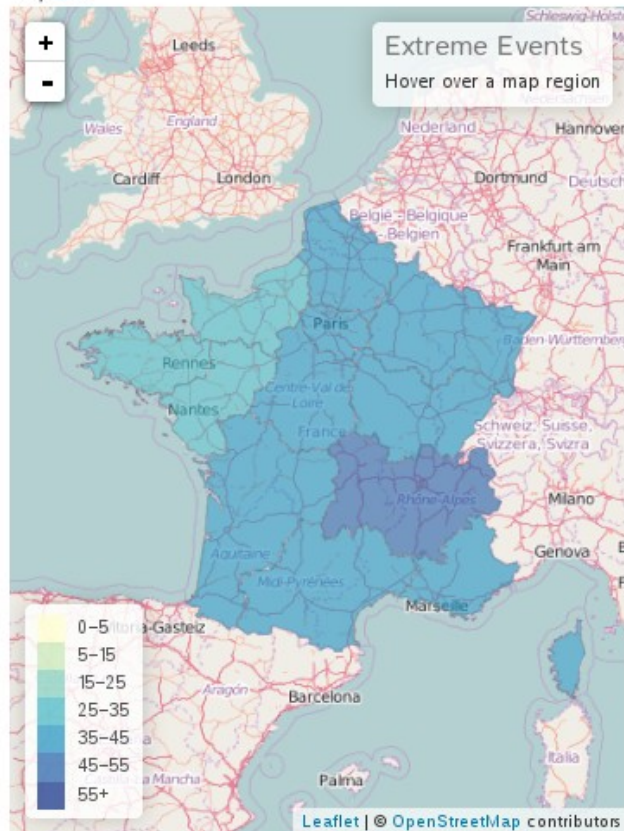
Extremoscope: Evolution of extremes in France [Help](#) [Contact](#)

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

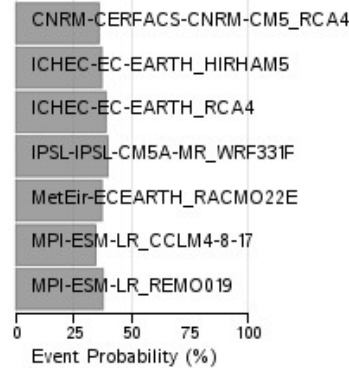
[View](#) Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.

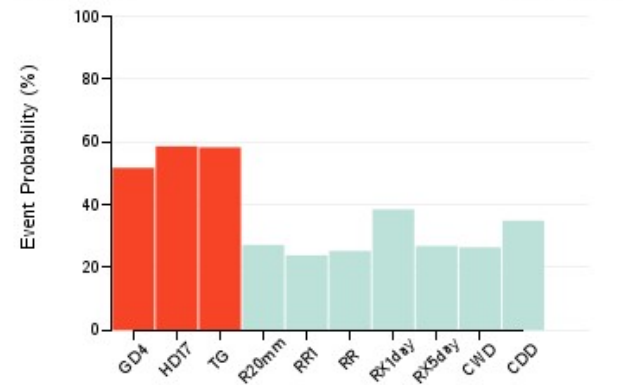
Map [reset](#)



Model



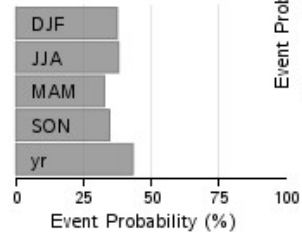
Climate Index



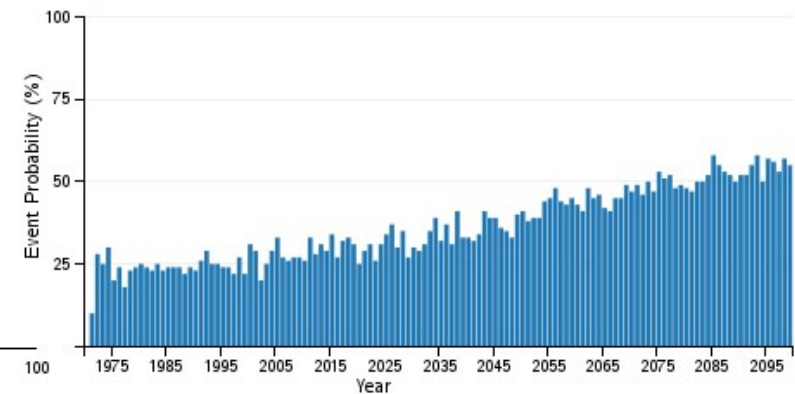
Category



Time Aggregate



Year range

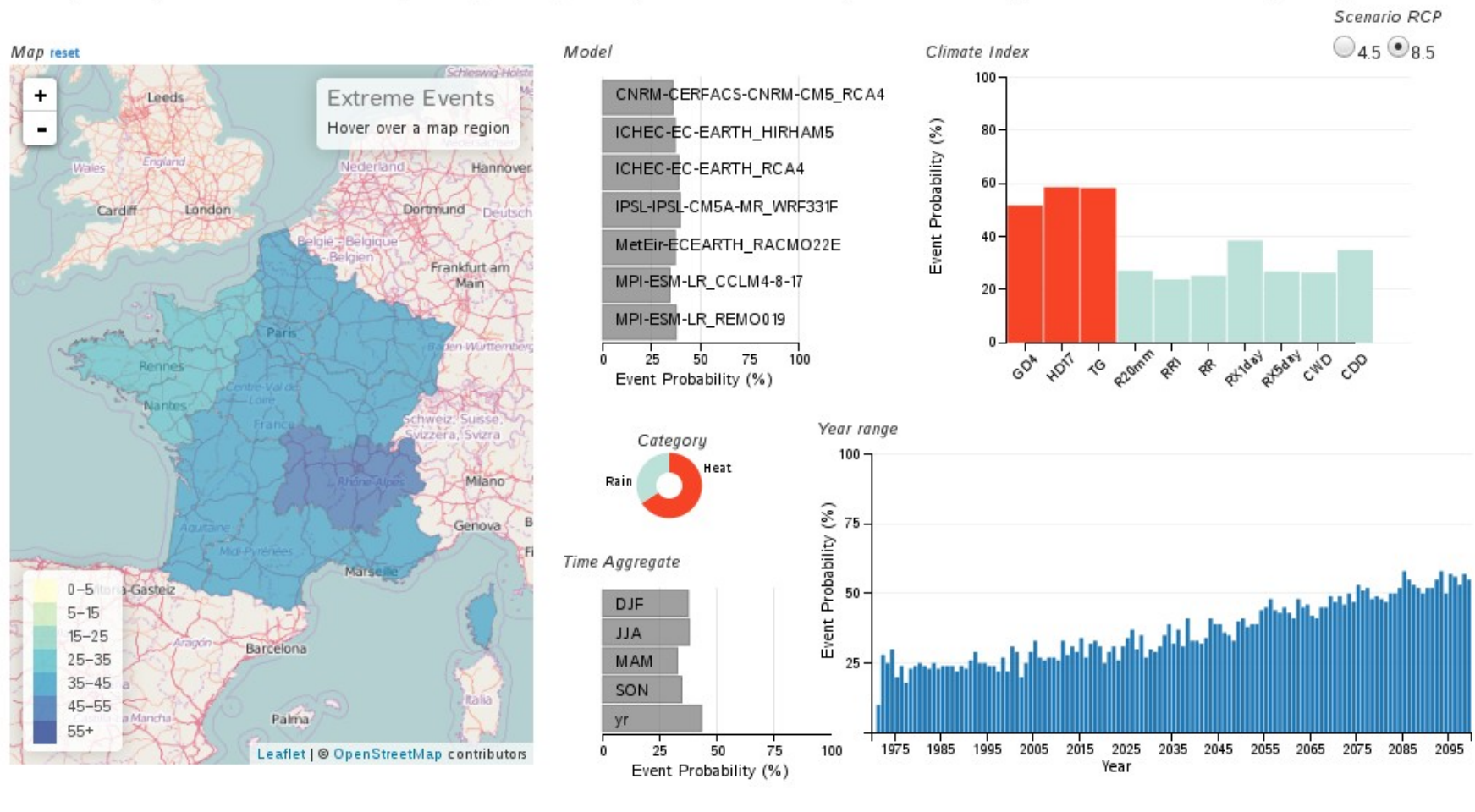


1: Les modèles semblent qu'ils sont d'accord

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

View Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.



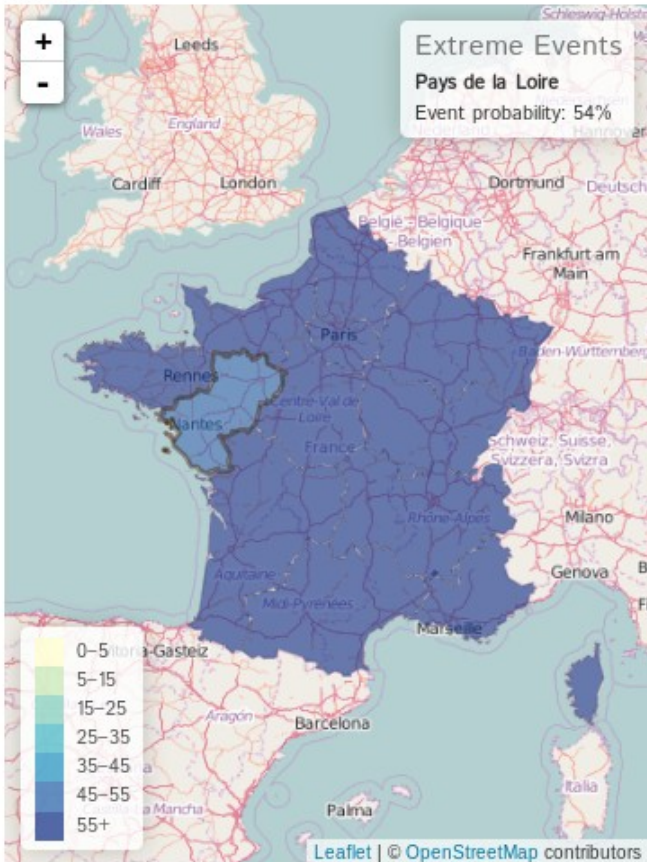
2: TG (mean daily temperature) augmente pour chaque période, année et dans tous les régions

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

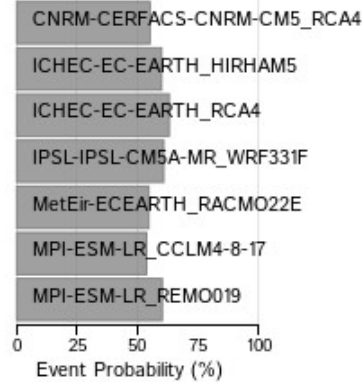
View Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.

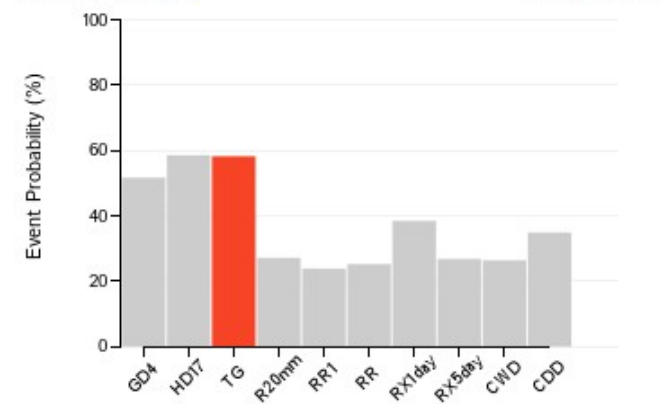
Map reset



Model



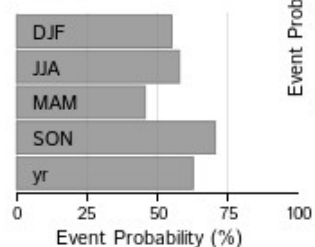
Climate Index reset



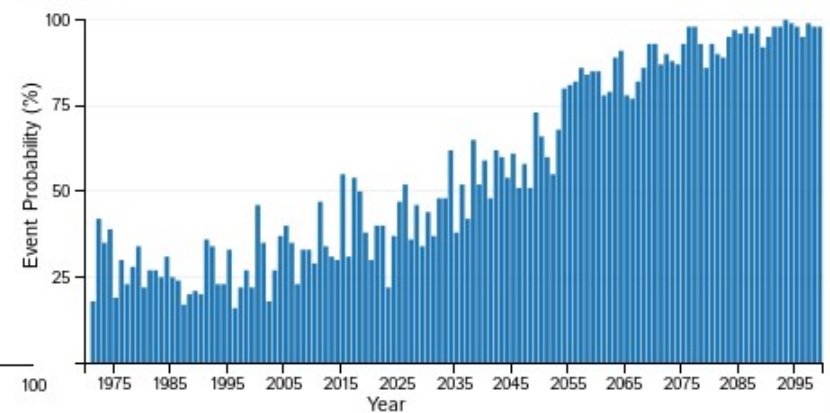
Category



Time Aggregate



Year range



3: Les indices de température et une indice de précip (RX1day) seront extrêmes dans l'avenir

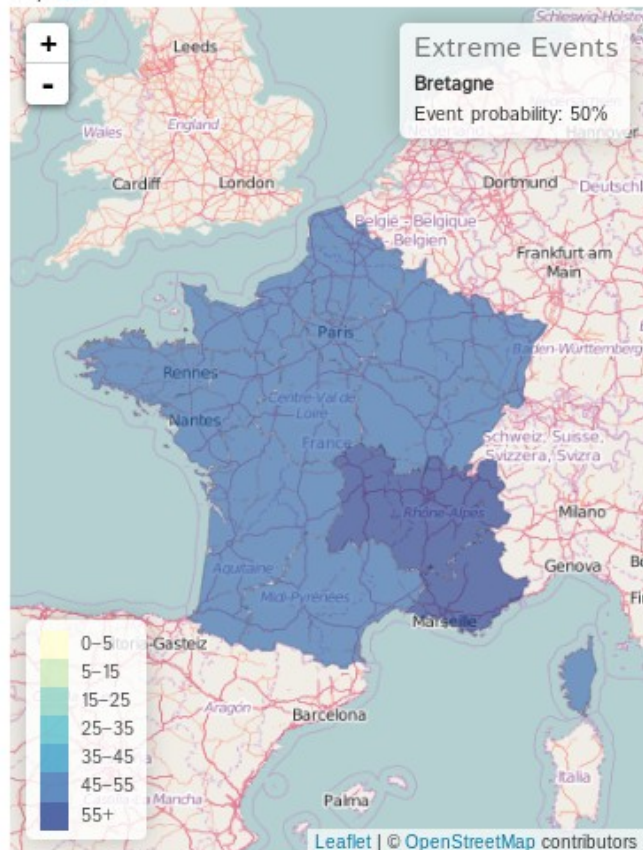
Extremoscope: Evolution of extremes in France Help Contact

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

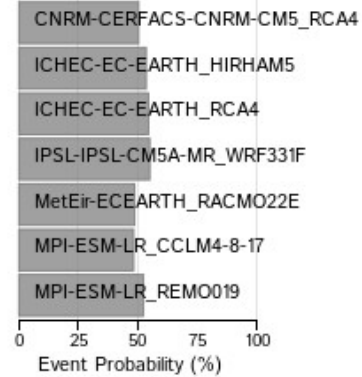
View Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.

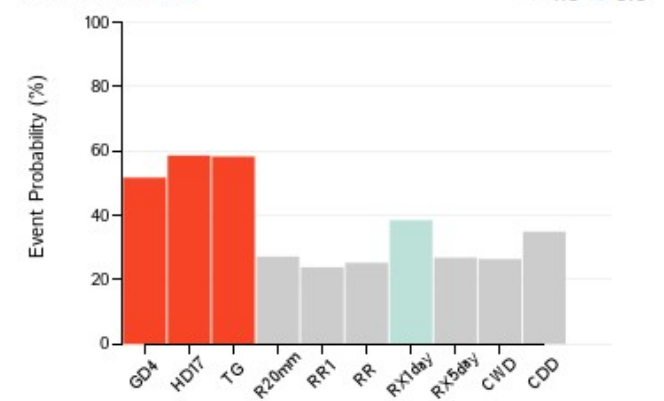
Map reset



Model



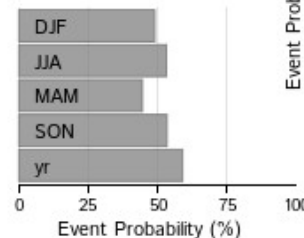
Climate Index reset



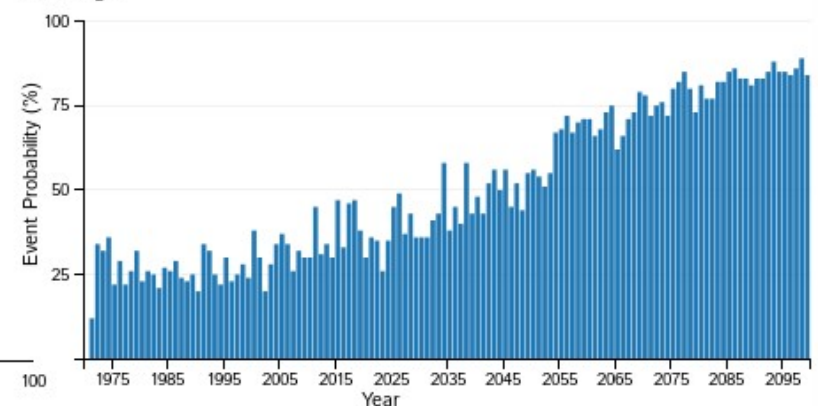
Category



Time Aggregate



Year range



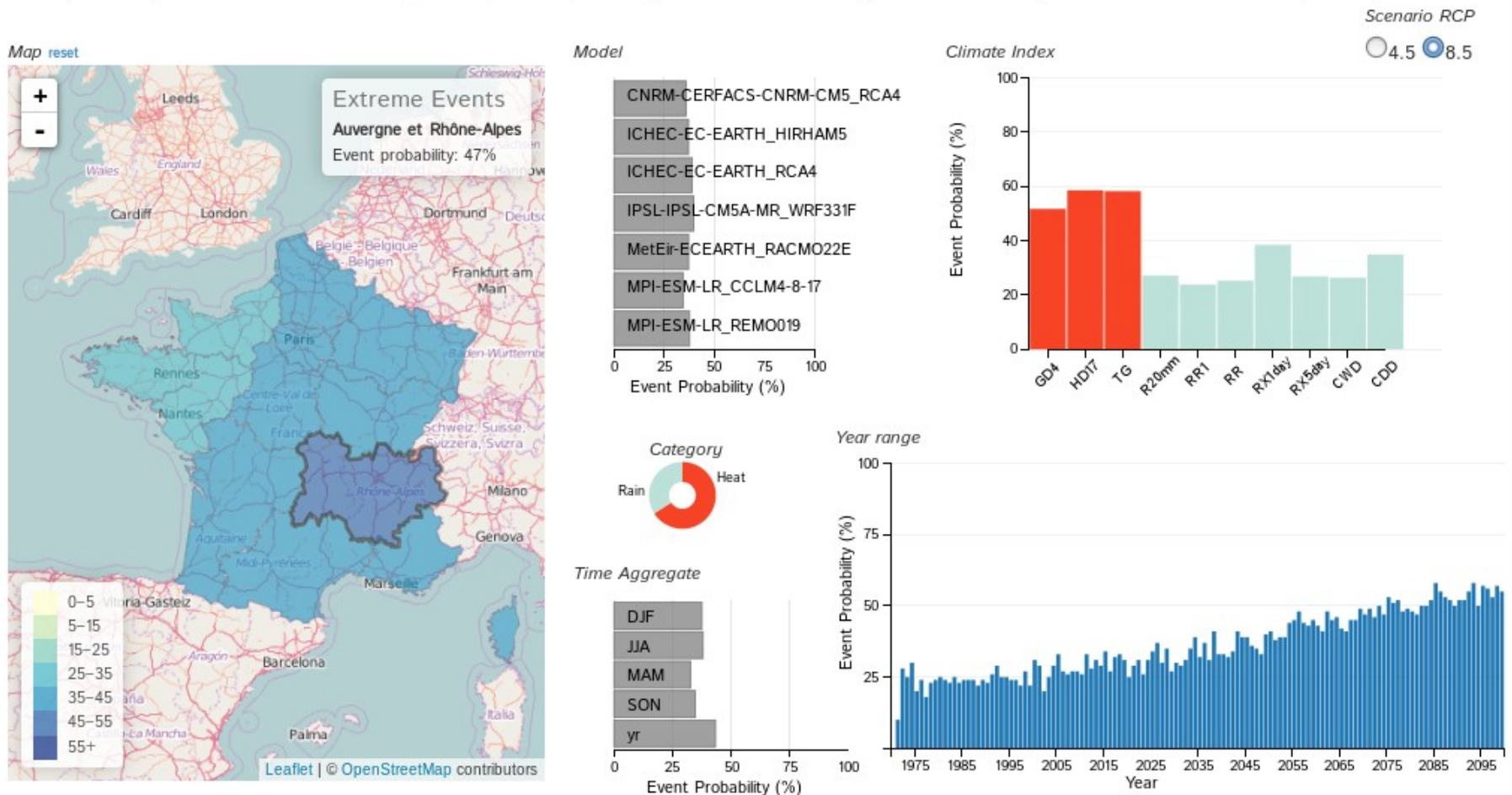
4. La région montagneuse (Auvergne et Rhône-Alpes) est la plus impactée par des extrêmes

Extremoscope: Evolution of extremes in France Help Contact

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

View Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.



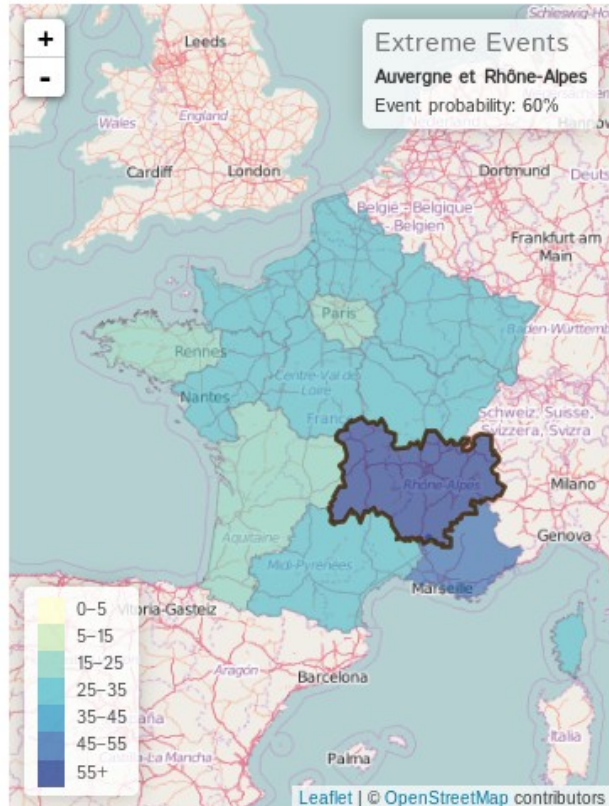
5: Pour la precip en Auvergne et Rhône-Alpes, l'hiver et le plus extrême saison

Projected evolution of extreme seasons and years in France from several indices based on 7 models. Extremes are characterized by values less than the 10th or above the 90th percentile using the reference period 1976 – 2005.

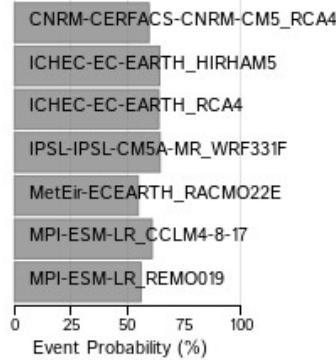
View Select **one** index and **one** region to view the timeseries.

Start by selecting one or more climate indices, then explore the probability of extreme seasons or years in different regions for different models and year ranges.

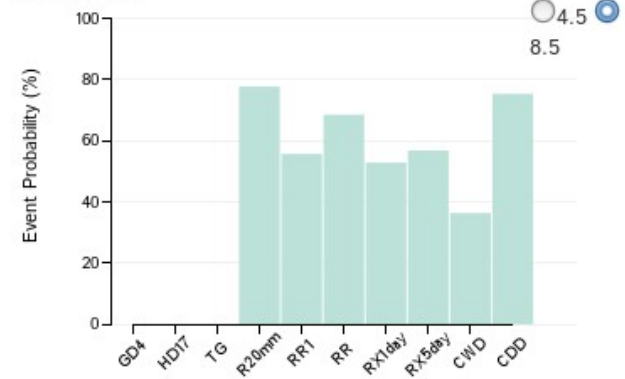
Map reset



Model



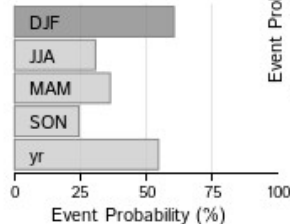
Climate Index



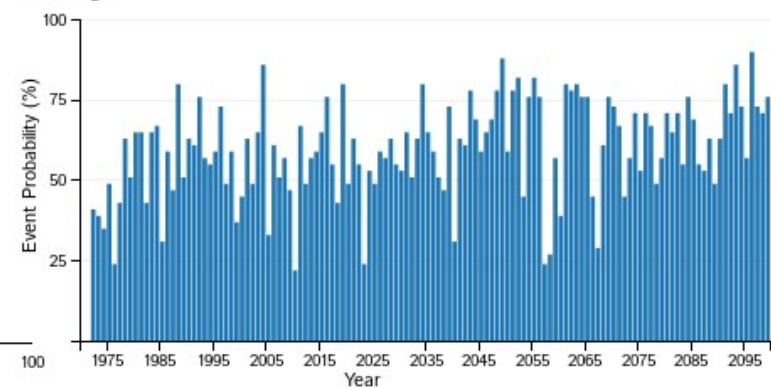
Category reset



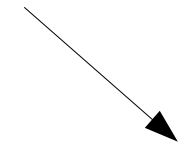
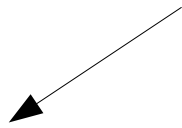
Time Aggregate reset



Year range



Mais ces technologies web autrefois réservées aux développeurs sont accessibles au monde scientifique



Les développeurs

Embedding D3 in an IPython Notebook

Python + Leaflet
= folium



Python bokeh +
crossfilter



networkD3:

D3 JavaScript Network Graphs from R



Les scientifiques

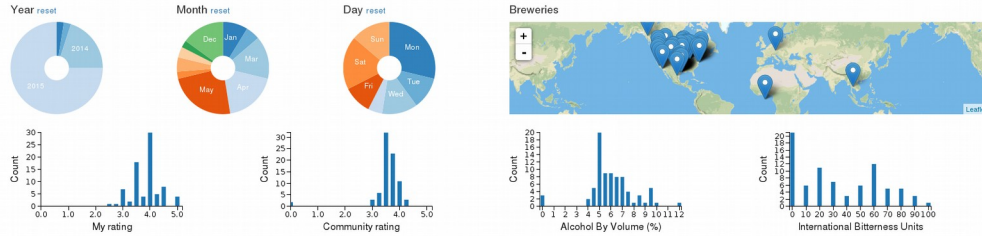


Toujours de plus en plus !

dc.js pur

La même chose dans un ipython notebook (+ folium)

Beer History 80 selected out of 80 records | Reset All



Brewery	Beer	Style	My Rating	Community Rating	ABV %	IBU
Goose Island Beer Co.	Harvest Ale	Harvest Ale	5	3.25	5.5	30
Firestone Walker Brewing Company	Double Jack	Imperial / Double IPA	5	4.25	9.5	80
Lagunitas Brewing Company	A Little Sumpin' Sumpin' Ale	American Pale Wheat Ale	5	4	7.5	60
Dogfish Head Craft Brewery	90 Minute IPA	Imperial / Double IPA	5	4	9	90
Lagunitas Brewing Company	Undercover Investigation Shut-Down Ale	American Strong Ale	4.5	4	9.5	60
Stone Brewing Co.	Stone Enjoy By 12.05.14 IPA	Imperial / Double IPA	4.5	4.25	9.5	80
Boulevard Brewing Co.	Bourbon Barrel Quad	Belgian Strong Dark Ale	4.5	4	12	10
Boulevard Brewing Co.	Nutcracker Ale	Winter Warmer	4.5	3.5	8	30
Great Divide Brewing Company	Yeti	American Imperial / Double Stout	4.5	4	9.5	70
Dogfish Head Craft Brewery	60 Minute IPA	American IPA	4.5	4	6	60
Stone Brewing Co.	Stone Enjoy By 04.20.15 IPA	Imperial / Double IPA	4.5	4.25	9.5	80
Sury Brewing Company	Furious	American IPA	4.5	4	6.5	100
Odell Brewing Co.	90 Shilling	Scotch Ale / Wee Heavy	4.25	3.75	5.5	20

Jupyter nbviewer

JUPYTER FAQ

```

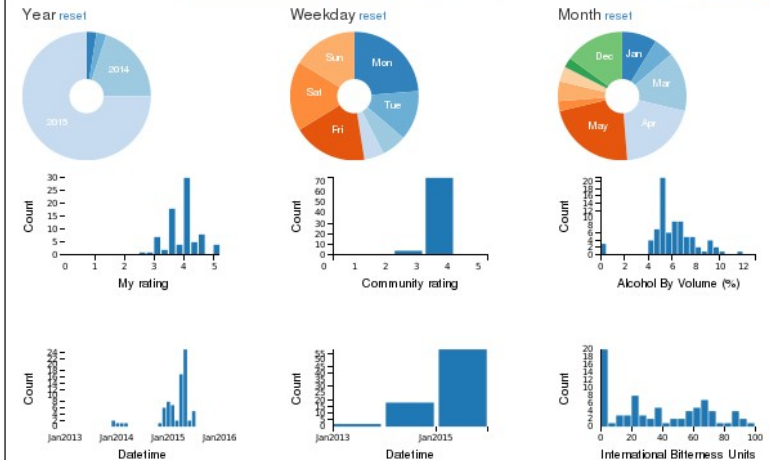
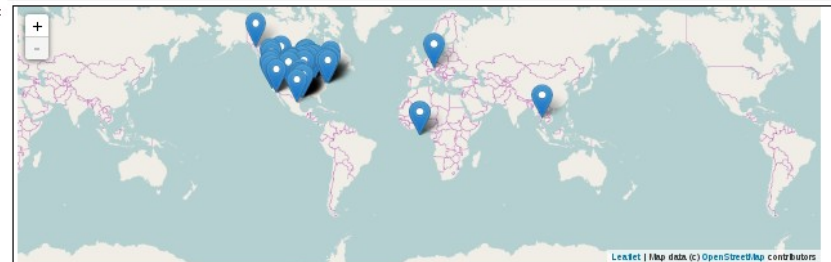
width=150, height=150, inner_radius=20,
order=x['order'], class_='col-xs-4 pie-chart'))

for i, x in enumerate(paraml['barChart']):
    row = row2 if i<3 else row3
    row.add_children(BarFilter(
        c,
        x['column'],
        width=250,
        height=150,
        bar_padding=0.1,
        domain=x['domain'],
        groupby=x['groupby'],
        xlabel=x['xAxisLabel'],
        ylabel=x['yAxisLabel'],
        margins=x['margins'],
        xticks=x['xAxisTickValues'],
        time_format=x.get('timeFormat', None),
        class_='col-xs-4 dc-chart'))

row4.add_children(TableFilter(c,
    ['brewery_brewery_name', 'beer_beer_name', 'beer_beer_style', 'rating_score',
    'beer_rating_score', 'beer_beer_abv', 'beer_beer_ibu'
    ],
    class_='table table-bordered table-striped'))

```

Out[7]:



brewery_brewery_name	beer_beer_name	beer_beer_style	rating_score	beer_rating_score	beer_beer_abv	beer_beer_ibu
Hops & Grain	Porter Culture	Baltic Porter	3.5	3.777	6.56	0

L'évolution de la data viz est arrivée au navigateur et a ouvert une grande gamme de possibilités



Il y a plein d'outils pour visualiser les données de manière interactive. Ouverts aussi aux scientifiques !

Merci à :

Patrick Brockmann, LSCE

Nils Hempelmann, LSCE

Robert Vautard, LSCE

Pascal Yiou, LSCE