



Data Engineering

Sylvain BARTHELEMY

Qui suis-je?

gwen**lake**.

Build next-gen apps using AI

We help organizations by operationalizing AI, machine learning and data analytics

Déroulement du cours

Objectifs

Construction et mise en application d'un
projet data / ML opérationnel.

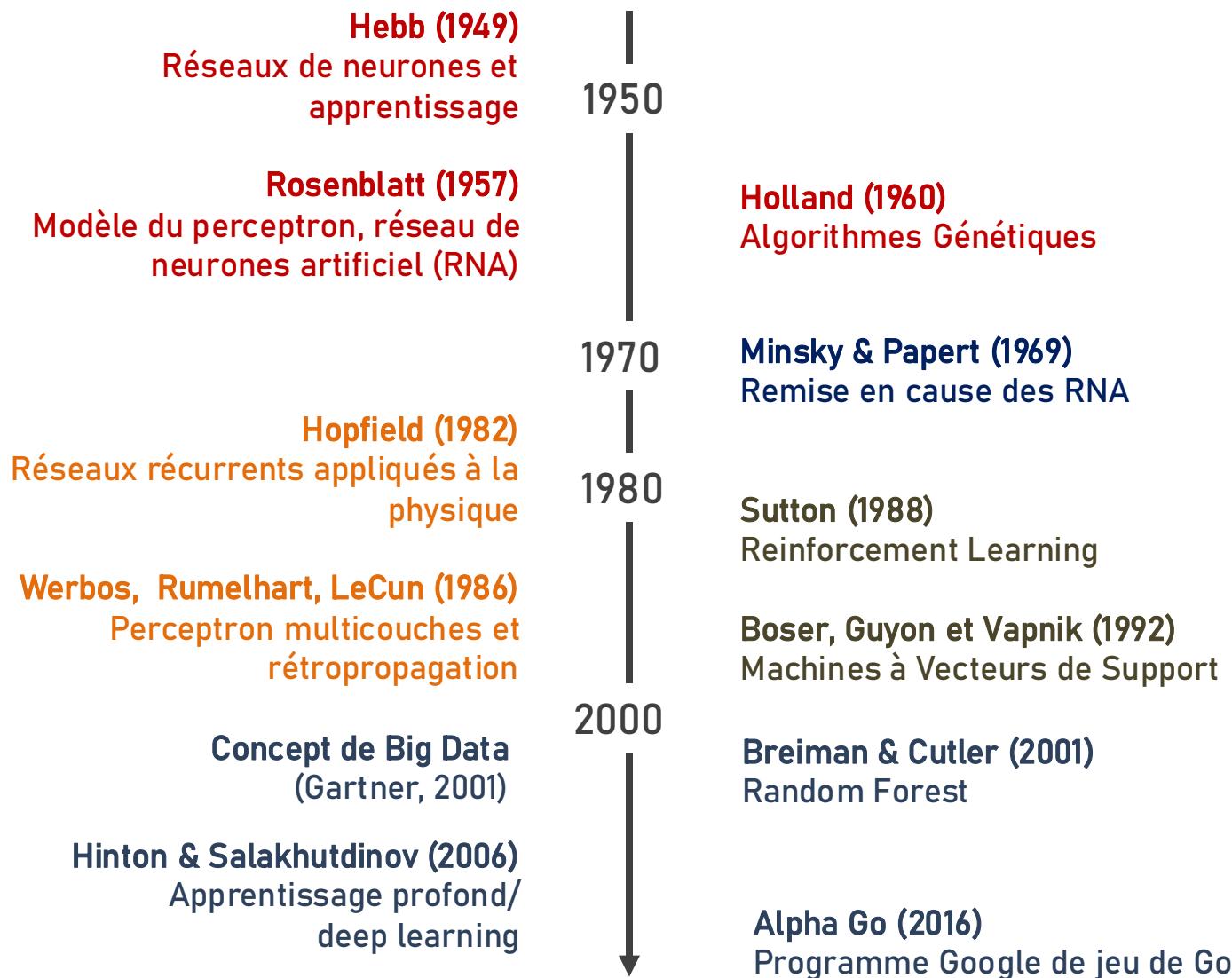
Introduction des concepts suivants

- Introduction
- Découverte de Docker
- Création d'une base de données pour stocker les données
- Récupération et upload de données dans la base
- Création d'indicateurs pertinents
- Création d'une API pour interroger le système
- Interrogations sous Postman
- Présentation d'un projet

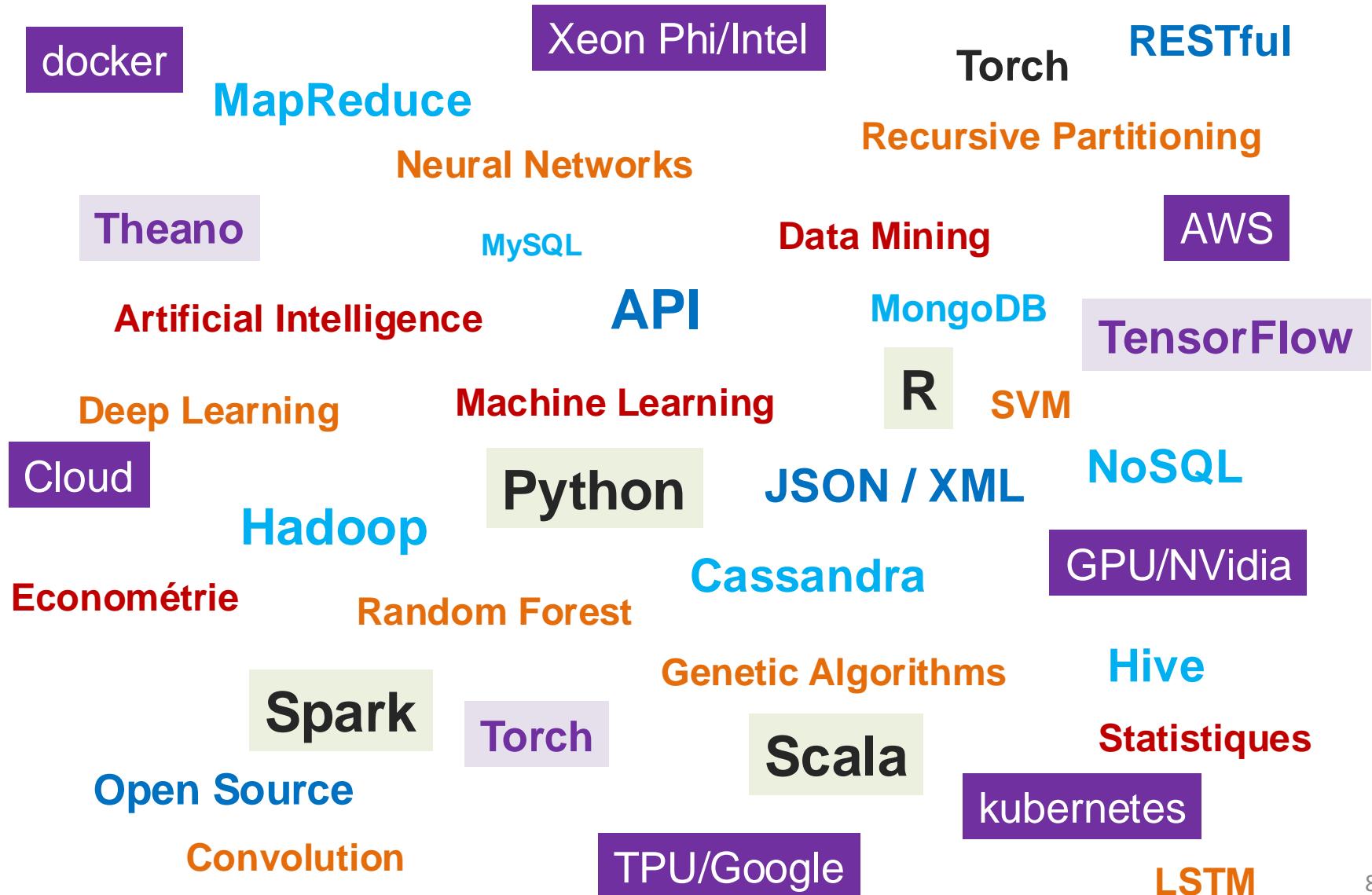
A dark blue background featuring a complex network of light blue dots connected by thin white lines, forming a mesh-like structure.

De nouveaux concepts et de nouvelles compétences

Une brève histoire du machine learning et de l'IA

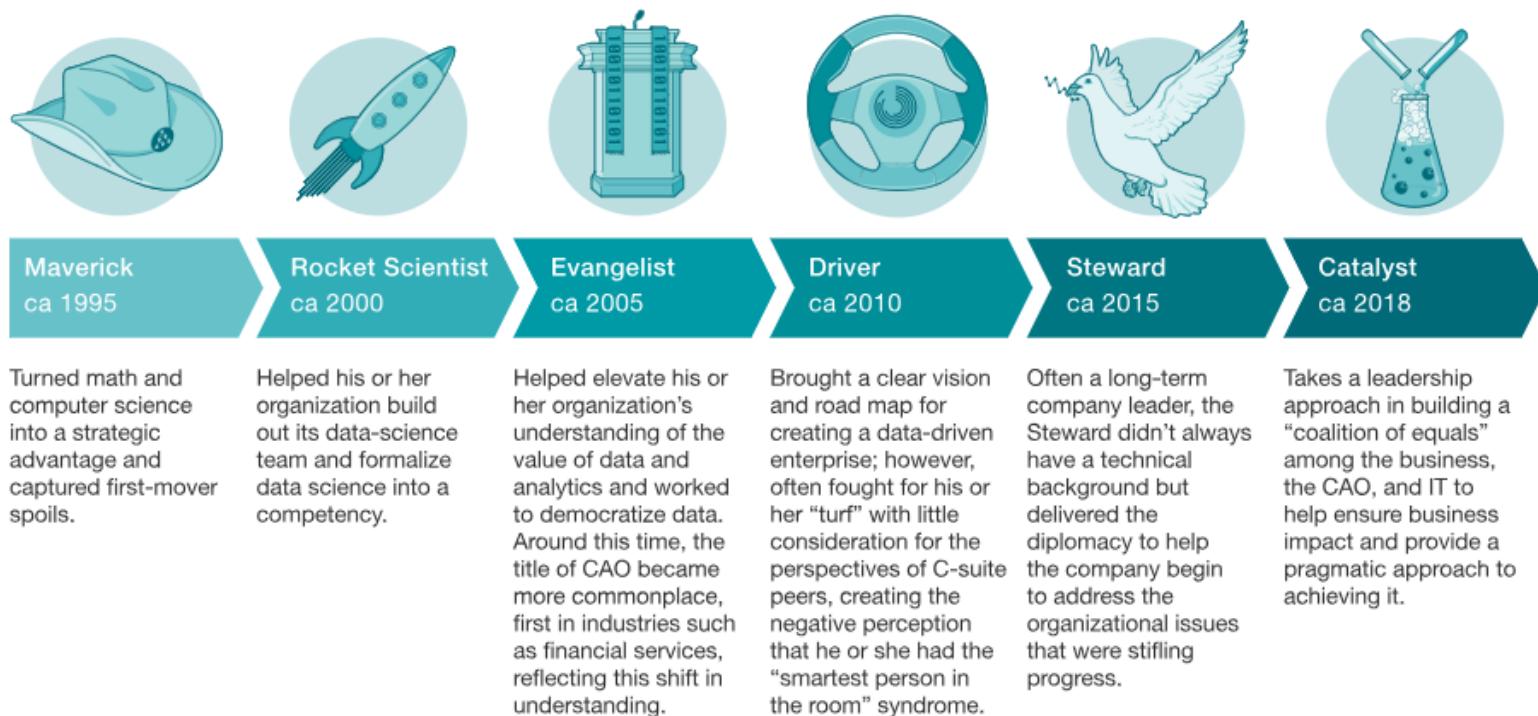


Big data and AI

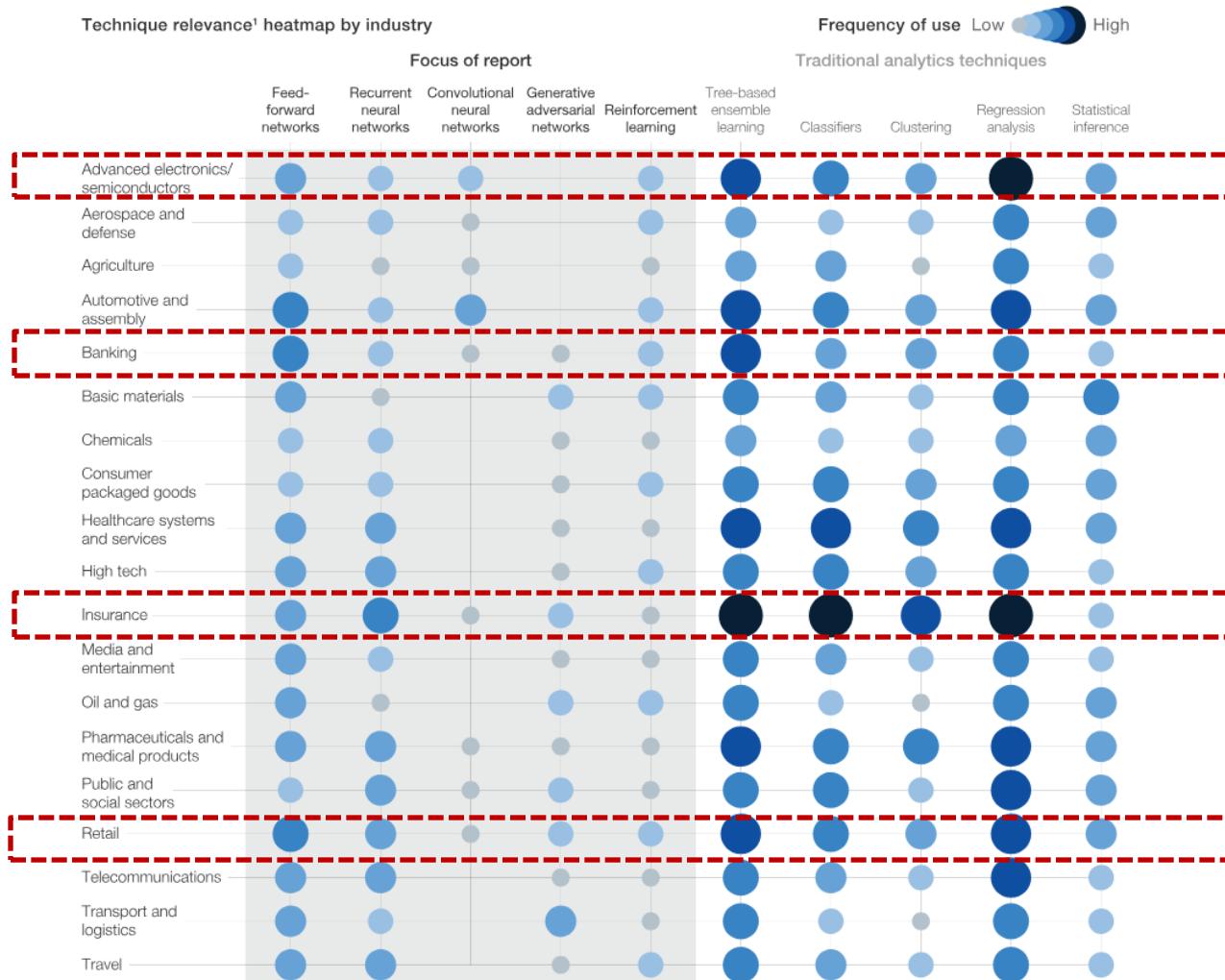


The evolving role of the Chief Analyst Officer

Dominant CAO personas emerged as data and analytics advanced over the past quarter century.



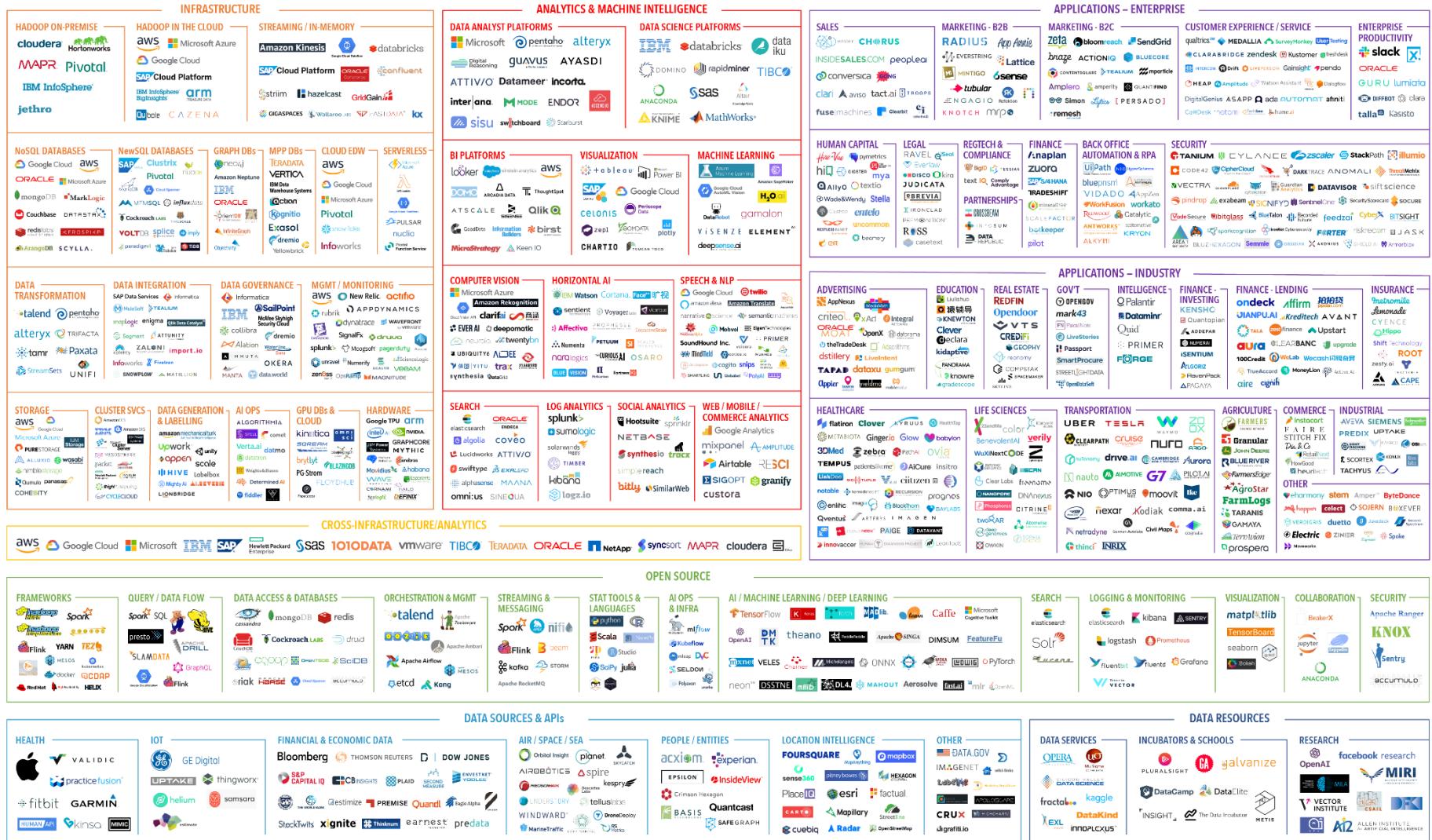
Traditional and advanced analytics techniques



¹Relevance refers to frequency of use in our use case library, with the most frequently found cases marked as high relevance and the least frequently found as low relevance. Absence of circles indicates no or statistically insignificant number of use cases.
Note: List of techniques is not exhaustive.

L'écosystème du « datascience » et du « big data »

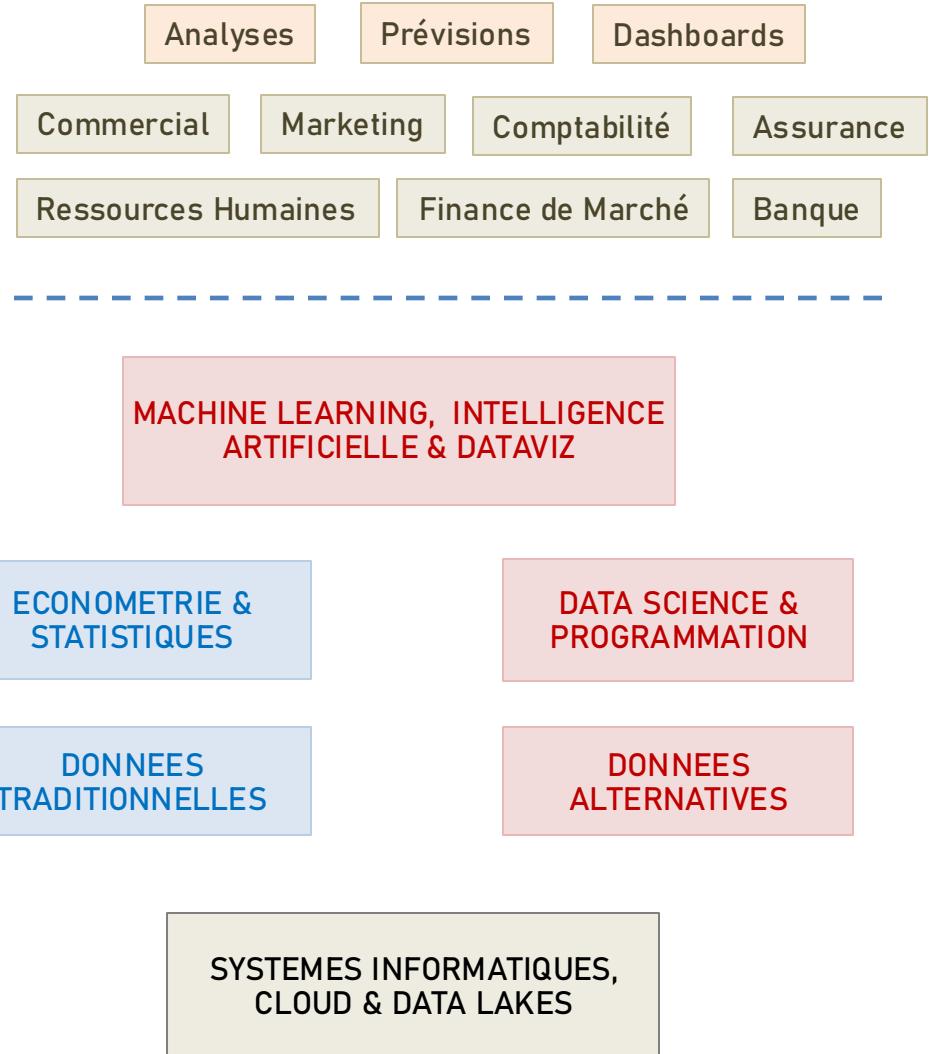
DATA & AI LANDSCAPE 2019



Is it useful for me?

**Even if you are not data-scientists, one day or another,
you will have to work with data-scientists**

Le syndrome de l'iceberg



Les langages

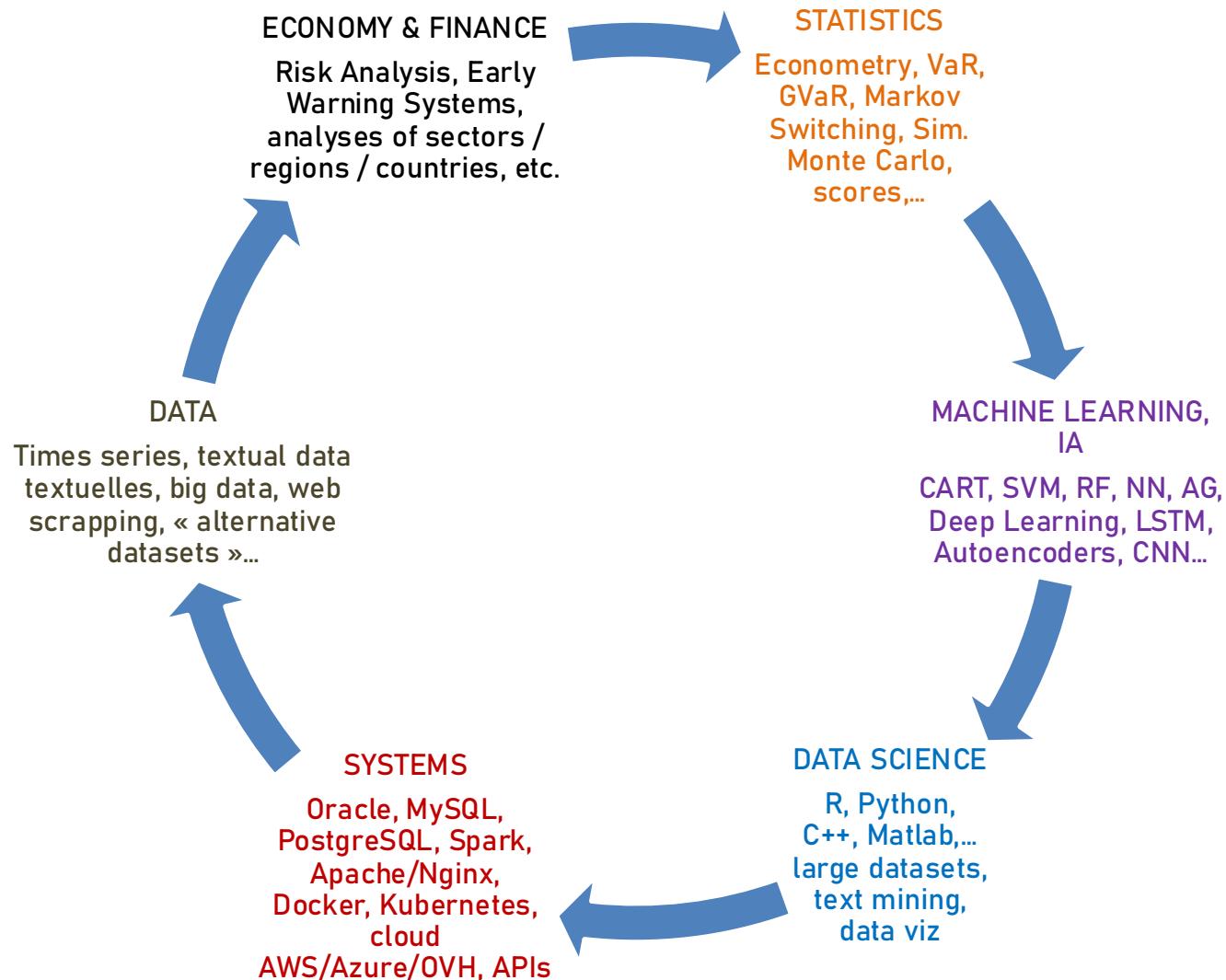
Classement IEEE Spectrum Top Programming Languages 2021

IEEE Spectrum Top Programming Languages				Q	T
Rank	Language	Type	Score		
1	Python▼	🌐💻⚙️	100.0		
2	Java▼	🌐📱💻	95.4		
3	C▼	📱💻⚙️	94.7		
4	C++▼	📱💻⚙️	92.4		
5	JavaScript▼	🌐	88.1		
6	C#▼	🌐📱💻⚙️	82.4		
7	R▼	💻	81.7		
8	Go▼	🌐💻	77.7		
9	HTML▼	🌐	75.4		
10	Swift▼	📱💻	70.4		
11	Arduino▼	⚙️	68.4		

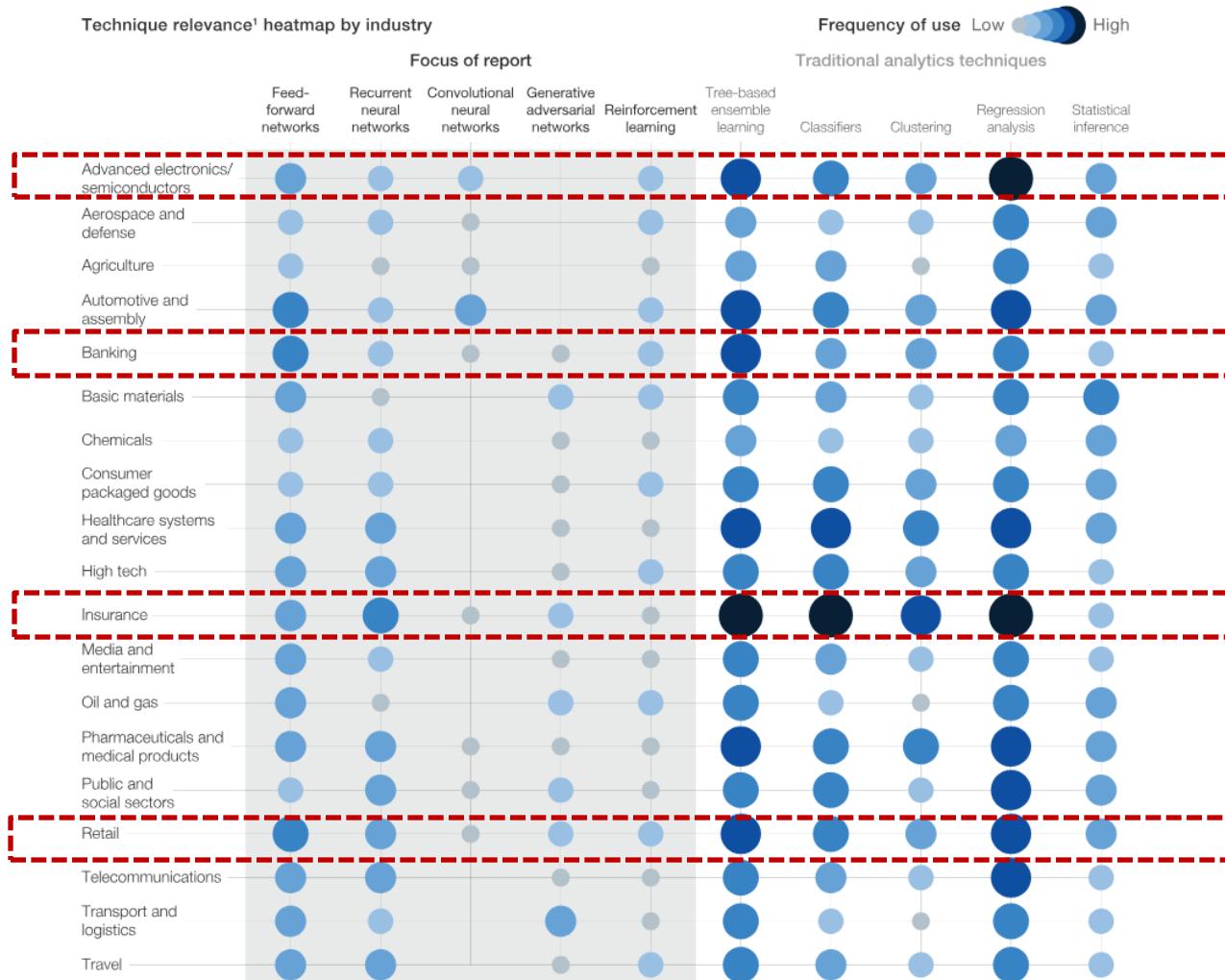
Data Engineering ?

Elaboration et structuration des flux de données
pour une exploitation optimale.

Data Engineering et Data Scientists « full stack »



Traditional and advanced analytics techniques



Crises Economiques & Financières

Prévoir les crises économiques et financières (EWS)

- De nombreuses études empiriques sur le risque pays et les indicateurs avancés de crises économiques et financières: Krugman (1979), Obstfeld (1994), Cantor and Packer (1996), Eichengreen et al. (1996), Frankel and Rose (1996), Goldstein (1996), Goldstein and Turner (1996), Kaminsky and Reinhart (1999), Komulainen and Lukkarila (2003) , ...
- Mais malgré les classifications existantes, un très grand nombre de crises économiques et financières ont laissé les observateurs perplexes.
- Le manque de relations causales homogènes et des interactions complexes rendent l'identification ex-ante des facteurs de risque extrêmement difficile.

Supervised & Unsupervised Methods

SUPERVISED

UNSUPERVISED

A quoi servent ces algorithmes ?

- Très performants pour identifier des patterns, des combinaisons critiques (non linéaires): cas du skieur, modèles à deux variables, signaux faibles et apprentissage en temps réel.
- Beaucoup d'entre eux sont presque entièrement « automatisables » (à la différence des approches économétriques traditionnelles)
- Attention au risque de sur-apprentissage, au « machine learning sauvage », aux risques de bases de données « poubelles » (données erronées, pas de notion d'intégration/stationnarité, liens économiques), phénomènes de « taches solaires », attention au sampling, etc...
- Certains algorithmes sont des « boites noires »... mais pas tous !

From econometry to data-science

Term in Statistics/Econometry	Equivalent in Machine Learning
Independent Variable, X	Input Feature, attribute, pattern
Dependent Variable, Y	Output Feature, response, label
In-Sample	Training Set
Out-of-Sample	Test Set
Estimate a Model	Learn a Model
Model Parameters	Model Weights
Regression	Supervised Learning
Clustering	Unsupervised Learning
Accuracy, R2	Sensitivity, Specificity, ROC, Likelihood

Supervised & Unsupervised Methods

SUPERVISED

Explain relationships between inputs and targets

Examples: economic crises, inflationary episodes, corporate defaults, etc....

Tools: econometry, linear discriminant analysis, neural networks/perceptrons, SVM, Random Forests, deep learning, etc...

UNSUPERVISED

Analyze a dataset that has not been « labelled »

Examples: analogies between countries, classifying companies (beyond sector & size), discovering macroeconomic patterns,

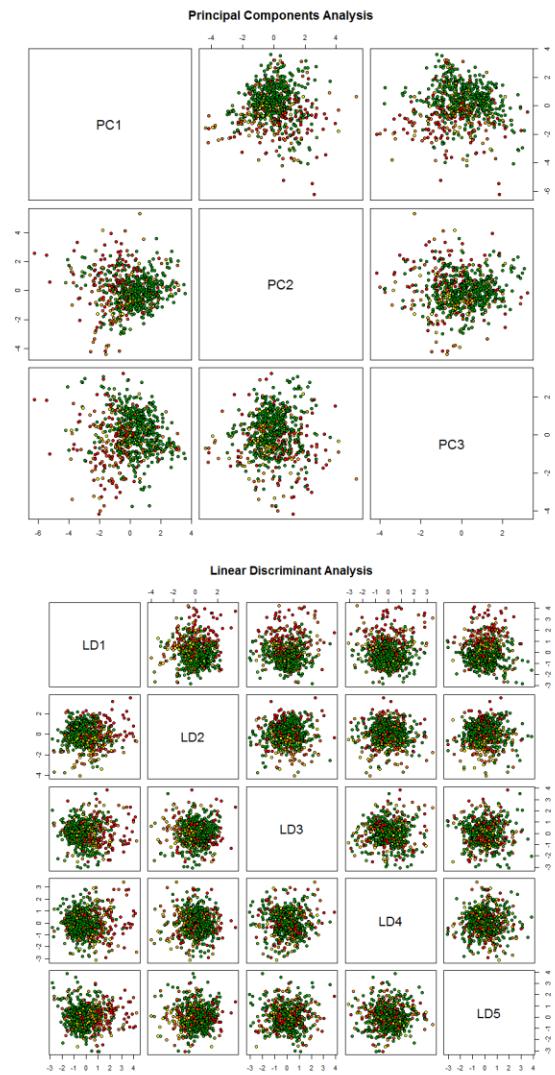
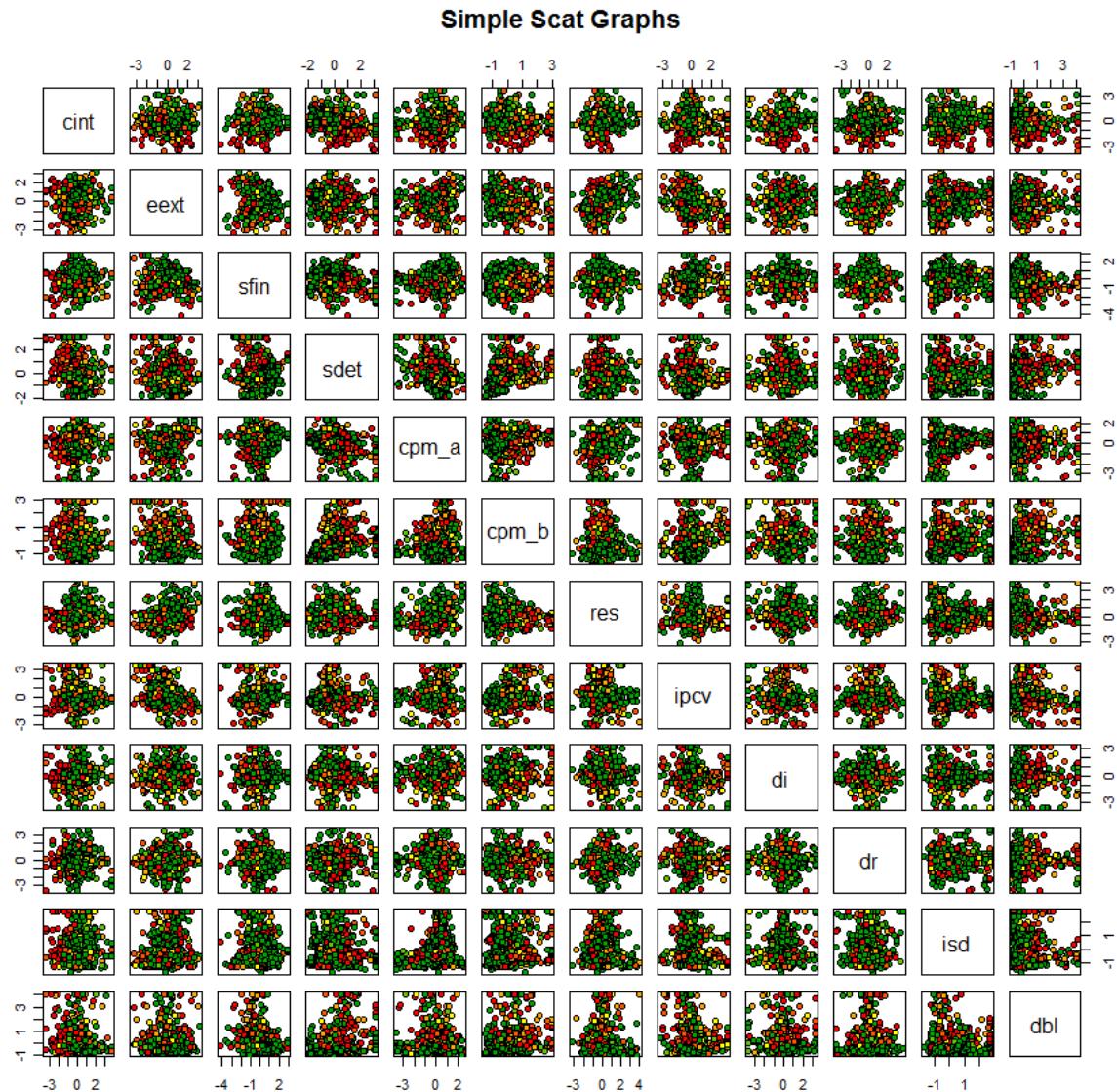
Tools: principal components analysis, k-means, clustering algorithms, self organizing maps, etc...

TAC ECONOMICS/RiskMonitor Fundamental Balances

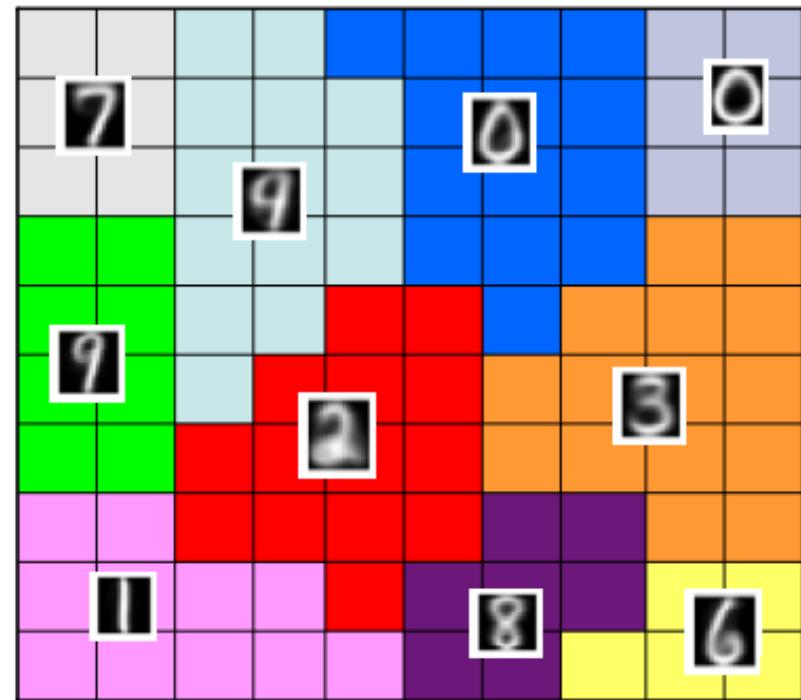
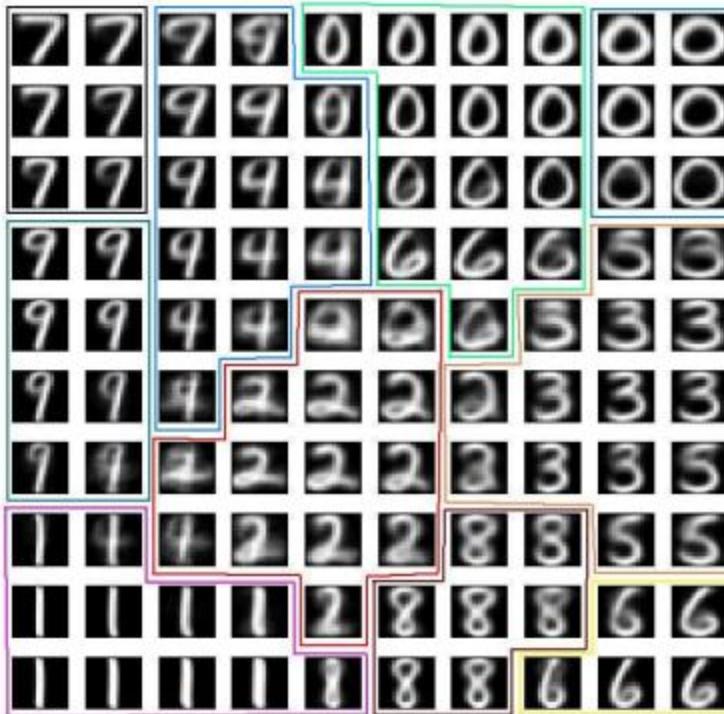
Macroeconomic indicators of the RiskMonitor fundamental balances

Indicator	Periodicity	Description
Economic growth	Annual	GDP growth
External balance	Annual	External balance sustainability
Financing stability	Annual	Stability of FDI inflows
Debt service	Annual	External financing
Forex liquidity	Quarterly	Foreign currency situation
Maximum potential service	Quarterly	Short-term foreign currency liabilities
Forex reserves quality	Quarterly	Dynamics in forex reserves
Exch. rate competitiveness	Quarterly	International competitiveness of exchange rate
Monetary pressure	Quarterly	Quality of monetary policy
Real economic pressure	Quarterly	Momentum of domestic activity
Domestic leverage	Quarterly	Activity and banks' health
Foreign financing	Quarterly	Dependence on foreign financing

Les méthodes traditionnelles



Self Organizing Maps... et Tessellations de Voronoi



Source

Data Analysis using Self-Organizing Maps

Marie Cottrell and Patrick Letrémy

https://samos.univ-paris1.fr/IMG/pdf_Porvoo_Kohonen_Data_Analysis_V3-2.pdf

World Poverty Map using Self Organizing Maps

An example on World Bank data, by Samuel Kaski

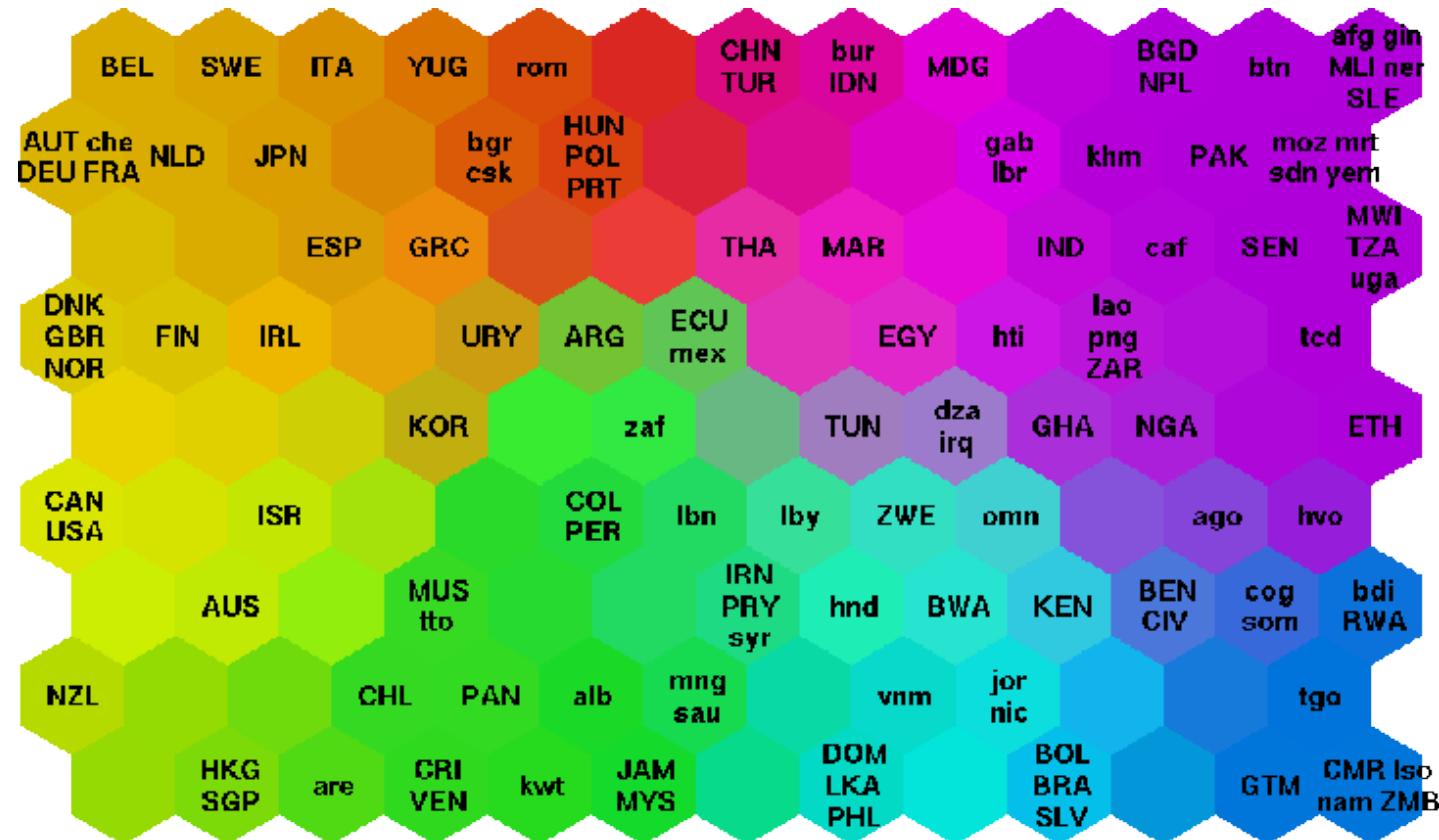
- 39 World Bank country indicators describing various quality-of-life factors, such as state of health, nutrition, educational services, etc, were used.
- The complex joint effect of these factors can be visualized by organizing the countries using the self-organizing map.
- The map consists of a regular grid of processing units, "neurons". A model of some multidimensional observation, eventually a vector consisting of features, is associated with each unit.
- The map attempts to represent all the available observations with optimal accuracy using a restricted set of models. At the same time the models become ordered on the grid so that similar models are close to each other and dissimilar models far from each other.

Source: <http://www.cis.hut.fi/research/som-research/worldmap.html>

World Poverty Map using Self Organizing Maps

An example on World Bank data, by Samuel Kaski

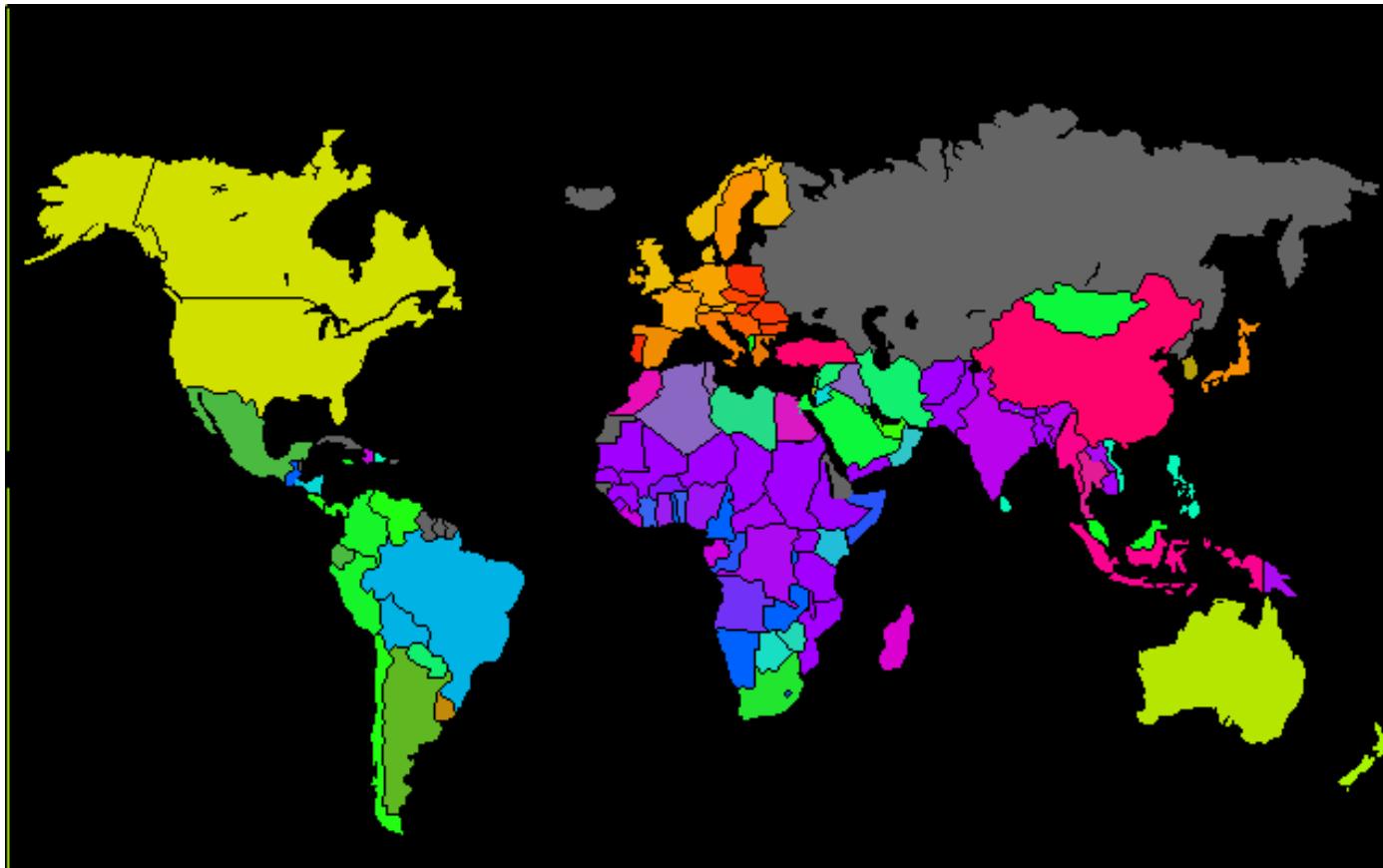
each country is automatically assigned a color describing its poverty type in relation to other countries



Source: <http://www.cis.hut.fi/research/som-research/worldmap.html>

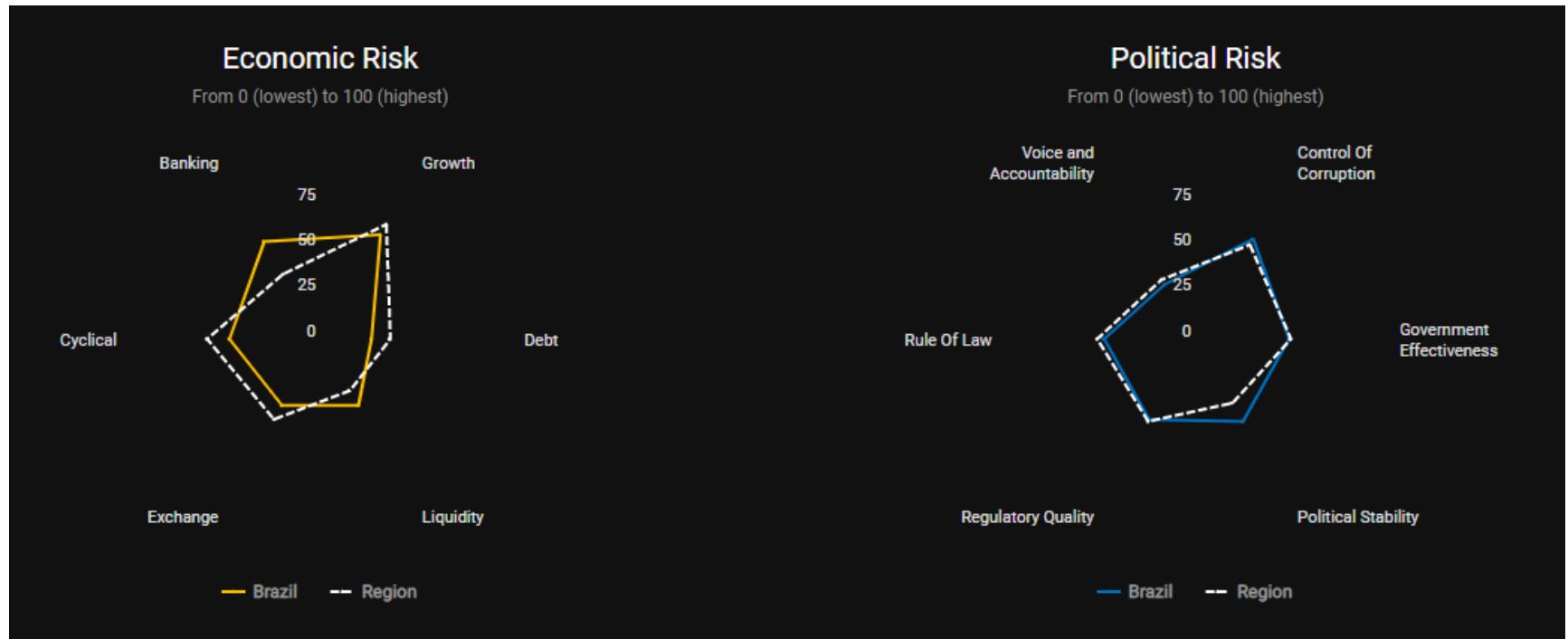
World Poverty Map using Self Organizing Maps

An example on World Bank data, by Samuel Kaski

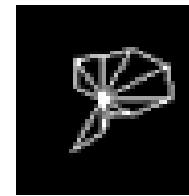
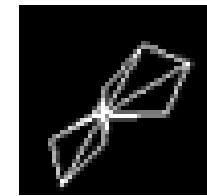
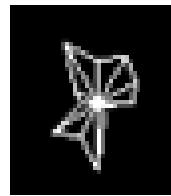


Source: <http://www.cis.hut.fi/research/som-research/worldmap.html>

Web charts on macroeconomic indicators



Des pays, des indicateurs, des années... et des papillons

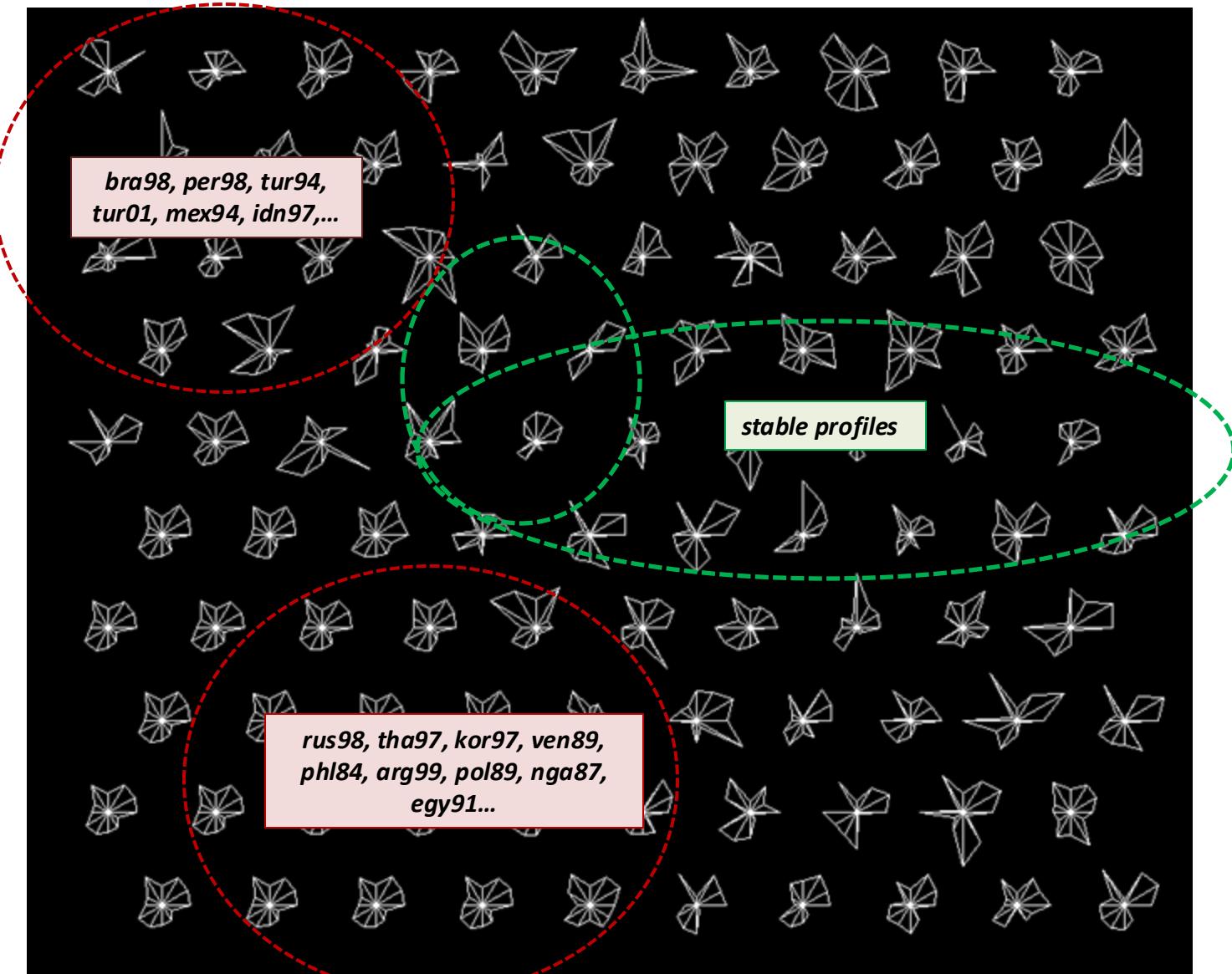


on more than 100 countries since the 80s

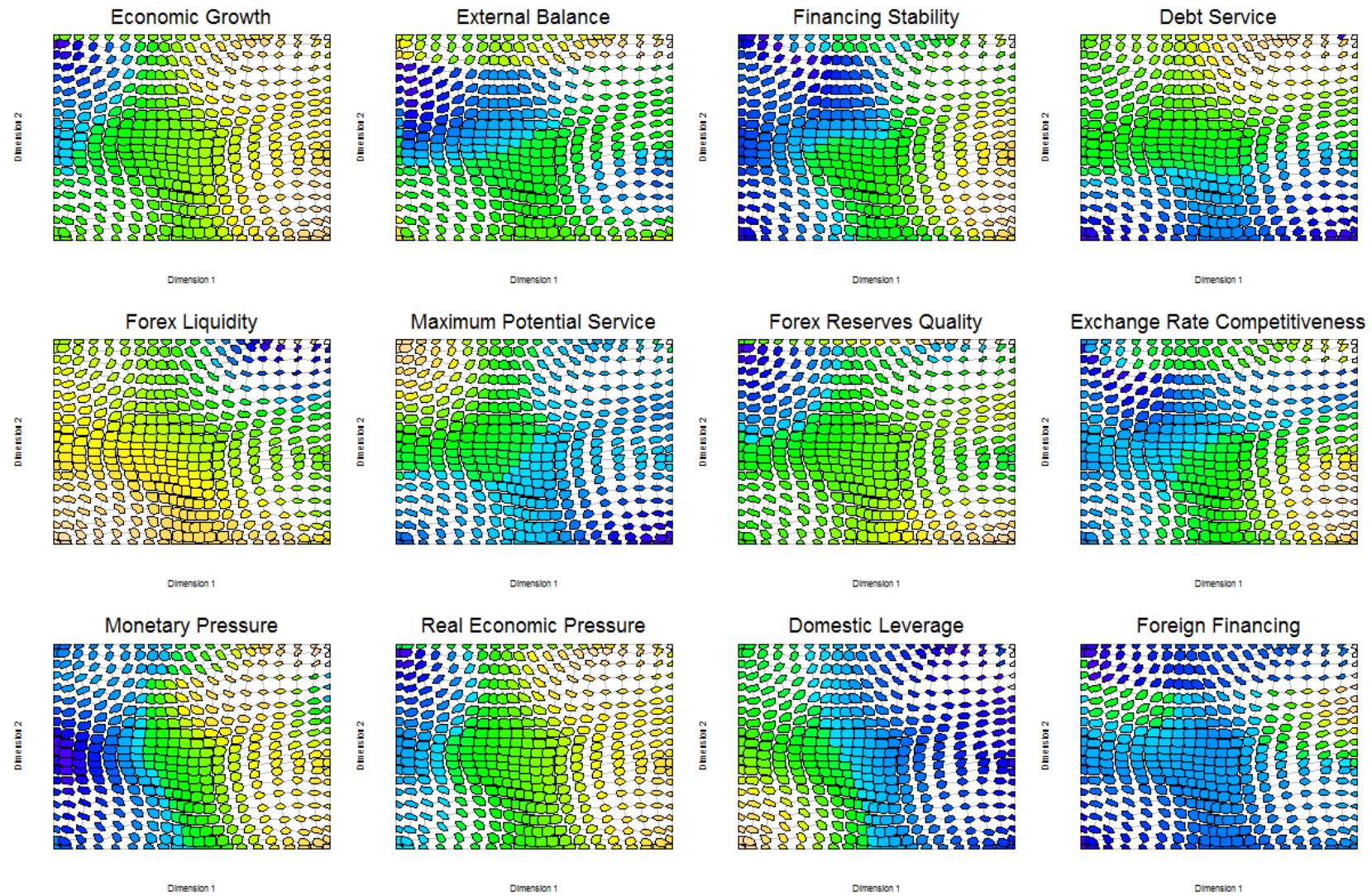
100 countries x 40 years x 4 quarters per year

16,000 country web charts to analyse!

Similarités macroéconomiques et réseaux de neurones



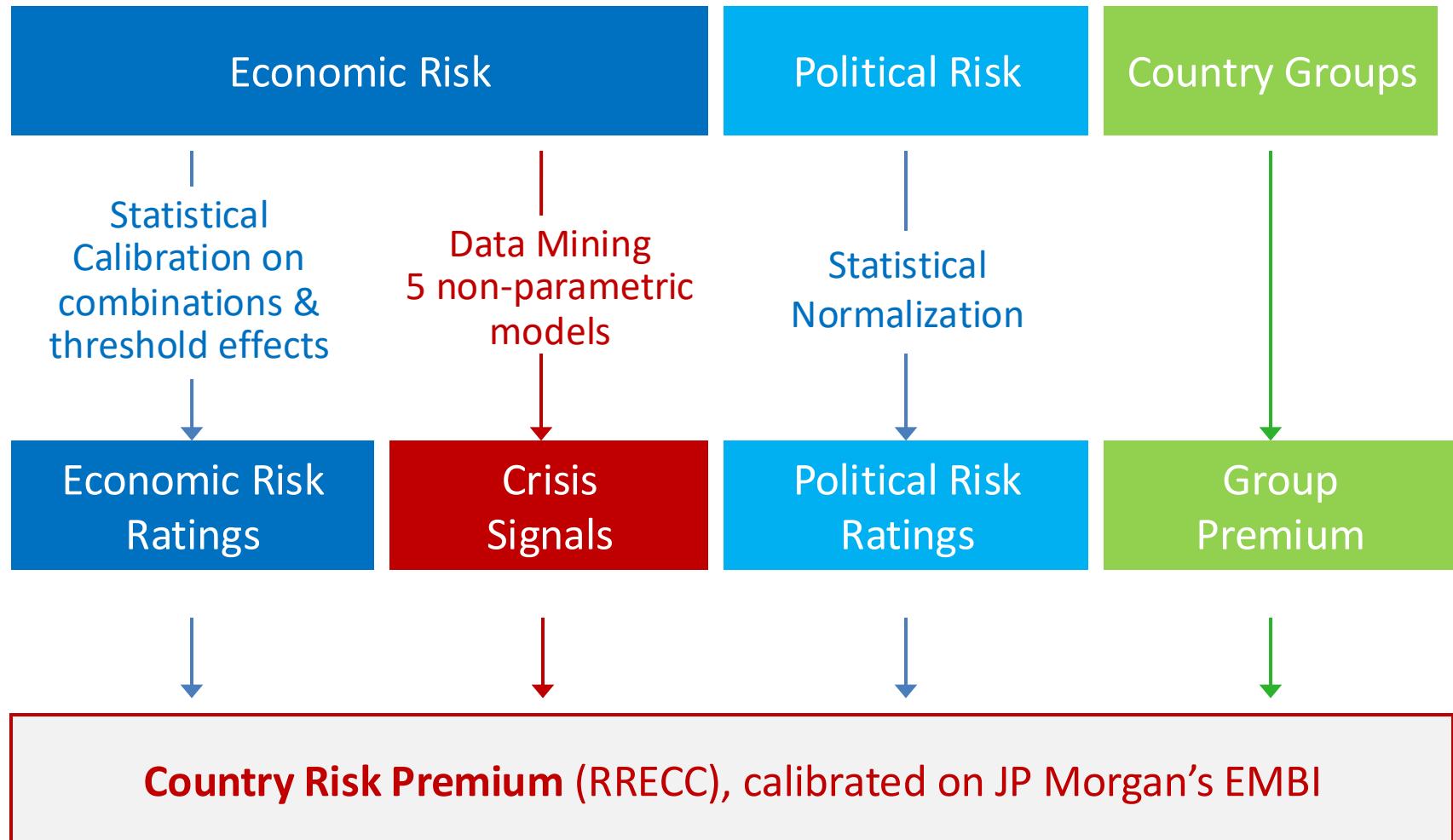
Influence des indicateurs différente selon le groupe



Performances des outils de EWS/IA développés

Avec moins d'une vingtaine d'indicateurs macroéconomiques « bien choisis » et un séquence de modèles d'IA relativement sophistiquée, nous arrivons ainsi prévoir près de 90% des crises économiques et financières, jusqu'à 2 ans à l'avance

Méthodologie RiskMonitor



ClientGate

TAC ClientGate

Publications Countries Data & Charts Search..

Brazil

Brazil

RISK PREMIUM basis points
563bp
Region: 477bp, World: 449bp

ECONOMIC RISK from 0 (lowest) to 100 (highest)
47
Region: 45, World: 45
Average

POLITICAL RISK from 0 (lowest) to 100 (highest)
51
Region: 50, World: 54
High

EARLY WARNING SIGNALS vulnerability to a significant shock
Currency
Watch List

RECENT DATA

GDP GROWTH Percentage Y/Y
1.1%
As of 2018
1.1% in 2017

INFLATION RATE Percentage Y/Y
4.7%
As of May 2019
2.9% in May 2018

EXCHANGE RATE LC per USD
3.85
As of Jul 02, 2019
3.91 in Jul 2018

FOREX RESERVES USD
372bn
As of May 2019
365bn in May 2018

POLICY Percentage
6.5%
As of Jun 2019
6.5% in

IN THE NEWS

Thousands Of Anti-corruption Protesters Rally In Brazil In Support Of Justice Minister Sergio Moro Wake Of 'car Wash' Probe
Jul 1st, 2019 - SCMP

Innovative Time-share Model In São Paulo Allows Apartment To Have Several Owners
Jun 29th, 2019 - Rio Times

LATEST PUBLICATIONS

Country Snapshots BRAZIL A Monthly

Exchange Rate
y/y growth of LC per USD (positive + depreciation). Last Data Available

Country	Y/Y Growth (%)
Turkey	20.8
Korea	5.1
China	3.7
South Africa	3.1
Russia	0.4
India	0.1
Poland	-0.3
Indonesia	-1.4
Others	1.1

Exchange Rate Index
average for 10 emerging markets, weighted by GDP (LC per USD)

Brazil:

LATEST PUBLICATIONS

The Era Of Independent Central Banks Coming To An End: Independence, No More Than A Moment In A Sometimes Multi-Secular History, Somewhat Middle-Headed Dispute, Ideological Dispute
Jun 20th, 2019 | PDF

The Banker's Comment
Turkey: Adjustment Spoiled By Politics Leading To Further Troubles For The Corporate Sector And Currency Risks, Introducing RiskMonitor's Monthly Dashboard & Heatmap, A Visual Check On All Key Changes On Our Quantitative Outputs
Jun 20th, 2019 | PDF

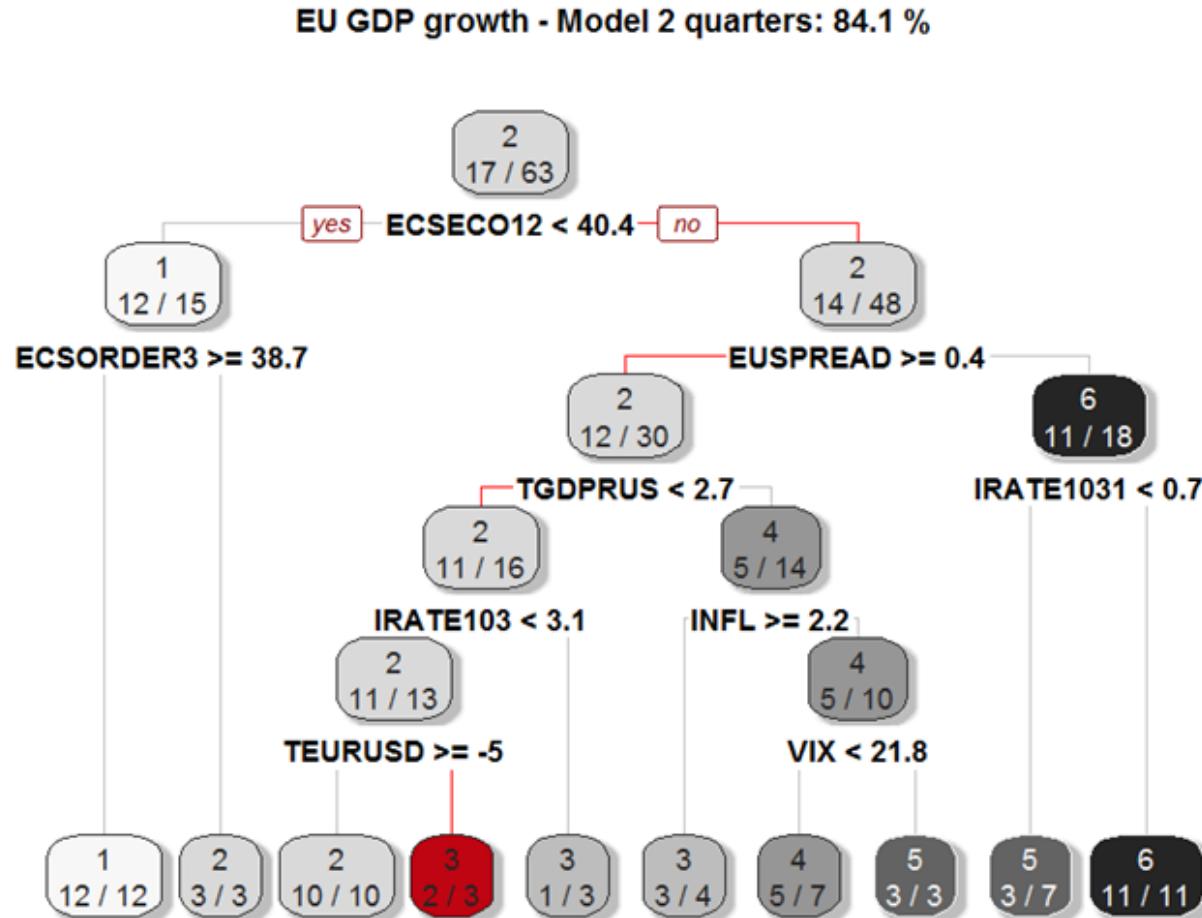
Flash Comment ARGENTINA
Argentina: Economic Pressures On Tight Political Agenda
May 29th, 2019 | PDF

Flash Comment INDIA
India: PM Modi Reelected Amidst Economic Slowdown
Jun 5th, 2019 | PDF

Flash Comment NAMIBIA
Namibia: A High Power Of Attraction For A Small And Almost Peaceful Nation | Half-Half Of Hispaniola Suffering | India-Pakistan-Afghanistan: Haunting Borders
May 16th, 2019 | PDF

36

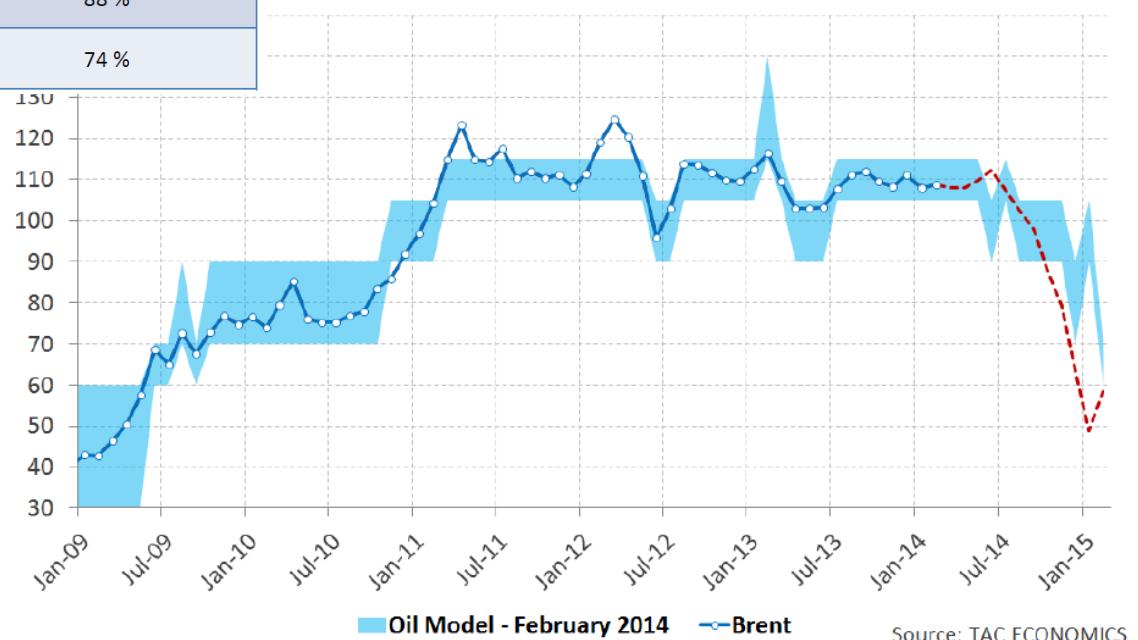
CART Applied to Real GDP Growth Forecast



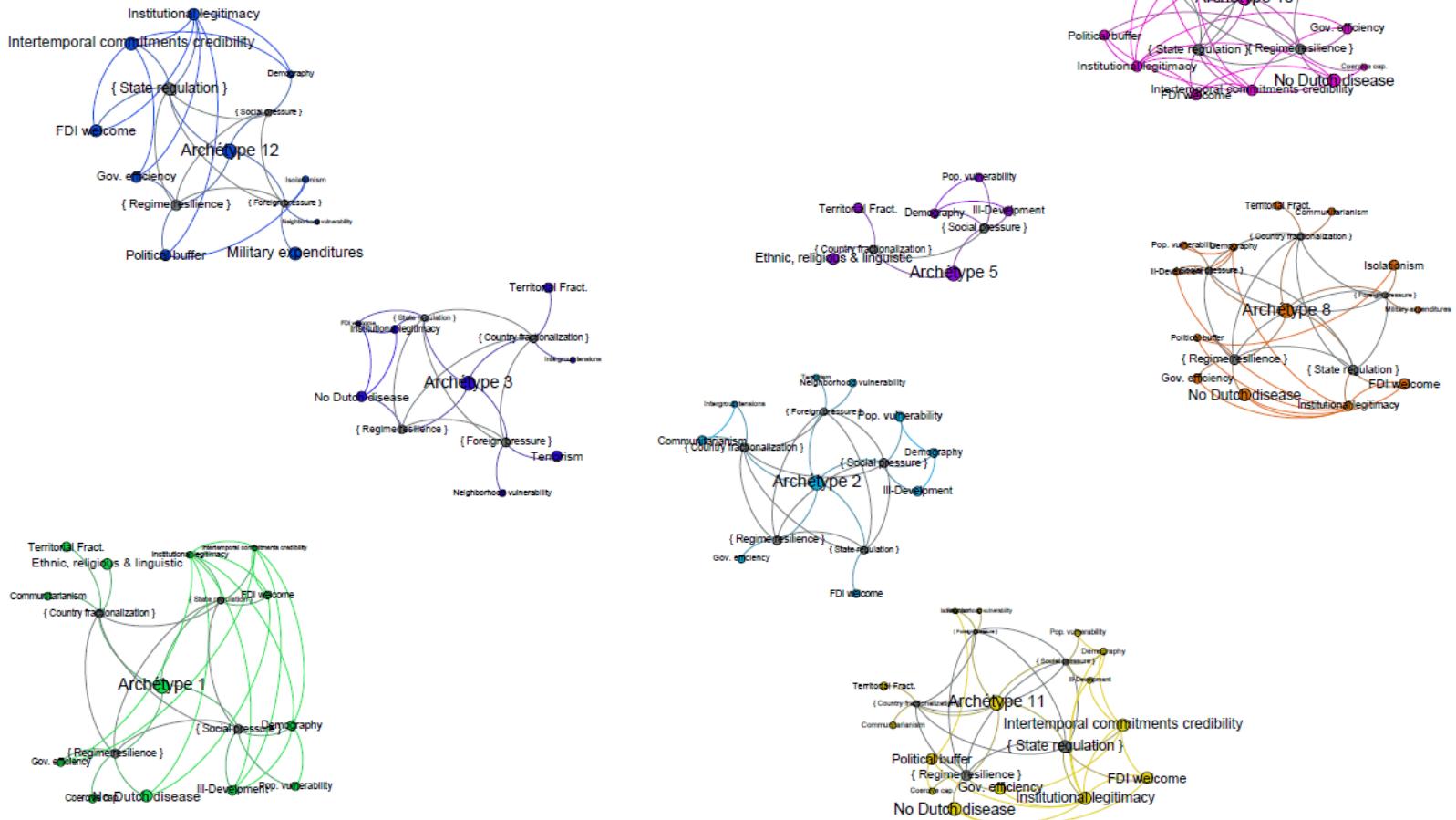
Machine Learning on Oil Price Forecasts

	Average accuracy (training dataset)	Average accuracy (testing, dataset 1)
Naive Bayes	96 %	88 %
Tree Bagging	99.9 %	90 %
Gradient Boosted Machine	100 %	90 %
Supervised SOM	86 %	75 %
Neural Network multilayer perceptron	82 %	74 %
Random Forest	100 %	90 %
Support Vector Machine	96 %	88 %
k-nearest neighbors	84 %	74 %

TAC Brent short-term Projections (\$/bl)



Archetypal of Political Risk



Ruptures sur les Marchés Financiers

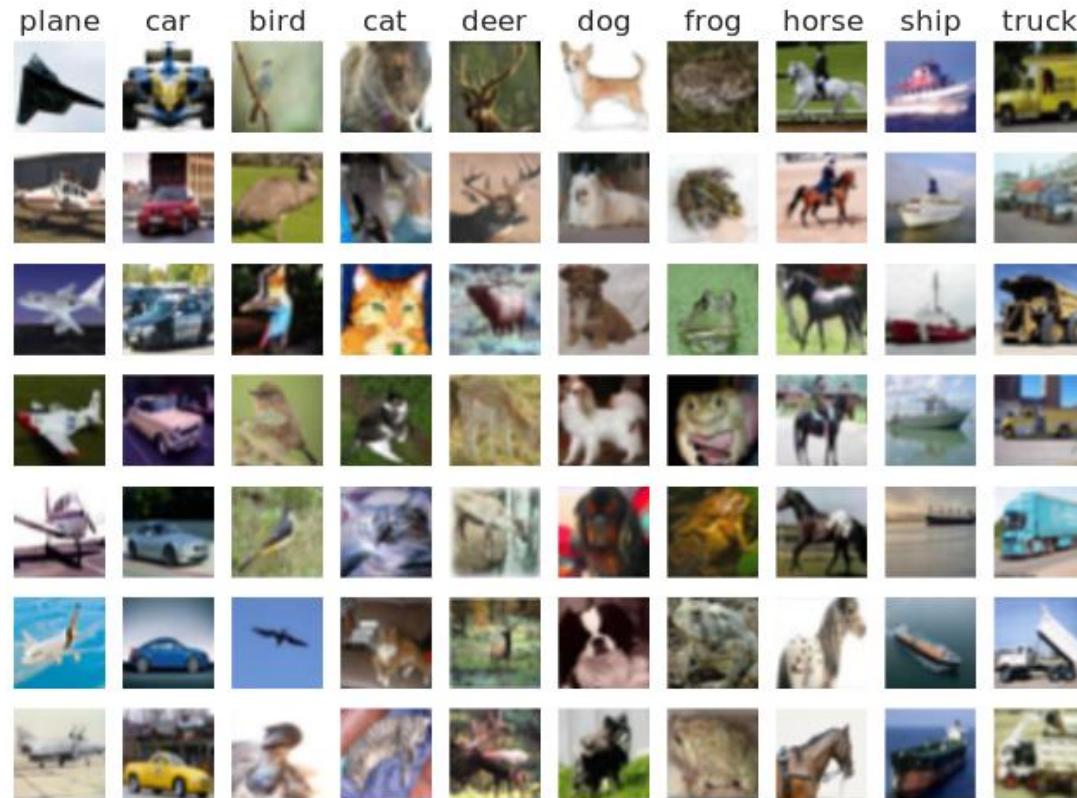
Petit retour sur les travaux historique

- Réseaux de neurones en finance d'abord comme substitut aux systèmes experts, dans les années 90.
- Prévision des cours boursiers dès Kimoto et Yoda (1993), sur le Tokyo Stock Index avec... 5 variables d'entrée ! Mais aussi sur l'or ou le S&P500 (Grudnitski et Obsburn, 1993, Quang Do 1995).
- Mais aussi des prévisions de taux de change, dès le milieu des années 90 (Rawani 1993, Azoff 1994, Avouyi Dovi 1995).
- Depuis, de très nombreux travaux sur l'utilisation de méthodes de machine learning en finance.

Les plus gros « hedge funds » du monde

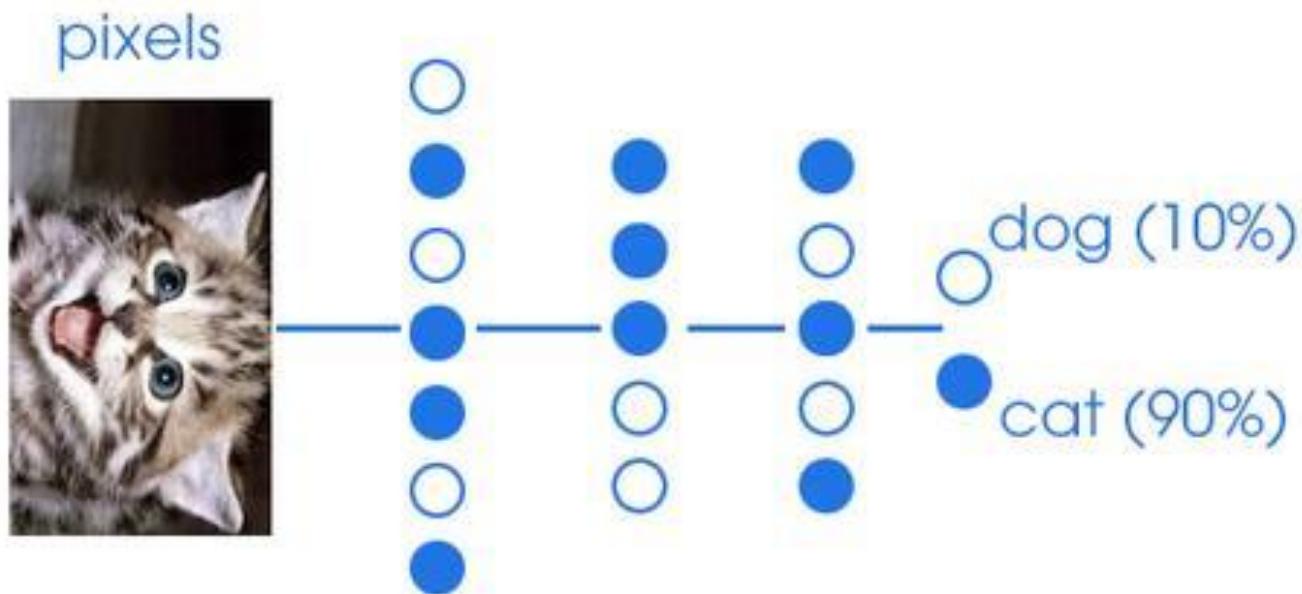
	Actifs sous Gestion en mds de \$ en 2017	Quantitatif ?
Bridgewater Associates	122.2	(Non)
AQR Capital Management	69.6	Oui
JPMorgan AM	45.0	(Oui)
Renaissance Technologies	42.0	Oui
Two Sigma	38.9	Oui
De Shaw & Co	34.7	Oui
Man Group	33.9	Oui
Millennium Management	33.9	Oui
Och-Ziff Capital Management	33.5	(Oui)
Winton Group	32.0	(Oui)

Computer Vision

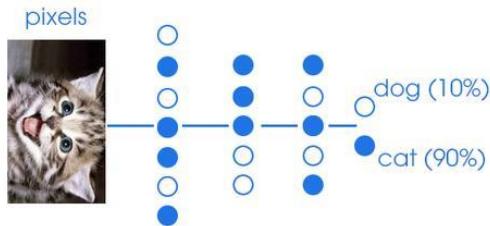


Source: dominodatalab

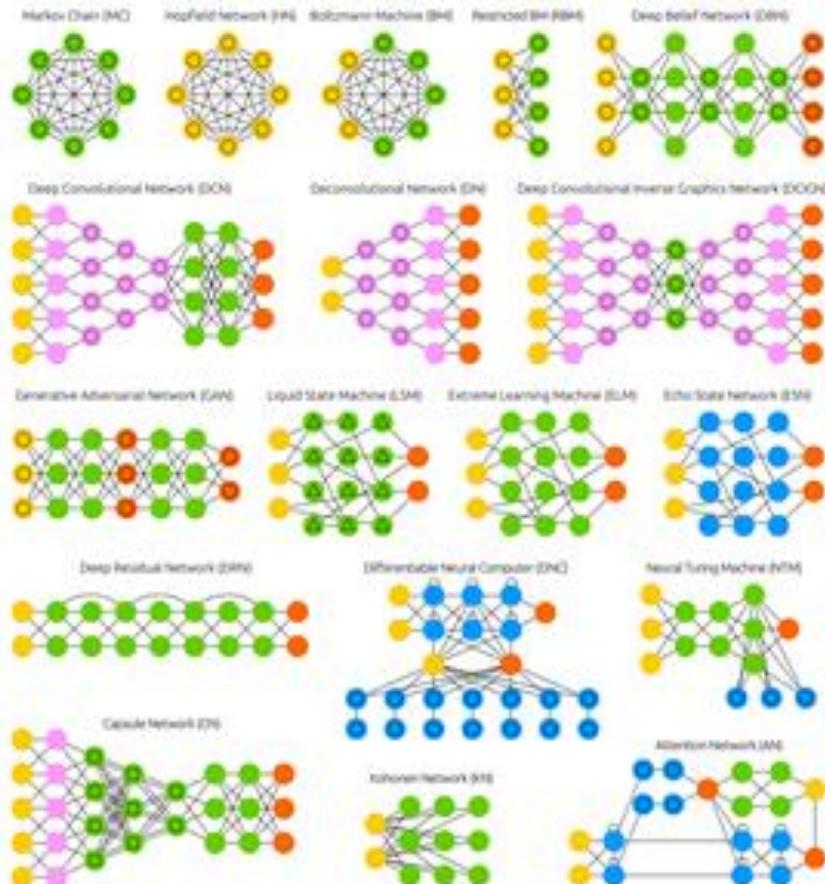
Réseaux de neurones « traditionnels »



Du perceptron au « deep learning »



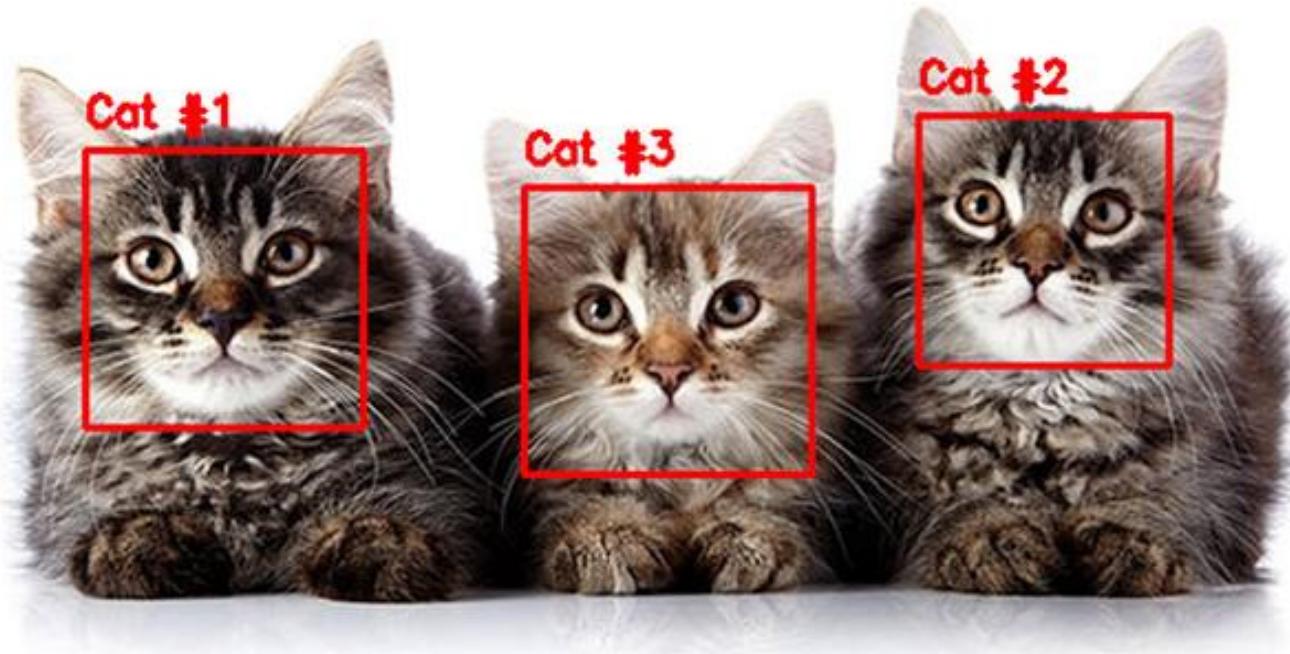
Le Neural Network Zoo du Asimov Institute



...parfois compliqué: distinguer les chiens des muffins



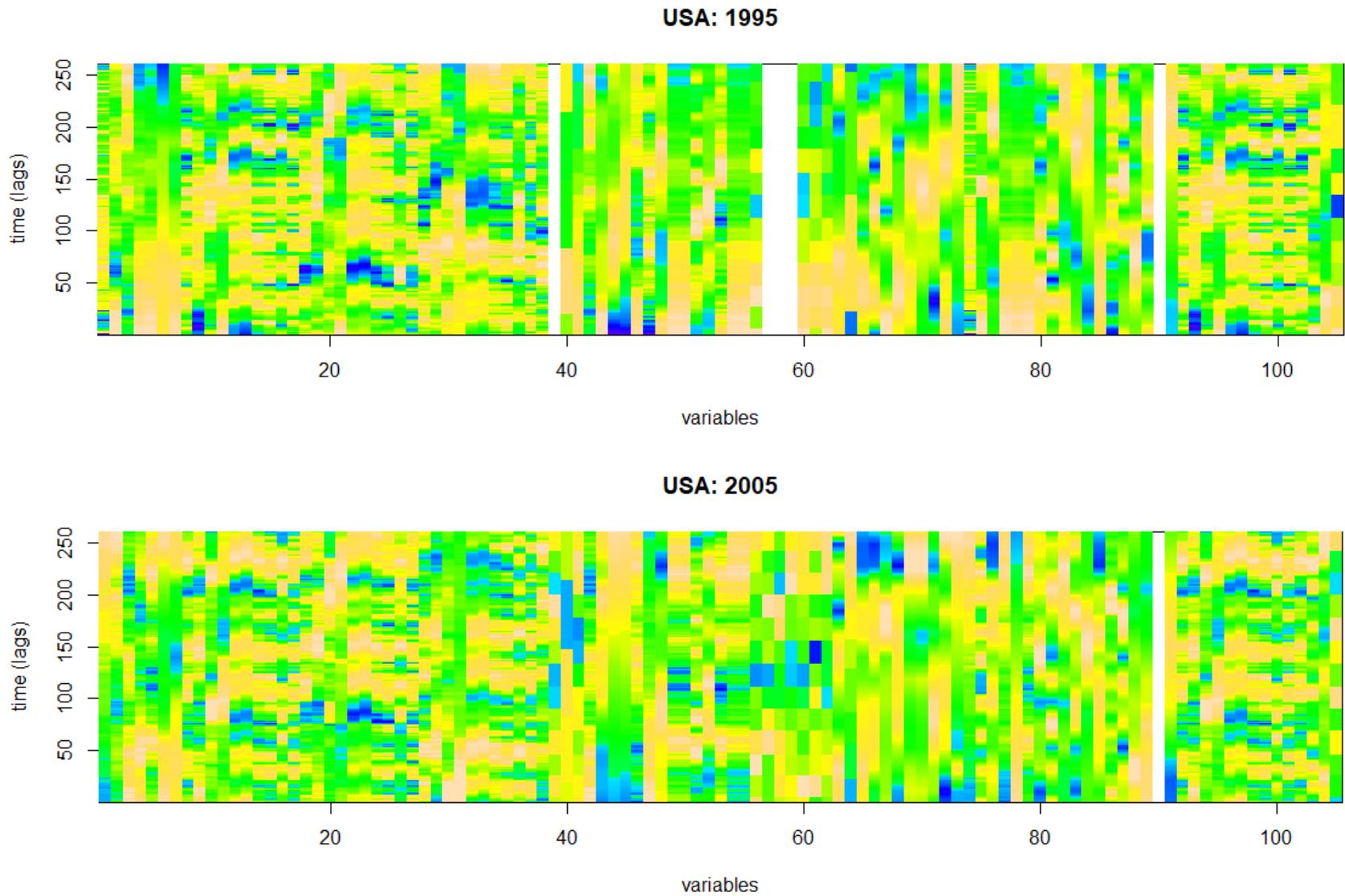
La convolution ou la « révolution du petit chat »...



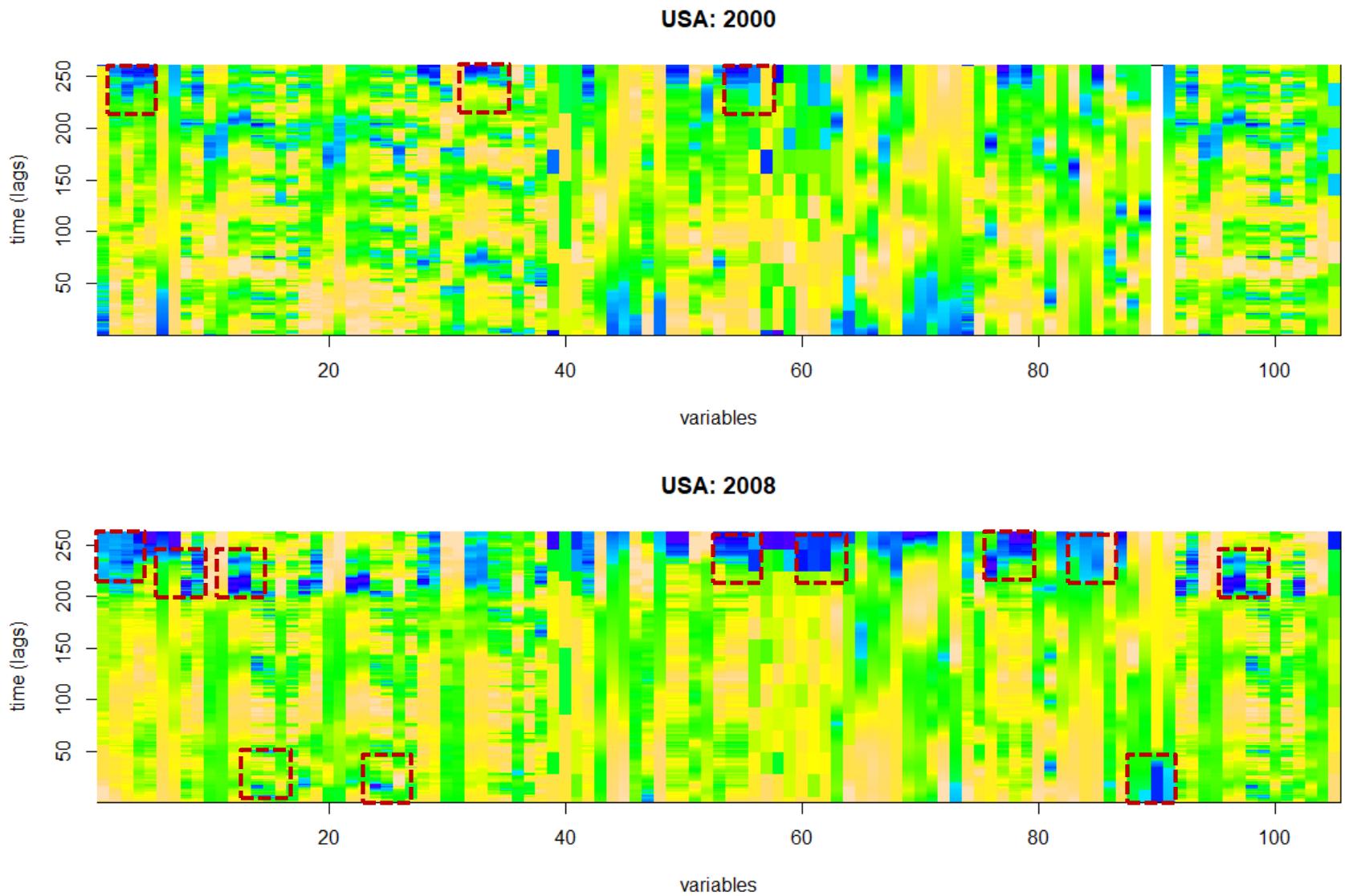
Quantitative Market Alert (QMA)

- Outil de détection des tendances et des ruptures sur plus de 20 marchés internationaux (actions, obligations, spreads corporates).
- Signaux sur les tendances, et calibrage automatique des poids des différents actifs dans un portefeuille.
- Possibilité d'ajuster les fonctions objectifs sur la base de cibles « combinées ».
- Outil mixant à la fois des outils de machine learning « simples », à des outils plus puissants de deep learning appliqués à des échantillons de plus de 200 données macroéconomiques et financières sur des périodes de 40 jours (identification de « patterns »).
- Difficulté majeure: bien gérer les échantillons, et l'instabilité des performances historiques.

« Patterns » du QMA et convolution



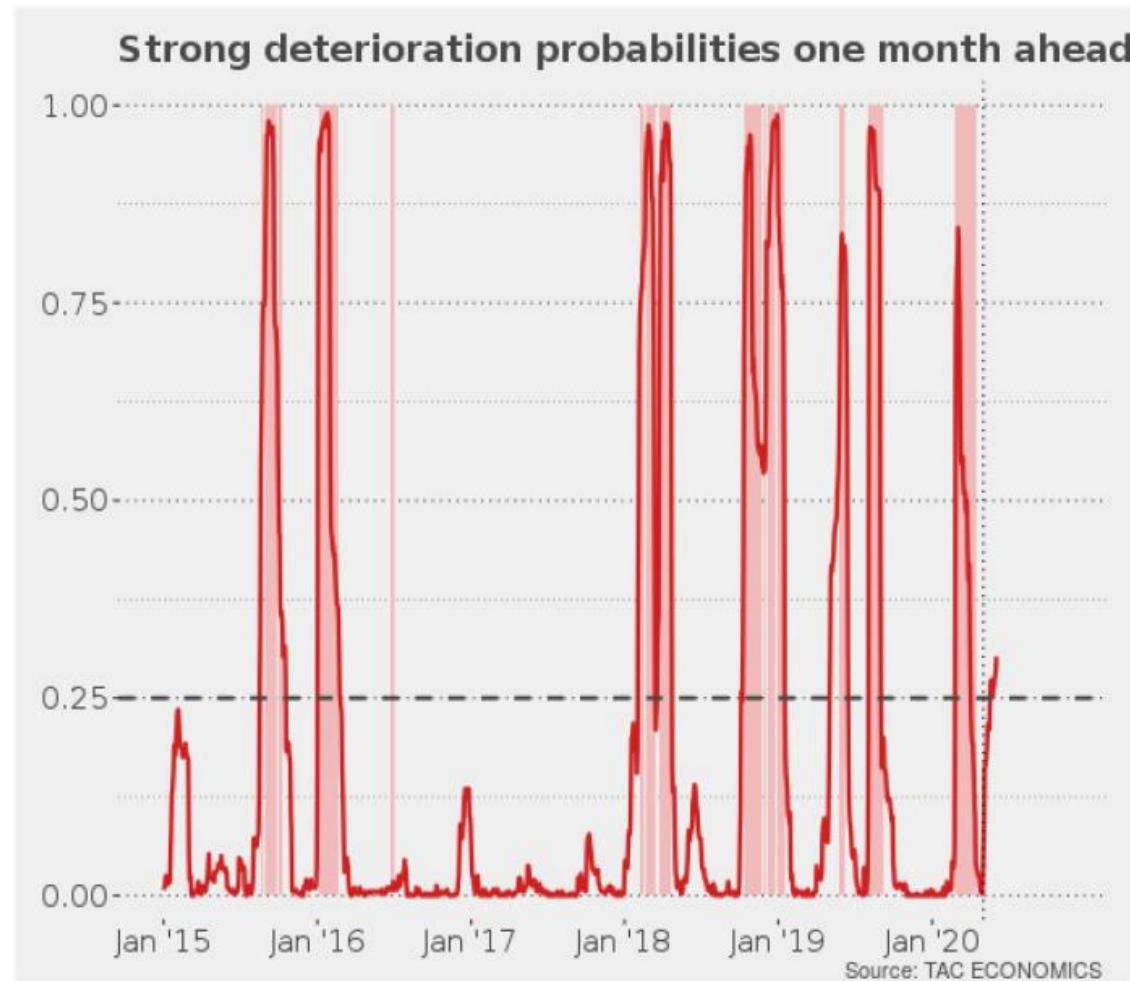
« Patterns » du QMA et convolution



Performances récente du QMA : les Fair Values en Jan. 2020

	Market Level (Jan. 03)	Fair Value November 19	Gap to Fair Value	Expected short-term direction
Equity Indices				
S&P 500	3 235	2 849	386	Decrease
CAC 40	6 044	5 344	700	Decrease
DAX 30	13 219	12 103	1116	Decrease
FTSE 100	7 622	7 268	354	Decrease
Nikkei 225	23 657	21 578	2079	Decrease

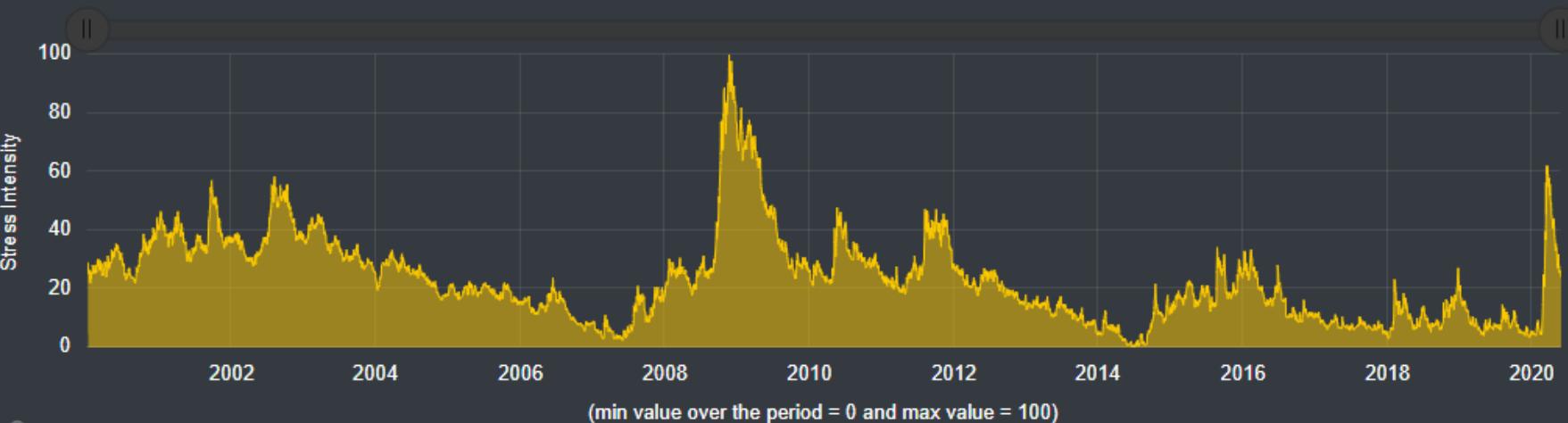
Performances et signaux du QMA sur le S&P500



Financial Index « Covid-19 »

Financial Index

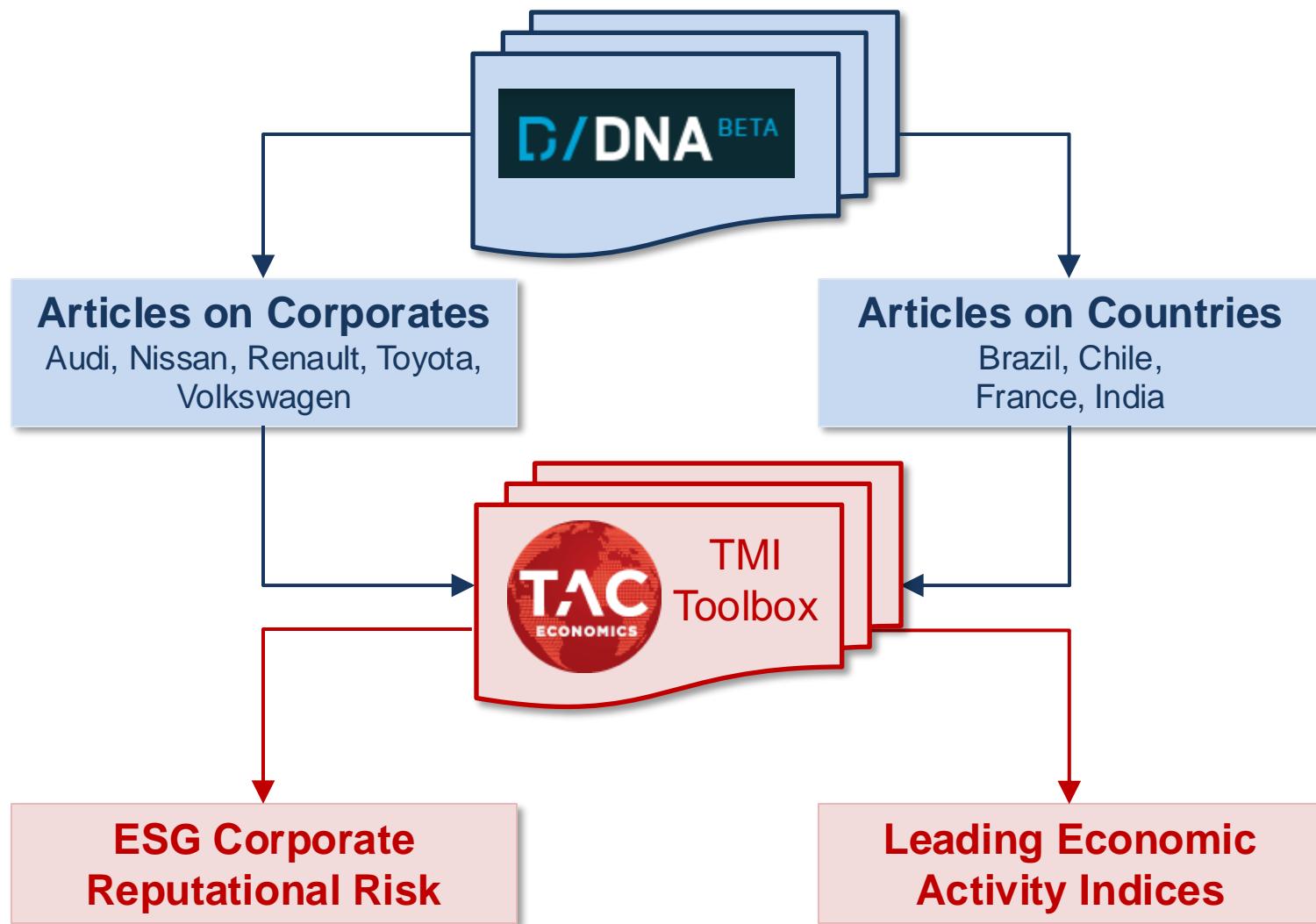
The Financial Index is calculated using long-run measures on historical drawdowns on financial markets, corporate and sovereign spreads, the VIX index and a volatility of the EUR/USD parity.



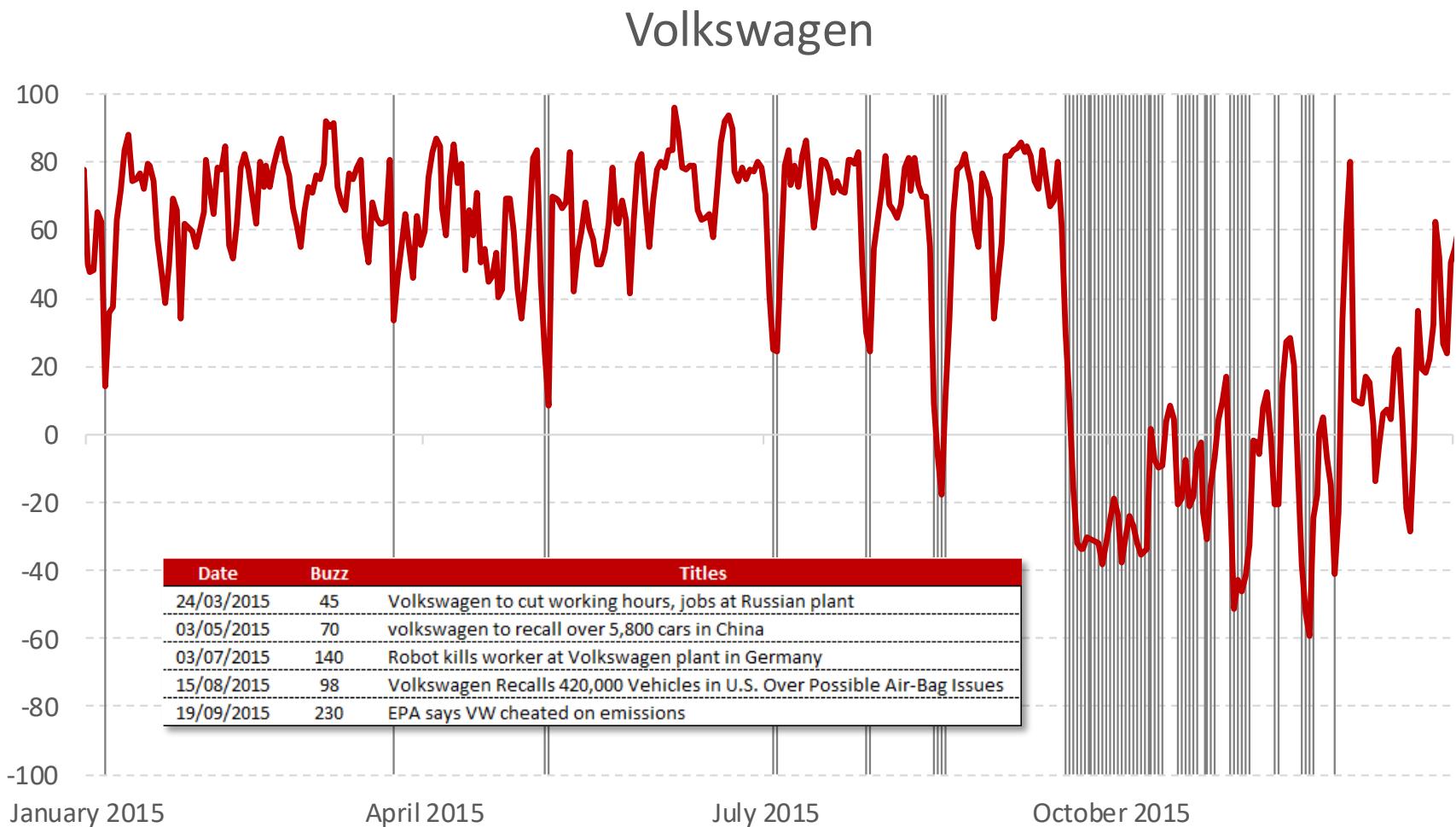
<https://www.taceconomics.com/covid19/>

Natural Language Processing

Text mining, topics, sentiments et risque réputationnel

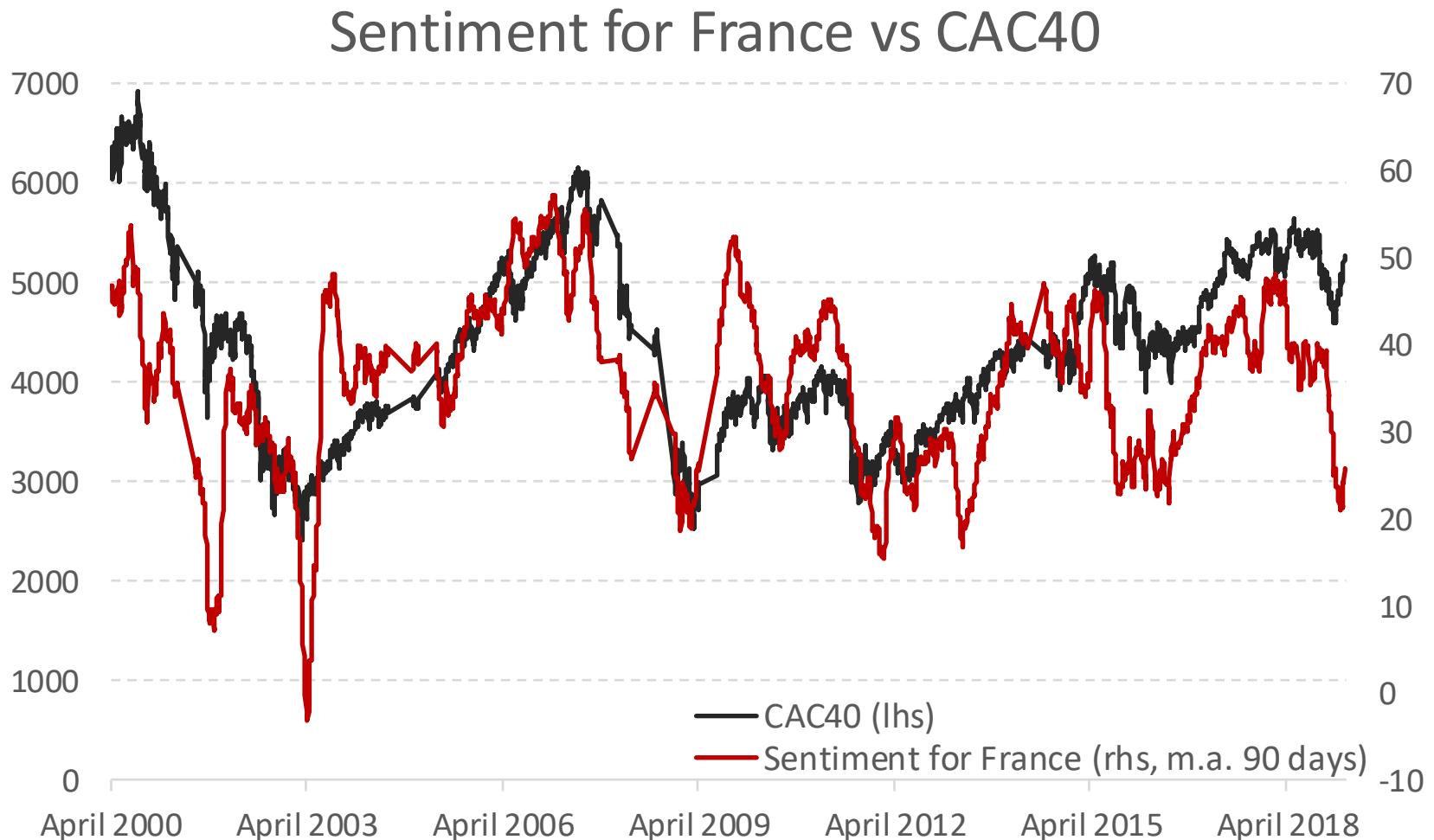


Risque réputationnel: Volkswagen & the Diesel Gate

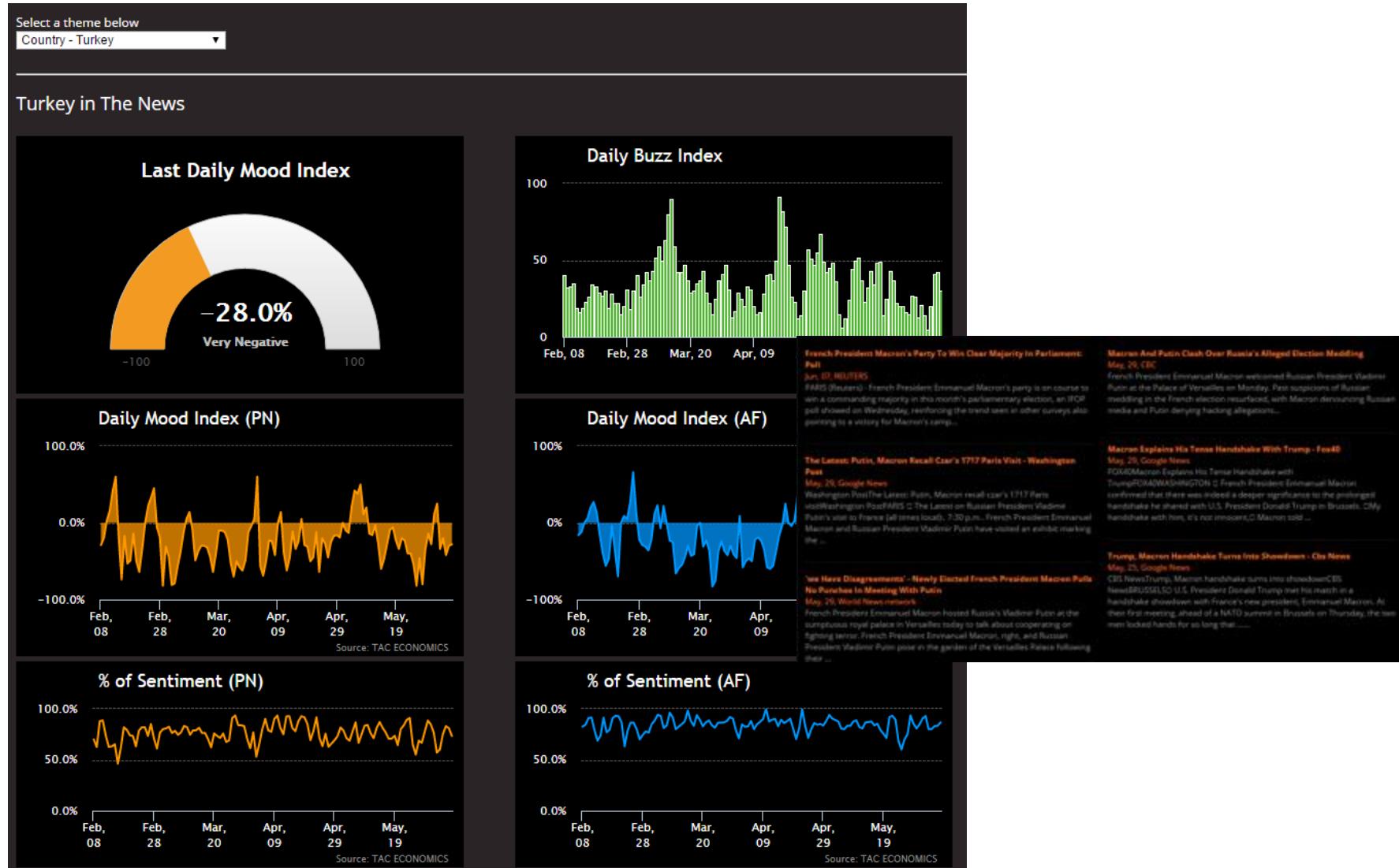


Source: TAC ECONOMICS

Indicateurs de sentiment et indices boursiers



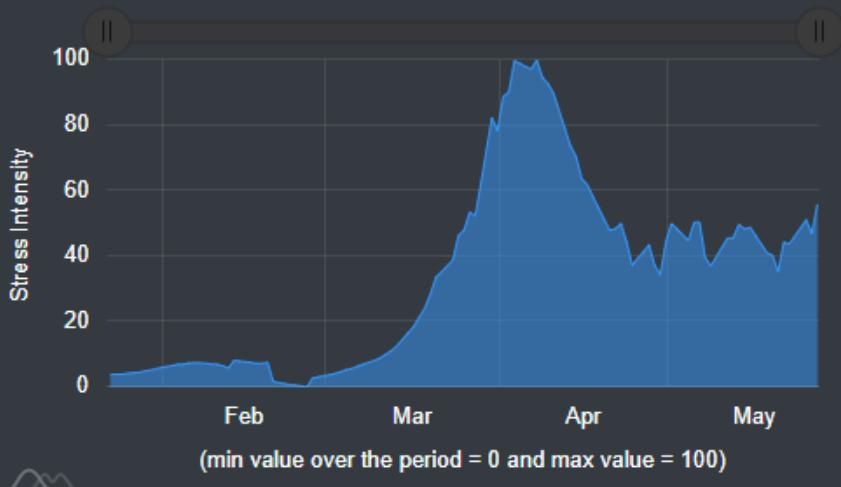
Sentiment Analysis, Topics & Finance



Sentiment Index « Covid-19 »

Epidemiologic Index

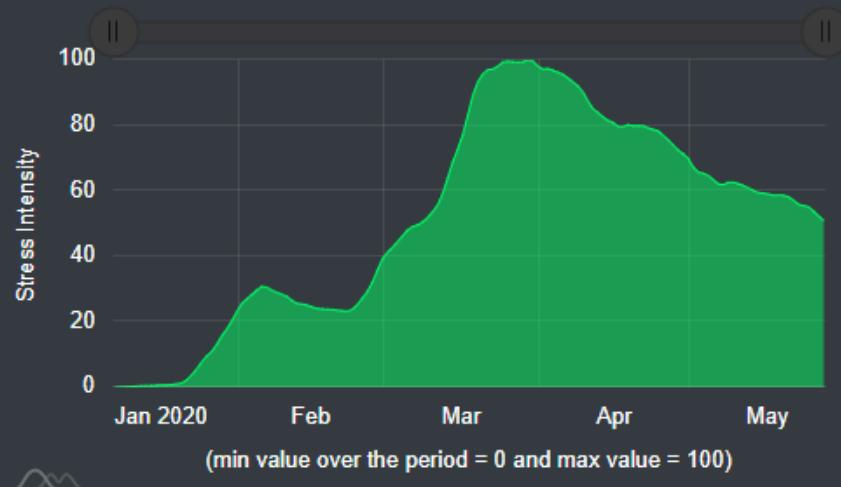
The Epidemiologic Index is calculated using indicators based on speed and acceleration of cases and deaths due to the Covid-19.



Last update: May 27, 2020

Sentiment Index

The Sentiment Index is calculated using text-mining techniques applied to the analysis of thousands press articles published on a daily basis.



Last update: May 27, 2020

<https://www.taceconomics.com/covid19/>

Choix des sujets globaux

Déroulement des cours



Introduction d'un ou plusieurs concepts



Mise en application pratique



Amélioration / construction d'un projet global

Mode de fonctionnement

- Groupes de 2 personnes max
- Préparation du travail pendant les cours, avec l'aide de l'enseignant
- Objectif de réaliser un outil qui sera 100% fonctionnel en fin de séance

Sujets

1. Indicateurs sur le risque ESG avec données Banque Mondiale
2. Analyse du sentiment sur articles du New York Times
3. Création d'indicateurs de risque thématiques à partir de GDELT
4. Indicateurs de risque journaliers sur les actions du S&P500
5. Analyse des données INSEE des communes et visualisation des disparités et spécificités géographiques sur la France
6. Analyse du marché des cryptos
7. Optimisation de stratégie de référencement (SEO)
8. Analyse des données de data.gouv.fr
9. Lecture de la presse et analyse des entités nommées
10. Autres sujets au choix, à la discréction des élèves

Déroulement du travail

1. Cadrage (1h)

1. Revue de la littérature et recherches sur Internet (rapide)
2. Cadrage du sujet et objectifs
3. Partage des taches

2. Travaux Quantitatifs

1. Création d'une base de stockage sous Docker (MySQL, MongoDB, InfluxDB, OpenSearch)
2. Récupération, traitement et stockage des données (Postman / ETL automatisé, script)
3. Analyse quantitative (modèles et dataviz les plus adaptés)

3. Présentation

1. Quelques slides pour décrire les objectifs et les étapes du travail réalisé
2. Quels sont les résultats clés ? Quelles prochaines étapes suggérez-vous ?
3. Présentation des résultats devant la classe (10min)

Planning des cours

Jour	Activité	Salles	Enseignants	Etudiants	Durée (h)	Début	Fin
Lundi 02/09/2024	Data ingeneering	B02 - Salle info 10 (101)	BARTHELEMY Sylvain	M2 SD SP fa, M2 SD SP fi, M2 SD DP fa, M2 SD SE fa, M2 SD SE fi	03h00	09h00	12h00
Mardi 03/09/2024	Data ingeneering	B02 - Salle info 10 (101)	BARTHELEMY Sylvain	M2 SD SP fa, M2 SD SP fi, M2 SD DP fa, M2 SD SE fa, M2 SD SE fi	03h00	13h30	16h30
Mercredi 04/09/2024	Data ingeneering	B02 - Salle info 10 (101)	BARTHELEMY Sylvain	M2 SD SP fa, M2 SD SP fi, M2 SD DP fa, M2 SD SE fa, M2 SD SE fi	03h00	09h00	12h00
Lundi 09/09/2024	Data ingeneering	B02 - Salle info 10 (101)	BARTHELEMY Sylvain	M2 SD SP fa, M2 SD SP fi, M2 SD DP fa, M2 SD SE fa, M2 SD SE fi	03h00	09h00	12h00
Mardi 10/09/2024	Data ingeneering	B02 - Salle info 10 (101)	BARTHELEMY Sylvain	M2 SD SP fa, M2 SD SP fi, M2 SD DP fa, M2 SD SE fa, M2 SD SE fi	03h00	09h00	12h00

Démarrage

Data + ETL

Indicateurs

Analyses + Viz

Présentations

Virtualisation, Cloud et Stockage des Données

Stockage des données

- **Les bases de données traditionnelles SGDB**
MySQL, PostgreSQL, Oracle
notion ACID (Atomicité, Cohérence, Isolation, Durabilité)
- **Les bases de données objets ou orientées colonnes (NoSQL)**
MongoDb, Cassandra
- **Les bases en RAM**
Redis
- **Les bases textuelles**
ElasticSearch, OpenSearch
- **Les bases de séries temporelles**
InfluxDB, M3
- **Le stockage objet**
AWS S3 et stockage clés/valeurs
- **Les clusters big data et les « events »**
Hadoop, Hive et Kafka

Nouvelles bases de données et nouveaux outils



Clouds et Systèmes



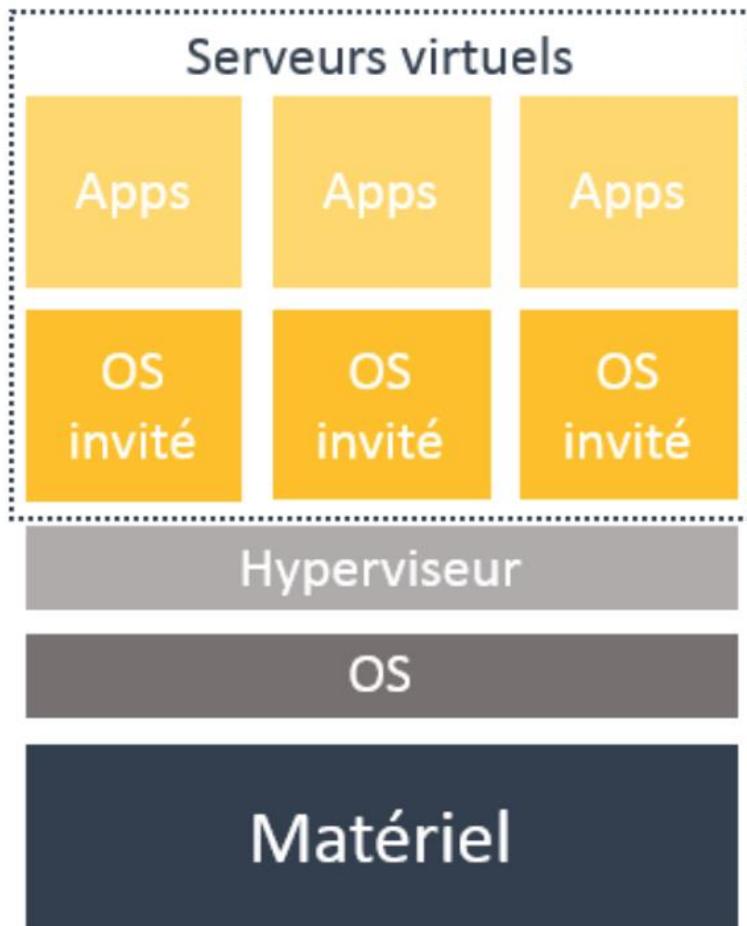
Google Cloud

Clouds et Systèmes

- Que proposent AWS, Azure, Google et OVH
- Concept de bare metal, virtualisation, conteneurisation et SaaS
- VMs et connexion et fonctionnement d'un serveur sous linux.
- Choisir le bon type de VM/machine/cout
- La notion de cluster
- Terraform et le IAC

Installation de Docker

Docker ?



source: <https://www2.itroom.fr/docker-en-un-coup-d-oeil/>

Installation de Docker sous Windows

The screenshot shows a screenshot of the Docker documentation website. The top navigation bar includes links for Home, Guides, Manuals, Reference, Samples, and Contribute. Below the navigation is a breadcrumb trail: Home / Manuals / Docker Desktop / Install Docker Desktop / Install on Windows. A sidebar on the left contains a list of links: Understand permission requirements for Mac, **Install on Windows** (which is selected), Understand permission requirements for Windows, Install on Apple silicon, Install on Linux, Installation per Linux distro, Quick Start Guide and sign in, Explore Docker Desktop, Change settings, Troubleshoot and diagnose, Additional resources, Dev Environments (Beta), Extensions (Beta), Extensions SDK (Beta), Containerd Image Store (Beta), FAQs, Give feedback, Release notes, Previous versions, and Docker Engine. The main content area has a title "Install Docker Desktop on Windows" with an estimated reading time of 9 minutes. It features a "Docker Desktop terms" section with a note about commercial use requiring a paid subscription. Below this is a "Welcome" section with a "Download Docker Desktop for Windows" button. A note for checksums links to release notes. The "System requirements" section states that Windows must meet certain requirements. It lists two backend options: "WSL 2 backend" (selected) and "Hyper-V backend and Windows containers". The "WSL 2 backend" section includes a bulleted list of system requirements for Windows 11 and Windows 10.

Understand permission requirements for Mac

Install on Windows

Understand permission requirements for Windows

Install on Apple silicon

Install on Linux

Installation per Linux distro

Quick Start Guide and sign in

Explore Docker Desktop

Change settings

Troubleshoot and diagnose

Additional resources

Dev Environments (Beta)

Extensions (Beta)

Extensions SDK (Beta)

Containerd Image Store (Beta)

FAQs

Give feedback

Release notes

Previous versions

Docker Engine

Install Docker Desktop on Windows

Estimated reading time: 9 minutes

Docker Desktop terms

Commercial use of Docker Desktop in larger enterprises (more than 250 employees OR more than \$10 million USD in annual revenue) requires a paid subscription.

Welcome to Docker Desktop for Windows. This page contains information about Docker Desktop for Windows system requirements, download URL, instructions to install and update Docker Desktop for Windows.

Download Docker Desktop for Windows

Docker Desktop for Windows

For checksums, see [Release notes](#)

System requirements

Your Windows machine must meet the following requirements to successfully install Docker Desktop.

WSL 2 backend Hyper-V backend and Windows containers

WSL 2 backend

- Windows 11 64-bit: Home or Pro version 21H2 or higher, or Enterprise or Education version 21H2 or higher.
- Windows 10 64-bit: Home or Pro 21H1 (build 19043) or higher, or Enterprise or Education 20H2 (build 19042) or higher.
- Enable the WSL 2 feature on Windows. For detailed instructions, refer to the [Microsoft documentation](#).
- The following hardware prerequisites are required to successfully run WSL 2 on Windows 10 or Windows 11:

Mon premier Docker: serveur Web NGINX

The screenshot shows the Docker Hub interface for the official NGINX Docker image. At the top, there's a navigation bar with links for Explore, Pricing, Sign In, and Register. Below the bar, a search bar says "Search for great content (e.g., mysql)". The main content area shows the NGINX logo and the text "nginx" followed by "DOCKER OFFICIAL IMAGE", "1B+", and "10K+". It also says "Official build of Nginx.". On the right, there's a button labeled "docker pull nginx" with a copy icon. The page has tabs for "Overview" (which is selected) and "Tags". The "Quick reference" section contains a bulleted list: "Maintained by: the NGINX Docker Maintainers" and "Where to get help: the Docker Community Forums, the Docker Community Slack, or Stack Overflow". The "Supported tags and respective Dockerfile links" section lists tags: "1.23.1", "mainline", "1", "1.23", "latest", and "1.23.1", "mainline", "1", "1.23", "1", "1", "1". To the right, there's a "Recent Tags" sidebar with tags like "latest", "stable-perl", "stable", "perl", "mainline-perl", "mainline", "1.23.1-perl", "1.23.1", "1.23-perl", and "1.23". Below that is an "About Official Images" section stating that Docker Official Images are curated open source repositories, and a "Why Official Images?" section explaining their benefits.

Explore Pricing Sign In Register

Search for great content (e.g., mysql)

nginx DOCKER OFFICIAL IMAGE 1B+ 10K+

Official build of Nginx.

docker pull nginx

Overview Tags

Quick reference

- Maintained by:
the NGINX Docker Maintainers
- Where to get help:
the Docker Community Forums, the Docker Community Slack, or Stack Overflow

Supported tags and respective Dockerfile links

- 1.23.1, mainline, 1, 1.23, latest
- 1.23.1, mainline, 1, 1.23, 1, 1, 1, 1

Recent Tags

latest stable-perl stable perl mainline-perl
mainline 1.23.1-perl 1.23.1 1.23-perl 1.23

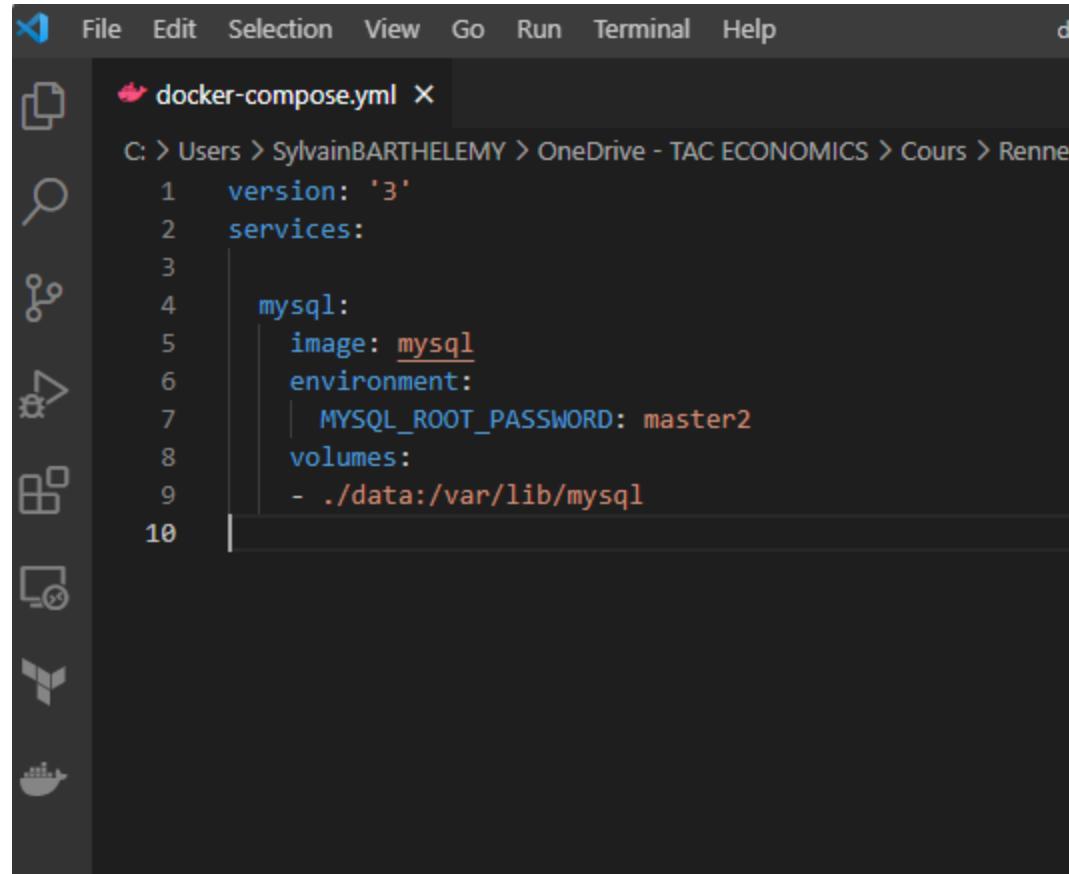
About Official Images

Docker Official Images are a curated set of Docker open source and drop-in solution repositories.

Why Official Images?

These images have clear documentation, promote best practices, and are designed for the most common use

Base de données MySQL

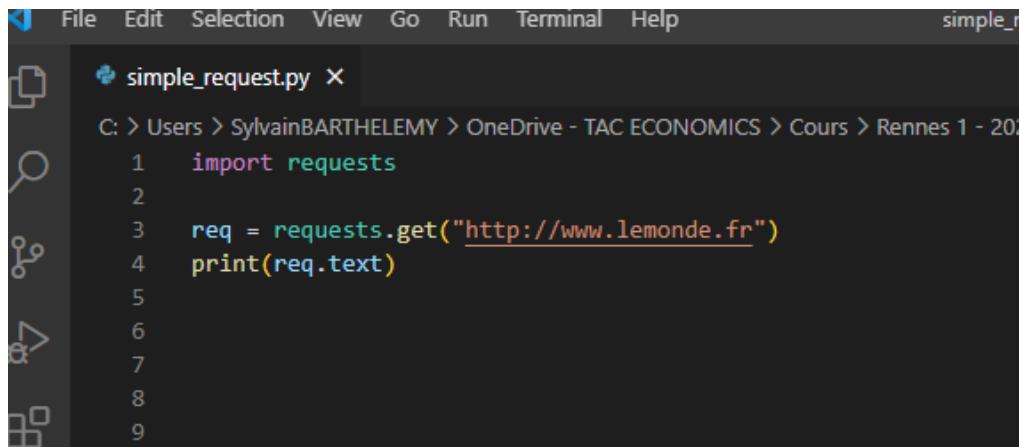


A screenshot of a code editor window showing a `docker-compose.yml` configuration file. The file defines a service named `mysql` using the `mysql` image, setting the root password to `master2`, and mounting the local directory `./data` to `/var/lib/mysql`.

```
version: '3'
services:
  mysql:
    image: mysql
    environment:
      MYSQL_ROOT_PASSWORD: master2
    volumes:
      - ./data:/var/lib/mysql
```

Construction des premiers ETL

Introduction au Web Scraping RSS en Python

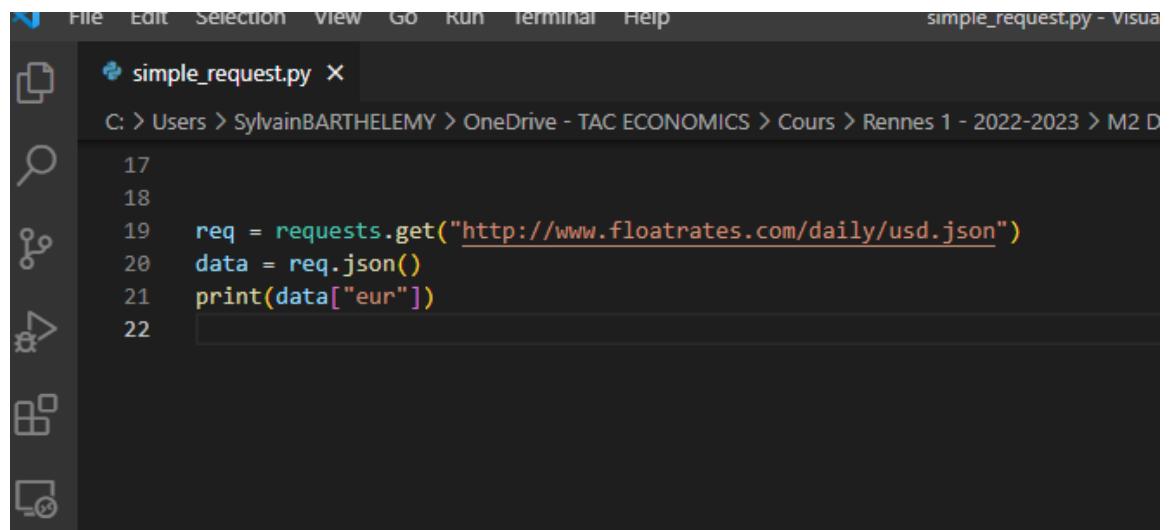


```
File Edit Selection View Go Run Terminal Help simple_request.py

simple_request.py X

C: > Users > SylvainBARTHELEMY > OneDrive - TAC ECONOMICS > Cours > Rennes 1 - 2022-2023 > M2 D

1 import requests
2
3 req = requests.get("http://www.lemonde.fr")
4 print(req.text)
5
6
7
8
9
```



```
File Edit Selection View Go Run Terminal Help simple_request.py - Visual Studio Code

simple_request.py X

C: > Users > SylvainBARTHELEMY > OneDrive - TAC ECONOMICS > Cours > Rennes 1 - 2022-2023 > M2 D

17
18
19 req = requests.get("http://www.floatrates.com/daily/usd.json")
20 data = req.json()
21 print(data["eur"])
22
```

Introduction au Web Scraping

Lecture d'articles sur le New York Times

{CT} Developers

[Home](#) [APIs](#) [Covid-19 Data](#) [Get Started](#) [Sign In](#)

The New York Times Developer Network

All the APIs Fit to Post

[GET STARTED](#)

Get Started

Learn how to sign up for an API key.

APIs

Learn about and try out NYT's APIs.

FAQ

Get answers to frequently asked questions.

Branding

Download NYT's logo and icons

Covid-19 Data

Download NYT Covid-19 Data

Sylvain BARTHELEMY

sylvain.barthelemy@gwenlake.com

<https://gwenlake.com>