



**Bienvenue à toutes et à tous,
nous allons commencer dans un instant.**

QUEL(S) OUTILS POUR DE LA VISUALISATION DE DONNEES EFFICIENTE ?



**DATA
VISUALISATION
EN ACTUARIAT**

28 nov. au 12 déc.

LE CHALLENGE

BY Institut. 

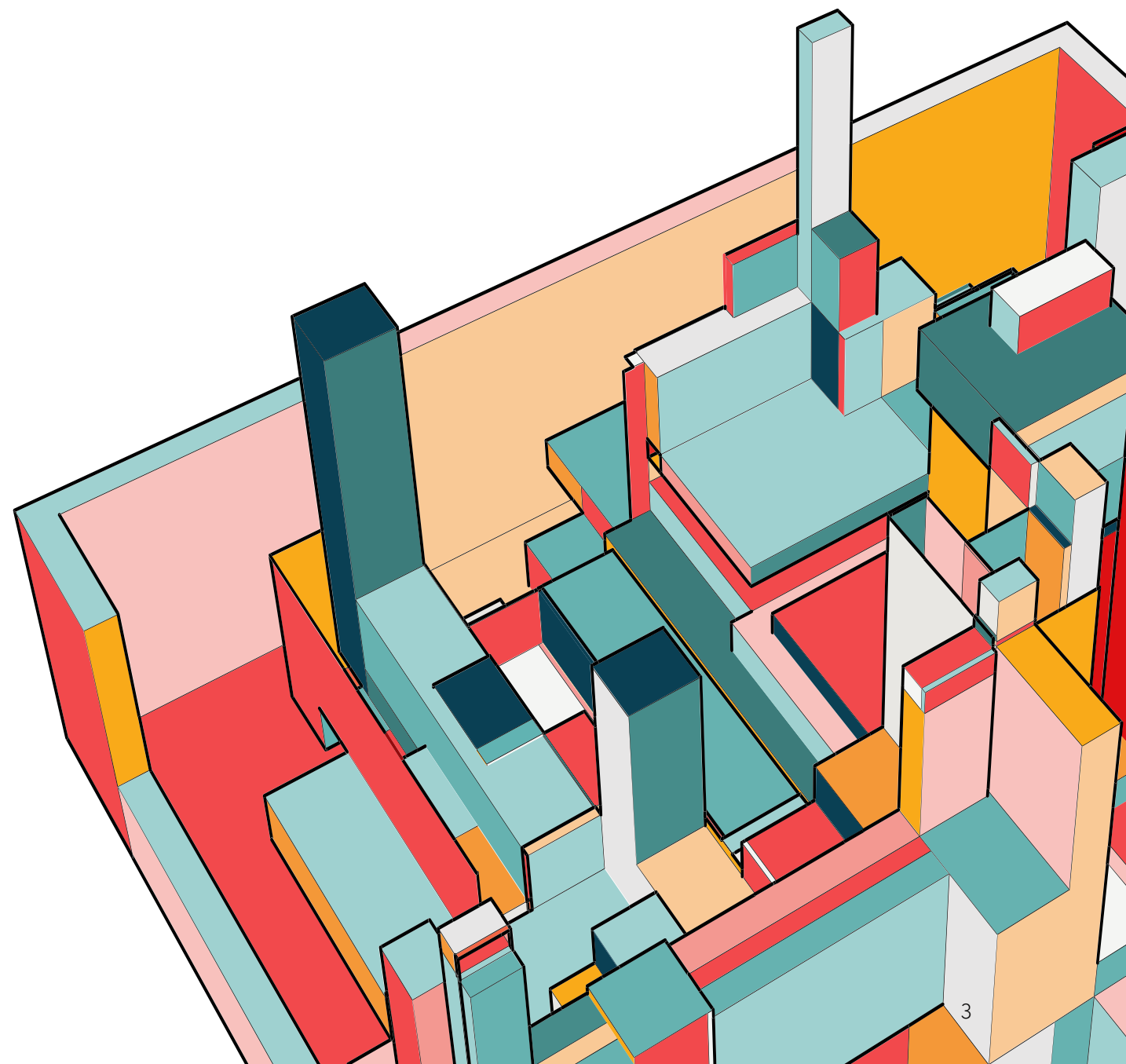
ABOUT ME

Jean-Claude RAZAFINDRAKOTO

SCOR Data Analytics Advocate



15/11/2022



PRÉSENTATION DU CLUB ALGO

1

ANIMATRICE



ANIMER

Fédérer le réseau des
anciens

Animer le réseau

des membres

Créer des **synergies**

15

MEMBRES
ACTIFS



ECHANGER

Se réunir **une fois par mois**

Présentation, échanges,

réflexion

6

ANS
D'ACTIVITES

1

FOIS PAR MOIS



SE PERFECTIONNER

Dans le domaine algo

Participer aux concours
type Kaggle et événements

100% DS, ateliers

AGENDA

Introduction

Qu'est-ce qu'une
visualisation de
données...efficiente ?

Zooms

Focus sur quelques outils

Démonstration

Exemples en pratique

AGENDA

Introduction

Qu'est-ce qu'une
visualisation de
données...efficiente ?

Zooms

Focus sur quelques outils

Demo

Exemples en pratique

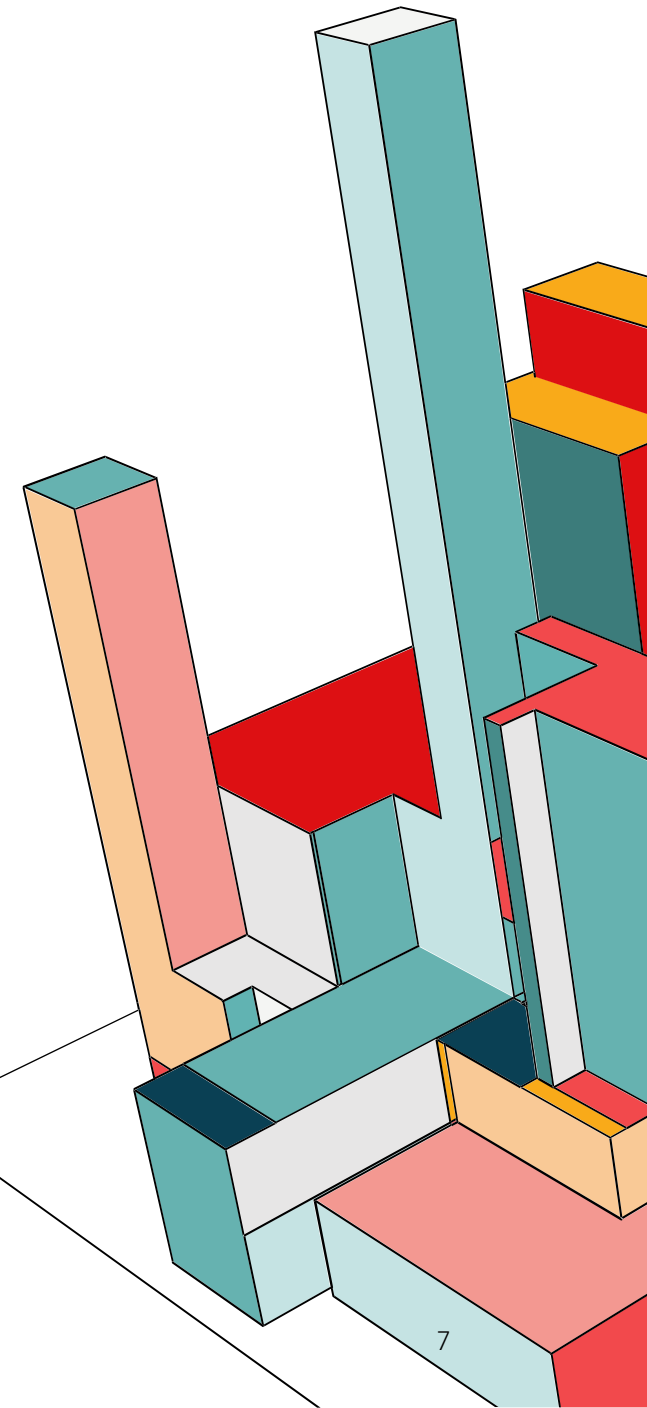
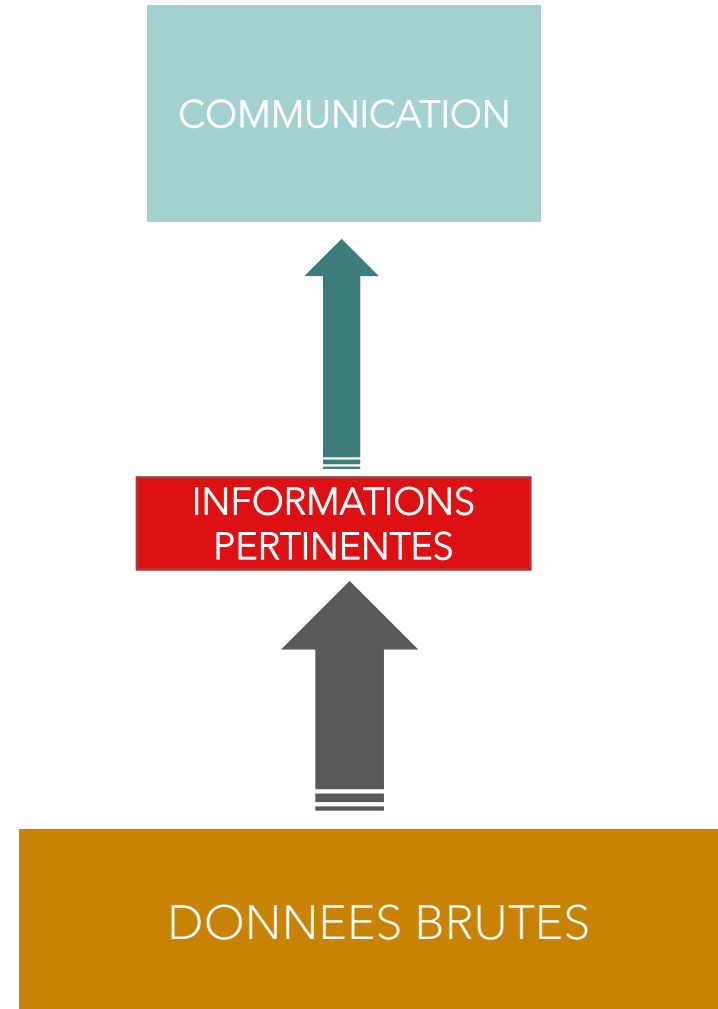
VISUALISATION DE DONNEES

UNE DEFINITION (POSSIBLE)

La data visualisation est la discipline qui consiste à mettre en forme de manière lisible et ergonomique un ensemble de données brutes, par l'intermédiaire de graphiques, de tableaux ou de cartes colorées... **dans le but d'éclairer et mettre en évidence des informations en apparence complexes ou noyées dans une grande quantité de paramètres.**

SUJETS CONNEXES

- Meilleurs choix de représentations (cf. Zooms)
- Meilleures pratiques en visualisation de données
- (Data) Story telling



NOTRE OBJECTIF DU JOUR

VISUALISATION DE DONNEES EFFICIENTE

La data visualisation est la discipline qui consiste à mettre en forme de manière lisible et ergonomique un ensemble de données brutes, par l'intermédiaire de graphiques, de tableaux ou de cartes colorées... dans le but d'éclairer et mettre en évidence des informations en apparence complexes ou noyées dans une grande quantité de paramètres.

CONTRAINTES

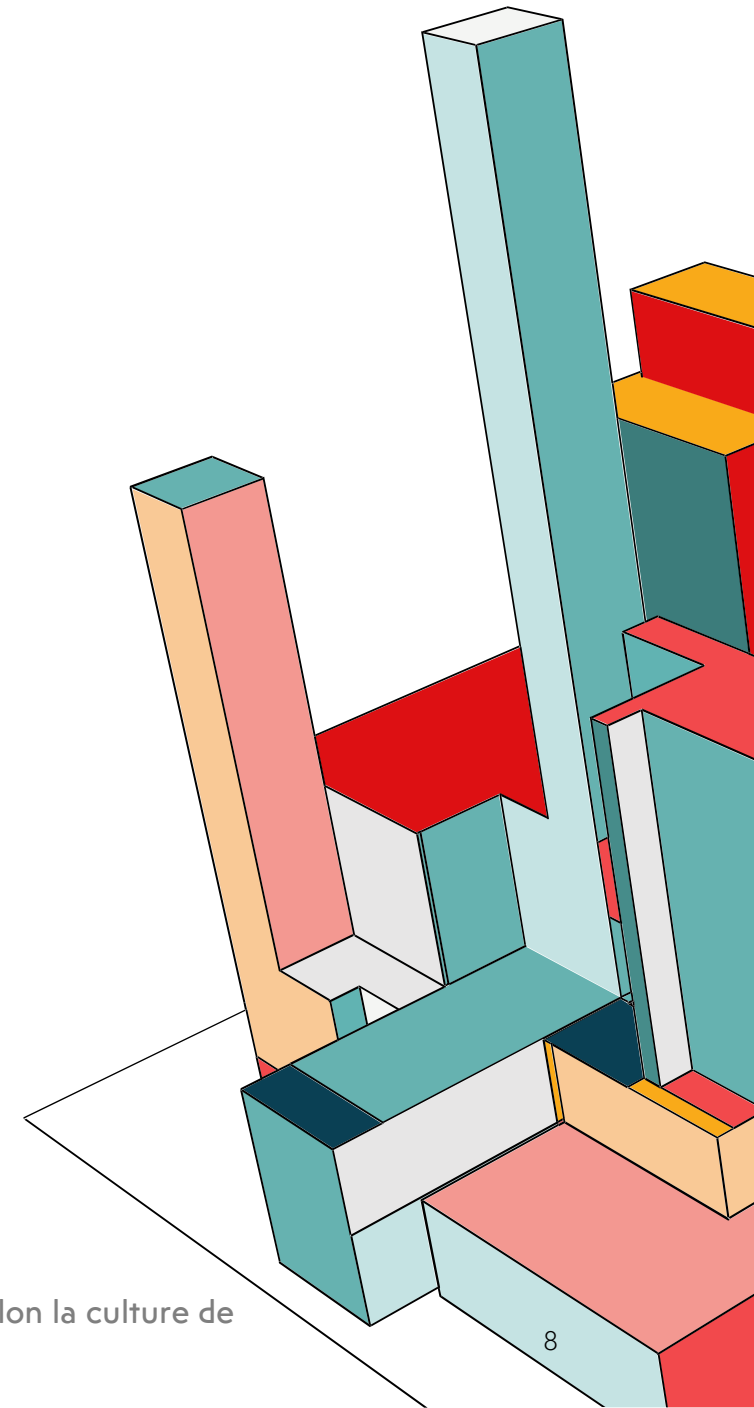
- R ou Python
- Basé sur du code uniquement (exit Tableau et PowerBI)

CRITERES D'EFFICIENCE*

- Transition aisée d'un prototype à une application "production-ready"
- Transition aisée de graphes statiques vers des graphes interactifs
- Ecosystème robuste (y compris support commercial et/ou communauté open source)

15/11/2022

*vous pouvez rajouter d'autres critères selon la culture de votre équipe/organisation





RECOMMANDATIONS GENERALES

UN SEUL LANGAGE

S'appuyer sur vos compétences R ou Python

RESTER PRAGMATIQUE ET SIMPLE

- Ne pas céder aux sirènes de l'"exotisme" (cf. Zooms sur les meilleures représentations)
- S'appuyer sur l'écosystème et la communauté dans laquelle vous évoluez
- Ne pas hésiter à copier (sans plagier)

INTEGRER LA VISUALISATION DANS VOTRE PROCESSUS GLOBAL D'ANALYSE

La visualisation des données est présente depuis l'exploration des données jusqu'à la publication d'un rapport/d'un tableau de bord ("dashboard")

AGENDA

Introduction

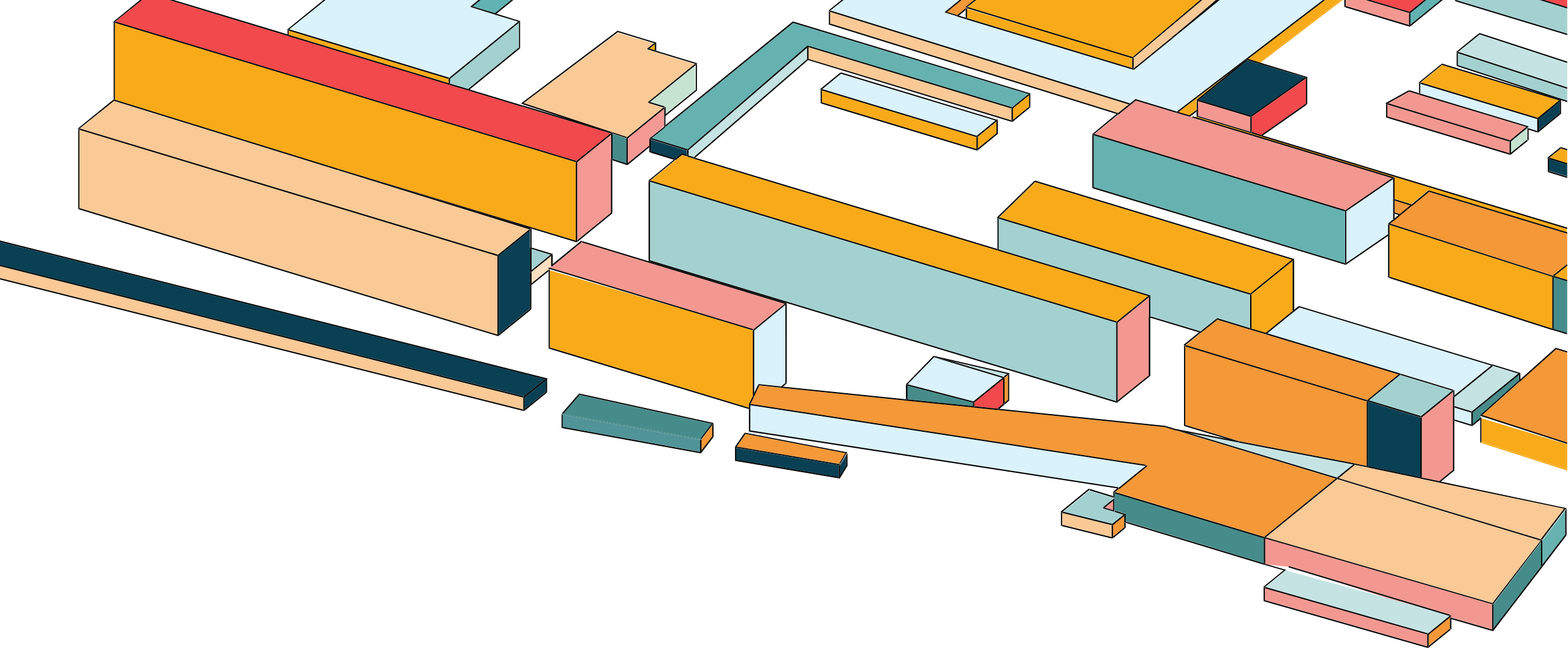
Qu'est-ce qu'une
visualisation de
données...efficiente ?

Zooms

Focus sur quelques outils

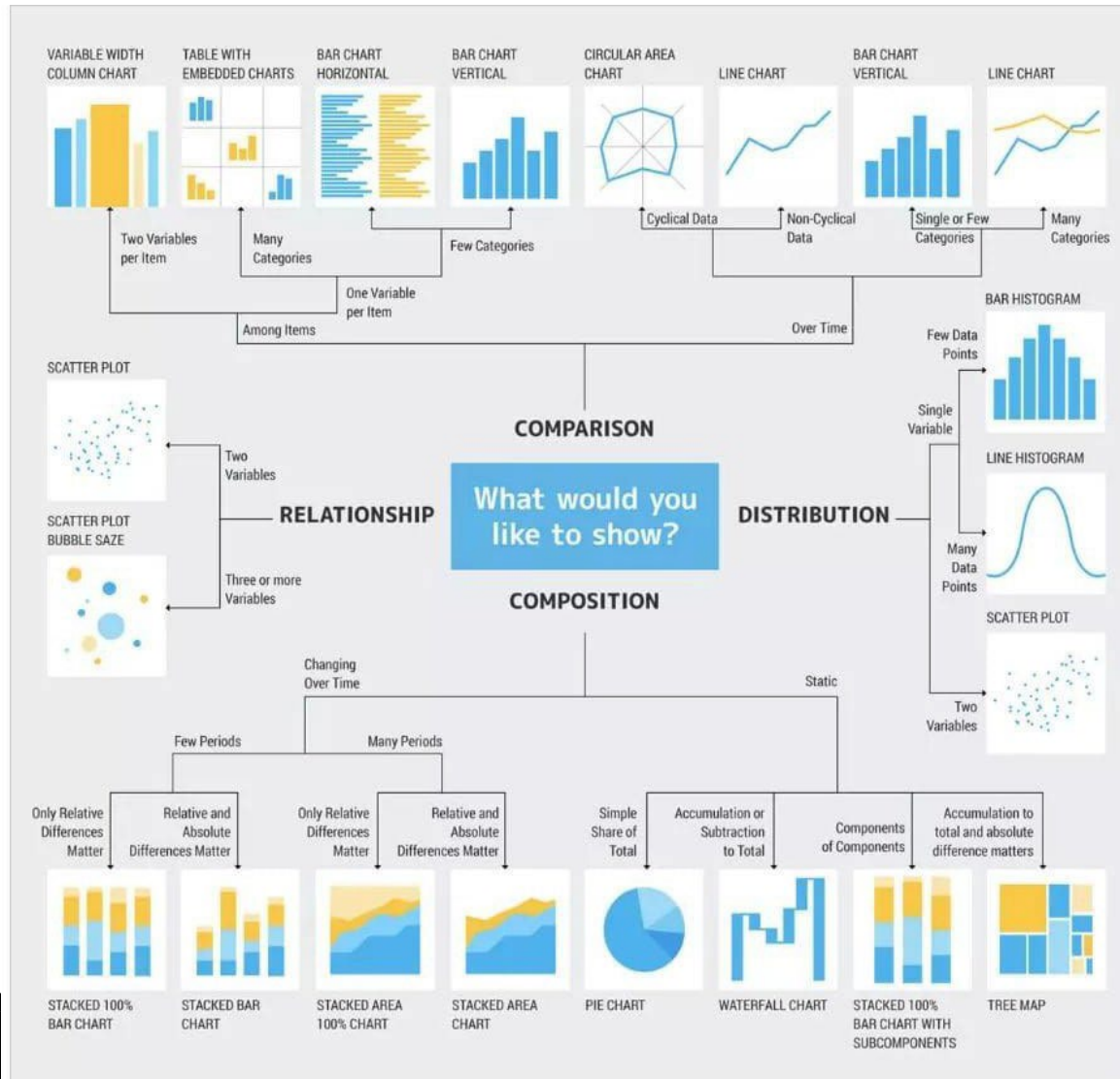
Demo

Exemples en pratique



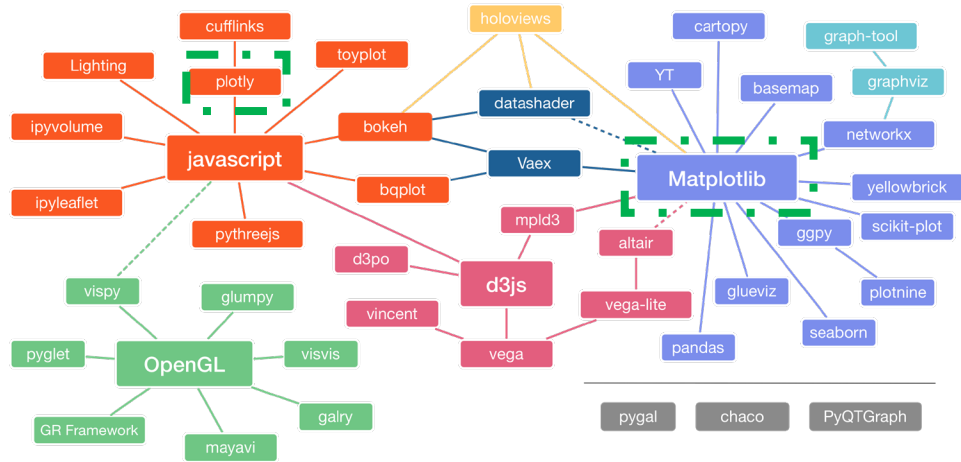
ZOOMS

CHOIX PERTINENTS DE REPRESENTATIONS



Best choice:
<https://www.data-to-viz.com/>

LES INCONTOURNABLES





matplotlib



Shiny

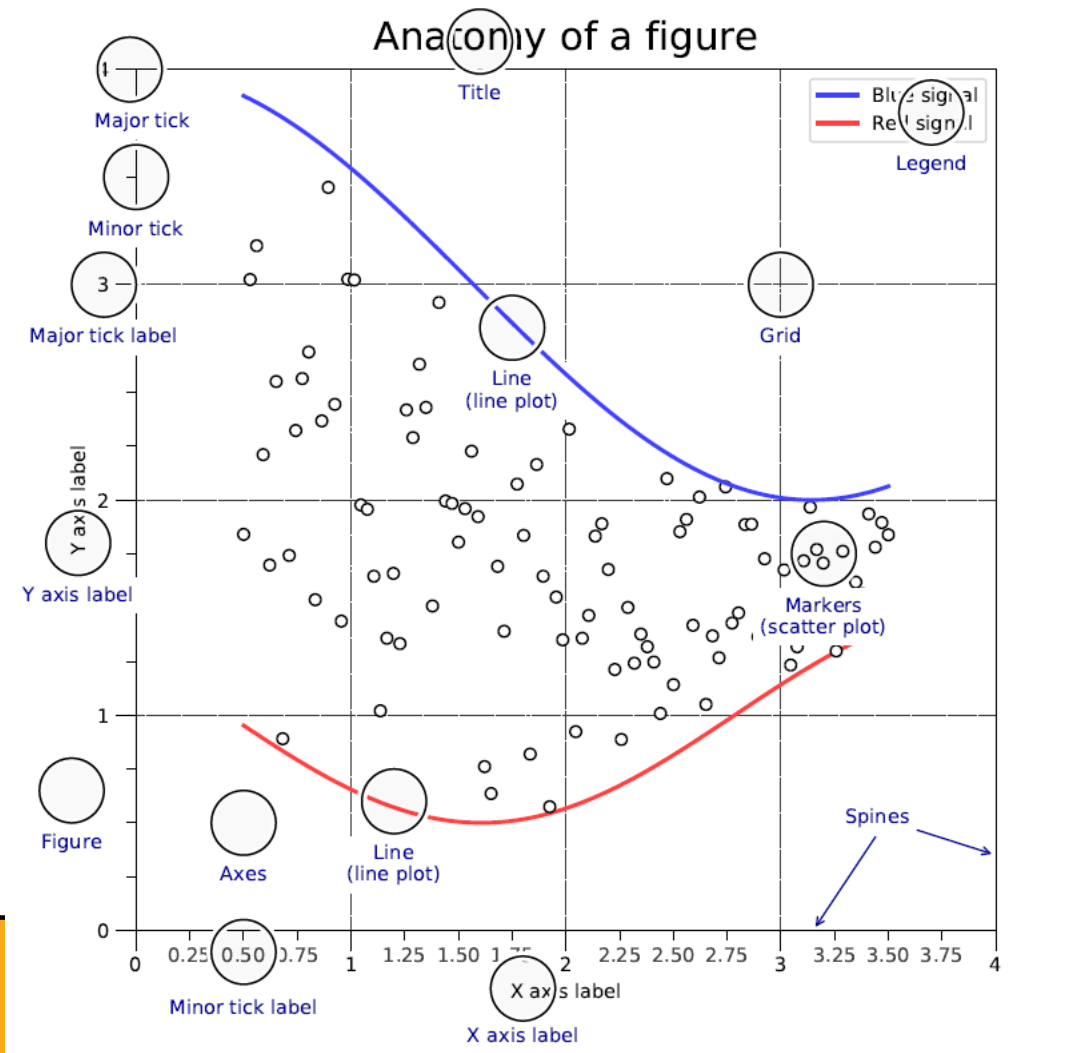


Non traités

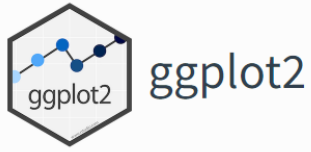
| | |
|---|--|
|  Plotly Python Open Source Graphing Library Star 12,425 |  Plotly R Open Source Graphing Library Star 2,280 |
|---|--|



UN PEU DE “REVISION”



R

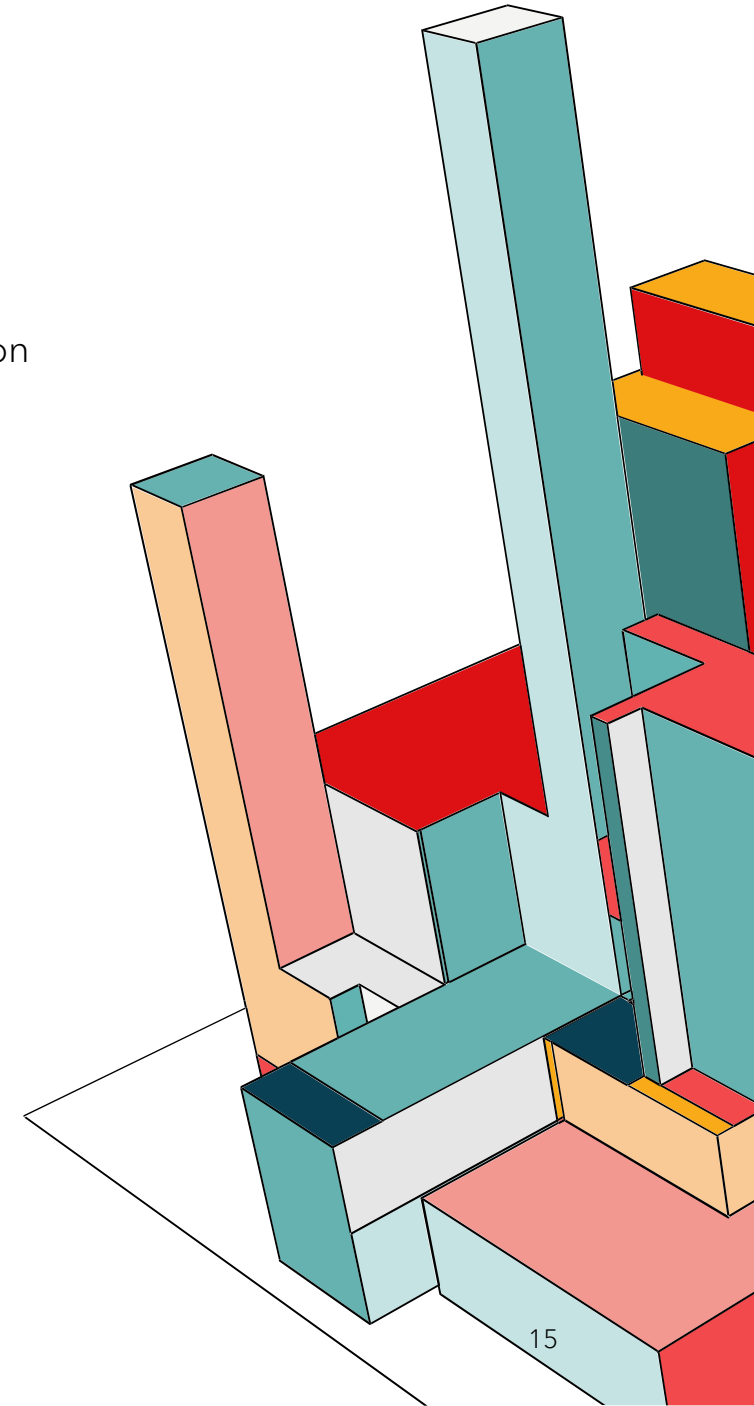


Shiny

CRITERES D'EFFICIENCE

- Transition aisée d'un prototype à une application "production-ready" +++
- Transition aisée de graphes statiques vers des graphes interactifs +++
- Ecosystème robuste (y compris support commercial et/ou communauté open source) +++

- Prototypage (RMarkdown) → application Shiny
- Ggplot2 → extensions existent pour rendre les graphiques interactifs
- Ecosystème et communauté



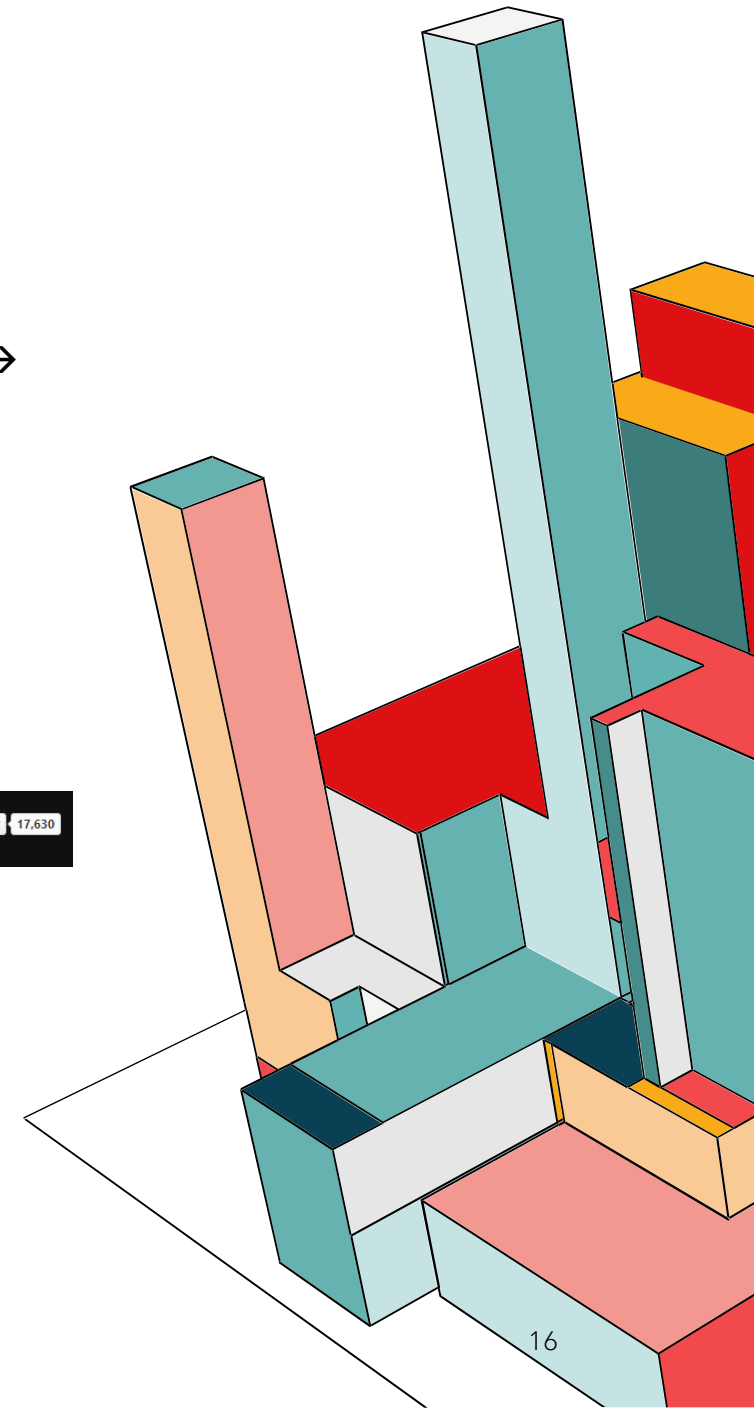
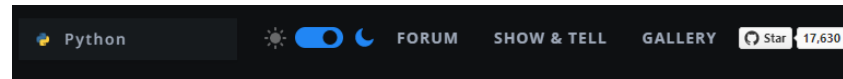
PYTHON



CRITERES D'EFFICIENCE

- Transition aisée d'un prototype à une application "production-ready" +++
- Transition aisée de graphes statiques vers des graphes interactifs +++
- Ecosystème robuste (y compris support commercial et/ou communauté open source) +++

- Prototypage (Jupyter+ Plotly express) → application Dash
- Plotly : tout est interactif
- Ecosystème et [communauté](#)



AGENDA

Introduction

Qu'est-ce qu'une
visualisation de
données...efficiente ?

Zooms

Focus sur quelques outils

Démonstration

Exemples en pratique

R



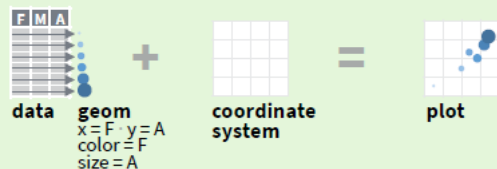
Shiny

Basics

ggplot2 is based on the **grammar of graphics**, the idea that you can build every graph from the same components: a **data** set, a **coordinate system**, and **geoms**—visual marks that represent data points.



To display values, map variables in the data to visual properties of the geom (**aesthetics**) like **size**, **color**, and **x** and **y** locations.



Complete the template below to build a graph.

```
ggplot (data = <DATA> +
<GEOM_FUNCTION> (mapping = aes (<MAPPINGS>),
stat = <STAT>, position = <POSITION>) +
<COORDINATE_FUNCTION> +
<FACET_FUNCTION> +
<SCALE_FUNCTION> +
<THEME_FUNCTION>
```

required

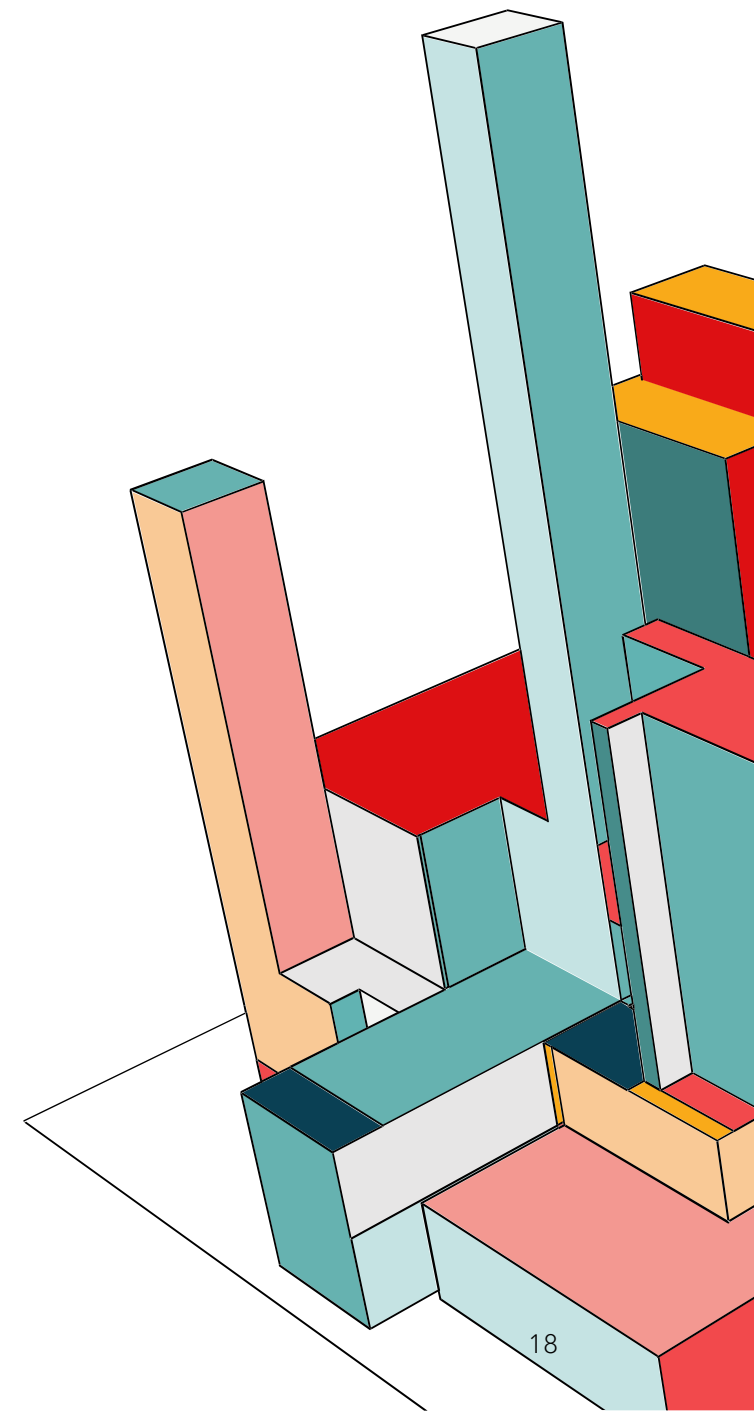
Not required, sensible defaults supplied

DEMO

- Prototypage (RMarkdown) → application Shiny

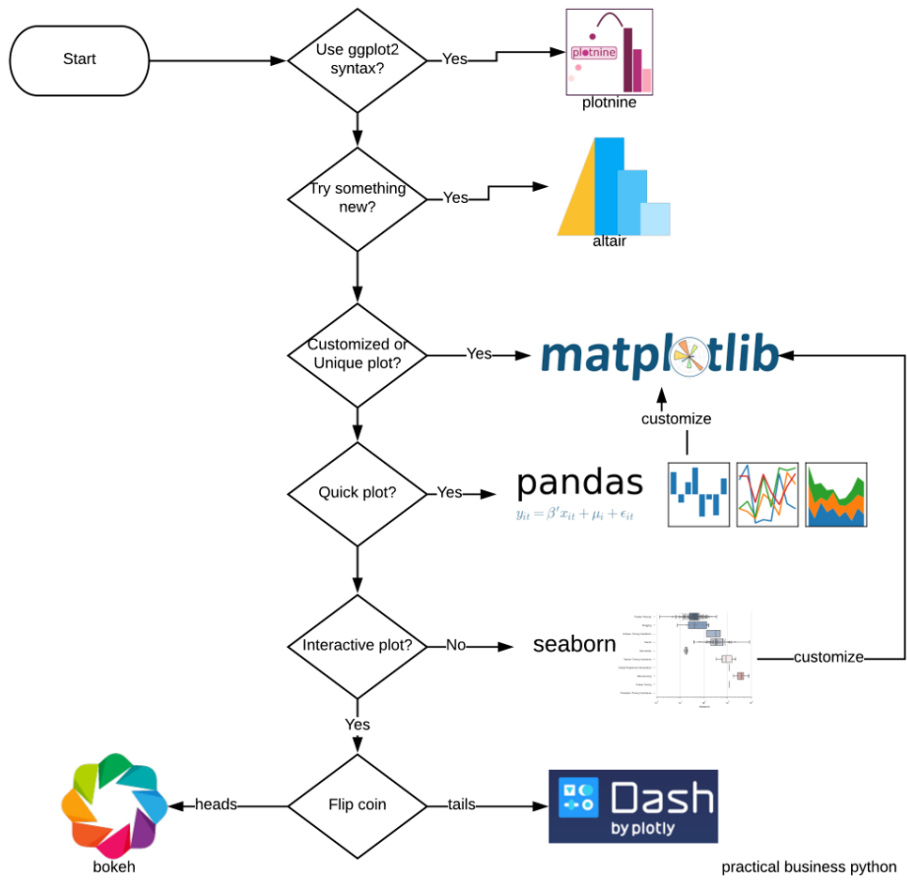
<https://ggplot2.tidyverse.org/>

- Ggplot2 → extensions ganimate (<https://exts.ggplot2.tidyverse.org/gallery/>) OU ggirpah (<https://davidgohel.github.io/ggiraph/index.html>)



PYTHON

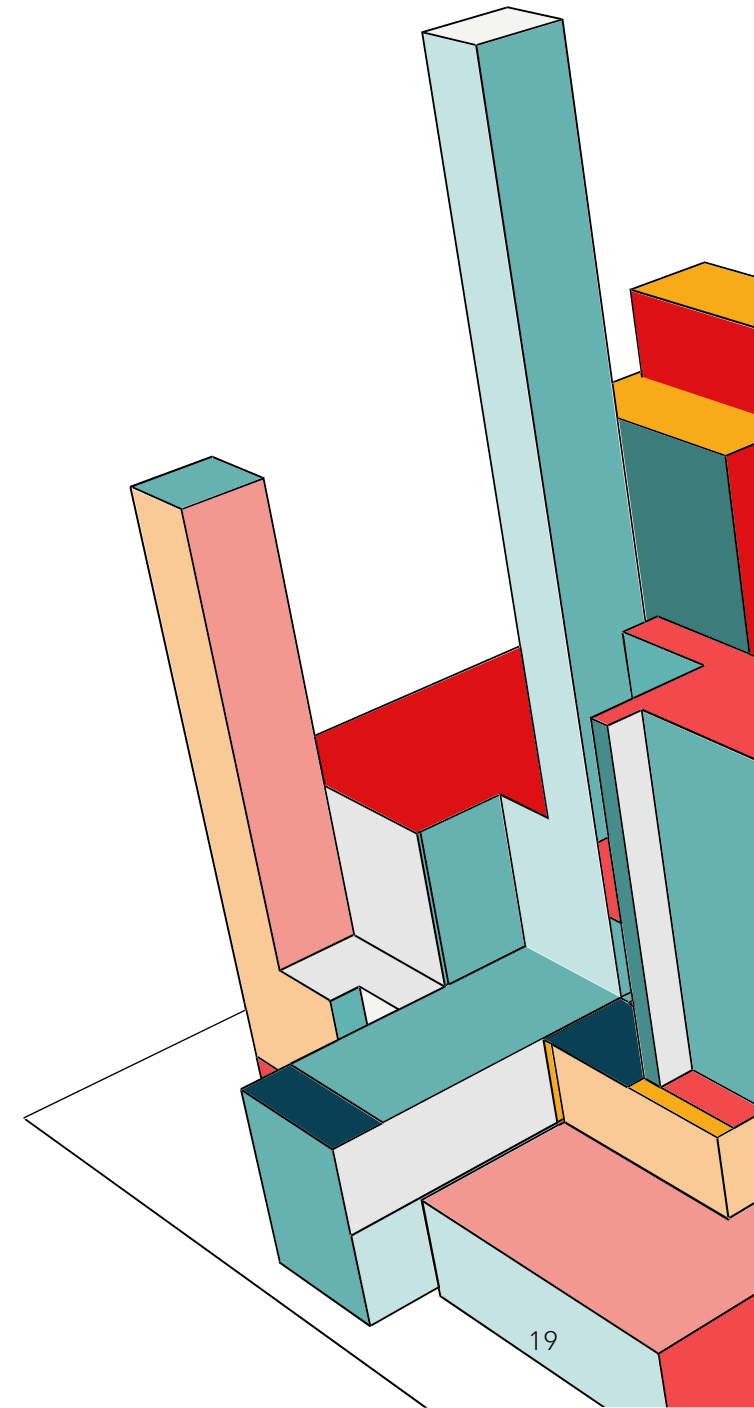
Choosing a python visualization tool



DEMOS:

- Seaborn
- Pandas + Plotly
- Pandas + Plotly + Dash

- Ecosystème et [communauté](#)

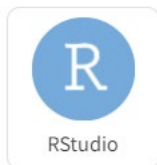


Comment choisir pour Python ?

BONUS ! QUARTO, LE PETIT NOUVEAU

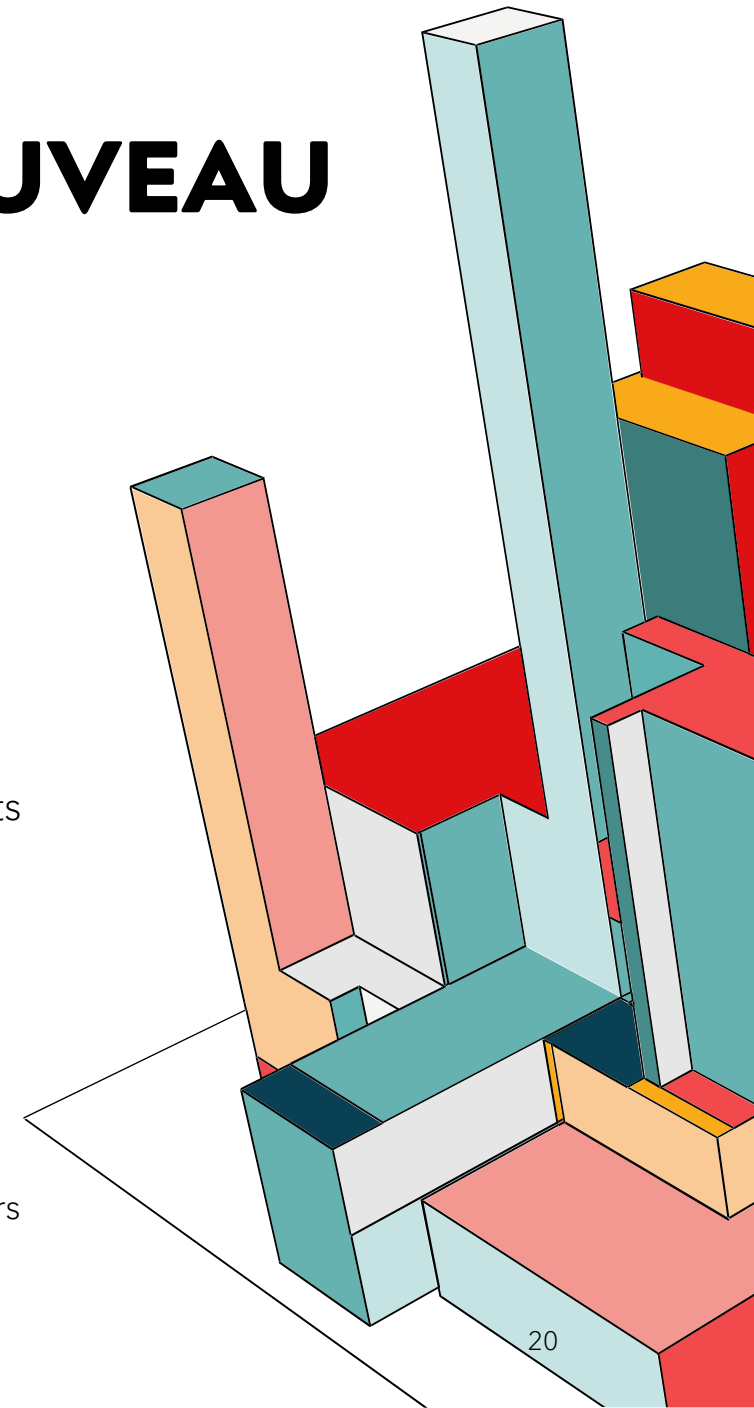
CRITERES D'EFFICIENCE

- Transition aisée d'un prototype à une application "production-ready" + + +
- Transition aisée de graphes statiques vers des graphes interactifs + + +
- Ecosystème robuste (y compris support commercial et/ou communauté open source) +



A TESTER

- "Beyond R Shiny"
- Similarités avec Rmarkdown mais couvre aussi Python en plus de R
 - Base = notebook Jupyter (ou format spécifique .qmd)
- Outputs :
 - Reports, articles, documents interactifs (Word, html)
 - Sites web
 - Livres
- Outillage dans VS Code en plus de Rstudio ou Jupyter → intéressant pour les utilisateurs de Python



POUR S'INSCRIRE AU CLUB ALGO SUR LE SITE DE L'INSTITUT DES ACTUAIRES



The screenshot shows the website's navigation bar with social media icons (LinkedIn, Instagram, YouTube, Email) and a "NEWSLETTERS" link. On the right, there are buttons for "ADHÉRER" and "SE CONNECTER". The main header includes the Institut des ACTUAIRES logo and the tagline "Éclairer les risques, tracer l'avenir". A dark navigation bar contains menu items: "DÉCOUVRIR L'INSTITUT", "DEVENIR ACTUAIRE", "SE PERFECTIONNER", "CONTRIBUER À L'INSTITUT", "SE DOCUMENTER", and "CHERCHER UN EMPLOI".

The main content area is titled "Rejoindre un groupe de travail" and includes a breadcrumb trail: "Accueil > Contribuer à l'institut > Commissions techniques et groupes de travail > Rejoindre un g...". A yellow warning box states: "⚠ L'accès à cet aux membres de". Below this is a login form with the heading "J'AI UN COMPTE" and a field for "Identifiant ou email".

A sidebar menu lists options: "CONTRIBUER À L'INSTITUT", "AGENDA DES ÉVÉNEMENTS", "LIVRET D'ACCUEIL DU BÉNÉVOLE", "DEVENIR JURY OU PRÉNOTEUR", "COMMISSIONS TECHNIQUES ET GROUPE DE TRAVAIL" (highlighted with a blue arrow), "REJOINDRE UN COMITÉ D'ORGANISATION", and "REJOINDRE UN CLUB". The main content area also features buttons for "REJOINDRE UNE COMMISSION TECHNIQUE" and "REJOINDRE UN GROUPE DE TRAVAIL" (highlighted with a mouse cursor).

At the bottom, a yellow banner contains the text "SE JOINDRE UN CLUB" and "Le jour et groupe de travail/rejoindre-un-groupe-de-travail-2001". A small note at the bottom right reads "tre nom de naissance le cas échéant".

MERCI

Jean-Claude

RAZAFINDRAKOTO

SCOR Data Analytics Advocate

RESSOURCES

R (ggplot2, Shiny)

- <https://ggplot2.tidyverse.org/>
- Extensions ggplot2, dont ggiraph
<https://exts.ggplot2.tidyverse.org/gallery/>

Python

- Seaborn <https://seaborn.pydata.org/index.html>
- Plotly <https://plotly.com/python/plotly-fundamentals/>
- Plotly Express <https://plotly.com/python/plotly-express/>

Quarto

- <https://quarto.org/>

