

Big Data



Big Data

CM1 – Introduction

Nathan Sanchiz-Viel

Université de Limoges

2025

Big Data

Big Data ?

Pourquoi Big Data ≠ Data classique ?

Big Data = Beaucoup de données ?

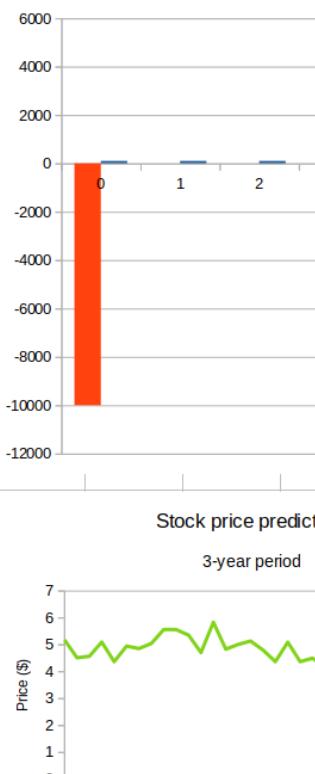
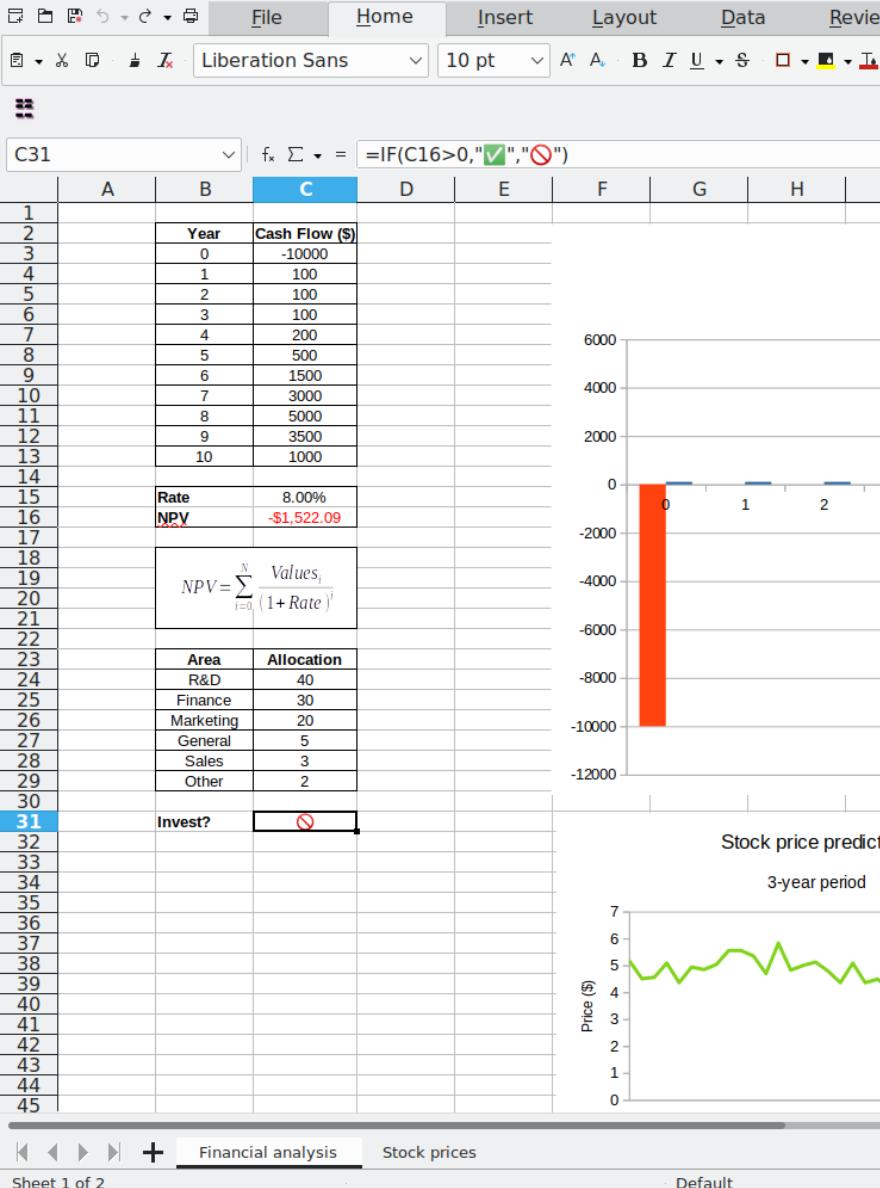
Big Data = Beaucoup de données ?

Nop



#





```

1 command = ""
2 while command != "quit":
3     command = input("> ").lower()
4     if command == "start":
5         print('Car starts.')
6     elif command == "stop":
7         print('Car stopped.')
8     elif command == "help":
9         print("""
10            start - to start the car.
11            stop - to stop the car.
12            quit - quit the game.
13        """)
14    else:
15        print("Sorry we don't understand.")

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL JUPYTER

```

> &c:/Python/Python38/python.exe c:/Users/sybef/Opdracht.py
> stop
> &c:/Python/Python38/python.exe c:/Users/sybef/Opdracht.py
> start
> stop
> ssdf
> help
> &c:/Python/Python38/python.exe c:/Users/sybef/Opdracht.py
> start
> &c:/Python/Python38/python.exe c:/Users/sybef/Opdracht.py
> start
> start
>

```

Big Data, c'est quand :

- le volume
- la vitesse
- la distribution

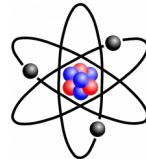
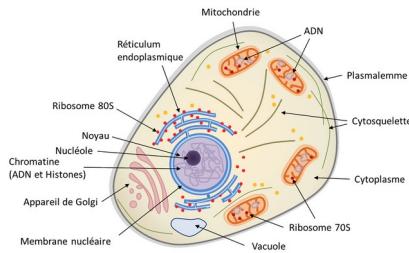
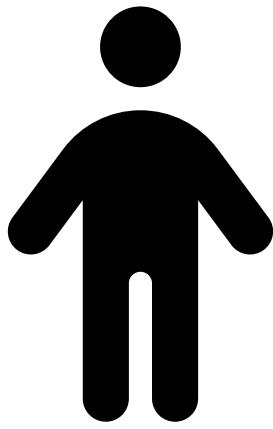
changent la façon de penser les
algorithmes

Pourquoi le langage Python ?

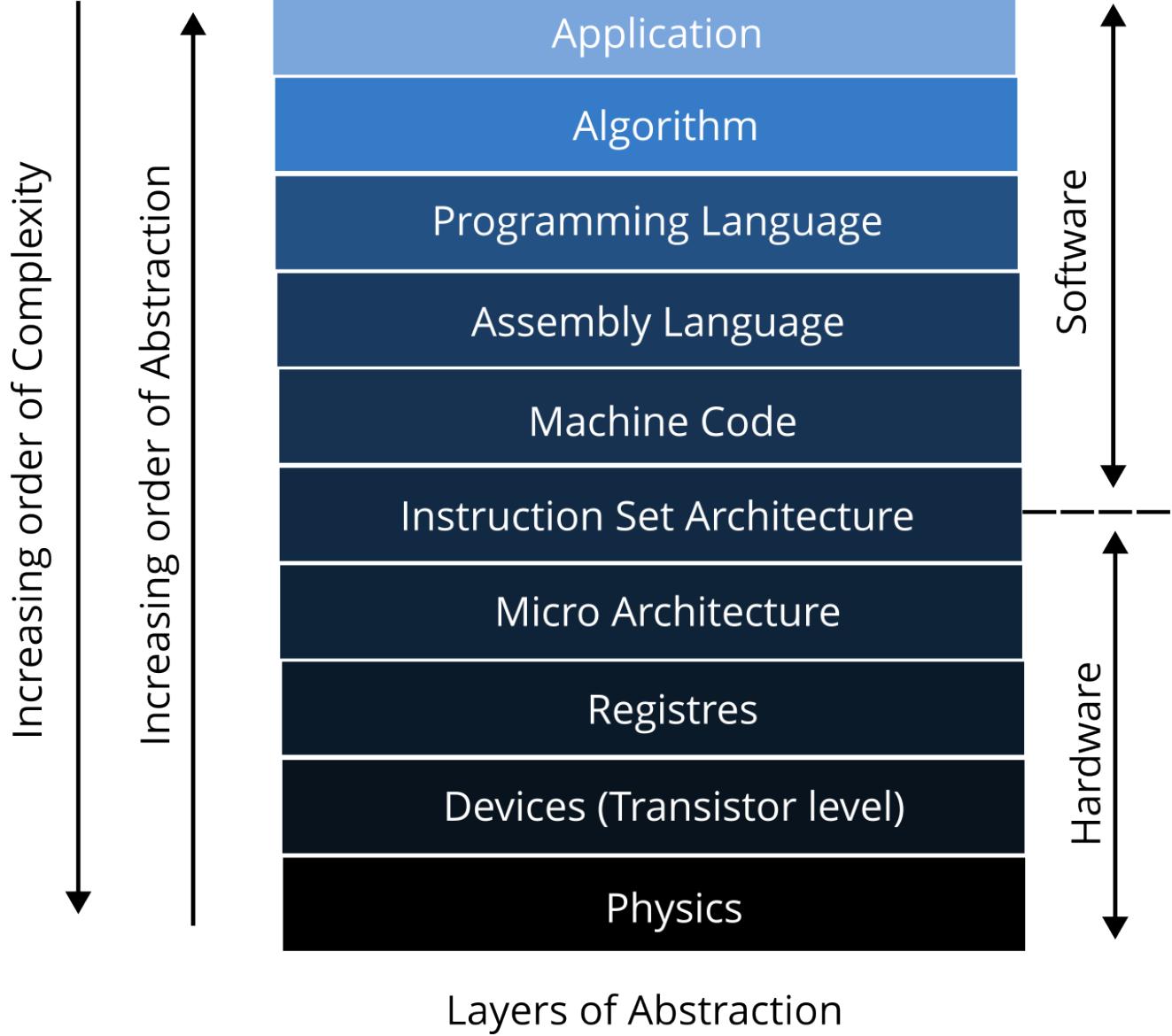
Python
= haut niveau d'abstraction

Python
= quoi ?

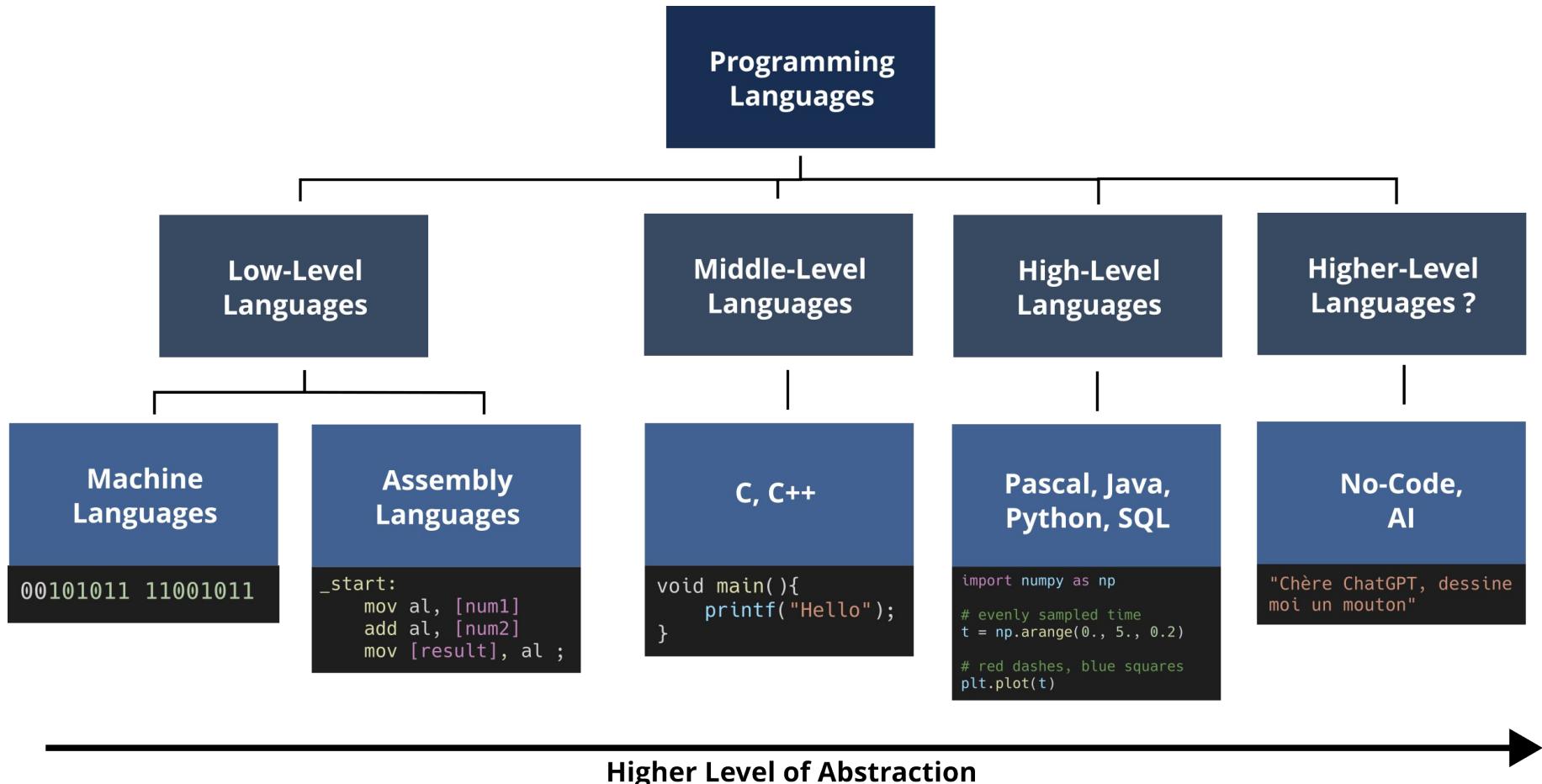
Niveaux d'abstraction



Niveaux d'abstraction

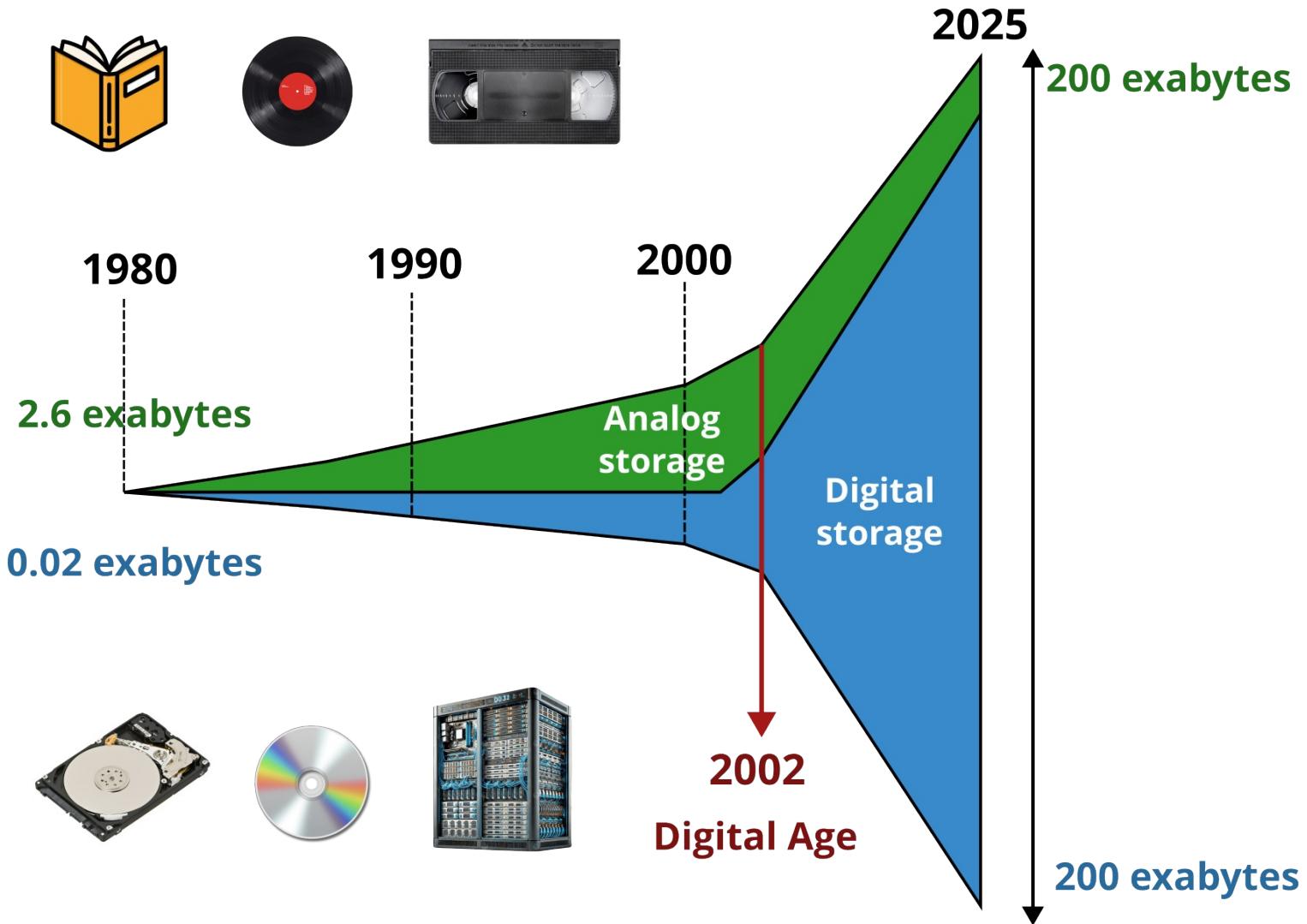


Python = haut niveau d'abstraction ?



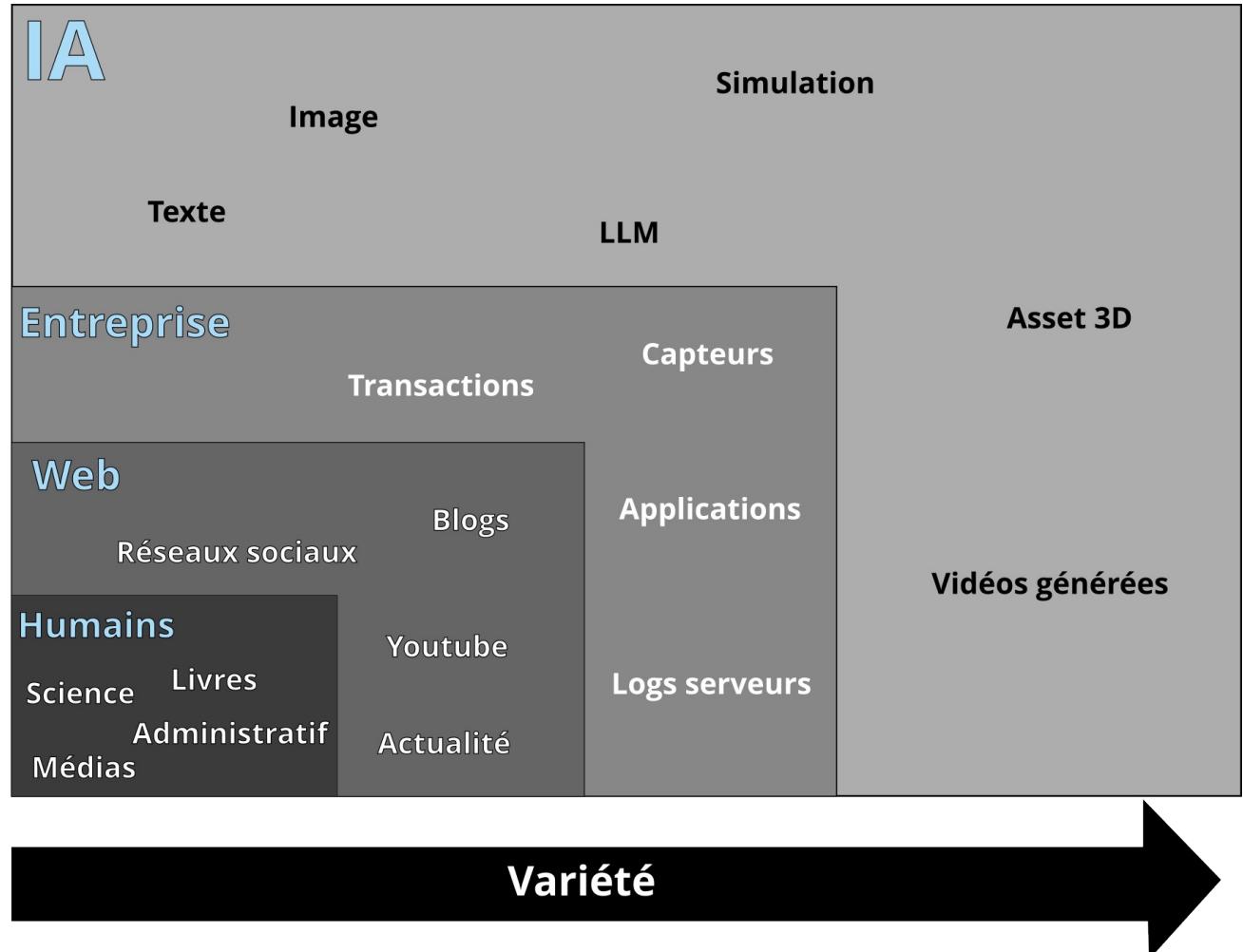
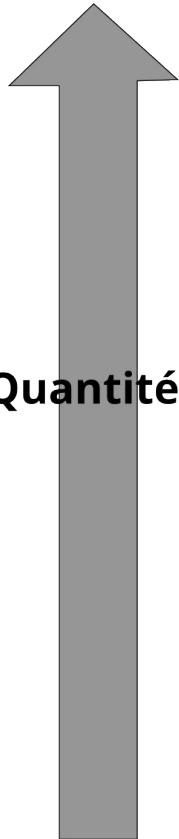
Big ?

Big Big data

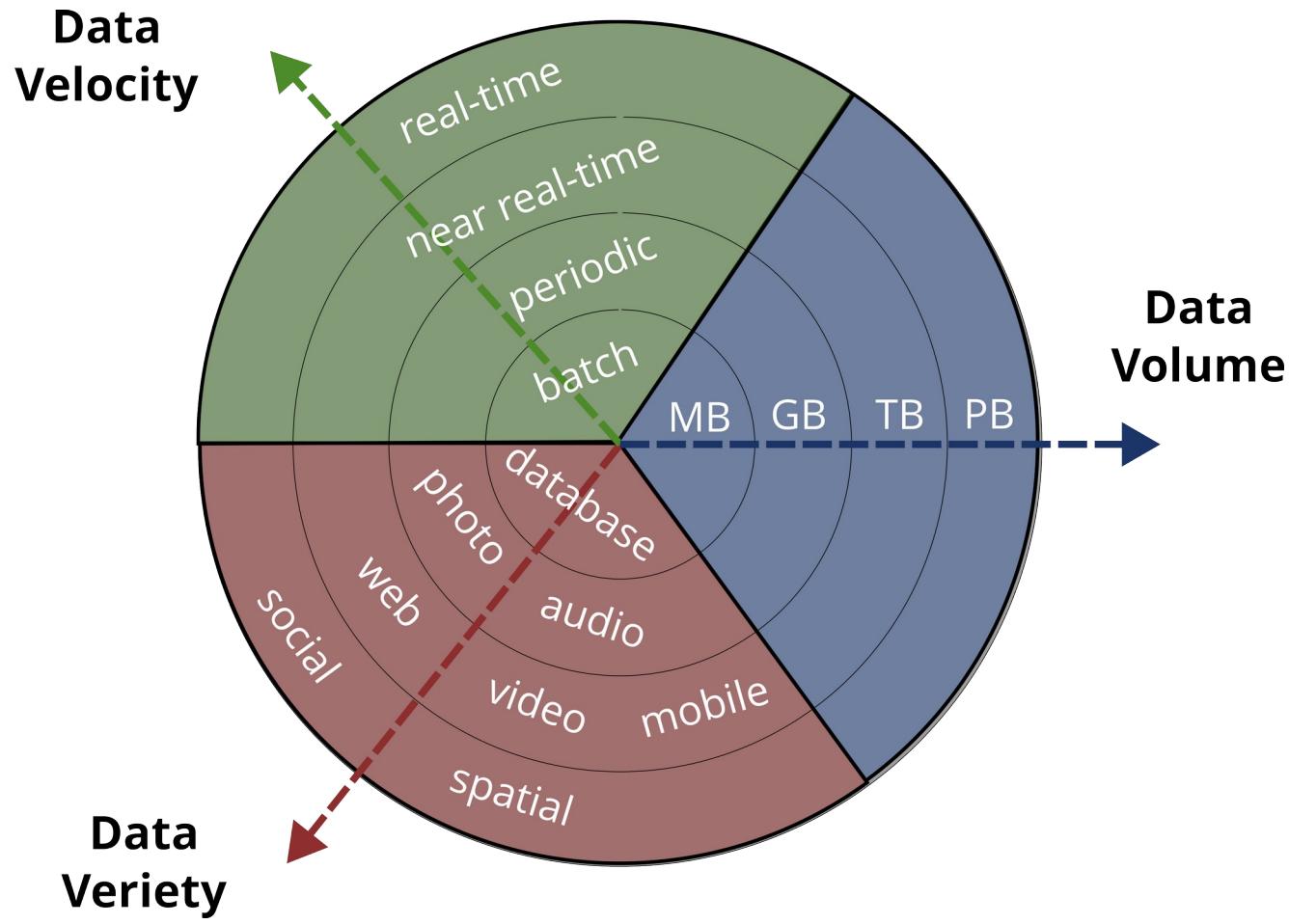


Big Big data

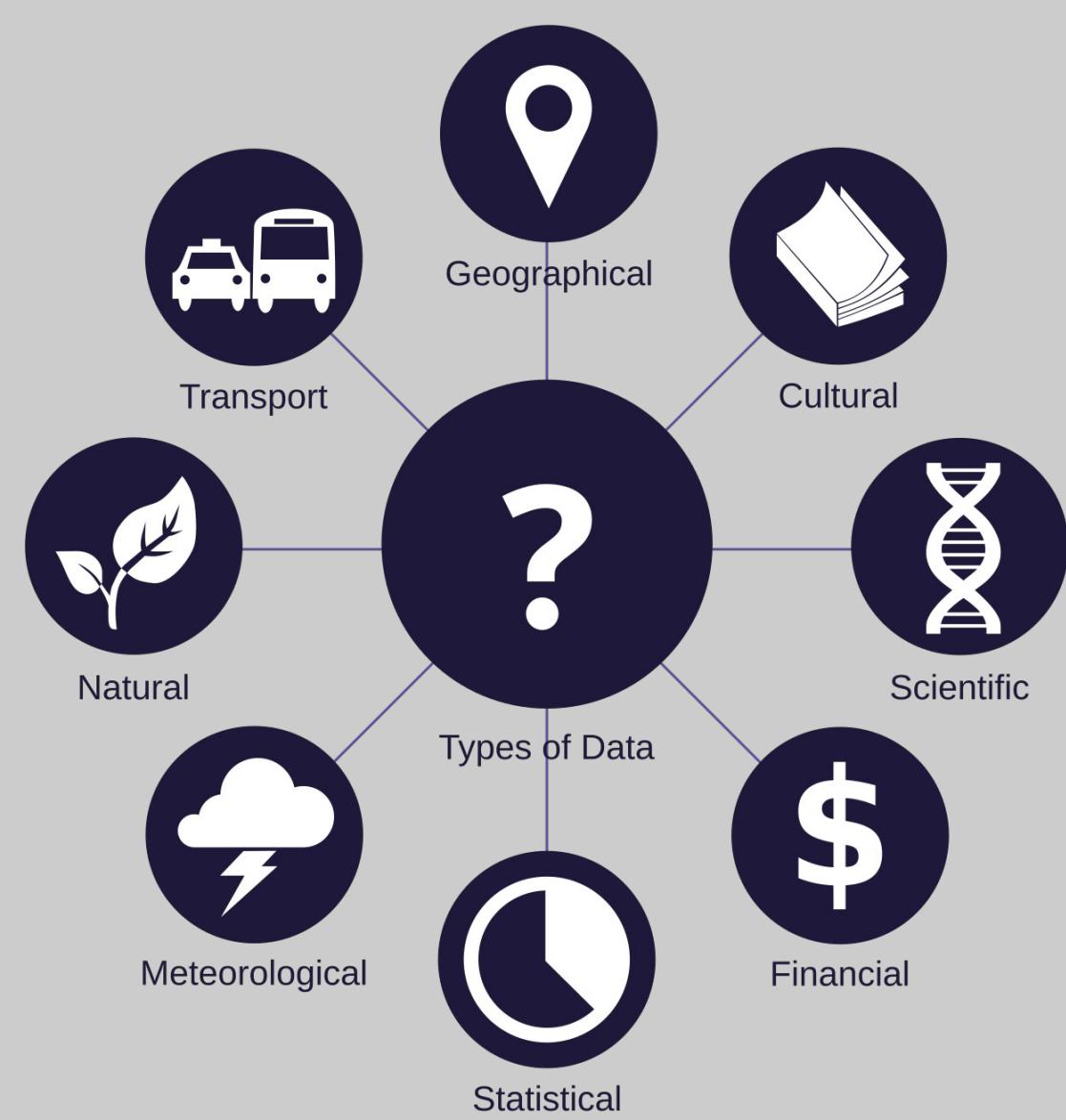
Quantité



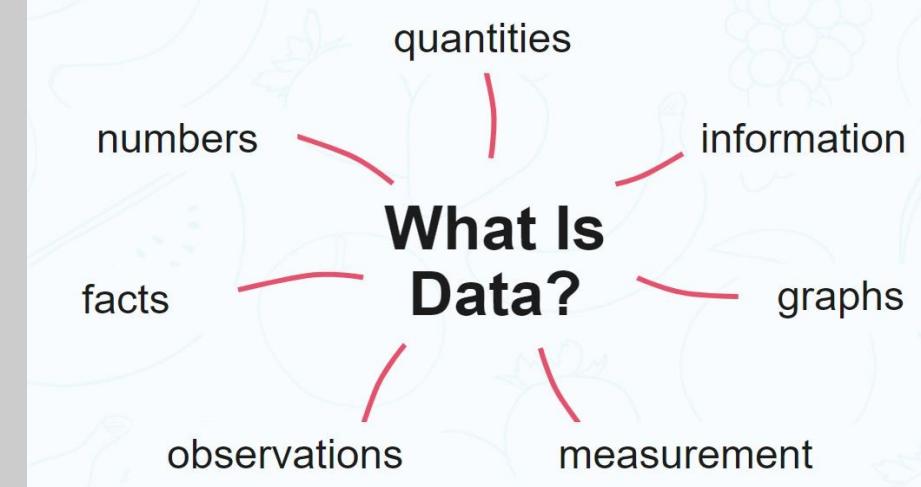
Multidimensional Big data



Data ?



A	B	C	D	E	F	G
Case Number	Date	Year	Type	Country	Area	Location
2017.06.11	2017-06-11	2017	Unprovoked	AUSTRALIA	Western Australia	Point Casuarina, Bunbury
2017.06.10.b	2017-06-10	2017	Unprovoked	AUSTRALIA	Victoria	Flinders, Mornington Peninsula
2017.06.10.a	2017-06-10	2017	Unprovoked	USA	Florida	Ponce Inlet, Volusia County
2017.06.07.R	Reported 07-JR	2017	Unprovoked	UNITED KINGDOM	South Devon	Bantham Beach
2017.06.04	2017-06-04	2017	Unprovoked	USA	Florida	Middle Simbo Reef off Boca Chica, Monroe Co.
2017.06.02	2017-06-02	2017	Unprovoked	BAHAMAS	New Providence	Athol Island
2017.05.30	2017-05-30	2017	Provoked	USA	South Carolina	Awendaw, Charleston County
2017.05.28	2017-05-28	2017	Unprovoked	USA	Florida	Off Jupiter
2017.05.27	2017-05-27	2017		AUSTRALIA	New South Wales	Evans Head
2017.05.12	2017-05-12	2017	Unprovoked	UNITED ARAB EMIRATES	Sharjah,	Khor Fakkan
2017.05.06.R	Reported 06-M	2017	Provoked	AUSTRALIA	Queensland	Weipa
2017.05.06	2017-05-05	2017	Unprovoked	MEXICO	Baja California Sur	Los Arbolitos, Cabo Pulmo
2017.05.03	2017-05-03	2017	Invalid	USA	California	Sunset Beach, Orange County
2017.04.29.d	2017-04-29	2017	Unprovoked	USA	California	San Onofre, San Diego County
2017.04.29.c	2017-04-29	2017	Unprovoked	SOUTH AFRICA	Western Cape Province	Keurbooms, Plettenberg Bay
2017.04.29.b	2017-04-29	2017	Unprovoked	USA	South Carolina	Folly Beach, Charleston County
2017.04.29.a	2017-04-29	2017	Unprovoked	REUNION	Saint-Leu	Pointe au Sal
2017.04.27	2017-04-27	2017	Unprovoked	NEW ZEALAND	South Island	Porpoise Bay, Catlins
2017.04.26	2017-04-26	2017	Invalid	USA	Florida	
2017.04.22	2017-04-22	2017	Unprovoked	ST HELENA, British overseas territory	Ascension Island	English Bay
2017.04.20	2017-04-20	2017	Invalid	USA	South Carolina	Georgetown County
2017.04.17.b	2017-04-17	2017	Unprovoked	USA	Florida	Daytona Beach, Volusia County
2017.04.17.a	2017-04-17	2017	Unprovoked	AUSTRALIA	Western Australia	Kelpies near Wylie Bay
2017.04.14	2017-04-14	2017	Unprovoked	USA	Hawaii	Kekaha Beach, Kauai
2017.04.13	2017-04-13	2017	Unprovoked	USA	Florida	Hana Park, Jacksonville, Duval County
2017.04.12.b	2017-04-12	2017	Unprovoked	USA	Florida	St. Augustine
2017.04.12.a	2017-04-12	2017	Invalid	SOUTH AFRICA	KwaZulu-Natal	Protea Banks
2017.04.11	2017-04-11	2017	Unprovoked	USA	Florida	Ormond Beach, Volusia County
2017.04.10.b	2017-04-10	2017	Unprovoked	USA	Florida	Melbourne Beach, Brevard County
2017.04.10.a	2017-04-10	2017	Unprovoked	USA	Florida	Melbourne Beach, Brevard County
2017.04.06	2017-04-06	2017	Unprovoked	USA	Florida	Daytona, Volusia County
2017.04.05	2017-04-05	2017	Unprovoked	USA	Florida	New Smyrna Beach, Volusia County
2017.04.02.b	2017-04-02	2017	Unprovoked	AUSTRALIA	Queensland	Moreton Bay
2017.04.02.a	2017-04-02	2017	Unprovoked	USA	Florida	Destin, Okaloosa County
2017.03.27	2017-03-27	2017	Unprovoked	USA	Florida	New Smyrna Beach, Volusia County
2017.03.19	2017-03-19	2017	Unprovoked	AUSTRALIA	New South Wales	The Farm
2017.03.18	2017-03-18	2017	Unprovoked	USA	California	Monterey Bay
2017.03.09	Reported 09-M	2017	Unprovoked	BAHAMAS	Great Exuma	
2017.02.25	2017-02-25	2017	Unprovoked	AUSTRALIA	Western Australia	Mauds Point



Data ?

= Observation encodée

	A	B	C	D	E	F	
1	Case Number	Date	Year	Type	Country	Area	Location
2	2017.06.11	2017-06-11	2017	Unprovoked	AUSTRALIA	Western Australia	Point Casuarina, Bunbury
3	2017.06.10.b	2017-06-10	2017	Unprovoked	AUSTRALIA	Victoria	Flinders, Mornington Peninsula
4	2017.06.10.a	2017-06-10	2017	Unprovoked	USA	Florida	Ponce Inlet, Volusia County
5	2017.06.07.R	Reported 07-J	2017	Unprovoked	UNITED KINGDOM	South Devon	Banham Beach

Data = Observation encodée

Ok mais **data = information ?**

Data = Observation encodée

Ok mais **data = information ?**

Nop

Data ≠ information

Information = Interprétation utile

Qu'est ce que l'information ?

- **in-** → « dans »
- **formare** → « donner forme »
- **-actio** → « action de »

“ Action de donner une forme à quelque chose ”

Data ≠ information

Information = Interprétation utile

= **Structure extraite à partir d'un ensemble de données**

Data = Observation encodée

information = Interprétation utile

Et pour finir

Information = connaissance ?

Data = Observation encodée

information = Interprétation utile

Et pour finir

Information = connaissance ?

Nop

Data = Observation encodée

information = Interprétation utile

connaissance = information intégrée

= capacité potentiel d'en tirer profis

Data = Observation encodée

information = Interprétation utile

connaissance = information intégrée

= capacité potentiel d'en tirer profis

= Potentiel d'action / d'agir sur le monde

donnée = Observation encodée

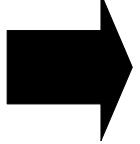
information = Interprétation utile

connaissance = Potentiel d'action

Cas pratique

A	B	C	D	E
1	Case Number	Date	Year Type	Country Area
2	2017.06.11	2017-06-11	2017-Unprovoked	AUSTRALIA Western Australia
3	2017.06.10.b	2017-06-10	2017-Unprovoked	AUSTRALIA Victoria
4	2017.06.10.a	2017-06-10	2017-Unprovoked	USA Florida
5	2017.06.07.R		2017-Unprovoked	UNITED KINGDOM South Devon
6		2017-06-04	2017-Unprovoked	AUSTRALIA Tasmania
7		2017-06-02	2017-Unprovoked	BAHAMAS New Providence
8	2017.05.30	2017-05-30	2017-Provoked	USA South Carolina
9	2017.05.29	2017-05-28	2017-Unprovoked	USA Florida
10	2017.05.27	2017-05-27	2017-Unprovoked	AUSTRALIA New South Wales
11	2017.05.12	2017-05-12	2017-Unprovoked	UNITED ARAB EMIRATES Sharjah,
12	2017.05.06.R		2017-Provoked	AUSTRALIA Queensland
13	2017.05.06	2017-05-05	2017-Unprovoked	MEXICO Big California Sur
14	2017.05.03	2017-05-03	2017-Invalid	USA California
15	2017.04.29.d	2017-04-29	2017-Unprovoked	USA California
16	2017.04.29.c	2017-04-29	2017-Unprovoked	SOUTH AFRICA Western Cape Province
17	2017.04.29.b	2017-04-29	2017-Unprovoked	USA South Carolina
18		2017-04-29	2017-Unprovoked	REUNION Sainte-Lucie
19	2017.04.27	2017-04-27	2017-Unprovoked	NEW ZEALAND South Island
20	2017.04.26	2017-04-26	2017-Invalid	USA Florida
21	2017.04.22	2017-04-22	2017-Unprovoked	ST HELENA, British overseas territory Ascension Island
22	2017.04.20	2017-04-20	2017-Invalid	USA South Carolina
23	2017.04.17.b	2017-04-17	2017-Unprovoked	USA Florida
24	2017.04.17.a	2017-04-17	2017-Unprovoked	AUSTRALIA Western Australia
25	2017.04.14	2017-04-14	2017-Unprovoked	USA Hawaii
26	2017.04.13	2017-04-13	2017-Unprovoked	USA Florida
27	2017.04.12.b	2017-04-12	2017-Unprovoked	USA Florida
28	2017.04.12.a	2017-04-12	2017-Invalid	SOUTH AFRICA KwaZulu-Natal
29	2017.04.11	2017-04-11	2017-Unprovoked	USA Florida
30	2017.04.10.b	2017-04-10	2017-Unprovoked	USA Florida
31	2017.04.10.a	2017-04-10	2017-Unprovoked	USA Florida
32	2017.04.06	2017-04-06	2017-Unprovoked	USA Florida
33	2017.04.05	2017-04-05	2017-Unprovoked	USA Florida
34	2017.04.02.b	2017-04-02	2017-Unprovoked	AUSTRALIA Queensland
35	2017.04.02.a	2017-04-02	2017-Unprovoked	USA Florida
36	2017.03.27	2017-03-27	2017-Unprovoked	USA Florida
37	2017.03.19	2017-03-19	2017-Unprovoked	AUSTRALIA New South Wales
38	2017.03.18	2017-03-18	2017-Unprovoked	USA California
39	2017.03.09		2017-Unprovoked	BAHAMAS Great Exuma
40	2017.02.29	2017-02-29	2017-Unprovoked	AUSTRALIA Western Australia
41	2017.02.24	2017-02-24	2017-Unprovoked	AUSTRALIA Western Australia
42	2017.02.21	2017-02-21	2017-Unprovoked	REUNION Saint-André
43	2017.02.17	2017-02-17	2017-Unprovoked	AUSTRALIA Queensland
44	2017.02.11	2017-02-11	2017-Unprovoked	USA Florida
45	2017.02.06.b	2017-02-06	2017-Prevoked	BRAZIL Pernambuco
46	2017.02.06.a	2017-02-06	2017-Boat	SOUTH AFRICA KwaZulu-Natal
47	2017.02.01.b	2017-02-01	2017-Unprovoked	USA South Carolina
48	2017.02.01.a	2017-02-01	2017-Unprovoked	BAHAMAS Bimini
49	2017.01.29	2017-01-29	2017-Unprovoked	USA Florida
50	2017.01.21	2017-01-21	2017-Unprovoked	AUSTRALIA Queensland
51	2017.01.13.b	2017-01-13	2017-Unprovoked	AUSTRALIA Queensland
52	2017.01.13.a	2017-01-13	2017-Unprovoked	USA Florida
53	2017.01.09	2017-01-09	2017-Unprovoked	INDONESIA Bali
54	2017.01.08.R		Invalid	AUSTRALIA Queensland
55	2017.01.05	2017-01-05	2017-Unprovoked	USA Florida
56	2016.12.30	2016-12-30	2016-Unprovoked	AUSTRALIA New South Wales
57	2016.12.27	2016-12-27	2016-Unprovoked	USA Florida
58	2016.12.24	2016-12-24	2016-Unprovoked	AUSTRALIA Western Australia
59	2016.12.19	2016-12-19	2016-Unprovoked	SOUTH AFRICA Western Cape Province
60	2016.12.14	2016-12-14	2016-Unprovoked	AUSTRALIA Western Australia
61	2016.12.11	2016-12-11	2016-Unprovoked	USA Florida
62	2016.12.10	2016-12-10	2016-Unprovoked	ENGLAND Suffolk
63	2016.12.06	2016-12-06	2016-Prevoked	AUSTRALIA New South Wales
64	2016.12.01	2016-12-01	2016-Unprovoked	AUSTRALIA New South Wales
65	2016.11.28	2016-11-28	2016-Prevoked	USA Victoria
66	2016.11.14	2016-11-14	2016-Unprovoked	USA Hawaii
67	2016.10.30	2016-10-30	2016-Unprovoked	USA Florida
68	2016.10.29	2016-10-29	2016-Unprovoked	USA Florida
69	2016.10.24	2016-10-24	2016-Unprovoked	AUSTRALIA New South Wales
70	2016.10.21	2016-10-21	2016-Unprovoked	USA Hawaii
71	2016.10.15	2016-10-15	2016-Unprovoked	AUSTRALIA Queensland
72	2016.10.14	2016-10-14	2016-Unprovoked	USA Hawaii
73	2016.10.13.R		2016-Invalid	MEXICO
74	2016.10.12	2016-10-12	2016-Unprovoked	AUSTRALIA New South Wales
75	2016.10.10	2016-10-10	2016-Unprovoked	USA Oregon
76	2016.10.02	2016-10-02	2016-Unprovoked	USA Florida
77	2016.10.01	2016-10-01	2016-Unprovoked	USA Florida
78	2016.09.26	2016-09-26	2016-Unprovoked	AUSTRALIA New South Wales
79	2016.09.18.c	2016-09-18	2016-Unprovoked	USA Florida
80	2016.09.15.d	2016-09-15	2016-Unprovoked	USA Florida
81	2016.09.15.b	2016-09-15	2016-Unprovoked	USA Florida
82	2016.09.17.b	2016-09-17	2016-Unprovoked	USA California
83	2016.09.17.a	2016-09-17	2016-Unprovoked	AUSTRALIA Victoria
84	2016.09.15	2016-09-16	2016-Unprovoked	AUSTRALIA Victoria
85	2016.09.15.R		2016-Boat	AUSTRALIA Western Australia
86	2016.09.11	2016-09-11	2016-Unprovoked	USA Florida
87	2016.09.07	2016-09-07	2016-Unprovoked	USA Hawaii
88	2016.09.06	2016-09-06	2016-Unprovoked	NEW CALEDONIA North Province
89	2016.09.05.b	2016-09-05	2016-Unprovoked	USA South Carolina
90	2016.09.05.a	2016-09-05	2016-Unprovoked	AUSTRALIA Western Australia
91	2016.09.04	2016-09-04	2016-Unprovoked	USA Florida
92	2016.09.01	2016-09-01	2016-Unprovoked	MEXICO California
93	2016.08.30	2016-08-30	2016-Invalid	MEXICO California
94	2016.08.29.b	2016-08-29	2016-Unprovoked	USA Florida
95	2016.08.29.a	2016-08-29	2016-Unprovoked	USA Florida
96	2016.08.27	2016-08-27	2016-Unprovoked	REUNION
97	2016.08.25	2016-08-25	2016-Unprovoked	USA Florida
98	2016.08.07	2016-08-07	2016-Unprovoked	BAHAMAS New Providence Island
99	2016.08.06	2016-08-06	2016-Unprovoked	HAWAII Hawaii
100	2016.08.04	2016-08-04	2016-Unprovoked	USA Florida
101	2016.08.03	2016-08-03	2016-Unprovoked	INDONESIA Raja Ampat

Data



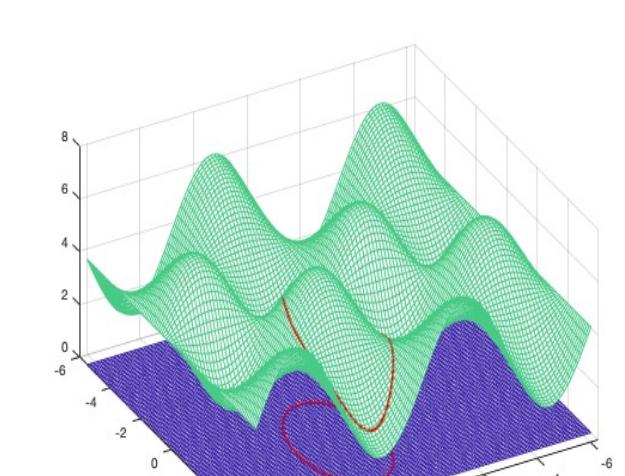
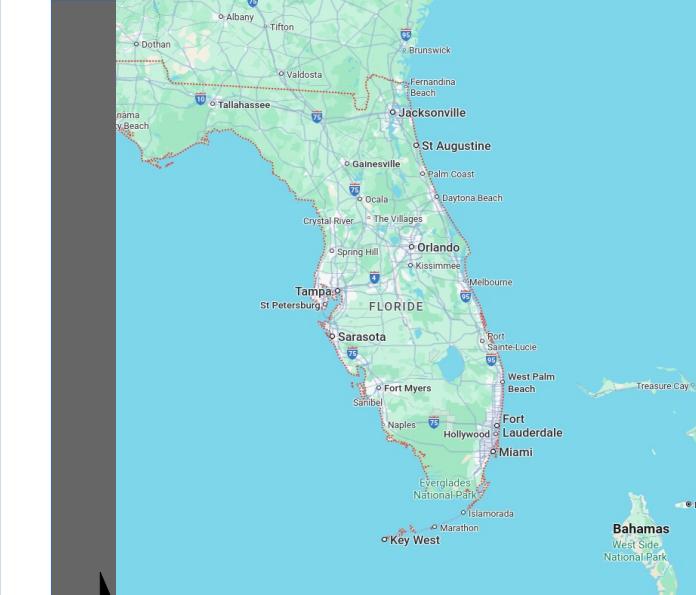
- # Base de données sur les attaques de requin sur les 100 dernières années
- Zone géographique
 - Date
 - Activité de la victime
 - Blessure

A	B	C	D	E
1	Case Number	Date	Year Type	Country Area
2	2017.06.11	2017-06-11	2017-Unprovoked	AUSTRALIA Western Australia
3	2017.06.10.b	2017-06-10	2017-Unprovoked	AUSTRALIA Victoria
4	2017.06.10.a	2017-06-10	2017-Unprovoked	USA Florida
5	2017.06.07.R	2017	2017-Unprovoked	UNITED KINGDOM South Devon
6	2017.06.06	2017-06-04	2017-Unprovoked	USA Oregon
7	2017.06.02	2017-06-02	2017-Unprovoked	BAHAMAS New Providence
8	2017.05.30	2017-05-30	2017-Provoked	USA South Carolina
9	2017.05.29	2017-05-28	2017-Unprovoked	USA Florida
10	2017.05.27	2017-05-27	2017	AUSTRALIA New South Wales
11	2017.05.12	2017-05-12	2017-Unprovoked	UNITED ARAB EMIRATES Sharjah,
12	2017.05.06.R	2017	2017-Provoked	AUSTRALIA Queensland
13	2017.05.06	2017-05-05	2017-Unprovoked	MEXICO Big California Sur
14	2017.05.03	2017-05-03	2017-Invalid	USA California
15	2017.04.29.d	2017-04-29	2017-Unprovoked	USA California
16	2017.04.29.c	2017-04-29	2017-Unprovoked	SOUTH AFRICA Western Cape Province
17	2017.04.29.b	2017-04-29	2017-Unprovoked	USA South Carolina
18	2017.04.29.a	2017-04-29	2017-Unprovoked	REUNION Sainte-Lucie
19	2017.04.27	2017-04-27	2017-Unprovoked	NEW ZEALAND South Island
20	2017.04.26	2017-04-26	2017-Invalid	USA Florida
21	2017.04.22	2017-04-22	2017-Unprovoked	ST HELENA, British overseas territory Ascension Island
22	2017.04.20	2017-04-20	2017-Invalid	USA South Carolina
23	2017.04.17	2017-04-17	2017-Unprovoked	USA Florida
24	2017.04.17.b	2017-04-17	2017-Unprovoked	AUSTRALIA Western Australia
25	2017.04.14	2017-04-14	2017-Unprovoked	USA Hawaii
26	2017.04.13	2017-04-13	2017-Unprovoked	USA Florida
27	2017.04.12.b	2017-04-12	2017-Unprovoked	USA Florida
28	2017.04.12.a	2017-04-12	2017-Invalid	SOUTH AFRICA KwaZulu-Natal
29	2017.04.11	2017-04-11	2017-Unprovoked	USA Florida
30	2017.04.10.b	2017-04-10	2017-Unprovoked	USA Florida
31	2017.04.10.a	2017-04-10	2017-Unprovoked	USA Florida
32	2017.04.06	2017-04-06	2017-Unprovoked	USA Florida
33	2017.04.05	2017-04-05	2017-Unprovoked	USA Florida
34	2017.04.02.b	2017-04-02	2017-Unprovoked	AUSTRALIA Queensland
35	2017.04.02.a	2017-04-02	2017-Unprovoked	USA Florida
36	2017.03.27	2017-03-27	2017-Unprovoked	USA Florida
37	2017.03.19	2017-03-19	2017	AUSTRALIA New South Wales
38	2017.03.18	2017-03-18	2017-Unprovoked	USA California
39	2017.03.09	2017	2017-Unprovoked	BAHAMAS Great Exuma
40	2017.02.29	2017-02-25	2017-Unprovoked	AUSTRALIA Western Australia
41	2017.02.24	2017-02-24	2017-Unprovoked	AUSTRALIA Western Australia
42	2017.02.21	2017-02-21	2017-Unprovoked	REUNION Saint-André
43	2017.02.17	2017-02-17	2017-Unprovoked	AUSTRALIA Queensland
44	2017.02.11	2017-02-11	2017-Unprovoked	USA Florida
45	2017.02.06.b	2017-02-06	2017-Prevoked	BRAZIL Pernambuco
46	2017.02.06.a	2017-02-06	2017-Prevoked	SOUTH AFRICA KwaZulu-Natal
47	2017.02.01.b	2017-02-01	2017-Prevoked	USA South Carolina
48	2017.02.01.a	2017-02-01	2017-Prevoked	BAHAMAS Bimini
49	2017.01.27	2017-01-27	2017-Unprovoked	USA Florida
50	2017.01.21	2017-01-21	2017-Unprovoked	AUSTRALIA Queensland
51	2017.01.13.b	2017-01-13	2017-Unprovoked	AUSTRALIA Queensland
52	2017.01.13.a	2017-01-13	2017-Unprovoked	USA Florida
53	2017.01.09	2017-01-09	2017-Unprovoked	INDONESIA Bali
54	2017.01.08.R	2017	2017-Invalid	AUSTRALIA Queensland
55	2017.01.05	2017-01-05	2017-Unprovoked	USA Florida
56	2016.12.30	2016-12-30	2016-Unprovoked	AUSTRALIA New South Wales
57	2016.12.27	2016-12-27	2016-Unprovoked	USA Florida
58	2016.12.24	2016-12-24	2016-Unprovoked	AUSTRALIA Western Australia
59	2016.12.19	2016-12-19	2016-Unprovoked	SOUTH AFRICA Western Cape Province
60	2016.12.14	2016-12-14	2016-Unprovoked	AUSTRALIA Western Australia
61	2016.12.11	2016-12-11	2016-Unprovoked	USA Florida
62	2016.12.10	2016-12-10	2016-Unprovoked	ENGLAND Suffolk
63	2016.12.06	2016-12-06	2016-Prevoked	AUSTRALIA New South Wales
64	2016.12.01	2016-12-01	2016-Unprovoked	AUSTRALIA New South Wales
65	2016.11.28	2016-11-28	2016-Prevoked	USA Victoria
66	2016.11.14	2016-11-14	2016-Unprovoked	USA Hawaii
67	2016.10.30	2016-10-30	2016-Prevoked	USA Florida
68	2016.10.29	2016-10-29	2016-Prevoked	USA Florida
69	2016.10.24	2016-10-24	2016-Unprovoked	AUSTRALIA New South Wales
70	2016.10.21	2016-10-21	2016-Unprovoked	USA Hawaii
71	2016.10.15	2016-10-15	2016-Unprovoked	AUSTRALIA Queensland
72	2016.10.14	2016-10-14	2016-Unprovoked	USA Hawaii
73	2016.10.13.R	2016	2016-Invalid	MEXICO
74	2016.10.12	2016-10-12	2016-Prevoked	AUSTRALIA New South Wales
75	2016.10.10	2016-10-10	2016-Prevoked	USA Oregon
76	2016.10.02	2016-10-02	2016-Unprovoked	USA Florida
77	2016.10.01	2016-10-01	2016-Unprovoked	USA Florida
78	2016.09.26	2016-09-26	2016-Unprovoked	AUSTRALIA New South Wales
79	2016.09.18.c	2016-09-18	2016-Prevoked	USA Florida
80	2016.09.16.b	2016-09-16	2016-Prevoked	USA Florida
81	2016.09.15.a	2016-09-15	2016-Prevoked	USA Florida
82	2016.09.17.b	2016-09-17	2016-Prevoked	USA California
83	2016.09.17.a	2016-09-17	2016-Prevoked	AUSTRALIA Victoria
84	2016.09.15	2016-09-16	2016-Unprovoked	AUSTRALIA Victoria
85	2016.09.15.R	2016	2016-Invalid	AUSTRALIA Western Australia
86	2016.09.11	2016-09-11	2016-Prevoked	USA Florida
87	2016.09.07	2016-09-07	2016-Prevoked	USA Hawaii
88	2016.09.06	2016-09-06	2016-Unprovoked	NEW CALEDONIA North Province
89	2016.09.05.b	2016-09-05	2016-Unprovoked	USA South Carolina
90	2016.09.05.a	2016-09-05	2016-Unprovoked	AUSTRALIA Western Australia
91	2016.09.04	2016-09-04	2016-Unprovoked	USA Florida
92	2016.09.01	2016-09-01	2016-Prevoked	USA California
93	2016.08.30	2016-08-30	2016-Prevoked	MEXICO
94	2016.08.29.b	2016-08-29	2016-Unprovoked	USA Florida
95	2016.08.29.a	2016-08-29	2016-Unprovoked	USA Florida
96	2016.08.27	2016-08-27	2016-Unprovoked	REUNION
97	2016.08.25	2016-08-25	2016-Unprovoked	USA Florida
98	2016.08.07	2016-08-07	2016-Prevoked	BAHAMAS New Providence Island
99	2016.08.06	2016-08-06	2016-Unprovoked	HAWAII Hawaii
100	2016.08.04	2016-08-04	2016-Unprovoked	USA Florida
101	2016.08.03	2016-08-03	2016-Prevoked	NEW CALEDONIA Noumea Province



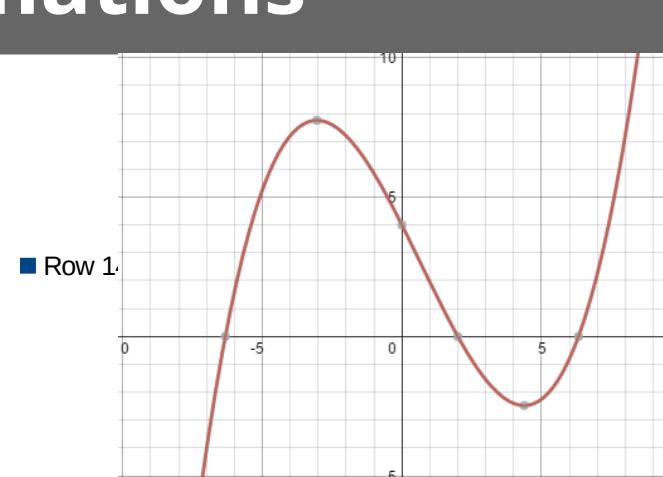
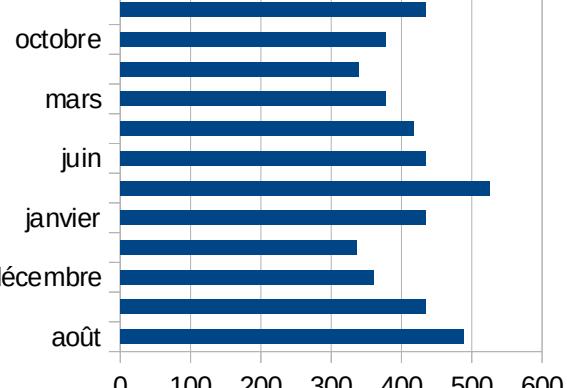
Ok mais et alors ?

A	B	C	D	E	
Case Number	Date	Year	Type	Country	Area
1	2017.06.11	2017	Unprovoked	AUSTRALIA	Western Australia
2	2017.06.10.b	2017	Unprovoked	AUSTRALIA	Victoria
4	2017.06.10.a	2017	Unprovoked	USA	Florida
5	2017.06.07.R	2017	Unprovoked	UNITED KINGDOM	South Devon
6	2017.06.04	2017	Unprovoked	AUSTRALIA	Queensland
7	2017.06.02	2017	Unprovoked	BAHAMAS	New Providence
8	2017.05.30	2017	Provoked	USA	South Carolina
9	2017.05.29	2017	Unprovoked	USA	Florida
10	2017.05.27	2017	Unprovoked	AUSTRALIA	New South Wales
11	2017.05.26.R	2017	Unprovoked	UNITED ARAB EMIRATES	Sharjah
12	2017.05.26.R	2017	Unprovoked	AUSTRALIA	Queensland
13	2017.05.05	2017	Unprovoked	MEXICO	Baja California Sur
14	2017.05.03	2017	Invalid	USA	California
15	2017.04.29.d	2017	Unprovoked	USA	California
16	2017.04.29.c	2017	Unprovoked	SOUTH AFRICA	Western Cape Province
17	2017.04.29.b	2017	Unprovoked	USA	South Carolina
18	2017.04.29.a	2017	Unprovoked	REUNION	Reunion
19	2017.04.27	2017	Unprovoked	NEW ZEALAND	South Island
20	2017.04.26	2017	Invalid	USA	Florida
21	2017.04.22	2017	Unprovoked	ST HELENA, British overseas territory	Ascension Island
22	2017.04.20	2017	Invalid	USA	South Carolina
23	2017.04.17.b	2017	Unprovoked	USA	Florida
24	2017.04.17.a	2017	Unprovoked	AUSTRALIA	Western Australia
25	2017.04.14	2017	Unprovoked	USA	Hawaii
26	2017.04.13	2017	Unprovoked	USA	Florida
27	2017.04.12.b	2017	Unprovoked	USA	Florida
28	2017.04.12.a	2017	Invalid	SOUTH AFRICA	KwaZulu-Natal
29	2017.04.11	2017	Unprovoked	USA	Florida
30	2017.04.10.b	2017	Unprovoked	USA	Florida
31	2017.04.10.a	2017	Unprovoked	USA	Florida
32	2017.04.06	2017	Unprovoked	USA	Florida
33	2017.04.05	2017	Unprovoked	USA	Florida
34	2017.04.02.b	2017	Unprovoked	AUSTRALIA	Queensland
35	2017.04.02.a	2017	Unprovoked	USA	Florida
36	2017.03.27	2017	Unprovoked	USA	Florida
37	2017.03.19	2017	Unprovoked	AUSTRALIA	New South Wales
38	2017.03.18	2017	Unprovoked	USA	California
39	2017.03.09	2017	Unprovoked	BAHAMAS	Great Exuma
40	2017.02.29	2017	Unprovoked	AUSTRALIA	Western Australia
41	2017.02.24	2017	Unprovoked	AUSTRALIA	Western Australia
42	2017.02.21	2017	Unprovoked	REUNION	Saint-André
43	2017.02.17	2017	Unprovoked	AUSTRALIA	Queensland
44	2017.02.11	2017	Unprovoked	USA	Florida
45	2017.02.06.b	2017	Prevoked	BRAZIL	Pernambuco
46	2017.02.06.a	2017	Prevoked	BRAZIL	Kwazulu-Natal
47	2017.02.01.b	2017	Boat	USA	South Carolina
48	2017.02.01.a	2017	Unprovoked	BAHAMAS	BBinimi
49	2017.01.29	2017	Unprovoked	USA	Florida
50	2017.01.21	2017	Unprovoked	AUSTRALIA	Queensland
51	2017.01.13.b	2017	Unprovoked	AUSTRALIA	Queensland
52	2017.01.13.a	2017	Unprovoked	USA	Florida
53	2017.01.09	2017	Unprovoked	INDONESIA	Bali
54	2017.01.08.R	2017	Invalid	AUSTRALIA	Queensland
55	2017.01.05	2017	Unprovoked	USA	Florida
56	2016.12.30	2016	Unprovoked	AUSTRALIA	New South Wales
57	2016.12.27	2016	Unprovoked	USA	Florida
58	2016.12.24	2016	Unprovoked	AUSTRALIA	Western Australia
59	2016.12.19	2016	Unprovoked	SOUTH AFRICA	Western Cape Province
60	2016.12.14	2016	Unprovoked	AUSTRALIA	Western Australia
61	2016.12.11	2016	Unprovoked	USA	Florida
62	2016.12.10	2016	Unprovoked	ENGLAND	England
63	2016.12.06	2016	Prevoked	AUSTRALIA	New South Wales
64	2016.12.01	2016	Unprovoked	AUSTRALIA	New South Wales
65	2016.11.26	2016	Prevoked	AUSTRALIA	Victoria
66	2016.11.14	2016	Unprovoked	USA	Hawaii
67	2016.10.30	2016	Unprovoked	USA	Florida
68	2016.10.29	2016	Unprovoked	USA	Florida
69	2016.10.24	2016	Unprovoked	AUSTRALIA	New South Wales
70	2016.10.21	2016	Unprovoked	USA	Hawaii
71	2016.10.15	2016	Unprovoked	AUSTRALIA	Queensland
72	2016.10.14	2016	Unprovoked	USA	Hawaii
73	2016.10.13.R	2016	Invalid	MEXICO	
74	2016.10.12	2016	Unprovoked	AUSTRALIA	New South Wales
75	2016.10.10	2016	Unprovoked	AUSTRALIA	Oregon
76	2016.10.02	2016	Unprovoked	USA	Florida
77	2016.10.01	2016	Unprovoked	USA	Florida
78	2016.09.26	2016	Unprovoked	AUSTRALIA	New South Wales
79	2016.09.18.c	2016	Unprovoked	USA	Florida
80	2016.09.18.b	2016	Unprovoked	USA	Florida
81	2016.09.18.d	2016	Unprovoked	USA	Florida
82	2016.09.17.b	2016	Unprovoked	USA	California
83	2016.09.17.a	2016	Unprovoked	AUSTRALIA	Victoria
84	2016.09.15	2016	Unprovoked	AUSTRALIA	Victoria
85	2016.09.15.R	2016	Boat	AUSTRALIA	Western Australia
86	2016.09.11	2016	Unprovoked	USA	Florida
87	2016.09.07	2016	Unprovoked	USA	Hawaii
88	2016.09.06	2016	Unprovoked	NEW CALEDONIA	North Province
89	2016.09.05.b	2016	Unprovoked	USA	South Carolina
90	2016.09.05.a	2016	Unprovoked	AUSTRALIA	Western Australia
91	2016.09.04	2016	Unprovoked	USA	Florida
92	2016.09.01	2016	Unprovoked	USA	California
93	2016.08.30	2016	Unprovoked	MEXICO	
94	2016.08.29.b	2016	Unprovoked	USA	Florida
95	2016.08.29.a	2016	Unprovoked	USA	Florida
96	2016.08.27	2016	Unprovoked	REUNION	
97	2016.08.25	2016	Unprovoked	USA	Florida
98	2016.08.07	2016	Unprovoked	BAHAMAS	New Providence Island
99	2016.08.06	2016	Unprovoked	USA	Hawaii
100	2016.08.04	2016	Unprovoked	USA	Florida
101	2016.08.03	2016	Unprovoked	REUNION	Reunion

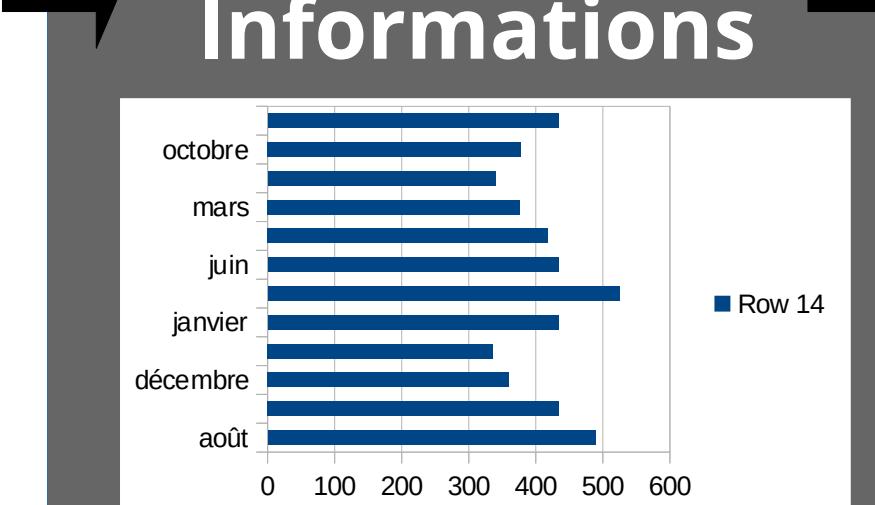
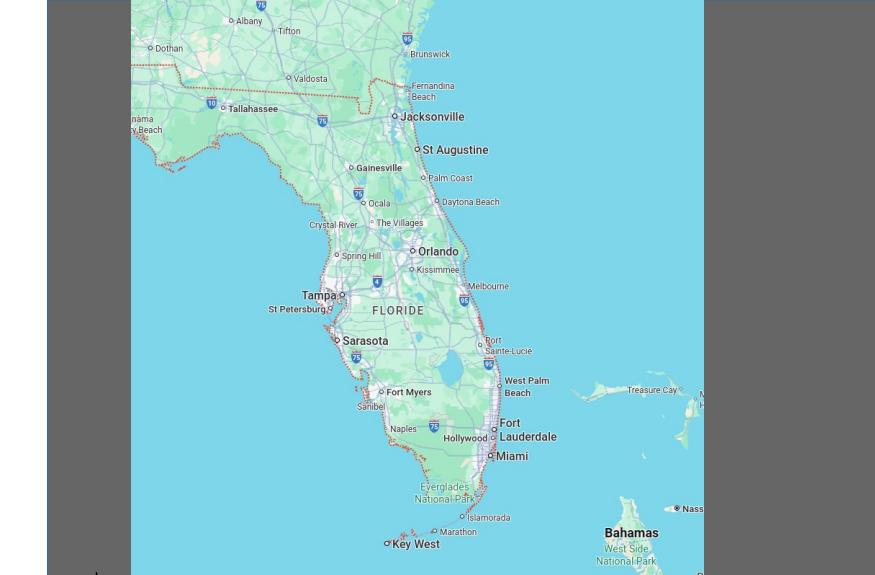


Data

Informations



A	B	C	D	E	
Case Number	Date	Year	Type	Country	Area
1	2017.06.11	2017	Unprovoked	AUSTRALIA	Western Australia
2	2017.06.10.b	2017	Unprovoked	AUSTRALIA	Victoria
3	2017.06.10.a	2017	Unprovoked	USA	Florida
4	2017.06.09.R	2017	Unprovoked	UNITED KINGDOM	South Devon
5	2017.06.07.R	2017	Unprovoked	UNITED KINGDOM	Devon
6	2017.06.06	2017	Unprovoked	MEXICO	Quintana Roo
7	2017.06.02	2017	Unprovoked	BAHAMAS	New Providence
8	2017.05.30	2017	Provoked	USA	South Carolina
9	2017.05.29	2017	Unprovoked	USA	Florida
10	2017.05.27	2017	Unprovoked	AUSTRALIA	New South Wales
11	2017.05.12	2017	Unprovoked	UNITED ARAB EMIRATES	Sharjah
12	2017.05.09.R	2017	Unprovoked	AUSTRALIA	Queensland
13	2017.05.06	2017	Unprovoked	MEXICO	Baja California Sur
14	2017.05.03	2017	Unprovoked	USA	California
15	2017.04.29.d	2017	Unprovoked	USA	California
16	2017.04.29.c	2017	Unprovoked	SOUTH AFRICA	Western Cape Province
17	2017.04.29.b	2017	Unprovoked	USA	South Carolina
18	2017.04.29.a	2017	Unprovoked	REUNION	Reunion
19	2017.04.27	2017	Unprovoked	NEW ZEALAND	South Island
20	2017.04.26	2017	Unprovoked	USA	Florida
21	2017.04.22	2017	Unprovoked	ST HELENA, British overseas territory	Ascension Island
22	2017.04.20	2017	Unprovoked	USA	South Carolina
23	2017.04.17.b	2017	Unprovoked	USA	Florida
24	2017.04.17.a	2017	Unprovoked	AUSTRALIA	Western Australia
25	2017.04.14	2017	Unprovoked	USA	Hawaii
26	2017.04.13	2017	Unprovoked	USA	Florida
27	2017.04.12.b	2017	Unprovoked	USA	Florida
28	2017.04.12.a	2017	Unprovoked	SOUTH AFRICA	KwaZulu-Natal
29	2017.04.11	2017	Unprovoked	USA	Florida
30	2017.04.10.b	2017	Unprovoked	USA	Florida
31	2017.04.10.a	2017	Unprovoked	USA	Florida
32	2017.04.06	2017	Unprovoked	USA	Florida
33	2017.04.05	2017	Unprovoked	USA	Florida
34	2017.04.02.b	2017	Unprovoked	AUSTRALIA	Queensland
35	2017.04.02.a	2017	Unprovoked	USA	Florida
36	2017.03.27	2017	Unprovoked	USA	Florida
37	2017.03.19	2017	Unprovoked	AUSTRALIA	New South Wales
38	2017.03.18	2017	Unprovoked	USA	California
39	2017.03.09	2017	Unprovoked	BAHAMAS	Great Exuma
40	2017.02.29	2017	Unprovoked	AUSTRALIA	Western Australia
41	2017.02.24	2017	Unprovoked	AUSTRALIA	Saint-André
42	2017.02.21	2017	Unprovoked	REUNION	Reunion
43	2017.02.17	2017	Unprovoked	AUSTRALIA	Queensland
44	2017.02.11	2017	Unprovoked	USA	Florida
45	2017.02.06.b	2017	Unprovoked	BRAZIL	Pernambuco
46	2017.02.06.a	2017	Unprovoked	BOA ISLAND	KwaZulu-Natal
47	2017.02.01.b	2017	Unprovoked	USA	South Carolina
48	2017.02.01.a	2017	Unprovoked	BAHAMAS	Exuma
49	2017.01.27	2017	Unprovoked	USA	Queensland
50	2017.01.21	2017	Unprovoked	AUSTRALIA	Queensland
51	2017.01.13.b	2017	Unprovoked	AUSTRALIA	Queensland
52	2017.01.13.a	2017	Unprovoked	USA	Florida
53	2017.01.09	2017	Unprovoked	INDONESIA	Bali
54	2017.01.05	2017	Unprovoked	AUSTRALIA	Queensland
55	2017.01.05	2017	Unprovoked	USA	Florida
56	2017.01.03	2017	Unprovoked	AUSTRALIA	New South Wales
57	2016.12.27	2016	Unprovoked	USA	Florida
58	2016.12.24	2016	Unprovoked	AUSTRALIA	Western Australia
59	2016.12.19	2016	Unprovoked	SOUTH AFRICA	Western Cape Province
60	2016.12.14	2016	Unprovoked	AUSTRALIA	Western Australia
61	2016.12.11	2016	Unprovoked	USA	Florida
62	2016.12.10	2016	Unprovoked	ENGLAND	England
63	2016.12.06	2016	Unprovoked	AUSTRALIA	New South Wales
64	2016.12.01	2016	Unprovoked	AUSTRALIA	New South Wales
65	2016.11.28	2016	Unprovoked	AUSTRALIA	Victoria
66	2016.11.14	2016	Unprovoked	USA	Hawaii
67	2016.10.30	2016	Unprovoked	USA	Florida
68	2016.10.29	2016	Unprovoked	USA	Florida
69	2016.10.24	2016	Unprovoked	AUSTRALIA	New South Wales
70	2016.10.21	2016	Unprovoked	USA	Hawaii
71	2016.10.15	2016	Unprovoked	AUSTRALIA	Queensland
72	2016.10.14	2016	Unprovoked	USA	Hawaii
73	2016.10.13.R	2016	Unprovoked	MEXICO	Boat
74	2016.10.12	2016	Unprovoked	AUSTRALIA	New South Wales
75	2016.10.10	2016	Unprovoked	AUSTRALIA	Oregon
76	2016.10.02	2016	Unprovoked	USA	Florida
77	2016.10.01	2016	Unprovoked	AUSTRALIA	Florida
78	2016.09.26	2016	Unprovoked	AUSTRALIA	New South Wales
79	2016.09.18.c	2016	Unprovoked	USA	Florida
80	2016.09.18.b	2016	Unprovoked	AUSTRALIA	Florida
81	2016.09.18.d	2016	Unprovoked	USA	Florida
82	2016.09.17.b	2016	Unprovoked	USA	California
83	2016.09.17.a	2016	Unprovoked	AUSTRALIA	Victoria
84	2016.09.15	2016	Unprovoked	AUSTRALIA	Victoria
85	2016.09.15.R	2016	Unprovoked	AUSTRALIA	Western Australia
86	2016.09.11	2016	Unprovoked	USA	Florida
87	2016.09.07	2016	Unprovoked	USA	Hawaii
88	2016.09.06	2016	Unprovoked	NEW CALEDONIA	North Province
89	2016.09.05.b	2016	Unprovoked	USA	South Carolina
90	2016.09.05.a	2016	Unprovoked	AUSTRALIA	Western Australia
91	2016.09.04	2016	Unprovoked	USA	Florida
92	2016.09.01	2016	Unprovoked	USA	Florida
93	2016.08.30	2016	Unprovoked	MEXICO	Boat
94	2016.08.29.b	2016	Unprovoked	USA	Florida
95	2016.08.29.a	2016	Unprovoked	USA	Florida
96	2016.08.27	2016	Unprovoked	REUNION	
97	2016.08.25	2016	Unprovoked	USA	Florida
98	2016.08.07	2016	Unprovoked	BAHAMAS	New Providence Island
99	2016.08.06	2016	Unprovoked	USA	Hawaii
100	2016.08.04	2016	Unprovoked	USA	Florida
101	2016.08.03	2016	Unprovoked	REUNION	Réunion
102	2016.08.02	2016	Unprovoked	USA	Florida
103	2016.08.01	2016	Unprovoked	USA	Florida

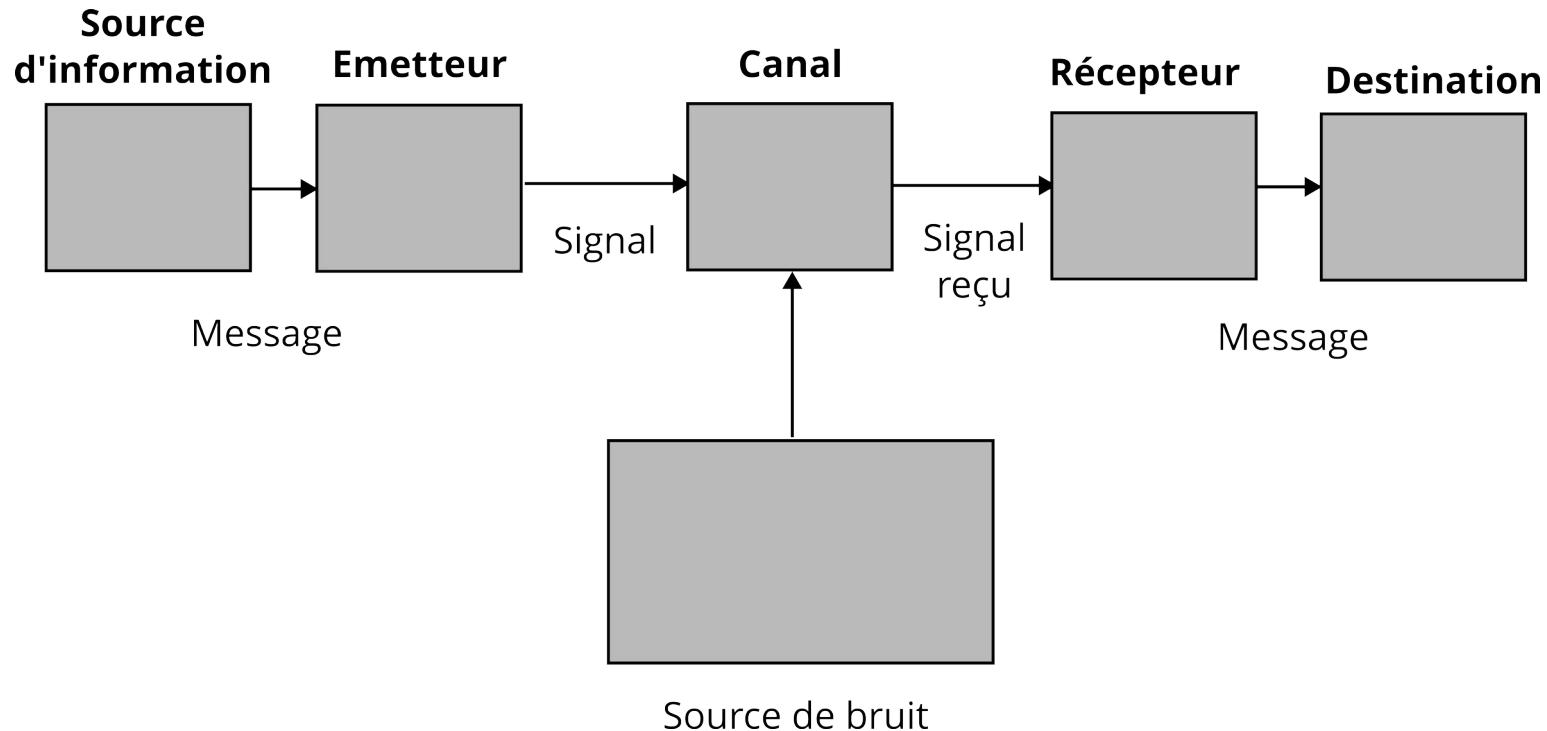


Connaissance

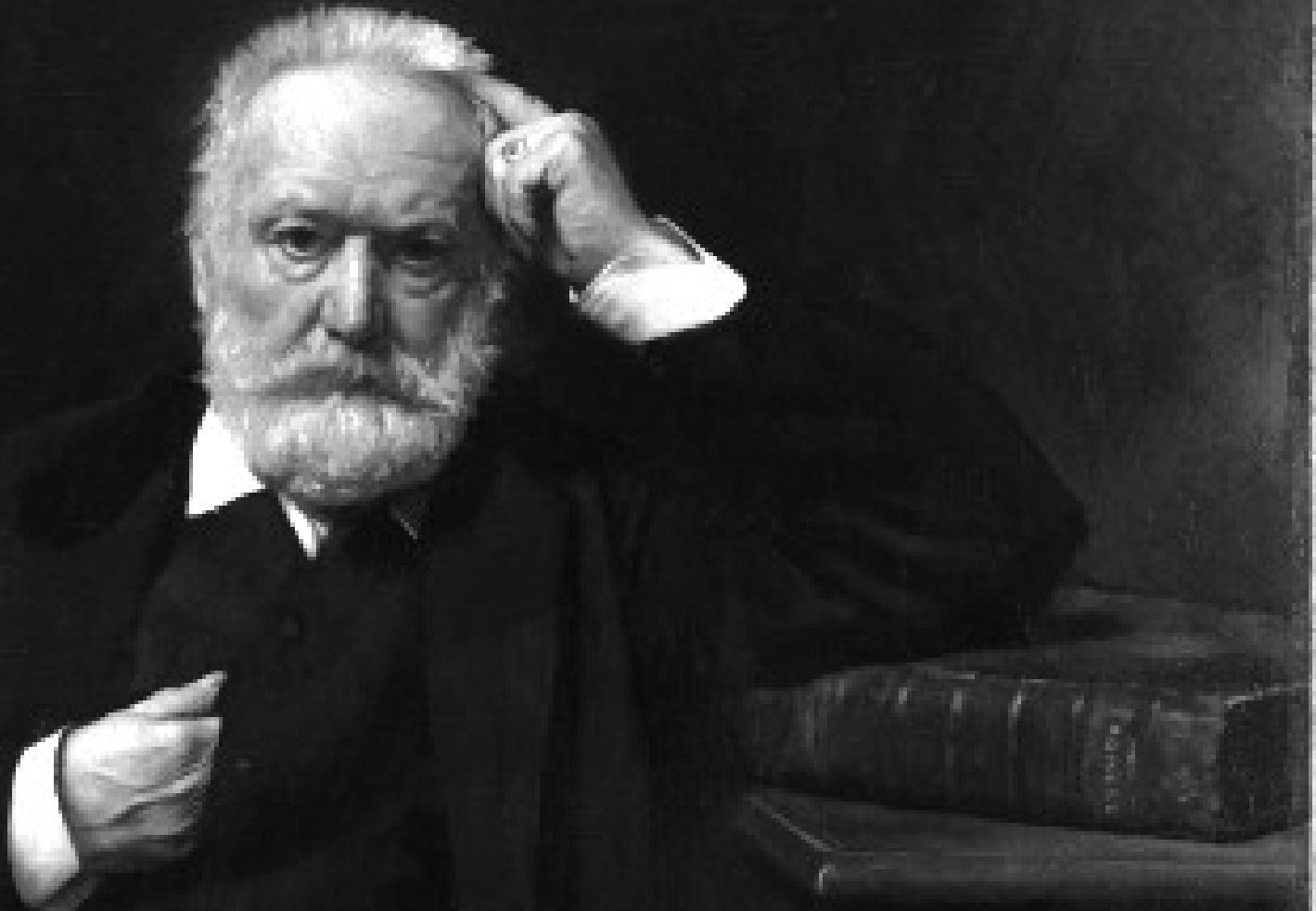
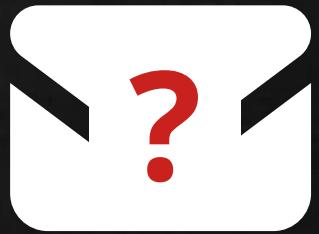
Si vous surfez en octobre dans l'après-midi sur la plage de New Smyrna Beach, en Floride vous avez plus de chance de finir en confit de canard pour requin que de prendre une bonne vague.

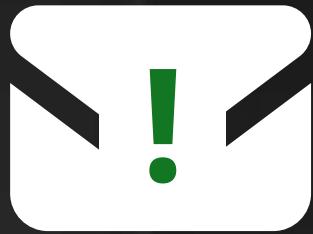
L'information a besoin d'un **canal** de transmission

Theorie de l'information de Claude Shannon

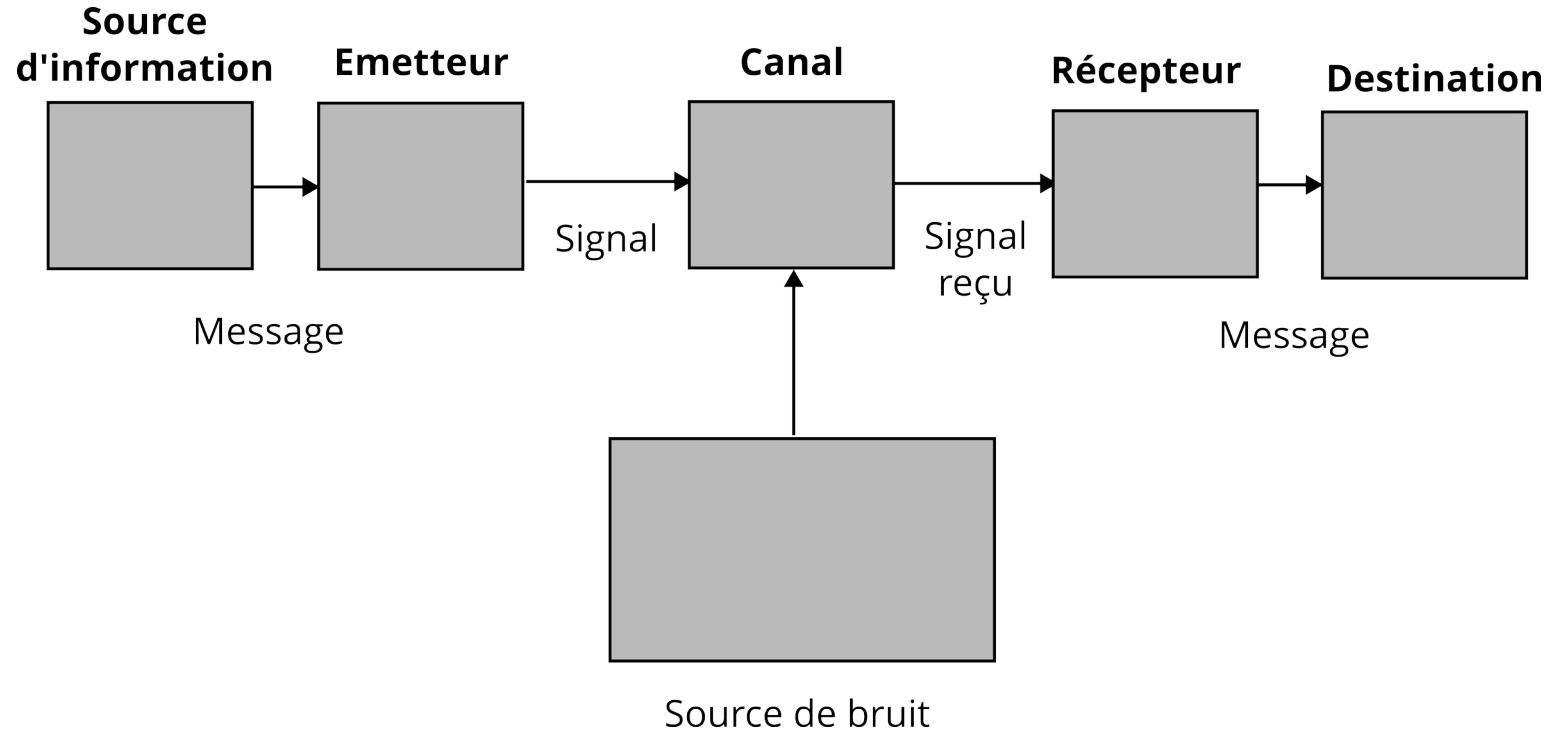


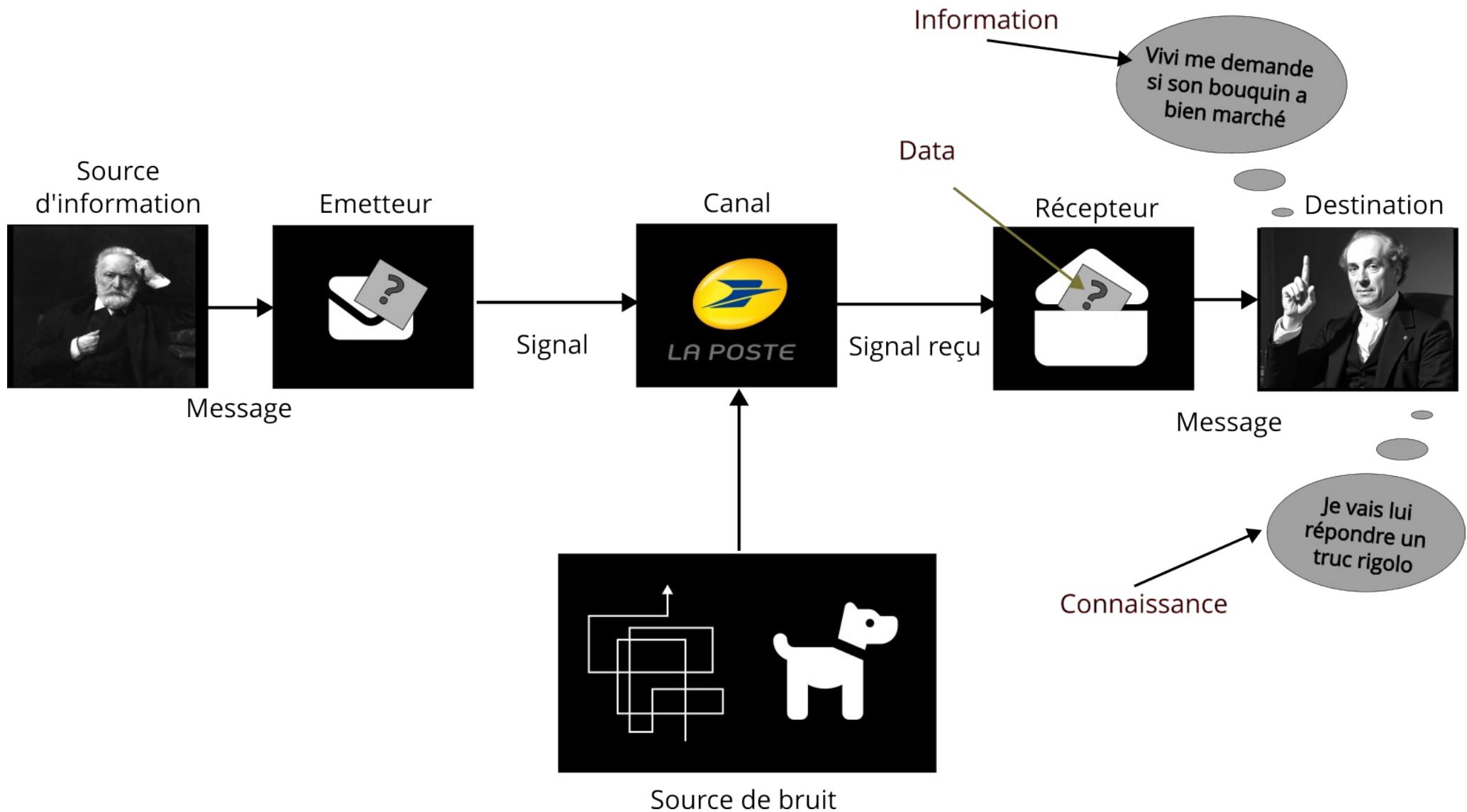
Cas pratique





?







Data Science



Is a branch of **computer science** where we study **how to store, use and analyse data** for deriving **information** from it.

OK

Nettoyage des données

Quelques exemples

- caractères **mal encodés** dans le texte des tweets ;
- **valeurs manquantes** dans certaines colonnes (localisation, fuseau horaire, etc.) ;
- **formats hétérogènes** pour une même information (ex. lieux, États, fuseaux horaires) ;
- informations **non pertinentes** selon la question étudiée.

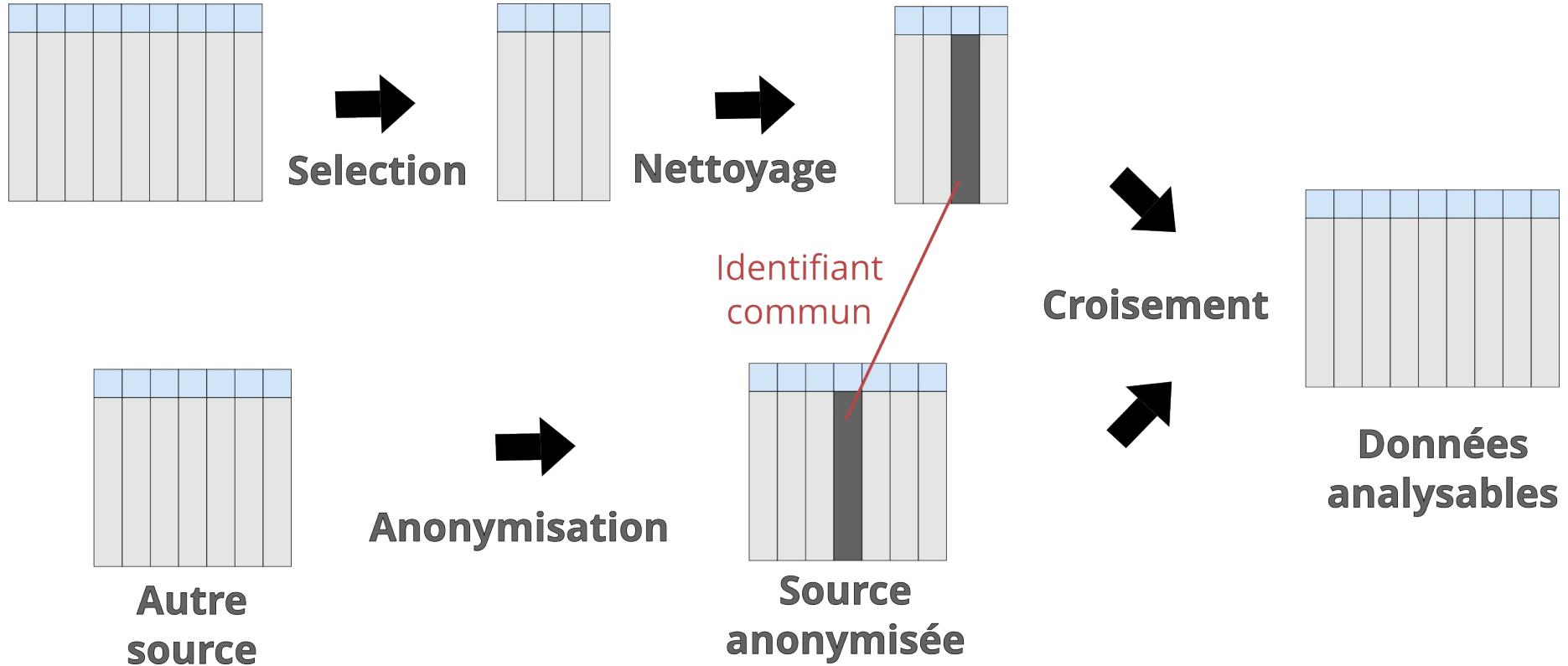
Ce n'est pas une étape **mécanique**,
mais une étape de **réflexion**

→ **il faut rendre les données exploitables**

Selon la question d'analyse retenue, certaines variables peuvent être conservées, transformées ou écartées.

→ **Toute décision est acceptable à condition d'être cohérente avec l'objectif de l'analyse et explicitement justifiée.**

Nettoyage des données



ATLANTA

Casse incohérente

Orlando, Florida-Sunshine
State

Texte parasite

%Û¢detroit%Û¢

Problème d'encodage

mn

Abréviation ambiguë

In service to others – San
Diego

Texte non géographique

Brookhaven/Chester
Heights

Plusieurs villes

ATLANTA

Orlando, Florida-Sunshine
State

%^odetroit%

mn

In service to others – San
Diego

Brookhaven/Chester
Heights

Atlanta

Orlando

Detroit

Minnesota

San Diego

Brookhaven

Les outils



- Créer en **1991**
- Version **3.14** (octobre 2025)



```
In [1]: print("Hello, World!")
```

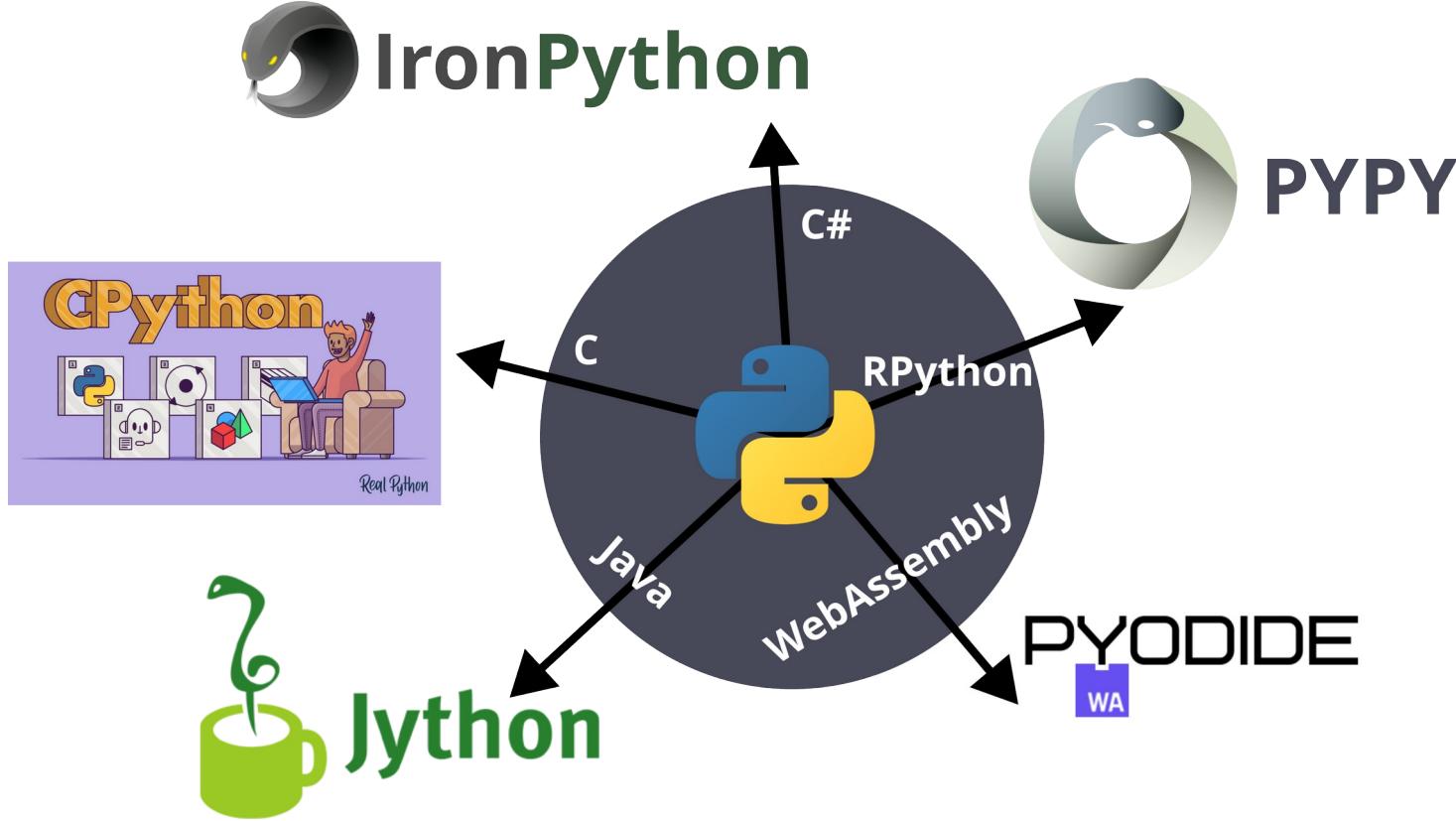
```
Out [1]: Hello, World!
```



Python

Plein d'implémentations !

- Compilateur + Interpréteur + Runtime



Python

= 2 façons de coder



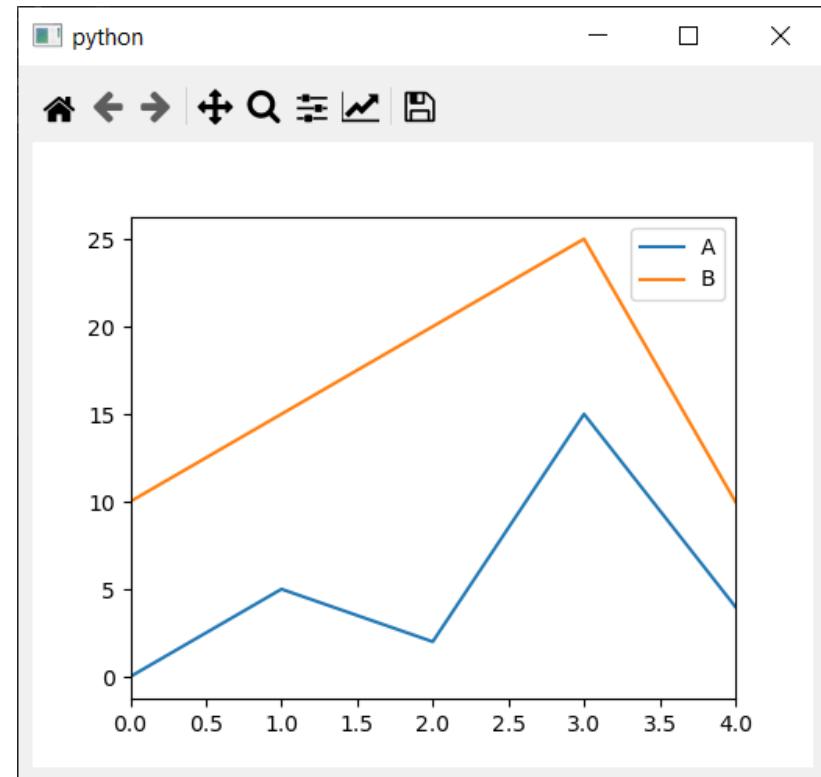
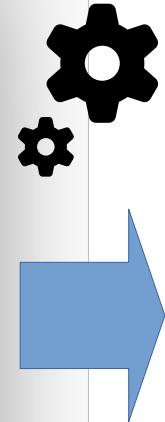
```
# Python program to swap two variables

x = 5
y = 10

# To take inputs from the user
#x = input('Enter value of x: ')
#y = input('Enter value of y: ')

# create a temporary variable and swap the values
temp = x
x = y
y = temp

print('The value of x after swapping: {}'.format(x))
print('The value of y after swapping: {}'.format(y))
```



I

Python

Interactive Python

- ⚡ Auto-complétion
- 🧠 Historique
- pip install ipython



```
In [1]: a = 5
```

```
In [2]: b = a
```

```
In [3]: b  
Out[3]: 5
```

quarto-jupyter - JupyterLab

localhost:8888/lab/tree/quarto-jupyterlab.ipynb

File Edit View Run Kernel Tabs Settings Help

quarto-jupyterlab.ipynb

Code Python 3 (ipykernel)

```
---
```

```
title: Matplotlib Demo
author: Norah Smith
date: 'May 22nd, 2021'
format:
  html:
    code-fold: true
keep-ipynb: true
---
```

Polar Axis

For a demonstration of a line plot on a polar axis, see @fig-polar.

```
[2]: #| label: fig-polar
#| fig.cap: A line plot on a polar axis

import numpy as np
import matplotlib.pyplot as plt

r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(subplot_kw={'projection': 'polar'})
ax.plot(theta, r)
ax.set_rticks([0.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

90°
135°
45°
0

Terminal 1

```
author: Norah Smith
date: 'May 22nd, 2021'

Output created: quarto-jupyterlab.html

Watching files for changes
```

Simple Python 3 (ipykernel) | Idle Mode: Edit Ln 3, Col 1 quarto-jupyterlab.ipynb



Déroulé du cours

2 TP non notés

mais à rendre tout de
même

499

Résolutions du nouvelle
an 2015



tweet_created	tweet_text	tweet_category	tweet_topics	Tweet_location	tweet_state	tweet_region
2014-12-21 16:11:00	#NewYearsResolution to not put the parking lot ticket directly in my mouth right after getting it from the machine #ew	Humor	Humor about Personal Growth and Interests Resolutions	City of Angels, CA	CA	West
2014-12-21 16:16:00	I got the hair. I just need to work on the body to pull off Lara Croft next Halloween #NewYearsResolution http://t.co/Ihqh5Z7prx	Health & Fitness	Improve my body	Arlington, TX	TX	South
2014-12-21 16:30:00	Self improvement!.. Mentally, physically, and financially. #NewYearsResolution	Personal Growth	Improve my attitude	Las Vegas, NV	NV	West
2014-12-21 17:51:00	I've decided to get pregnant in #2015 #NewYearsResolution	Humor	Humor about Family/Friends/Relationships Resolutions	Ft. Luderdale, Florida	FL	South
2014-12-21 18:09:00	All I want for Christmas is to be Carrie Underwood. #NewYearsResolution	Personal Growth	Other	South Carolina	SC	South

2 TP non notés

mais à rendre tout de même

499

Résolutions du nouvelle
an 2015

TP1 → Nettoyage du jeu de données

TP2 → Analyse exploratoire

Monôme ou binôme



Et **2** projets

Et pas d'examens

Projet 1 = **1/2** note finale

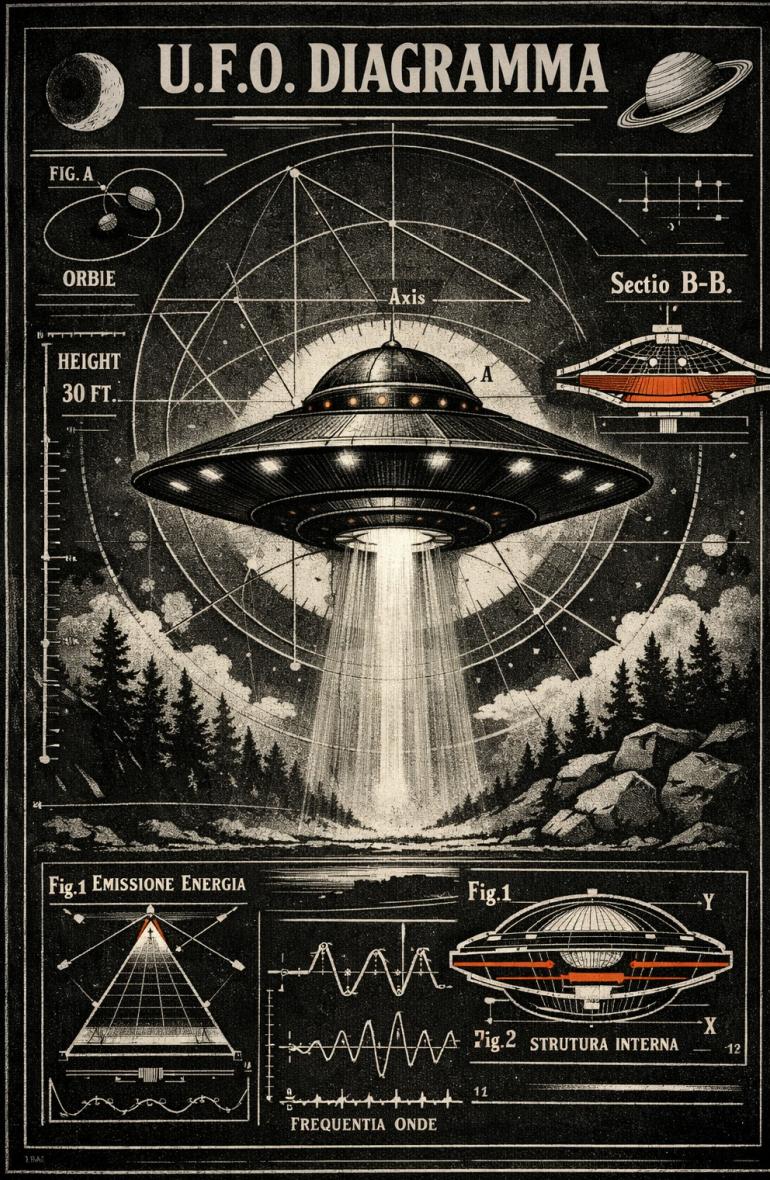
Objectif : **Analyser une base de données**
& Produire une connaissance

4 pages MAX

A rendre pour le **28 Février 23h59** au plus tard

80 000+ enregistrements de témoignages d'ovni

- Quelles formes apparaissent le plus souvent ?
- Il y a t'il de plus en plus d'apparitions ?



Projet 2 = 1/2 note finale

Objectif : Répondre à une question & Produire une connaissance

Quels facteurs influencent la propagation d'une maladie comme le COVID-19 dans un pays ?

Quels facteurs influencent le prix des actions d'une entreprise sur une période donnée ?

Quels hashtags génèrent le plus d'engagement sur Twitter pour un événement donné ?

Comment les températures évoluent-elles selon les villes et saisons ?

/!\ A vous de choisir /!

Projet 2 = **1/2** note finale

Objectif : **Répondre à une question**
& Produire une connaissance

Rapport à rendre pour le **31 Mars 23h59 au plus tard**
Avec les bases de données, sources & codes

Projet 2 = 1/2 note finale

Objectif : **Répondre à une question**
& Produire une connaissance

Présentations orales du **16 au 27 mars** sur 3 séances

- **Organisez vous dès maintenant**
- **+1 point pour ceux qui passeront la première séance**

	Nombre de pages	Présentation
Monôme	5	5min
Binôme	7-8	7min
Trinôme	9-10	10min

Critère	Excellent	Bien	Passable	Insuffisant	Pondération
Maîtrise Python & coding	Code structuré, modulaire, efficace	Code clair et fonctionnel	Code fonctionnel mais peu structuré	Code non fonctionnel / incompréhensible	20 %
Méthodes Big Data	Choix pertinents, pipeline efficace	Méthodes adaptées, résultats corrects	Méthodes limitées ou peu efficaces	Méthodes incorrectes ou manquantes	20 %
Analyse & Interprétation	Analyse complète, conclusions argumentées, insights pertinents	Analyse correcte, conclusions correctes	Analyse partielle, conclusions limitées	Analyse absente ou incorrecte	20 %
Présentation orale	Présentation claire, engageante, répond aux questions	Présentation claire, répond aux questions	Présentation peu claire ou incomplète	Présentation absente ou incompréhensible	20 %
Rendu écrit / rapport	Rapport clair, structuré, professionnel	Rapport correct	Rapport incomplet ou peu clair	Rapport absent / illisible	20 %

Big Data - CM1

C'est tout pour aujourd'hui

