

Festival Online de la Data



PROCHAINES DATES

Jeudi 04 Février de 15h à 15h30

Automatisation et pipeline CI/CD

Glenn CARROY, Développeur Web, Synaltic

Fabien OGER, Développeur Web, Synaltic

Jeudi 18 Février de 15h à 15h30

Intégrer du data storytelling dans votre stack décisionnelle

Ludovic Narayanin, DATAROCKS

Charly Clairmont, Synaltic

Jeudi 11 Février de 15h à 15h30

Intégrer rapidement vos données dans Google Cloud

Victor Coustenoble, Trifacta

Charly Clairmont, Synaltic

Jeudi 25 Février de 15h à 15h30

Tableau + R

John BONTIT, Data Analyst, Synaltic

SYNALTIC EN QUELQUES MOTS

Acteur innovant, dénicheur de solution et une équipe de collaborateurs engagés



30

30 collaborateurs formés et certifiés contribuant aux communautés d'utilisateurs.

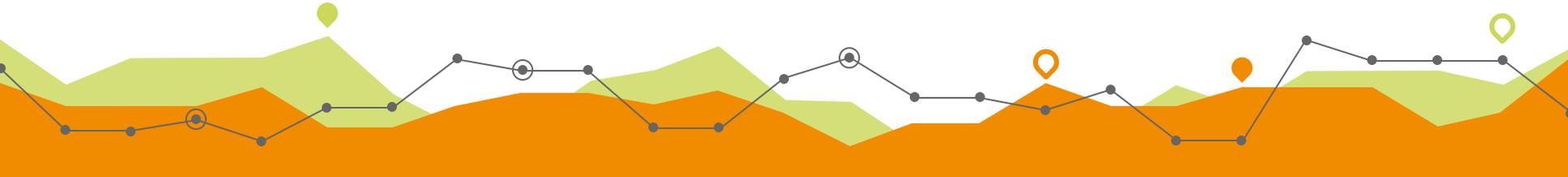
15

Spécialiste en Data Management depuis plus de 15 ans et plus de **250 projets** réalisés.

160

Plus de **160 clients** dont certains s'appuient sur nous avec succès depuis plus de **8 ans**.

Experts en Data Management, passionnés d'Open Source et d'Innovation !



La Data Littératie, l'atout clé des entreprises du 21ème siècle !



Charly CLAIRMONT

CTO, Synaltic

@egwada

#OpenSource #BI #Hadoop #DataLake #Analytics #Innovation

1999 : développeur web, site d'informations sportives à large audience

2004 : création de Altic qui deviendra Synaltic

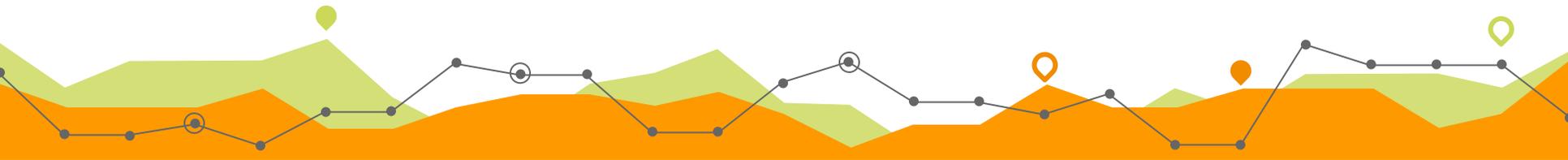
2011 : enseignant universitaire (Paris 6, Paris 13, UTT)

2012 : création du Hadoop User Group / Paris Spark Meetup

2019 : Synaltic une trentaine d'employés
Hug France : +1500 membres
Paris Spark Meetup : +3000 membres

“La donnée” est partout !

2020 consacre la donnée



Tous les français parlent Data ! n'est-ce pas ?

Situation épidémique

36 437 Nouveaux cas depuis 24h

58% Lits occupés en réanimation

35 785 Décès depuis le début de l'épidémie

Covid 19, nous avons aussi dénombré 3000 plus de la moitié des capacités nationales.

RAPPEL DE NOTRE STRATÉGIE SCÉNARIO AU 28 MARS 2020

SANS LE CONFINEMENT

AVEC LE CONFINEMENT

Édouard Philippe Premier ministre

déconfinement, les précisions d'Édouard Philippe et Olivier Véran

Situation épidé

4 900 patients en réanimation le 16 novembre 2020

4 300 patients en réanimation aujourd'hui

nts hospitalisés le 16 novembre, nous avons e décré. Le nombre de personnes en réa

GARANTIR DES MASQUES POUR LES SOIGNANTS

STOCK 117

BESOIN HEBDOMADAIRE MINIMUM = 45 MILLIONS

Période	Production nationale	Importations
Janvier	5	7
Du 30 Mars au 5 Avril	3	34
Du 6 au 13 Avril	1	63
Du 14 au 19 Avril	1	81

Production nationale (rouge), Importations (bleu)

GARANTIR DES MASQUES POUR LES SOIGNANTS

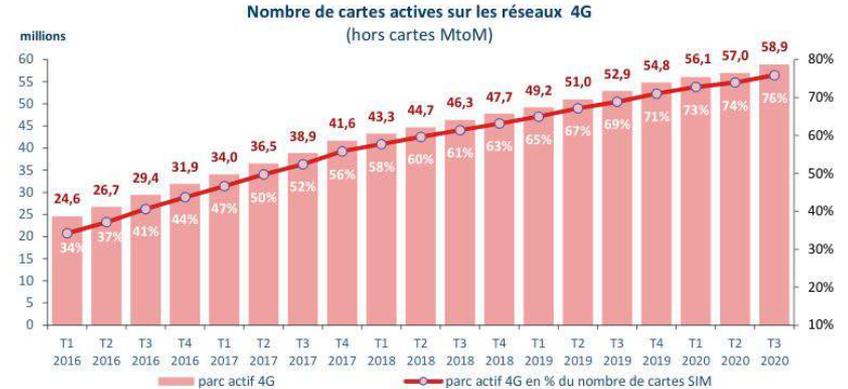
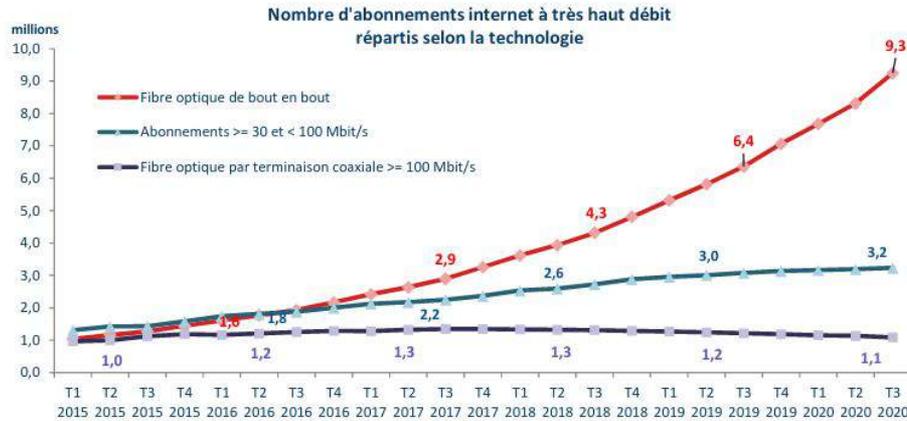
STOCK 117

BESOIN HEBDOMADAIRE MINIMUM = 45 MILLIONS

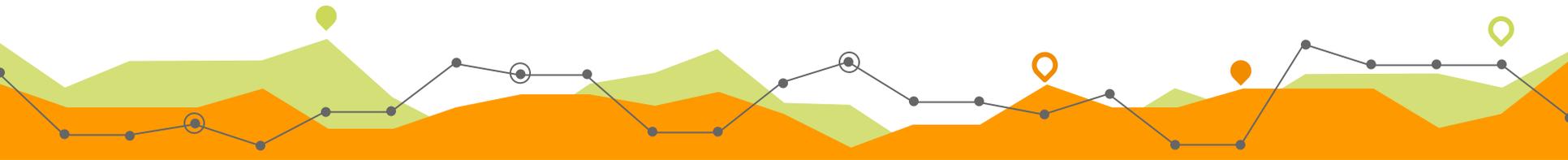
Période	Production nationale	Importations
Janvier	5	7
Du 30 Mars au 5 Avril	3	34
Du 6 au 13 Avril	1	63
Du 14 au 19 Avril	1	81

Production nationale (rouge), Importations (bleu)

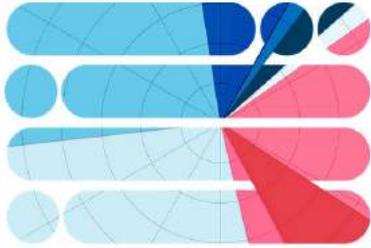
Et nous sommes toujours tous plus nombreux à être connectés !



https://www.arcep.fr/fileadmin/cru-1611251965/reprise/observatoire/3-2020/obs-marches-t3-2020_140121.pdf



Dans les entreprises, parlons-nous Data ?

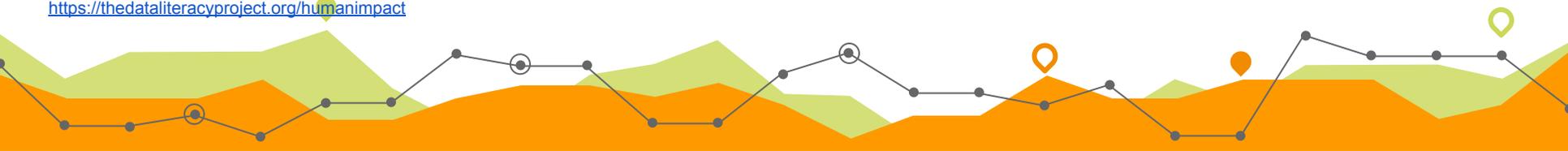


"Nous devons restaurer la confiance des individus et leur montrer qu'ils ne devraient pas avoir peur des données"

Emma Alexander, Samsung Electronics

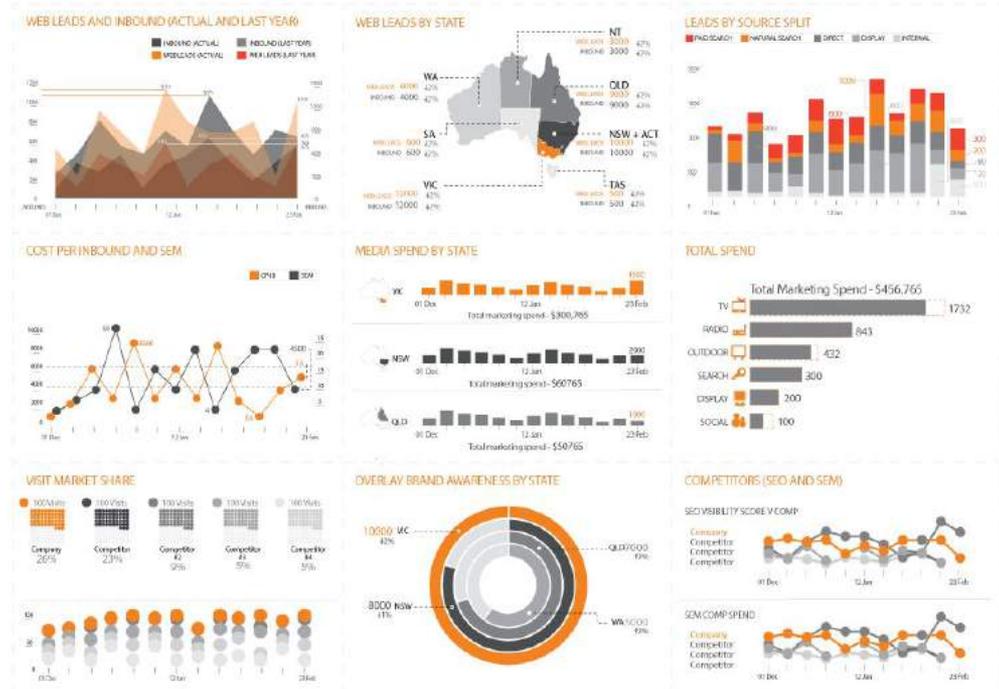


<https://thedataliteracyproject.org/humanimpact>



La Data Literacy

Je m'arrête un instant pour que vous intégrez ce graphique et les chiffres associés !



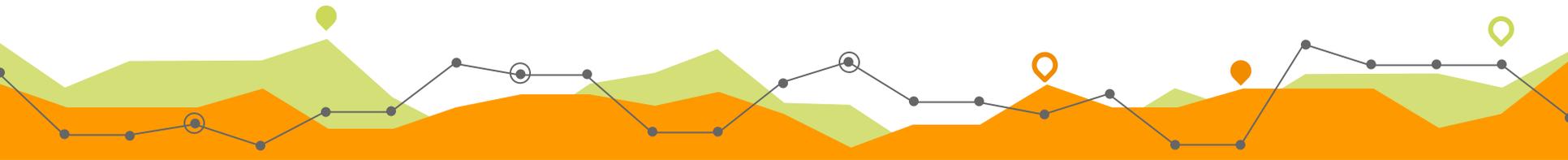
Littératie des données



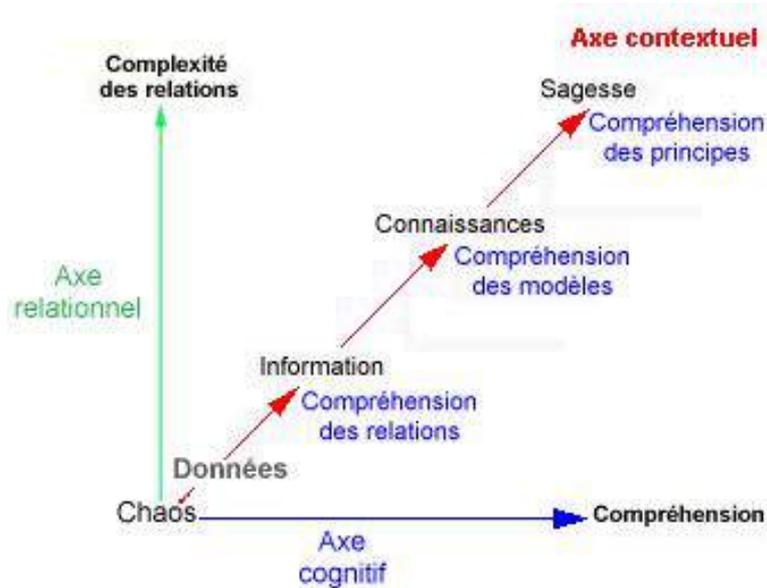
WIKIPEDIA

La **littératie des données** (du terme anglais *data literacy*) ou **culture des données** est la capacité d'**identifier**, de **collecter**, de **traiter**, d'**analyser** et d'**interpréter** des données afin de **comprendre** les phénomènes, les processus, les comportements qui les ont générées en sachant faire preuve d'esprit critique.

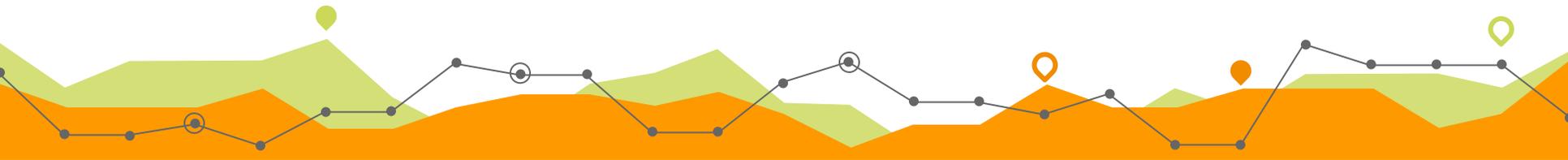
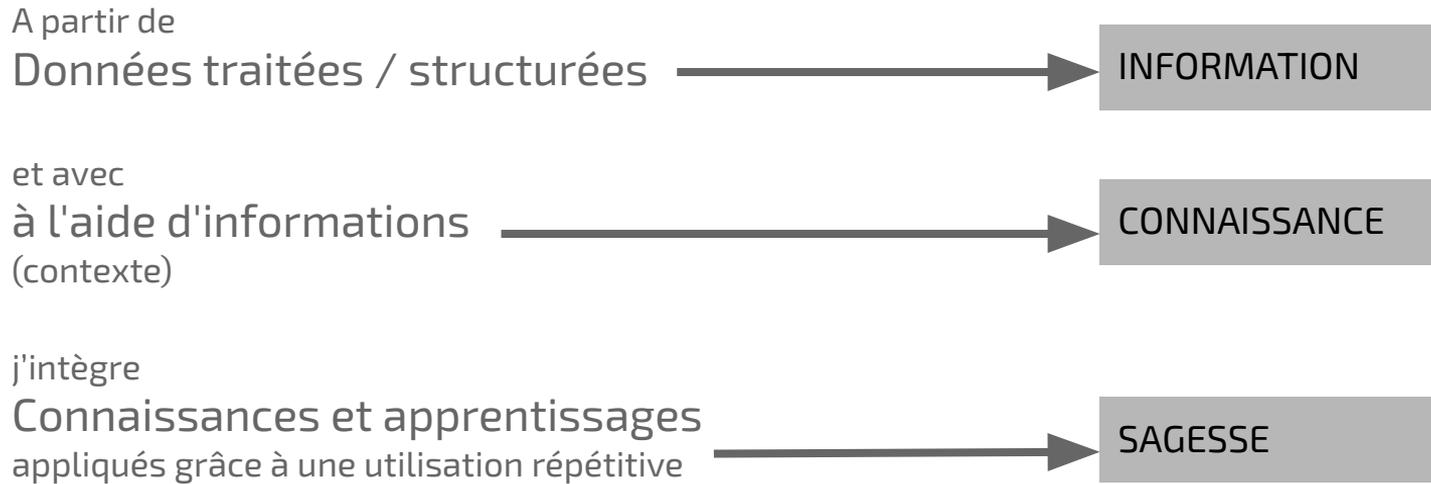
https://fr.wikipedia.org/wiki/Litt%C3%A9ratie_des_donn%C3%A9es



Le spectre du savoir : données, information, connaissances et sagesse (1)

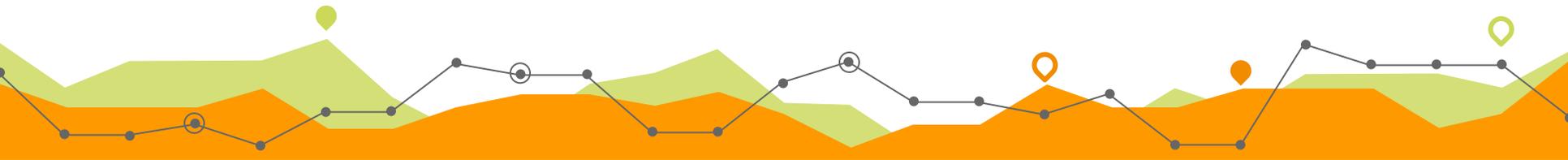


Le spectre du savoir : données, information, connaissances et sagesse (2)



« Plus les employés savent lire, écrire, analyser et argumenter avec les données, plus leur contribution sera significative, ce qui aura un impact positif sur le futur de leur organisation. »

Bernard Marr, Data Strategist et auteur à succès



Quelles données ?

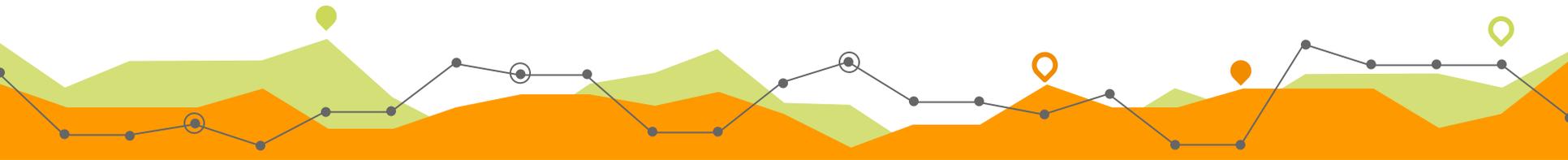
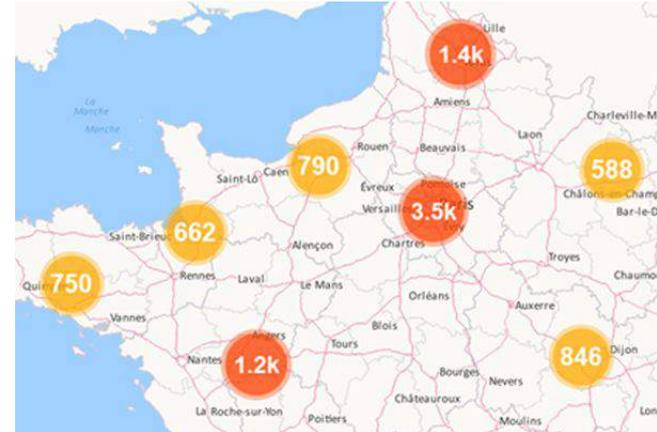
Temporelles



1. En USD-PPA constant de 2010.

Source : OCDE (2019), *Statistiques de l'OCDE des comptes nationaux* (base de données).

Géographique



Quelles données ?

Quels types de données ?

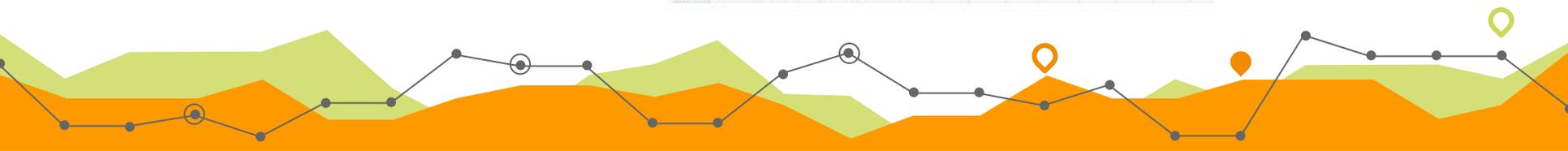
QUALITATIVE

QUANTITATIVE



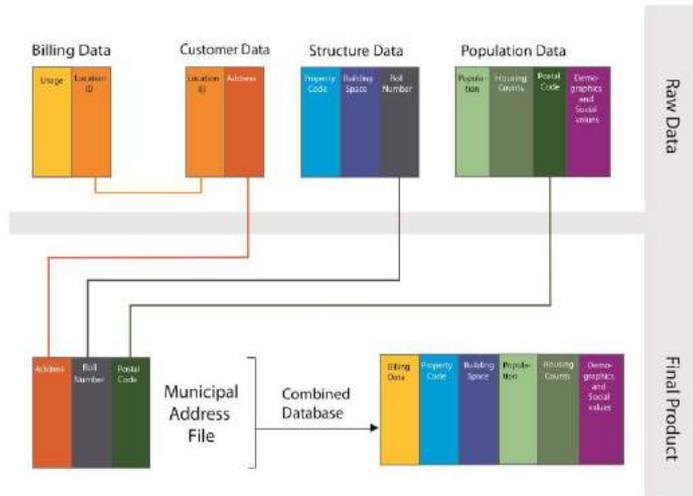
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Average
1950	48.88	48.49	51.87	51.74	51.8	51.38	51.58	51.93	51.2	51.24	51.13	51.48	51.89
1951	51.44	50.39	54.71	51.72	51.89	51.18	51.34	51.4	51.4	51.04	51.18	50.9	51.56
1952	51.42	50.47	53.96	51.62	52.53	51.89	51.18	51.92	51.18	51.04	51.5	51.24	51.41
1953	51.48	51.37	54.56	51.64	52.58	51.87	51.47	51.74	51.4	51.58	50.1	51.81	51.54
1954	51.34	50.56	51.88	52.43	51.51	50.55	51.58	51.74	51.45	51.42	51.44	51.7	51.45
1955	51.54	50.58	54.86	52.39	52.24	52.47	51.74	51.77	51.25	51.84	51.72	51.81	51.86
1956	51.57	49.99	49.99	49.99	52.24	51.88	51.44	51.82	51.67	51.67	51.79	51.79	51.79
1957	51.44	50.44	53.89	52.13	52.15	51.87	51.38	51.89	51.76	51.73	51.87	51.84	51.84
1958	52.42	51.59	52.39	52.87	54.51	52.75	52.28	52.37	51.64	51.1	51.79	52.88	52.26
1959	52.86	51.5	53.04	53.42	53.5	53.09	52.55	52.92	51.51	52.72	52.91	52.91	52.26
1960	52.57	52.15	52.49	52.66	52.57	52.56	51.14	51.03	50.42	51.45	52.31	52.36	52.31
1961	52.8	50.42	52.04	52.66	52.34	50.59	52.58	52.65	52.32	52.78	52.87	52.12	52.13
1962	52.83	50.39	53.47	52.15	52.67	52.66	51.38	51.49	52.1	52.18	52.98	52.12	52.58
1963	52.12	52.68	52.78	52.75	52.92	52.92	52.57	52.54	52.94	52.57	52.8	52.81	52.82
1964	52.71	51.85	52.75	52.72	53.07	52.93	52.88	52.32	51.43	52.58	52.5	52.75	52.74
1965	52.95	52.96	53.5	53.5	53.44	53.07	53.05	53.0	52.7	52.71	52.71	52.84	52.88
1966	52.95	52.61	53.84	52.45	53.81	53.29	53.18	52.4	52.41	52.41	52.41	52.41	52.41
1967	52.85	53.11	54.64	53.11	53.94	53.4	53.4	53.4	53.4	53.4	53.4	53.4	53.4
1968	53.19	53.66	53.5	53.96	54.87	54.18	53.08	53.84	52.4	52.84	53.0	53.0	53.0
1969	53.83	53.42	53.17	53.42	53.18	52.27	53.92	53.2	53.14	53.24	53.18	53.18	53.18
1970	53.49	53.19	53.4	53.35	53.41	53.78	53.4	53.48	53.76	53.2	53.1	53.1	53.1
1971	53.21	53.36	53.46	53.89	53.47	53.84	53.7	53.68	53.81	53.86	53.8	53.8	53.8
1972	53.41	53.38	54.09	54.07	54.15	54.17	53.58	53.5	53.48	53.62	53.7	53.7	53.7
1973	53.23	53.47	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18
1974	54.25	54.11	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27
1975	54.15	54.11	54.64	54.11	54.94	54.4	54.4	54.4	54.4	54.4	54.4	54.4	54.4
1976	54.19	53.66	53.5	53.96	54.87	54.18	53.08	53.84	52.4	52.84	53.0	53.0	53.0
1977	54.83	53.42	53.17	53.42	53.18	52.27	53.92	53.2	53.14	53.24	53.18	53.18	53.18
1978	54.49	53.19	53.4	53.35	53.41	53.78	53.4	53.48	53.76	53.2	53.1	53.1	53.1
1979	54.21	53.36	53.46	53.89	53.47	53.84	53.7	53.68	53.81	53.86	53.8	53.8	53.8
1980	54.41	53.38	54.09	54.07	54.15	54.17	53.58	53.5	53.48	53.62	53.7	53.7	53.7
1981	54.23	54.07	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18	54.18
1982	54.25	54.11	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27	54.27
1983	54.12	54.02	54.11	54.11	54.11	54.11	54.11	54.11	54.11	54.11	54.11	54.11	54.11
1984	54.17	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15
1985	54.17	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15	54.15
1986	54.18	54.06	54.06	54.06	54.06	54.06	54.06	54.06	54.06	54.06	54.06	54.06	54.06
1987	54.82	54.47	54.42	54.42	54.42	54.42	54.42	54.42	54.42	54.42	54.42	54.42	54.42
1988	54.42	54.19	54.19	54.19	54.19	54.19	54.19	54.19	54.19	54.19	54.19	54.19	54.19
1989	54.76	54.17	54.17	54.17	54.17	54.17	54.17	54.17	54.17	54.17	54.17	54.17	54.17
1990	54.86	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7
1991	54.72	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7
1992	54.86	54.72	54.72	54.72	54.72	54.72	54.72	54.72	54.72	54.72	54.72	54.72	54.72
1993	54.81	54.71	54.71	54.71	54.71	54.71	54.71	54.71	54.71	54.71	54.71	54.71	54.71
1994	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
1995	54.87	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7	54.7
1996	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
1997	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
1998	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
1999	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2000	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2001	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2002	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2003	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2004	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2005	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2006	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2007	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81
2008	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81	54.81

Years show highest monthly consumption adjusted to equivalent 2005 hours in the 12th day of each month, and are parts per million by volume (ppmv) measured in the 2005-2006 seasonal cycle. Figure shows the "total average" is the arithmetic mean of the highest 12 monthly values where monthly values are missing.



Où est-elle ?

Data Integration



Data Catalog

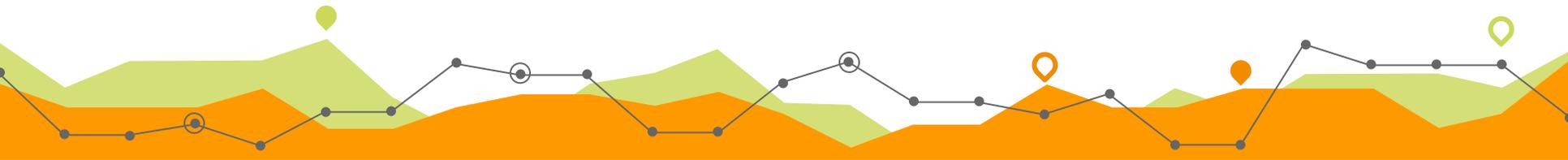
Rechercher 	Tracer les changements 	Construire un réseau agile 
Où est la table, le tableau de bord ? Quel en est le contenu ?	J'ai changé un schéma d'une table qui est le responsable et quels sont les utilisateurs ?	Je voudrais suivre un utilisateur clé de mon équipe
Est-ce que cette analyse existe déjà ?	La livraison du jeu de données est en retard. Je voudrais notifier tout le monde	Je voudrais mettre en favoris une table et être notifié des modifications de schémas, Des nouvelles données, des incidents.

Les analystes et data scientists passe 1/3 de leur temps à découvrir les données.

Eviter le "select **" : trop de ressources utilisées, trop de colonnes à comprendre

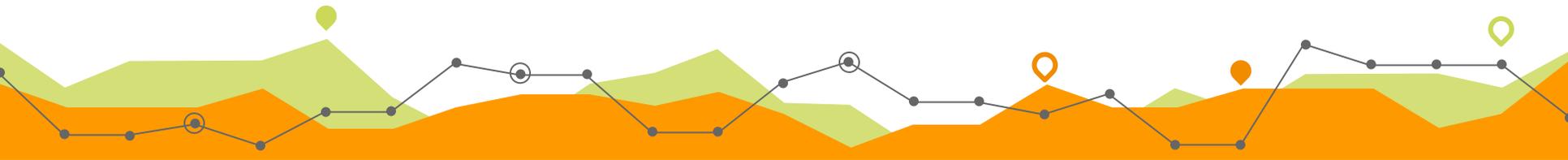
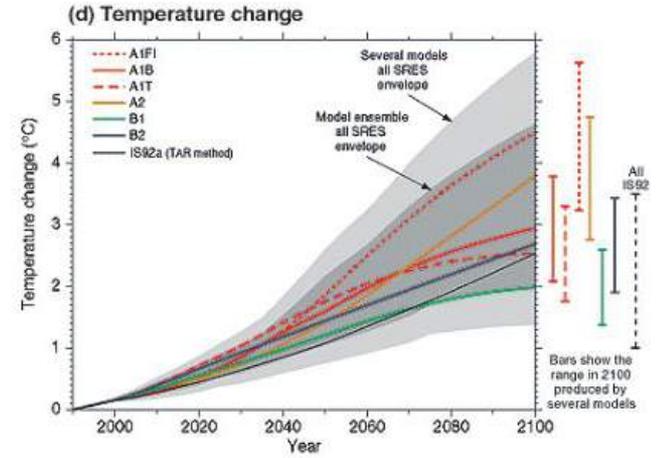
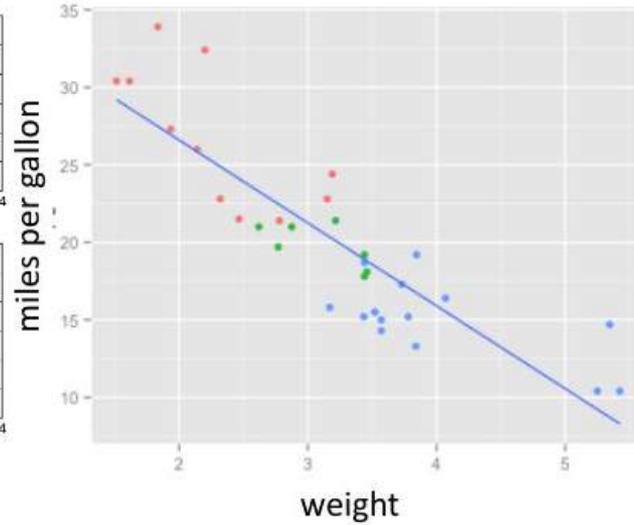
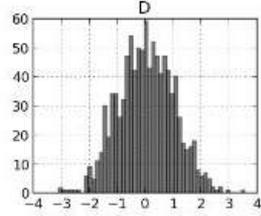
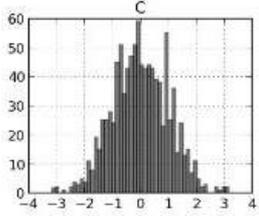
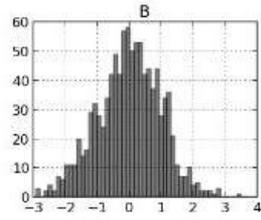
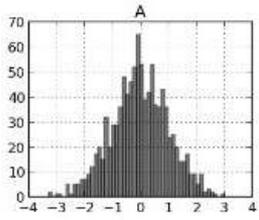
Efficacité à trouver les jeux de données utiles aux analyses.

= Augmenter la productivité



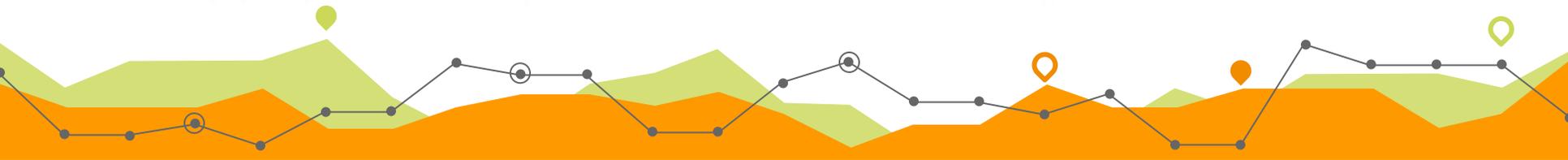
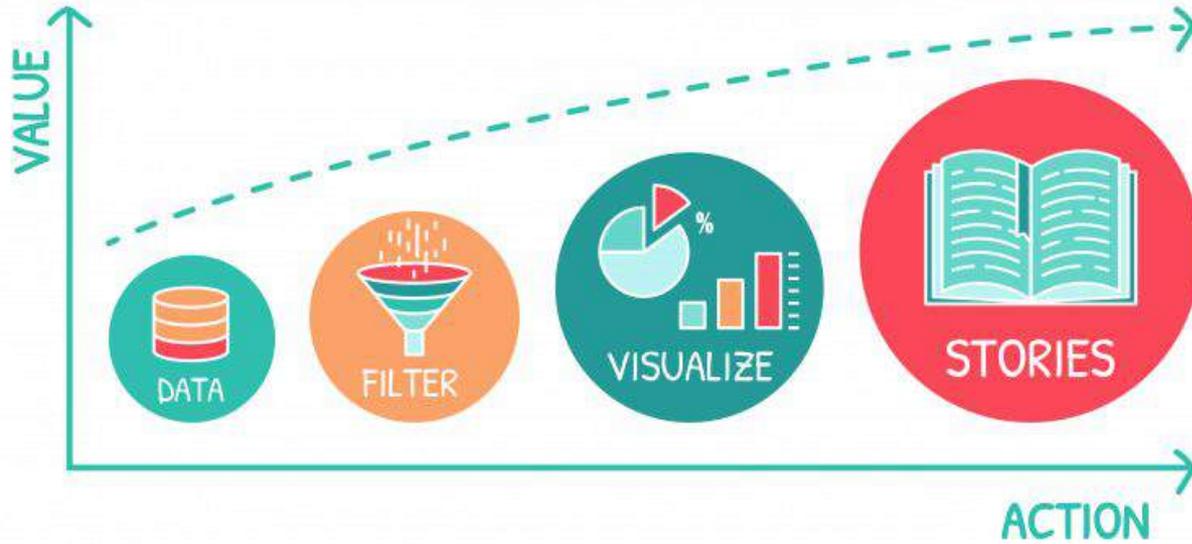
Qu'en déduisons ?

Data Analysis



Comment diffuse-t-on la donnée ?

Data Storytelling



La culture des données, levier de la transformation numérique des organisations

“Inclure” chaque membre de l’organisation !



10
LES COMMANDEMENTS
DE L'EXPERIMENTATION DATA

LES 10 COMMANDEMENTS DE L'EXPERIMENTATION DATA SONT UN TRAVAIL COLLECTIF DE LA COMMUNITE D'INNOVATION DE LA PAFI AU ZENITH, AVEC LA PARTICIPATION DE TOUTES LES MAIRIES DE LA REGION. ILS SONT LE FRUIT D'UN TRAVAIL COLLECTIF ET DE LA PARTICIPATION DE TOUTES LES MAIRIES DE LA REGION. ILS SONT LE FRUIT D'UN TRAVAIL COLLECTIF ET DE LA PARTICIPATION DE TOUTES LES MAIRIES DE LA REGION.

Retrouvez ce document en ligne :

www.mairie.fr/legislation

- 1. À APPRENDRE, PAS FORCÉMENT À RÉUSSIR, TU CHERCHERAS**
- 2. DES HYPOTHÈSES TU FORMULERAS OU L'EXPLORATION TU ASSUMERAS**
- 3. TES RÉUSSITES COMME TES ÉCHECS TU PARTAGERAS**
- 4. À LA DIVERSITÉ DES PROFILS ET CONTRIBUTEURS TU VEILLERAS**
- 5. FRUGALITÉ, AGILITÉ, SIMPLICITÉ TU CHERIRAS**
- 6. UN ACCÈS AUX DONNÉES TU OBTIENDRAS**
- 7. DE L'INTÉRÊT DES DONNÉES TU NE PRÉJUGERAS PAS**
- 8. DANS DES UNIVERS NOUVEAUX, LES DONNÉES TU CHERCHERAS**
- 9. FACE AUX DONNÉES, UN ESPRIT CRITIQUE TU GARDERAS**
- 10. À CES PRINCIPES TOUT LE MONDE ADHÉRERA**

#SPECIFIQUE DATA

17 caractéristiques clés de la maîtrise des données

Ben Jones

Connaissances

ce que nous savons

Attitudes

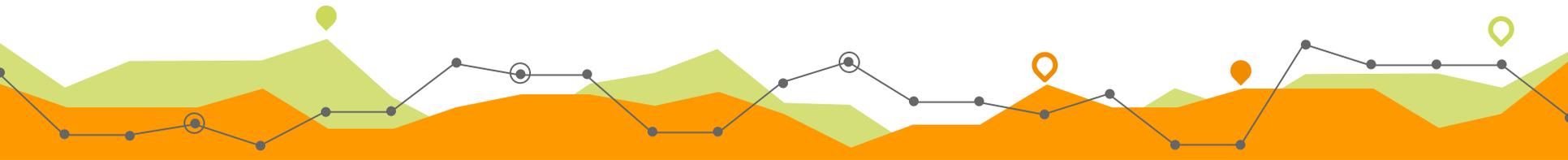
comment nous pensons et ressentons

Compétences

ce que nous pouvons faire

Comportements

comment nous agissons

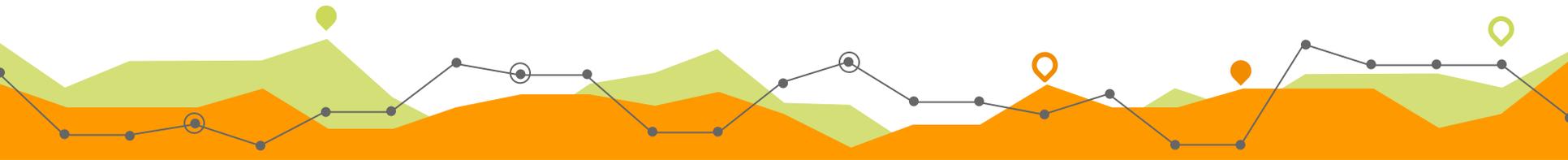


17 caractéristiques clés de la maîtrise des données

Ben Jones

Connaissances

1. Savoir **distinguer** les différentes **formes élémentaires** de données
2. Connaître la **manière** dont les données sont **collectées**, **structurées** et **stockées**
3. Comprendre les **principes fondamentaux** de l'**analyse** et des **statistiques** et quand ils s'appliquent
4. Comprendre les **moyens** de **visualiser** les données et leurs avantages et inconvénients respectifs

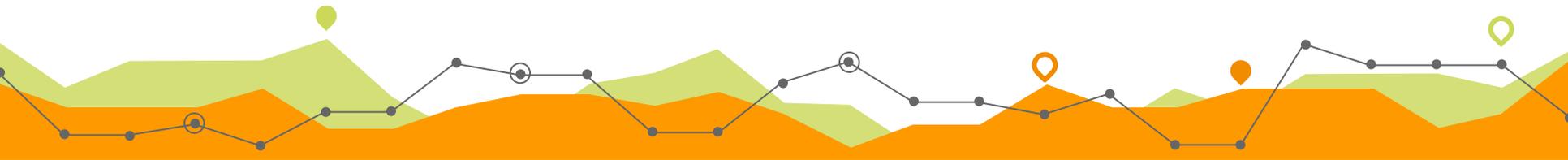


17 caractéristiques clés de la maîtrise des données

Ben Jones

Compétences

5. Lire et comprendre les affichages visuels des données **créées par d'autres**
6. Nettoyer les données **incorrectes** et combiner **plusieurs ensembles** de données pour l'analyse
7. Explorer les ensembles de données et identifier les **faits, modèles** et **tendances** pertinents
8. Créer des affichages **visuels clairs** des données pour **révéler des informations aux autres**
9. Communiquer efficacement en utilisant les données et **informations consolidées**

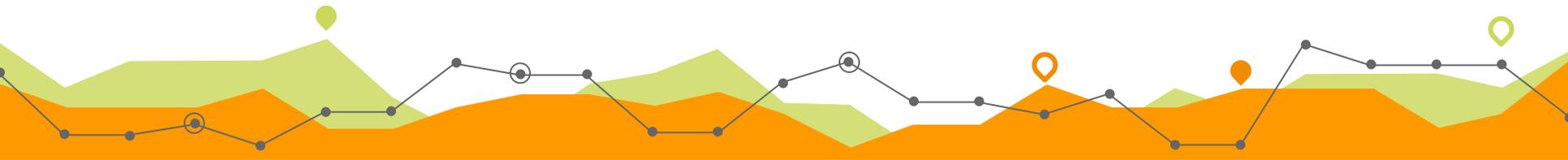


17 caractéristiques clés de la maîtrise des données

Ben Jones

Attitudes

10. Estimer que les données sont **une langue** que tout le monde peut **apprendre à lire et à parler**
11. Se sentir **confiant** en utilisant les données comme **ressources** pour répondre aux questions et en identifier de nouvelles
12. Être attentif à ses **biais cognitifs** et aux pièges courants dans lesquels les gens tombent lorsqu'ils travaillent avec des données
13. Considérer que l'**utilisation éthique** des données et l'**impact sur la société** sont de la plus haute importance

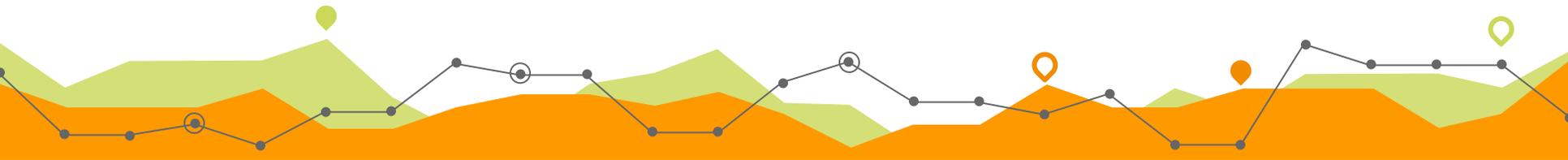


17 caractéristiques clés de la maîtrise des données

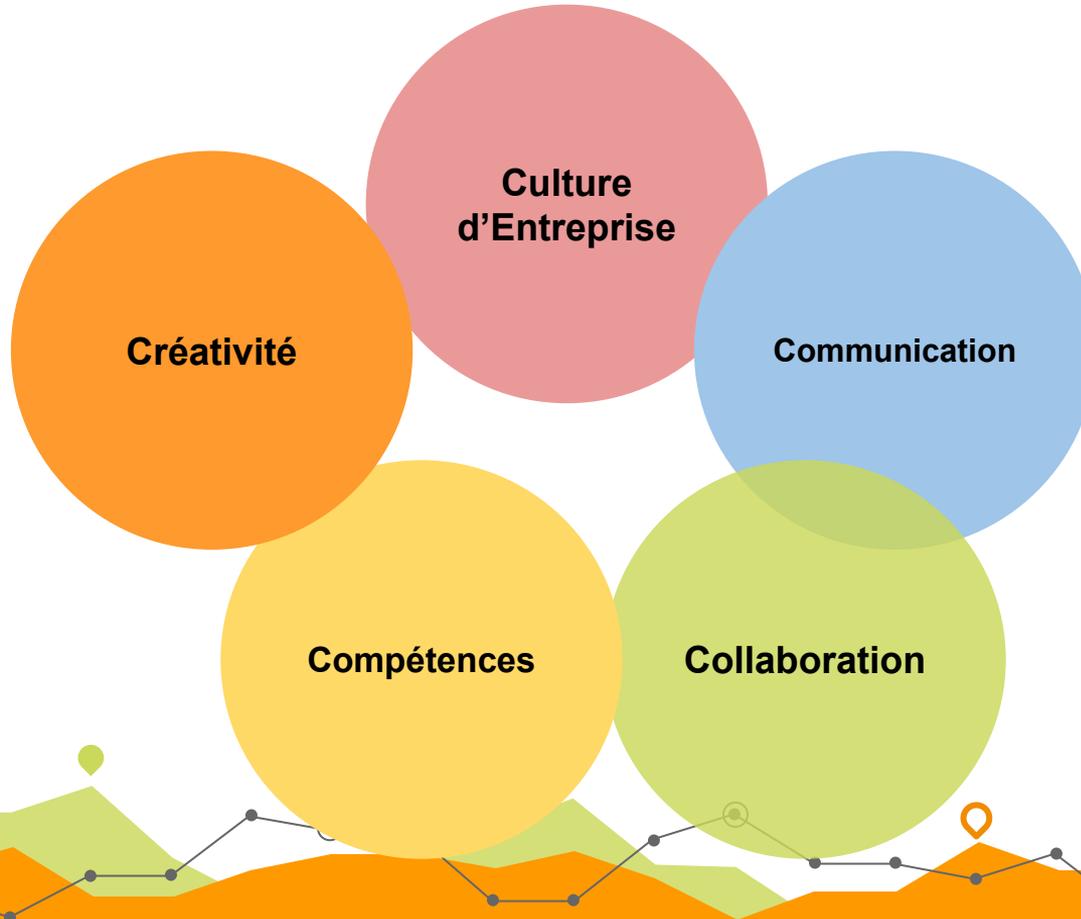
Ben Jones

Comportements

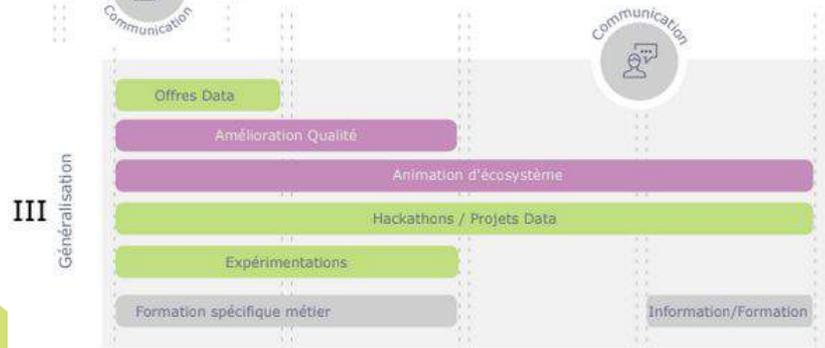
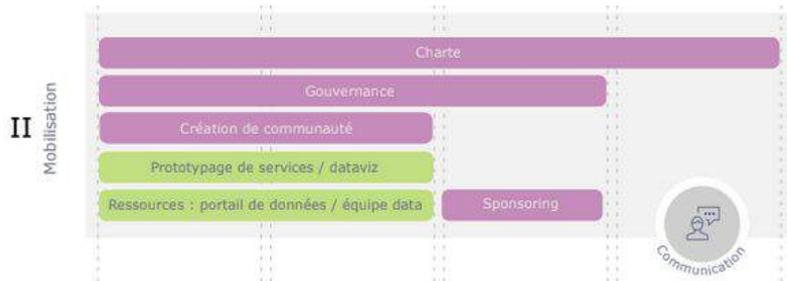
14. Chercher et créer des données comme moyen de **recueillir des informations**
15. Identifier les domaines d'**amélioration** des données et de l'analyse associée
16. Plaider pour l'utilisation efficace des données dans la **communication** et la **prise de décision**
17. Propager la maîtrise des données grâce à une **utilisation active et compétente de la langue**



La culture de la donnée dans les organisations



Appropriation de la culture de la donnée dans les organisations

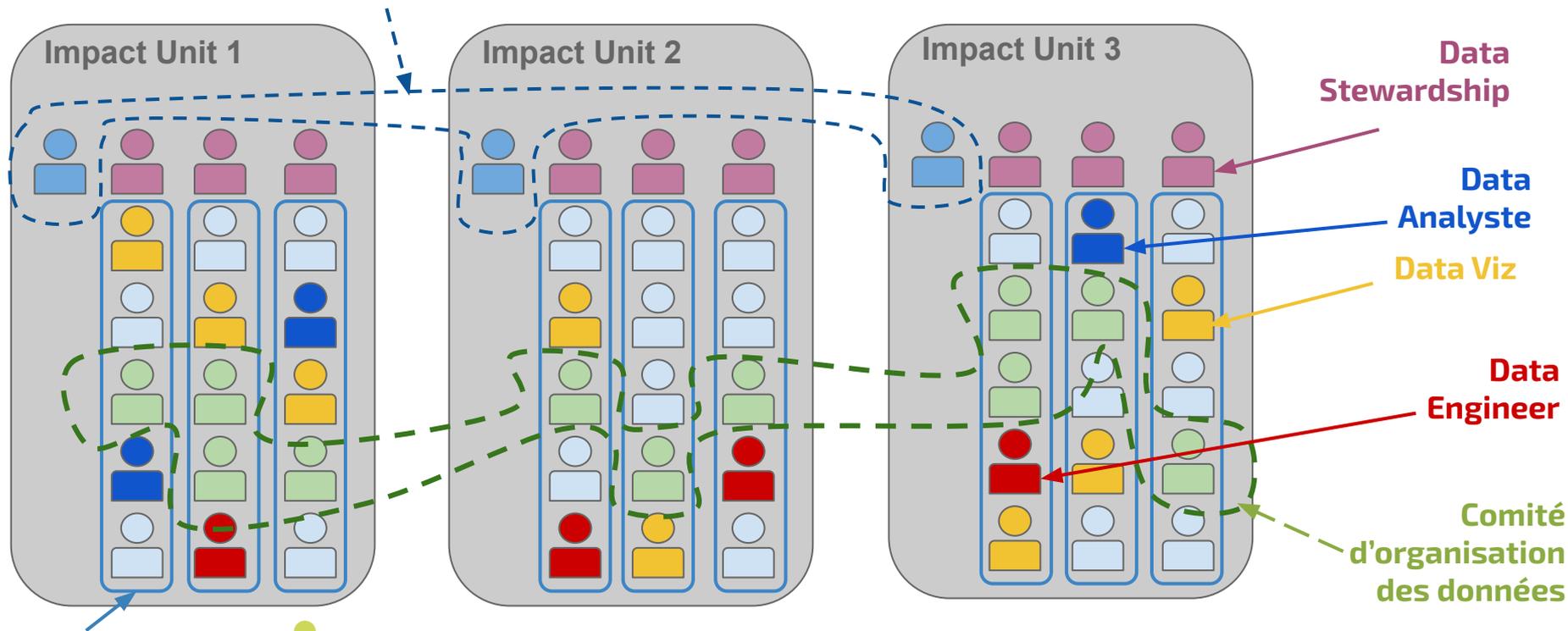


Communauté

1. Construire
2. Eduquer
3. Animer
4. Mesurer

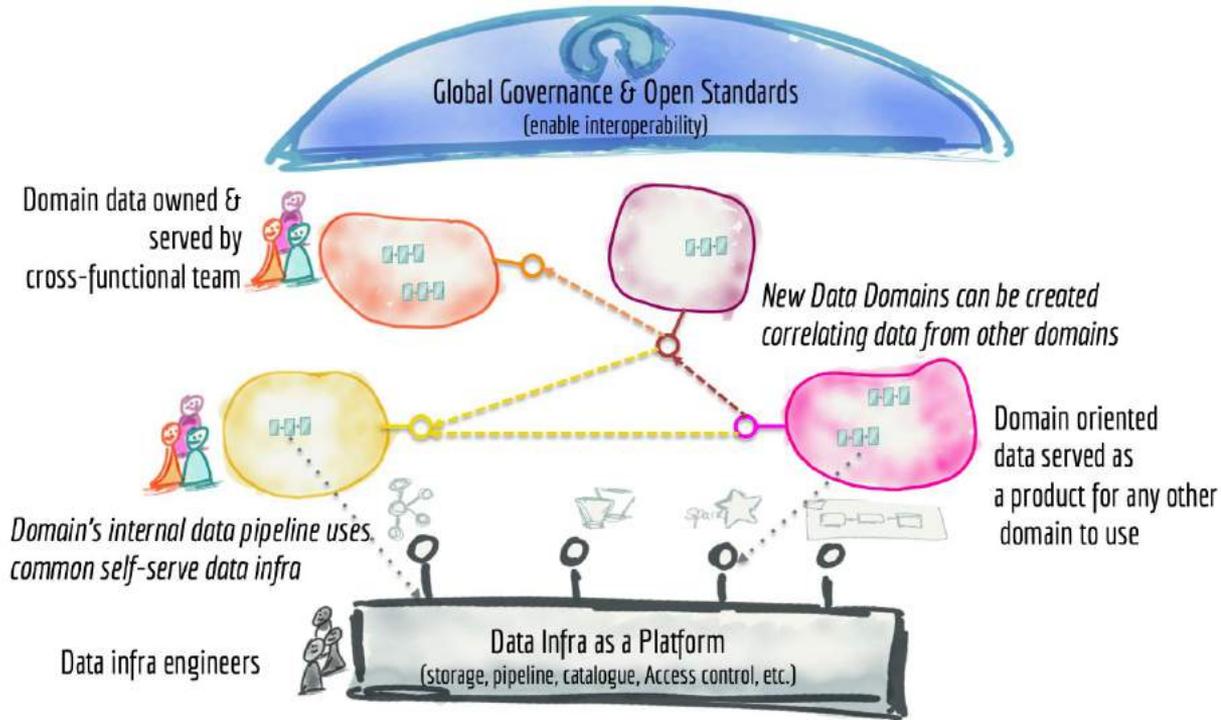
Favoriser la Data Gouvernance Agile des données

Conseil de Gouvernance des Données



Favoriser une approche libre, agile, indépendante des projets

Data Mesh



Qualité de donnée et Observabilité

DataOps



DataOps Process



DataOps est une méthodologie automatisée et orientée processus, utilisée par les équipes d'analyse et de données, pour améliorer la qualité et réduire les délais de cycle de l'analyse des données. Bien que le mouvement "DataOps" a commencé tel un ensemble de meilleures pratiques, il est maintenant devenu une approche nouvelle et indépendante de l'analyse des données

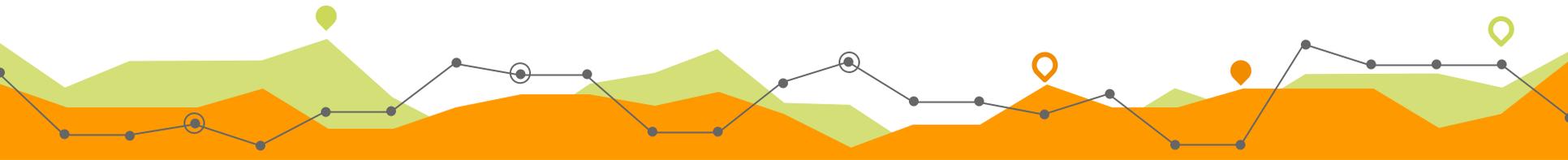


Système Opérationnel	Data Lake / Data Warehouse			Solution Décisionnelle / Data Science			
Données Opérationnelles	Spécification Entreprise		Spécification Métier	Spécification Appliquée		Spécification Restitution	
Bac à Sable	Données Brutes		Données Enrichies	Données Validées		Données Présentées	
Spillier	Construire Infrastructures	Ingestion Ordonner Structure	Exploit, Profiler	Test et valider	Modèle	Optimiser	Gouverner
Data Scientist / Business User	Data Analyst	Data Engineer	Data Scientist	Data Engineer	Data Scientist	Data Analyst	Business User
	Infrastructure	JOB, ETL	Fufling, Filings	JOB, ETL	Modèle ML	Pipeline, Présentations	Data Analyst
TESTS QUALITE							
CODE COLLABORATION							

Pour aller plus loin...

De nombreux autres sujets pour enrichir votre culture de donnée

- Organisation, processus métier,
- Système d'information et Infrastructure : cloud, conteneurisation, cloud native, stockage
- Open Data
- Données Privées
- API
- ...



Adopter un “Data Analytique Facilitateur”

Faciliter l'évolution de vos pratiques avec la donnée

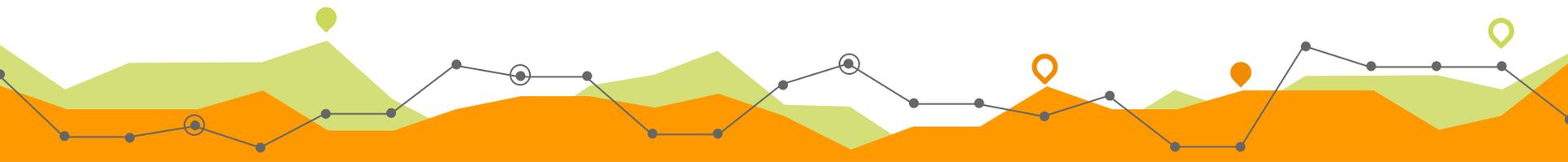
Connaissances

Culture de la donnée
Système d'information
Processus métier
Agit selon les méthodes agiles
Créativité et innovation

Animateur / Négociateur

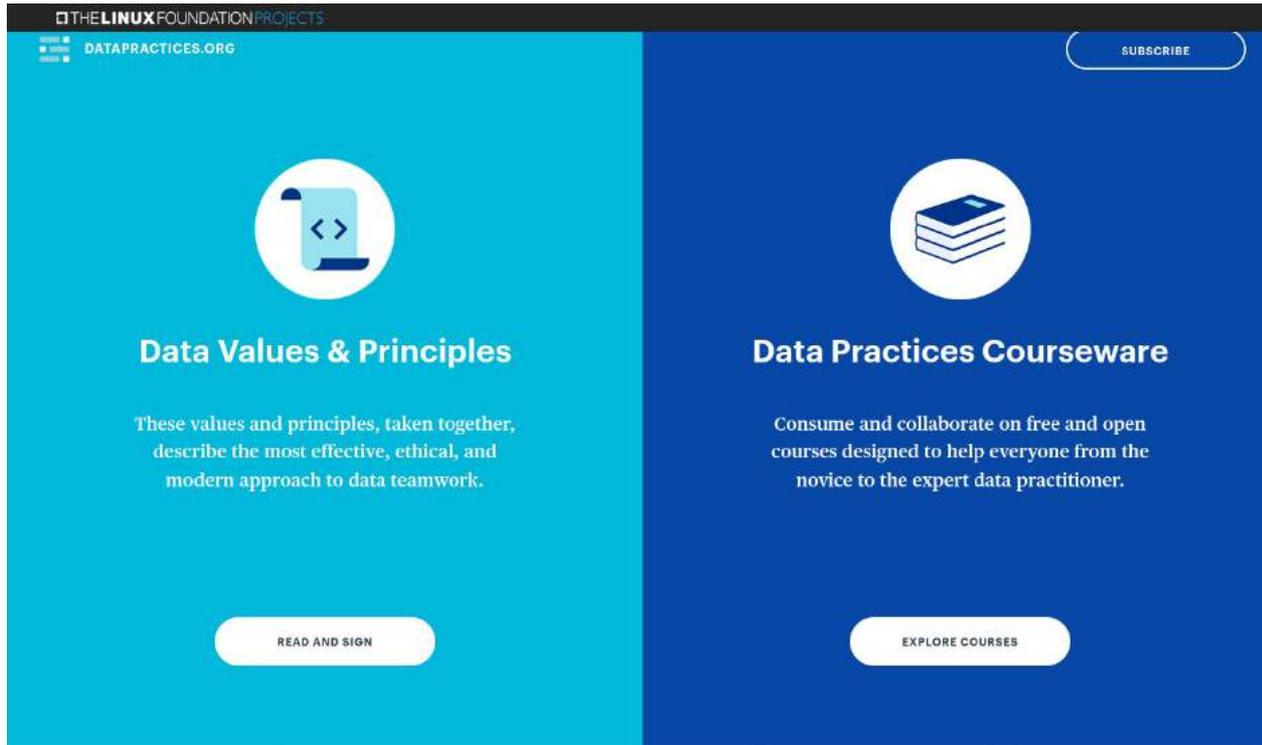
Favoriser l'engagement de vos équipes,
Promouvoir la culture de la donnée
Assurer la transparence des choix,
Organiser les rendus,
Transformer les idées en projets, les projets en produit

Contactez nous, nous avons spécialement conçu ce rôle pour vous accompagner !



Signer le manifeste de “Data Practices” de la Fondation Linux

<https://datapactices.org>



THE **LINUX** FOUNDATION PROJECTS

DATA PRACTICES
DATA PRACTICES.ORG

SUBSCRIBE



Data Values & Principles

These values and principles, taken together, describe the most effective, ethical, and modern approach to data teamwork.

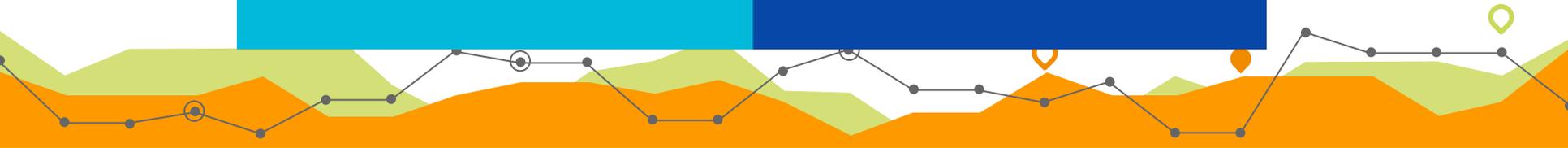
READ AND SIGN



Data Practices Courseware

Consume and collaborate on free and open courses designed to help everyone from the novice to the expert data practitioner.

EXPLORE COURSES

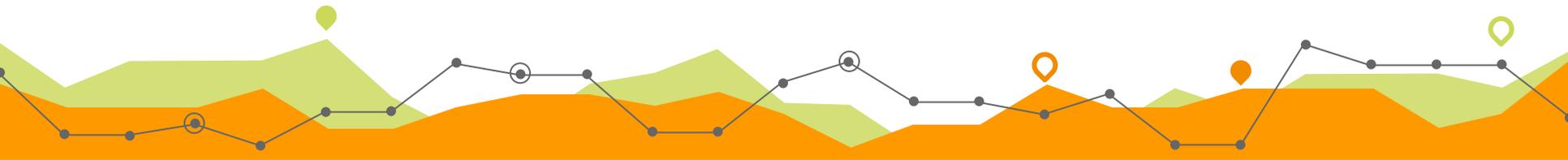


Merci !

Questions?

[@Synaltic](#) / [@egwada](#)

<https://www.synaltic.fr>



Festival Online de la Data



PROCHAINES DATES

Jeudi 04 Février de 15h à 15h30

Automatisation et pipeline CI/CD

Glenn CARROY, Développeur Web, Synaltic

Fabien OGER, Développeur Web, Synaltic

Jeudi 18 Février de 15h à 15h30

Intégrer du data storytelling dans votre stack décisionnelle

Ludovic Narayanin, DATAROCKS

Charly Clairmont, Synaltic

Jeudi 11 Février de 15h à 15h30

Intégrer rapidement vos données dans Google Cloud

Victor Coustenoble, Trifacta

Charly Clairmont, Synaltic

Jeudi 25 Février de 15h à 15h30

Tableau + R

John BONTIT, Data Analyst, Synaltic