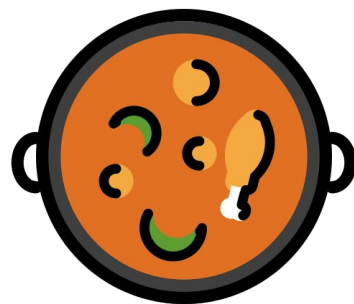
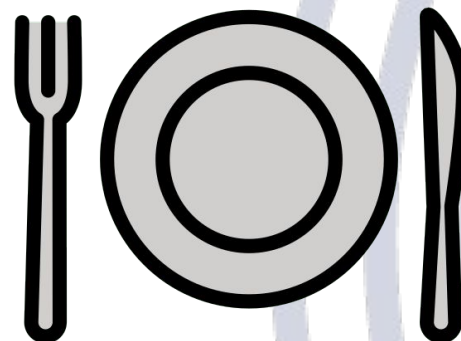


Quand la dataviz interroge le menu des cantines



Sponsors GOLD :



www.salondata.fr |
#SalonData2021



1. Qui sommes-nous et quelle est notre démarche
2. Le jeu de données
3. Premières études visuelles
4. Trouvailles
5. Conclusion

1. Qui sommes-nous et quelle est notre démarche
2. Le jeu de données
3. Premières études visuelles
4. Trouvailles
5. Conclusion : points clés data

Florian Melki

Data analyst pour
l'agence digitale

Dernier Cri

Co-organisateur du
meetup dataviz Nantes



Edith Maulandi

Dev front & dataviz chez



Membre du meetup
dataviz Toulouse

Réaliser une étude *open bar* sur des données ouvertes :

- sans a priori
- sans obligation de résultat
- dans l'idée de présenter une **démarche**
- dans l'idée de profiter de la loi Egalim (octobre 2018) et de l'ouverture des données des menus de cantine

1. Qui sommes-nous et quelle est notre démarche
- 2. Le jeu de données**
3. Premières études visuelles
4. Trouvailles
5. Conclusion

Toulouse



- Très bien formaté
- Disponible que sur une fenêtre temporelle glissante

Nantes



- Disponible sur 2011-2019
- Non standardisé

9990 lignes

Deux colonnes :

- date
- plat

Menus des cantines à Nantes de 2011 à 2019

Informations Tableau Calendrier Export API Commentaires (0)

	Date	Plat
1	3 janvier 2011	Pamplemousse
2	6 janvier 2011	Mâche au thon
3	6 janvier 2011	Croziflette
4	11 janvier 2011	Taboulé
5	12 janvier 2011	Salade carnaval (maïs, tomate, c...
6	20 janvier 2011	Potage tomate / vermicelle
7	20 janvier 2011	Camembert
8	24 janvier 2011	Potage carottes
9	25 janvier 2011	sauté de veau Marengo
10	26 janvier 2011	Pâté
11	26 janvier 2011	Haché de volaille
12	27 janvier 2011	poulet fermier
13	31 janvier 2011	salade verte
14	2 février 2011	Crêpes au sucre
15	3 février 2011	Emincé de dinde au caramel
16	8 février 2011	Toulouse grillée
17	10 février 2011	Carottes râpées
18	10 février 2011	Cœur de boeuf à la dijonnaise
19	11 février 2011	tomate / concombre
20	11 février 2011	poisson meunière
21	11 février 2011	Brocolis beurre Provençal
22	11 février 2011	St Bricet
23	14 février 2011	Blanquette de veau
24	17 février 2011	cordons bleus de volaille
25	17 février 2011	st paulin

- Y a-t-il des cycles ?
- Quels types de légumes ?
- Présence de produits transformés ?
- Quid du bio ?
- Présence de sucres ajoutés ?
- Evolution des proportions entre 2011 et 2019
- Evolution de la composition d'un repas au cours du temps (*apparition légumineuses, repas végé...*)

Google Spreadsheet

viande blanche	veau					
viande blanche	dinde					
viande blanche	canard					
viande blanche	pintade					

Jupyter Notebook partagé (Python)

Mains, courage et dévotion

```
[ ] dfAlimentsGroup['regex'] = dfAlimentsGroup.apply(
    lambda row: '|'.join(row['aliment'])
    , axis = 1)

▼ Tag des plats par aliments

[ ] def flagCategorie(df, categorie, alimentRegex):
    df[categorie] = df["Plat"].str.lower().replace("é", "e", regex=True).replace("è", "e", regex=True).str.c

[ ] for index, row in dfAlimentsGroup.iterrows():
    flagCategorie(df2, row['categorie'], row['regex'])

/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html

/usr/local/lib/python3.7/dist-packages/pandas/core/strings.py:2001: UserWarning: This pattern has match gr
return func(self, *args, **kwargs)

[ ] len(df2.index)

9866

[ ] df2.head(3)
```

	Date	Plat	autre protéine	feculent	fruit	laitage	legume	plat préparé	poisson fruit de mer	produit industriel	sauce	sucré
0	2011-01-04	Crêpe au fromage	False	False	False	True	False	True	False	False	False	F

Création de listes d'aliments parmi :

- viande rouge
- viande blanche
- poisson / fruits de mer
- légume
- féculent
- fruit
- laitage

1	categorie	aliment
152	feculent	Flageolets
153	feculent	Haricots
154	feculent	Haricots rouges
155	feculent	Haricots blancs
156	feculent	Mogettes
157	feculent	Haricots noirs
158	feculent	Lentilles
159	feculent	Pois chiche
160	feculent	Pois jaune
161	feculent	Pois vert
162	feculent	Avoine
163	feculent	Blé
164	feculent	Épeautre
165	feculent	Maïs
166	feculent	Millet
167	feculent	Orge
168	feculent	Riz
169	feculent	Seigle
170	feculent	Igname
171	feculent	Manioc
172	feculent	Patate douce
173	feculent	Pomme de terre
174	feculent	Topinambour
175	fruit	abricot

Vérification pour chaque plat de la présence d'un aliment dans les listes créées

Tag de chaque plat



Objectif : taggage d'au moins 95% des plats



Difficultés rencontrées :

- variété d'orthographe : boeuf / bœuf
- fautes : omellette
- accents : Pâtes vs pâté
- plats hors catégories : hachis parmentier, ratatouille...
- évènements : sapin de Noël, grève, pentecôte, nouvel an chinois...



Complétion des listes d'aliments :

- produit industriel
- plat préparé
- sous-catégorisation des plats préparés et produits industriels en aliments primaires
- évènements à exclure

categorie	aliment	à ajouter dans	
laitage	Port Salut		
laitage	emmental 220gr		
laitage	entrammes en 240gr		
laitage	St Bricet		
laitage	Saint-Nectaire		
laitage	suisse aro		
laitage	Fleur d'Anjou bio		
plat préparé	hachis parmentier	viande rouge,feculent	
plat préparé	parmentier de poisson	poisson fruit de mer,feculent	
plat préparé	omelette	autre protéine	
plat préparé	Omelette	autre protéine	
plat préparé	Potage	legume	
plat préparé	Pain perdu	feculent	
plat préparé	Gaspacho	legume	
plat préparé	chili con carné	viande rouge,legume,feculent	

sapin de noel		
PA lundi		
PA jeudi		
Journée internationale sans viande		
reprise scolaire		
REPAS POUR LA PLANETE		
MANGER BOUGER		
GRANDS MENAGES		

Le jeu de données Catégorisation

Prise en compte des nouvelles catégories

Mise en place d'une pondération des catégories d'aliments

Agrégation par jour

Taggage des plats déclarés "bio"

	Date	Wautre protéine	Wfeculent	Wfruit	Wlaitage	Wlegume	Wsucrerie	Wsauce	Wpoisson	fruit de mer	Wviande blanche	Wviande rouge	Wbio	Windustriel
0	2011-01-03	0.000	0.250	0.25	0.000	0.250	0.125	0.0	0.00	0.00	0.0	0.125	0.00	0.50
1	2011-01-04	0.000	0.200	0.20	0.300	0.300	0.000	0.0	0.00	0.00	0.0	0.000	0.00	0.00
2	2011-01-05	0.000	0.200	0.10	0.200	0.300	0.000	0.0	0.00	0.00	0.2	0.000	0.00	0.00
3	2011-01-06	0.000	0.000	0.00	0.200	0.300	0.000	0.0	0.30	0.30	0.2	0.000	0.20	0.00
4	2011-01-07	0.000	0.125	0.00	0.375	0.375	0.125	0.0	0.00	0.00	0.0	0.000	0.00	0.25
...
2046	2020-01-13	0.200	0.400	0.20	0.200	0.000	0.000	0.0	0.00	0.00	0.0	0.000	0.60	0.40
2047	2020-01-14	0.000	0.250	0.25	0.125	0.375	0.000	0.0	0.00	0.00	0.0	0.000	0.00	0.00
2048	2020-01-15	0.000	0.000	0.25	0.250	0.250	0.000	0.0	0.25	0.25	0.0	0.000	0.00	0.25
2049	2020-01-16	0.125	0.250	0.00	0.375	0.125	0.000	0.0	0.00	0.00	0.0	0.125	0.25	0.25
2050	2020-01-17	0.000	0.200	0.20	0.200	0.200	0.000	0.1	0.10	0.10	0.0	0.000	0.20	0.20

2051 rows x 13 columns

Tag par catégorie

Plat	Jour	Féculent	Viande	Legumes
1	1	true	false	false
2	1	true	false	true
3	1	true	true	true

Pondération par plat

Plat	Jour	Féculent	Viande	Legumes
1	1	1	0	0
2	1	0.5	0	0.5
3	1	0.33	0.33	0.33

Agrégation par jour

Jour	Féculent	Viande	Legumes
1	61 %	11 %	28 %
2

1. Qui sommes-nous et quelle est notre démarche
2. Le jeu de données
- 3. Premières études visuelles**
4. Trouvailles
5. Conclusion



Outil “no code”



Rapide pour la recherche exploratoire



Interactif



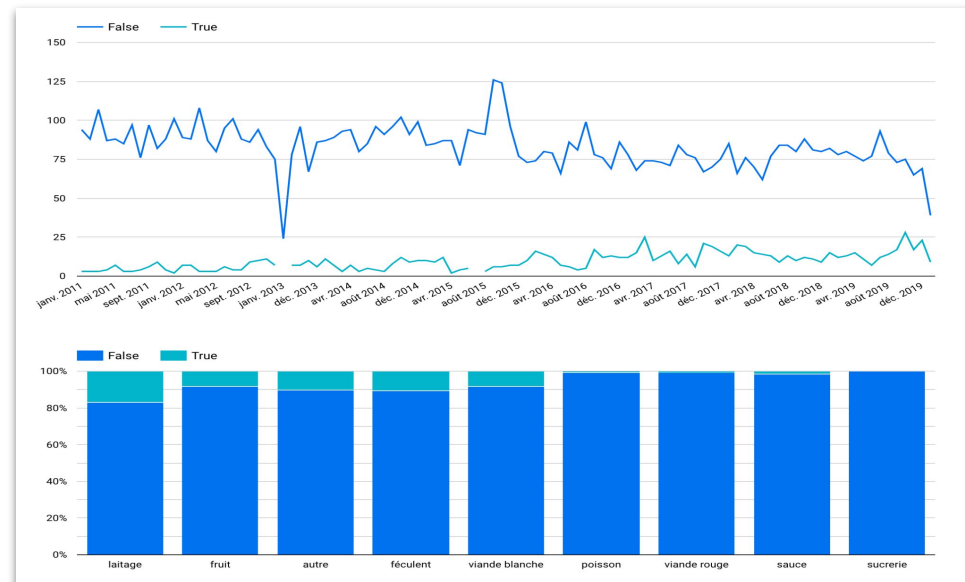
Gratuit



Limité dans les représentations offertes



Pas toujours clair pour les calculs





Notebook reactif dans le navigateur,
rapide à mettre en place



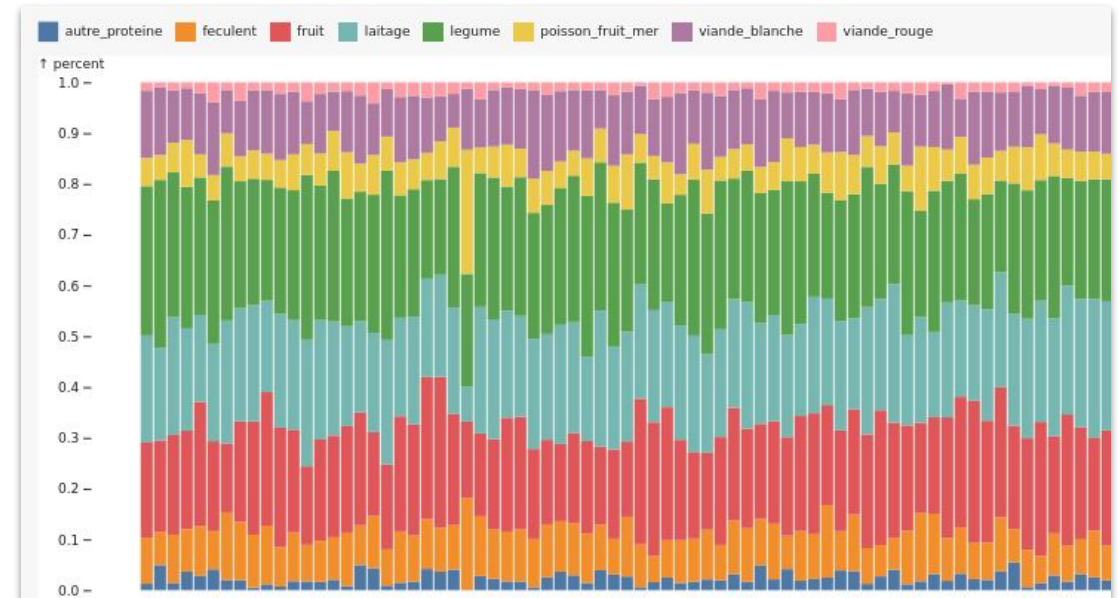
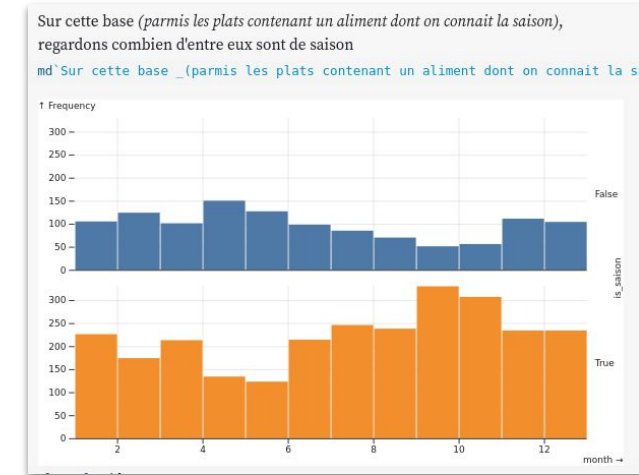
Des librairies disponibles pour
explorer les données facilement +
multiples exemples



Expérimentation et intégration
future avec d3 facilité



Nécessite des connaissances de
bases en Javascript

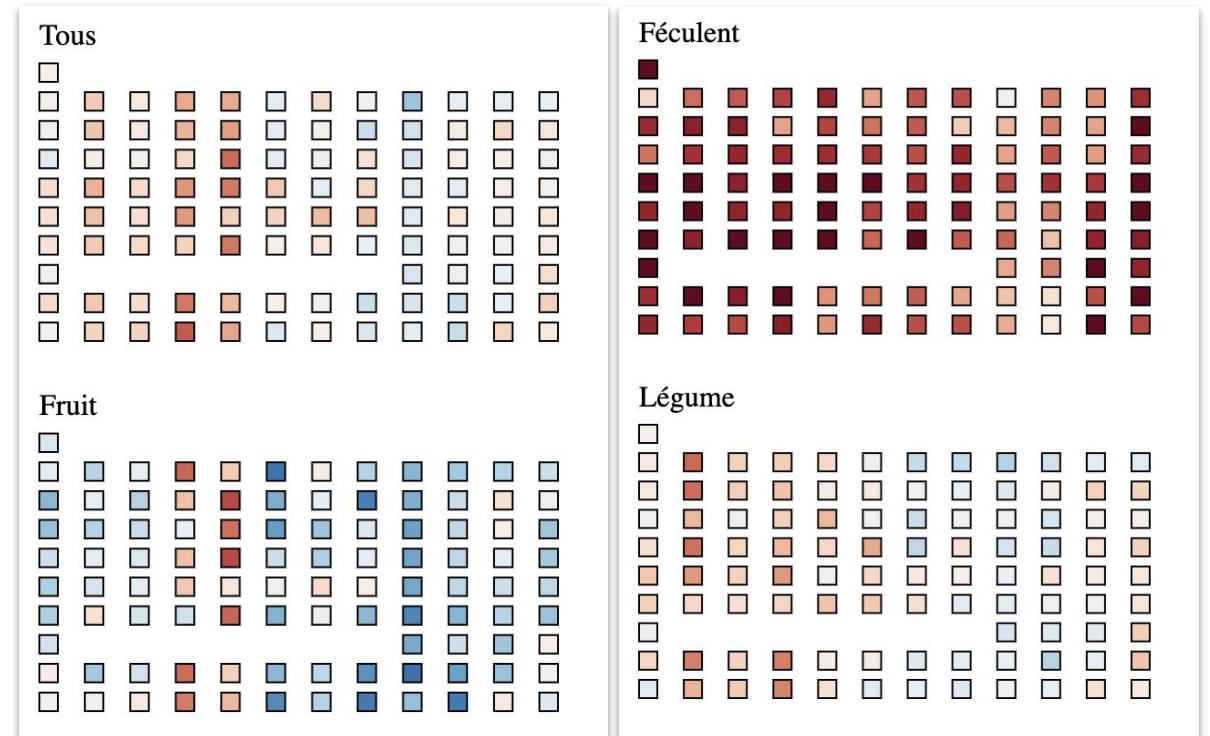




Rapide pour prototyper des représentations spécifiques

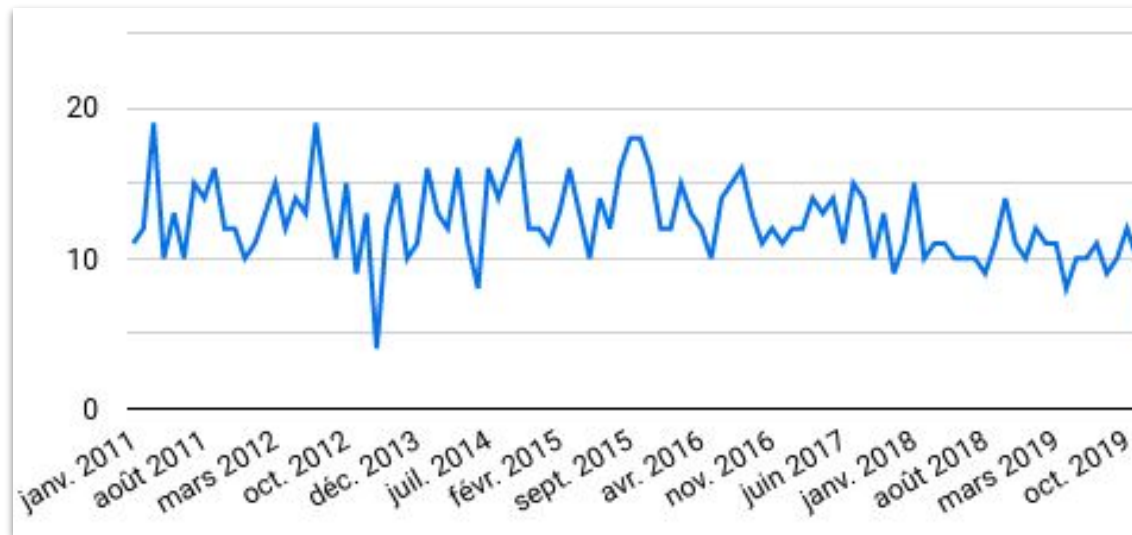


Long pour rendre un produit abouti



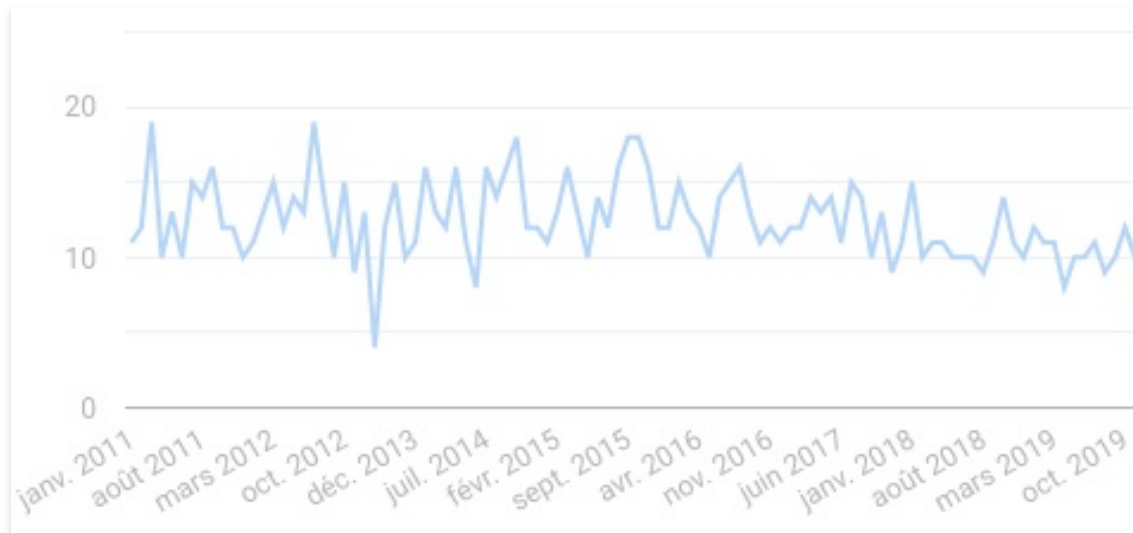
1. Qui sommes-nous et quelle est notre démarche
2. Le jeu de données
3. Premières études visuelles
- 4. Trouvailles**
5. Conclusion

Nombre de plat par mois contenant de la viande



Vue calendrier

Nombre de plat par mois contenant de la viande

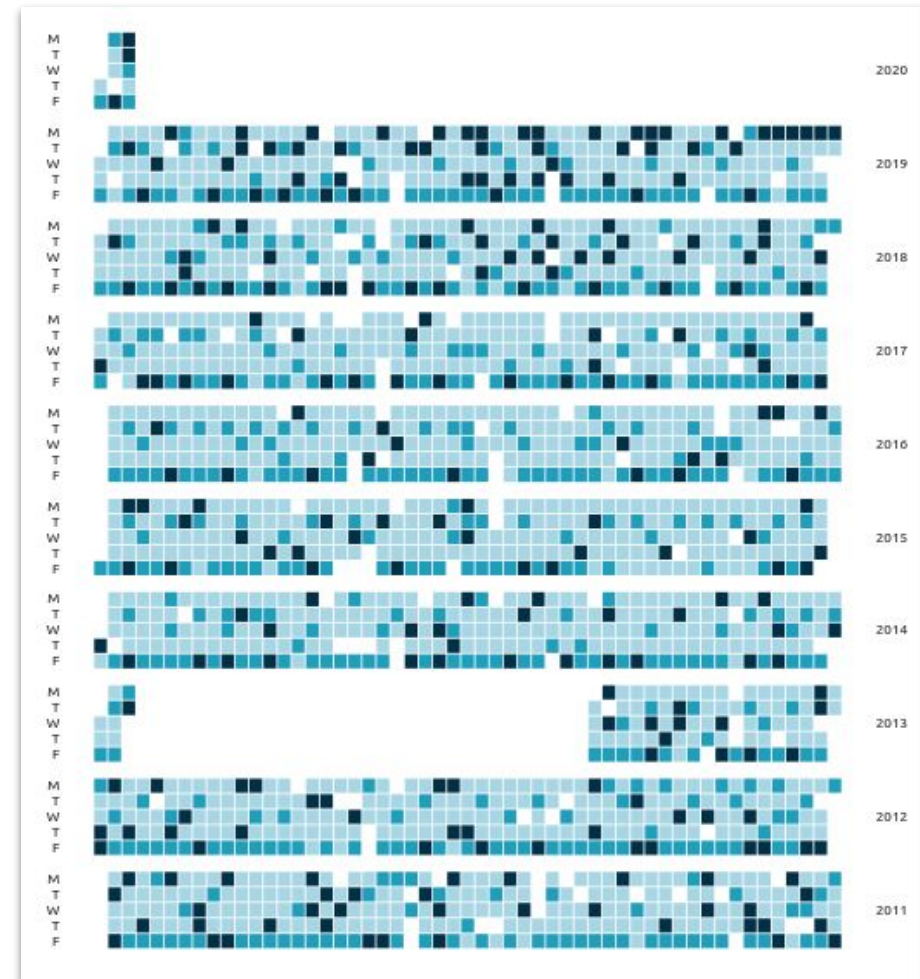


Type de repas

viange rouge / blanche

que poisson

sans protéine animale

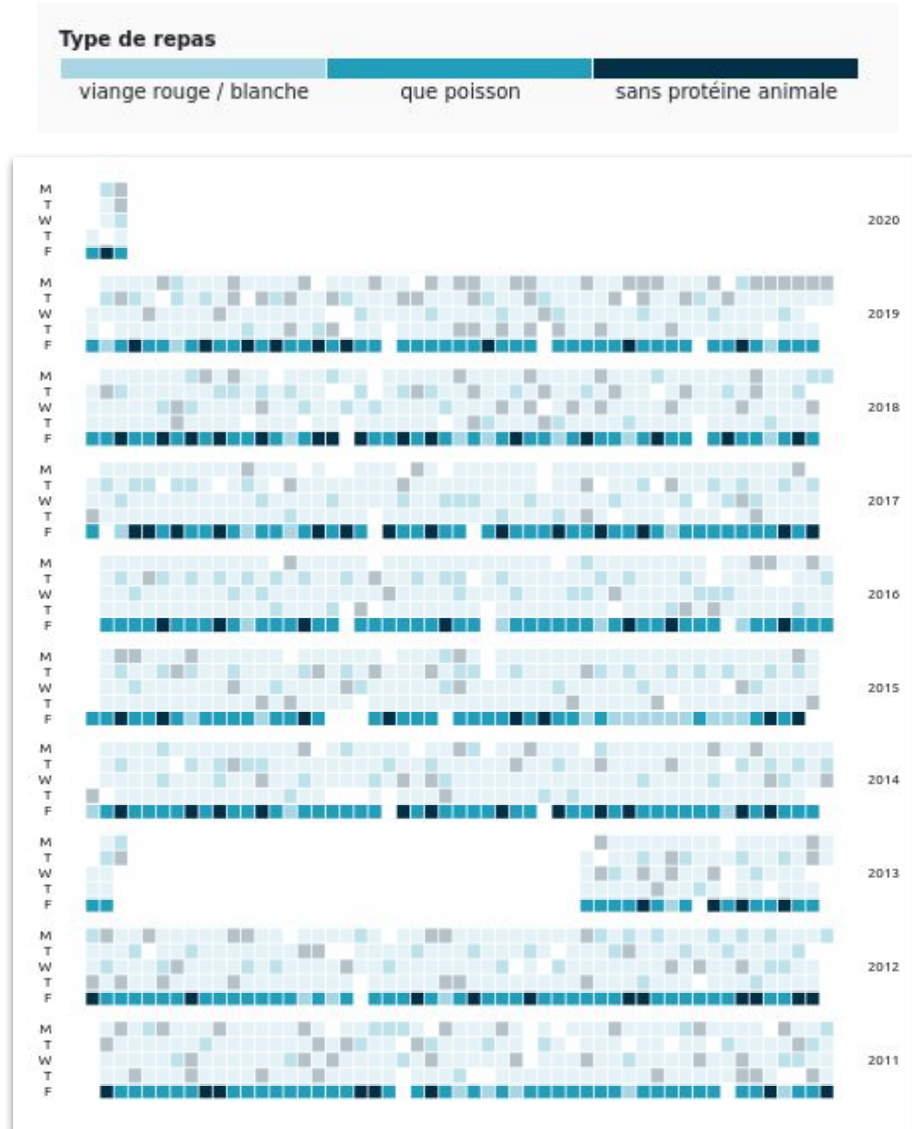
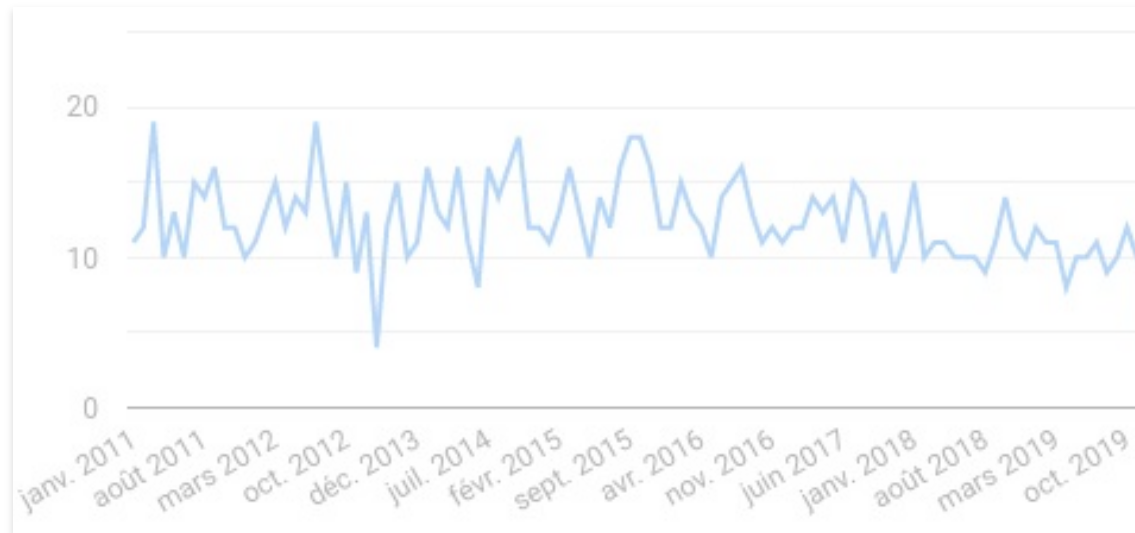


2020

2011

Vue calendrier

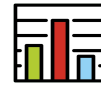
Nombre de plat par mois contenant de la viande





Les limites

- éventuels trous dans les données visible ou non selon la granularité
- jours sans données, deux viandes : vérifier le processus de catégorisation



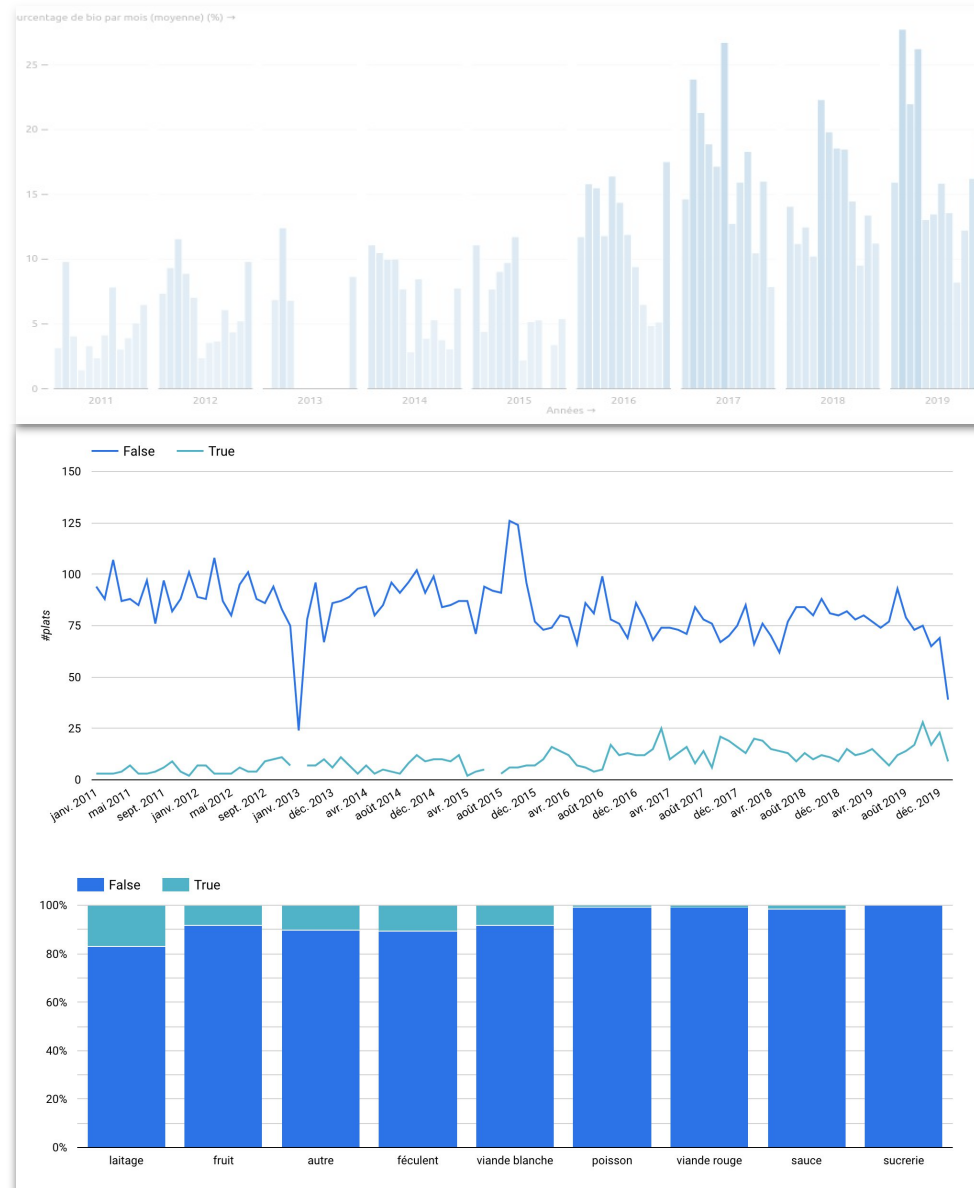
Côté dataviz

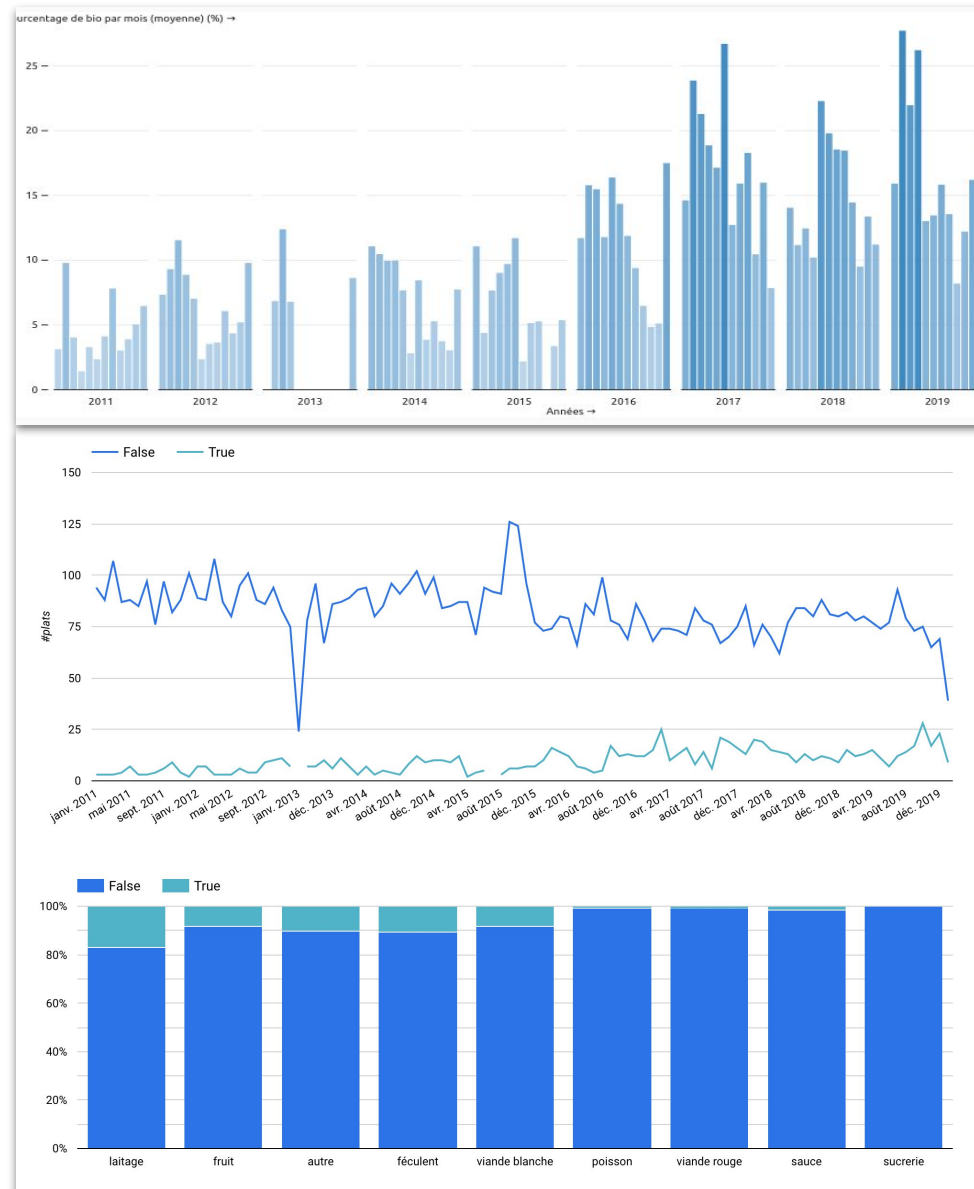
- Vue statistique vs vue détaillée des fois plus facile à appréhender
- Explorer plusieurs niveaux pour découvrir des patterns



Côté data

- Créer des variables dérivées pour réduire la complexité / souligner un point spécifique







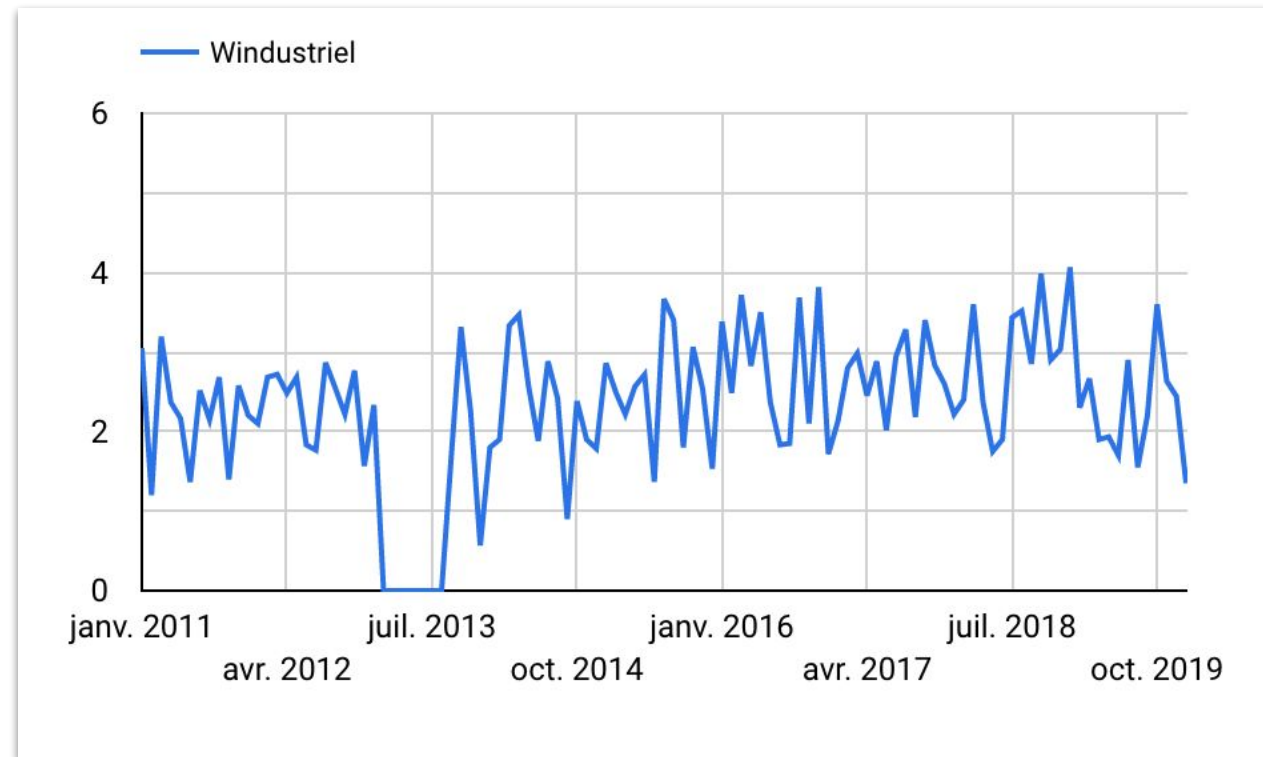
Les limites

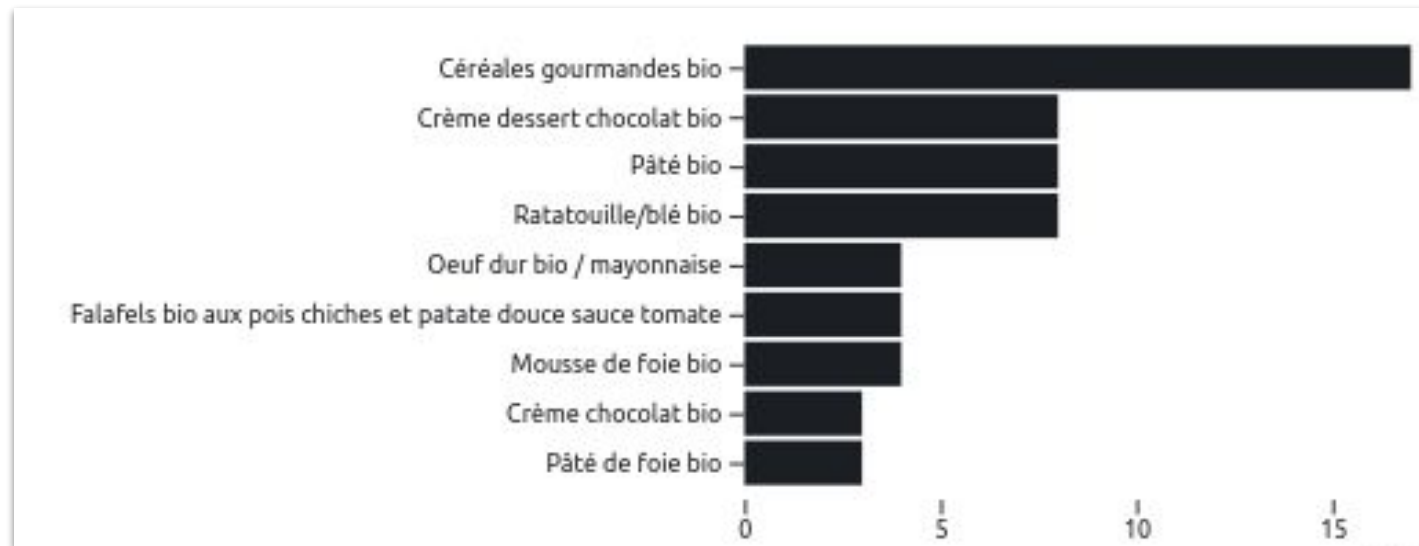
La mention “bio” est
uniquement déclarative

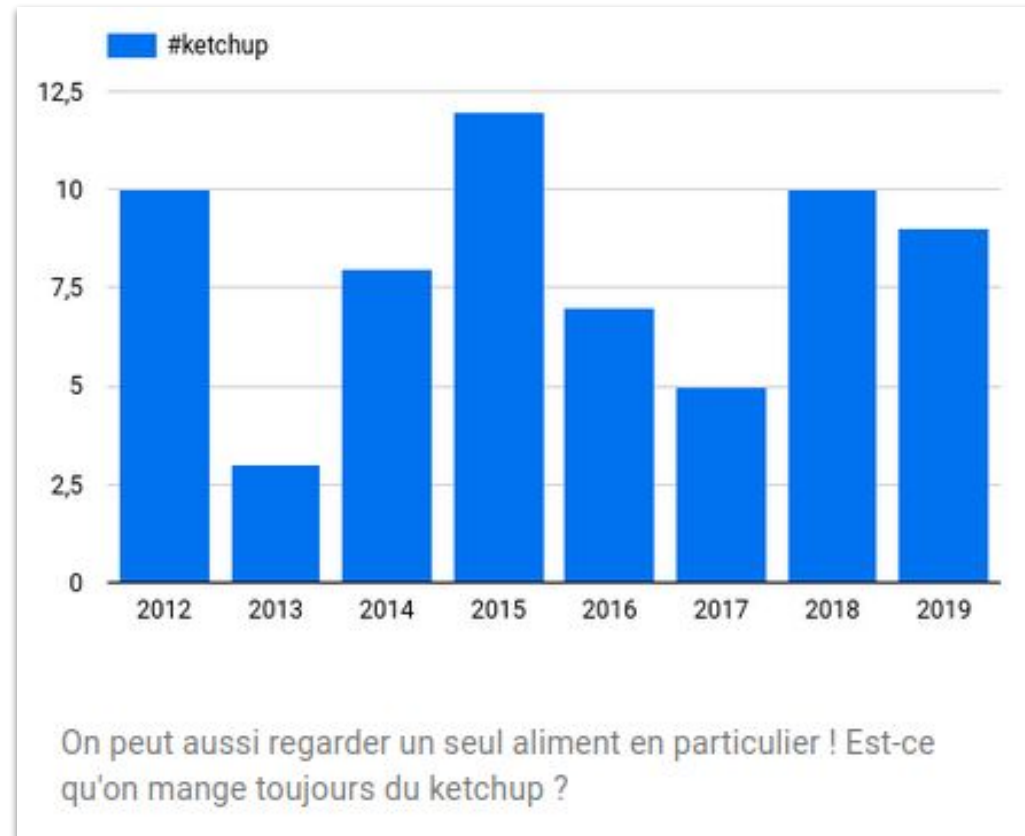


Côté dataviz

Nombre de plats vs
pourcentage de plats









Les limites

- Qu'est-ce qu'un produit industriel ?
- Gestion des cas limites :
*"Poêlée ratatouille Bonduelle",
 "ratatouille", "ratatouille maison"*



Côté dataviz

Plus pertinent de faire un focus sur un type de produit en particulier



Côté data

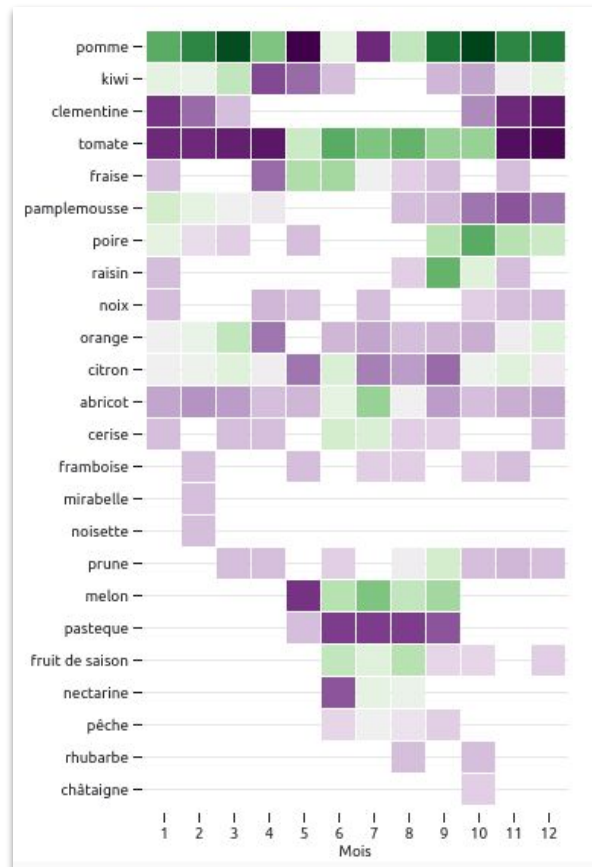
Taggage manuel

de saison

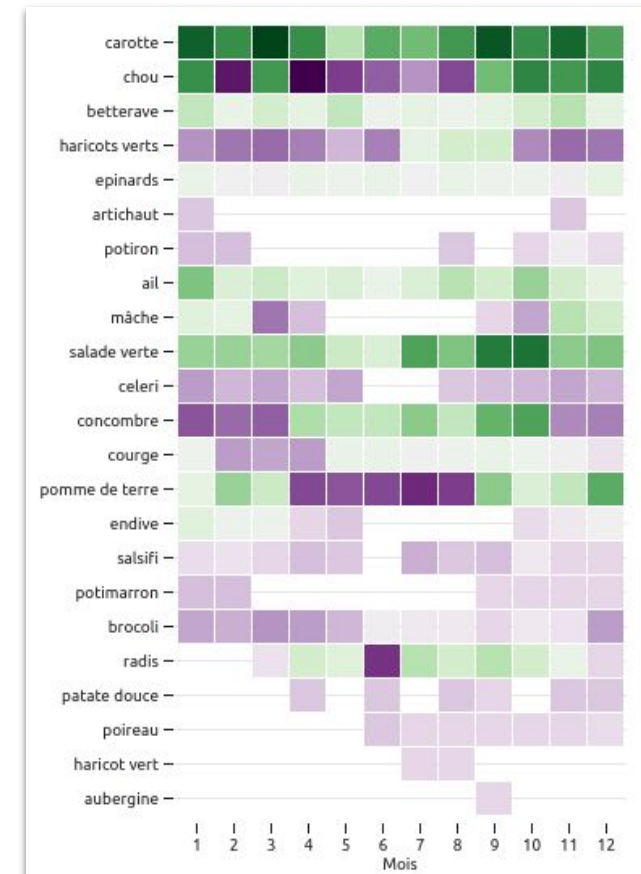


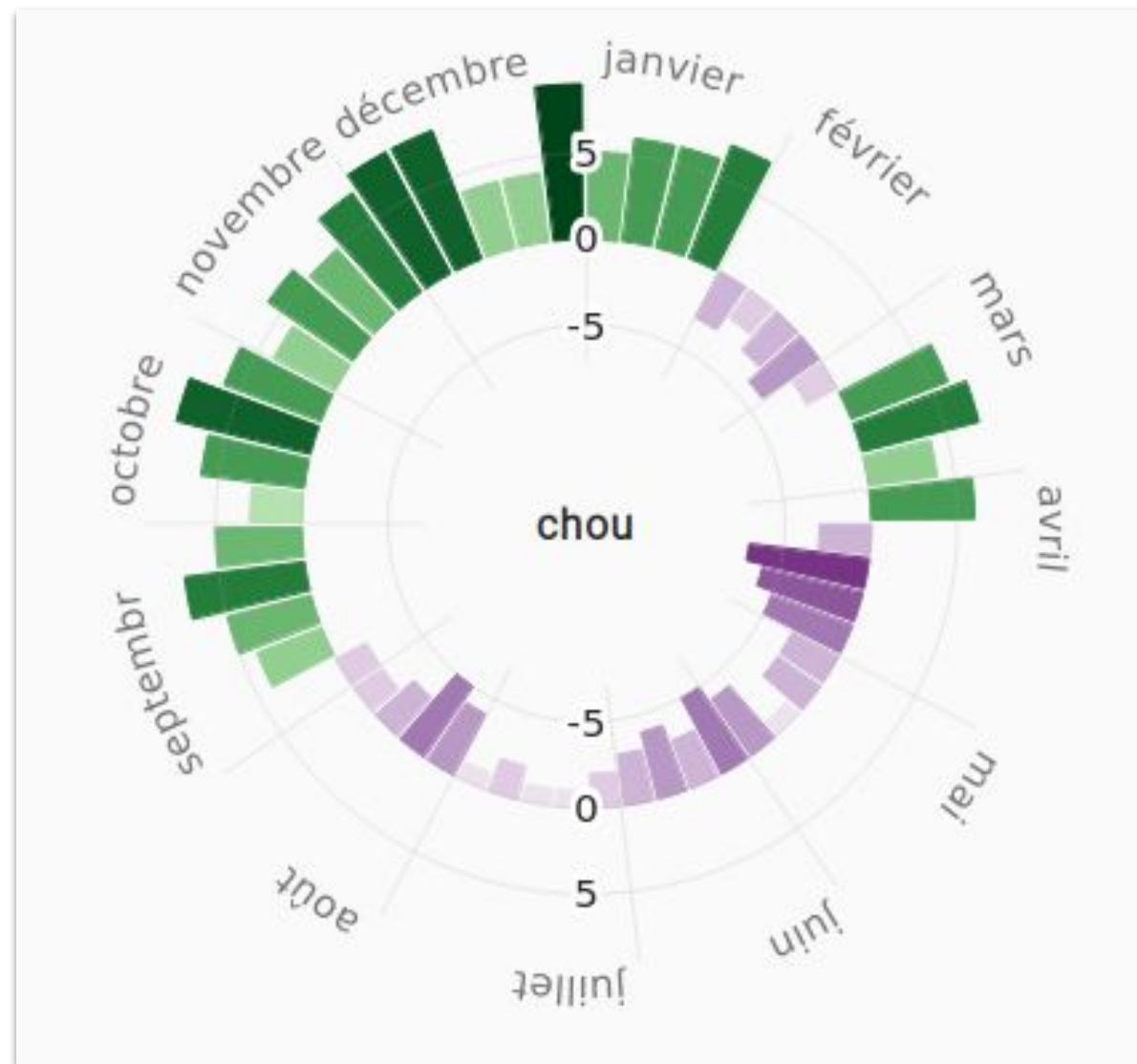
fruits

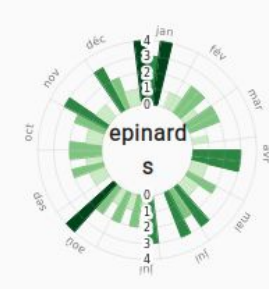
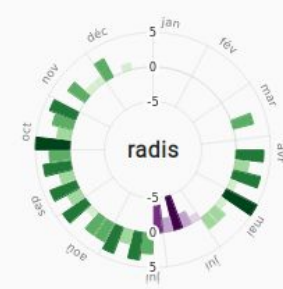
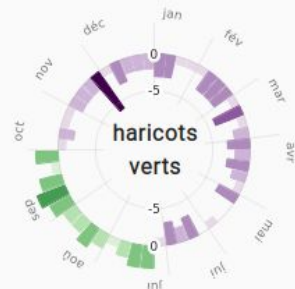
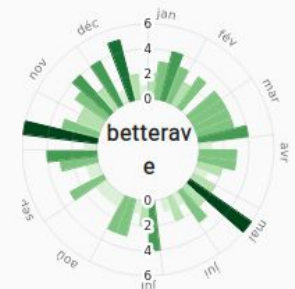
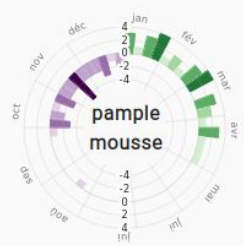
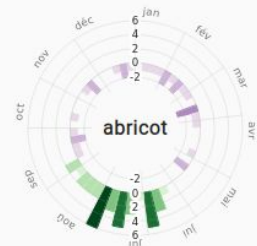
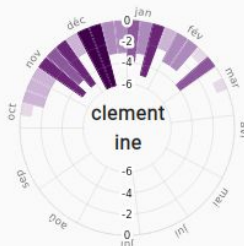
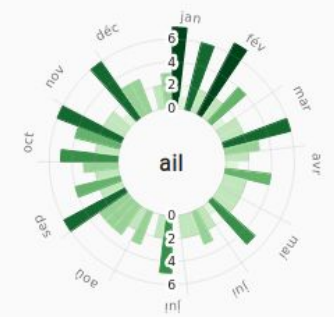
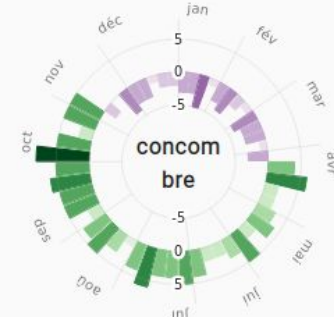
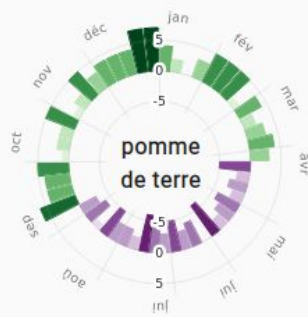
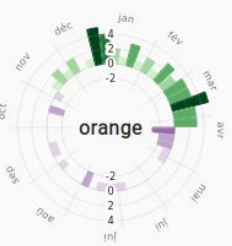
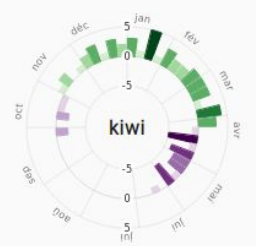
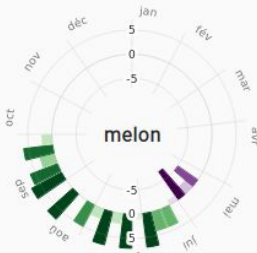
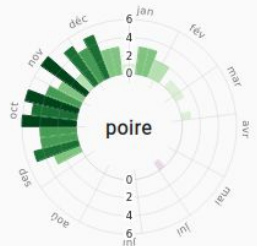
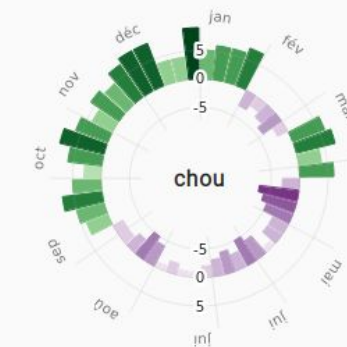
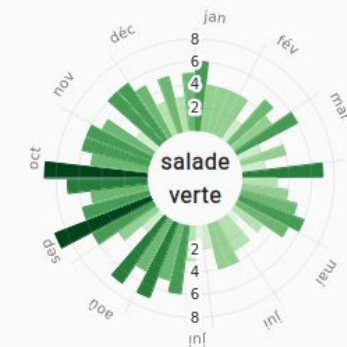
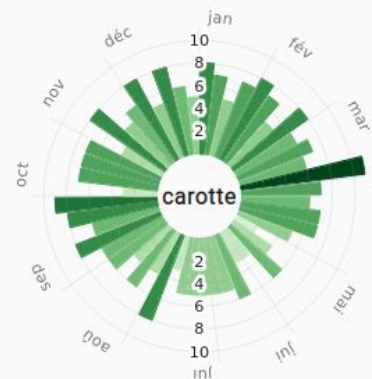
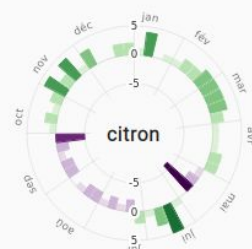
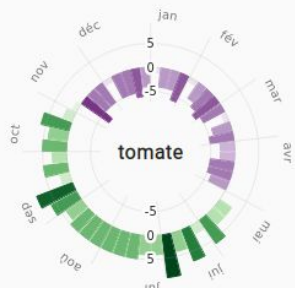
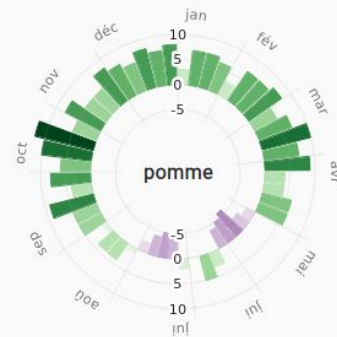
plus
consommés



légumes









Les limites

- couverture = 67% des plats
(plats préparés et industriel non pris en compte)
- banane / ananas hors France ➡ différentes notions de saisonalités
- fruit / légume identifié dans le texte sans nuances
(compote de pomme)



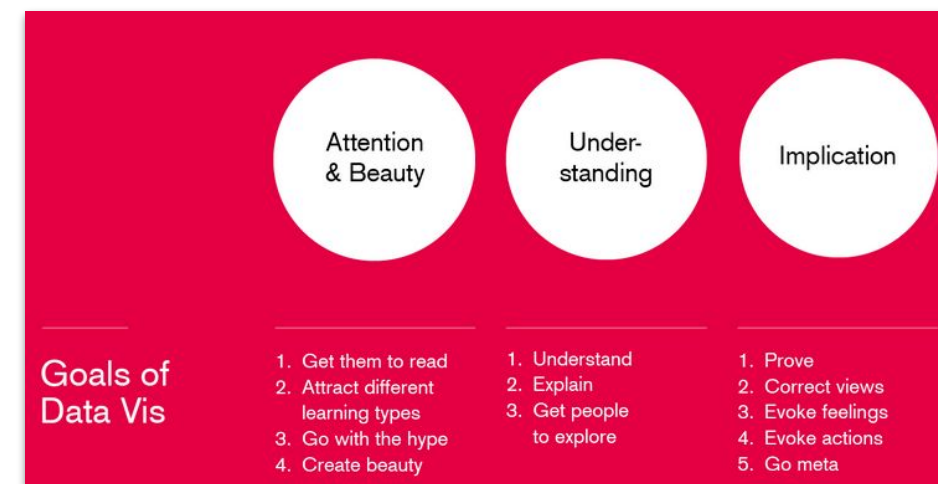
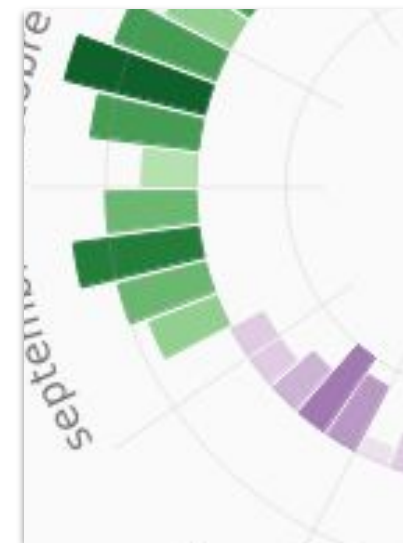
Les limites

- couverture = 67% des plats (*plats préparés et industriel non pris en compte*)
- banane / ananas hors France ➡ différentes notions de saisonnalités
- fruit / légume identifié dans le texte sans nuances (*compote de pomme*)



Côté dataviz

- Vue radiale “mensongère”
- De la représentation simple à plus travaillée :
 - quels objectifs ?
 - quels formats ?
 - quel public ?





Les limites

- couverture = 67% des plats (*plats préparés et industriel non pris en compte*)
- banane / ananas hors France ➡ différentes notions de saisonalités
- fruit / légume identifié dans le texte sans nuances (*compote de pomme*)



Côté dataviz

- Vue radiale “mensongère”
- De la représentation simple à plus travaillée : quels objectifs ? quels formats ? quel public ?



Côté data

- Croiser les jeux de données pour enrichir les analyses
- Pas de jeux de données disponibles ? Dans certains cas, on peut en construire soi-même

1. Qui sommes-nous et quelle est notre démarche
2. Le jeu de données
3. Premières études visuelles
4. Trouvailles
5. Conclusion

Data

- **Etre conscient des limites** du jeu de données / de sa propre analyse / de son interprétation
- **Croiser** avec les autres jeux de données (voire les construire)
- **La dataviz est un bon outil** pour explorer des pistes, infirmer des hypothèses
- **Parler avec des gens du métier** : producteur de la donnée, praticien, politique...

Dataviz

- **Concevoir ses visualisations** de données en fonction de son **auditoire** et de son **objectif**
- Expérimenter et tester différents niveaux de **granularité**
- Retravailler les données pour être plus **pertinent**
- Différents **types de rendus** : statique (image, poster), exploratoire (dashboard), site web scrollytelling ... alignés avec ses objectifs



Menu des cantines

Explorer l'open data avec la dataviz

Qu'est-ce que l'on peut apprendre sur le menu des cantines grâce à l'open data et à la dataviz ? C'est la question que l'on s'est posé (Florian Melki et Edith Maulandi).

Le but de ce projet : réaliser une étude open bar sur des données ouvertes sans a priori (*et sans obligation de résultat*). Nous allons voir avec vous comment cet exemple peut nous éclairer sur l'utilisation des jeux de données en open data et de l'usage de la dataviz pour les comprendre et les valoriser.

 **Comment choisir un jeu de données et le manipuler ?**

Sélection

Nous avons parcouru les portails open data de plusieurs villes en cherchant

menu-cantine.netlify.app

github.com/emaulandi/menu-cantines

Florian Melki

florian@derniercri.io

 @FloDataviz

dernier cri



Edith Maulandi

edith.maulandi@makina-corpus.com

 @edithmaulandi



 MAKINA CORPUS