



# Club Power BI

Mardi 9 mars 2021 sur Teams

Retour d'expérience **Azure** + Power BI chez

**arte**



/Club-Power-BI



@ClubPowerBI



/ClubPowerBI



/ClubPowerBI

# Pendant la session

The screenshot shows a Microsoft Teams meeting window. The top bar includes the 'Microsoft Teams' logo, a search bar with the text 'Recherchez ou tapez une commande', and a user profile 'GUSS'. The left sidebar contains navigation icons for 'Activité', 'Conversation', 'Équipes', 'Calendrier', and 'Appels'. The main content area displays a presentation slide titled 'Power Saturday' with a date '6 juin 2020, Online' and a background image of the Eiffel Tower. Below the slide is a toolbar with icons for video, microphone, screen sharing, and chat. A yellow box with an arrow points to the chat icon, and another yellow box with an arrow points to the screen sharing icon. A third yellow box with an arrow points to the microphone icon. A fourth yellow box with an arrow points to the video icon. A fifth yellow box with an arrow points to the chat icon. A sixth yellow box with an arrow points to the chat icon. A seventh yellow box with an arrow points to the chat icon. A large yellow box with an arrow points to the chat window on the right. The chat window is titled 'Conversation de la réunion' and shows a message from 'Joel CREST (Invité)' at 20:28: 'Je suis tenté de rejoindre ... mais non'. Below the chat window is a text input field labeled 'Répondre' and a row of icons for attachments, emojis, GIFs, and other features.

Microsoft Teams

Recherchez ou tapez une commande

GUSS

Activité

Conversation

Équipes

Calendrier

Appels

Power Saturday

6 juin 2020, Online

La conférence #Data, #PowerBI, #PowerPlatform, #O365, #M365, #Teams et bien plus encore

02:18

Levez la main pour demander la parole.  
Ouvrir la fenêtre de discussion

Nous demandons de désactiver Caméra et Micro

Ne partagez pas votre écran

Espace de conversation où vous pouvez poser vos questions

Conversation de la réunion

Joel CREST (Invité) 20:28  
Je suis tenté de rejoindre ... mais non

20:29  
Lol

Répondre

# Les news communautaires

Philippe GEIGER





## PARIS



Isabelle



Jean-Pierre



Tristan

## LILLE



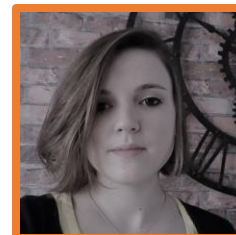
Guillaume

## STRASBOURG

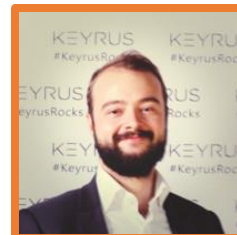


Philippe

## LYON



Marie



Thomas

## NANTES



Paul



Mohamed



## BORDEAUX (COMING SOON)



Gaëlle



Christophe



Yasuteru

## AIX-MARSEILLE



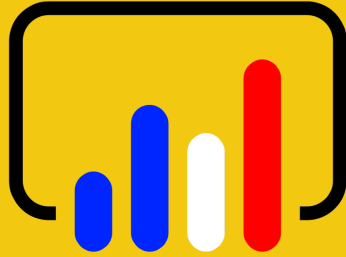
Joël



Franck

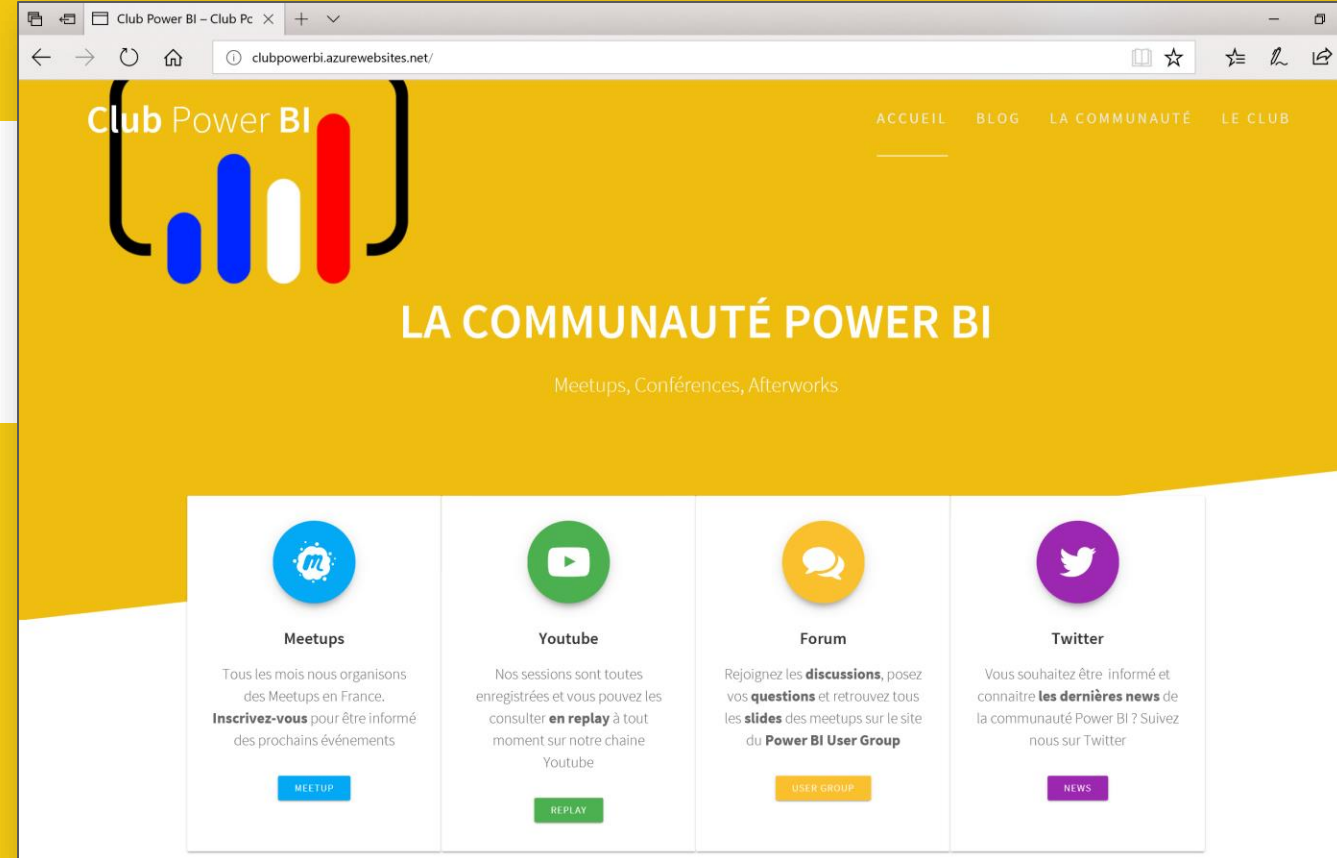
# Club Power BI

L'équipe de la saison 5



# Club Power BI

<http://clubpowerbi.com>



# Les prochains Meetups

## Online

**15 mars**

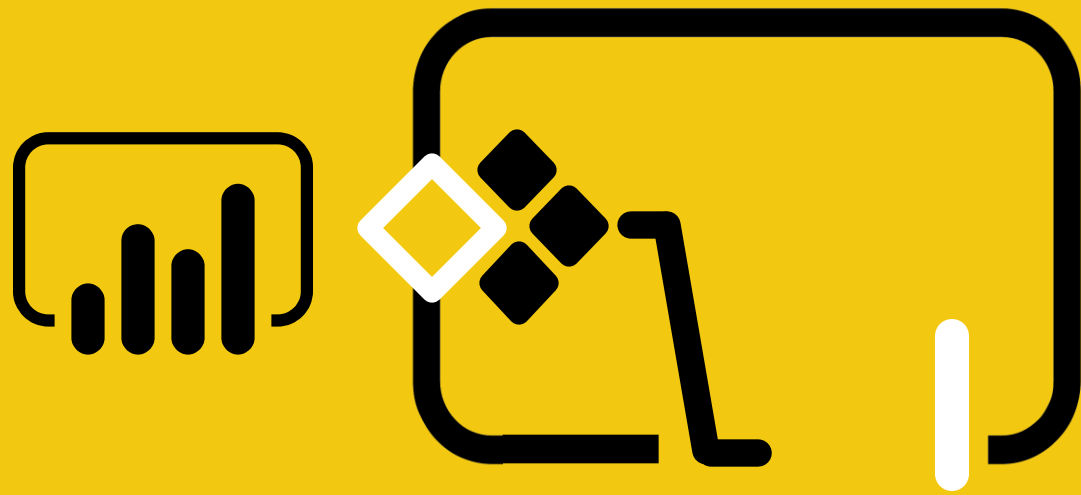
DAX & Fonctions  
Temporelles

Jean-Pierre &  
Tristan + Mathieu

**22 mars**

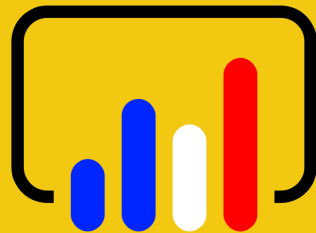
Les calculations groupées -  
Pour qui et quand ?

Guillaume



# Power Saturday

**5 juin 2021 - Online**



Une conférence des communautés Club Power BI, GUSS, aMS

Un succès depuis  
plusieurs années.

Une nouvelle  
édition Online en  
**2021.**



- **700** participants **connectés** en 2020
- **45** conférenciers internationaux
- Des animations pour un événement atypique
- **40%** d'utilisateurs métiers

Les 3 premières éditions du Power Saturday ont été un succès. Notre ambition est de renouveler cette réussite pour cette 4<sup>ème</sup> édition 2021 toujours en version Online



# L'événement

**5 juin 2021**

**900 personnes attendues**



- **6 tracks**
- **40 sessions**
- **40 conférenciers**

[Power Saturday 2021: Call for Speakers/Papers @ Sessionize.com](#)

# Les nouveautés Power BI

Philippe GEIGER



# Agenda

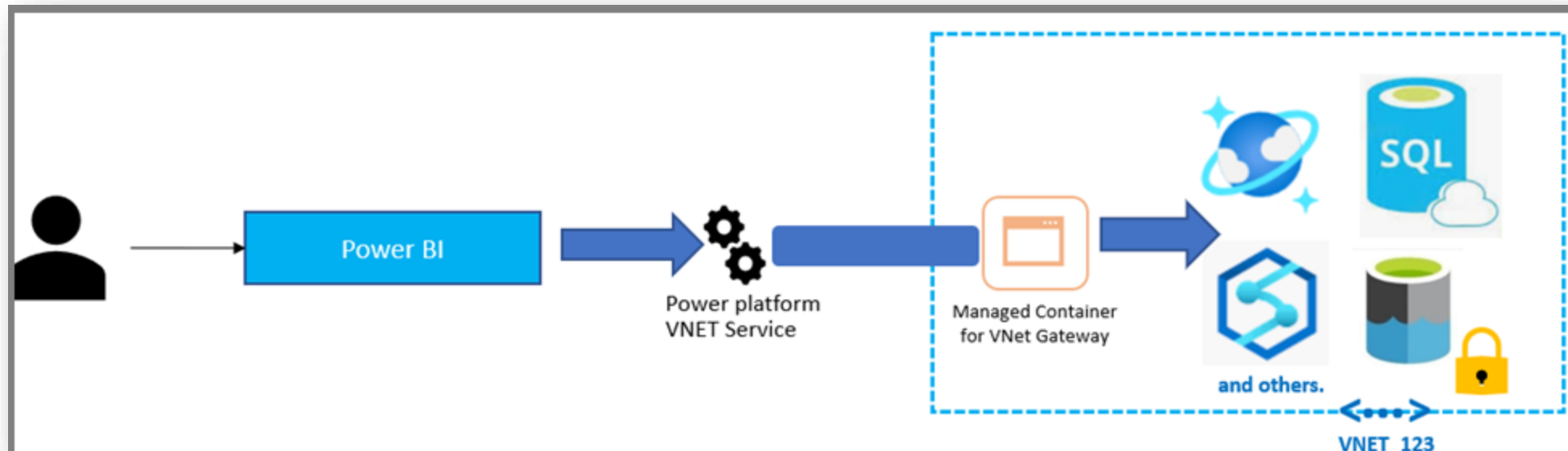
- Sécurité au niveau de l'objet dans Power BI (Preview)
- Connectivité VNet pour les ensembles de données (Preview)
- Power BI Premium par utilisateur
- Application d'analyse de l'utilisation de Power BI Premium Gen 2
- Mise à l'échelle de Power BI Premium Gen 2 (Preview)
- Power BI Embedded Analytics Playground (Preview)

# Sécurité au niveau de l'objet dans Power BI (Preview)

- La sécurité au niveau de l'objet (OLS) permet aux auteurs de modèles de sécuriser des tables ou des colonnes spécifiques du rapport.
- Du point de vue de l'utilisateur, la table ou la colonne n'existe tout simplement pas.
- Actuellement, les définitions OLS ne sont pas créées nativement dans Power BI Desktop, mais via des outils externes comme Tabular Editor
- Différent de RLS
- Référence : <https://powerbi.microsoft.com/en-us/blog/object-level-security-ols-now-available-for-public-preview-in-power-bi-premium/>

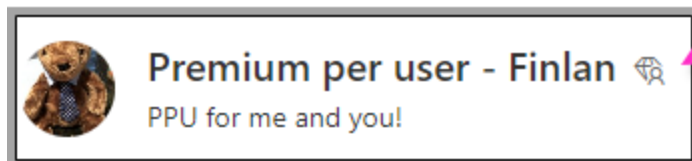
# Connectivité VNet pour les ensembles de données (Preview)

- Connectivité des services Azure (Cosmos DB, SQL Database, Synapse Analytics, etc) via VNet
- Référence : <https://docs.microsoft.com/en-us/data-integration/vnet/overview>



# Power BI Premium par utilisateur

- Tarif : 16,90 euros par utilisateur et par mois
- GA : 2 avril 2021



## Power BI Premium Feature Comparison

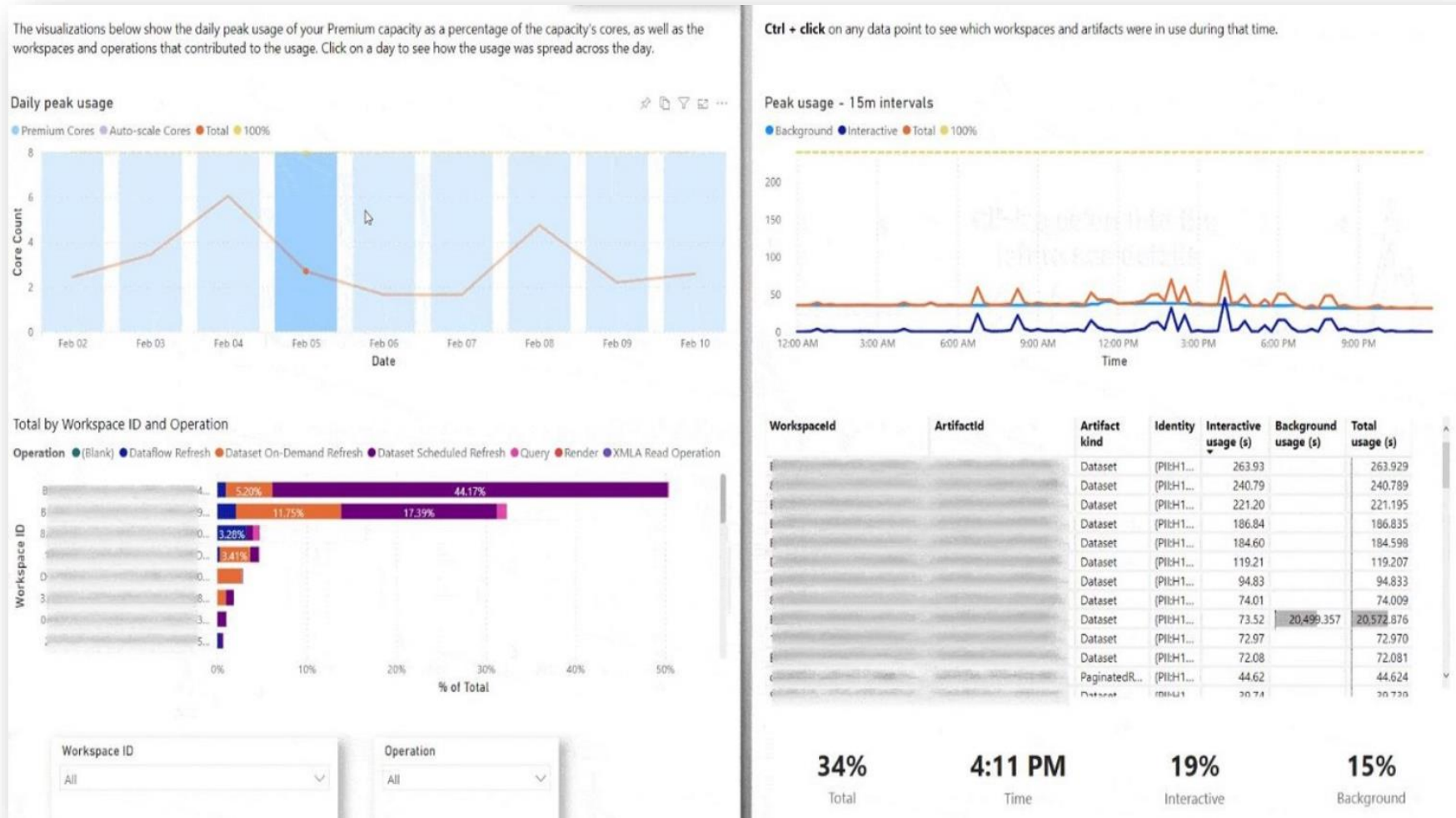
	Premium	
	Per User	Per Capacity
Model Size Limit	10GB**	10GB**
Refresh rate	48/day	48/day
Paginated Reports	✓	✓
AI Capabilities (AutoML, Impact Analysis, Cognitive Services)	✓	✓
Advanced Dataflows Features (Direct Query, etc.)	✓	✓
Usage-based Aggregate Optimization	✓	✓
Application Lifecycle Management (Deployment Pipelines)	✓	✓
XMLA Endpoint Connectivity	✓	✓
Enhanced automatic page refresh	✓	✓
Multi-Geo Support	×	✓
Unlimited Distribution	×	✓
Power BI reports on-premises	×	✓
Bring Your Own Key	×	✓

\*\* Using preview of large models provides higher model sizes for each SKU level

	Users with this license type can view/access the content			
		Pro	Premium Per User	Premium Per Capacity User (no named license)
If you create/share Power BI content in the Power BI service in this type of workspace/app	Pro	Yes	Yes	No
	Premium Per User	No	Yes	No
	Premium Per Capacity	Yes	Yes	Yes

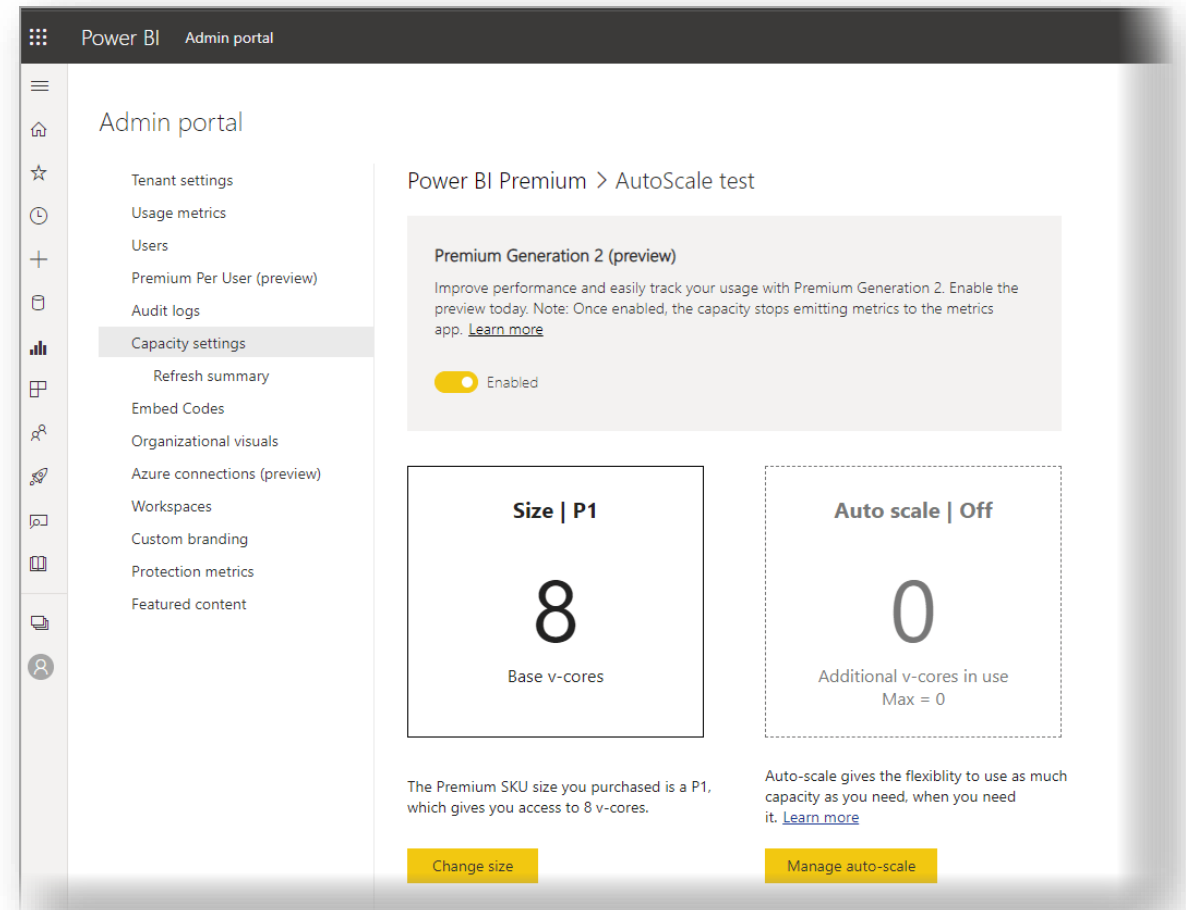
Référence : <https://docs.microsoft.com/en-us/power-bi/admin/service-premium-per-user-faq>

# Application d'analyse de l'utilisation de Power BI Premium Gen 2



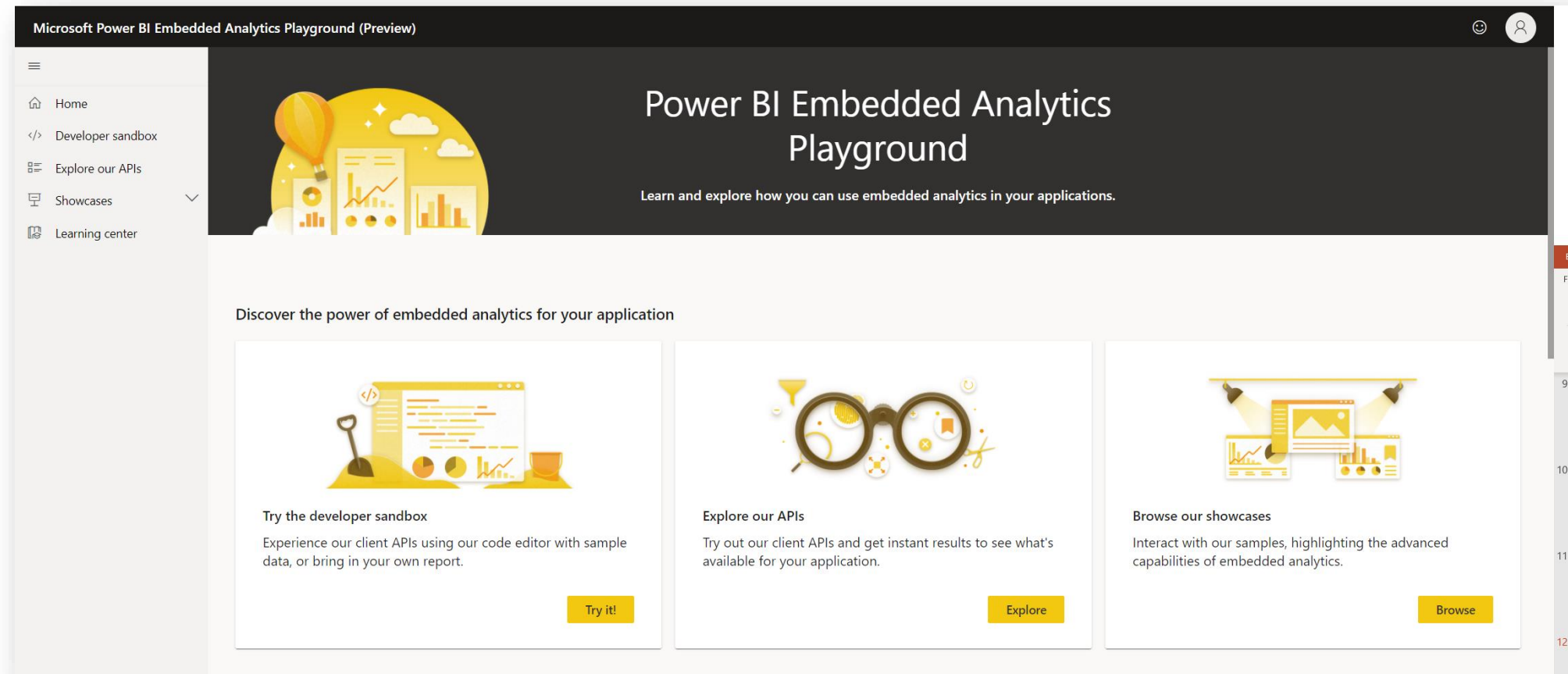
# Mise à l'échelle de Power BI Premium Gen 2 (Preview)

- P1 : 8 V-Cores
- 71,68 euros par Vcore/24 heures
- Possibilité de limiter le nombre de vCores supplémentaires
- Référence : <https://docs.microsoft.com/en-us/power-bi/admin/service-premium-auto-scale>





# Power BI Embedded Analytics Playground (preview)



<https://playground.powerbi.com/>

# Retour d'expérience Azure+Power BI

## arte

Paul PETON | Georges BOUTROS | Andra CIACIRU | Pierre ROLLINGER

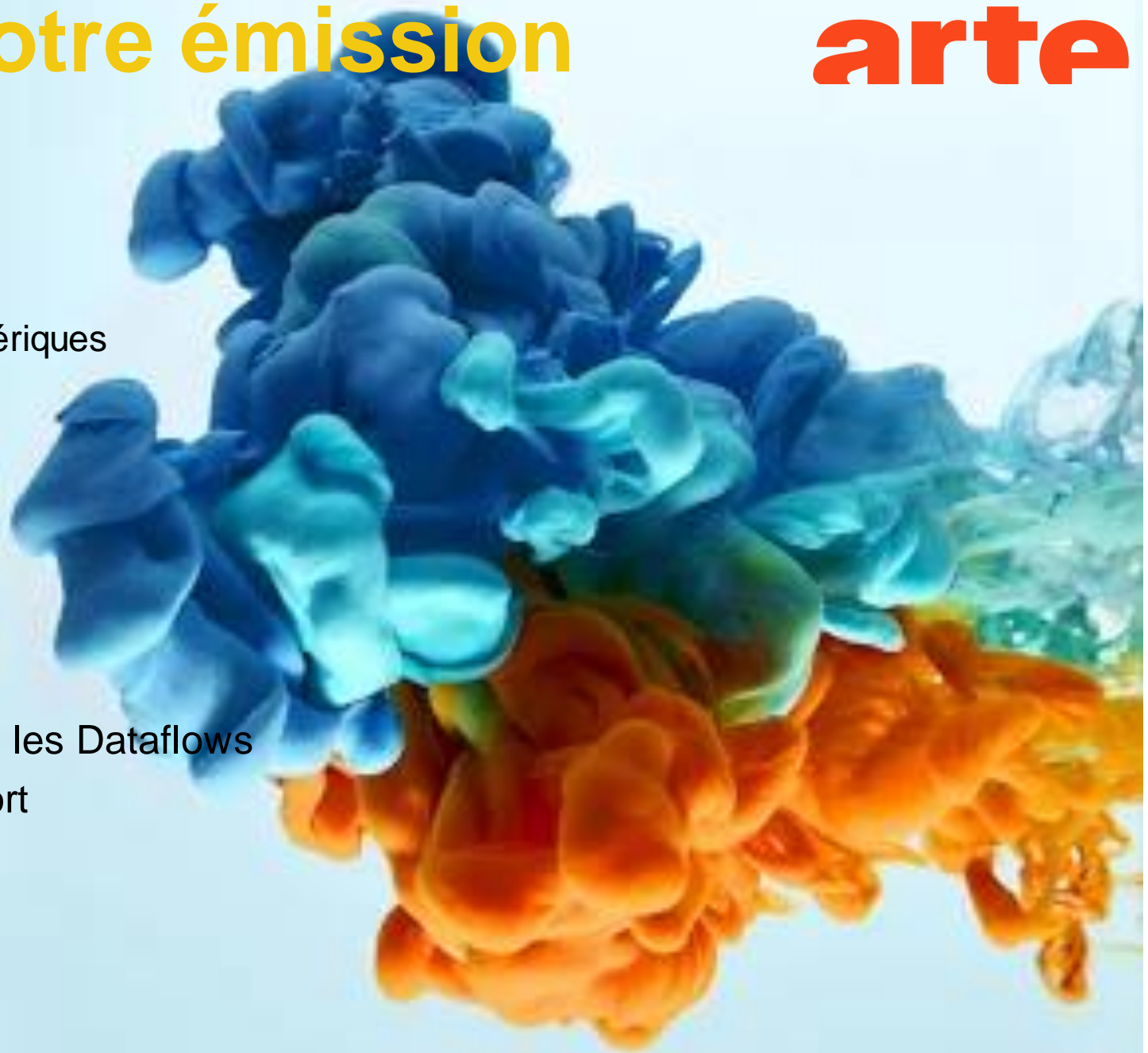


# Au sommaire de notre émission

arte

arte

- ❖ « 28 minutes »
  - Contexte Arte
    - Le rapport des audiences numériques
  - Architecture Azure
    - Focus Data Factory
  - Interroger le Data Lake ?
    - 4 scénarios
- ❖ « Thema »
  - Actualisation incrémentielle avec les Dataflows
  - Séparer jeu de données et rapport
  - Pipeline de déploiement



**Le contexte  
Arte G.E.I.E.**





- ❖ Données liées à la diffusion de nos contenus numériques sur de multiples environnements
  - données de trafic (site et application)
  - réseaux sociaux
- ❖ Cahier des charges :
  - Mise en place d'un *dashboard* automatisé qui agrège toutes nos données
  - Interface consultable sur PC, tablettes et smartphones (*responsive*)
  - Possibilité de datavisualisation
  - Possibilité de comparaison entre deux périodes
    - Comparaison avec une autre période spécifique
    - Comparaison avec la même période mois-1, semaine-1 et jour-1
- ❖ Profils d'utilisateurs : direction, marketing, audiences, partenaires (ARD, ZDF...)

# Les sources de données

arte

## ❖ Des bases de références « *on premises* »

- Les *metadata* de tous nos programmes

ORACLE®

## ❖ Les mesures d'audience (eStat)

- Médiamétrie
- Molotov
- MyCanal
- Free



## ❖ Les médias sociaux

- Facebook (40 pages)
- YouTube (42 chaînes)
- Instagram (7 comptes)

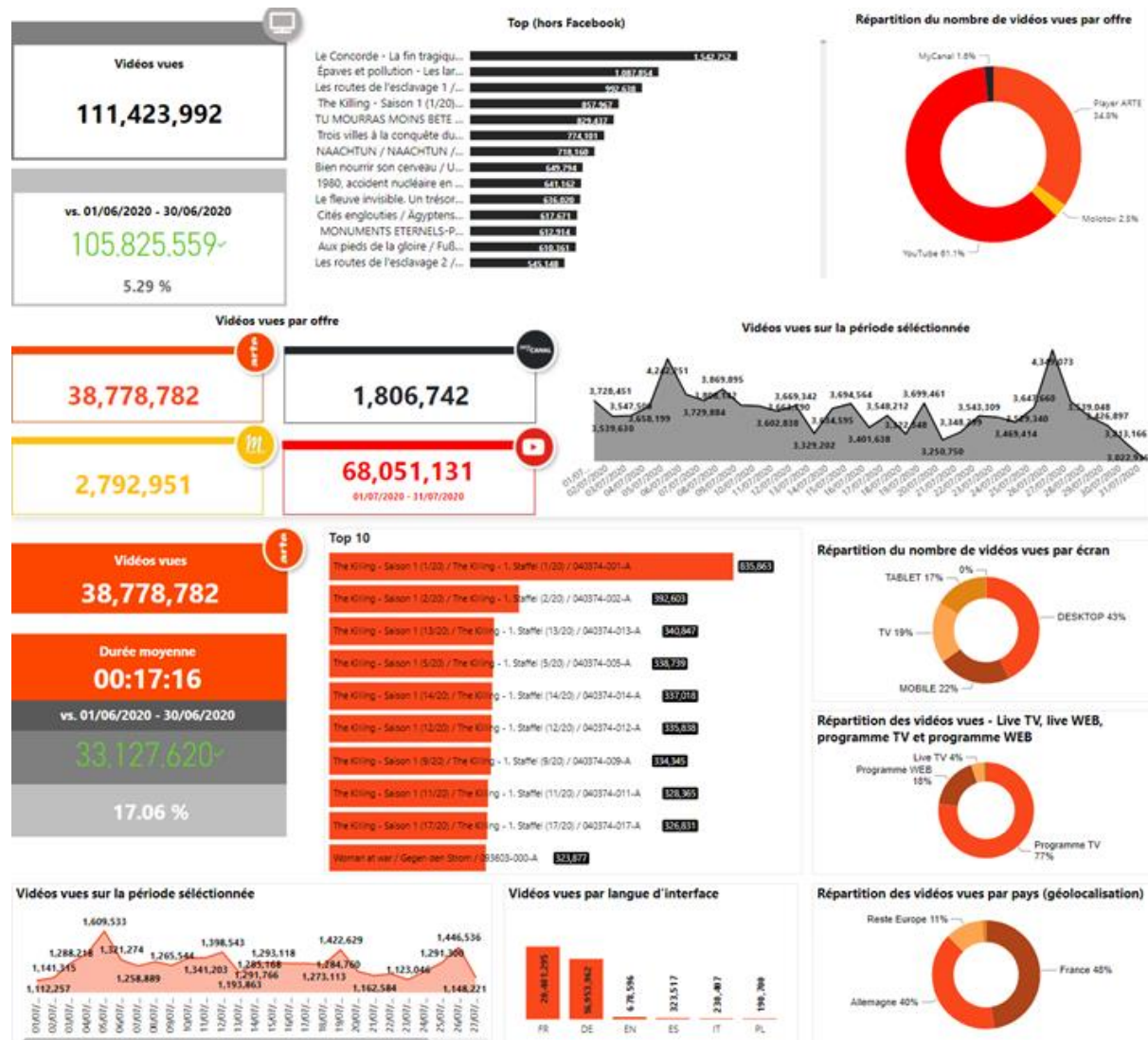


# Rapport des audiences numériques

arte



# Rapport des audiences numériques





# Rapport des audiences numériques

arte

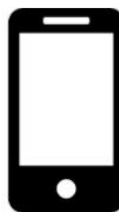
arte



VERSION  
BUREAU  
(PC)

AUDIENCES  
NUMÉRIQUES  
ONLINE-QUOTEN

Cliquer ici !



VERSION  
MOBILE

AUDIENCES  
NUMÉRIQUES  
ONLINE-QUOTEN

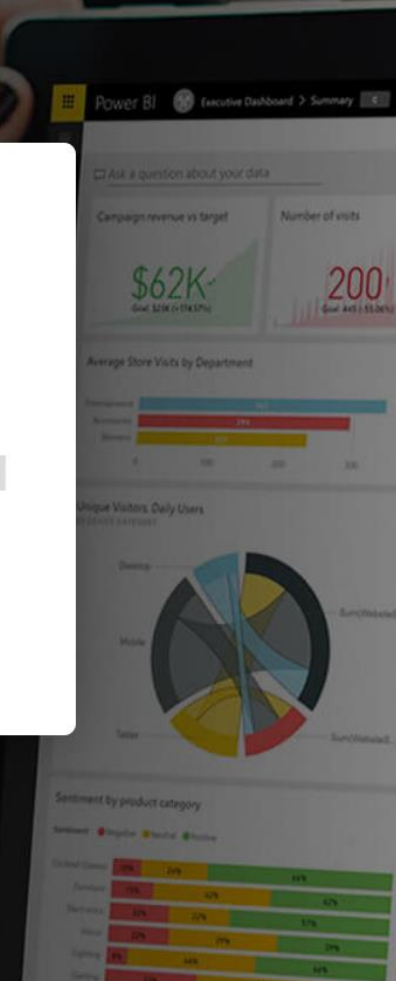
Cliquer ici !



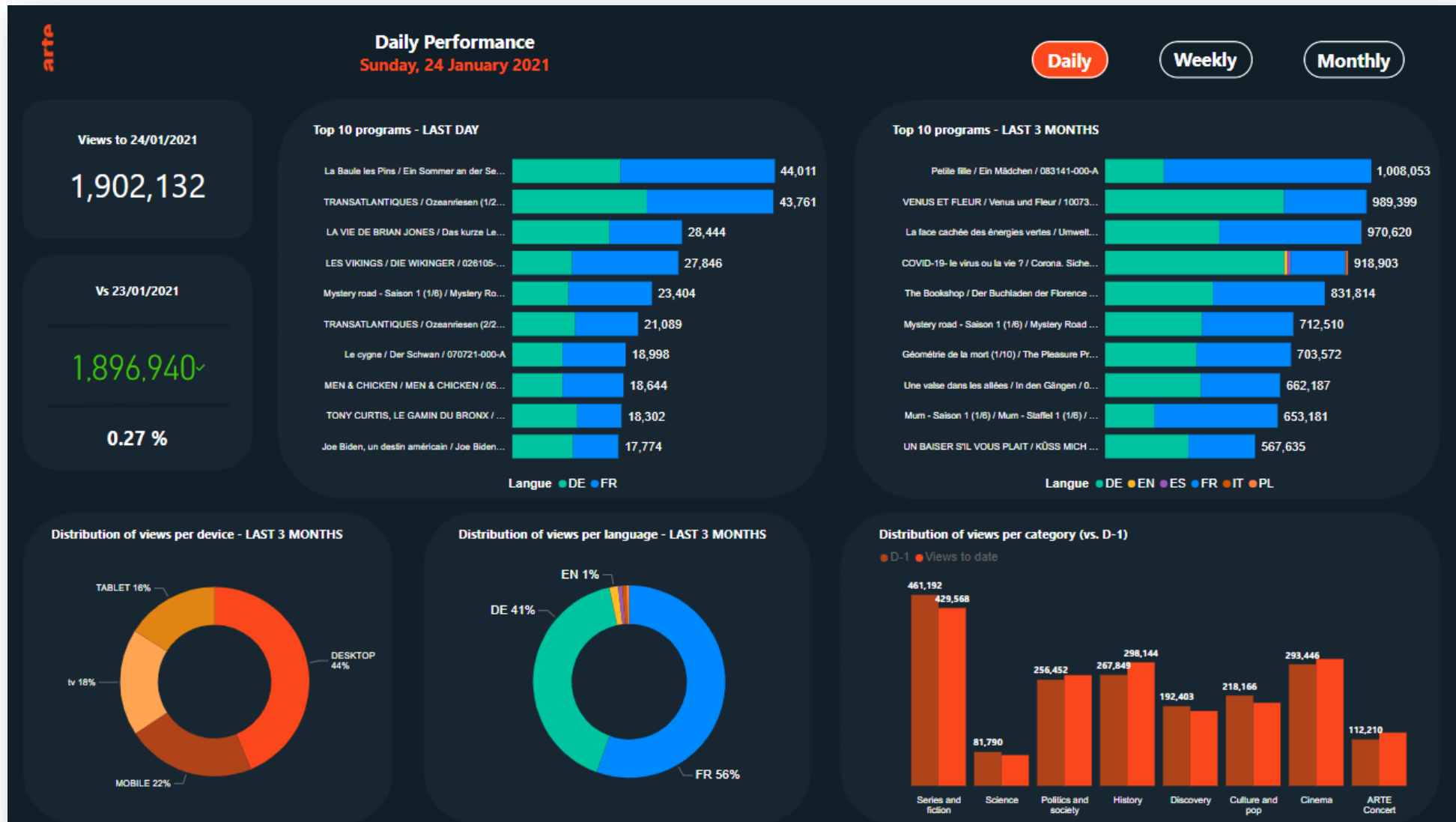
SUPPORT /  
DOCUMENTATION



AUDIENCES NUMÉRIQUES  
ONLINE-QUOTEN



# Audiences – chiffres clés



# Les étapes du projet Azure



# Les étapes du projet Azure



## Phase 1

- Un seul environnement (entre bac à sable et production)
  - Centré autour du Data Lake
- Exposition des premiers rapports

## Phase 2

- Environnements séparés et stratégie CI/CD
- Génération et audit de logs
- Sécurité réseau
- Stratégie de backup des données

## Phase 3

- Supervision des coûts et optimisations
- Data cataloging



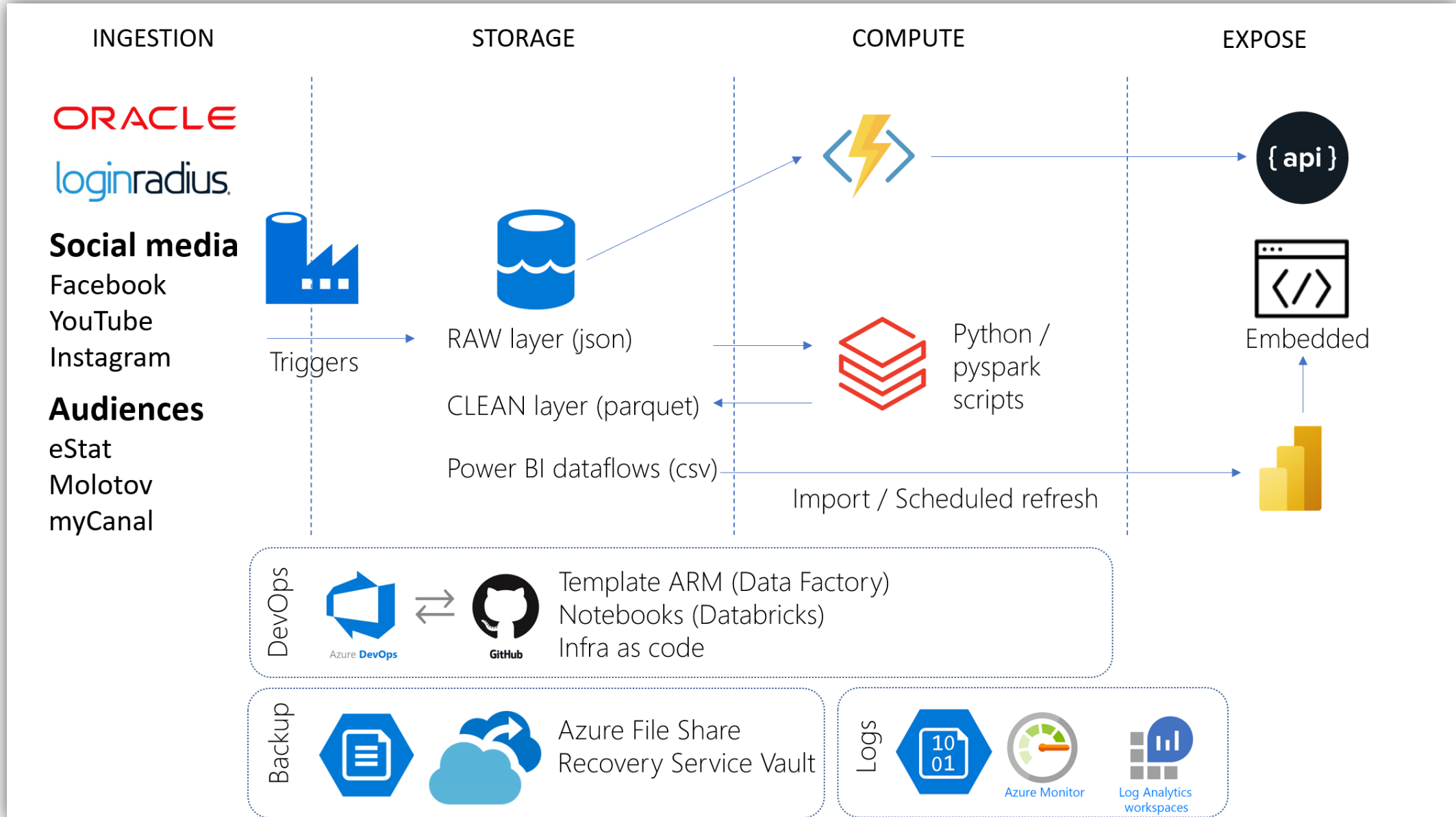
# 'approche « Lakehouse »

- Les données sont archivées dans le *data lake* dans plusieurs couches
  - RAW : les données brutes (sans transformation) telles qu'elles sont transmises
  - CLEAN : les données nettoyées et enregistrées dans un format optimisé (parquet)
  - AGG : les données agrégées selon le besoin de reporting



- La BI peut tirer profit des couches CLEAN et AGG
- La Data Science utilisera les couches RAW ou CLEAN selon le besoin

# Architecture Azure



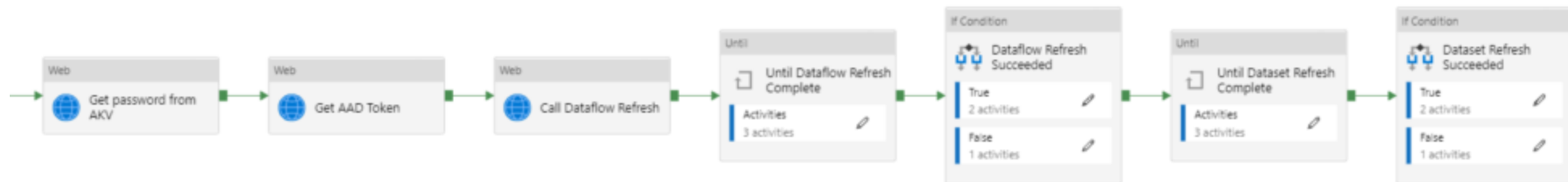


# Enchainer les étapes

❖ Avec Azure Data Factory et l'API REST de Power BI



Déclencher les actualisations d'un *dataflow* puis d'un *dataset* Power BI



<http://methodidacte.org/2021/01/azure-data-factory-pilote-les-actualisations-power-bi/>



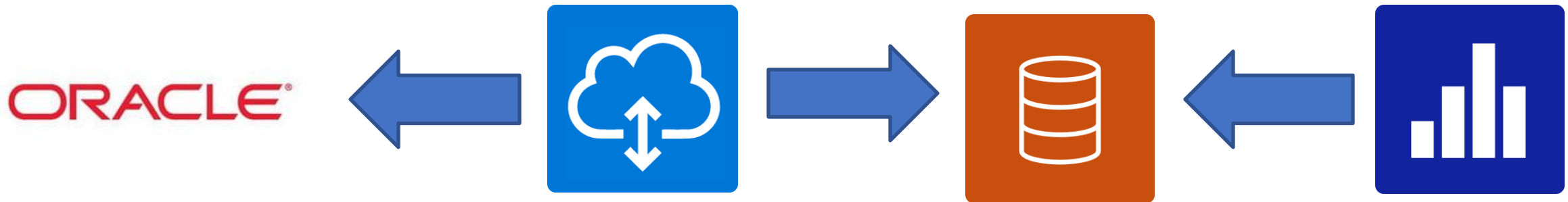
Du Data Lake  
au dataset...





# Dataset version 1

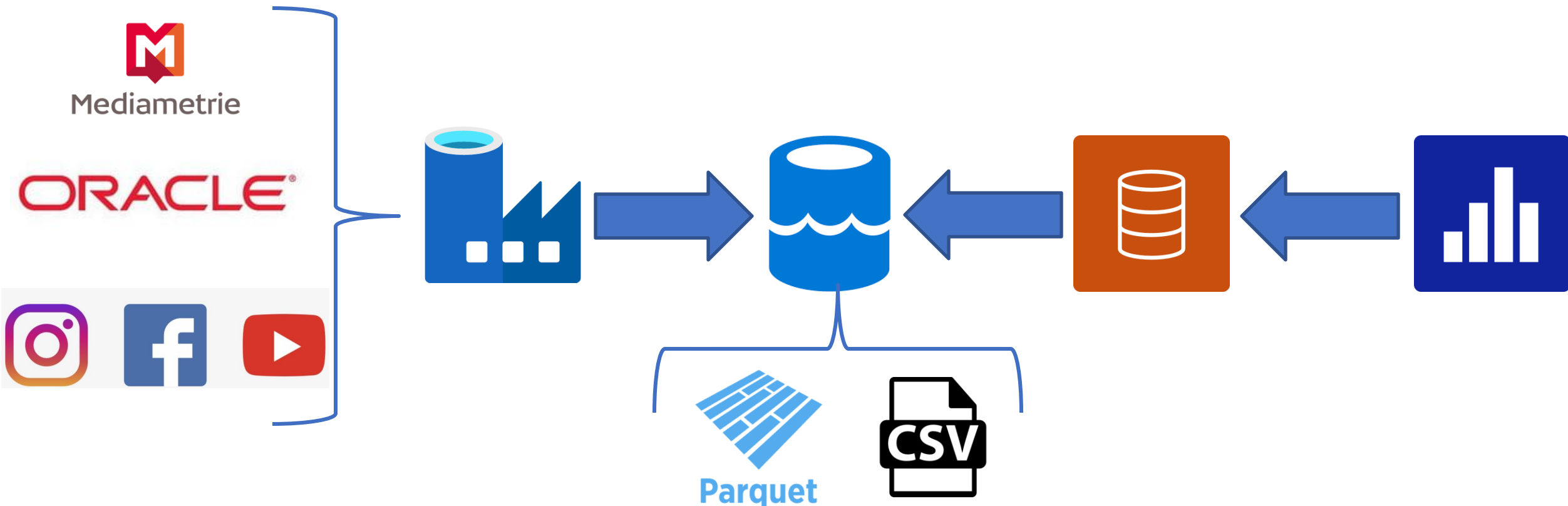
- ❖ Connexion à la base Oracle *on-premises*
- ❖ Nécessité d'une passerelle de données
  - *data gateway* sur une VM locale



# Dataset version 2

## ❖ Connexion aux fichiers du Data Lake

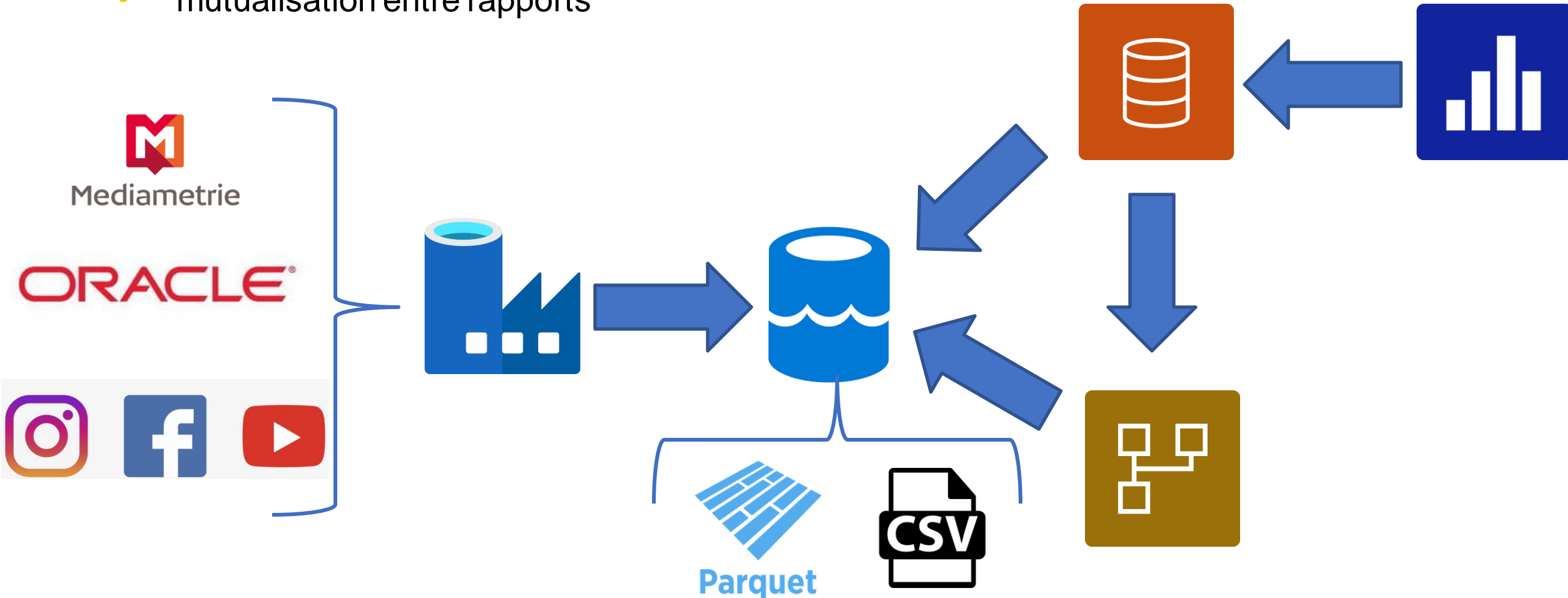
- Pas de connecteur Parquet à l'époque (CSV uniquement)
- Actualisation depuis le portail Power BI, pas d'actualisation incrémentielle
- Pas de mutualisation entre rapports



# Dataset version 3

## ❖ Création d'un *dataflow*

- actualisation incrémentielle MAIS pas au niveau du dataset
- mutualisation entre rapports



# Incremental refresh dans un dataflow **arte**

- ❖ Sur source Azure Blob Storage
- ❖ Idée : limiter les fichiers sur leur date de création
- ❖ Astuce :
  - Ajouter une actualisation incrémentielle
  - La déplacer au moment du listing des fichiers
    - Avant la combinaison des fichiers

The screenshot displays the Arte Dataflow interface. The main window shows a query with the following SQL logic:

```
Table.SelectRows("#Changed column type", each DateTime.From([DateFile]) >= RangeStart and DateTime.From([DateFile]) < RangeEnd)
```

The query results are displayed in a table with columns: Content, Name, Extension, Date accessed, Date modified, Date created, Attributes, and Folder Path. The results show a list of files with their creation dates and paths.

On the right, the 'Query settings' panel is visible, showing the following configuration:

- Properties:**
  - Name: fs-clicktracking
  - Entity type: Custom
- Applied steps:**
  - Source
  - Navigation
  - Lignes filtrées
  - Fichiers masqués filtrés
  - AddDate
  - Changed column type
  - Filter
  - Appeler une fonction p...
  - Colonnes renommées
  - Autres colonnes suppli...
  - Colonne de table dével...
  - Type de colonne changé
  - Colonnes supprimées
  - Type de colonne chang...
  - fs-clicktracking-446174...

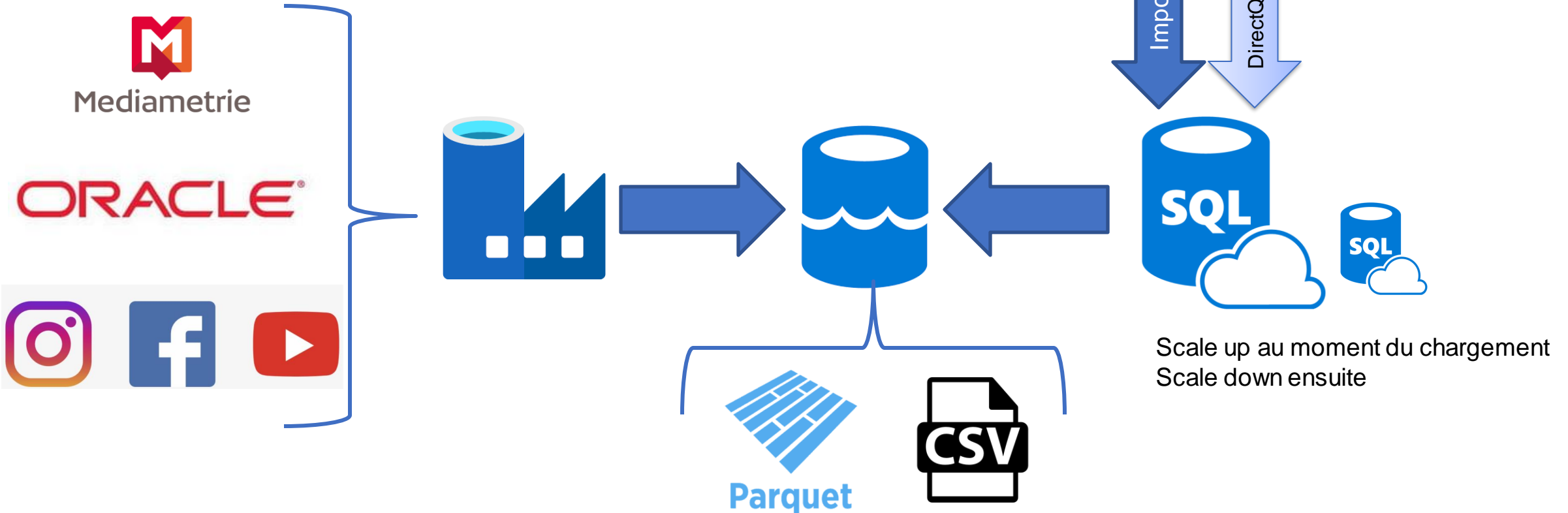
At the bottom, the status bar indicates 'Completed (1.85 s) Columns: 9 Rows: 99+'. The 'Save & close' button is highlighted in yellow.

# Dataset version 4

arte

## ❖ Utilisation d'un serveur Azure SQL DB

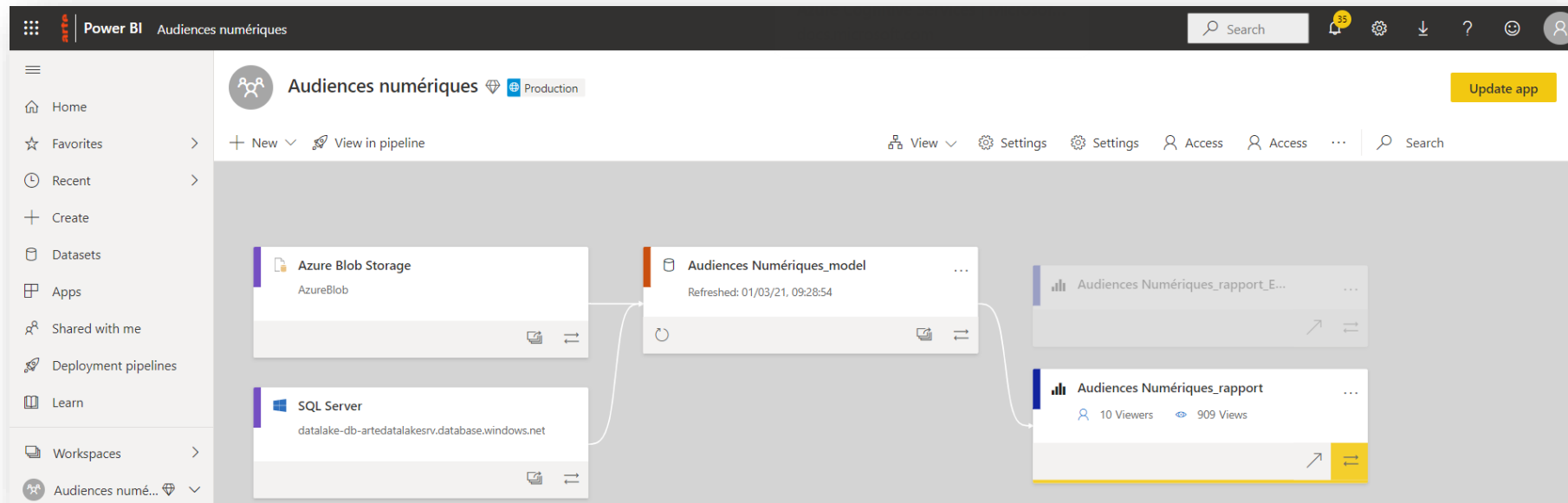
- actualisation incrémentielle
- Possible de faire un dataset en mode hybride :
  - Import sur une période récente
  - (DirectQuery sur l'historique)



# Séparer le jeu de données du rapport **arte**

## ❖ Avantages

- Un jeu de données centralisé
- Plusieurs rapports sans dupliquer le modèle
  - gain en volumétrie sur le service
  - Moins de maintenance
- Simplification du déploiement par pipeline







**Les pipelines de déploiement**



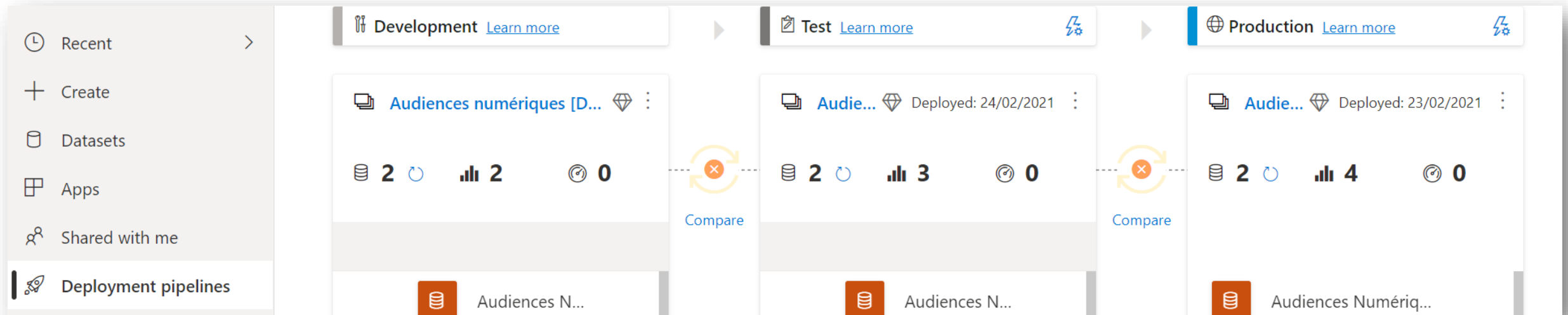
# Les pipelines de déploiement

## ❖ Avantages

- Droits d'accès dissociés
- Paramétrables
  - Ex. : source de données
- Comparaison et « déploiement arrière »

## ❖ Limites

- Ne prennent pas en compte les dataflows
- Augmentent (x3) le nombre d'espaces de travail
- Uniquement en Premium





# Les évolutions à venir

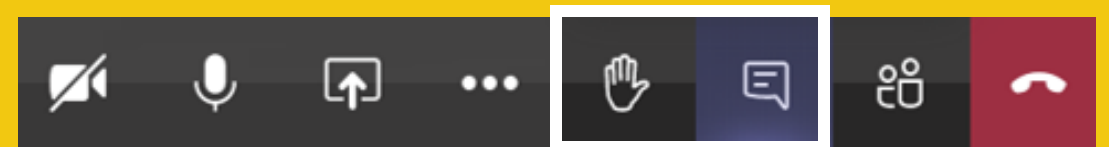


- ⇒ améliorer les performances, offrir de nouvelles fonctionnalités
- ⇒ réduire les coûts

- Tester le connecteur Parquet
  - <https://methodidacte.org/2021/01/poser-du-parquet-dans-le-cloud-azure-oui/>
- Exploiter le DirectQuery sur les dataflows Power BI
  - Bénéficier des performances de Power BI Premium Gen2
- Disposer de l'IP du nœud Premium pour ajout au *firewall*
- Exploiter le *data discovery* et l'obtention des linéages avec Azure Purview ?
- Utiliser le *soft delete* ou Azure Recovery Vault Services sur le Data Lake
  - Sans passer par une copie coûteuse sur Azure File Share

# Q&A

Et maintenant, échangeons. Vous pouvez poser vos questions dans le chat ou demander la parole



Levez la main pour demander la parole.  
Ouvrir la fenêtre de discussion

# Les prochains Meetups

## Online

**15 mars**

DAX & Fonctions  
Temporelles  
Jean-Pierre &  
Tristan

**22 mars**

Les calculations groupées -  
Pour qui et quand ?  
Guillaume

**Merci !**