omega TECH-DZ

OMEGATECH DZ
Support de formation professionnelle

NIVEAU AVANCÉ

Power BI Avancé : modèles d'entreprise, DAX expert, performance, gouvernance

Moteur VertiPaq · modélisation avancée · DAX expert · optimisation · M avancé · ALM. Pour industrialiser vos solutions de BI.

SECTEUR DU JEU DE DONNÉES
🛒 **Distribution Électro & Multimédia**

Durée	Public	Prérequis
3 jours · 21 h	**Analystes seniors, architectes BI, responsables de plateformes**	**Niveau Intermédiaire validé**

Bienvenue : de concepteur à architecte de modèles

Vos objectifs en 3 jours

- Comprendre le **moteur** (VertiPaq, Formula / Storage Engine) pour concevoir juste dès le départ
- Modéliser des cas **complexes** : bridge, SCD, parent-enfant, multi-devises, groupes de calcul
- Écrire du **DAX expert** : transition de contexte, TREATAS, fenêtres, Pareto, cohortes
- **Mesurer et optimiser** : Performance Analyzer, DAX Studio, VertiPaq Analyzer, BPA, incrémental, agrégations
- **Industrialiser** : M avancé, pipelines, Git / PBIP / TMDL, gouvernance

PRÉREQUIS

Prérequis : niveau Intermédiaire validé (Power Query, modèle en étoile, CALCULATE, Time Intelligence).

Le fil conducteur

Connecter → Nettoyer → Modéliser → Visualiser → Publier

CNMVP

Chaque **N**ouveau **M**odèle **V**ise la **P**erfection — ici, à l'échelle de l'entreprise.

PRÉPARATION

Les ateliers de performance supposent un modèle volumineux : générer les données avec `Outils\generer_donnees.py --echelle 20` (≈ 480 000 lignes de transactions).

Outils à installer pour ce niveau

Boîte à outils

Outil	Usage	Coût
DAX Studio	Requêtes DAX, Server Timings, VertiPaq Analyzer	Gratuit
Tabular Editor 2	Édition du modèle, Best Practice Analyzer, scripts	Gratuit (v2)
ALM Toolkit	Comparer / fusionner des modèles	Gratuit
VS Code + extension TMDL	Éditer les fichiers .tmdl du projet	Gratuit
Git	Versionner le projet .pbip	Gratuit
PowerShell + MicrosoftPowerBIMgmt	Automatiser le Service	Gratuit

⚠️ **PRÉVOYEZ LA DSI**

Installez les outils **avant** la formation : les restrictions de sécurité des postes (App Control, antivirus) bloquent parfois les outils externes.

🧠 **À RETENIR EN 10 SECONDES**

Outils externes : onglet « Outils externes » de Desktop (apparaît après installation).

Planning des 3 jours (21 h)

Jour 1 — Architecturer

Créneau	Module
J1 matin	Accueil · **1** Architecture et moteur (1 h 15) · **2** Modélisation avancée (début)
J1 après-midi	**2** suite (3 h 30 au total) : bridge, SCD, parent-enfant, devises, calc groups, RLS dynamique

Jour 2 — Calculer

Créneau	Module
J2 matin	**3** DAX avancé (6 h) — contexte, TREATAS, tables virtuelles
J2 après-midi	**3** suite : fenêtres, Pareto, cohortes, calendriers perso

Jour 3 — Optimiser et industrialiser

Créneau	Module
J3 matin	**4** Performance (3 h 30) · **5** Power Query avancé / M (2 h 30)
J3 après-midi	**6** Gouvernance & ALM (1 h 45) · **7** Design avancé (1 h) · **8** Capstone (1 h 30)

TOTAL

Durées nettes : M1 1 h 15 · M2 3 h 30 · M3 6 h · M4 3 h 30 · M5 2 h 30 · M6 1 h 45 · M7 1 h · M8 1 h 30 = **21 h**.

Sommaire

01 Architecture et moteur analytique 8

02 Modélisation avancée 16

03 DAX avancé 29

04 Performance : mesurer, comprendre, optimiser 46

05 Power Query avancé et langage M 56

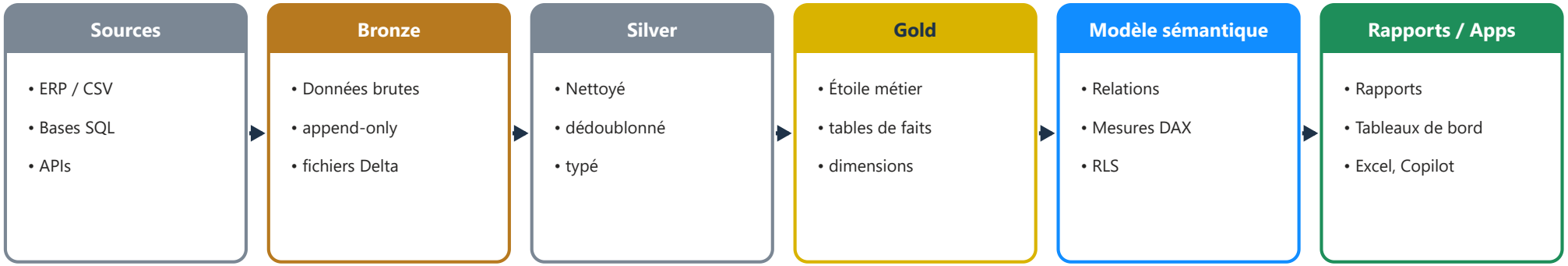
06 Gouvernance, ALM et industrialisation 65

07 Design avancé et storytelling d'entreprise 73

08 Projet capstone et examen blanc 78

Panorama d'une architecture analytique d'entreprise

Architecture analytique moderne : de la source au rapport (couches « médaillon »)



Où fait-on quoi ?

Power Query (Dataflow Gen2) ou notebooks / pipelines : Bronze → Silver → Gold · Power BI : modèle sémantique + rapports sur le Gold

Plus la transformation est lourde, plus elle doit être faite EN AMONT et partagée (une seule vérité).

Référence : « Power BI Architecture End-to-End » (M. Coates) — composants : outils de création, modes de stockage, Service, passerelle, capacités, dataflows.

Les noms de produits Microsoft Fabric évoluent : vérifiez la terminologie et les licences en vigueur.

Du fichier source au rapport : où fait-on quoi ?

Lire les couches

- **Bronze** : on garde la donnée telle que reçue (traçabilité, retour arrière).
- **Silver** : nettoyée, typée, dédoublonnée — une version « propre » par entité.
- **Gold** : modélisée pour le métier (faits, dimensions) — c'est elle que consomme Power BI.

À RETENIR EN 10 SECONDES

Plus tôt on corrige, moins cher c'est.
Power BI excelle sur le Gold : n'y mettez pas les 10 chantiers de qualité de données.

VertiPaq : stocker en colonnes et compresser

VertiPaq : stockage en colonnes, compressé par dictionnaire et répétitions

Stockage en lignes (classique)

P012	Alger	1	Validée
P012	Alger	1	Validée
P012	Alger	1	Validée
P012	Alger	1	Validée
P012	Alger	1	Validée

On lit TOUTE la ligne pour sommer une colonne

Stockage en colonnes (VertiPaq)

Statut	Wilaya	Quantite
Validée	Alger	1
Validée	Alger	2
Validée	Alger	1
Annulée	Alger	3
Validée	Oran	1

Seules les colonnes utiles sont lues · ≈ 1 million de lignes par segment · compression ×5 à ×

Trois techniques

- **Dictionnaire** — valeurs distinctes → petits entiers
- **Valeur** — nombres → décalage + moins de bits
- **RLE** — séquences répétées → (valeur, longueur)

Conséquence n°1 : la taille du modèle dépend de la CARDINALITÉ des colonnes (nombre de valeurs distinctes), pas seulement du nombre de lignes.

Un ID_Ligne unique (24 000 valeurs) coûte beaucoup plus qu'un Statut (3 valeurs) ; un horodatage à la seconde plus qu'une date seule.

→ Supprimer les colonnes inutiles, séparer date / heure, arrondir les décimaux, trier le fait pour allonger les répétitions.

La cardinalité détermine la taille du modèle

Que faire avec ça ?

- Colonnes à **faible cardinalité** (Statut, Région) : minuscules, quasi gratuites.
- Colonnes **uniques** (ID_Ligne, horodatage à la seconde) : le dictionnaire devient aussi gros que la colonne.
- Trier le fait sur une colonne répétitive **allonge les séquences RLE** et réduit la taille.

ASTUCE DE PRO

Dans DAX Studio : **Advanced › View Metrics** classe les colonnes par taille. Commencez toujours par les 5 plus grosses.

Deux moteurs : Formula Engine et Storage Engine

Deux moteurs dans le modèle : Formula Engine (FE) et Storage Engine (SE)

Visuel
requête DAX

→

FORMULA ENGINE (FE)

- Mono-thread
- Évalue les formules DAX
- Itérateurs, logique, contexte
- Utilise le cache du SE

xmSQL →

STORAGE ENGINE (SE)

- Multi-thread, très rapide
- Scanne les colonnes VertiPaq
- Agrégations, filtres, jointures
- Mode DirectQuery : requête SQL

Lecture d'un diagnostic (DAX Studio › Server Timings)

SE dominant : modèle / cardinalité / relations → optimiser le modèle.
FE dominant : formule (itérateur imbriqué, FILTER sur table) → réécrire le DAX.

Beaucoup de requêtes SE courtes ?
→ itérateur qui rappelle le SE ligne à ligne
(« callbacks »)

Qui travaille ? Lire un diagnostic de performance

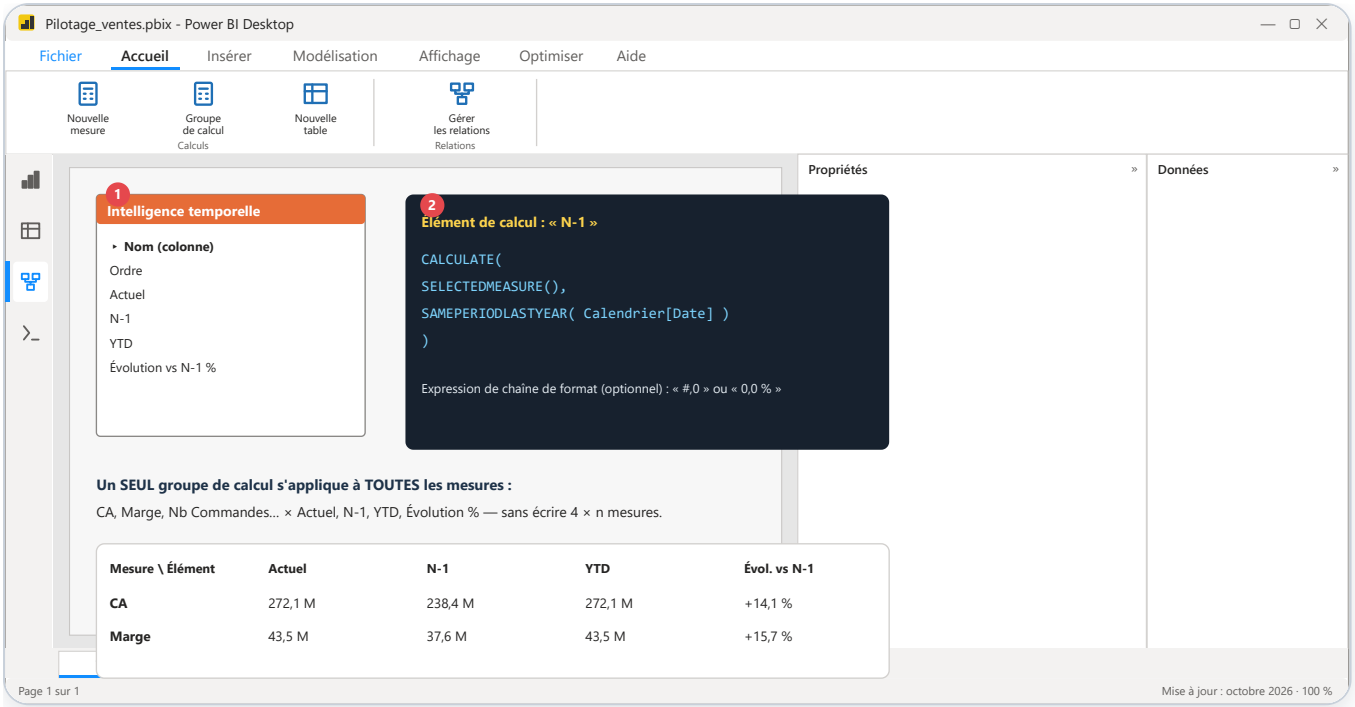
À RETENIR EN 10 SECONDES

FE réfléchit (un seul cœur), SE compte (tous les cœurs). Une bonne requête laisse le SE faire le gros du travail.

POUR ALLER PLUS LOIN

Les requêtes que le FE envoie au SE sont en **xmSQL** (visible dans DAX Studio). Un nombre élevé de petites requêtes signale un itérateur coûteux.

Groupes de calcul : écrire une logique, l'appliquer à toutes les mesures



Vue Modèle : un groupe de calcul « Intelligence temporelle »

Pas à pas

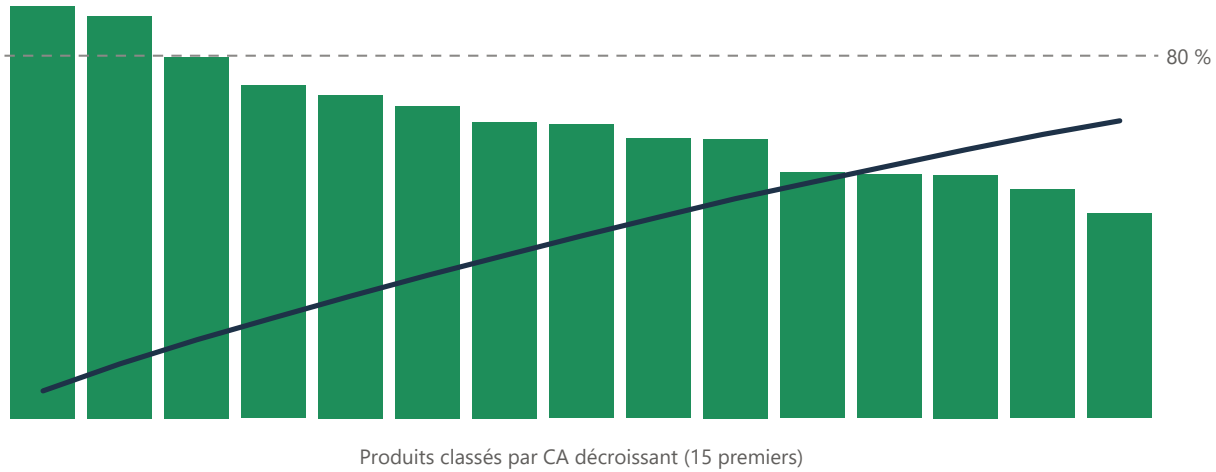
1. Vue Modèle › **Nouveau groupe de calcul** (ou Tabular Editor).
2. Ajouter les éléments : Actuel, N-1, YTD, Évolution %.
3. Chaque élément = une expression CALCULATE autour de `SELECTEDMEASURE()`.
4. Tester dans une matrice : mesures en lignes, éléments en colonnes.

⚠ **EFFET DE BORD**

Un groupe de calcul désactive les **mesures implicites** : toutes les valeurs doivent venir de mesures explicites.

Analyse ABC / Pareto

Analyse ABC (Pareto) : quelques produits portent l'essentiel du chiffre d'affaires



Mesure du cumul % + classe

```
Cumul % CA =
VAR _ca = [CA]
VAR _p = ALLSELECTED( Produits[Nom_Produit] )
VAR _cumul = SUMX( FILTER( _p, [CA] >= _ca ), [CA] )
RETURN DIVIDE( _cumul, SUMX( _p, [CA] ) )

Classe ABC = SWITCH( TRUE(),
[Cumul % CA] <= 0.8, "A", <= 0.95, "B", "C" )
```

- A : ≈ 80 % du CA
- B : les 15 % suivants
- C : la longue traîne

Les produits de classe A portent environ 80 % du CA

Comment l'interpréter ?

- **Lecture** : les barres sont le CA par produit ; la courbe est le cumul en % ; la ligne pointillée marque 80 %.
- **Usage** : rationaliser le catalogue (classe C), sécuriser le stock (classe A), cibler les négociations.
- **ALLSELECTED** : le Pareto se recalcule sur la sélection de l'utilisateur (une région, une année).

À RETENIR EN 10 SECONDES

Cumul % = part du CA des produits au moins aussi vendeurs que celui-ci.

omega
TECH-DZ

FIN DE L'EXTRAIT

Power BI · Avancé

Le support complet compte 84 pages : 8 modules, ateliers corrigés, aide-mémoire.

Il est remis à chaque participant de la formation, avec les jeux de données de son secteur.

Vous avez suivi la formation ou souhaitez le support seul ? Commandez-le en ligne :

https://omegatech-dz.com/site/formation/power-bi-avance