



# OBJECTIFS

## Objectifs pédagogiques

Bienvenue dans ce projet qui va s'étaler sur 3 semaines de formation.

Vous allez avoir l'honneur de travailler sur des données issues de Grenoble Alpes Invest qui n'ont jamais été mises à la disposition du grand public. Ces données ont été rassemblées pour un projet de Base d'entreprises de la région grenobloise et voironnaise.

Vous allez pouvoir travailler sur une problématique assez large de Business Intelligence. Vous présenterez votre travail en fin de module devant les 2 formateurs de ce module et sans doute des invités surprise.

Voilà une trame simplifiée afin que vous puissiez mieux comprendre le déroulé de votre travail.

Nb : une itération équivaut à 4 ou 5 jours de travail soit une semaine de formation.

## Démarche pédagogique

**Itération 1 :** découverte de la stack technique, tutoriel DBT, construction d'un datawarehouse simplifié, création d'un dashboard très simple

**Itération 2 :** Analyse exploratoire des données de Grenoble Alpes Invest, modélisation du datawarehouse, création du projet DBT correspondant, réalisation d'un ou plusieurs dashboards sur PowerBI.

**Itération 3 :** Approfondissement de certaines problématiques avec des ajouts d'outils ou de données, finalisation de vos travaux + préparation de vos présentations orales.

## Compétences développées :

Vous allez être confrontés à une vraie problématique de Business Intelligence. Vous avez déjà été confronté aux données autour des entreprises dans le module précédent d'extraction de données. Si vous n'avez pas eu la possibilité de récupérer l'ensemble des données, on peut vous les mettre à disposition.

Au cours de ce module vous allez apprendre à :

- Analyser des données et en faire émaner des problématiques

- Créer un modèle avec vos données afin d'enrichir un Datawarehouse
- Utiliser DBT
- Transformer des données en SQL
- Créer un ou plusieurs dashboards pérennes afin de répondre aux problématiques choisies.
- Restituer votre travail et l'expliquer intelligiblement aux propriétaires des données analysées.

# MODALITÉS

## Livrables

1 Dashboard PowerBI, 1 support de présentation au format PowerPoint, 1 documentation (format de votre choix).

Une présentation orale en binôme d'une durée de 20 min suivie d'une dizaine de minutes de questions.

## Durée

Module : 3 itérations de 5 jours soit 15 jours au total

Lancement le 30/06/2025 et clôture le 21/07/2025.

## Formateur(s)

Virginie Desharnais

Jules Tison

# TRAME

		Planning		Jour	Activités
30/06	MODULE	Formateur 1	Formateur 1	1	Itération 1
01/07	MODULE	Formateur 1	Formateur 1	2	Itération 1
02/07	MODULE	Formateur 1	Formateur 1	3	Itération 1
03/07	MODULE	autonomie	autonomie	4	Itération 1
04/07	MODULE	Formateur 2	Formateur 2	5	Itération 2
07/07	MODULE	Formateur 2	Formateur 2	6	Itération 2
08/07	MODULE	Formateur 1	Formateur 1	7	Itération 2
09/07	MODULE	Formateur 2	Formateur 2	8	Itération 2
10/07	MODULE	autonomie	autonomie	9	Itération 2
11/07	MODULE	Formateur 2	Formateur 2	10	Itération 3
15/07	MODULE	Formateur 2	Formateur 2	11	Itération 3
16/07	MODULE	autonomie	autonomie	12	Itération 3
17/07	MODULE	Formateur 1	Formateur 1	13	Itération 3
18/07	MODULE	Formateur 1	Formateur 1	14	Itération 3
21/07	MODULE	Formateur 1	Formateur 1		Itération 3

# ITÉRATIONS 2 & 3

## Projet Invest in Grenoble Alpes

### Modalités

- Travail en binôme
- 4 jours en présentiel

### Livrables

Présentation orale et trace écrite de la compréhension des données ainsi que de plusieurs problématiques traitées. Un Datawarehouse (pas forcément finalisé) : 3 schémas différents (G1\_ods, G1\_stg et G1\_core par exemple) avec plusieurs tables par schéma.

### Objectifs

L'objectif de cette itération est d'explorer les données transmis par Invest in Grenoble Alpes, de choisir un ou plusieurs axes d'analyse, de modéliser une version 1 de votre datawarehouse et éventuellement commencer un ou plusieurs dashboards illustrant vos choix d'analyse. Une bonne compréhension des données implique une conscience du potentiel de ces dernières mais aussi leurs limites (d'un point de vue qualitatif et quantitatif), la phase d'analyse exploratoire est donc particulièrement importante.

Une fois les données comprises, votre but va être de créer un datawarehouse modélisé de façon optimale pour la visualisation des résultats d'analyse. Ce datawarehouse doit être réfléchi afin qu'il soit le plus ergonomique possible dans l'injection des données dans PowerBI mais aussi lorsque l'on devra importer de nouvelles données. Afin de construire ce datawarehouse, vous devrez utiliser un outil de transformation de données, DBT découvert en itération 1.

### Compétences

- Comprendre des données ainsi que leur contexte associé.
- Nettoyer les données brutes jusqu'à obtenir un ensemble de fichiers CSV facilement intégrable en base de données.
- Concevoir la modélisation de votre datawarehouse sur papier

- Créer le projet DBT représentant votre datawarehouse
- Utiliser une ou plusieurs fonctions PostgreSQL dans un modèle DBT
- Réaliser un premier dashboard PowerBI

## JOUR 1

### 1.1 – Analyse exploratoire des données

7h – Présentiel

Vous trouverez sur le drive un ensemble de données source dans le dossier data. Téléchargez-les et analysez-les en binôme avec l'outil de votre choix.

Ces données devront être nettoyées et un peu transformées, via un script Python.

Une fois le nettoyage effectué, prenez le temps de comprendre, d'expérimenter, de tester pour mieux réfléchir aux problématiques qui vous intéressent.

Vous trouverez également un document Contexte afin de vous guider dans le choix de votre problématique.

A l'issue de cette analyse, vous devez être capable de formuler des problématiques auprès du formateur. Appelez-nous lorsque vous voulez nous montrer le fruit de vos réflexions.

#### RESSOURCES

- Les fichiers partagés sur le drive
- Le doc « Contexte ».

#### COMPÉTENCES ASSOCIÉES

- Comprendre des données ainsi que leur contexte associé.
- Nettoyer les données brutes via un script Python réutilisable (Attention à la qualité de votre code...)

---

## JOUR 2

### 2.1 – Intégration des données dans votre Base de Données

3h – Présentiel

Vous allez maintenant pouvoir vous connecter à votre BDD.

IIIData met à votre disposition une base de données PostgreSQL par binôme accessible avec les identifiants ci-dessous.

ATTENTION ! Aucune sauvegarde n'a été prévue, donc évitez de supprimer des schémas ou des tables...

Host : 51.159.205.157

Port : 15030

Database name : db\_binome\*\*chiffredubinôme\*\*

User/password :

binome1	%K7XXP<]y??dR2yKS]&3
binome2	rIY>0QzFd;vwlo~mM~m]
binome3	7pWX{)V&MQsWc/mct?-?
binome4	p6LRN,+ORizF+-A/+~jW
binome5	^*7h9,9pNY49hn~nK5-S
binome6	Gzn4LDT;[[dP%(nnx2_*
binome7	w1V3upX&3ROfw,,#K?Y?
binome8	6FE[CBis}KnsjLiMH8%g

Afin de vous connecter, nous vous conseillons d'installer DBeaver community.

Enfin, trouvez le moyen (EN PYTHON) d'injecter vos fichiers sous forme de table dans un schéma que l'on appellera "raw". Prenez soin de vérifier le typage, l'encodage et autre de vos données, une fois les tables créées dans votre BDD.

## RESSOURCES

- DBeaver : <https://dbeaver.io/>

## 2.2 – Modélisation de votre datawarehouse.

4h – Présentiel

Lors de l'itération 1, vous avez découvert la modélisation dimensionnelle en étoile ou en flocon. Vous allez devoir mettre en pratique ces notions afin de construire votre datawarehouse. A l'issue

de cette phase, vous présenterez sur papier les 3 couches de transformation constituant votre datawarehouse (raw, staging, core) et la couche CORE sera modélisée de manière optimale.

#### COMPÉTENCES ASSOCIÉES

- Concevoir la modélisation de votre datawarehouse sur papier

---

## JOUR 3

---

### 3.1 — Réalisation de la modélisation dans un projet DBT

7h — Présentiel

Lors de l'itération 1, vous avez créé un projet DBT simple pour découvrir l'outil : il est maintenant temps d'appliquer ces notions à un projet plus complexe : à vous de jouer !...

L'objectif du jour est clair : le datawarehouse que l'on avait imaginé sur papier doit être figolé et finalisé. Donc la demi-journée est dédiée à cela.

Attention ! Cette journée remplie de PostgreSQL doit s'achever avec l'utilisation d'une ou plusieurs fonctions intégrées à vos modèles DBT. Ne partez pas avant d'avoir appelé le formateur et montré une ou plusieurs fonctions et prouvé leur bon fonctionnement dans vos modèles.

Par ailleurs, gardez en tête que de nouvelles données pourraient venir s'ajouter aux données source, donc pensez développement propre, réutilisable, pro !

#### FACULTATIF

Pour les plus rapides, soyez curieux, étudiez la doc de DBT, creusez du côté des macros...

#### COMPÉTENCES ASSOCIÉES

- Utiliser une ou plusieurs fonctions postgresSQL dans un modèle DBT
- 

## JOUR 4

### 4.1 — Conception et réalisation des dashboards PowerBI

7h — Présentiel

Une fois le datawarehouse finalisé on peut passer à la suite. Ouvrez PowerBI et trouvez un moyen de connecter votre couche CORE du datawarehouse.

Vous pouvez commencer à concevoir et maquetter les dashboards de PowerBI.

**Ce qu'on attend de vous :** vous avez fixé une ou plusieurs problématiques en début de modules.

Le dashboard est censé être le support qui vous permettra d'y répondre. Encore faut-il que tout soit clair dans votre tête pour ne pas faire quelque chose de contre-productif.

Alors on reprend un stylo et on répond aux questions ci-dessous :

- A qui je veux que ce dashboard s'adresse ?
- Quelle problématique est étudiée/approfondie dans ce dashboard ?
- Plusieurs onglets ? Si oui, un onglet = une problématique ? Autre ?
- Sentez-vous libre de brainstormer sur d'autres points qui vous paraissent utiles à anticiper.



A la fin de la journée, vous devez avoir au moins un dashboard fonctionnel répondant à des problématiques simples, il est donc maintenant temps d'aborder des problématiques plus complexes.

Les personnes de Grenoble Invest Alpes souhaitent aussi rapprocher les données terrain et compléter les informations manquantes sur les effectifs, les CA, le statut juridique ainsi que les potentielles levées de fonds. Étudiez des méthodes de prédiction d'effectif par exemple.

Soyez créatifs : de nouveaux axes d'analyse sont toujours les bienvenus !

---

#### COMPÉTENCES ASSOCIÉES :

- Réaliser un premier dashboard PowerBI