

29/05/2021



**Rapport du projet : la réalisation d'un data warehouse et reporting.**

Travail réalisé par : SABRI Abdelilah.

Encadré par : DR HILAL Imane.

Figure 1: structure de OLTP .....	3
Figure 2:la structure en étoile du datawarehouse .....	5
Figure 3:dimension student .....	5
Figure 4:dimension cours .....	6
Figure 5:table de fait .....	6
Figure 6:impact de vle sur le résultat .....	8
Figure 7:le nombre des étudiants inscrits dans chaque cours .....	9
Figure 8:le nombre des étudiants par sexe .....	9
Figure 9:le pourcentage des étudiants handicapés .....	10

# Introduction

---

La BI (Business Intelligence ou DSS pour Decision Support System ou plus simplement Informatique Décisionnelle en français) permet d'extraire des données informatisées afin d'aider les décideurs à prendre des décisions alimentées par des chiffres. Il s'agit donc d'outils d'aide à la décision qui agrègent des données hétérogènes, les rend comparables et exploitables et génèrent des tableaux de bord avec graphiques et rapports.

Ce mini-projet vient pour m'aider à pratiquer tous ce que nous avons vu dans les séances du cours, et maîtriser les différents outils existants dans ce domaine.

Pour mieux profiter de mon travail dans ce mini-projet j'ai mis des objectifs à atteindre quand je le finirai.

- Habituer à travailler avec des jeu de données volumineuses.
- Faire la conception d'un datawarehouse.
- Maîtriser un outil d'intégration des données (ETL).
- Utiliser un outil de reporting .
- Automatiser les taches (exécution des jobs).

# L'exploration de la source OLTP

Learning Analytics se concentre sur la collecte et l'analyse des données des apprenants afin d'améliorer leur expérience d'apprentissage en fournissant des conseils éclairés et d'optimiser le matériel d'apprentissage. Pour soutenir la recherche dans ce domaine, nous avons développé un ensemble de données contenant des données provenant de cours présentés à **l'Open University (OU)**. Ce qui rend l'ensemble de données unique, c'est le fait qu'il contient des données démographiques ainsi que des données agrégées en flux de clics sur les interactions des élèves dans l'environnement d'apprentissage virtuel (VLE). Cela permet d'analyser le comportement des élèves, représentés par leurs actions. L'ensemble de données contient l'information sur **22 cours**, **32 593** étudiants, leurs résultats d'évaluation et les journaux de leurs interactions avec le VLE représentés par des résumés quotidiens des clics des étudiants (**10 655 280** entrées).

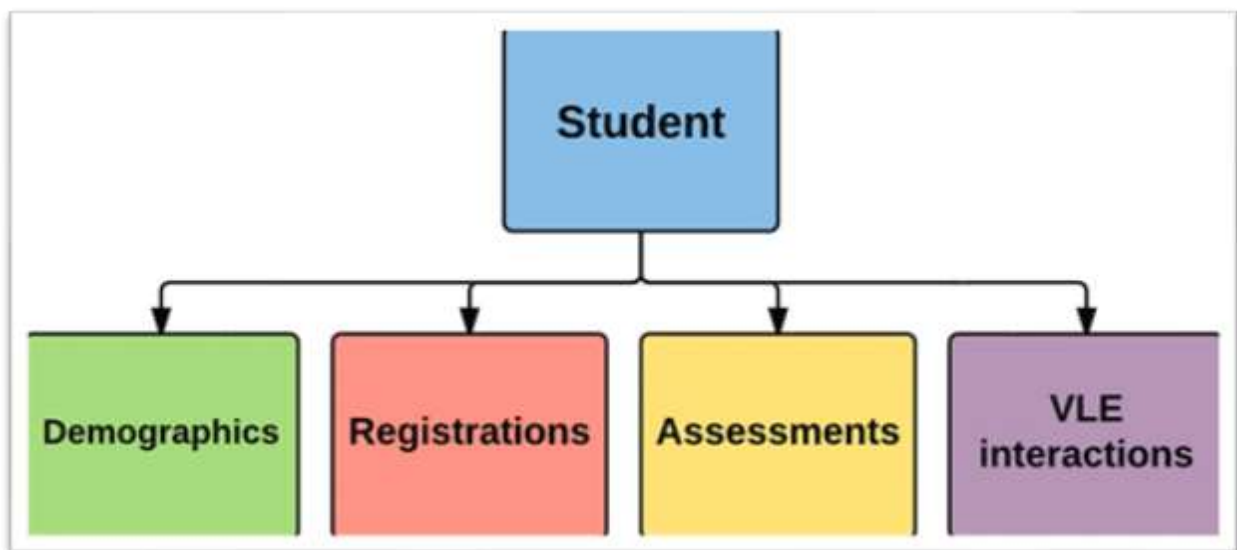


Figure 1: structure de OLTP

# La création de datawarehouse

---

L'étape de la conception et la création d'un datawarehouse dans un projet décisionnel est indispensable et fondamentale ,en effet nous permet de bien répondre au besoin de nos décideurs avec la manière le plus efficace ,pour cela j'ai consacré une partie suffisante du temps dans sa création ,et comme j'ai déjà indiqué le datawarehouse vient pour nous aide au répondre à des questions décisionnelles , et comme notre projet est académique ,j'ai essayé d'inventé moi même des questions sur lesquelles je me base ma conception du datawarehouse.

- **Quel est l'impact de l'accès régulière à l'environnement d'apprentissage virtuel sur le résultat final ?**
- Quel est le nombre des étudiants inscrits dans chaque cours ?
- Quel est le nombre des étudiants par sexe ?
- Quel est le pourcentage des étudiants handicapés ?

Après analyser les données et comprendre la structure du OLTP, j'ai opté pour la structure en **étoile** pour créer mon datawarehouse, en alimentant 2 dimensions et une table de fait.

**NB** :J'ai utilisé « talend » comme outil d'intégration des données.

## Table des matières

Aucune entrée de table des matières n'a été trouvée.

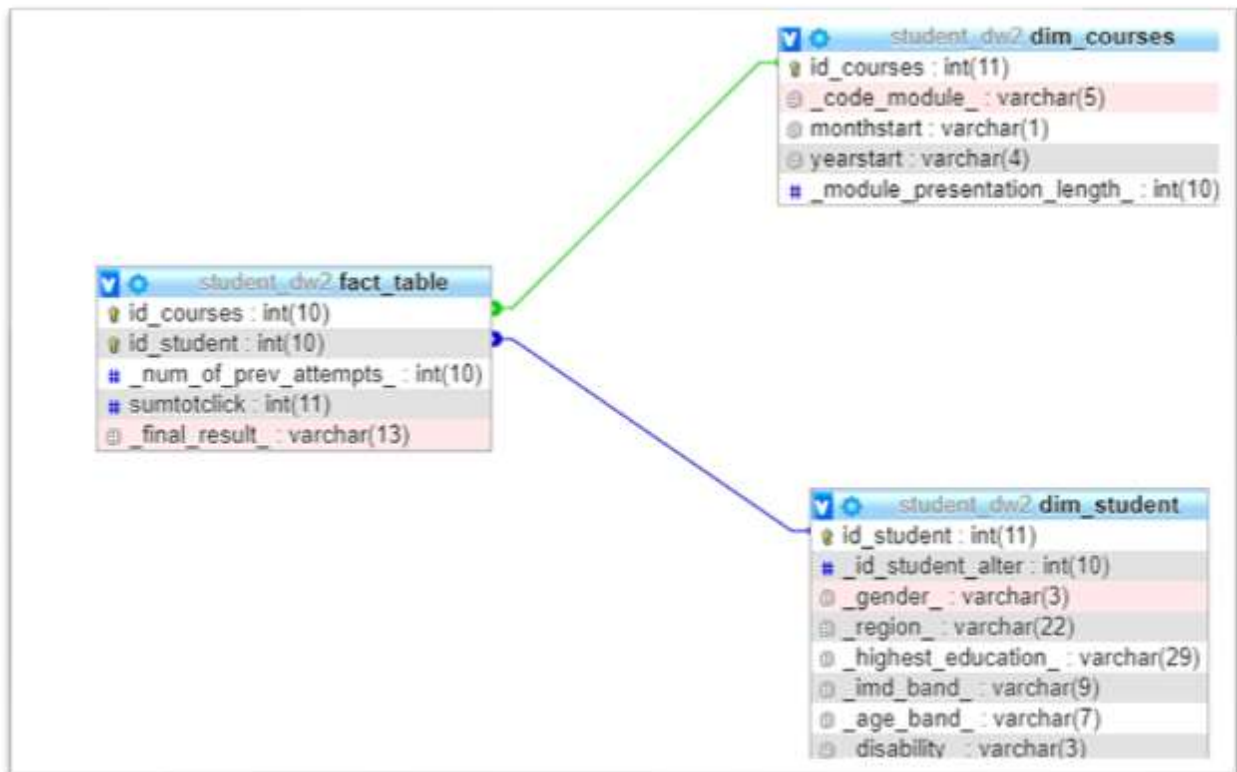


Figure 2: la structure en étoile du datawarehouse

Premièrement j'ai commencé par l'alimentation de la dimension « dim\_student » qui contient les informations démographiques sur les étudiants.

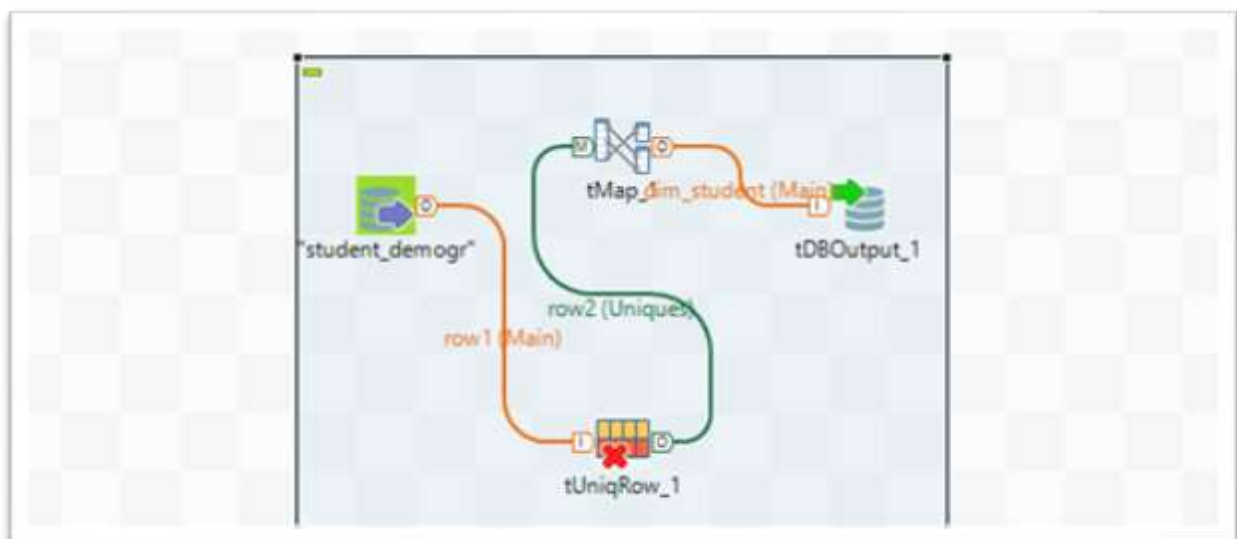


Figure 3: dimension student

Ensuite j'ai créé la deuxième dimension « dim\_courses » qui stocke les informations des différents cours.



Figure 4:dimension cours

Et finalement j'ai alimenté ma table de fait

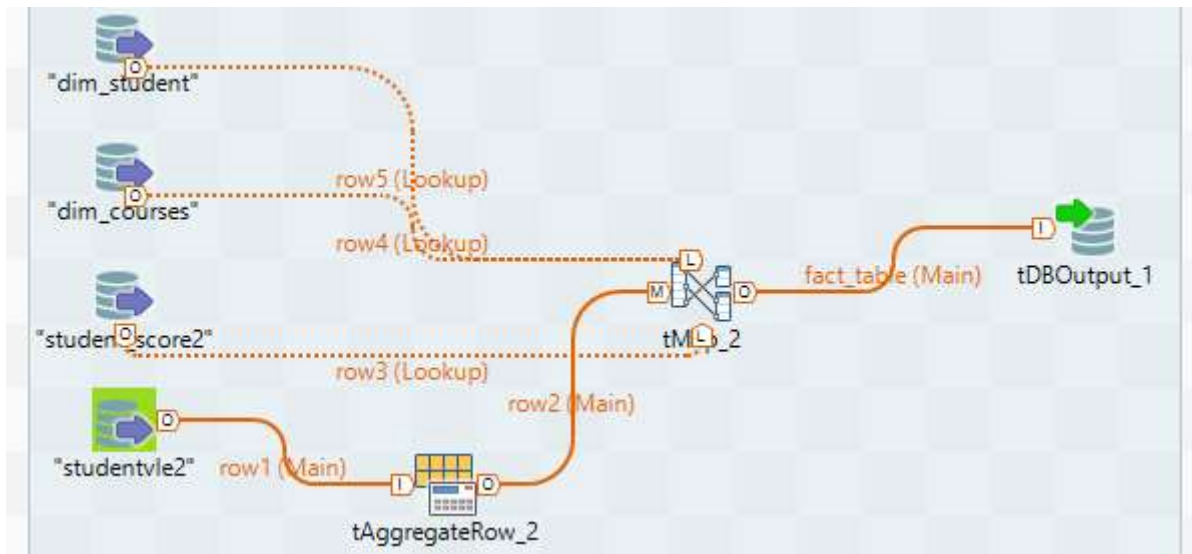


Figure 5:table de fait

Et puisque la taille des données à alimenter à chaque exécution du job est très grande, j'ai décidé de automatiser l'exécution des jobs d'alimentation en des créneaux de temps bien choisis pour ne pas bloquer le serveur d'exécution.

Et donc j'ai utilisé « **le planificateur des tâches** » de Windows.

Propriétés de talend\_fact\_job (Ordinateur local)

Général Déclencheurs Actions Conditions Paramètres Historique (désactivé)

Nom : talend\_fact\_job

Emplacement : \

Auteur : DESKTOP-BE6PB1A\hp

Description :

Options de sécurité

Utiliser le compte d'utilisateur suivant pour exécuter cette tâche :

hp Utilisateur ou groupe...

☐ N'exécuter que si l'utilisateur est connecté

☒ Exécuter même si l'utilisateur n'est pas connecté

☒ Ne pas enregistrer le mot de passe. La tâche n'accède qu'aux ressources locales.

Propriétés de talend\_fact\_job (Ordinateur local)

Général Déclencheurs Actions Conditions Paramètres Historique (désactivé)

Lorsque vous créez une tâche, vous pouvez spécifier les conditions qui la déclenchent.

Déclenchement	Détails	Statut
Une fois	À 16:02 le 24/05/2021	Activé



# Reporting

Je vais essayer dans cette étape de créer quelques diagrammes graphiques représentatifs pour répondre aux questions posées au départ (voir page 3).

J'ai utilisé **Power BI Desktop** pour faire ce travail et j'ai le partagé en ligne pour qu'il soit accessible pour les décideurs.

- Quel est l'impact de l'accès régulière à l'environnement d'apprentissage virtuel sur le résultat final ?

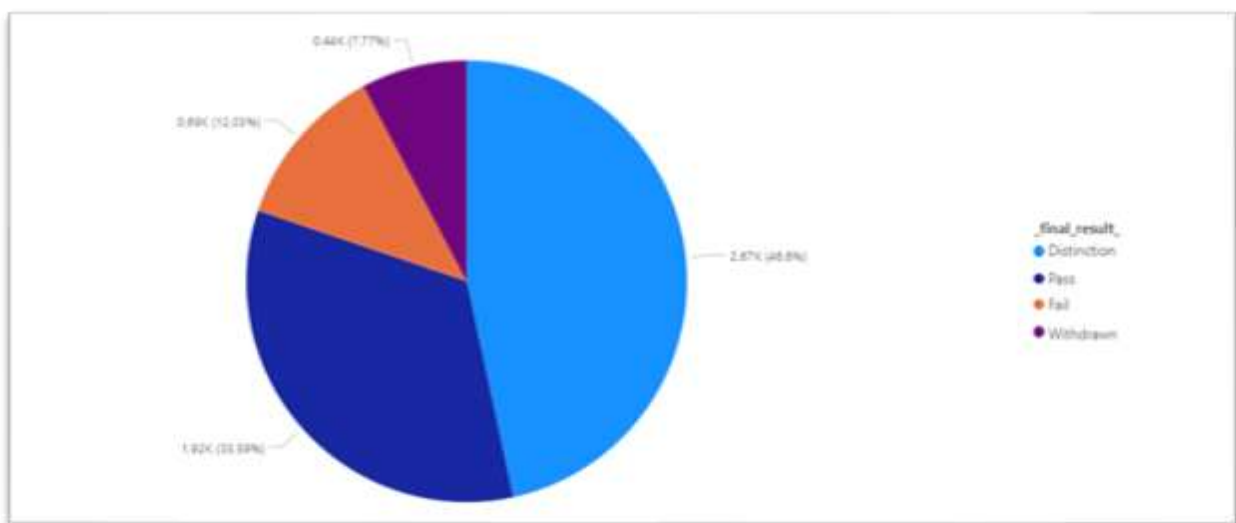


Figure 6: impact de vle sur le résultat

- Quel est le nombre des étudiants inscrits dans chaque cours ?

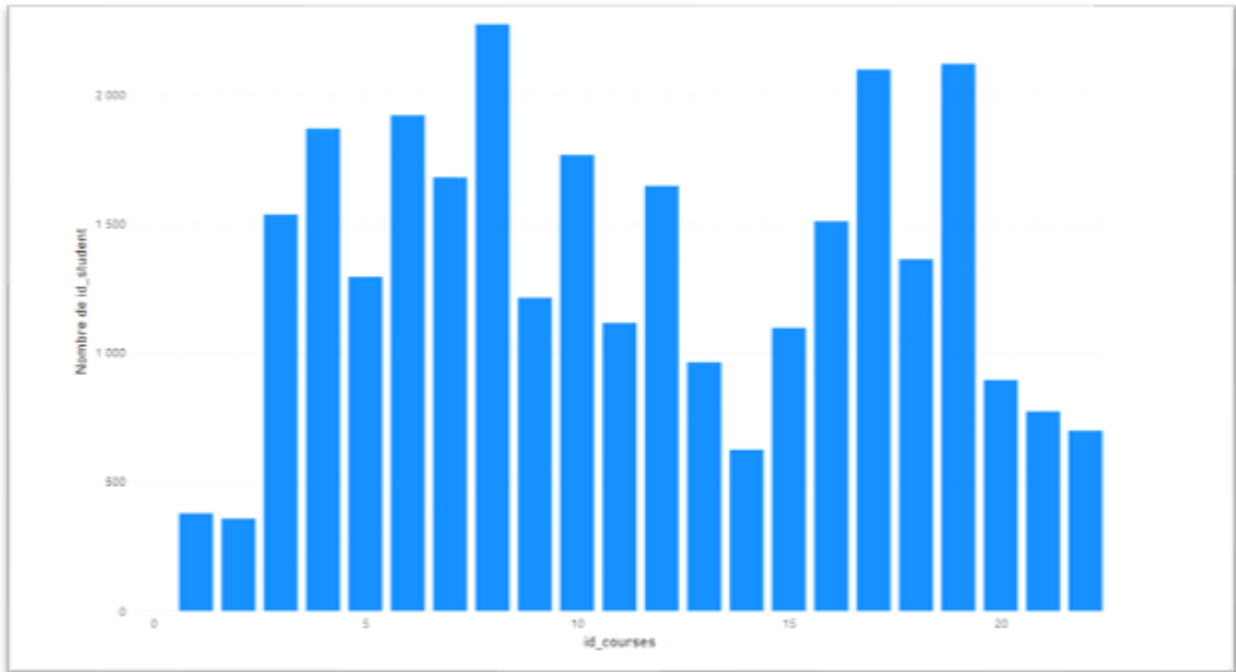


Figure 7:le nombre des étudiants inscrits dans chaque cours

- Quel est le nombre des étudiants par sexe ?

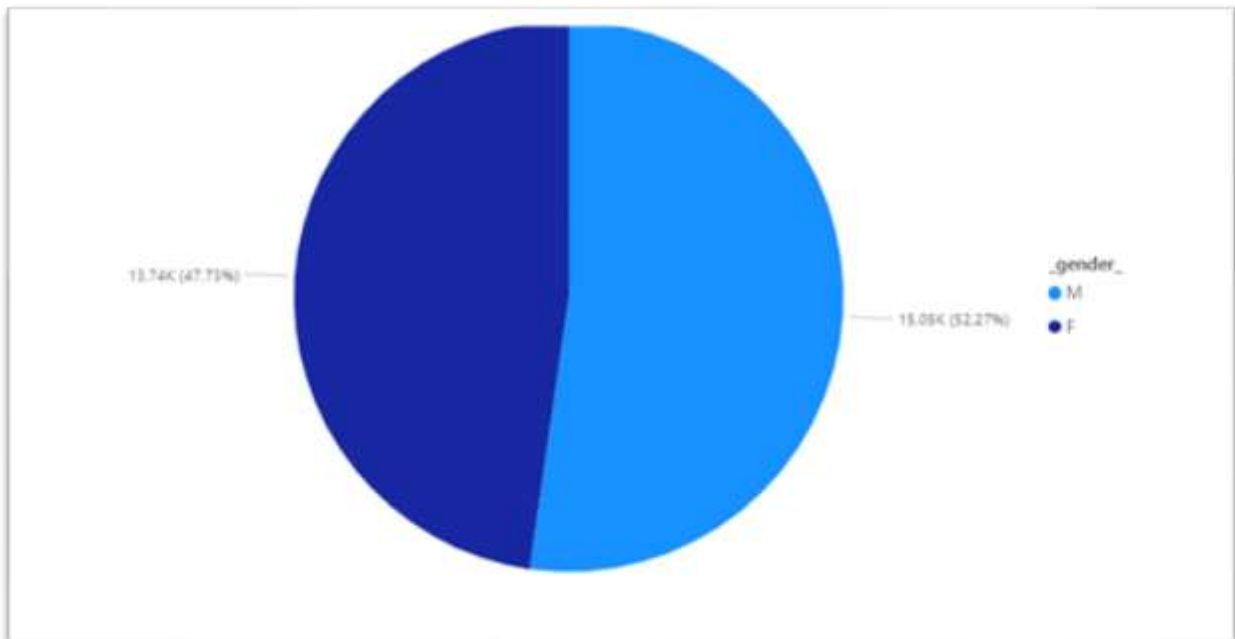


Figure 8:le nombre des étudiants par sexe

- Quel est le pourcentage des étudiants handicapés ?

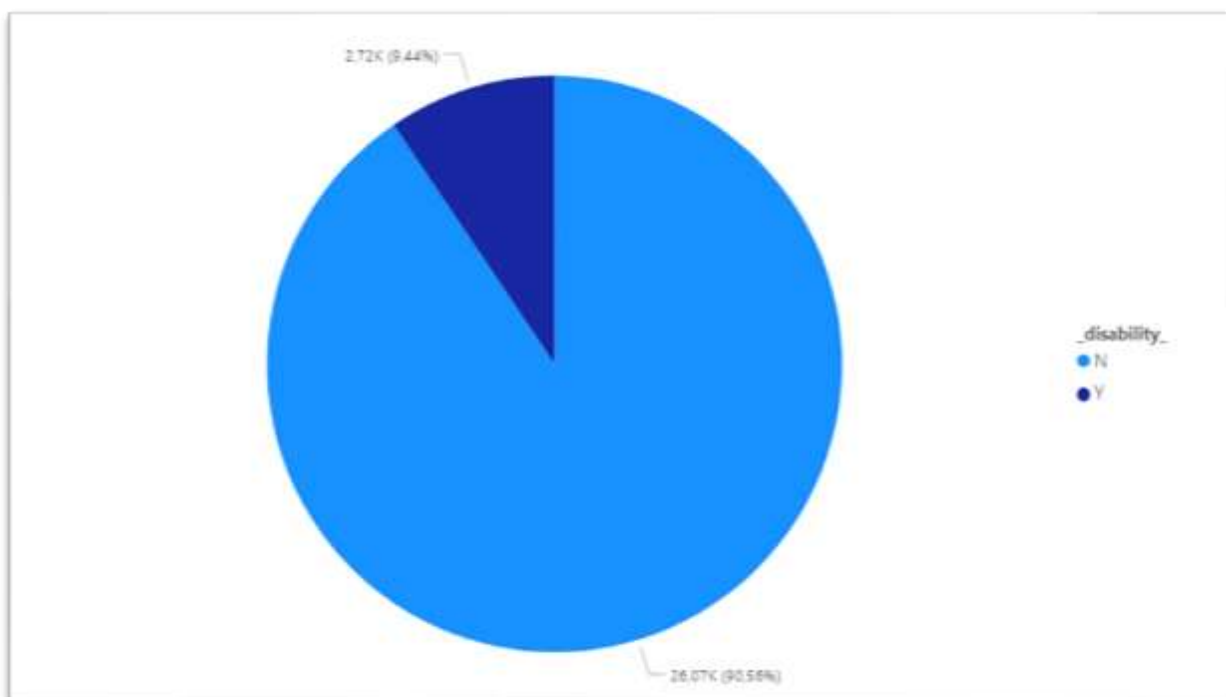


Figure 9:le pourcentage des étudiants handicapés

Finalement voilà le lien pour voir ces diagrammes en ligne : [powerBI - Power BI](#)