

Description du Rapport final et Soutenance

A. Consignes

- **Dépôt :**
 - Le rapport est à déposer sur Moodle(fichier zippé) au plus tard le lundi **20/01/2025**, à 23h00
 - Le fichier zippé comprends :
 - Les sources de données
 - Fichier power BI
 - Rapport final
- **Structure du rapport :** Le rapport doit être organisé selon les étapes ci-dessous, avec des titres clairs pour chaque section. Vous pouvez ajouter d'autres sections qui vous semblent pertinentes.
- **Github :**
 - Veuillez créer un compte GitHub si vous n'en possédez pas déjà un(Voir ci-dessous).
 - Assurez-vous de m'avoir ajouté comme collaborateur sur votre projet GitHub, en utilisant l'adresse email "jeanmuabila90@gmail.com".
- **Soutenance :**
 - La soutenance est prévue le mercredi 22/01/2024
 - Préparer une présentation de 5 min contenant :
 - Le contexte
 - Objectifs
 - Étapes d'intégration (Modélisation des SID, Intégration(ETL), Restitution)
 - Conclusion
-

B. Description rapport

1. **Présentation du projet (à inclure sur github)**
 - a. Contexte
 - b. Problématique
 - c. Objectifs
2. **Modélisation du SID**
 - a. Choix du processus métier à modéliser
 - b. Définition de la granularité
 - c. Choix des dimensions
 - d. Identification des faits(Mesures)
3. **Intégration de données(Processus ETL)**
 - a. **Description du processus d'extraction**

Détailler comment les données sont extraites, incluant les sources de données utilisées

b. Transformation des données

Expliquer les transformations appliquées aux données, telles que la suppression de colonnes non pertinentes, la modification des types de données, la séparation ou la fusion de colonnes. Chaque transformation doit être justifiée par rapport aux objectifs de l'analyse.

c. Gestion des erreurs et des données manquantes

Discuter des techniques utilisées pour gérer les erreurs et les données manquantes, telles que la suppression des doublons, la gestion des valeurs nulles, et les ajustements de données erronées.

d. Création de colonnes personnalisées et conditionnelles

Décrire comment et pourquoi des colonnes personnalisées ont été créées, y compris l'utilisation de formules spécifiques dans Power Query pour calculer de nouvelles valeurs

e. Analyse de la qualité des données

Analyser et rapporter sur la qualité des données après transformations, en utilisant des outils comme le profil de colonne pour montrer la distribution des données et identifier des anomalies potentielles

f. Création et utilisation de tables de temps

Expliquer comment une table de temps (ou table calendrier) est créée, soit manuellement avec des fonctions comme CALENDAR ou CALENDARAUTO dans DAX ou autres,

g. Création des mesures(fait) avec DAX

Utilisez les requêtes DAX pour créer des nouvelles mesures dynamiques(indicateurs, faits), qui seront utilisées dans la partie visualisation.

h. Description des modèles de données utilisés.

Expliquez clairement le modèle de données utilisé dans votre projet en lien avec le processus métier et choix des dimensions.

4. Restitution

- a. Reporting :
 - i. Présenter la visualisation de votre rapport suivant les objectifs établis lors de la partie présentation.

- b. Tableaux de bords

- Création d'un Compte github:

- GitHub est une plateforme de développement collaboratif utilisée pour héberger des projets utilisant le système de contrôle de version [Git](#).
- Vous êtes appelé à créer un compte github(avec votre mail personnel, de préférence) si vous en avez pas, sur lequel vous allez remplir le fichier Readme.md.
- Une seule personne est chargée de créer un nouveau projet([détails ici](#)), nommé SAE SID, et ajoutez d'autres collaborateurs([Détails ici](#)).
- M'envoyer une invitation comme collaborateur afin d'avoir accès à votre projet
- Pour créer un compte, voici le lien [Créer un compte Github](#)
- Pour un premier temps, le fichier Readme doit contenir la **Présentation de votre projet**.
- *Indiquez le lien github de votre projet dans la partie présentation.*