

Projet ATP

Auteur: Nessim LIAMANI

Date: Vendredi 28/06/2024

Table des matières

1. [Introduction](#)
 2. [Gestion de Projet](#)
 - [Planification Initiale](#)
 - [Exemples de Tâches](#)
 - [Gestion des Risques](#)
 3. [Configuration de l'Environnement](#)
 - [Outils Nécessaires](#)
 4. [Préparation des Données et Processus ETL](#)
 - [Télécharger et Organiser les Données](#)
 - [Créer la Base de Données Principale et les Tables](#)
 - [Charger les Données dans la Base de Données Principale](#)
 - [Créer l'Entrepôt de Données et Peupler les Données](#)
 5. [Analyse et Visualisation des Données](#)
 - [Nettoyage et Transformation des Données](#)
 - [Charger les Données dans Power BI](#)
 - [Exemples de Questions Métier](#)
 - [Visualisations Power BI](#)
 6. [Conclusion](#)
-

Introduction

Cette documentation décrit les étapes nécessaires pour mener à bien le projet ATP, impliquant la configuration des bases de données, le nettoyage des données, l'exécution des processus ETL et la création de visualisations.

Gestion de Projet

Planification Initiale

- Définir la portée du projet
- Identifier les parties prenantes
- Définir les objectifs du projet

Exemples de Tâches

Tâche	Durée
Planification Initiale	0.5 jour(s)
Acquisition de Données	0.5 jour(s)
Configuration de l'Environnement	0.2 jour(s)
Configuration de la Base de Données	0.5 jour(s)
Chargement des Données	1 jour(s)
Création de l'Entrepôt de Données	1 jour(s)
Nettoyage et Transformation des Données	1 jour(s)
Analyse et Visualisation des Données	1 jour(s)
Test et Validation	0.5 jour(s)
Documentation et Rapport	0.5 jour(s)
Total : 6,7 jour(s)	

Gestion des Risques

- Incohérences de données : Vérifications régulières de validation des données
 - Erreurs de script : Contrôle de version et test des scripts
-

Configuration de l'Environnement

Outils Nécessaires

- SQL Server 2022
- Python (avec pandas et numpy)
- PowerShell

- Windows Task Scheduler
-

Préparation des Données et Processus ETL

Télécharger et Organiser les Données

- Créer un répertoire pour le projet
- Télécharger les données des matchs et des joueurs ATP
- Enregistrer les fichiers dans le répertoire désigné

Créer la Base de Données Principale et les Tables

Script SQL pour créer la base de données principale et les tables (ATP_0_Main):

```
sql
Copy code
-- Créer la base de données principale
-- ...

-- Utiliser la base de données
-- ...

-- Créer les tables
-- ...
```

Charger les Données dans la Base de Données Principale

Script Python pour charger les données CSV dans SQL Server en utilisant pandas:

```
python
Copy code
# Placeholder pour le script Python
```

Script PowerShell pour exécuter le script Python:

```
powershell
Copy code
# Placeholder pour le script PowerShell
```

Créer l'Entrepôt de Données et Peupler les Données

Script SQL pour créer l'entrepôt de données et les tables (ATP_1_DWH):

```
sql
Copy code
-- Placeholder pour le script SQL
```

Script Python pour peupler l'entrepôt de données à partir de la base de données principale:

```
python
Copy code
# Placeholder pour le script Python
```

Script PowerShell pour exécuter le script Python pour l'entrepôt de données:

```
powershell  
Copy code  
# Placeholder pour le script PowerShell
```

Analyse et Visualisation des Données

Nettoyage et Transformation des Données

- Identifier les incohérences de données
- Écrire des scripts pour le nettoyage et la transformation des données
- Valider les données transformées

Charger les Données dans Power BI

- Connecter Power BI à l'entrepôt de données
- Créer des rapports pour répondre aux questions métier

Exemples de Questions Métier

- Qui a remporté le plus de matchs entre 2010 et 2017 ?
- Quelle est la durée moyenne d'un match en 3 sets ?
- Qui a réussi le plus d'aces en 2010, en 2011, etc. ?

Visualisations Power BI

- Concevoir les visualisations dans Power BI
 - Créer des visualisations pour chaque question métier
 - Valider les visualisations par rapport aux données
-

Conclusion

- Compiler le rapport final avec les conclusions et les visualisations
- Revoir le projet pour s'assurer de son exhaustivité
- Présenter le rapport final aux parties prenantes