

Note de cadrage : POC Avantages Sportifs

1 - Objectifs du POC

- ⇒ Tester la faisabilité technique de la solution.
- ⇒ Déterminer les données à collecter pour évaluer l'activité physique des salariés.
- ⇒ Calculer l'impact financier sur l'entreprise des avantages proposés.

2 - Avantages proposés pour les salariés

Prime sportive : 5% du salaire annuel brut pour les salariés venant au bureau en pratiquant une activité physique (vélo, trottinette, course à pied, marche, etc.). Il faut que le déplacement prenne une forme sportive la majorité du temps pour pouvoir être éligible et c'est fait avec le déclaratif des salariés (tu trouveras cette information dans le fichier RH).

5 journées bien-être : Accordées aux salariés ayant une activité physique en dehors du travail. Pour être éligible, il faut au **minimum 15 activités physiques** dans l'année. Pour le moment, nous allons demander au salarié de déclarer les différentes activités dans un google doc, mais nous souhaitons à terme utiliser une application comme Strava pour récupérer directement les données.

Les paramètres ne sont pas fixes. En effet, en fonction de l'évolution des pratiques et du projet, nous serons amenés à modifier certains paramètres.

3 - Périmètre du POC

1. **Création de l'Infrastructure de données** : Mise en place d'une base de données sécurisée.
2. **Création des données** : Créer automatiquement des données comme l'API [Strava](#) pour pouvoir générer des messages sur Slack et avoir des informations sur les pratiques sportives des employés (nécessaire pour le calcul des jours supplémentaires).
3. **Test sur les données** : documenter les tests de cohérence et de fonctionnalité (ex. vérifier que les distances ne sont pas négatives, que les dates sont valides, etc.). GreatExpectation ou SODA sont parfaits pour ce type de test.

4. **Développement du pipeline de données** : Extraire, transformer et charger (ETL) les données.
5. **Monitoring des des flux de données** : installer un outil de monitoring pour surveiller la volumétrie et l'état d'exécution du pipeline.
6. **Interface et restitution des résultats PowerBI** : Visualiser les KPI les plus importants pour ce projet (coût des solutions, nombre de jours supplémentaires, pratique sportive, etc.) dans un outil de visualisation type PowerBI.

4 - Contraintes et Exigences

- **Technologies** : Liberté sur le choix des outils et technologies.
- **Robustesse et sécurité** : Intégrité et protection des données (ce sont des données RH donc sensibles)
- **Documentation** : Code disponible sur un repo GitHub avec un ReadMe détaillé.

5 - Spécificité pour le POC

- 1) Valider que l'ensemble des déclarations des employés sur le mode de déplacement pour se rendre au travail est correct. Pour cela, il va falloir calculer automatiquement la distance entre le domicile et l'adresse de l'entreprise en fonction du mode de déplacement (Adresse de l'entreprise : 1362 Av. des Platanes, 34970 Lattes) en utilisant l'API de Google Maps

Règle à tester (à faire évoluer en fonction des résultats) :

- Marche/Running ⇒ Maximum 15 km
- Vélo/Trottinette/Autres ⇒ Maximum 25 km

Exemple : Si un salarié déclare venir en marchant et qu'il habite à 50 km de l'entreprise. C'est sûrement une erreur de déclaration ou un changement de situation. Nous souhaitons faire remonter les erreurs.

- 2) Toujours en se basant sur la pratique d'un sport, générer pour le test les flux de données pour les 12 derniers mois qui vont pouvoir

alimenter notre channel slack.

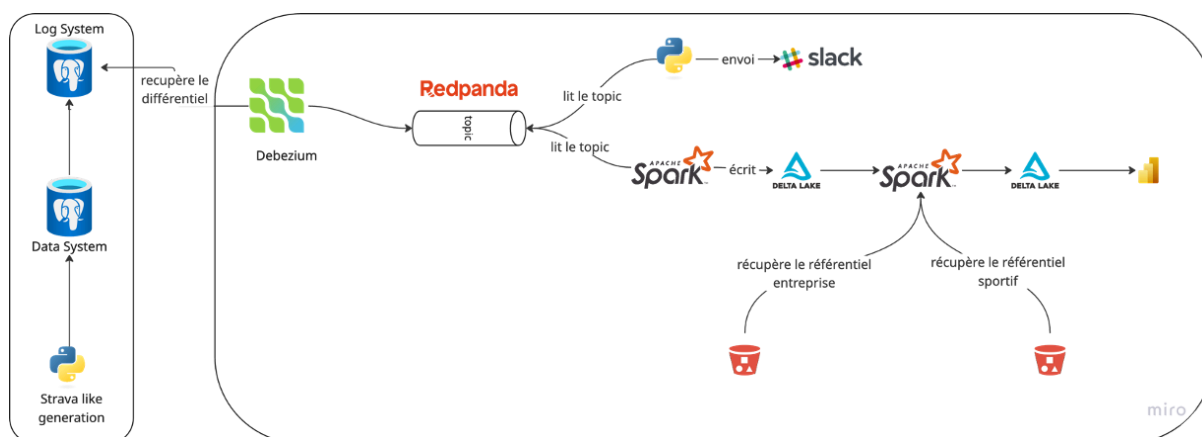
3) Exemple de message slack :

"Bravo Juliette Mendes ! Tu viens de courir 10,8 km en 46 min ! Quelle énergie ! 🏃‍♀️🏆"

"Magnifique Laurence Morvan ! Une randonnée de 10 km terminée et un nouveau spot à découvrir ! 🌄 ("Randonnée de St Guilhem le désert, je vous la conseille c'est top") 🥾"

ID	ID salarié	Date de début	sport type	Distance (m)	Temps écoulé (s)	Commentaire
1	43015	09/12/2023 09:09	Course à pied	10879	4617	
2	43542	07/12/2023 17:31	Course à pied	1304	6993	Reprise du sport :)
3	66425	03/12/2023 09:18	Course à pied	17503	7462	
4	91916	02/12/2023 09:41	Course à pied	3457	2170	
5	35731	01/12/2023 17:06	Randonnée	10108	17546	Randonnée de st Guilhem le desert, je vous la conseille c'est top

6 - Exemple d'architecture



Pour la phase finale, nous devrions pouvoir nous connecter directement aux données Strava des employés, afin de suivre concrètement leur activité physique. En attendant, merci de générer une simulation cohérente de ces données, couvrant un historique des 12 derniers mois. Cette simulation devra s'appuyer sur les informations issues des fichiers RH.

Les métadonnées attendues pour chaque enregistrement sont les suivantes :

ID ; ID salarié ; Date de début de l'activité ; Type ; Distance (en mètres, à laisser vide si non pertinent – par exemple pour l'escalade) ; Date de fin de l'activité ; Commentaire.

Nous aurons besoin de plusieurs milliers de lignes pour cet historique. Je te recommande de stocker ces données dans une base de données afin de faciliter l'injection de résultats "live" que nous souhaitons tester, notamment pour l'envoi du message Slack.