



IMT Atlantique

Bretagne-Pays de la Loire

École Mines-Télécom



BARBOSA PEREIRA Guilherme Augusto

Mémoire de Stage

« Améliorer la fiabilité du système de prévision de la demande carbone Safran en identifiant les données nécessaires, la gouvernance, la gestion de données, et en appliquant des analyses avec l'appui des outils data science »

Rapport de stage effectué 12/04/2021 au 08/10/2021

Tutrice Ecole : Cécile Bothorel

Tuteur Entreprise : David Lalandre

Établissement : IMT Atlantique, 655 Avenue du Technopole, 29280, Plouzané - France

Entreprise d'accueil : Safran Landing Systems, 7 Rue Général Valérie André, 78140 Vélizy-Villacoublay - France

Paris - FR
2021

Remerciement

Je tiens tout d'abord à remercier mon établissement de formation IMT Atlantique, et particulièrement Monsieur LALANDRE David et Madame BOTHOREL Cécile pour m'avoir fait confiance. Je vous remercie pour le temps passé ensemble et pour le partage de votre expertise. L'écoute et l'accompagnement dont j'ai bénéficié m'ont permis de trouver rapidement un stage, dans le but d'affiner mon projet professionnel et d'obtenir mon diplôme de formation BAC +5.

Je remercie également Monsieur SIMON Julien, directeur Supply Chain Division roues et freins, et Monsieur HAQUIN Arnaud, directeur industriel division roues et freins, qui a cru en mon potentiel et m'a accueilli au sein de son entreprise. Ce stage a été une très bonne occasion de découvrir le métier de la data science à travers des missions très intéressantes.

Concernant ce titre, je souhaiterais remercier tout particulièrement Monsieur LALANDRE David, qui m'a épaulé et conseillé et qui m'a surtout transmis son expertise. Ainsi que son équipe, notamment Madame DIATTA Angélique, chargée de la demande carbone.

Ce stage m'a permis d'affiner certaines pistes pour bâtir mon projet d'orientation professionnel et constitue l'aboutissement de mon année scolaire.

Je n'oublie pas non plus mes proches qui m'ont sans cesse soutenu dans l'élaboration de mon projet professionnel et m'ont aidé à chaque étape de ce rapport de stage.

Résumé

Dans le but d'améliorer l'utilisation de données et d'optimiser les processus de décision dans l'entreprise, le stage dans le domaine de la data science peut apporter de valeur au service Safran. L'entreprise Safran est un groupe international de haute technologie qui a comme missions de contribuer à la performance du transport aérien, améliorer la protection des citoyens et faciliter l'accès à l'espace. Le principal défi pour l'intégration de la data science au service de la compagnie est d'ajuster la gouvernance de ses données. Cela peut être résolu à travers la mise en œuvre d'une bonne gestion de données et des analyses avec l'appui des outils de traitement de données, tels que l'application Power BI. Le manuscrit ici présenté vient exposer les travaux associés à la conception des projets capables de fiabiliser une bonne gouvernance de données et d'automatiser les processus d'analyse ou de prise la décision, afin qu'ils puissent réduire le temps d'exécution de certains services manuels et offrir des meilleures organisations et visualisations de données avec des tableaux de bord via Power BI pour aider à rendre également plus efficace les décisions stratégiques. Les idées des projets sont résultats de plusieurs discussions avec des experts de différents domaines qui travaillent dans la division où s'est déroulé mon stage. Dans un point de vue de la data science, on se concentre sur la méthodologie CRISP-DM pour le développement des projets, tout en cherchant l'utilisation d'un modèle simple et en même temps efficace. De plus, il y a eu également des projets qui ont déjà été développés et, dans ceux-ci, des améliorations ont été apportées conformément au cahier de charge des projets, et correction de certaines anomalies. Le mémoire se concentrera sur un projet spécifique appelé « Dashboard HSE & Traffic », le projet le plus important développé qui a apporté plus de valeur ajoutée à Safran dans le processus de décision face aux conséquences de la crise du coronavirus. Ce projet sera utilisé principalement par le Directeur de la Division Carbone et pour le Top Management Safran Roues et Freins.

Abstract

In order to improve the use of data and optimize decision-making processes in the company, the internship in the field of data science can add value to the Safran department. The Safran Company is an international high-tech group whose missions are to contribute to the performance of air transport, improve the protection of citizens and facilitate access to space. The main challenge for the integration of data science in the service of the company is to adjust the governance of its data. This can be solved through the implementation of good data management and analytics with the support of data processing tools, such as the Power BI application. The manuscript presented here presents the work associated with the design of projects capable of making data governance more reliable and of automating analysis or decision-making processes, so that they can reduce the execution time of certain manuals services and deliver better data organizations and data visualizations with dashboards through Power BI to help make strategic decisions more efficient as well. The ideas for the projects are the result of several discussions with experts from different fields who work in the division where the internship took place. From a data science perspective, we focus on the CRISP-DM methodology for project development, while seeking the use of a simple and at the same time efficient model. In addition, there have also been projects that have already been developed and, in these, improvements have been made according to the specifications of the projects or the correction of certain anomalies. The report will focus on a specific project, called "Dashboard HSE & Traffic", the most important project developed which brought more added value to Safran in the decision-making process faced with the consequences of the coronavirus crisis. This project will be used mainly by the Director of the Carbon Division and for Safran Wheels and Brakes Top Management.

Table des matières

REMERCIEMENT	2
RESUME	3
ABSTRACT	3
TABLE DES MATIERES	4
LISTE DE FIGURE	6
GLOSSAIRE	6
1 INTRODUCTION	7
2 CONTEXTE	8
2.1 Entreprise.....	8
2.2 Mission.....	8
2.3 Méthodologie CRISP-DM.....	9
2.4 Power BI ^[4]	10
2.4.1 <i>Power BI Report Server</i>	11
2.5 DHA.....	11
2.6 Tiscam	11
2.7 Zementer.....	12
2.8 SAP.....	12
2.9 LPO et HSE	12
2.10 Données Carbon Supply Chain	12
2.11 Projets.....	14
3 DASHBOARD HSE & TRAFFIC	17
3.1 Introduction.....	17
3.2 Business Understanding.....	17
3.3 Data Understanding.....	17
3.3.1 <i>Budget</i>	18
3.3.2 <i>Cycle</i>	18
3.3.2.1 <i>Fonction « GetCycleMonthAndYear »</i>	18
3.3.2.2 <i>Fonction « GetAllCycleByMonthAndYear »</i>	19
3.3.3 <i>Data Fleet</i>	19
3.3.4 <i>Data SAP</i>	19
3.3.5 <i>Hubs Allocation</i>	20

3.3.6	<i>Immat.....</i>	20
3.3.7	<i>S&OP</i>	20
3.4	<i>Data Préparation.....</i>	21
3.4.1	<i>Hubs Allocation</i>	21
3.4.2	<i>Date Last Refreshed.....</i>	21
3.4.3	<i>Correspondence S&OP & Budget.....</i>	22
3.4.4	<i>Correspondence Role & Program</i>	22
3.4.5	<i>S&OP +M</i>	22
3.4.6	<i>S&OP -M.....</i>	23
3.4.7	<i>Calendar</i>	24
3.4.8	<i>Cycle</i>	24
3.4.9	<i>Data SAP.....</i>	26
3.4.10	<i>Match Cycle & Data SAP.....</i>	26
3.4.11	<i>Match Cycle</i>	27
3.4.12	<i>Budget.....</i>	27
3.4.13	<i>Match Data SAP.....</i>	28
3.4.14	<i>Site PP.....</i>	28
3.4.15	<i>Match Program & Data SAP.....</i>	28
3.4.16	<i>Data Fleet</i>	28
3.5	<i>Modeling.....</i>	28
3.6	<i>Evaluation.....</i>	30
3.7	<i>Deployment.....</i>	33
4	CONCLUSION	35
4.1	<i>Perspectives</i>	35
4.2	<i>Retour d'expérience & bilan de compétences</i>	36
	BIBLIOGRAPHIE	37

Liste de figure

Figure 1. Variation du trafic aérien en Europe entre 2021-2020 [1].....	7
Figure 2. Diagramme CRISP-DM [2].....	9
Figure 3. Diagramme avec les bases de données et leurs informations contenues	13
Figure 4. Diagramme des bases de données, ses experts et la division qu'elles sont utilisées	14
Figure 5. Schéma d'organisation des données par projets et les informations de ses extractions	16
Figure 6. Code d'extraction de l'ensemble de données Budget.....	18
Figure 7. Code de la définition de la fonction « GetCycleMonthAndYear ».....	19
Figure 8. Code de la définition de la fonction « GetAllCycleByMonthAndYear »	19
Figure 9. Code de l'extraction de l'ensemble de données Data Fleet.....	19
Figure 10. Code de l'extraction de l'ensemble de données Data SAP.....	20
Figure 11. Code de l'extraction de l'ensemble de données Hubs Allocation	20
Figure 12. Code pour construire la table Date Last Refreshed.....	21
Figure 13. Code pour intégrer la table « Correspondence S&OP & Budget » dans le projet PBI	22
Figure 14. Code pour intégrer la table « Correspondence Role & Program » dans le projet PBI.....	22
Figure 15. Code pour construire la table de données « S&OP +M ».....	23
Figure 16. Code pour construire la table de données « S&OP –M ».....	24
Figure 17. Code pour construire la table « Calendar »	24
Figure 18. Code pour construire la table « Data SAP ».....	26
Figure 19. Code de construction de la table « Match Cycle »	27
Figure 20. Code pour la construction de la table « Budget »	27
Figure 21. Code de construction pour intégrer la table « Data Fleet ».....	28
Figure 22. Relations entre les données des tables dans la modélisation	29
Figure 23. Design de la modélisation réalisée pour le projet	30
Figure 24. Graphique principal de l'onglet Dashboard HSE & Traffic	31
Figure 25. Table « Percentages » de l'onglet Dashboard HSE & Traffic.....	31
Figure 26. Table « HSE and Traffic » de l'onglet Dashboard HSE & Traffic.....	32
Figure 27. Filtres de l'onglet Dashboard HSE & Traffic	33
Figure 28. Dashboard HSE and Traffic.....	34
Figure 29. Onglet avec des informations générales qui résume le projet.....	34

Glossaire

SLS : Safran Landing Systems ;

Cycle : Durée de vie du puit de chaleur du frein de l'avion ;

LPY : Mesure d'atterrissages par an pour un avion ;

LPO : Mesure d'atterrissages par révision pour un dissipateur thermique ;

HSE : Mesure de la quantité d'échange de dissipateur de chaleur du frein de l'avion ;

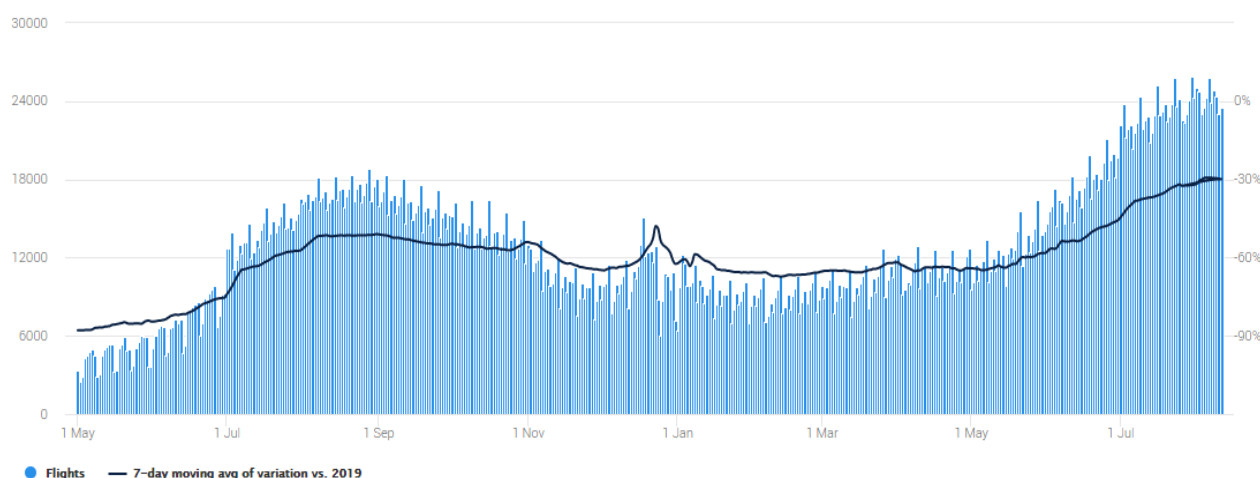
BDD : Base de données ;

PBI : Power BI ;

SAA : Filiale data science de Safran ;

1 Introduction

En 2020 et 2021, la pandémie de coronavirus a causé de nombreuses inquiétudes pour de nombreuses entreprises, principalement celles qui dépendaient du trafic aérien. Selon « *Supporting European Aviation* » (EUROCONTROL), il y a eu des moments où le trafic aérien européen a chuté de plus de 80% par rapport à la même période en 2019 [1]. Dans ce contexte, il est d'une grande importance pour ces entreprises qui dépend du trafic aérien pour leur activité d'étudier bien leurs données avec la data science afin d'optimiser les meilleures stratégies qui doivent être mises en place sur la base d'une analyse de la situation globale à travers une bonne gouvernance de données. Dans la figure 1 illustre les conséquences du coronavirus pour le trafic aérien en Europe, par exemple.



Le mémoire présente le développement d'un stage de fin d'études de la data science pour améliorer la fiabilité du système de prévision de la demande carbone en identifiant les données nécessaires, la gouvernance, la gestion de données, et en appliquant des analyses avec l'appui des outils de traitement de données chez Safran Landing Systems, un groupe international de haute technologie qui a comme missions de contribuer à la performance du transport aérien, améliorer la protection des citoyens et faciliter l'accès à l'espace. Le stage est réalisé par Guilherme Augusto Barbosa Pereira, étudiant brésilien en double diplôme de la formation BAC+5 en Data Science à l'IMT Atlantique.

Cette expérience a été l'opportunité de percevoir comment une entreprise spécialisée dans le secteur aérien se développe et de relever avec elle le défi à laquelle elle était confrontée et notamment face à la question du coronavirus.

Afin de rendre compte de manière fidèle et analytique les 6 mois au sein de SLS, il apparaît logique de présenter au préalable le contexte de l'entreprise, puis d'envisager le cadre du stage, les méthodologies de développement, les aspects plus technique et les informations des projets travaillés. Enfin, il sera précisé les différentes missions et tâches que j'ai pu effectuer et les qualités professionnelles et personnelles que j'ai pu en tirer.

2 Contexte

2.1 Entreprise

La compagnie Safran a comme principaux domaines d'activité la propulsion aéronautique et spatiale, les équipements aéronautiques, les intérieurs d'avions et la défense. Elle est la troisième mondiale en aéronautique, hors avionneurs, en plus, elle est également la première mondiale sur d'autres sujets spécifiques, comme par exemple les moteurs d'hélicoptères et les systèmes d'interconnexions électriques aéronautiques. Chez Safran travaille plus de 84 000 salariés dans 30 pays et 270 sites différents. Son chiffre d'affaires a atteint 24,6 milliards d'euros début 2020 avec 3,8 milliards d'euros de résultat opérationnel et 1,7 milliard d'euros de réinvestissement en R&D. Actuellement, Safran compte déjà plus de 1200 déposes de brevets. Par ailleurs, concernant la quantité des produits qui sont sur le marché, plus de 35 000 de ses moteurs d'avion court et moyen-courrier sont en service dans le monde et plus de 27 000 aéronefs sont équipés de ses systèmes d'atterrissage dans le monde entier.

Le stage s'est déroulé au siège de Safran Landing Systems (SLS) à Vélizy. Il y a la Direction de la filiale, les fonctions supports (RH, IT, Qualité, Support Clients), la direction des divisions produits, Trains d'atterrissages, Systèmes de Freinage, Roues et Freins, Maintenance et Réparations, avec leurs fonctions commerciales et technique. Le secteur responsable du stage est la « Supply Chain Centrale » de la division Roues et Freins. Il y a 5 subdivisions :

- *Carbon Supply Chain (Demand and Planner);*
- *Wheels & Brakes Supply Chain ;*
- *Program Production Manager ;*
- *Wheels Brakes Supply Chain(Demand and Planner);*
- *Inventories / Risks / Transport / Innovative Solution Logistics.*

Le stage s'est déroulé au sein de la « Carbon Supply Chain ». Ces freins mentionnés ont besoin de carbone parce qu'ils le constituent pour absorber une quantité d'énergie considérable dans le processus de freinage (le carbone a cet avantage par rapport à l'acier). Safran peut ainsi disposer de la meilleure technologie pour un freinage aéronautique de haute performance. Par exemple, par rapport aux disques de freins en acier, les disques en carbone ont de multiples qualités, qui se traduisent en autant d'avantages. Dans l'ensemble, les freins en carbone ont les avantages d'être efficaces, performants, légers (gains de masse de plusieurs centaines de kilos sur avion), durables (durée de vie 3 fois supérieures à celle de l'acier) et économiques (temps de renouvellement allongés).

2.2 Mission

La mission du stage est d'utiliser les connaissances de la data science pour assurer la bonne gouvernance de données de demande carbone et améliorer la fiabilité de la gestion des données de la part compagnie. Pour remplir cette mission, les principaux processus ont été :

- La familiarisation avec les données de la compagnie ;
- Connaître et comprendre sa gouvernance et sa gestion de données actuelle ;
- Explorer les moyens d'augmenter la fiabilité de cette gestion de données ;
- Réaliser des analyses avec l'appui des outils de traitement de données afin d'automatiser les processus et les interactions entre les données, en optimisant la productivité et efficacité du Safran.

L'objectif du stage est d'évaluer si la data science peut apporter des valeurs pertinentes au service Safran et appliquer ces valeurs dans l'entreprise, et l'élaboration de ce rapport a pour

principale source la pratique journalière des missions qui m'étaient affectées, mise en parallèle avec les enseignements théoriques de ma formation.

2.3 Méthodologie CRISP-DM

Pour développer un projet, il existe plusieurs manières de démarrer, ces méthodes peuvent être des méthodologies qui ont déjà été étudiées et testées, de sorte que les chances que les projets fonctionnent bien et soient plus organisés sont plus grandes. Parmi les différentes méthodes qui peuvent être utilisées en développement, il y a la méthodologie CRISP-DM, qui est très efficace et agile pour les projets qui nécessitent une étude de données. Cet acronyme signifie « *Cross Industrie Process pour Data Mining* » et a été publié en 1999 pour standardiser les processus d'exploration de données dans tous les secteurs.

La méthodologie est un modèle de processus en six phases qui décrit naturellement le cycle de vie de la science des données. Il s'agit d'un ensemble de processus ordonnés utilisés pour planifier, organiser et mettre en œuvre le projet de science des données. Les six phases sont: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, *Deployment*. Pour chaque phase, on peut définir une question simple pour faciliter le début du raisonnement [2]. La figure 2 illustre le cycle de la méthodologie CRISP-DM.

- *Business Understanding* – De quoi l'entreprise a-t-elle besoin ?;
- *Data Understanding* – De quelles avons-nous besoin ? C'est propre ?;
- *Data Preparation* – Comment organisons-nous les données pour la modélisation ?;
- *Modeling* – Quelles techniques de modélisation appliquer ?;
- *Evaluation* – Quel modèle répond le mieux aux objectifs de l'entreprise ?;
- *Deployment* – Comment les parties prenantes accèdent-elles aux résultats ?.

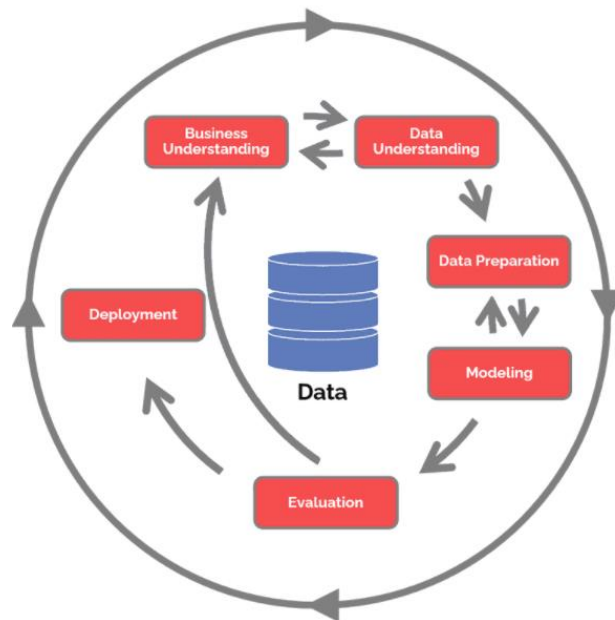


Figure 2. Diagramme CRISP-DM [2]

La première étape (*Business Understanding*) du processus CRISP-DM consiste à comprendre ce qu'il faut accomplir d'un point de vue business. L'organisation peut avoir des objectifs et des contraintes concurrentes qui doivent être correctement équilibrés. L'objectif de cette étape du processus est de découvrir les facteurs importants qui pourraient influencer l'issue du projet [3]. Cette étape peut se résumer en 4 processus :

- Quels sont les résultats souhaités du projet ?;
- Evaluer la situation actuelle ?;

- Déterminer les objectifs d'exploration de données ;
- Produire le plan de projet.

La deuxième étape (*Data Understanding*) oblige à acquérir les données répertoriées dans les ressources du projet. Cette collecte initiale peut comprendre le chargement des données. S'il y a plusieurs sources, il faut déterminer comment et quand les intégrer [3]. Cette étape peut se résumer en 4 processus :

- Décrire les données ;
- Explorer les données ;
- Vérifier la qualité des données ;
- Rapport sur la qualité des données.

La troisième étape (*Data Preparation*) est celle où la préparation et l'analyse des ensembles de données finaux pour la modélisation. Cette étape peut se résumer en 5 processus :

- Sélectionner les données ;
- Nettoyer les données ;
- Construire des données ;
- Intégrer les données ;
- Formater les données.

Dans la quatrième étape (*Modeling*) il faut créer et évaluer divers modèles basés sur plusieurs techniques différentes. Cette étape peut se résumer en 4 processus :

- Sélectionner les techniques de modélisation ;
- Générer une conception de test ;
- Modèle de construction ;
- Évaluer le modèle.

La cinquième étape (*Evaluation*) est la phase qui se concentre sur l'évaluation du modèle technique. Il y a des examens plus larges sur quel modèle correspond le mieux à l'entreprise et ce qu'il faut faire ensuite. Cette étape peut se résumer en 3 processus :

- Évaluer les résultats ;
- Révision ;
- Déterminer les étapes suivantes.

La sixième étape (*Deployment*) est le déploiement des résultats du projet. Elle est utile pour le client qui peut accéder à ses résultats et consiste de :

- Planifier le déploiement ;
- Produire le rapport final ;
- Réviser le projet.

Le méthode CRISP-DM ne précise pas ce qu'il faut faire après le déploiement du projet, mais si le projet est mis en production, il faut assurer de maintenir le modèle en production avec des surveillances constantes et des réglages occasionnels [2].

2.4 Power BI [4]

Il existe des applications pour aider le développement de projets qui comprennent la science de données et la business intelligence, un d'entre eux est l'application Power BI. Cette application permet obtenir des analyses en libre-service à l'échelle de l'entreprise, en réduisant les coûts supplémentaires, la complexité et les risques de sécurité de plusieurs solutions avec une plateforme d'analyse qui s'étend des individus à l'organisation dans son ensemble.

Pour interpréter les résultats ou les données d'un projet de façon plus intelligente, le Power BI utilise des outils intelligents pour trouver et partager des informations significatives avec des centaines de visualisations de données, des capacités d'Intelligence artificielle intégrées, une intégration Excel étroite et des connecteurs de données prédéfinis et personnalisés.

Dans le contexte de la sécurité, l'application aide à protéger les données d'analyse grâce à des capacités de sécurité des données de pointe, notamment l'étiquetage de sensibilité, le chiffrement de bout en bout et la surveillance des accès en temps réel.

En résumé, cette application Microsoft permet d'extraire directement et automatiquement les données des différentes bases, de les analyser et de les visualiser de manière personnalisée et interactive avec une interface simple, dynamique et pratique.

Pour effectuer les publications des projets PBI chez Safran, ils utilisent le système Power BI Report Server.

2.4.1 Power BI Report Server

Le Power BI Report Server est un portail de publication de rapports Power BI Desktop, avec lequel il est possible d'envoyer le projet au système de serveur Safran et de permettre à d'autres personnes d'y accéder, en pouvant choisir qui peut ou non y accéder.

Le portail permet la publication de rapport Power BI Desktop, il s'agit d'une publication interne à Safran, pas de stockage externe à Safran (e.g. cloud Microsoft). Le portail est également mutualisé entre toutes les sociétés du groupe. Chaque société et chaque métier disposera d'un espace de publication sur le portail et il est organisé par processus One Safran, puis par division et site. Normalement, les personnes vont pouvoir avoir l'accès juste dans le site et division qu'elle travaille. Par ailleurs, la sécurité des rapports publiés est gérée par le responsable du rapport et par défaut, chaque collaborateur de Safran Landing Systems a un accès en visualisation au portail. Pour obtenir les droits de publier, la demande est à formuler auprès de votre RSIP/Key User ou votre référent Power BI DSI.

2.5 DHA

L'acronyme DHA est défini comme Datahub Aero et il y a une API pour extraire ses informations. Cette API sert pour exposer les données du projet DHA. Le Datahub Aero est un échange de données avec un flux de données sans friction en son cœur. Elle peut être décrite comme solution composée de différents domaines : data science, *warehouse* et ingénierie.

Toutes les demandes faites par les requêtes de cette API doivent utiliser des dates, parce que dans le Datahub Aero il y a beaucoup de données et il doit restreindre la requête pour obtenir une réponse dans un timing correct. De plus, le projet du Dashboard HSE & Traffic qui a été utilisé l'API DHA et n'était qu'un type de requête utilisé, la « datahub_cycle », qui extrait toutes les données pertinentes pour obtenir des informations sur le trafic aérien mondial.

2.6 Tiscam

Le logiciel Tiscam a été développé par l'application qui s'appelle Omnis pour solutionner le problème de traçabilité des disques carbone Safran (fabrication, site, lot, date, opérations, opérateurs, clients, livraisons). Il était mis en place en 2003 et il est un développement entièrement maison et utilisé par tous les sites de productions de roues et freins Safran et par Vélizy (SLS). A travers du logiciel est possible de suivre la vie du disque de puits de chaleur Safran, c'est-à-dire, les informations de livraison à telle date au client, rénovation du disque, son retour à telle date, ses défauts ou suivre sa performance en service chez les compagnies (par exemple, durée de vie, LPO, temps passée chez le client ou raison de la dépose).

2.7 Zemeter

Le logiciel Zemeter est un ERP (*Enterprise Resource Planning*) et de prévision pour petites entreprises, mais utilisé par le département de roues et freins chez SLS comme un Excel ou un PBI puissant. Il est pareil à Tiscam, cependant, avec une diffusion mondiale plus restreinte. Il était mis en production en 2012 et son utilisation chez SLS est pour :

- Suivre la flotte équipée de freins carbone Safran ;
- Suivre des opérations de la flotte Safran (par exemple, atterrissages) ;
- Suivre la prévision de déposes sur 5 ans.

2.8 SAP

Le logiciel SAP est également un ERP (*Enterprise Resource Planning*) numéro 1 mondial, utilisé pour les grosses entreprises comme, par exemple, Safran, Peugeot et Moët Hennessy Louis Vuitton SE. Normalement SAP permet de consolider facilement toutes les données de toutes les filiales et de tous les sites d'un groupe. Chez SLS ses fonctionnalités sont :

- Management de la demande ;
- Production ;
- Administration des ventes ;
- Finance ;
- Suivre les informations de déposes, clients, compagnies, LPOs et toutes d'autres aspects commerciaux liés aux déposes de puit de chaleur (commandes, livraison, stocks, etc...).

2.9 LPO et HSE

L'acronyme LPO (*Landings per Overhaul*) est la mesure d'atterrissages par révision pour un dissipateur thermique chez Safran. Cette mesure a été largement utilisée dans le projet Dashboard LPO. Elle peut être calculé par la suivante :

$$LPO = Cycle * LPY$$

L'acronyme HSE (*Heat Sink Exchange*) est la mesure de la quantité d'échange de dissipateur de chaleur du frein de l'avion chez Safran. Cette mesure a été largement utilisée dans le projet Dashboard HSE & Traffic. Elle peut être calculé par la suivante :

$$HSE = Nombre\ de\ frein * Nombre\ d'avion * \left(\frac{LPY}{LPO} \right)$$

2.10 Données Carbon Supply Chain

Au début du stage, avec Monsieur LALANDRE David, il a été faite une l'étude sur les informations les plus pertinentes sur l'entreprise et sur certaines bases de données du département *Carbon Supply Chain*. Ces bases de données sont complexes et contiennent une grande quantité de données. Il fallait donc savoir quelles données sont nécessaires aux métiers pour mener à bien ses activités. Les idées des projets à travailler pendant le stage ont été discutées et il a été convenu de commencer à réaliser des projets concrets qui permettraient de se familiariser avec les données, dans le but d'efficacité et de productivité. Par ailleurs, pendant chaque projet il y avait un expert

dans le domaine du projet afin de conseiller et orienter la bonne réalisation, en aidant à comprendre les données utilisées.

Au cours du stage, il a été possible de comprendre chaque information contenue dans les bases de données investiguées, à la fois grâce aux résultats obtenus à partir de plusieurs analyses de données qui ont été nécessaires pour pouvoir terminer les projets. Ainsi, la figure 3 montre un schéma qui a été faite afin d'illustrer d'une manière claire les relations entre les données utilisées dans la « Supply Chain Carbon ».

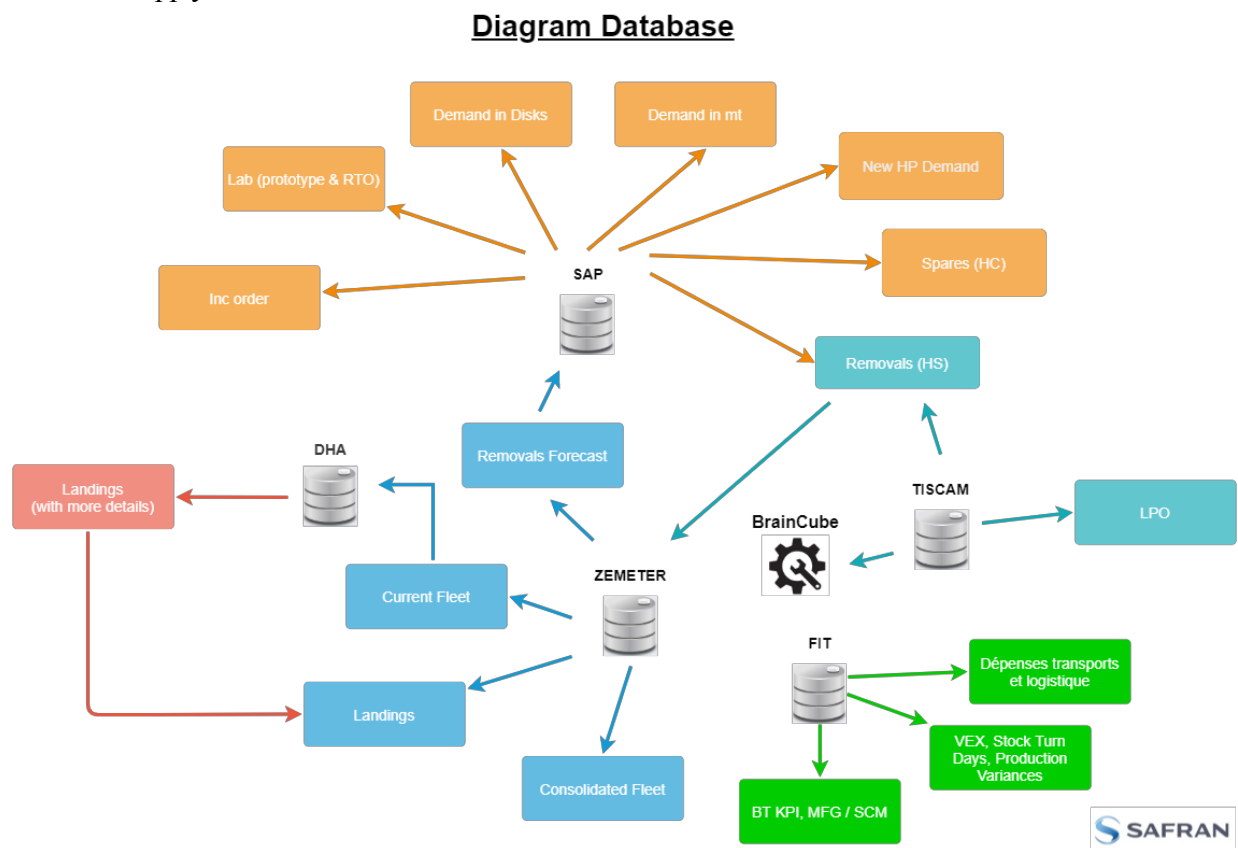


Figure 3. Diagramme avec les bases de données et leurs informations contenues

En plus de ce diagramme des bases de données, voici un autre diagramme créé dans le stage montrant la relation entre les bases de données, les experts et ses expertises. Ainsi avec toutes ces fonctionnalités et en réalisant des projets pour se familiariser plus avec les aspects techniques de l'entreprise, le processus d'étude de la gouvernance de données démarre une investigation pour augmenter la fiabilité de cette gestion, en automatisant les extractions, les analyses et les interconnexions. Dans la figure 4 illustre le diagramme fait pour illustrer les relations entre les BDD et ses experts.

Diagram Database & Experts

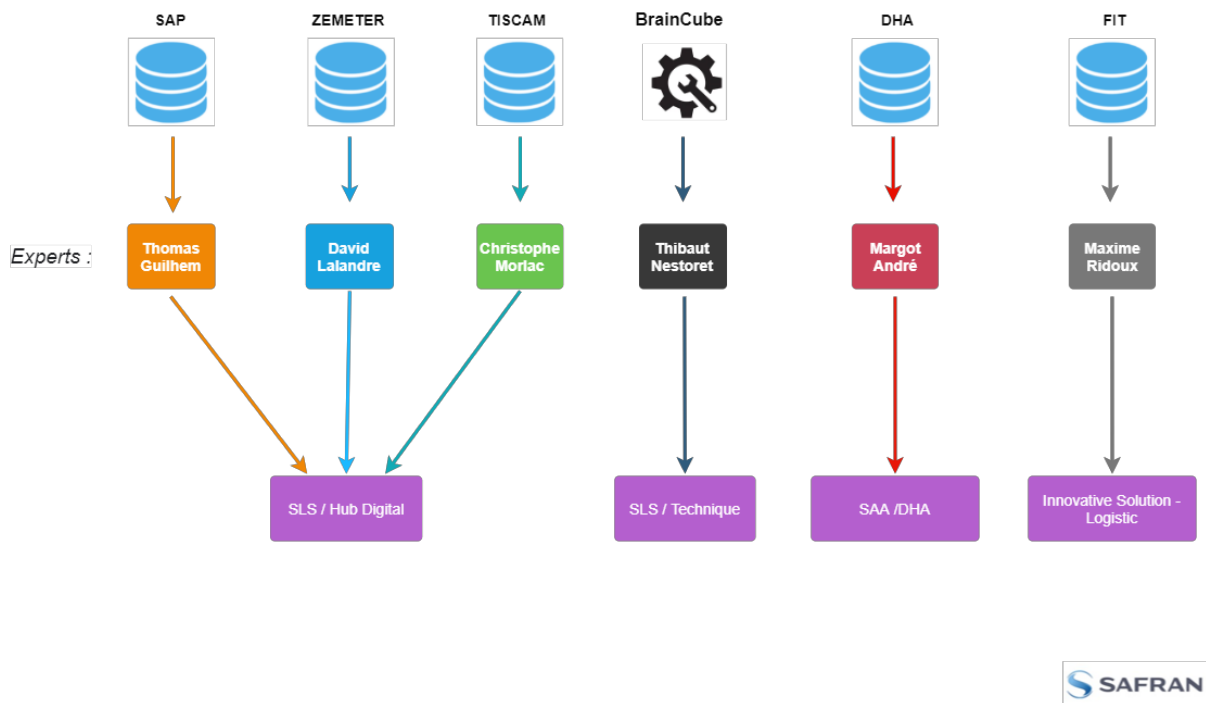


Figure 4. Diagramme des bases de données, ses experts et la division qu'elles sont utilisées

2.11 Projets

Pendant le stage, il y a eu l'opportunité de réaliser quatre projets différents et un d'entre eux est toujours en développement. La méthodologie CRISP-DM a été utilisée pour le développement de tous ces projets. Le résumé des cinq projets sont présentés ci-dessous avec certains détails.

- **Algorithme nettoyage de données :** Un projet pour réaliser le nettoyage des données de la base de données Tiscam, plus précisément, les informations de programme avion Zemeter et code la compagnie Zemeter. Ces informations sont le lien entre les données enregistrées sur Tiscam et les données enregistrées sur le Zemeter. Pendant le stage, une entreprise externalisée finissait ce projet et il fallait l'aider juste dans les dernières étapes du projet, c'est-à-dire, résoudre certains *bugs* et ajuster les bons paramètres et configurations pour exécuter le projet sur réseau interne et l'ordinateur Safran. En résumé, le projet utilise simplement programmation dans la langage python 3.8 et son installation était configurée pour être automatique à travers d'une exécution d'un fichier « .bat » et l'utilisation du projet par un autre exécutable « .bat ». A partir des informations du fichier d'entrée issu du Tiscam, un algorithme d'apprentissage *Machine Learning* qui utilise la technique « Random Forest » l'analysera et générera un fichier avec les prédictions des programmes avions et des codes clients nettoyés dans un fichier tabulé. Il y a aussi comme résultat un fichier avec sa matrice de confusion et ses informations de performances pour regarder la précision du résultat de nettoyage de chaque donnée. Le temps qu'il fallut dans le stage pour résoudre ce projet était d'environ 1 mois.

- Anomalies de consolidation LPO : Un projet pour corriger les erreurs d'une Macro Excel complexe qui est utilisée pour nettoyer les données LPO qui seront consolidées pour être enregistrées sur le logiciel Tiscam. Pour résoudre le problème, Il fallait refaire quelques parties de la macro et effectuer plusieurs tests afin de trouver la raison de l'erreur et ensuite, une bonne solution. Avec la solution réalisée, SLS a pu à nouveau consolider les bonnes données de LPO de façon plus vite et plus efficace. Le temps qu'il fallut dans le stage pour résoudre ce projet était d'environ 1 mois et le projet a permis de me familiariser avec les données LPO et Tiscam grâce le contact avec ces aspects.
- Dashboard LPO : Un tableau de bord Power BI qui utilise majoritairement des données avec les informations d'atterrissage d'avions, ses informations particulières (compagnie, date d'arrivée, modèle d'avion, type et fournisseur du frein etc.) et des mesures du LPO. La définition du LPO est « Landings Per Overhaul », c'est-à-dire la durée de vie du carbone du frein (puit de chaleur) en nombre d'atterrissages. De plus, les résultats du projet de correction de « Anomalies de Consolidation LPO » ont été un autre facteur qui a contribué à l'extraction de données propre, parce que il a aidé à effectuer le nettoyage dans les données Tiscam. Ce rapport a comme objectif principal de réduire le temps du processus de demande d'information et l'optimiser grâce à la possibilité de mise à jour les données automatiquement dans le dashboard et son filtrage. C'est la première fois qu'un PBI était connecté à cette base de données dans ce service. En règle générale, les commerciaux lors des négociations contrat avec les compagnies aériennes avait besoin de données LPO et d'autres informations connexes. Avant, il fallait demander à la « Central Supply Chain » d'envoyer ces informations et la subdivision « Carbon Supply Chain » devait chercher dans un fichier Excel, utiliser des filtres et faire des calculs pour envoyer les réponses au « Commercial ». Actuellement avec ce projet le département commercial lui-même peut avoir toutes les informations qu'il veut et avec plus de détails qu'avant (graphiques et autres mesures). Le temps qu'il fallut dans le stage pour résoudre ce projet était d'environ 1 mois et demi.
- Dashboard Traffic & HSE : Un tableau de bord Power BI qui permet au management haut plus au niveau du siège de Safran Landing Systems d'accéder aux informations pour construire des hypothèses de prévisions de reprise d'activité suite à la crise Covid. Les données utilisées sont très sensibles et confidentielles. Le dashboard présente principalement le trafic aérien et les rechanges carbones « Heat Sink Exchange », activités contribuant majoritairement à la marge de la filiale par rapport à 2019, référentiel avant crise. C'est-à-dire le comportement de la quantité d'échange de carbone et le comportement du trafic aérien dans le temps en comparaison avec 2019. Le temps qu'il fallut dans le stage pour résoudre ce projet était d'environ 2 mois et demi et le mémoire le présentera plus en détail ultérieurement.
- Dashboard HSE-Fleet-LPY-LPO : Un tableau de bord Power BI qui donnera la vue complète de la prévision HSE et des paramètres de suivi de sa prévision. Après avoir

le finaliser, il sera possible regarder les évolutions et des variations par rapport la flotte, les opérations (atterrissages) et LPO, en pouvant faire des comparaisons facilement avec les données du passé. Pour construire ce tableau de bord, il faut utiliser principalement les données « cycles » de Datahub Aero(DHA), les données de flotte de la base de données Zemeter, les données LPO de la base de données Tiscam et les données HSE de la base de données SAP. De plus, il sera possible de filtrer toutes ces informations par compagnies aériennes, types de programme avion et de réaliser des comparaisons mensuelles entre les plusieurs données qui sont extraites automatiquement par le Power BI. L'équipe *Carbon Supply Chain* chez Safran pourra facilement regarder la valeur moyenne des derniers mois de ces données, avoir un écart du dernier mois par rapport à cette moyenne, regarder les top 5 plus grosses variations en plus et en moins. Concernant les avantages que le projet peut apporter, avec lui Safran pourra partager ses informations avec l'ensemble des acteurs de la demande avec une automatisation dans le mise à jour de données et un support pour analyser plus rapidement et précisément, en aidant à encore mieux comprendre la demande carbone et d'améliorer la précision Forecast avec, évidemment, du gain de temps. Il y a d'autres analyses qui sont en discussion et sont encore en construction, puisque le stage se terminera le 8 octobre 2021 et ce projet sera le dernier à être livré. Le temps qui a été estimé dans l'étape pour résoudre ce projet était d'environ 1 mois et demi.

A titre d'illustration, il a été construit une relation entre les données utilisés pour les 4 projets déjà réalisés, ses extractions, la façon d'effectuer ces extractions et son temps idéal de rafraîchissement. La figure 5 montre les relations mentionnées.

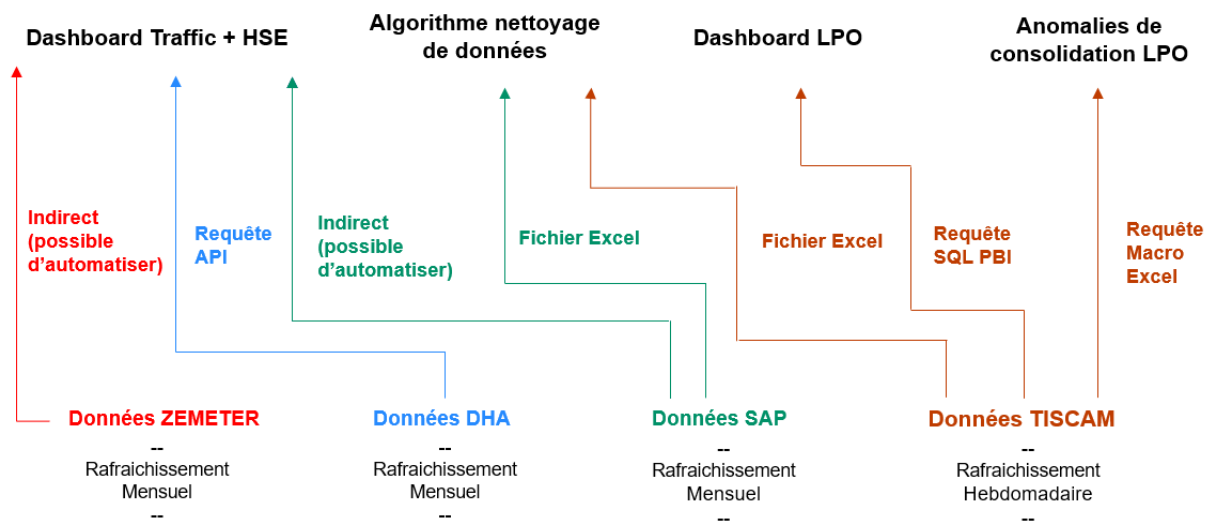


Figure 5. Schéma d'organisation des données par projets et les informations de ses extractions

Parmi les projets travaillés, il y a un projet qui a été réalisé pendant le stage, en partant de zéro, avec l'outil spécifique de traitement des données, appelé Power BI. Ce projet était le plus important du stage, le « Dashboard Traffic & HSE ». Dans cette perspective, le mémoire envisagera davantage le développement de ce projet, qui a apporté plus des valeurs à l'entreprise et rempli les objectifs et missions du stage de fin d'études.

3 Dashboard HSE & Traffic

3.1 Introduction

Pendant la formation du projet, il a fallu une lecture des présentations sur le système aéronautique Safran que le tuteur David Lalandre a réalisé. Ensuite, il a fallu identifier les données nécessaires pour le construire. Après cette identification, le processus d'extraction a été nécessaire, le nettoyage, l'intégration, etc... A chaque étape du processus de création et de réflexion, il y a eu plusieurs discussions entre le stagiaire et le tuteur d'entreprise pour avoir la certitude que la voie de développement était la bonne.

Le développement du projet a suivi la méthodologie CRISP-DM, ainsi, afin d'être plus claire l'explication, sa présentation dans le mémoire suivra les topics de cette méthodologie.

3.2 Business Understanding

Les équipes de management Safran utilisent des ressources techniques et ses connaissances pour mettre en place des activités stratégiques au sein de la compagnie. Dans cette optique, l'idée de créer un tableau de bord Power BI dynamique avec les analyses claires et pertinentes qui permet une mise à jour de données automatiquement et un filtrage efficace est une ressource très puissante pour aider la business intelligence. Le « dashboard Traffic + HSE » vient pour collaborer justement comme une ressource technique de traitement des données intéressantes aux managers.

Ce dashboard doit utiliser des données du trafic aérien, le nombre d'échange de puits de chaleur, les prévisions de ces échanges à travers un algorithme de forecast et les informations intrinsèques de chaque avion (modèles, compagnies, date d'arrivée aux aéroports, etc.). Toutes ces données au fil du temps et en mettant l'accent sur la comparaison avec 2019 et aujourd'hui.

L'innovation que l'entreprise a besoin réside dans la mise en place d'une API se connectant directement depuis PBI à une base de données de trafic aérien de Safran Analytics (SAA, filiale data science de Safran). Le but est de rendre les équipes autonomes et que le Dashboard se mette automatiquement à jour. Avant il fallait que SAA extrait des données sous format csv, envoyé à SLS, qui l'incorporait dans un PBI, puis la SC faisait un extrait csv pour l'injecter dans un fichier Excel.

3.3 Data Understanding

Pour construire le projet il y a eu besoin des informations issues de différentes bases de données. Premièrement, en résumé, les données utilisées dans le projet étaient extraites des documents en format Excel qui sont sur le réseau interne Safran et, il y a d'autres qui proviennent d'une API qui fait une extraction depuis les données mises à disposition par DataHub Aero (DHA). De plus, une API est utilisée pour extraire les données issues de la DHA par la requête « datahub_cycle ». Avec cette requête était possible d'extraire les données cycles par année et par mois et les cycles sont calculés à partir du segment de vol du datahub, en utilisant deux fonctions « getCycleByMonthAndYear » et « getAllCycleByMonthAndYear », qui permettent récupérer toutes les données mises à disposition pour la période souhaitée. Par conséquent, c'est possible de faire toutes les actualisations des données de façon automatique.

Dans les prochains topics, les données extraites dans le projet sont décrites séparément avec de détails sur leur exploration, leur extraction et leur qualité.

Observation : Les images des tables ne peuvent pas être publiées en raison de la présence de données confidentielles, en plus, il y aura des parties des codes qui seront cachés pour préserver les informations sensibles.

3.3.1 Budget

Les informations de budget peuvent se représenter en HSE et il y a un document Excel qui a été créé avec ces relations. Son importance est de permettre faire connaissance à la fréquence de changement de puit de chaleur par rapport au budget. La figure 6 montre le code utilisé pour réaliser cette extraction.

```
Source = Excel.Workbook(File.Contents("XXXXXXXXXXXXXXXXXXXXX"), null, true),
DATA Sheet = Source{[Item="DATA",Kind="Sheet"]}[Data],
```

Figure 6. Code d'extraction de l'ensemble de données Budget

Après avoir fait l'étude du document complet et discuté avec les experts du domaine, il a été identifié que ce document dispose de 4 ensembles de données importantes, elles sont présentées ci-dessous :

- PROGRAM : Informer le type de programme d'avion auquel les données correspondent ;
- SITE PP : La région que la donnée fait référence (Europe, Amérique, Asie) ;
- Date : La date de référence du donnée HSE ;
- HSE : La valeur de HSE.

3.3.2 Cycle

Le Cycle dispose des informations sur le trafic aérien global et sa pertinence est de pouvoir analyser le scénario global aérien, pouvant être comparé aux activités Safran. Ses ensembles de données les plus importantes sont :

- `ARRIVAL_DATE` : La date d'arrivée pour réaliser la dépose HSE ;
- `aircraft_registration` : Le nombre de registre d'avions, un code de référence pour les avions du trafic aérien mondial ;
- `operator_icao` : Le tricode qui représente en abrégé l'opérateur du vol, la compagnie aérienne.

Ses données proviennent du Datahub Aero. Il y a eu une requête réalisée pour les extraire. Les deux fonctions utilisées sont présentées ci-dessous.

3.3.2.1 Fonction « GetCycleMonthAndYear »

Le but de cette fonction est d'effectuer l'extraction des données « Cycle » via la requête « datahub_cycle » de l'API DHA pour des années et mois spécifiques. La figure 7 présente le code créé pour construire cette fonction.

```

let
Source = (year as number, month as number, page as number) as table => let
_request = "https://www.bls.gov/charts/tables/tables.aspx?lang=en&table=1000&year=" & year & "&month=" & month & "&page=" & page & "&format=json";
Source = Json.Document(Web.Contents(_request)),
_items = Source[_items],
DataRow = if List.Count(_items) > 1 then Table.FromList(_items, Splitter.SplitByNothing(), null, null, ExtraValues.Error) else
Error.Record
in
DataRow
in
Source

```

Figure 7. Code de la définition de la fonction « GetCycleMonthAndYear »

3.3.2.2 Fonction « GetAllCycleByMonthAndYear »

Le but de cette fonction est d'effectuer une demande la fonction « GetCycleMonthAndYear » plusieurs fois afin d'obtenir toutes les données « Cycle » et pas juste des années et mois spécifiques. La figure 8 montre le code crée pour construire cette fonction.

```
(year as number, month as number) as table =>
let
    Source = List.Generate( ()=> [Result= try getCycleByMonthAndYear(year,month,1) otherwise null, Page = 1], each [Result] <> null,
        each [Result = try getCycleByMonthAndYear(year,month,[Page]+1) otherwise null, Page = [Page]+1], each [Result]),
    RawData = Table.FromList(Source, Splitter.SplitByNothing(), null, null, ExtraValues.Error)
in
    RawData
```

Figure 8. Code de la définition de la fonction « GetAllCycleByMonthAndYear »

3.3.3 Data Fleet

Le Data Fleet dispose d'informations sur la flotte équipée de freins carbone Safran et des opérations de ces flottes, par exemple, les atterrissages. L'ensemble de données le plus importante pour le projet de ces données est :

- **Registration** : Nombre de registre de référence d'avions Safran.

Les données sont extraites d'un fichier Excel qui a obtenu les informations à partir de l'application Zemeter, parce que, malheureusement, pour le moment il n'est pas possible faire la connexion directe avec le Power BI pour les extraire, le fichier Excel est donc créé et les données sont extraites de lui. La figure 9 affiche le code qui réalise cette extraction.

```
let
Source = Excel.Workbook(File.Contents("C:\Users\jamesr\OneDrive\Documents\2022\2022 Data Fleet export Zm PBI_Sheet.xlsx"), null, true),
#"Data Fleet export Zm PBI_Sheet" = Source{[Item="Data Fleet export Zm PBI",Kind="Sheet"]} [Data],
```

Figure 9. Code de l'extraction de l'ensemble de données Data Fleet

3.3.4 Data SAP

Le « Data SAP » a les informations exclusivement des avions qui sont liées au service Safran, ainsi il sera possible d'analyser le HSE et ses aspects de région, date de dépose, type de flotte, etc...

Les ensembles de données les plus importantes sont :

- **Hub Allocation** : La région que la donnée fait référence (Europe, Amérique, Asie) ;
- **ARRIVAL_DATE** : La date d'arrivée pour réaliser la dépose HSE ;
- **PROG ZM** : La référence de codage de l'application Zemeter pour la donnée ;
- **HSE** : La valeur de HSE.

Les données sont extraites d'un fichier Excel qui dispose de données de la BDD SAP, parce que, malheureusement, pour le moment il n'est pas possible faire la connexion directe entre le Power BI et SAP. La figure 10 montre la partie du code qui réalise l'extraction de données du fichier Excel.

```
let
Source = Excel.Workbook(File.Contents(" [REDACTED ]"), null,
true),
#"Data SAP Export PBI_Sheet" = Source[[Item="Data SAP Export PBI",Kind="Sheet"]][Data],
```

Figure 10. Code de l'extraction de l'ensemble de données Data SAP

3.3.5 Hubs Allocation

Ensemble de données avec les informations des types de freins Safran, compagnies aériennes et ses références, par exemple, ses codes Zemeter, ses codes Tiscam, ses codes SAP, ses codes de programme avion, etc...

L'importance du « Hubs Allocation » est de permettre réaliser des tables de correspondance entre les différents ensembles de données et faciliter la modélisation du projet.

Pour effectuer son extraction, les commandes ont été utilisées et la source était un fichier Excel avec toutes les ensembles de données avec les références. La figure 11 montre la partie de son code qui réalise l'extraction.

```
let
Source = Excel.Workbook(File.Contents(" [REDACTED ]"), null, true),
#"Hubs Allocation_Sheet" = Source[[Item="Hubs Allocation",Kind="Sheet"]][Data],
```

Figure 11. Code de l'extraction de l'ensemble de données Hubs Allocation

3.3.6 Immat

Le « Immat » est une table avec les correspondances entre le code « Registration » des données qui proviennent du Zemeter et le code « aircraft_registration » qui proviennent de la DHA. Avec cette table, il est possible de modéliser la liaison entre les deux ensembles de données « Data Fleet » et « Cycle » par ses codes de registre.

Pour l'obtenir, il a été fait une extraction d'un fichier Excel créé avec ces informations. Cependant, comme l'emplacement de ce document n'a pas été concrètement défini et peut généralement être modifié, une variable « input » a été créée pour définir l'emplacement du fichier et il suffit juste de changer cette variable pour mettre à jour, sans avoir à toucher au code.

3.3.7 S&OP

Les données S&OP sont les prévisions forecast de HSE qui ont été réalisés pour étudier les possibles performances des activités de l'avenir. L'importance de cet ensemble de données dans le projet est de permettre visualiser les résultats de façon plus claire et, aussi les comparer avec les données HSE réelles.

Il a été extrait d'un document Excel, ainsi que d'autres données précédemment mentionnées dans le mémoire. Ce document Excel est fréquemment mis à jour de nouvelles valeurs de prévision par SLS.

Les ensembles de données plus importantes sont :

- Date : La date de référence de la prévision HSE calculé ;
- SITE PP : La région que la donnée fait référence (Europe, Amérique, Asie) ;

- PROGRAM : Informer le type de programme d'avion auquel les données correspondent ;

3.4 Data Préparation

Pour traiter les données, Power BI dispose de l'éditeur power query, qui permet d'effectuer de code en langage M, DAX, Python ou R. Le langage utilisé pour développer les codes du projet était « M » (Mashup), parce qu'il est une bonne solution ETL (*Extracting, Transforming, Loading*) en mode libre-service et largement utilisé pour automatiser le processus d'exportation de données de différentes sources et d'importation dans Excel ou Power BI. De plus, il peut facilement transformer les données ainsi importées afin qu'elles puissent convenir aux besoins d'analyse et charger ces données dans une ou des tables, qui seront éventuellement utilisées par un modèle de données. Autrement dit, le langage M qui a été choisi pour le traitement est un langage de programmation structuré, étape par étape, où chaque étape va généralement lire dans la précédente.

Une fois les sélections, les nettoyages, les constructions des données et les intégrations d'eux, d'autres formations et modifications peuvent être effectuées en dehors de la fenêtre de « éditeur power query », dedans directement du projet dans l'onglet, par exemple, « Données » du PBI. Le langage de programmation couramment utilisé dans cet environnement est DAX, qui est un langage d'expression de formule qui comprend des fonctions, des opérateurs et des valeurs qui permettent d'effectuer des requêtes et des calculs complexes sur des données de colonnes et tables déjà intégrées dans le projet ^[5].

Les données extraites et les données construites sont organisées par table dans l'application Power BI et dans le total, il y a 18 tables créées. Ensuite, chaque table de données est décrite avec des détails sur les parties les plus précieuses pour le projet.

Observation : Les images des tables ne peuvent pas être publiées en raison de la présence de données confidentielles, en plus, il y aura des parties des codes qui seront cachés pour préserver les informations sensibles.

3.4.1 Hubs Allocation

Dans cette table, il a été créé les valeurs de référence avec les différents codes de « Key Tiscam_Country » et de « Key Zm_Country » pour réaliser les créations des données de référence pour lier les tables « Data SAP » et « Cycle ». Pour construire les deux colonnes avec ces valeurs, les formules sont :

```
Key Tiscam_Country = 'Hubs Allocation'[Key TisCam code]&'Hubs Allocation'[Country]
```

```
Key Zm_Country = 'Hubs Allocation'[Key TisCam code]&'Hubs Allocation'[Country]
```

3.4.2 Date Last Refreshed

Il est important d'avoir la date de la dernière mise à jour du projet, afin de savoir toujours à quel jour, mois, année et à quelle heure ses données présentées se réfèrent. À cette fin, ces données ont été créées à l'aide du code illustré à la figure 12.

```
let
Source = #table(type table[Date Last Refreshed=datetime], {{DateTime.LocalNow()}})
in
Source
```

Figure 12. Code pour construire la table Date Last Refreshed


```

let
Source = Excel.Workbook(File.Contents(" [REDACTED]"), null, true),
DATA_Sheet = Source[Item="DATA",Kind="Sheet"][Data],
#"Type modifié" = Table.TransformColumnTypes(DATA_Sheet,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4",
type any}, {"Column5", type text}, {"Column6", type text}, {"Column7", type any}, {"Column8", type any}, {"Column9", type any}, {"Column10",
type any}, {"Column11", type any}, {"Column12", type any}, {"Column13", type any}, {"Column14", type any}, {"Column15", type any},
{"Column16", type any}, {"Column17", type any}, {"Column18", type any}, {"Column19", type any}, {"Column20", Int64.Type}, {"Column21", type
any}, {"Column22", type any}, {"Column23", type any}, {"Column24", type any}, {"Column25", type any}, {"Column26", type any}, {"Column27",
type any}, {"Column28", type any}, {"Column29", type any}, {"Column30", type any}, {"Column31", type any}, {"Column32", type any},
{"Column33", type number}, {"Column34", type any}, {"Column35", type any}, {"Column36", type any}, {"Column37", type any}, {"Column38", type
any}, {"Column39", type any}, {"Column40", type any}, {"Column41", type any}, {"Column42", type any}, {"Column43", type any}, {"Column44",
type any}, {"Column45", type any}, {"Column46", type number}, {"Column47", type any}, {"Column48", type any}, {"Column49", type any},
{"Column50", type any}, {"Column51", type any}, {"Column52", type any}, {"Column53", type any}, {"Column54", type any}, {"Column55", type
any}, {"Column56", type any}, {"Column57", type any}, {"Column58", type any}, {"Column59", type number}, {"Column60", type number},
{"Column61", type number}, {"Column62", type number}, {"Column63", type number}}},
#"Colonnes supprimées" = Table.RemoveColumns(#"Type modifié",{"Column59", "Column60", "Column61", "Column62", "Column63", "Column46", "Column33",
"Column20"}),
#"Premières lignes supprimées" = Table.Skip(#"Colonnes supprimées",2),
#"En-têtes promus" = Table.PromoteHeaders(#"Premières lignes supprimées", [PromoteAllScalars=true]),
#"Type modifié1" = Table.TransformColumnTypes(#"En-têtes promus",{{"Data type", type text}, {"S&OP", type text}, {"Trafic Curve", type text},
{"Prog", type any}, {"Role type SC", type text}, {"SITE PP", type text}, {"PROGRAM", type any}, {"01/01/2019", Int64.Type}, {"01/02/2019",
Int64.Type}, {"01/03/2019", Int64.Type}, {"01/04/2019", Int64.Type}, {"01/05/2019", Int64.Type}, {"01/06/2019", Int64.Type}, {"01/07/2019",
Int64.Type}, {"01/08/2019", Int64.Type}, {"01/09/2019", Int64.Type}, {"01/10/2019", Int64.Type}, {"01/11/2019", Int64.Type}, {"01/12/2019",
Int64.Type}, {"01/01/2020", type number}, {"01/02/2020", type number}, {"01/03/2020", type number}, {"01/04/2020", type number}, {"01/05/
2020", type number}, {"01/06/2020", type number}, {"01/07/2020", type number}, {"01/08/2020", type number}, {"01/09/2020", type number},
{"01/10/2020", type number}, {"01/11/2020", type number}, {"01/12/2020", type number}, {"01/01/2021", type any}, {"01/02/2021", type number},
{"01/03/2021", type number}, {"01/04/2021", type number}, {"01/05/2021", type number}, {"01/06/2021", type number}, {"01/07/2021", type
number}, {"01/08/2021", type number}, {"01/09/2021", type number}, {"01/10/2021", type number}, {"01/11/2021", type number}, {"01/12/2021",
type number}, {"01/01/2022", type number}, {"01/02/2022", type number}, {"01/03/2022", type number}, {"01/04/2022", type number}, {"01/05/
2022", type number}, {"01/06/2022", type any}, {"01/07/2022", type any}, {"01/08/2022", type any}, {"01/09/2022", type any}, {"01/10/2022",
type any}, {"01/11/2022", type any}, {"01/12/2022", type any}}},
#"Supprimer le tableau croisé dynamique des autres colonnes" = Table.UnpivotOtherColumns(#"Type modifié1", {"PROGRAM", "SITE PP", "Role type SC",
"Prog", "Trafic Curve", "S&OP", "Data type"}, "Attribut", "Valeur"),
#"Colonnes renommées" = Table.RenameColumns(#"Supprimer le tableau croisé dynamique des autres colonnes",{{"Attribut", "Date"}}),
#"Type modifié2" = Table.TransformColumnTypes(#"Colonnes renommées",{{"Date", type date}}),
#"Colonnes renommées1" = Table.RenameColumns(#"Type modifié2",{{"Valeur", "HSE"}}),
#"Colonnes supprimées1" = Table.RemoveColumns(#"Colonnes renommées1",{"Data type"}),
#"Lignes filtrées1" = Table.SelectRows(#"Colonnes supprimées1", each Text.StartsWith([#"S&OP"], "S&OP-") and not Text.Contains([#"S&OP"], "+")),
#"Valeur remplacée" = Table.ReplaceValue(#"Lignes filtrées1", "S&OP-", "", Replacer.ReplaceText, {"S&OP"}),
#"Colonnes renommées2" = Table.RenameColumns(#"Valeur remplacée",{{"S&OP", "S&OP -M"}}),
#"Type modifié3" = Table.TransformColumnTypes(#"Colonnes renommées2",{{"HSE", Int64.Type}})
in
#"Type modifié3"

```

Figure 16. Code pour construire la table de données « S&OP-M »

3.4.7 Calendar

Cette table est faite à partir de données déjà intégrées dans le projet, c'est-à-dire, la table n'est pas faite par de code en langage M dans la fenêtre « éditeur power query » du PBI. Elle a été construite en langage DAX et pour regrouper les différentes dates des tables qui présentent ce type de données dans le projet, il y aura donc une table regroupant les données en commun des différentes tables, facilitant ainsi son utilisation en modélisation. La figure 17 affiche le code pour la construction de cette table.

```
Calendar = DISTINCT(UNION(DISTINCT(Cycle[ARRIVAL_DATE]), DISTINCT('Data SAP'[ARRIVAL_DATE]), DISTINCT('S&OP -M'[Date]), DISTINCT('S&OP -M'[Date]), DISTINCT(Budget[Date]))))
```

Figure 17. Code pour construire la table « Calendar »

3.4.8 Cycle

Les données « Cycle » ont des différentes colonnes créées à partir d'autres tables intégrées dans le projet avec l'objectif principal de référencer les données des activités Safran avec les données des activités globales, c'est-à-dire, réaliser la liaison entre les informations de « Data Fleet » et de « Cycle » avec interception des données de registre « Immat ».

Afin de pouvoir effectuer les modélisations avec le but d'identifier les flottes Safran dans l'ensemble de données cycle, les colonnes créées sont :

- Brake Type : Des données de type de frein issues de « Data Fleet » à travers de la relation qu'il y aura entre les « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule de la construction de cette colonne est :

Brake Type = RELATED('Data Fleet'[Brake Type])

- Current Op : Des données de compagnies aériennes issues de « Data Fleet » à travers de la relation qu'il y aura entre le « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule pour la construction de cette colonne est :

Current Op = RELATED('Data Fleet'[Current Op])

- Current Op Ctry : Les données des pays que l'activité aérienne fait référence issues de « Data Fleet » à travers de la relation qu'il y aura entre le « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule pour la construction de cette colonne est :

Current Op Ctry = RELATED('Data Fleet'[Current Op Ctry])

- Cur Op Wreg : Les données du continent que l'activité aérienne fait référence issues de « Data Fleet » à travers de la relation qu'il y aura entre le « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule pour la construction de cette colonne est :

Cur Op Wreg = RELATED('Data Fleet'[Cur Op Wreg])

- Current Brakes : Les données de freins qui sont en service dans l'activité aérienne issues de « Data Fleet » à travers de la relation qu'il y aura entre le « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule pour la construction de cette colonne est :

Current Brakes = RELATED('Data Fleet'[Current Brakes])

- Code Tiscam : Le code de référence Tiscam. La formule pour la construction de cette colonne est :

Code Tiscam = [Brake Type]&[operator_icao]

- Registration : Le nombre de registre d'avions qui sont liée aux activités Safran. Si la valeur est vide, c'est-à-dire que l'avion n'appartient pas à Safran. La formule pour la construction de cette colonne est :

Registration = RELATED(immat[0_Registration])

- Landings : La quantité d'atterrissage que chaque avion a fait. La formule ci-dessous pour construire cette colonne est simple, les données étaient juste copiées d'autre colonne de la table « Cycle » comme valeurs entières, parce que la quantité d'atterrissage doit être entière:

Landings = INT(Cycle[cycle])

- **Current Brakes** : Les type de freins en service qui sont issues de « Data Fleet » à travers de la relation qu'il y aura entre le « Data Fleet » et « Cycle » dans la modélisation, en référence au registre des aéronefs entre les deux. La formule pour la construction de cette colonne est :

`Current Brakes = RELATED('Data Fleet'[Current Brakes])`

- **ARRIVAL_DATE** : La date d'arrivée de l'avion et la colonne a été créée à partir des colonnes existantes déjà de l'année et du mois dans le « Cycle ». La formule pour la construction de la colonne est :

`ARRIVAL_DATE = date(Cycle[year],Cycle[month],1)`

- **Key Tiscam_Country** : Un code créé pour lier les informations du codage Tiscam et des pays d'une même donnée d'activité aérienne, afin de pouvoir réaliser la modélisation entre les tables « Cycle » et « Data SAP ». La formule pour construire cette colonne est :

`Code Tiscam = [Brake Type]&[operator_icao]`

3.4.9 Data SAP

Cette table contient les informations de l'ensemble de données « Data SAP » récupéré pour traitement. Le document Excel extrait était déjà propre, parce qu'il y avait un processus de nettoyage et d'organisation du format idéal que ce document devait avoir afin que les données de la BDD SAP soient constamment mises à jour facilement. Le traitement donc n'est pas de tout nécessaire, comme il est possible de regarder dans la figure 18.

```
let
Source = Excel.Workbook(File.Contents("C:\Users\user\AppData\Local\Microsoft\Excel\Excel12.xlsm"), null, true),
#"Data SAP Export PBI_Sheet" = Source[[Item="Data SAP Export PBI",Kind="Sheet"]][Data],
#"En-têtes promus" = Table.PromoteHeaders("#Data SAP Export PBI_Sheet", [PromoteAllScalars=true]),
#"Type modifié" = Table.TransformColumnTypes("#En-têtes promus",{{"HP_ID", type text}, {"PART_NUMBER", type text}, {"ARRIVAL_DATE", type date}, {"COMPANY", type text}, {"REPAIR_ORDER", type any}, {"Hub allocation", type text}, {"AIRCRAFT_SERIAL", type any}, {"PROG ZM", type any}, {"Code Zm", type text}, {"A/L designation", type text}, {"Country", type text}, {"Area", type text}, {"Regroupement Tiscam", type any}, {"Key Data SAP", type text}, {"HSE", Int64.Type}})
in
#"Type modifié"
```

Figure 18. Code pour construire la table « Data SAP »

Après avoir intégrés la table dans le projet, il a été créé une autre colonne importante pour réaliser le processus de modélisation de données entre « Data SAP » et « Cycle ». Les valeurs de cette colonne sont l'ensemble de code de référence Zemeter et le pays de la donnée. Sa formule de construction est :

`Key Zm_Country = 'Data SAP'[Key Data SAP]&'Data SAP'[Country]`

3.4.10 Match Cycle & Data SAP

Une table créée pour organiser les correspondances entre « Key Tiscam_Country » issu de la table « Cycle » et « Key Zm_Country » issu de la table « Data SAP ». Avec cela, il sera possible construire des relations entre ces données dans le processus de modélisation du projet. La formule pour construire cette table est :

`Match Cycle_DataSAP = DISTINCT(CALCULATETABLE(SELECTCOLUMNS('Hubs Allocation', "Key Tiscam_Country", [Key Tiscam_Country], "Key Zm_Country", [Key Zm_Country]), 'Hubs Allocation'[Key Tiscam_Country] <> BLANK(), 'Hubs Allocation'[Key Zm_Country] <> BLANK()))`

3.4.11 Match Cycle

Afin de réaliser la liaison entre les données « Cycle » et « Data SAP » dans la modélisation, il y a des éléments qui intercèdent la connexion entre eux, parce que des valeurs uniques sont nécessaires dans la table qui coupe la liaison entre « Cycle » et « Match Cycle & Data SAP ». La donnée de la table « Cycle » qui est importante pour effectuer la liaison est le « Key Tiscam_Country ». La table « Match Cycle » est donc construite pour disposer juste les différentes valeurs de « Key Tiscam_Country ». La figure 19 affiche le code de construction de cette table.

```
Match Cycle = DISTINCT(CALCULATETABLE(SELECTCOLUMNS(Cycle, "Code Tiscam_Country", [Key Tiscam_Country], Cycle[Key Tiscam_Country] <> BLANK())))
```

Figure 19. Code de construction de la table « Match Cycle »

3.4.12 Budget

Les données extraites pour construire la table « Budget » e ses formations sont intégrés au projet à travers du code affiché dans la figure 20.

```
let
Source = Excel.Workbook(File.Contents("XXXXXXXXXXXXXXXXXXXX"), null, true),
DATA_Sheet = Source[Item="DATA",Kind="Sheet"][Data],
#"Type modifié" = Table.TransformColumnTypes(DATA_Sheet,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type any}, {"Column5", type text}, {"Column6", type text}, {"Column7", type any}, {"Column8", type any}, {"Column9", type any}, {"Column10", type any}, {"Column11", type any}, {"Column12", type any}, {"Column13", type any}, {"Column14", type any}, {"Column15", type any}, {"Column16", type any}, {"Column17", type any}, {"Column18", type any}, {"Column19", type any}, {"Column20", type any}, {"Column21", type any}, {"Column22", type any}, {"Column23", type any}, {"Column24", type any}, {"Column25", type any}, {"Column26", type any}, {"Column27", type any}, {"Column28", type any}, {"Column29", type any}, {"Column30", type any}, {"Column31", type any}, {"Column32", type any}, {"Column33", type number}, {"Column34", type any}, {"Column35", type any}, {"Column36", type any}, {"Column37", type any}, {"Column38", type any}, {"Column39", type any}, {"Column40", type any}, {"Column41", type any}, {"Column42", type any}, {"Column43", type any}, {"Column44", type any}, {"Column45", type any}, {"Column46", type number}, {"Column47", type any}, {"Column48", type any}, {"Column49", type any}, {"Column50", type any}, {"Column51", type any}, {"Column52", type any}, {"Column53", type any}, {"Column54", type any}, {"Column55", type any}, {"Column56", type any}, {"Column57", type any}, {"Column58", type any}, {"Column59", type number}, {"Column60", type number}, {"Column61", type number}, {"Column62", type number}, {"Column63", type number}}},
#"Colonnes supprimées" = Table.RemoveColumns(#"Type modifié",{"Column59", "Column60", "Column61", "Column62", "Column63", "Column46", "Column33", "Column20"}),
#"Premières lignes supprimées" = Table.Skip(#"Colonnes supprimées",2),
#"En-têtes promus" = Table.PromoteHeaders(#"Premières lignes supprimées", [PromoteAllScalars=true]),
#"Type modifié1" = Table.TransformColumnTypes(#"En-têtes promus",{{"Data type", type text}, {"S&OP", type text}, {"Trafic Curve", type text}, {"Prog", type any}, {"Role type SC", type text}, {"SITE PP", type text}, {"PROGRAM", type any}, {"01/01/2019", Int64.Type}, {"01/02/2019", Int64.Type}, {"01/03/2019", Int64.Type}, {"01/04/2019", Int64.Type}, {"01/05/2019", Int64.Type}, {"01/06/2019", Int64.Type}, {"01/07/2019", Int64.Type}, {"01/08/2019", Int64.Type}, {"01/09/2019", Int64.Type}, {"01/10/2019", Int64.Type}, {"01/11/2019", Int64.Type}, {"01/12/2019", Int64.Type}, {"01/01/2020", type number}, {"01/02/2020", type number}, {"01/03/2020", type number}, {"01/04/2020", type number}, {"01/05/2020", type number}, {"01/06/2020", type number}, {"01/07/2020", type number}, {"01/08/2020", type number}, {"01/09/2020", type number}, {"01/10/2020", type number}, {"01/11/2020", type number}, {"01/12/2020", type number}, {"01/01/2021", type any}, {"01/02/2021", type number}, {"01/03/2021", type number}, {"01/04/2021", type number}, {"01/05/2021", type number}, {"01/06/2021", type number}, {"01/07/2021", type number}, {"01/08/2021", type number}, {"01/09/2021", type number}, {"01/10/2021", type number}, {"01/11/2021", type number}, {"01/12/2021", type number}, {"01/01/2022", type number}, {"01/02/2022", type number}, {"01/03/2022", type number}, {"01/04/2022", type number}, {"01/05/2022", type number}, {"01/06/2022", type any}, {"01/07/2022", type any}, {"01/08/2022", type any}, {"01/09/2022", type any}, {"01/10/2022", type any}, {"01/11/2022", type any}, {"01/12/2022", type any}}},
#"Supprimer le tableau croisé dynamique des autres colonnes" = Table.UnpivotOtherColumns(#"Type modifié1", {"PROGRAM", "SITE PP", "Role type SC", "Prog", "Trafic Curve", "S&OP", "Data type"}, "Attribut", "Valeur"),
#"Colonnes renommées" = Table.RenameColumns(#"Supprimer le tableau croisé dynamique des autres colonnes",{{"Attribut", "Date"}}),
#"Type modifié2" = Table.TransformColumnTypes(#"Colonnes renommées",{{"Date", type date}}),
#"Colonnes renommées1" = Table.RenameColumns(#"Type modifié2",{{"Valeur", "HSE"}}),
#"Colonnes supprimées1" = Table.RemoveColumns(#"Colonnes renommées1",{"Data type"}),
#"Lignes filtrées1" = Table.SelectRows(#"Colonnes supprimées1", each Text.StartsWith([#S&OP], "Budget")),
#"Colonnes renommées2" = Table.RenameColumns(#"Lignes filtrées1",{{"S&OP", "Data Budget"}}),
#"Type modifié3" = Table.TransformColumnTypes(#"Colonnes renommées2",{{"HSE", Int64.Type}}),
#"Removed Columns" = Table.RemoveColumns(#"Type modifié3",{"Data Budget"})
in
#"Removed Columns"
```

Figure 20. Code pour la construction de la table « Budget »

3.4.13 Match Data SAP

Cette table créée est chargée d'intercéder la connexion entre « Data SAP » et la table de correspondance de « Data SAP » et « Cycle ». Elle contient les différentes valeurs de la colonne « Key Zm_Country » de la table « Data SAP ». Sa formule de construction est :

```
Match DataSAP = DISTINCT(CALCULATETABLE(SELECTCOLUMNS('Data SAP', "Code Zm_Country", [Key  
Zm_Country])), 'Data SAP'[Key Zm_Country] <> BLANK()))
```

3.4.14 Site PP

Une table avec les différentes valeurs de « Site PP » que toutes les tables disposent, c'est-à-dire, les différentes valeurs de hubs allocation. La formule pour la création de la table est :

```
Site PP = CALCULATETABLE(DISTINCT(SELECTCOLUMNS('Data SAP', "SITE PP", [Hub allocation])),  
'Data SAP'[Hub allocation] <> "Montage")
```

3.4.15 Match Program & Data SAP

Cette table créée est chargée d'intercéder la connexion entre « Correspondance S&OP & Budget » et « Data SAP » avec les différentes valeurs de programme avion qu'il y a dans le « Data SAP ». La formule pour la construction est :

```
Match Program DataSAP = DISTINCT(SELECTCOLUMNS('Correspondance S&OP & Budget', "Program",  
[Programme DataSAP Original]))
```

3.4.16 Data Fleet

Pour réaliser la formation, construction et intégration de cette table, le code de la figure 21 a été utilisé.

```
let  
Source = Excel.Workbook(File.Contents("x"), null, true),  
#"Data Fleet export Zm PBI_Sheet" = Source[Item="Data Fleet export Zm PBI_Sheet",Kind="Sheet"][Data],  
#"En-têtes promus" = Table.PromoteHeaders(#"Data Fleet export Zm PBI_Sheet", [PromoteAllScalars=true]),  
#"Type modifié" = Table.TransformColumnTypes(#"En-têtes promus",{{"A/C Serial", type text}, {"Line Number", Int64.Type}, {"Model",  
type text}, {"Brake Type", type text}, {"A/C Program MBD", type text}, {"Current Op", type text}, {"Current_Operator_3L", type  
text}, {"Registration_Origine", type text}, {"Current Op Ctry", type text}, {"Cur Op Wreg", type text}, {"Owner", type text},  
{"Role", type text}, {"Original Delivery Date", type date}, {"Retrofit Date", type date}, {"Operator Delivery Date", type date},  
{"EIS Date", type date}, {"Original Brakes", type text}, {"Current Brakes", type text}, {"Retrofit_FAL", type text}, {"Comment",  
type text}, {"Status", type text}, {"Status_MBD", type text}, {"Key Data SAP", type text}, {"LigneOk", type text},  
{"Registration", type text}, {"zzz", type any}}),  
#"Lignes filtrées" = Table.SelectRows(#"Type modifié", each ([LigneOk] = "Ok") and ([Status] = "Active" or [Status] = "Parked") and (  
[Status_MBD] = "DELIVERED"))  
in  
#"Lignes filtrées"
```

Figure 21. Code de construction pour intégrer la table « Data Fleet »

3.5 Modeling

Dans le processus de modélisation des tables de données, plusieurs relations entre eux ont été réalisées, toutes avec l'objectif de les interconnecter afin de faire un projet organisé, efficace, interactif et dynamique.

La modélisation permet également d'effectuer les filtres, parce que des tables spécifiques qui n'ont que les valeurs distinctes d'un certain groupe ou type de données, peuvent être liées à

toutes les autres tables qui ont ce certain groupe ou type, telles que les tables : « Site PP » et « Calendar ».

Dans ce processus, les interconnexions entre les tables ont été effectuées, ce qui peut être vu sur la figure 22. Les ensembles de données qui ont fait la connexion dans chaque table sont représentés entre parenthèses sous forme de colonnes.

Active	De : Table (Colonne)	À : Table (Colonne)
✓	Budget (Date)	Calendar (ARRIVAL_DATE)
✓	Budget (PROGRAM)	Correspondence S&OP & Budget (Programme S&OP/Budget Original)
✓	Budget (SITE PP)	Site PP (SITE PP)
✓	Correspondence S&OP & Budget (Programme DataSAP Original)	Correspondence Role & Program (PROGRAM)
✓	Correspondence S&OP & Budget (Programme S&OP/Budget Original)	Match Program DataSAP (Program)
✓	Cycle (aircraft_registration)	immat (Registration - clean)
✓	Cycle (ARRIVAL_DATE)	Calendar (ARRIVAL_DATE)
✓	Cycle (Key Tiscam_Country)	Match Cycle (Code Tiscam_Country)
✓	Data Fleet (Registration)	immat (O_Registration)
✓	Data SAP (ARRIVAL_DATE)	Calendar (ARRIVAL_DATE)
✓	Data SAP (Hub allocation)	Site PP (SITE PP)
✓	Data SAP (Key Zm_Country)	Match DataSAP (Code Zm_Country)
✓	Data SAP (PROG ZM)	Match Program DataSAP (Program)
✓	Match Cycle (Code Tiscam_Country)	Match Cycle_DataSAP (Key Tiscam_Country)
✓	Match DataSAP (Code Zm_Country)	Match Cycle_DataSAP (Key Zm_Country)
✓	S&OP +M (Date)	Calendar (ARRIVAL_DATE)
✓	S&OP +M (PROGRAM)	Correspondence S&OP & Budget (Programme S&OP/Budget Original)
✓	S&OP +M (SITE PP)	Site PP (SITE PP)
✓	S&OP -M (Date)	Calendar (ARRIVAL_DATE)
✓	S&OP -M (PROGRAM)	Correspondence S&OP & Budget (Programme S&OP/Budget Original)
✓	S&OP -M (SITE PP)	Site PP (SITE PP)

Figure 22. Relations entre les données des tables dans la modélisation

Pour aider à visualiser ces interconnexions, la figure 23 présente la conception de la modélisation réalisée.

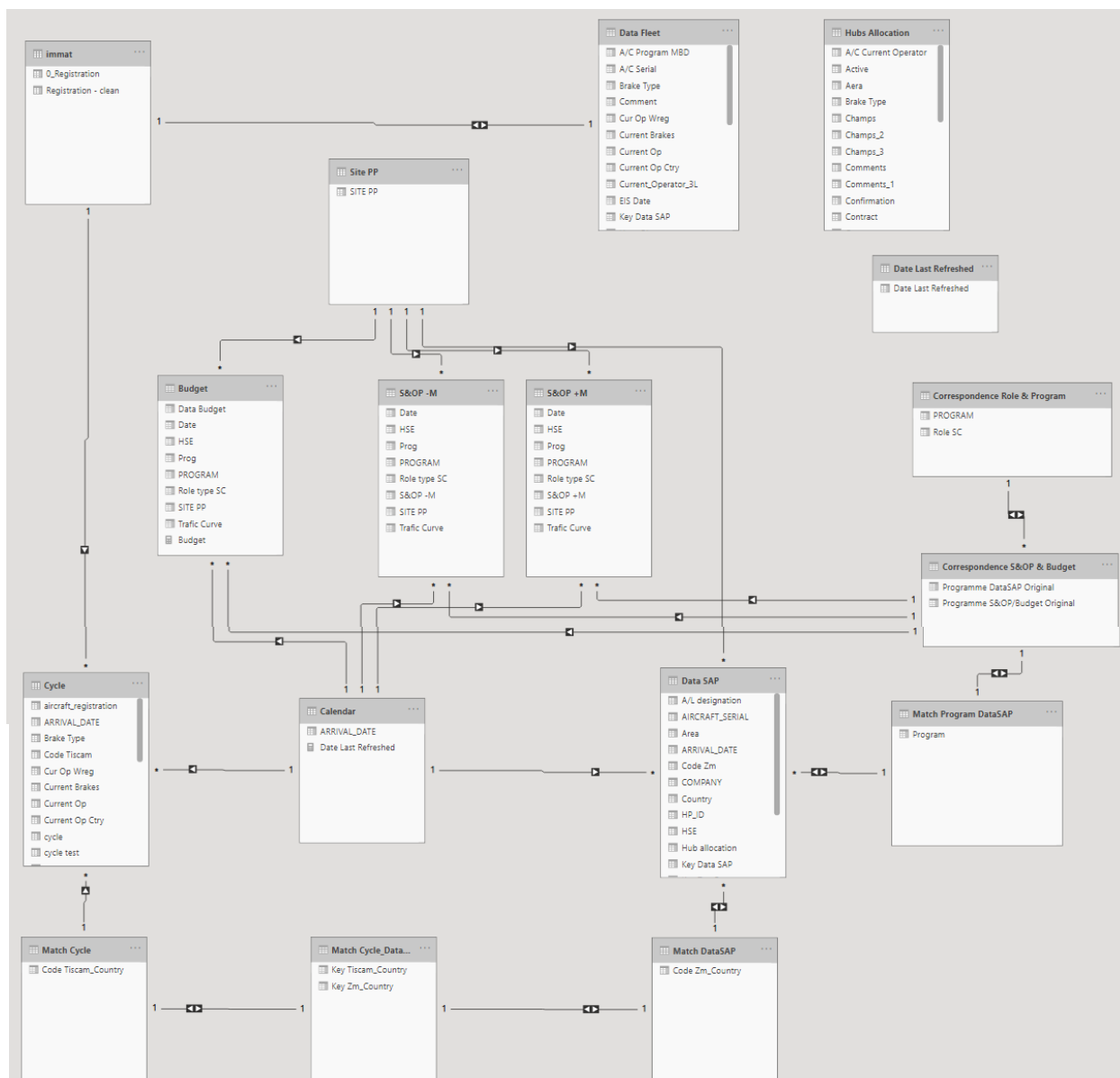


Figure 23. Design de la modélisation réalisée pour le projet

Un détail important que l'on peut noter est le lien entre « Data Fleet » et « Cycle » à travers la table « immat », cela permet aux colonnes créées dans la table « Cycle » de fonctionner, parce que il sera possible identifier les avions et trafiques qui sont liées à Safran dans le « Cycle ».

Les flèches à double sens dans certains cas dans les connexions signifient que les deux tables connectées peuvent s'influencer en cas de filtrage direct par le projet.

3.6 Evaluation

Après avoir évalué avec le tuteur que tous les processus de développement précédents étaient corrects, les étapes finales du développement du projet ont été réalisées avec toute le design approprié et conformément aux normes fixées par l'entreprise. Ensuite, les indicateurs ont été

réalisés, qui sont dérivés de mesures construites à partir des données des tables. Ces indicateurs sont expliqués ci-dessous.

- HSE : Représente la quantité d'échange de puit de chaleur ;
- S&OP +M : Représente les prévisions HSE les plus récentes de l'algorithme forecast ;
- Budget : Représente le Budget en HSE ;
- % HSE : Représente le pourcentage de HSE par rapport la moyenne HSE de 2019 ;
- % SLS Traffic : Représente le pourcentage de trafic de vols des compagnies aériennes qui travaillent avec Safran par rapport la moyenne de trafic aérienne de ces compagnies en 2019 ;
- % Traffic : Représente le pourcentage de vols des tous les compagnies aériennes par rapport la moyenne de trafic aérienne de ces compagnies en 2019 ;
- % S&OP –M : Représente le pourcentage des anciennes prévisions HSE de l'algorithme forecast par rapport le pourcentage de prévisions de 2019.

Pour réaliser la révision de chaque partie du projet et présenté, ils seront organisés séparément dans ce mémoire.

- **Graphique principal** : La visualisation du dashboard présente un graphique avec les 7 types d'indicateurs différents. La figure 24 affiche ce graphique avec les informations sensibles floutées.

S&OP -M: 2020-12 and S&OP +M: 2021-04

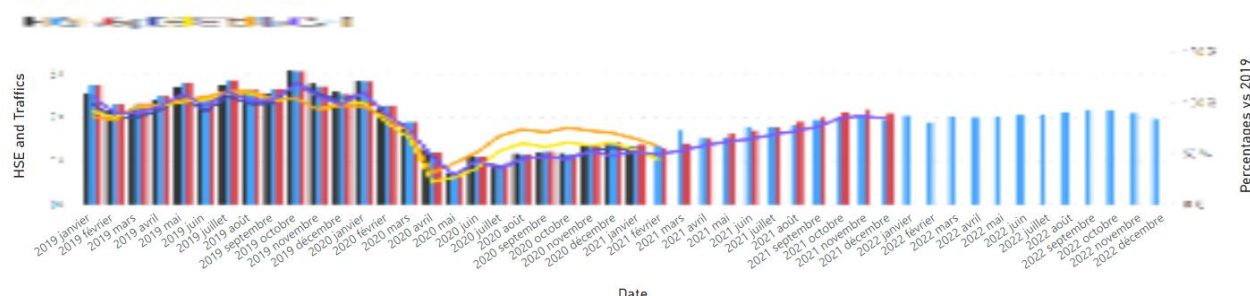


Figure 24. Graphique principal de l'onglet Dashboard HSE & Traffic

- **Table « Percentages »** : Une table qui présente les chiffres des indicateurs de pourcentage dans le temps, c'est-à-dire : « % SLS Traffic », « % S&OP –M », « % HSE » et « % Traffic » utilisés dans le graphique principal. Il est possible de glisser horizontalement pour voir tous les résultats qui datent depuis janvier 2019. La figure 25 présente la table « Percentages » avec les informations sensibles floutées.

Année	Mois	% SLS Traffic	% S&OP -M	% HSE	% Traffic
2019	janvier				
2019	février				
2019	mars				
2019	avril				
2019	mai				
2019	juin				
2019	juillet				

Figure 25. Table « Percentages » de l'onglet Dashboard HSE & Traffic

- **Table « HSE and Traffic »** : Une table dynamique qui dispose les valeurs des indicateurs qui ne sont pas des pourcentages. Les données datent depuis 2019 et il est possible de visualiser ces résultats par année, trimestre et mois, il suffit juste de cliquer sur le bouton « + » à gauche de la date dans la table. La figure 26 présente la table « HSE and Traffic » avec les informations sensibles floutées.

Année	S&OP +M	Budget	HSE
2019			
Trim 1			
Trim 2			
Trim 3			
Trim 4			
2020			
Trim 1			
Trim 2			

Figure 26. Table « HSE and Traffic » de l'onglet Dashboard HSE & Traffic

- **Filtres** : Il existe 5 types de filtres différents et ils peuvent interagir avec tous les composants et outils du projet. Par ailleurs, il est possible de choisir plus d'une option dans les filtres « Program », « SITE PP » et « Role SC ». Dans les autres filtres, c'est-à-dire, « S&OP –M » et « S&OP +M », il est possible de choisir juste une option et au début il y déjà des options auto sélectionnées. Par ailleurs, s'il est compliqué de trouver l'option de sélection dans les filtres, Il est possible d'effectuer une recherche par mot-clé dans l'espace appelé « Rechercher » dans ces filtres. Il est également possible d'effacer tous les sélections qui a été faites dans le filtre à travers du bouton « » qui apparaît lorsque le souris est sur l'un des filtres. De plus, il est possible de choisir plus qu'une option comme sélection pour tous les filtres et ses valeurs sont en ordre alphabétique.

Dans le filtre « S&OP –M », il y a la possibilité de choisir les dates de différentes prédictions HSE qui ont été faites précédemment. L'objectif de ce filtre est de pouvoir faire la comparaison des résultats de ces prédictions anciennes avec les données réelles ou avec les résultats des prédictions HSE plus récentes. Les options qu'il contient sont tous les dates de prédictions enregistrés dans la base de données du projet.

Dans le filtre « S&OP –M », il permet de choisir les dates des trois plus récentes prédictions HSE qui ont été faites et enregistrées dans la base de données du projet. L'objectif de ce filtre est de pouvoir faire la comparaison des résultats de ces prédictions récentes avec les données réelles ou avec les résultats des prévisions HSE plus anciennes.

Observation : Il existe des valeurs vides possibles dans les tables « Percentages » ou « HSE and Traffic » lors de l'utilisation de filtres, parce que lorsqu'il n'y a pas de données pour la date correspondante et pour le filtrage correspondant, les résultats seront affichés vides. La figure 27 affiche tous les filtres utilisés dans le projet.

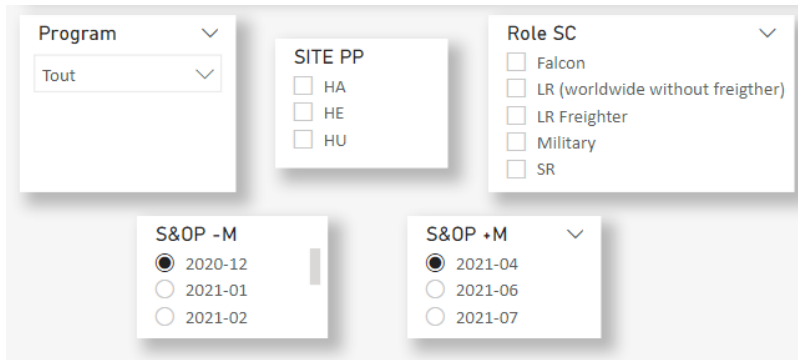


Figure 27. Filtres de l'onglet Dashboard HSE & Traffic

Lorsque le développement du projet était déjà bien structuré, il y a eu quelques réunions et présentations avec le « Supply Chain Director » pour entendre son avis et pouvoir améliorer le dashboard à partir de ses conseils.

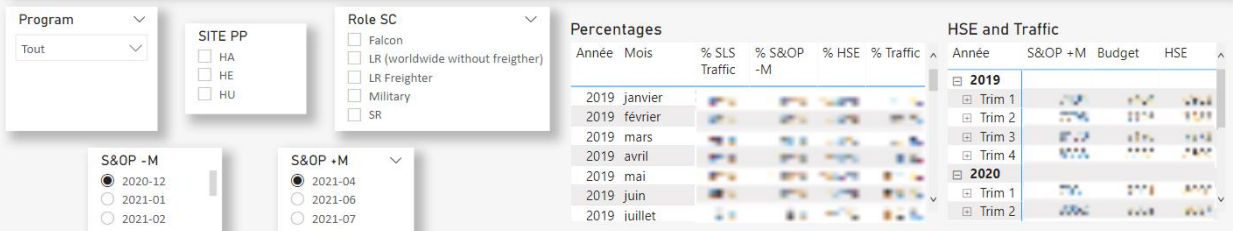
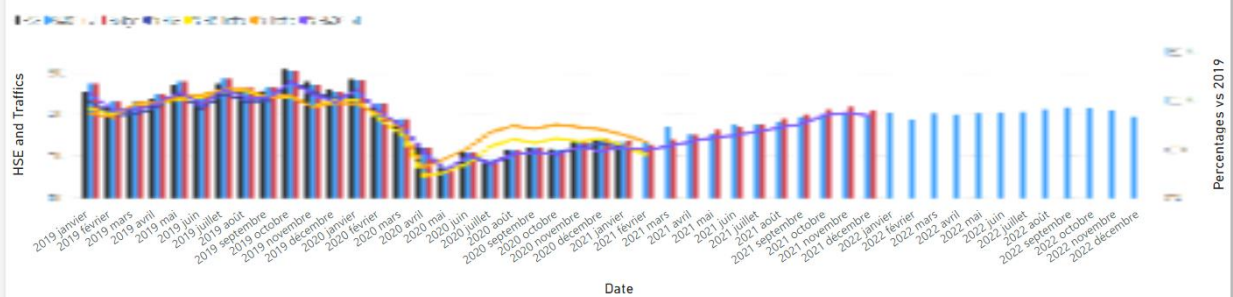
S'il y a eu des difficultés ou des obstacles au cours du projet, la première option était de chercher la réponse sans parler avec le tuteur d'entreprise, seulement avec ma connaissance technique et académique et ses recherches sur internet. En cas d'absence de solution, la prochaine étape est d'expliquer au tuteur d'entreprise la situation et le point bloquant. Si le problème persiste, la troisième voie consiste à rechercher un expert dans le domaine qui concerne la problématique.

Au final, une documentation détaillée du projet, un manuel d'utilisation et de publication sur le portail PBI ont été rédigés, tous ces documents à usage interne de Safran. Pour l'instant, pour que le projet soit abouti, il ne reste à Safran qu'un grand nettoyage de données historique côté dépose dans les données de la BDD SAP, dont l'activité n'est pas autorisée à exercer par un stagiaire.

3.7 Deployment

La visualisation du dashboard présente un graphique avec les informations explicitées dans le contexte. De plus sur le tableau de bord il y a 5 filtres pour pouvoir extraire des caractéristiques plus détaillées ou spécifiques. En bas à droite il y a deux petites tables, la première à gauche avec les valeurs de pourcentages par rapport à 2019 et, la deuxième avec les valeurs qui concernent le trafic aérien, échange de puit de chaleur et les prédictions de l'algorithme forecast interne Safran. La figure 28 montre le « Dashboard HSE and Traffic » avec les informations sensibles floutées.

S&OP -M: 2020-12 and S&OP +M: 2021-04



This document and the information it contains are the property of Safran. They must not be copied or communicated to a third party without the prior written consent of Safran.

Figure 28. Dashboard HSE and Traffic

Il y a un autre onglet (Cover) dans le projet qui a pour objectif d'organiser sous une forme résumée les informations sur les entrées, les sorties, les principaux objectifs, les personnes qui utiliseront du dashboard, les principales valeurs ajoutées à l'entreprise et les points encore à travailler pour améliorer le projet. La figure 29 montre cet onglet « Cover » qui est supplémentaire au Dashboard HSE and Traffic.

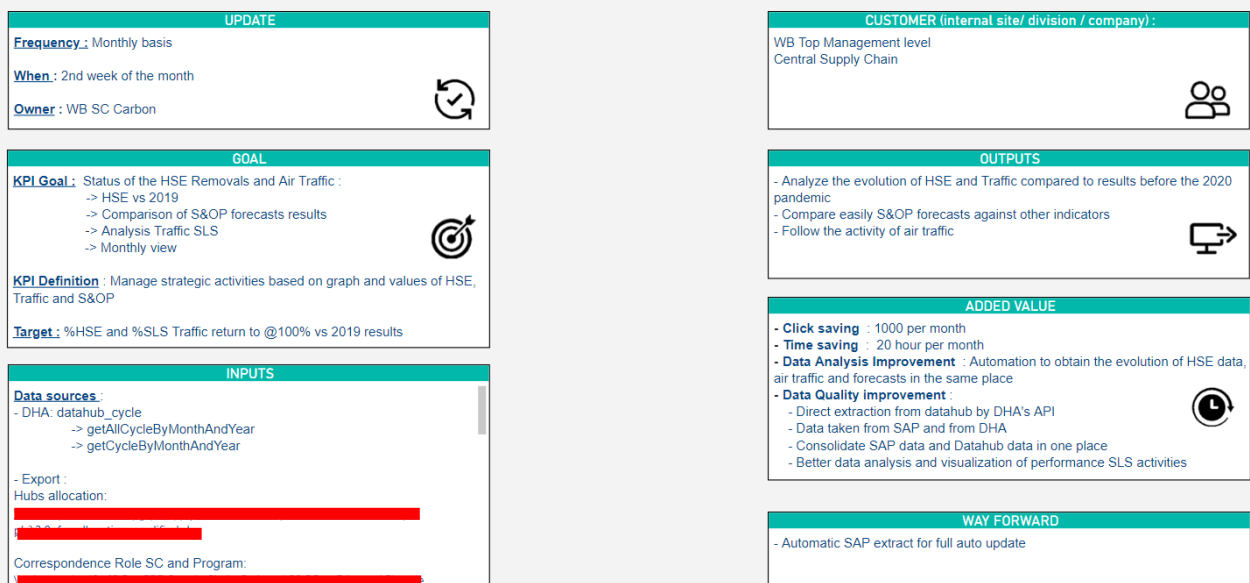


Figure 29. Onglet avec des informations générales qui résument le projet

La date et l'heure dans le coin supérieur droit de tous les onglets du projet, indiquent les informations de la dernière mise à jour des données, qui sont toujours programmée pour se mettre à jour automatiquement tous les mois.

Enfin, le déploiement a été fait par le portail PBI interne Safran, le Power BI Report Server. Dans la publication, le rapport publié doit avoir été créé dans une version de Power BI Desktop identique ou inférieure à la version de Power BI Report Server. La bonne version du Power BI Desktop est disponible dans le réseau interne Safran, parce que les 2 outils sont montés de version en même temps, c'est-à-dire qu'il n'y a pas de décalage entre la version de Power BI report Server et la version de Power BI Desktop. La version utilisée pour ce projet est « 2.86.1321.0 64-bit (Octobre 2020) »

Dans le portail PBI, un plan de mise à jour automatique des données a été créé pour être effectué toutes les deuxièmes semaines du mois.

4 Conclusion

Ce stage fut tout à fait enrichissant, parce qu'il m'a permis d'approfondir le domaine de la data science et de la business intelligence, ses acteurs, ses contraintes, mais aussi de participer concrètement à ses enjeux au travers de mes missions. Aujourd'hui, cette expérience vient confirmer le fait que j'ai fait le bon choix d'orientation et me permet d'affiner mon futur projet professionnel. En effet, je sais désormais qu'assurer la bonne gouvernance de données et appliquer les connaissances de la data science pour donner de valeur à l'entreprise me sont tout à fait adaptées.

Au cours du stage, il a été mis en place plusieurs projets qui permettent d'intégrer la data science aux processus de décision. Ainsi, la mission de la data science pour apporter de la valeur à l'entreprise a été remplie, parce que, principalement, à travers du « Dashboard HSE & Traffic », il a été possible de réduire le temps pris et les analyses faites par le processus de décision dans les analyses stratégiques faites par les managers Safran. Un autre projet qui a aussi collaboré à remplir la mission d'appliquer une bonne gouvernance des données était le « Dashboard LPO ». Ce projet a permis le département commercial d'accéder directement aux données et analyses qui étaient auparavant demandées. Par ailleurs, il y avait également des projets plus axés sur l'augmentation de la fiabilisation de données Safran, comme par exemple, « Anomalies de Consolidation LPO » et « Algorithme de nettoyage de données », deux projets très importants pour pouvoir avoir une gestion des données carbone saine et propre chez Safran.

4.1 Perspectives

Il y a certains aspects qui peuvent être améliorés dans le projet « Dashboard HSE & Traffic ». Actuellement, les données « Data SAP » et « Data Fleet » proviennent de documents Excel mis à jour manuellement à partir des applications SAP et Zemeter, respectivement. Suivant ce raisonnement, créer un lien direct entre le projet PBI et ces applications est un moyen d'optimiser encore plus le travail de mise à jour des données.

D'autre part, heureusement l'entreprise peut maintenant effectuer des bonnes analyses dynamiques avec des données fiables à utiliser, grâce aux projets développés pendant le stage, il sera plus simple de construire les prochains projets en se basant sur ceux déjà construits, principalement en se basant sur le Dashboard HSE & Traffic qui est un projet complexe et complet avec mise à jour automatique des données « Cycle » via l'utilisation de l'API DHA.

4.2 Retour d'expérience & bilan de compétences

Le projet « Dashboard HSE & Traffic » aborde des informations sensibles et très pertinentes pour l'entreprise. Avoir une connaissance claire et filtrée de la performance de ses activités par rapport aux données passées, aux crises ou aux prévisions, facilite grandement les processus de prise de décision d'amélioration. Pour cela, « Dashboard HSE & Traffic » constitue le projet le plus intéressant du stage, en plus d'être une excellente occasion de développer mes compétences dans le domaine de la data science. Le contenu technique abordé pendant la réalisation de ce projet a énormément contribué pour ma formation professionnelle. Ce stage m'a apporté une directrice pour ma formation et spécialisation, car les difficultés rencontrées m'ont fait comprendre ce dont j'ai besoin pour étudier davantage et ce dans quoi je suis vraiment bon. De plus, sur ce même sujet, on retrouve des différents domaines, en particulier, la business intelligence constitue le domaine sur lequel j'ai l'intérêt d'aller plus loin aussi.

Bibliographie

[1] COVID-19 impact on the European air traffic network, « Supporting European Aviation » 2021. [En ligne]. Disponible : <https://www.eurocontrol.int/covid19> [Accès le 24 août 2021].

[2] What is CRISP DM? , « Data Science Process Alliance » 2021. [En ligne]. Disponible : <https://www.datascience-pm.com/crisp-dm-2> [Accès le 01 Mai 2021].

[3] What is the CRISP-DM methodology? , « SmartVision Europe » 2021. [En ligne]. Disponible : <https://www.sv-europe.com/crisp-dm-methodology> [Accès le 20 Mai 2021].

[4] Power BI, « Microsoft » 2021. [En ligne]. Disponible : <https://powerbi.microsoft.com/en-us> [Accès le 30 Mai 2021]

[5] Vue d'ensemble de DAX, « Microsoft » 2021. [En ligne]. Disponible : <https://docs.microsoft.com/fr-fr/dax/dax-overview> [Accès le 05 Avril 2021]