

Conception SID

Cahier des Charges

-

Document fonctionnel

Table des matières

1.	Introduction.....	3
2.	Besoins utilisateurs.....	4
2.1	Données des vols des participants.....	4
2.2	Comportements des participants	4
2.3	Promotion sur les vols	4
2.4	Canaux de distribution.....	4
2.5	Monitoring et mesures clés.....	4
3.	Données d’entrées du système opérationnel.....	5
4.	Testes Technico-fonctionnel :.....	6
5.	Description des tables :.....	7
6.	Gestion projet	9
6.1	Méthodologie.....	9
6.2	Livrables attendus	9

1. Introduction

Le projet concerne un programme marketing pour les passagers réguliers d'une compagnie aérienne. L'objectif est d'analyser les vols de chaque participant pour mieux comprendre leur comportement et optimiser les offres, tout en assurant un suivi des miles, des surclassements et d'autres indicateurs de performance. Ce projet s'inscrit dans une stratégie de fidélisation et d'optimisation des revenus de la compagnie aérienne.

Pour ce faire, on souhaite mettre en place une infrastructure de « Business Intelligence ».

Cette infrastructure se composera des éléments suivants :

1. Système opérationnel

Le système opérationnel constitue la source des données initiales. Il regroupe toutes les applications fonctionnelles de la compagnie aérienne. Ces systèmes produisent des données brutes qui doivent être traitées.

2. STG/ODS (Staging Area ou Operational Data Store)

La couche STG/ODS a pour but de :

Collecter les données brutes en provenance des systèmes opérationnels.

Effectuer des validations de base pour garantir que les données extraites ne contiennent pas d'erreurs critiques (par exemple, vérification de formats, champs obligatoires, unicité des données, ...).

Stocker temporairement les données pour les transformer avant de les charger dans le DataWarehouse.

3. DataWarehouse (DWH)

Le **DataWarehouse** centralise et consolide les données issues de multiples systèmes opérationnels pour offrir une vision unique, fiable, homogène des données et c'est également dans le dwh qu'on fera les calculs des différentes mesures.

2. Besoins utilisateurs

Les utilisateurs souhaitent pouvoir analyser et suivre précisément les comportements des participants au programme marketing. Cette analyse doit couvrir une période de 3 ans d'historisation et inclure les fonctionnalités suivantes :

2.1 Données des vols des participants

- Identification des vols pris : Liste et suivi des vols empruntés par les participants, incluant les trajets et les dates.
- Analyse des appareils utilisés : Étude des types d'appareils (modèles d'avions) sur lesquels les participants voyagent, afin d'identifier les préférences et tendances.
- Étude des tarifs de base : Analyse des tarifs payés par les participants pour chaque vol, permettant de mieux comprendre les habitudes de dépenses.

2.2 Comportements des participants

Fréquences des surclassements : Suivi des habitudes de surclassement des participants.

Miles offerts et consommés : Monitoring des miles attribués (offres, promotions) et de leur utilisation par les participants.

2.3 Promotion sur les vols

Analyse de l'efficacité des promotions : Étude des périodes et types de promotions qui génèrent le plus d'engagement et d'achats parmi les visiteurs.

2.4 Canaux de distribution

Analyse des ventes par canal : Ventilation des ventes selon les différents canaux (vente directe, agences de voyage, plateformes en ligne, etc.) pour identifier les canaux les plus performants.

2.5 Monitoring et mesures clés

- Taux de conversion des vols : Calcul du taux de conversion pour chaque canal de distribution, défini comme le ratio : (Nombre de visiteurs via un canal / Nombre d'achats via ce canal).
- Durée des séjours : Suivi de la durée moyenne des séjours des participants après leurs vols, en lien avec les habitudes de voyage.
- Analyse des classes de voyage : Suivi de la répartition des participants par classes (économique, affaire) pour identifier les tendances.

3. Données d'entrées du système opérationnel

Le système opérationnel repose sur la gestion des opérations d'une compagnie aérienne, intégrant des fonctionnalités liées aux participants, vols, appareils, billets, promotions et canaux de distribution. Le système centralise les données afin d'assurer une gestion efficace et une analyse approfondie des différentes entités et relations. Les informations sont organisées autour des fichiers opérationnels suivants :

Participants et Miles

- **Fichier participant.csv** : Contient les informations personnelles des participants (nom, prénom, statut de fidélité, miles accumulés, coordonnées, etc.).
- **Fichier miles.csv** et **milesVol.csv** : Traitent les miles opérés et convertis par les participants en fonction de leur participation à des vols.

Vols et Appareils

- **Fichier vol.csv** : Gère les détails des vols, tels que les villes de départ et d'arrivée, la durée, le statut et les miles opérés.
- **Fichier appareil.csv** : Regroupe les données techniques des appareils utilisés, comme leur capacité, marque, modèle et année de fabrication.

Billets et Classes

- **Fichier billet.csv** : Rassemble les informations des billets émis, incluant le tarif de base, le prix acheté, la date d'achat et la classe de voyage associée.
- **Fichier classe.csv** : Contient les caractéristiques des classes passagers (nom, surclassement possible, avantages et tarifs supplémentaires).
- **Fichier surclassement.csv** : Permet de gérer les surclassements appliqués aux passagers, avec les éventuels frais associés.

Promotions

- **Fichier promotion.csv** et **promo.csv** : Décritent les promotions disponibles, incluant le taux de promotion, les dates de validité et les billets auxquels elles s'appliquent.

Canaux de Distribution et Visiteurs

- **Fichier canauxDistribution.csv** : Détaille les canaux de vente des billets, incluant le type de canal et le taux de conversion.
- **Fichier canalVisiteur.csv** : Relie les canaux aux visiteurs les ayant utilisés.
- **Fichier visiteur.csv** : Répertoire les informations des visiteurs (nom, prénom, coordonnées) pour le suivi des interactions.

Relations entre Entités

- Les relations entre les différentes entités sont clairement définies par le modèle conceptuel de données (MCD).

4. Testes Technico-fonctionnel :

Le processus de chargement des données dans l'ODS/STG a pour objectif principal de garantir la qualité et la conformité des données avant leur intégration dans les systèmes en aval. Les tests appliqués permettent d'assurer que seules les données valides et exploitables passent dans les tables de staging (STG).

Le traitement des données dans l'ODS/STG inclut les étapes et contrôles suivants :

1. **Validation des Unicités**
 - Les doublons sont détectés et filtrés à l'aide de tests d'unicité (tUniqRow).
 - Les enregistrements identifiés comme doublons sont rejetés et stockés dans un flux de rejet.
2. **Vérification de Conformité des Schémas**
 - Chaque fichier est validé pour s'assurer que sa structure respecte le schéma attendu (colonnes, types de données, valeurs obligatoires, etc.).
 - Les enregistrements non conformes (par exemple, colonnes manquantes ou types erronés) sont rejetés via le composant SchemaComplianceCheck.
3. **Contrôle des valeurs (LOV)**
 - Certains champs spécifiques (comme des catégories ou des statuts) sont soumis à des règles de validation basées sur des **Listes de Valeurs (LOV)** définies.
 - Les valeurs qui ne figurent pas dans la LOV sont rejetées à l'aide du composant tFilterRow et dirigées vers un flux de rejet.
4. **Tests de Cohérence Métier**
 - Des règles métier simples peuvent être appliquées au besoin (par exemple, vérifier que les dates sont dans une plage logique ou que certaines relations entre colonnes sont respectées).
5. **Nettoyage des Caractères Spéciaux**
 - Avant l'insertion dans les tables STG, une routine personnalisée appelée "Clean" est appliquée.
 - Cette routine nettoie les caractères spéciaux des données pour garantir leur qualité.
 - Cette étape est réalisée après tous les contrôles et avant l'insertion finale, afin de garantir que seules des données propres sont intégrées.
6. **Règles métier**
 - Avant insertion dans le Dwh , On vérifie plusieurs contrainte métier afin de garder une logique clair et précise dans l'entrepôt de données . Parmi ses testes , on trouve les testes concernant LOV (List Of View) c'est-à-dire que les données non présente dans le dictionnaire de données , ne seront pas pris en charge ni charger dans le DWH.

5. Description des tables :

Tables de Dimensions (DIM)

Les tables de dimensions contiennent des informations descriptives qui permettent de qualifier et d'enrichir les analyses.

DWH_DIM_APPAREILS

Contient des informations sur les appareils (avions) utilisés par la compagnie, comme leurs types, modèles, capacités, etc.

DWH_DIM_BILLET

Contient les informations liées aux billets, comme les types de billets (économique, affaires), le prix, les restrictions, etc.

DWH_DIM_CANALDISTRIBUTION

Fournit des informations sur les canaux par lesquels les billets ou services sont distribués (agences, en ligne, partenaires, etc.).

DWH_DIM_PARTICIPANTS

Décrit les participants au programme de fidélité ou les clients en général, incluant des détails comme le nom, les coordonnées, ou les niveaux de fidélité.

DWH_DIM_PROMOTION

Regroupe les détails des promotions ou campagnes marketing, telles que les réductions offertes ou les périodes de validité.

DWH_DIM_VOLS

Contient des informations sur les vols, comme les numéros de vol, les itinéraires, les durées, et les horaires.

Tables de Faits (FAIT)

Les tables de faits contiennent des données quantitatives, souvent utilisées dans les calculs et les analyses.

DWH_FAIT_CANALVISI

Contient des données quantitatives sur la performance des canaux de distribution, comme le volume des ventes par canal.

DWH_FAIT_MILES

Garde une trace des miles accumulés par les clients dans le cadre d'un programme de fidélité.

DWH_FAIT_PROMOTION

Mesure l'impact des promotions, incluant des métriques comme le nombre de billets vendus grâce à une promotion donnée.

DWH_FAIT_SURCLASSEMENT

Suit les surclassements des passagers, en indiquant combien de clients ont été surclassés, à quel coût ou pour quelle raison.

Tables Staging (STG)

Les tables STG (ou staging) servent à charger temporairement des données avant qu'elles ne soient nettoyées et insérées dans les tables finales (DWH).

STG_DIM_APPAREILS

Contient des données brutes sur les appareils avant transformation et intégration dans DWH_DIM_APPAREILS.

STG_DIM_BILLET

Données brutes sur les billets avant nettoyage pour DWH_DIM_BILLET.

STG_DIM_CANALDISTRIBUTION

Collecte des données initiales sur les canaux de distribution.

STG_DIM_CLASSE

Données sur les classes de billet (économique, affaires, première classe).

STG_DIM_PARTICIPANTS

Informations brutes sur les participants avant leur traitement pour la table dimensionnelle.

STG_DIM_PROMOTION

Rassemble des données initiales sur les promotions.

STG_DIM_VISITEUR

Contient des données brutes sur les visiteurs ou passagers avant transformation.

STG_DIM_VOLS

Enregistre les informations des vols dans leur état brut.

STG_FAIT_CANALVISI

Données factuelles brutes sur les performances des canaux de distribution.

STG_FAIT_MILES

Données initiales sur les miles accumulés.

STG_FAIT_PROMOTION

Mesure l'impact des promotions avant intégration dans la table des faits.

STG_FAIT_SURCLASSEMENT

Données brutes sur les surclassements des passagers.

6. Gestion projet

6.1 Méthodologie

Le soumissionnaire devra proposer une méthodologie de gestion de projet itérative et collaborative, garantissant un contrôle régulier et une validation des étapes clés tout au long du projet. Cette méthodologie s'appuiera sur le modèle du « cycle en V », tout en intégrant des approches itératives pour maximiser la flexibilité et l'adaptabilité.

Les étapes suivantes doivent être respectées au minimum :

- **Analyse des besoins** : Identification des besoins fonctionnels et non fonctionnels des utilisateurs, en étroite collaboration avec les parties prenantes.
- **Spécifications fonctionnelles et techniques** : Élaboration des spécifications détaillées décrivant les fonctionnalités et les exigences techniques du projet.
- **Développement** : Réalisation des composants techniques, incluant l'intégration et les tests unitaires.

6.2 Livrables attendus

Les livrables devront refléter une documentation claire, exhaustive et conforme aux standards.

Les éléments suivants sont attendus :

- **Documentation fonctionnelle** :
 - Un document fonctionnel succinct pour les spécifications initiales (mini cahier des charges, mini spécifications détaillées).
 - Un effort particulier devra être consacré aux spécifications fonctionnelles pour assurer une compréhension partagée des besoins.
- **Dictionnaire de données ou contrat d'interface** :
 - Une documentation précisant les structures de données, avec une préférence pour un contrat d'interface lorsque cela est pertinent.
- **Matrice décisionnelle** : Un tableau structurant les critères de décision pour valider les choix techniques ou fonctionnels.
- **Modèle conceptuel de données (MCD)** : Représentation des données et de leurs relations pour assurer une modélisation claire et cohérente.
- **Livrables techniques spécifiques** :
 - Livrables liés aux technologies et outils utilisés, notamment pour les projets Talend.