

UNIVERSITE DE MONTPELLIER

Rendu mini projet Entrepôt de données Air France

Belkassim BOUZIDI

Chakib ELHOUITI

Massili KEZZOUL



UNIVERSITÉ
DE MONTPELLIER



4 novembre 2020

1 Présentation de l'entreprise :

Air France est la compagnie aérienne nationale française, fondée le 7 octobre 1933. Ses activités principales sont le transport de passagers, de fret ainsi que la maintenance et l'entretien des avions. Elle dessert les principaux aéroports français ainsi que de nombreux aéroports étrangers.

2 Analyse complète d'Air France :

2.1 Objectifs de l'entreprise :

Le premier objectif d'Air France est de développer son réseau, de développer ses possibilités de destinations tout en réduisant les coûts. Au même temps assurer une prestation de service de haute qualité.

2.2 Position de l'entreprise sur le marché :

Air France est l'une des leaders du transport aérien européen. Elle est longtemps restée sur une image de chic à la française en proposant pour ses clients des services haut de gamme. Mais depuis l'arrivée des vols low-cost, Air France est en baisse de revenu ce qui l'a poussé à proposer pour tous ces vols des billets au prix réduit (seconde classe).

2.3 Quelques chiffres d'Air France :

- 104 Millions de passagers en 2019.
- 200 clients internationaux (maintenance).
- 391 destinations (activité cargo).
- 302 avions en 2019.
- 308 destinations (activité passager en 2019) dans 116 pays.
- L'activité passager représente 80% du chiffre d'affaire.
- 6 avions cargo.
- 15,8 Milliard d'euro de chiffre d'affaire en 2017.

2.4 Services proposés par l'entreprise :

- Air France propose plusieurs types de services, notamment :
- Transport de voyageurs (avec 80% du CA).
 - Transport de marchandise.
 - Maintenance et entretien des avions.

2.5 Formes de revenu et dépenses :

2.5.1 Revenu :

- Ventes de billets ;
- Ventes de services à bord de l'avion ;
- Transport de personnes ;
- Transport de marchandise ;
- Maintenance aéronautique

2.5.2 Dépenses :

- Carburant,
- Le coût de l'équipage,

- La maintenance de l'avion,
- L'achat de nouveaux avions...

2.6 Information aidant la prise de décision :

- Déterminer quel sont les liaisons les plus demander, en quel période.
- Pour un vol donner à quel prix peut se vendre un billet pour remplir au maximum l'avion.(les prix maximum pour remplir l'avion)
- Connaître la tranche d'âge de clients pour de promotions.
- Quels sont les destination en vogue en ce moment. (si une nouvelle destination apparait)
- Déterminer si un vol rapporte plus ou moins qu'un autre vol.
- Quelle est la position de chaque avion (Gérer au mieux la flotte d'avions disponible).
- la rentabilité de chaque type avion (Exemple : consommation trop forte de carburant d'un avion => le remplacer par un autre)

3 Q2 Les actions et opérations à tracer pour récupérer ces informations. :

- Ventes des billets/vols
 - la date de la vente.
 - le lieux de depart et d'arrivé(destination en vogue).
 - le prix de vente du billet.
 - Nombre de place occupé dans l'avion.
 - rentabilité du vol (revenu/dépenses)
 - La tranche de chaque client.
 - La fidelité de chaque client.
 - Nombre de billets enregistrés pour chaque client
- Gestion des stocks d'avions
 - position de chaque avion à la fin de la journée (aéroports)
 - consommation d'un avion (carburant et divers coûts)
 - nombre de vol effectués par un avion.

4 Q3 Traitement possibles :

- Ventes des billes/vols
 - Analyser la demande vers une destination (le nombre de billets vendu) selon la période
 - Analyser le ratio nombre de place occupé / capacité dans l'avion pour calculé le prix du billet
 - Comparer le ratio revenu / dépenses pour chaque vol (rentabilité)
 - analyser la tranche d'age des client.
 - analyser le nombre de billets enregistré pour chaque client.
 - analyser si un client a une carte de fidelité ou non.
- Gestion des stocks d'avions
 - Analyser la position de chaque avion à la fin d'une journée.
 - Analyser le coût moyen d'un avion en carburant.
 - Comparer les rentabilité des avions(CÂD : recettes - dépenses).

5 Q4 Ordonnez les actions par ordre d'importance / rentabilité potentielle :

1. Ventes des billets
2. Organisation des avions et vols

6 Q5 Identifiez les deux actions / opérations les plus importantes à analyser :

1. Ventes des billets
2. Organisation des avions et vols

7 Q6 et Q7 Data-Mart

7.1 Data-Mart ventes des billets

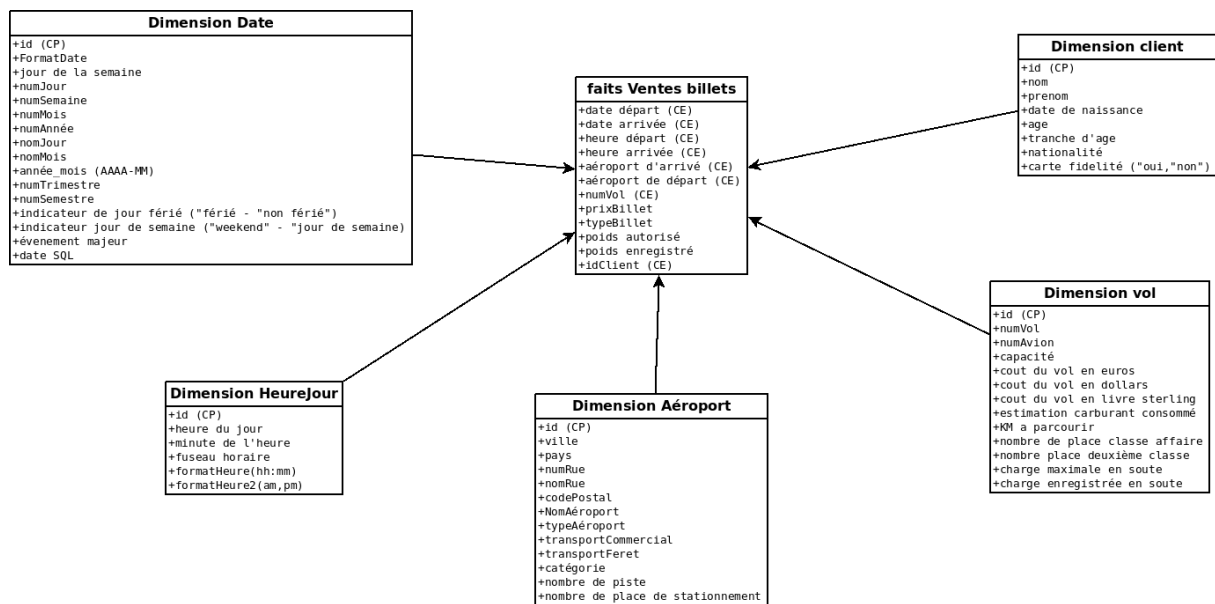


Schéma 1 – Data-Mart Ventes de Billets

Pour chaque billet vendu, on enregistre les mesures suivantes :

- Le prix du Billet - Additive
- Le type du Billet (1-ère class, seconde classe ..) - Non-Additive
- Le poids enregistré - Additive

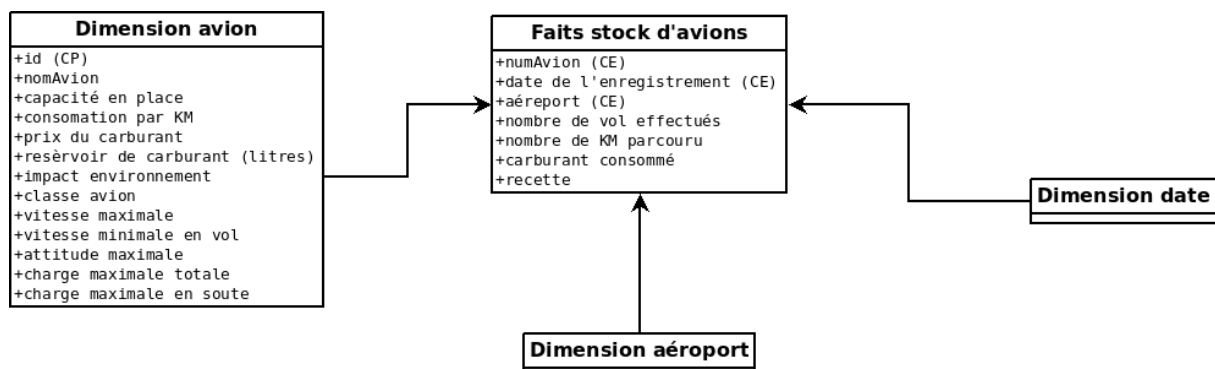


Schéma 2 – Data-Mart Gestion de la flotte

Une ligne de cette table des faits représente un avion à un certain moment. Les mesures sont :

- Le nombre de vol effectués par l'avion - Additive
- Le nombre de KM parcouru par l'avion - Additive

- Carburant consommé - Additive
- recette - Additive

8 Q8 Réponses au traitements

On peut répondre aux traitements qu'on vient d'indiquer avec le modèle qu'on a mis en place. Pour chaque traitement, on peut récupérer facilement l'information demandée, on va détaillé ci-dessous deux traitements (un traitement par action) :

- Si on veut analyser la demande vers une destination (le nombre de billets vendu) pour une période donnée, on doit récupérer tout les vols d'une période en faisant une jointure entre la table de faits Ventes de Billets avec les tables de dimension date, vol et aéroport. Ensuite en regroupant par destination, avec un count(*) et un order by, on aura le nombre de vol vers chaque destination dans cette période par ordre croissant. Par cette requête, on peut savoir la destination la plus demandée pendant une période donnée.
- L'analyse du coût moyen d'un avion en carburant se fait par une requête simple qui fait la jointure entre la table des faits (stocks d'avions) et la dimension avions en affichant pour chaque avions la moyenne du calcul suivant (carburant consommé * prix du carburant).

9 Q9 Exemple d'instance de l'Entrepôt de données

9.1 faits Ventes billets

date départ	date arrivée	heure départ	heure d'arrivée	aéroport de départ	aéroport d'arrivée	numVol	poids autorisé	poids enregistré	idClient
1	1	1	2	16	40	1	30	27	1
2	2	2	4	16	75	5	30	24	1
3	3	2	3	40	75	35	15	10	2

9.2 Dimension date

id	FormatDate	jour de la sem	numJour	numSemaine	numMois	numTri	indicateurJdS	Event Majeur	Date SQL
1	02/01/2020	3	2	1	1	1	1 SEMAINE		02:34:40
2	05/04/2020	2	5	1	4	4	2 SEMAINE		05:37:40
3	06/07/2020	3	6	1	7	3	3 WEEK		06:40:40

9.3 Dimension HeureJour

Id	Heure du Jour	minute de l'heu	fuseau horaire	formatHeure	FormatHeure2
1	5	35	UTC+1	05:35	05:35 am
2	9	36	UTC+0	09:36	09:36 am
3	10	37	UTC+0	10:37	10:37 am
4	17	0	UTC-5	17:00	05:00 pm

9.4 Dimension Aéroport

id(cp)	ville	pays	numRue	codePostal	NomAéroport	nombre de piste	nombre place stationnement
40	montpellier	france	15	34000	MTP	22	40
16	algers	algerie	16	16000	ALG	60	70
75	PARIS	FRANCE	19	75000	CDG	100	120

9.5 Dimension Vol

ID	numVol	numAvion	capacité	coût du vol en €	coût du vol en \$	coût du vol en £	Nombre de place 1-ère classe	nombre de place 2-nd classe	charge maximal en soute	charge enregistrée en soute
1	QQ256	1	150	500	509	486	26	124	60	90
5	AS308	2	220	780	788	767	25	180	80	100
35	AF568	3	500	390	394	370	70	300	90	75

9.6 Faits flotte d'avions

numAvion	date	aéroport	nombre de vol effectue	Nombre de KM parcouru	Carburant consommé	Recette
1	1	75	2	10000	290	6000
2	2	16	3	12000	340	9000
3	3	40	4	8000	220	4500

9.7 Dimension avion

id	nomAvion	capacité en pla	consommation par KM	réservoir de carburant	Impact environnement	classe Avion	altitude maximal	charge maximal en soute (en Tonne)	charge maximal total
1	Airbus A220	150	22	4000	Moyen	moyen-courrier	9000	60	600
2	AirBus A330	220	34	9000	Lourd	long-courrier	11000	500	950
3	AirBus A120	219	45	3500	Leger	cour-courrier	3000	80	800
4	boeing 747	500	65	3555	lourd	long-courrier	2500	250	125

10 Q10 Calcul de la taille du DataWarehouse

Table	Estimation nombre de lignes	Estimation pour une seule ligne	Total
Date	365	120 octets	43,8 Ko
HeureJour	14400	27 octets	38.8 Ko
Avions	554	113 octets	62 Ko
Vols	839500	73 octets	61.2 Mo
Aéroport	308	149 octets	45,9 Ko
Clients	20000000	98 octets	2 Go
Faits Ventes Billets	104000000	74 octets	7,7 Go
Faits stock avions	202210	35 octets	7 Mo

Schéma 3 – Estimation de la taille

11 Q11 Implémentation

Veuillez voir le code en annexe.

12 Q12 Requêtes analytique

requetes.sql

```

1  /* R1
2     la demande vers une destination donnée dans une période donnée
3  */
4  SELECT count(*) AS nb_billet_destination_période
5  FROM ventes_billets, dimension_aéroport, dimension_date
6  WHERE dimension_date.id = ventes_billets.date_depart
7  AND dimension_aéroport.id = ventes_billets.aéroport_arrivee
8  AND num_annee = 2019
9  AND ville = 'Montpellier';
10

```

```

11  /* R2
12      nombre de billets enregistrés par vol
13  */
14
15  SELECT num_vol,count(*) as nb_billets_vol
16  FROM ventes_billets, dimension_vol
17  WHERE dimension_vol.id = ventes_billets.num_vol
18  GROUP BY ventes_billets.num_vol;
19
20  /* R2
21      nombre de billtes enregistrés par vol comparé a la capacite du vol
22  */
23
24  SELECT num_vol,count(*) as nb_billets_vol, capacite
25  FROM ventes_billets, dimension_vol
26  WHERE dimension_vol.id = ventes_billets.num_vol
27  GROUP BY ventes_billets.num_vol;
28
29  /* R4
30      nombre de billtes enregistrés par client (pour faire une fidelité a ce
31          client)
32  */
33  SELECT num_client,count(*) as nb_billets_vol
34  FROM ventes_billets, dimension_client
35  WHERE dimension_client.id = ventes_billets.num_client
36  GROUP BY ventes_billets.num_client;
37
38  /* R5
39      avoir la tranche d'age de chaque client pour une destination donnée
40  */
41  SELECT tranche_age
42  FROM ventes_billets, dimension_client,dimension_aeroport
43  WHERE ventes_billets.num_client = dimension_client.id
44  AND dimension_aeroport.id = ventes_billets.aeroport_arrivee
45  AND ville = 'Montpellier';
46
47  /* R6
48      le nombre de clients qui n'ont pas une carte fidelité
49  */
50  SELECT count(*) as nb_client
51  FROM ventes_billets,dimension_client
52  WHERE ventes_billets.num_client = dimension_client.id
53  AND fidelite = 'non';
54
55
56  /* R7
57      le nombre de clients qui ont une carte fidelité
58  */
59
60  SELECT count(*) as nb_client
61  FROM ventes_billets,dimension_client
62  WHERE ventes_billets.num_client = dimension_client.id
63  AND fidelite = 'oui';
64
65  /* R8
66      Calculer la rentabilité de chaque avion
67  */
68  SELECT a.nom_avion , s.recette - s.carburant_consomme * a.prix_carburant
69  FROM dimension_avions a, stock_avions s
70  WHERE s.num_avion = a.id

```

```

71 GROUP BY s.num_avion;
72
73 /* R9
74     Calculer la moyenne de carburant consommé par avion par KM parcouru et la
       comparé avec la conso théorique
75 */
76 SELECT a.nom_avion , AVG(s.carburant_consomme / s.nombre_km), a.conso_par_km
77 FROM dimension_avions a, stock_avions s
78 WHERE s.num_avion = a.id
79 GROUP BY s.num_avion, s.date_jour;
80
81 /* R10
82     Calculer la moyenne du coût par vol pour chaque avion
83 */
84 SELECT a.nom_avion , s.carburant_consomme * a.prix_carburant / nombre_vol
85 FROM dimension_avions a, stock_avions stock_avions
86 WHERE s.num_avion = a.id
87 GROUP BY s.num_avion;

```

13 Q13 Les vues matérialisées

Dans notre cas, on a que la requête 3 qui peut utiliser une vue matérialisée (à partir de la requête 2) :

```

1 CREATE MATERIALIZED VIEW nb_billets_vol (num_vol,nb_billets_par_vol)
2 AS
3 SELECT num_vol,count(*)
4 FROM ventes_billets, dimension_vol
5 WHERE dimension_vol.id = ventes_billets.num_vol
6 GROUP BY ventes_billets.num_vol;

```

Après la création de cette vue matérialisée la 3ème requête devient :

```

1 SELECT dimension.id,nb_billets_par_vol,capacite
2 FROM nb_billets_vol, dimension_vol
3 WHERE dimension_vol.id = nb_billets_vol.num_vol;

```