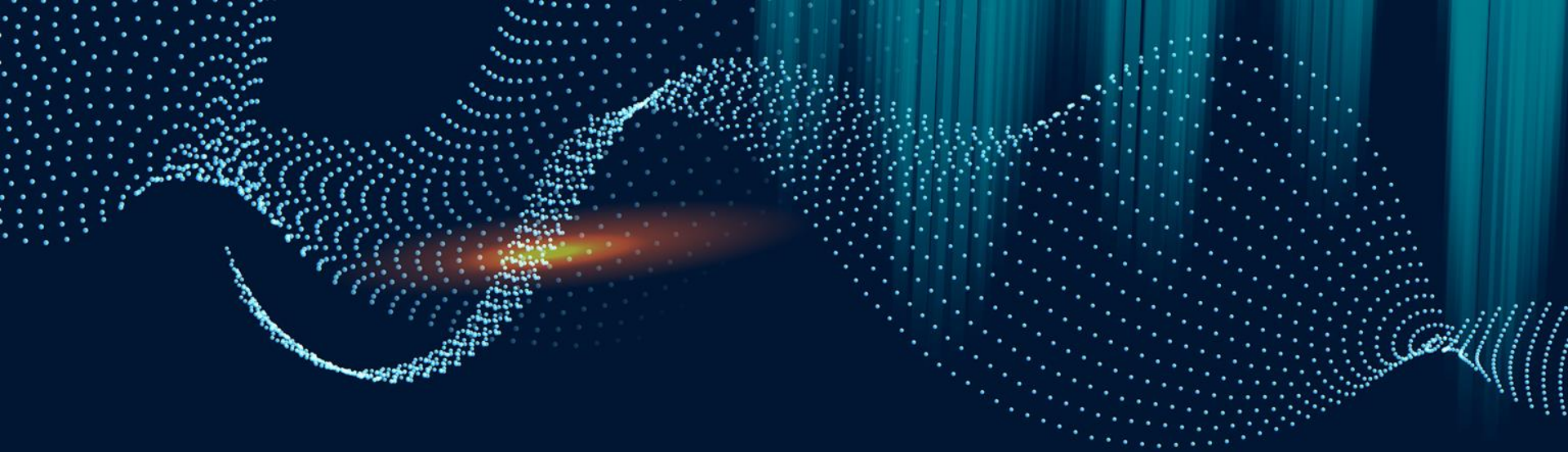
ENTREPÔT DE DONNÉES

AÉROPORT

LEKTATI MAHDI	**MONZEIN LEO**
MOKHTARI AHMED	**BADECHE M-AMINE**

2021/2022

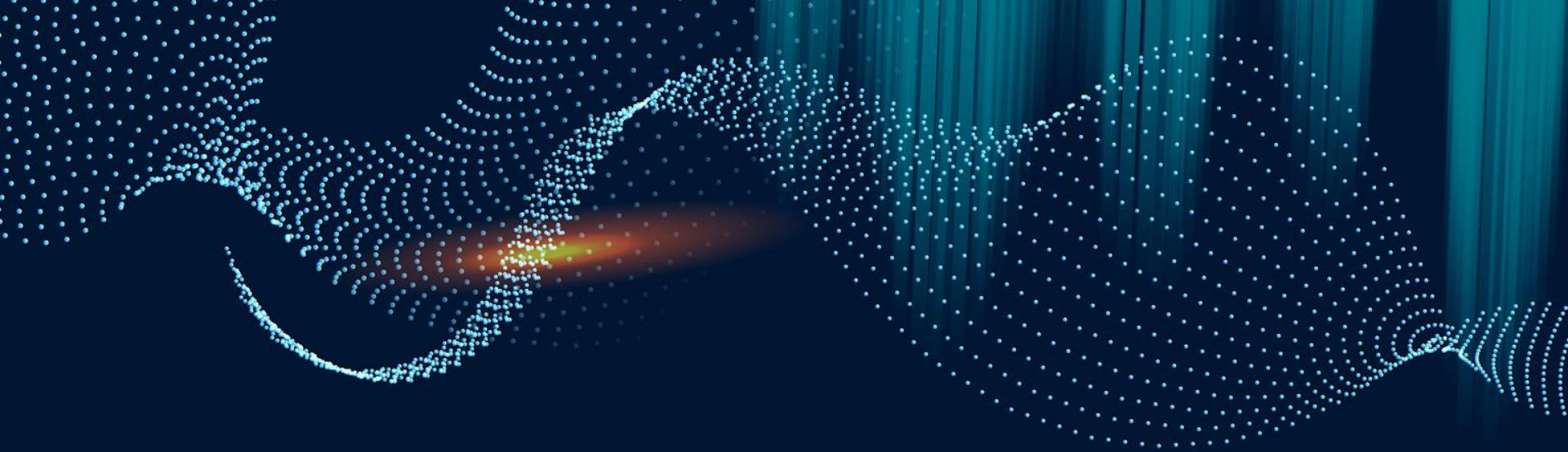
SOMMAIRE



01 PROBLEMATIQUE

PROBLEMATIQUE

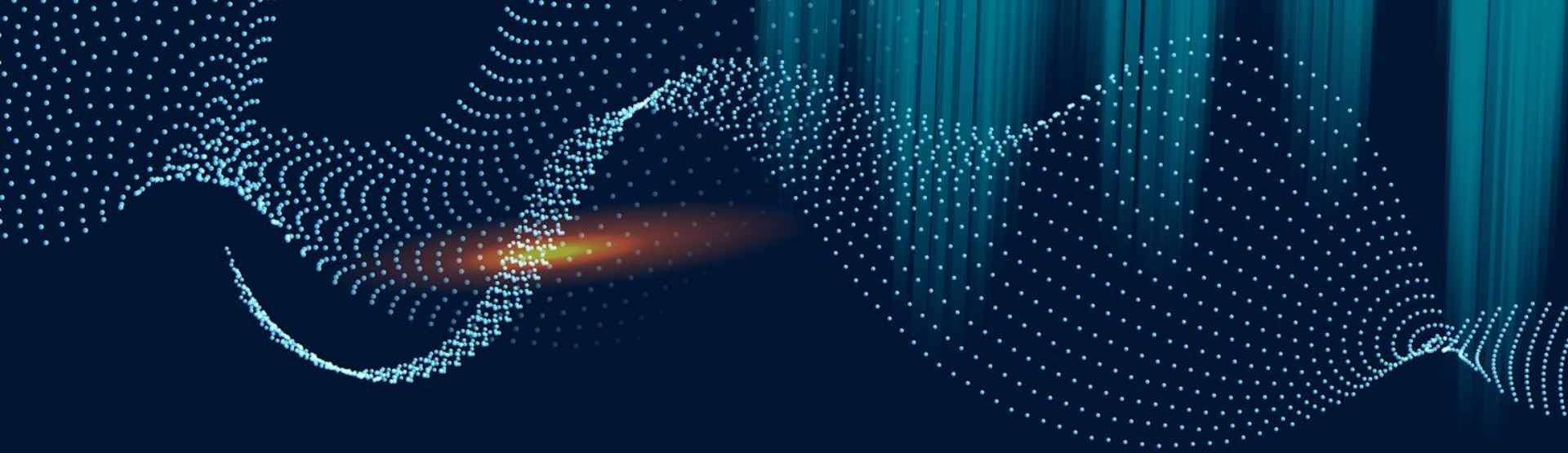
- Manque de qualité sur les indicateurs de performance.
- Dépenses en matière de coûts importantes.
- Limite sur la capacité de traitement de grandes données



02 OBJECTIFS

OBJECTIFS

- Construire un entrepôt de données pour le suivi des avions et les vols opérés.
- Implémenter l'entrepôt de données à partir d'un SGBD.
- Extraire les données et les interpréter selon les besoins des utilisateurs.



03 SOLUTION

SOLUTION : ARCHITECTURE

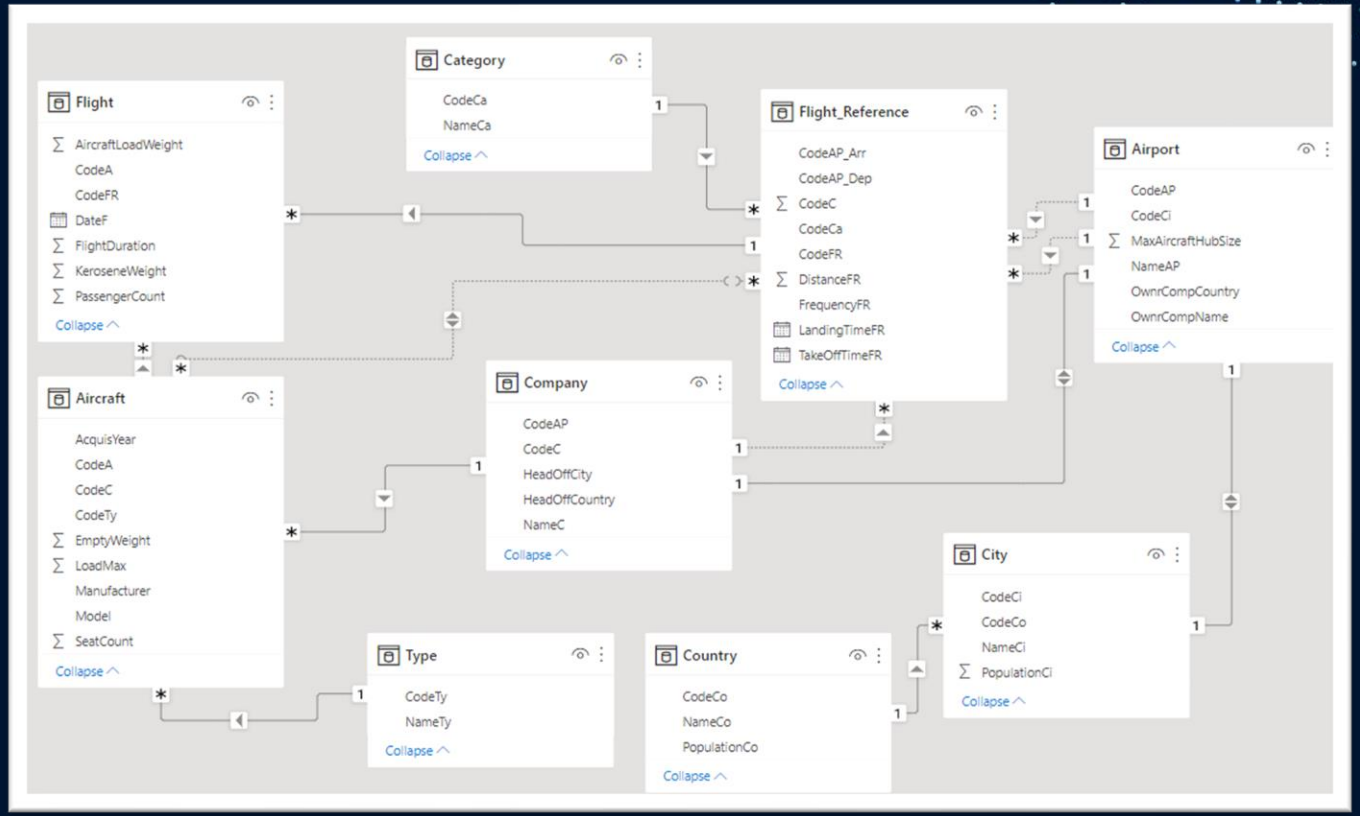
Base de données SQL

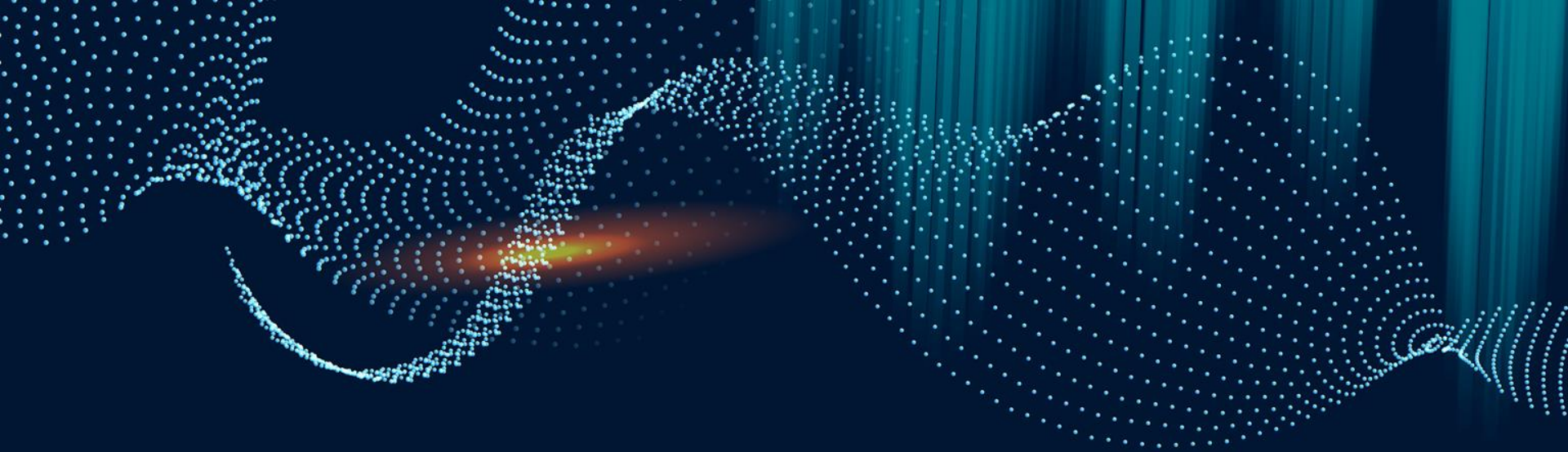
Transformation
Nettoyage

Power BI

Entrepôt de données

SOLUTION
MODELE DE DONNEES



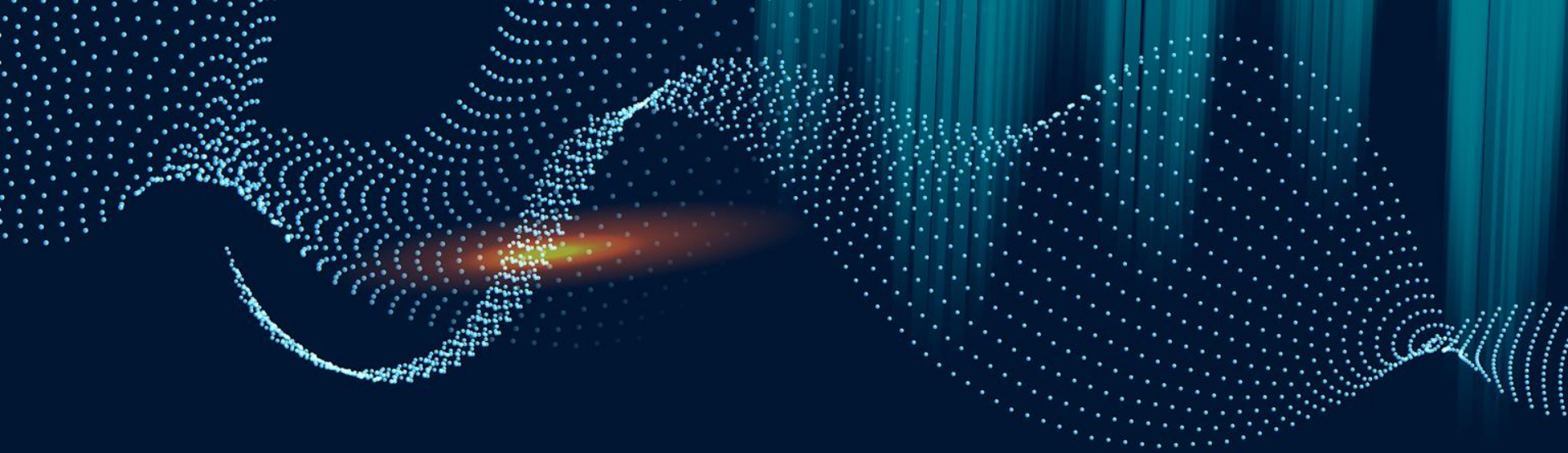


04 DONNEES, RESSOURCES

DONNEES, RESSOURCES

- Données prévenantes d'une base données relationnelles
- Données récupérées depuis des DataHub disponible en open source sur le web

NameC	CodeA	Model	AcquisYea	Manufactu	HeadOffCi	HeadOffCo	NameTy	FlightDura	KeroseneV	AircraftLo	EmptyWei	LoadMax	Passenger	SeatCount	Frequency	DistanceFl	Year	Month	Day
Angola Airl	AK3800	ABC	########	AIRBUS	Luanda	Angola	airliner	9	2000	6000	2500	4000	330	300	weekly	700	1899	December	30
AeroFranc	AZ3455	ZERT	########	KLMRT	Toulouse	France	airliner	3	550	1050	2000	3000	10	250	weekly	700	1899	December	30
Angola Airl	AK3800	ABC	########	AIRBUS	Luanda	Angola	airliner	11	2450	9000	2500	4000	610	300	daily	7742	1899	December	30
AeroFranc	AZ3455	ZERT	########	KLMRT	Toulouse	France	airliner	7	934	5000	2000	3000	275	250	daily	7742	1899	December	30
AeroFranc	DE3455	ZERT	########	DNTJF	Toulouse	France	regional	6	567	1500	1000	2000	45	50	daily	7742	1899	December	30
AeroFranc	DE3455	ZERT	########	DNTJF	Toulouse	France	regional	4	500	2000	1000	2000	50	50	monthly	9678	1899	December	30
Angola Airl	AK3800	ABC	########	AIRBUS	Luanda	Angola	airliner	2	500	1000	2500	4000	10	300	monthly	2345	1899	December	30



05 MISE EN OEUVRE

OUTILS

DOCKER

Manipulation et initialisation automatique de la base de données

SQL

Requêtage base de données

POWER BI

Gestion de l’entrepot de données et du reporting

LES ANALYSES POUR LES GESTIONNAIRES

(a) For each company (code, name and country of the head office) provide the total number of flights operated. Order alphabetically the results according to country names;

(b) List the companies (code and name) that only operate flights with the daily frequency during 2021. These companies may own aircrafts, operated by other companies and flown with frequencies other than daily. Each company should be displayed only once;

(c) Give the total number of flights during January 2021 for the aircrafts (code and model) operated by the company "Aero France" that flew at least 5 times;

4. Write also a new SQL query that would complete the current analyses of the users of this data warehouse. Justify your answer by stating the advantage(s) of this new analysis compared to the user's requirements mentioned above.

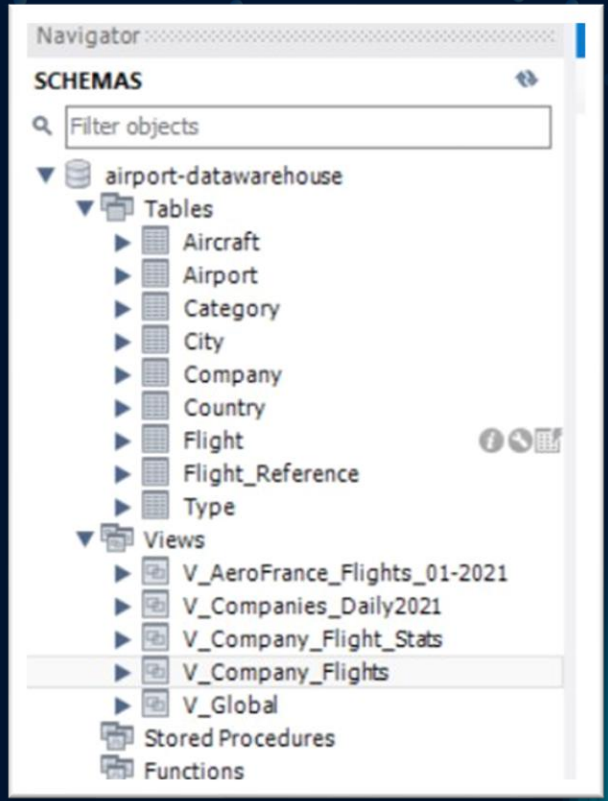
year	code	name	perc_daily_flights	perc_weekly_flights	perc_monthly_flights	avg_duration	avg_distance	total_number_flights_operated
2021	111111	AeroFrance	37.5000	50.0000	12.5000	2.6250	895.8750	8
2021	127658	AirFrance	100.0000	0.0000	0.0000	6.0000	500.0000	1
2021	198574	Angola Airline	100.0000	0.0000	0.0000	5.0000	2000.0000	1
2021	456789	Delta	0.0000	0.0000	100.0000	4.0000	9678.0000	1
2012	110234	AirCote	100.0000	0.0000	0.0000	6.0000	4568.0000	1

ETL : PASSAGE PAR LES VUES SQL

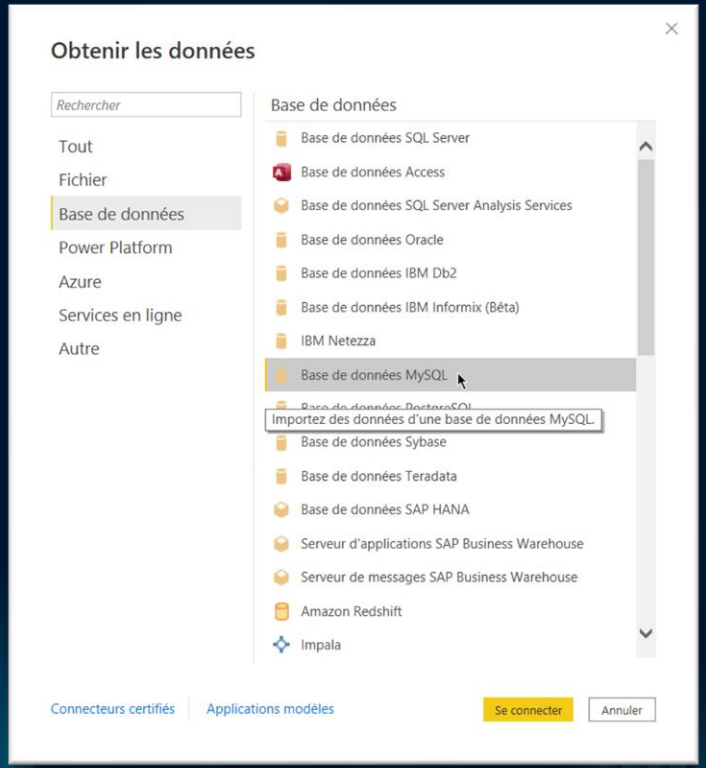
docker-compose.yaml | init.sql 9+

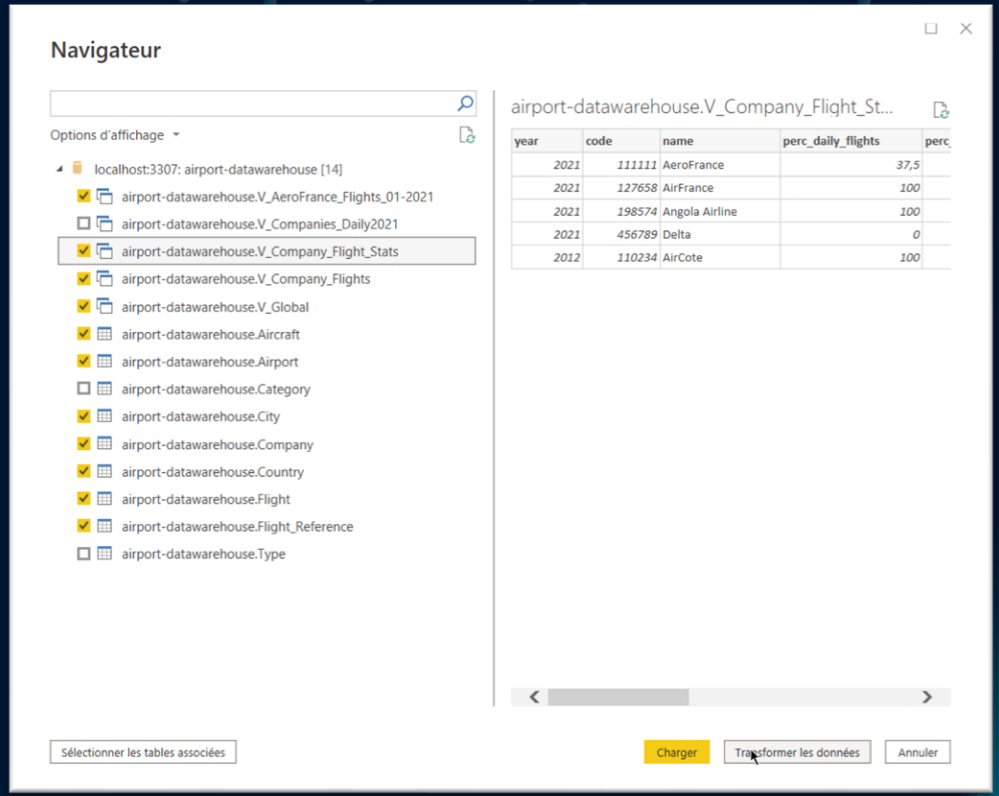
mysql > init.sql

```sql
CREATE  OR REPLACE VIEW `V_Company_Flight_Stats` AS
SELECT  YEAR(F.DateF) as year,
        C.CodeC as code,
        C.NameC as name,
        (SELECT  COUNT(*)
           FROM  Flight              F1,
                 Flight_Reference    FR1
          WHERE  FR1.CodeC = FR.CodeC
            AND  FR1.CodeFR = F1.CodeFR
            AND  YEAR(F1.DateF) = YEAR(F.DateF)
            AND  FR1.FrequencyFR IN ('daily', 'day only', 'daily working')) / COUNT(*) * 100 as perc_daily_flights,
        (SELECT  COUNT(*)
           FROM  Flight              F2,
                 Flight_Reference    FR2
          WHERE  FR2.CodeC = FR.CodeC
            AND  FR2.CodeFR = F2.CodeFR
            AND  YEAR(F2.DateF) = YEAR(F.DateF)
            AND  FR2.FrequencyFR = 'weekly') / COUNT(*) * 100 as perc_weekly_flights,
        (SELECT  COUNT(*)
           FROM  Flight              F3,
                 Flight_Reference    FR3
          WHERE  FR3.CodeC = FR.CodeC
            AND  FR3.CodeFR = F3.CodeFR
            AND  YEAR(F3.DateF) = YEAR(F.DateF)
            AND  FR3.FrequencyFR = 'monthly') / COUNT(*) * 100 as perc_monthly_flights,
        AVG(F.FlightDuration) as avg_duration,
        AVG(DistanceFR) as avg_distance,
        COUNT(*) as total_number_flights_operated
  FROM  Company C,
        Flight_Reference    FR,
        Flight              F
 WHERE  FR.CodeC = C.CodeC
   AND  F.codeFR = FR.codeFR
GROUP BY YEAR(F.DateF), C.CodeC, C.NameC
ORDER BY YEAR(F.DateF) DESC, C.NameC, COUNT(*);
```



IMPORT DES DONNEES





DEMONSTRATION

CONCLUSION

AVANTAGES

- Peut être exercé par un utilisateur réel
- Montée en compétence sur Power BI
- Travailler sur un autre domaine: Informatique decisionnelle

CONTRAINTES

- Le temps
- Manque de données
- Projet parallèle

MERCI !

QUESTION ?