

Toutes les données sont disponibles dans les 3 modèles dim_ que l'on a créés avec dbt. On peut donc les utiliser pour faire de l'EDA dans la base de données Snowflake.

1. Which aircraft has flown the most?

- ⇒ On s'appuie sur la table dim_aircraft qui contient le nombre de vols par modèle d'avion.
- ⇒ On peut donc faire une requête SQL pour récupérer les 5 modèles d'avion qui ont le plus volé.
- ⇒ Visualisation simple avec un graphique en barres.

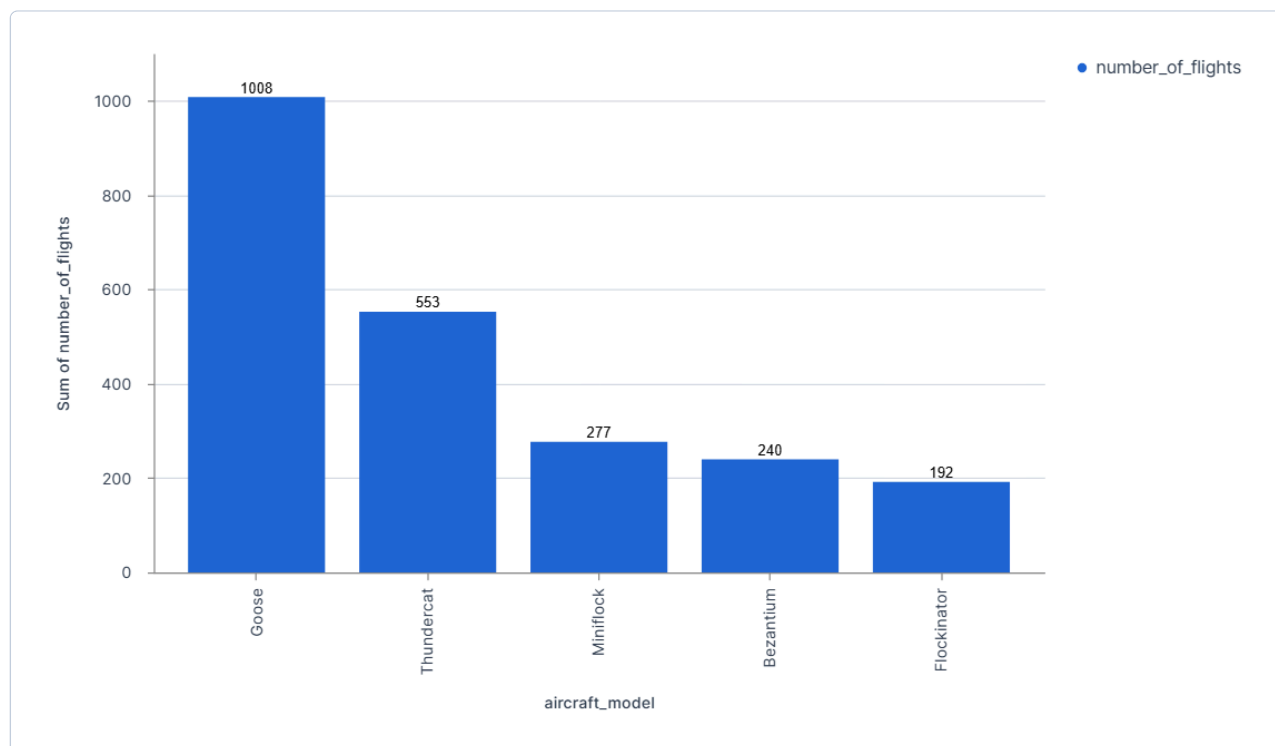
Réponse : c'est le modèle Goose qui a le plus volé, suivi du modèle Duck, puis du modèle Swan, puis du modèle Eagle, puis du modèle Heron.

⇒ On peut donc dire que le modèle Goose est le modèle d'avion qui a le plus volé, avec 1008 vols.

```
SELECT
  aircraft_model,
  number_of_flights
FROM ANALYTICS.DBT_NLEVASSEUR.DIM_AIRCRAFT
ORDER BY number_of_flights DESC
LIMIT 5
```

	aircraft_model o...	number_of_flights i	
0	Goose	1008	
1	Thundercat	553	
2	Miniflock	277	
3	Bezantium	240	
4	Flockinator	192	

5 rows, 2 cols 10 / page << < Page 1 of 1 > >> ↓



Question 2 :

Which airport has transported the most passengers?

- ⇒ On s'appuie sur la table dim_airport qui contient le nombre de passagers par aéroport.
- ⇒ On peut donc faire une requête SQL pour récupérer les 5 aéroports qui ont transporté le plus de passagers.
- ⇒ Visualisation simple avec un graphique en barres.

Réponse : c'est l'aéroport Amazon Mothership qui a transporté le plus de passagers, avec 2 423 000 passagers

```

SELECT
  airport_name,
  total_passengers_transport
FROM ANALYTICS.DBT_NLEVASSEUR.DIM_AIRPORTS
ORDER BY total_passengers_transport DESC
LIMIT 10

```

	airport_name obj...	total_passengers...	
0	Amazon Mothers...	2423400	
1	Nestland Airport	1999700	
2	Flocktopia	1685200	

3 rows, 2 cols

10

/ page

<<

<

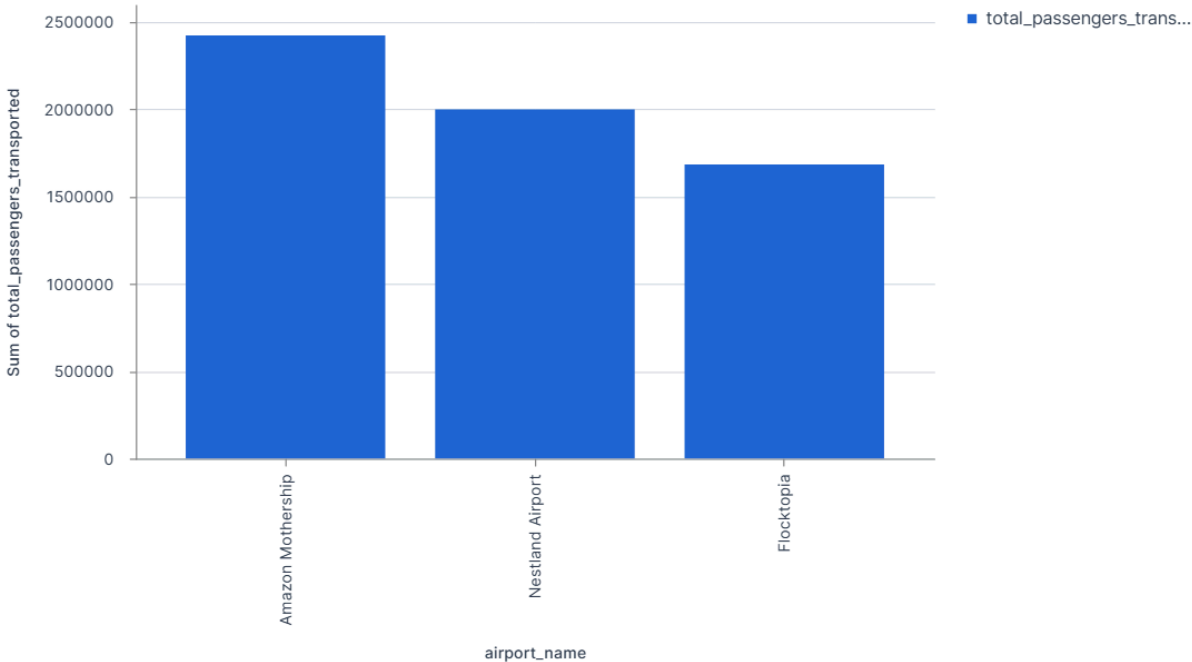
Page

1

>

>>

↓



Question 3 / Question 4 :

Quelle est la meilleure année RPM (Revenue Passenger-Miles) et meilleure année de croissance par compagnie aériennes ?

⇒ On s'appuie sur la table dim_airline qui contient le best_year_rpm et le best_year_growth par compagnie aérienne.

⇒ Visualisation simple en tableau étant donnée le peu de données à afficher.

Réponse : exemple pour Amazon Airlines ⇒ Meilleur année RPM 2013, meilleure année de croissance 2013...une corrélation ?

```

SELECT
  airline_name,
  best_year_rpm,
  best_year_growth
FROM ANALYTICS.DBT_NLEVASSEUR.DIM_AIRLINES
WHERE best_year_rpm IS NOT NULL -- Par sécurité, on filtre les compagnies inactives/sans données
ORDER BY airline_name ASC

```

	airline_name obj...	best_year_rpm o...	best_year_growth c	
0	Amazon Airlines	2013	2013	
1	Flock Air	2016	2016	
2	Goose Airways	2015	2016	

Preview

3 rows

25

/ page

<<

<

Page

1

>

>>

↓

Ci-dessous requêtes préalables pour vérifier les données sur les tables dispos dans raw.public, avant de réaliser les modèles dim_ dans dbt ⇒ On peut voir que les données sont relativement bien présentes et complètes.

Cependant, on ne peut pas identifier d'emblée les aéroports d'arrivée et de départ sur la table individual_flights avec une sélection de 10 lignes ⇒ on peut faire des tris sur cette table pour identifier les aéroports d'arrivée et de départ.

```
SELECT * FROM RAW.PUBLIC.AIRPORTS;
```

	index object	Airport_Code obj...	Airport_Name ob...	Airport_Employees	Airport_Size object	
0	0	FKT	Flocktopia	1000000	2000000	
1	1	NSA	Nestland Airport	20000	50000	
2	2	AMP	Amazon Mothers...	100000	10000	

[Preview](#) 3 rows 10 / page << < Page 1 of 1 > >> [Download](#)

```
SELECT * FROM RAW.PUBLIC.individual_flights;
```

	index int64 0 - 2269 	Flight_Id int64 100001 - 102270 	Airline_Code obj... GA 44.4% AA 34.9% FA 20.7%	Departure_Airpor... NSA 47% AMP 32.2% FKT 20.8%	Destination_Airp... FKT 53.7% AMP 27.8% 2 others 18.5%	Aircraft_Id object g72 44.4% t10 24.4% 3 others 31.2%	
0	0	100001	FA	FKT	NSA	12d	
1	1	100002	FA	FKT	NSA	12d	
2	2	100003	FA	FKT	NSA	12d	
3	3	100004	FA	FKT	NSA	12d	
4	4	100005	FA	FKT	NSA	12d	
5	5	100006	FA	FKT	NSA	12d	
6	6	100007	FA	FKT	NSA	12d	
7	7	100008	FA	FKT	NSA	12d	
8	8	100009	FA	FKT	NSA	12d	
9	9	100010	FA	FKT	NSA	12d	

2,270 rows, 6 cols 10 / page << < Page 1 of 227 > >> [Download](#)

```
SELECT *  
FROM RAW.PUBLIC.individual_flights  
WHERE "Destination_Airport_Code" NOT LIKE 'FKT'  
AND "Destination_Airport_Code" NOT LIKE 'AMP';
```

	index int64 0 - 2269 	Flight_Id int64 100001 - 102270 	Airline_Code obj... GA 56.1% AA 34.4% FA 9.5%	Departure_Airpor... AMP 58.9% FKT 41.1%	Destination_Airp... NSA 96.9% AA 3.1%	Aircraft_Id object g72 56.1% b23 27.9% 3 others 16%	
0	0	100001	FA	FKT	NSA	12d	
1	1	100002	FA	FKT	NSA	12d	
2	2	100003	FA	FKT	NSA	12d	
3	3	100004	FA	FKT	NSA	12d	
4	4	100005	FA	FKT	NSA	12d	
5	5	100006	FA	FKT	NSA	12d	
6	6	100007	FA	FKT	NSA	12d	
7	7	100008	FA	FKT	NSA	12d	
8	8	100009	FA	FKT	NSA	12d	
9	9	100010	FA	FKT	NSA	12d	

419 rows, 6 cols 10 / page << < Page 1 of 42 > >> [Download](#)

```
DESC TABLE RAW.PUBLIC.individual_flights;
```

	name object	type object	kind object	null? object	default object	primary key object	unique key object	c
0	index	NUMBER(38,0)	COLUMN	Y	None	N	N	N
1	Flight_Id	NUMBER(38,0)	COLUMN	Y	None	N	N	N
2	Airline_Code	VARCHAR(16777...	COLUMN	Y	None	N	N	N
3	Departure_Airpor...	VARCHAR(16777...	COLUMN	Y	None	N	N	N
4	Destination_Airpo...	VARCHAR(16777...	COLUMN	Y	None	N	N	N
5	Aircraft_Id	VARCHAR(16777...	COLUMN	Y	None	N	N	N

6 rows, 12 cols 10 / page << < Page 1 of 1 > >> [Download](#)

https://deepnote.com/app/aircraftproject/AircraftProject-60edcbf-bfd2-4ca8-8994-b830ca77e2d6?utm_source=app-settings&utm_medium=produ... 4/4