ÉCOLE 18.06

LA LIBERTÉ PAR LA SOUVERAINETÉ

Projet ETL

Supercar

Introduction

-Problématique: Fournir et trier des informations sur des sc

-Présentation du Dataset https://www.kaggle.com/datasets/sadiajavedd/sports-car-speed-power-and-pricing-dataset?resource=download

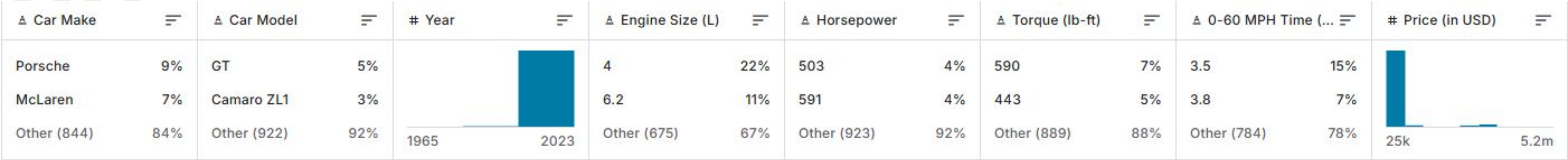
-Présentation du Pipeline Hop

-Présentation du Graph Neo4J

-Base de donnée MongoDB

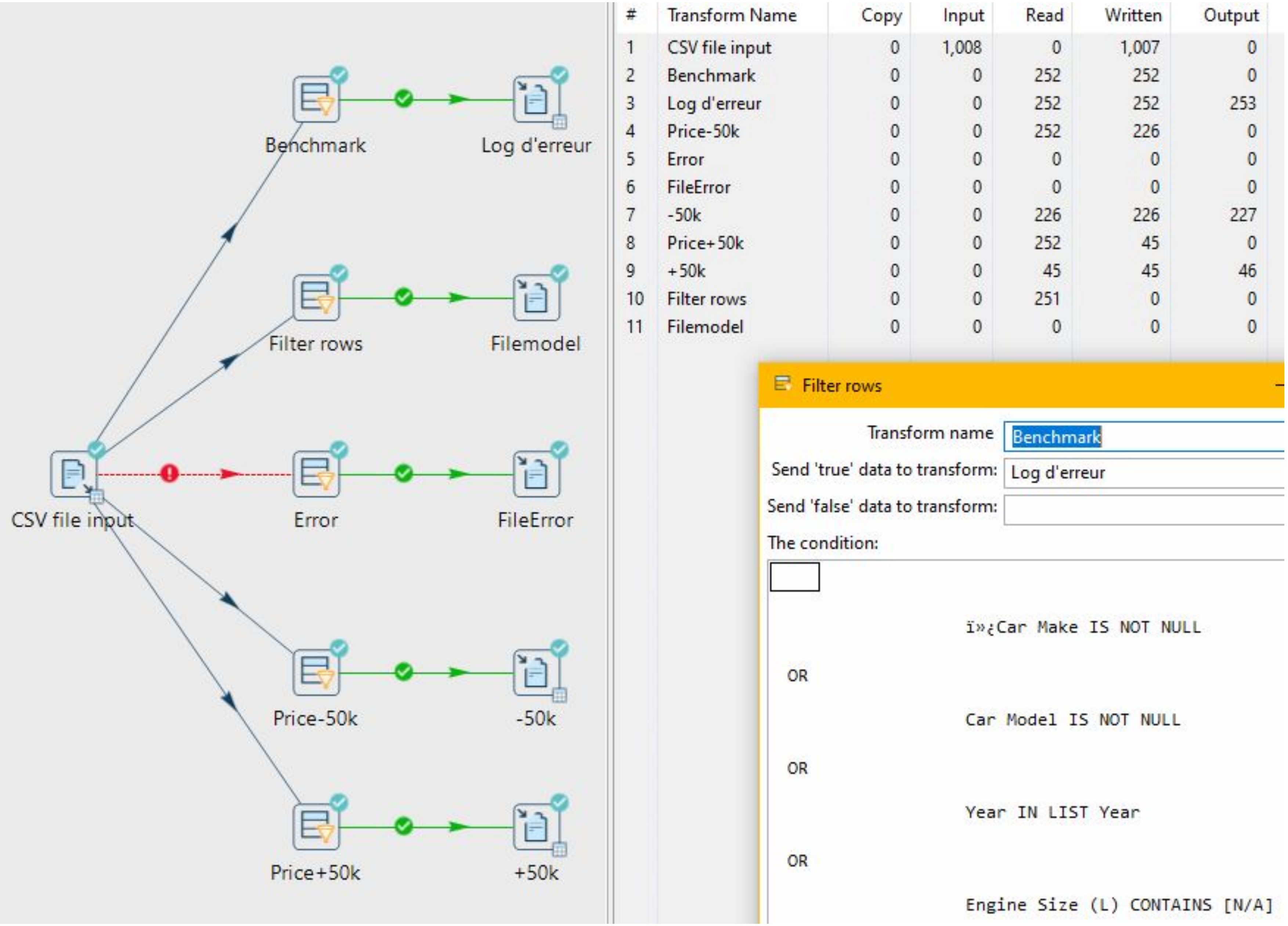
Dataset Supercar

Marque	Modèle	Année	Taille du moteur	Chevaux	Couple	0-100 km/h	Prix
Ex: Hyundai	Ex: i20	-	Litres	-	livre-pied(lb-ft)	Secondes	Dollars (USD)



Car Make		Car Model		# Year	Engine Size (L)		Horsepower		Torque (lb-ft)		0-60 MPH Time (...		# Price (in USD)
Porsche	9%	GT	5%		4	22%	503	4%	590	7%	3.5	15%	
McLaren	7%	Camaro ZL1	3%		6.2	11%	591	4%	443	5%	3.8	7%	
Other (844)	84%	Other (922)	92%	1965 2023	Other (675)	67%	Other (923)	92%	Other (889)	88%	Other (784)	78%	25k 5.2m

Pipeline Hop



#	Transform Name	Copy	Input	Read	Written	Output
1	CSV file input	0	1,008	0	1,007	0
2	Benchmark	0	0	252	252	0
3	Log d'erreur	0	0	252	252	253
4	Price-50k	0	0	252	226	0
5	Error	0	0	0	0	0
6	FileError	0	0	0	0	0
7	-50k	0	0	226	226	227
8	Price+50k	0	0	252	45	0
9	+50k	0	0	45	45	46
10	Filter rows	0	0	251	0	0
11	Filemodel	0	0	0	0	0

Filter rows

Transform name: Benchmark

Send 'true' data to transform: Log d'erreur

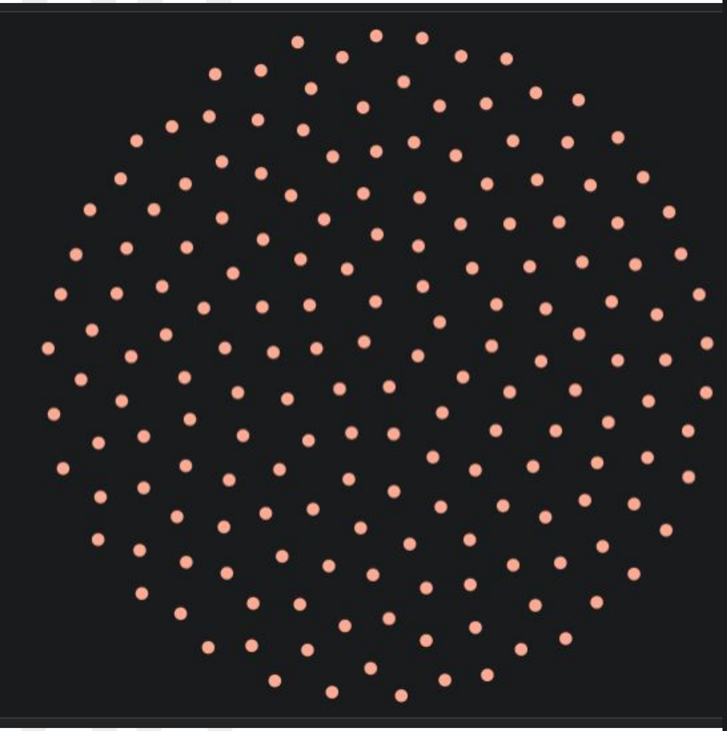
Send 'false' data to transform:

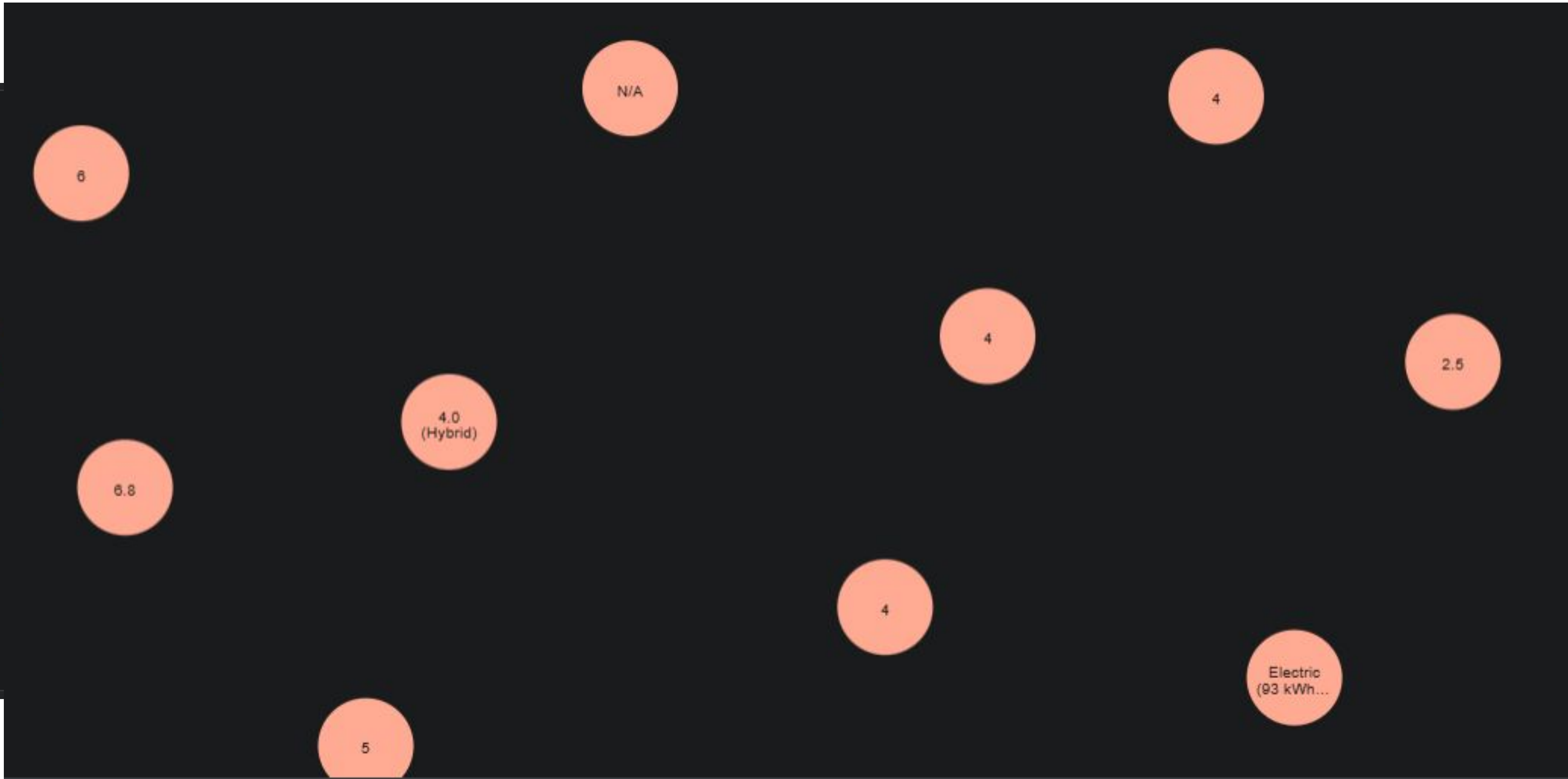
The condition:

```
                ï»¿Car Make IS NOT NULL
OR
                Car Model IS NOT NULL
OR
                Year IN LIST Year
OR
                Engine Size (L) CONTAINS [N/A]
```

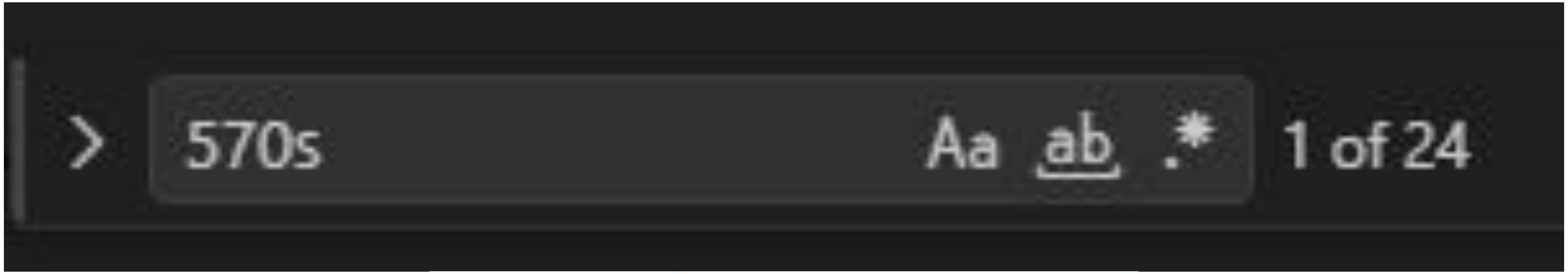
Graph Neo4J

Taille des moteurs pour tous les modèles



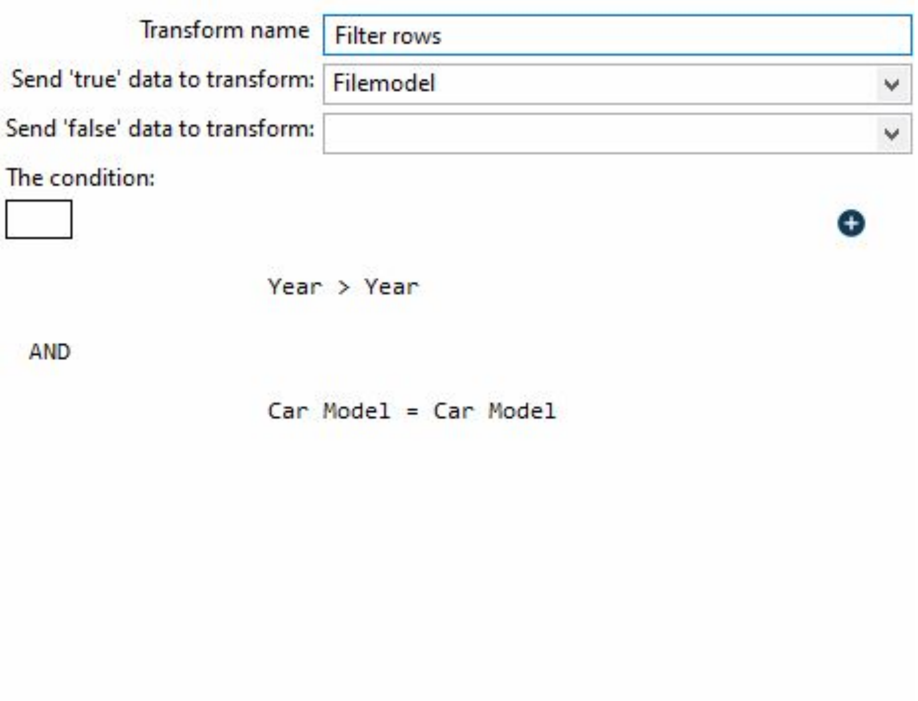


Autres difficultés rencontrés



570s

1 of 24



Transform name: Filter rows

Send 'true' data to transform: Filemodel

Send 'false' data to transform:

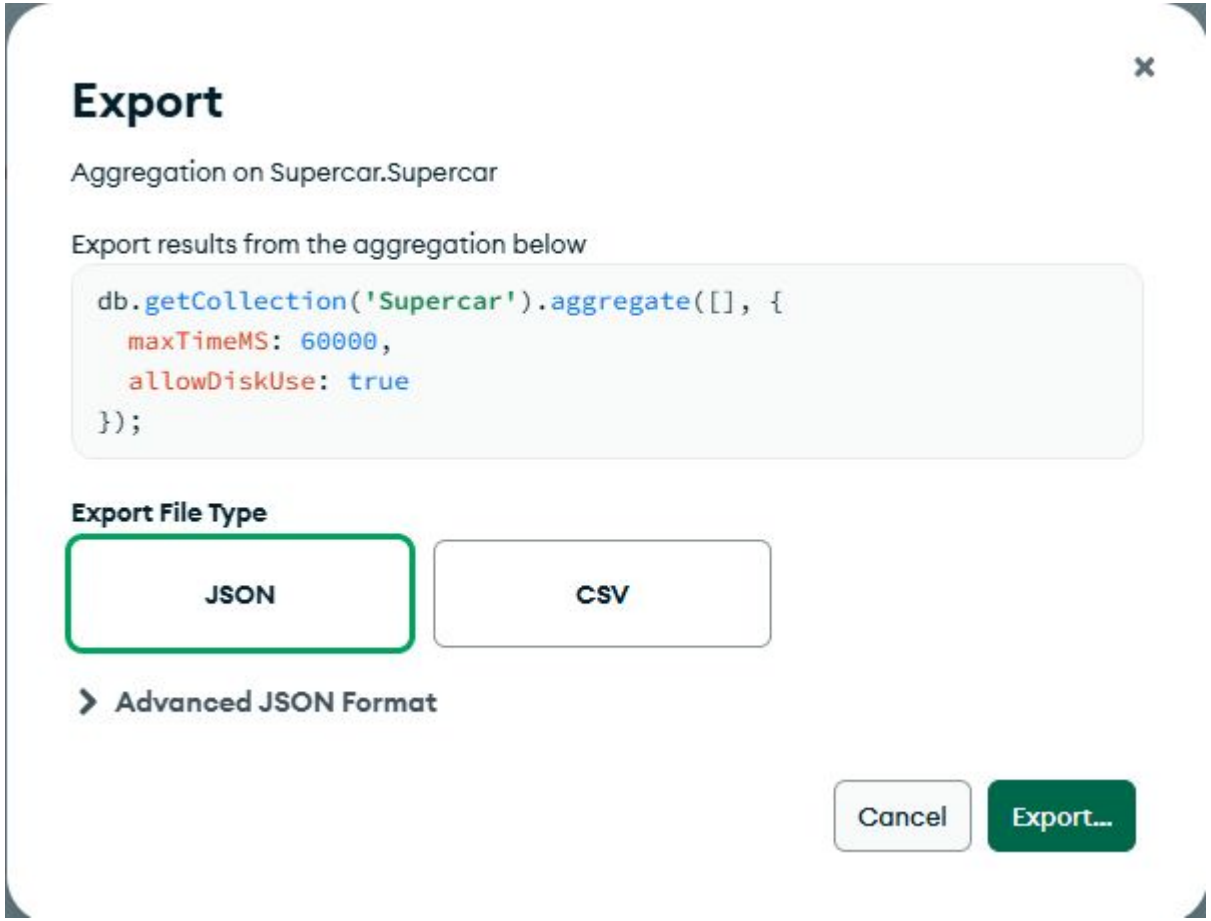
The condition:

```
                    Year > Year

AND

                    Car Model = Car Model
```

MongoDB ~~> JSON



Export

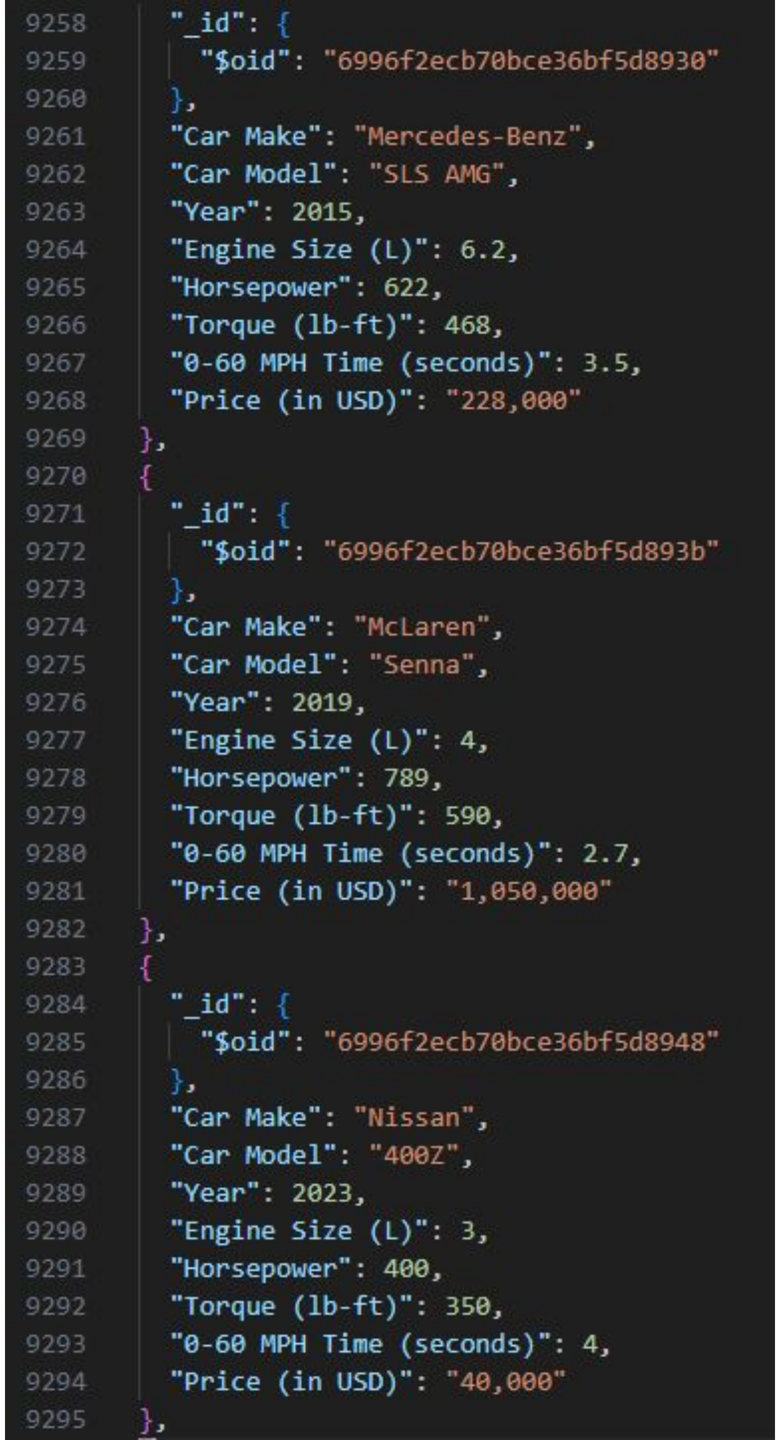
Aggregation on Supercar.Supercar

Export results from the aggregation below

```
db.getCollection('Supercar').aggregate([], {
  maxTimeMS: 60000,
  allowDiskUse: true
});
```

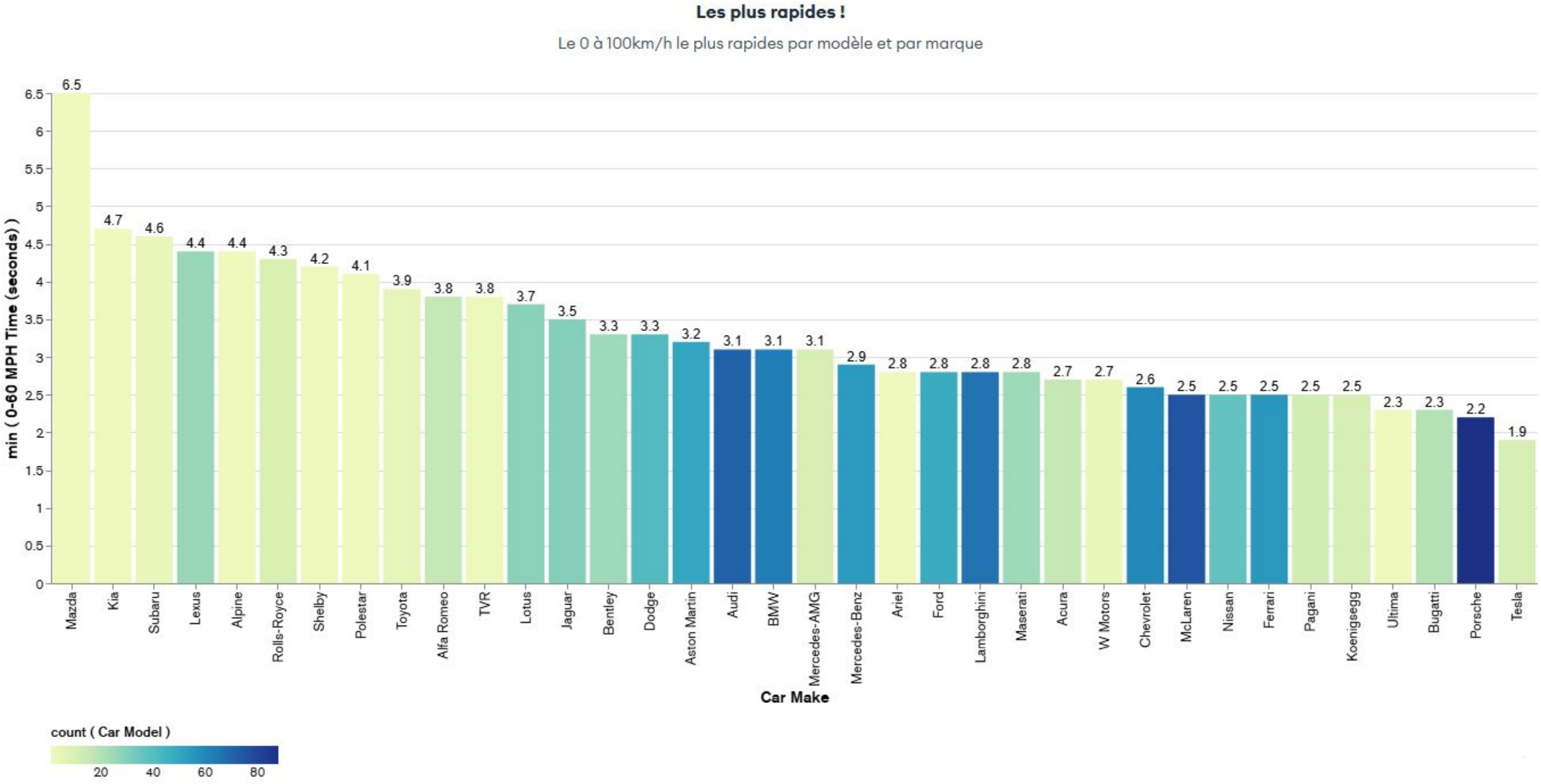
Export File Type

[✓ JSON] CSV



```
    "_id": {
      "$oid": "6996f2ecb70bce36bf5d8930"
    },
    "Car Make": "Mercedes-Benz",
    "Car Model": "SLS AMG",
    "Year": 2015,
    "Engine Size (L)": 6.2,
    "Horsepower": 622,
    "Torque (lb-ft)": 468,
    "0-60 MPH Time (seconds)": 3.5,
    "Price (in USD)": "228,000"
  },
  {
    "_id": {
      "$oid": "6996f2ecb70bce36bf5d893b"
    },
    "Car Make": "McLaren",
    "Car Model": "Senna",
    "Year": 2019,
    "Engine Size (L)": 4,
    "Horsepower": 789,
    "Torque (lb-ft)": 590,
    "0-60 MPH Time (seconds)": 2.7,
    "Price (in USD)": "1,050,000"
  },
  {
    "_id": {
      "$oid": "6996f2ecb70bce36bf5d8948"
    },
    "Car Make": "Nissan",
    "Car Model": "400Z",
    "Year": 2023,
    "Engine Size (L)": 3,
    "Horsepower": 400,
    "Torque (lb-ft)": 350,
    "0-60 MPH Time (seconds)": 4,
    "Price (in USD)": "40,000"
  },
```

BDD MongoDB



Les plus rapides !

Le 0 à 100km/h le plus rapides par modèle et par marque

Merci à VROUUUMMMM!