

# MindsDB

## [Introducing Minds](#)

### [doc](#)

#### Objectif :

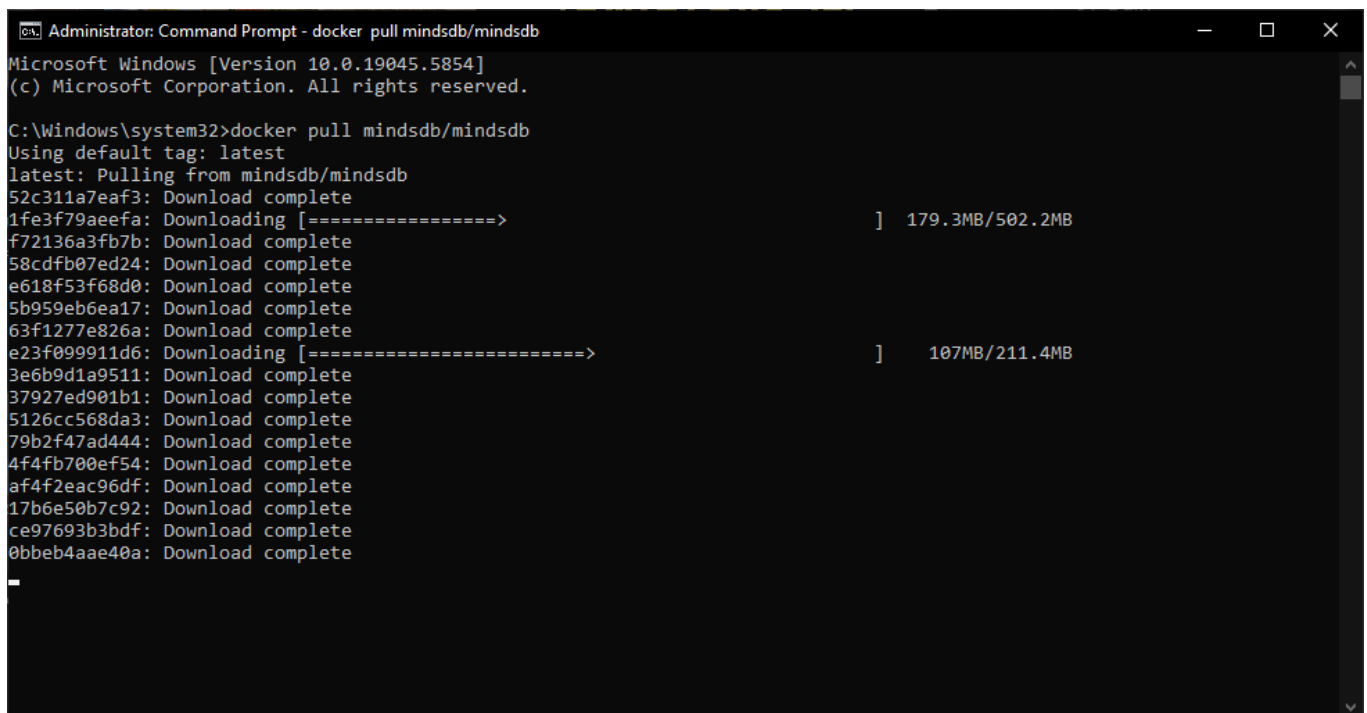
Prédire le montant des ventes ( Sales ) en fonction de certaines caractéristiques d'une commande comme la date, le mode de livraison, le segment client et la catégorie de produit.

## Setup MindsDB Environment

### Option A: Using Docker (Recommended)

1. Install Docker if you don't have it: <https://www.docker.com/products/docker-desktop>
2. Open terminal/command prompt and run:

```
docker pull mindsdb/mindsdb
docker run -p 47334:47334 -p 47335:47335 mindsdb/mindsdb
```

A screenshot of a Windows Command Prompt window titled "Administrator: Command Prompt - docker pull mindsdb/mindsdb". The window shows the output of the command "docker pull mindsdb/mindsdb". It indicates that the latest tag is being pulled from the mindsdb/mindsdb repository. The output shows a list of layers being downloaded, with progress bars and sizes. For example, "52c311a7eaf3: Download complete" and "1fe3f79aeefa: Downloading [=====>] 179.3MB/502.2MB". The final output shows "0bbeb4aae40a: Download complete".

```
Administrator: Command Prompt - docker pull mindsdb/mindsdb
Microsoft Windows [Version 10.0.19045.5854]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\system32>docker pull mindsdb/mindsdb
Using default tag: latest
latest: Pulling from mindsdb/mindsdb
52c311a7eaf3: Download complete
1fe3f79aeefa: Downloading [=====>] 179.3MB/502.2MB
f72136a3fb7b: Download complete
58cdfb07ed24: Download complete
e618f53f68d0: Download complete
5b959eb6ea17: Download complete
63f1277e826a: Download complete
e23f099911d6: Downloading [=====>] 107MB/211.4MB
3e6b9d1a9511: Download complete
37927ed901b1: Download complete
5126cc568da3: Download complete
79b2f47ad444: Download complete
4f4fb700ef54: Download complete
af4f2eac96df: Download complete
17b6e50b7c92: Download complete
ce97693b3bdf: Download complete
0bbeb4aae40a: Download complete
```

Access MindsDB GUI at <http://localhost:47334>

install pyarrow to upload files and lightwood to create models

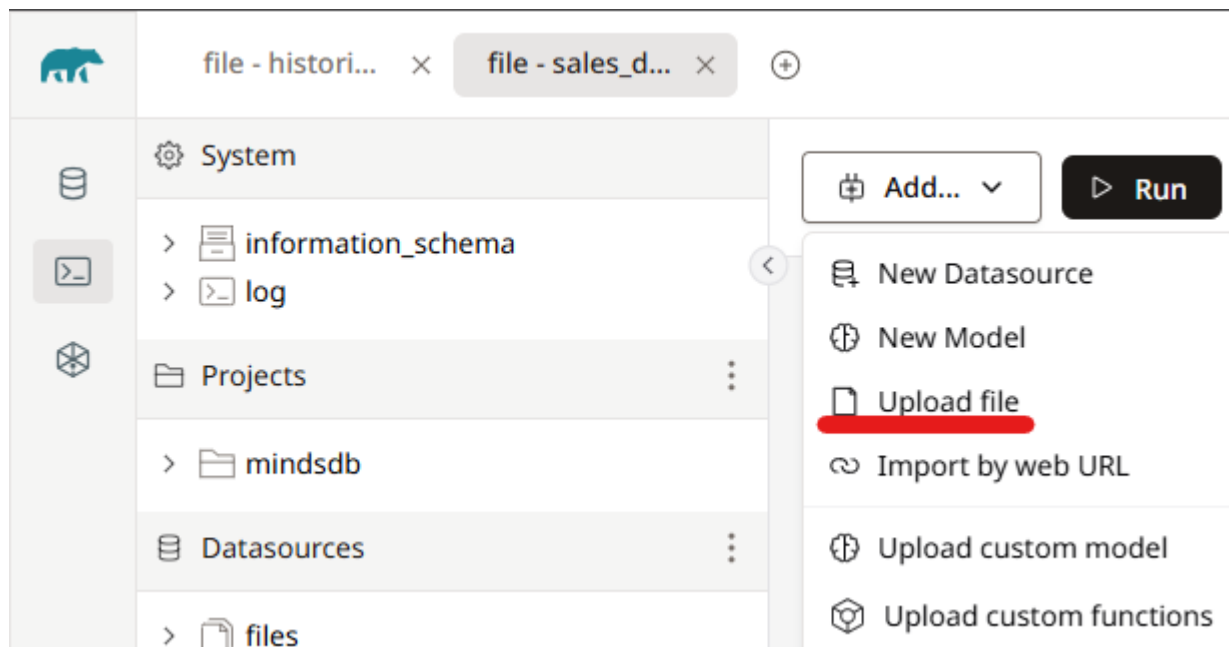
```
docker ps
docker exec -it <container ID> /bin/bash
root@<Container ID>:/mindsdb# pip install pyarrow
root@<Container ID>:/mindsdb# pip install lightwood
```

## Option B: Using MindsDB Cloud

1. Go to
2. Sign up for a free account

## Importation des Données

Once MindsDB is running (via Docker or MindsDB Cloud), upload your dataset. Each uploaded file becomes a table in the built-in `files` database.



## ← Upload file

Browser

URL

Import a file

↑ train.csv

×

Supported files: csv, xlsx, xls, sheet, json, parquet, pdf, txt

Save As...

sales\_data

Note: For spaces use an underscore \_

Save and Continue →

On Successful upload, you should be presented with the SQL Editor.

⚙ Add... ▾

▶ Run

✓ 0.2 seconds

ⓘ Lea

```
1
2 --- MindsDB ships with a filesystem database called 'files'
3 --- Each file you uploaded is saved as a table there.
4 ---
5 --- You can always check the list tables in files as follows:
6
7 SHOW TABLES FROM files;
8
9 --- These files can be queried as tables,
10 --- You just uploaded sales_data
```

Last Query: SHOW TABLES FROM files

Exp

#### Tables\_in\_files

1	historical_sales
2	historical_sales2
3	sales_data

Add... v

Run

0.1 seconds

Learning Hub

```
1
2 --- MindsDB ships with a filesystem database called 'files'
3 --- Each file you uploaded is saved as a table there.
4 ---
5 --- You can always check the list tables in files as follows:
6
7 SHOW TABLES FROM files;
8
9 --- These files can be queried as tables,
10 --- You just uploaded sales_data
11
12 | SELECT * FROM files.sales_data LIMIT 5;
13
```

Last Query: SELECT \* FROM files.sales\_data LIMIT 5

Export v

Row ID	Order ID	Order Date	Ship Date	Ship Mode	
1	CA-2017-152156	08/11/2017	11/11/2017	Second Class	(
2	CA-2017-152156	08/11/2017	11/11/2017	Second Class	(
3	CA-2017-138688	12/06/2017	16/06/2017	Second Class	I
4	US-2016-108966	11/10/2016	18/10/2016	Standard Class	:
5	US-2016-108966	11/10/2016	18/10/2016	Standard Class	:

## Understand the Data

This dataset contains eighteen columns, including the sales column. The order date, ship date, ship mode, segment, and category columns will be used to make predictions.

To see these columns, run the query below

Add...
Run
0.3 seconds

```

5 --- You can always check the list tables in files as follows:
6
7 SHOW TABLES FROM files;
8
9 --- These files can be queried as tables,
10 --- You just uploaded sales_data
11
12 SELECT * FROM files.sales_data LIMIT 5;
13
14 SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM files.sales_data;
15
16

```

Last Query: SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM files.sales\_data

	Order Date	Ship Date	Ship Mode	Segment	Category	Sales
1	08/11/2017	11/11/2017	Second Class	Consumer	Furniture	261.96
2	08/11/2017	11/11/2017	Second Class	Consumer	Furniture	731.94

Page Size: 50
1 to 50 of 9,800

What do these columns stand for?

- Order Date: The date the product was ordered by the consumer.
- Ship Date: The date the product was shipped by the company.
- Ship Mode: The shipment class that the customer has selected. There are four classes available here: Same Day, First Class, Standard Class, and Second Class.
- Segment: The product's segmentation. This field has three possible values: consumer, corporate, and home office.
- Category: The category to which the products belong. There are three values in this column: furniture, office supplies, and technology.

## Train a Predictor

we can now create a model (predictor) to forecast sales:

Add...
Run
24.3 seconds

```

15
16
17
18
19 CREATE PREDICTOR mindsdb.predict_sales
20 FROM files
21 (SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM sales_data)
22 PREDICT Sales;
23
24
25
26

```

Last Query: CREATE PREDICTOR mindsdb.predict\_sales FROM files (SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM

	NAME	ENGINE	PROJECT	ACTIVE	VERSION	STATUS	ACCURACY
1	predict_sales	lightwood	mindsdb	true	1	generating	[NULL]

The following syntax illustrates how the query works.

```
CREATE PREDICTOR mindsdb.[predictor_name]
FROM [integration_name]
    (SELECT [column_name, ...] FROM [table_name])
PREDICT [target_column];
```

- [predictor name] : The name of the model.
- [integration name] : The name of the database where your table is stored, for example, files.
- [column name,...] : The field to predict, as well as the fields to train for the prediction.
- [table name] : The table containing the columns, for example 'sales data'.
- [target column] : The column to be predicted, such as Sales.

## Making Predictions.

```
SELECT *
FROM information_schema.MODELS
LIMIT 50;
```

Query: SELECT \* FROM information\_schema.MODELS LIMIT 50

NAME	ENGINE	PROJECT	ACTIVE	VERSION	STATUS
daily_sales_predictor	lightwood	mindsdb	true	1	complete
predict_sales	lightwood	mindsdb	true	1	<u>complete</u>

Once training is complete, use this query to predict the sales for a specific set of values:

Consider the following question for this section.

A customer orders a product on March 28th, 2019, and the product is dispatched on March 31st, 2019. The product is from the consumer segment, the furniture category, and was shipped first class. How much did the item cost?

```

28
29 Select Sales_confidence, Sales_explain
30 FROM mindsdb.predict_sales
31 WHERE `Order Date` = '28/03/2019' AND
32       `Ship Date` = '31/03/2019' AND
33       `Ship Mode` = 'First Class' AND
34       Segment = 'Consumer' AND
35       Category = 'Furniture';
36
37

```

Last Query: Select Sales\_confidence, Sales\_explain FROM mindsdb.predict\_sales WHERE `Order Date` = '28/03/2019' AND `Ship Date` = '31/03/2019' AND `Ship Mode` = 'First Class' AND Segment = 'Consumer' AND Category = 'Furniture';

	Sales_confidence	Sales_explain
1	0.8	{"predicted_value": 292.72100830078125, "confidence": 0.8,

Using the newly constructed model, the product costs around \$292.72. The model has a confidence level of 80%, or 0.8.

```
SHOW TABLES FROM files;
```

```
--- These files can be queried as tables,
```

```
--- You just uploaded sales_data
```

```
SELECT * FROM files.sales_data LIMIT 5;
```

```
SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM
files.sales_data;
```



```
DROP MODEL IF EXISTS mindsdb.predict_sales;
```

```
CREATE PREDICTOR mindsdb.predict_sales
```

```
FROM files
```

```
(SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM  
sales_data)
```

```
PREDICT Sales;
```

```
SELECT *
```

```
FROM information_schema.MODELS
```

```
LIMIT 50;
```

```
Select Sales_confidence, Sales_explain
```

```
FROM mindsdb.predict_sales
```

```
WHERE `Order Date` = '28/03/2019' AND
```

```
`Ship Date` = '31/03/2019'AND
```

```
`Ship Mode` = 'First Class'AND
```

```
Segment = 'Consumer'AND
```

```
Category = 'Furniture';
```