

# mindsdb

## AI's Query Engine

2024/2025

Présenté par :

Abdessamad Misdak

Imade El Hilali

Supervisé par :

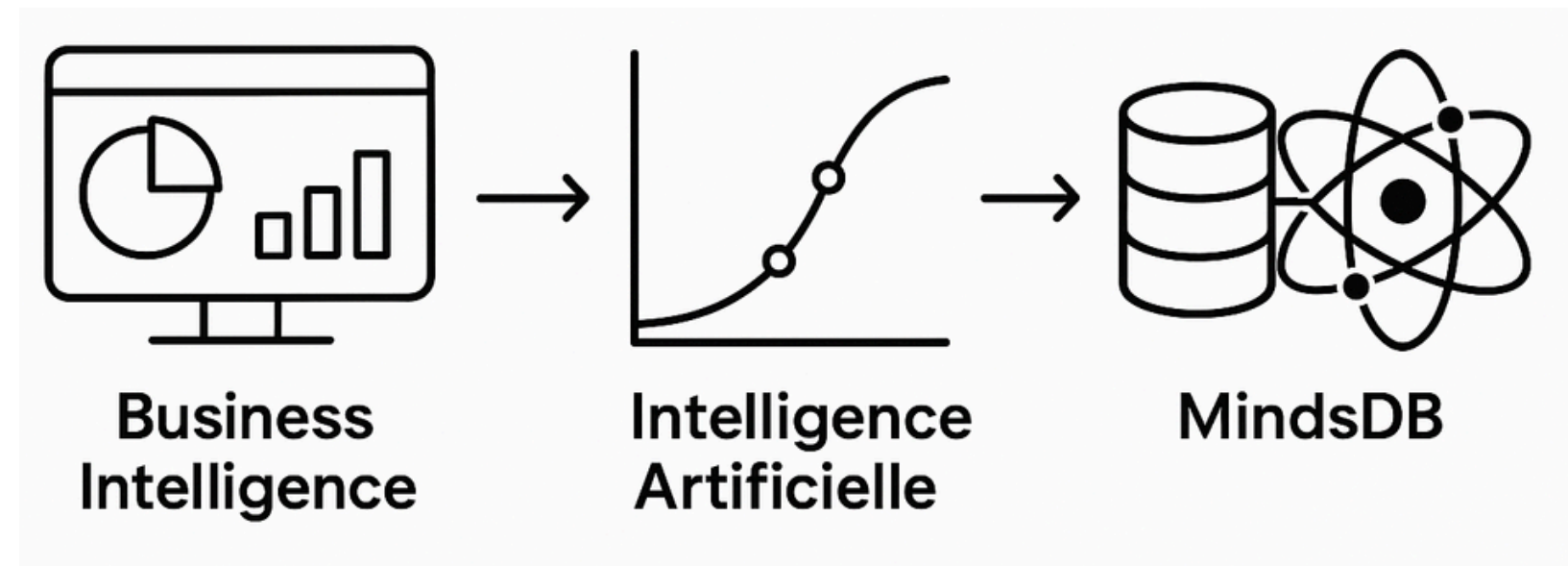
Pr. LOTFI NAJDI

# Plan

1. Pourquoi MindsDB en BI ?
2. Qu'est-ce que MindsDB ?
3. Architecture et fonctionnalités clés
4. Cas d'utilisation en Business Intelligence
5. Démonstration pratique
6. Perspectives et Conclusion

# **Pourquoi MindsDB en BI ?**

# Introduction : Pourquoi MindsDB en BI ?



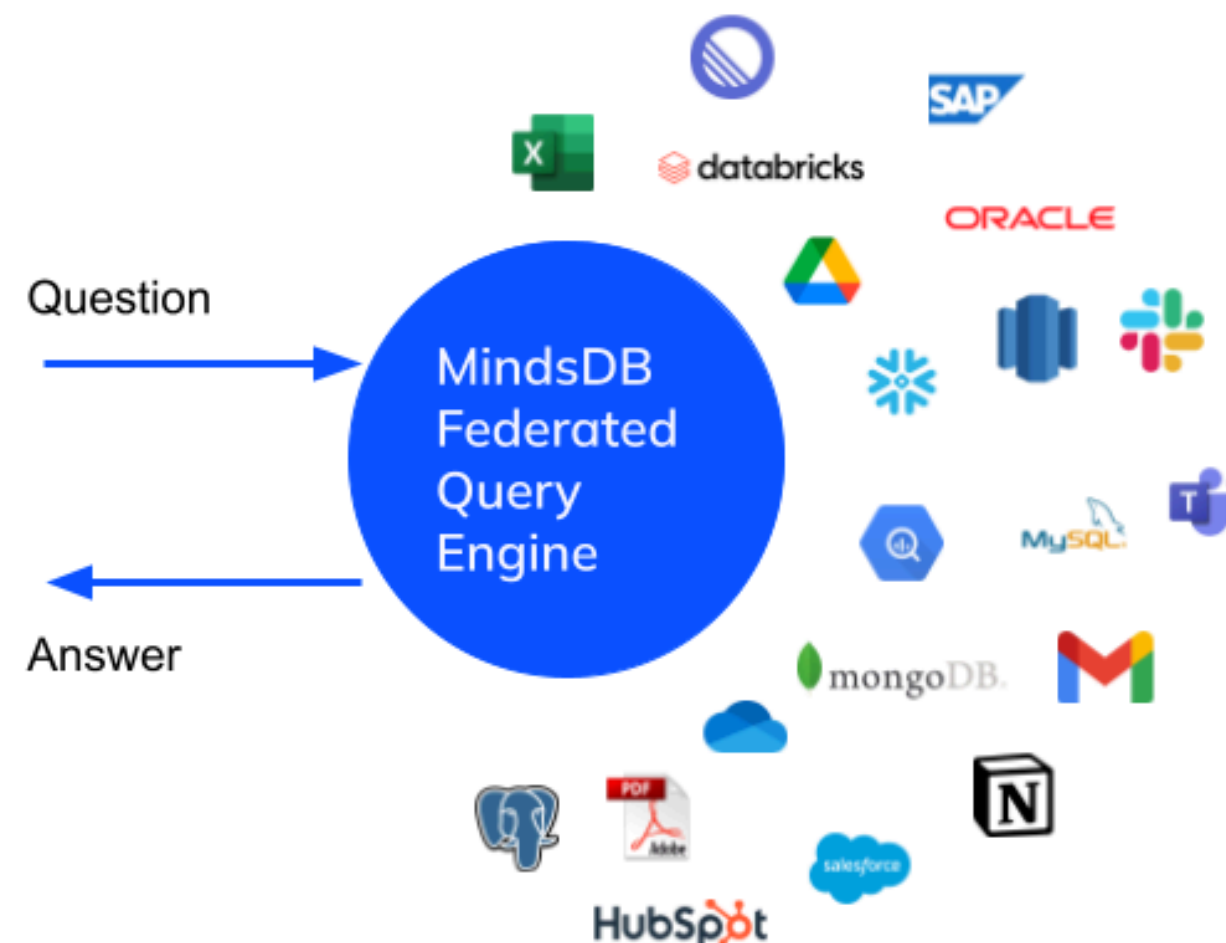
*MindsDB simplifie cette intégration en rendant l'IA accessible directement depuis les bases de données*

- Pour les professionnels de la BI qui maîtrisent déjà le SQL, c'est un grand avantage : ils peuvent utiliser la puissance de l'IA sans avoir besoin de compétences techniques avancées en Machine Learning.

**Qu'est-ce que  
MindsDB ?**

# Qu'est-ce que MindsDB ?

- MindsDB est une plateforme open-source créée en 2017.
- Son objectif principal est de simplifier l'intégration de l'IA avec les bases de données.
- Elle permet de rendre les prédictions accessibles en transformant la manière dont les développeurs implémentent des prédictions en utilisant des commandes SQL simples.



# Qu'est-ce que MindsDB ?

MindsDB est disponible en plusieurs éditions :

Critère	Édition Communautaire (CE)	Éditions Cloud/Commerciales (CC)
Déploiement	Auto-hébergé via Docker ou pip	Cloud géré (Demo, Starter, Pro), ou sur site
Support	Communautaire (Slack, GitHub)	Basique (Starter) , Assistance technique prioritaire (Pro)
Coût	Gratuit (open-source)	Starter payant, Pro/Enterprise sur devis
Scalabilité	Selon infrastructure utilisateur	Gérée par MindsDB, meilleure scalabilité et sécurité
Fonctionnalités	Fonctionnalités de base, intégrations BD, modèles ML pré-entraînés	CE + modèles avancés, intégrations supplémentaires

# **Architecture et fonctionnalités clés**



# Architecture et fonctionnalités clés

- L'architecture de MindsDB est conçue pour être un intergiciel flexible. Elle interagit avec diverses sources de données, frameworks ML, applications et autres outils de développement.
- Les trois capacités fondamentales sont :



**Connect**



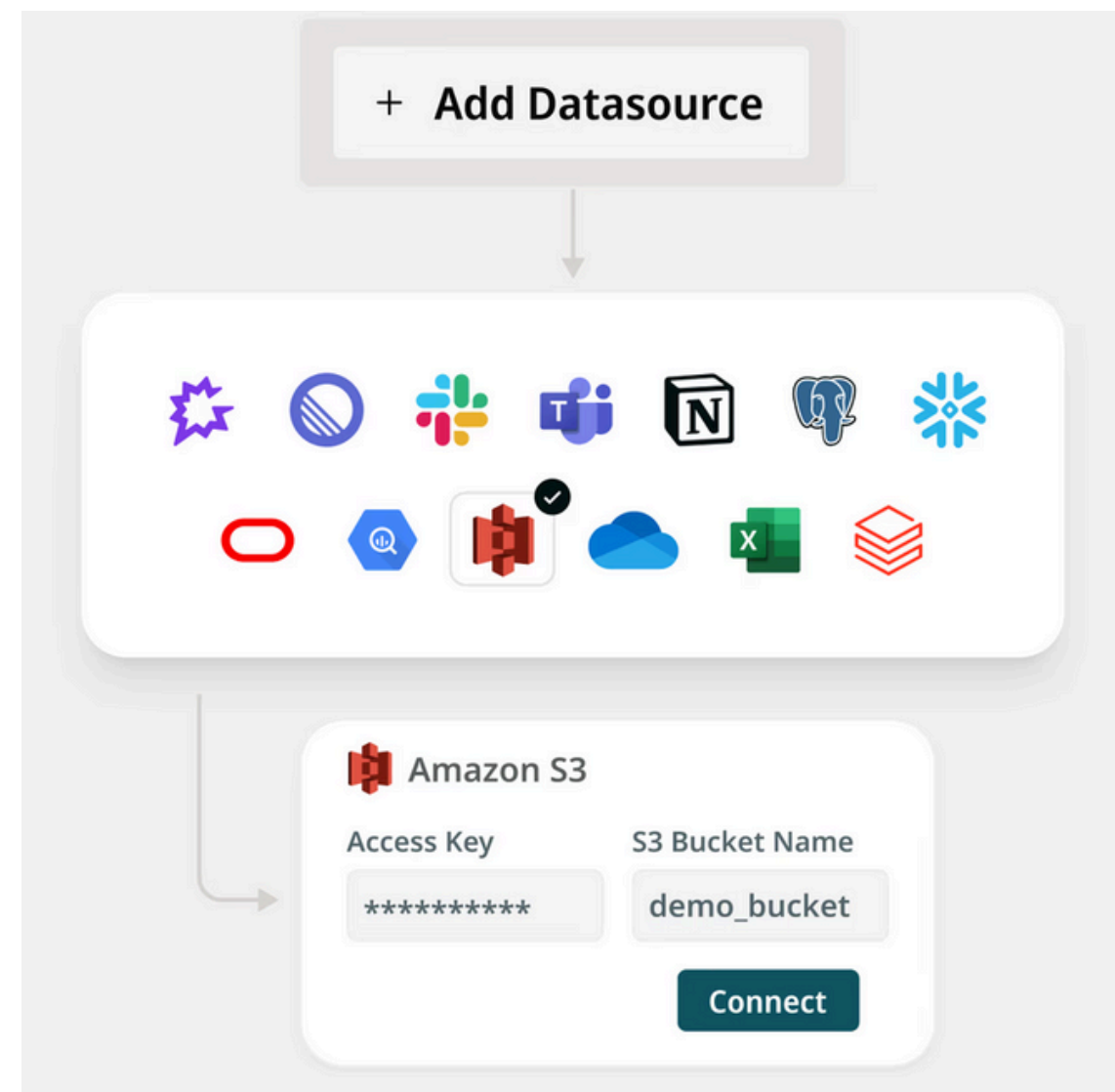
**Unify**



**Respond**

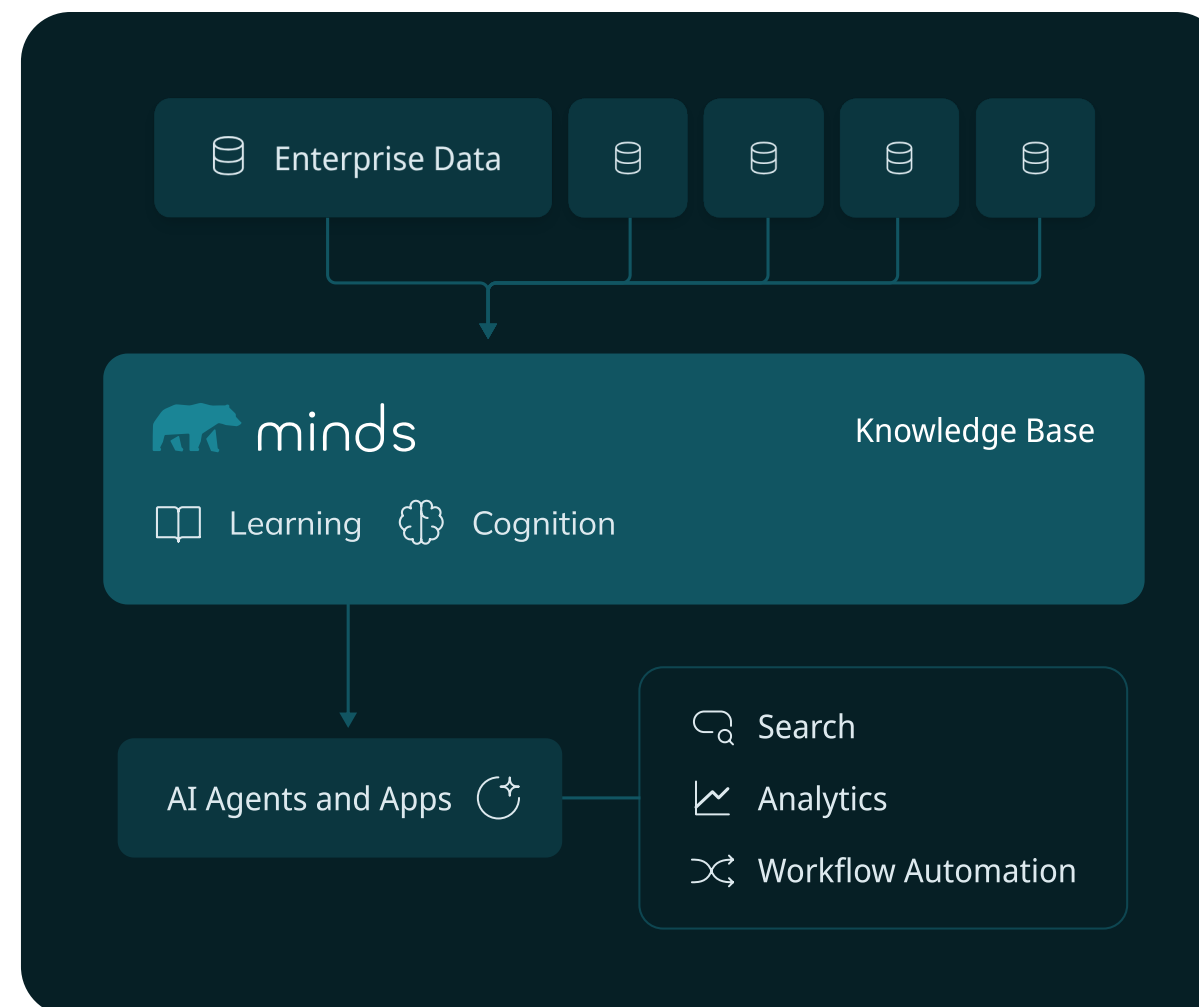
# Architecture et fonctionnalités clés

- **Connect** : Permet de connecter des données à partir de centaines de sources de données qui s'intègrent à MindsDB, incluant les bases de données, les entrepôts de données, les applications et les bases de données vectorielles.



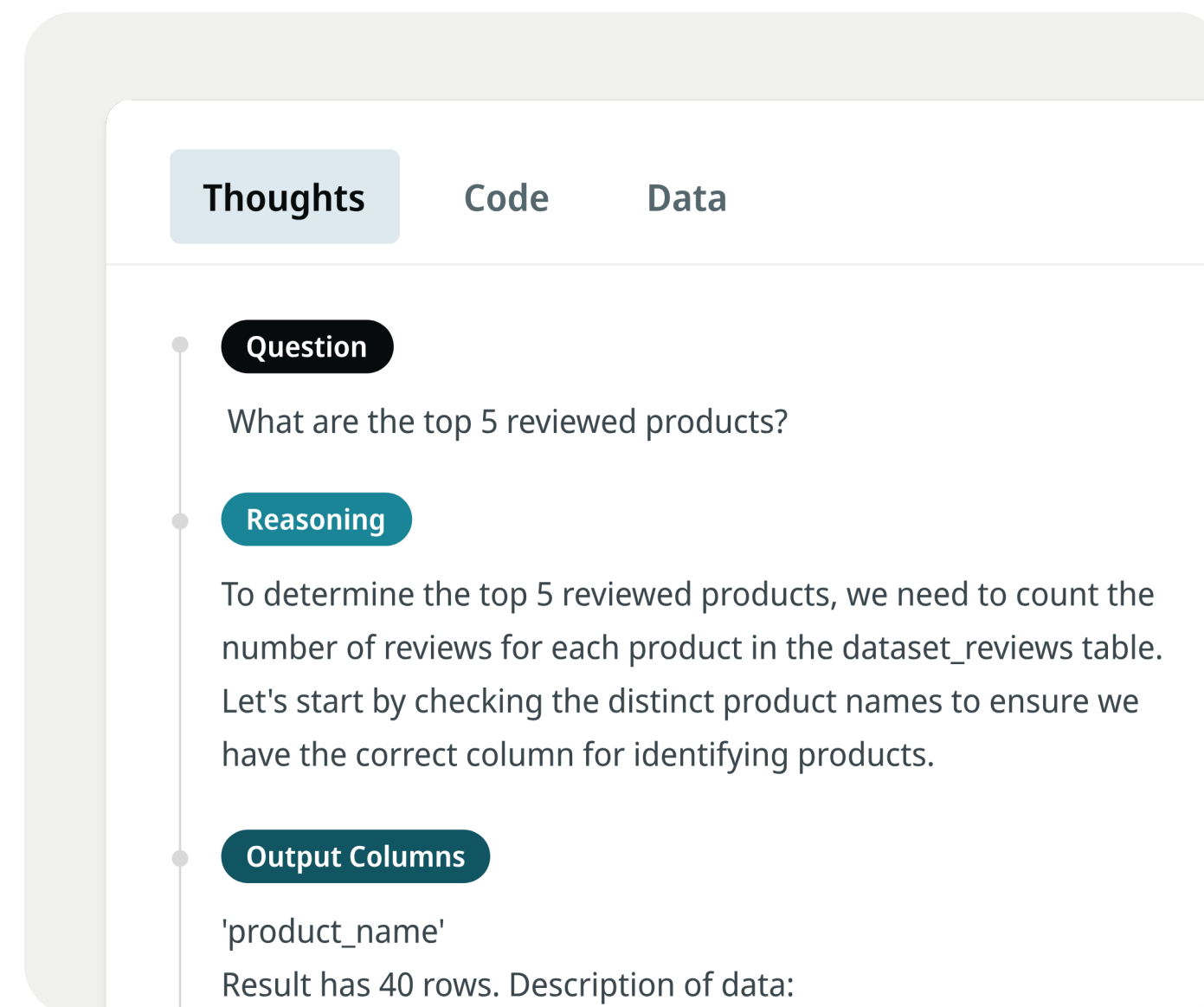
# Architecture et fonctionnalités clés

- **Unify** : Permet d'unifier les données provenant de plusieurs sources (structurées et non structurées) au sein de MindsDB. Cela permet d'effectuer des requêtes fédérées comme si toutes les données résidaient dans une base de données unique.



# Architecture et fonctionnalités clés

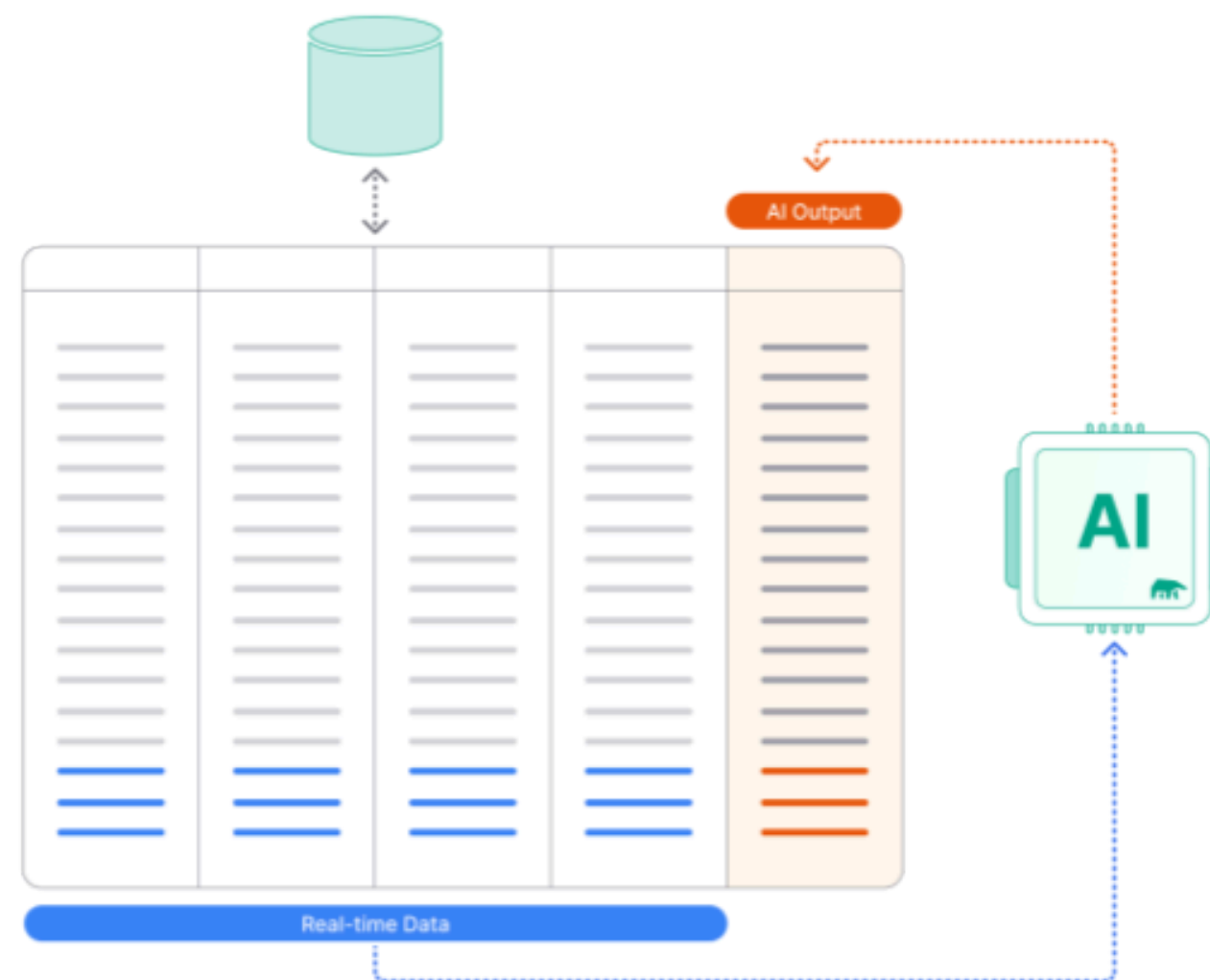
- **Respond** : Permet de générer des réponses précises et contextuelles à partir des données unifiées en utilisant des agents ou l'API MCP (Model Context Protocol).



# Architecture et fonctionnalités clés

- Le concept fondamental de MindsDB : AI Tables

Une AI Table est un modèle ML encapsulé dans une table virtuelle. Elle peut générer des prédictions en réponse à des requêtes, en s'appuyant sur un modèle ML sous-jacent et les données d'autres tables.



MindsDB has enhanced SQL by introducing **AI-Tables**, providing developers with AI capabilities to **solve problems where data and AI are closely intertwined** (coupled).

```
SELECT
    data.*, ai_model.output
FROM data
JOIN ai_model
```

# Architecture et fonctionnalités clés

- **Différence Fondamentale avec les Tables Standard**

Les tables de base de données traditionnelles sont conçues pour donner une réponse déterministe pour une entrée donnée.

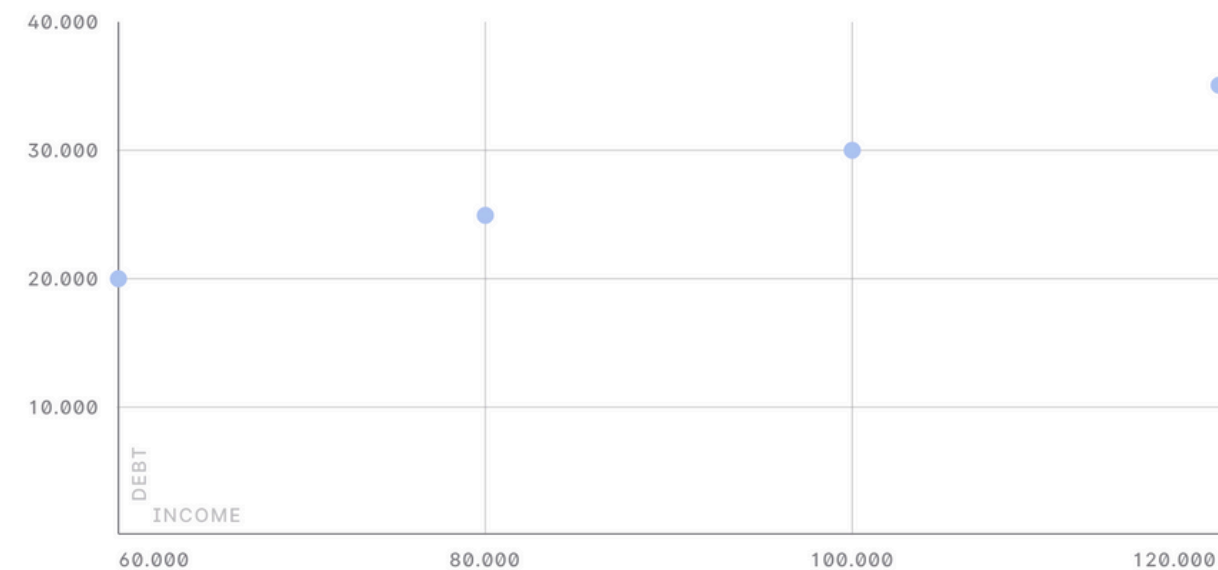
1

```
SELECT income, debt
FROM income_table;
```

```
+-----+-----+
|income|debt |
+-----+-----+
|60000 |20000 |
|80000 |25100 |
|100000|30040 |
|120000|36010 |
+-----+-----+
```

2

Income vs Debt



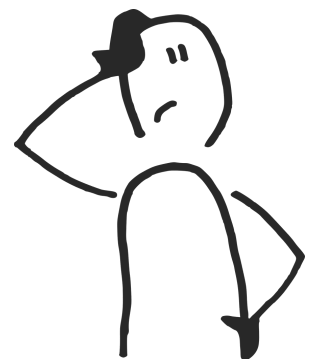
```
SELECT income, debt
FROM income_table
WHERE income = 80000;
```

```
+-----+-----+
|income|debt |
+-----+-----+
|80000 |25100 |
+-----+-----+
```

3

```
SELECT income, debt
FROM income_table
WHERE income = 90000;
```

Empty set (0.00 sec)



# Architecture et fonctionnalités clés

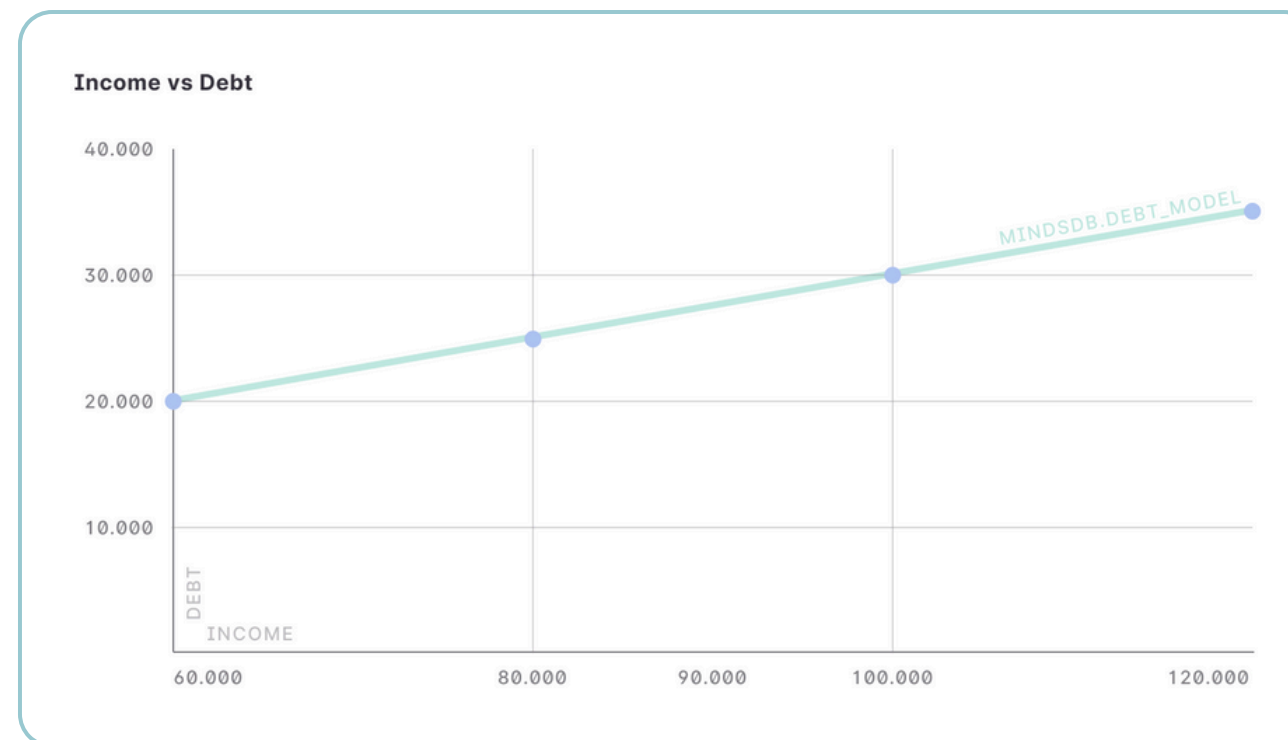
- **Différence Fondamentale avec les Tables Standard**

**Les Tables d'IA** (qui sont les modèles eux-mêmes créés via CREATE MODEL) sont conçues pour générer une réponse approximative pour une entrée donnée.

1

```
CREATE MODEL mindsdb.debt_model
FROM income_table
PREDICT debt;
```

2



3

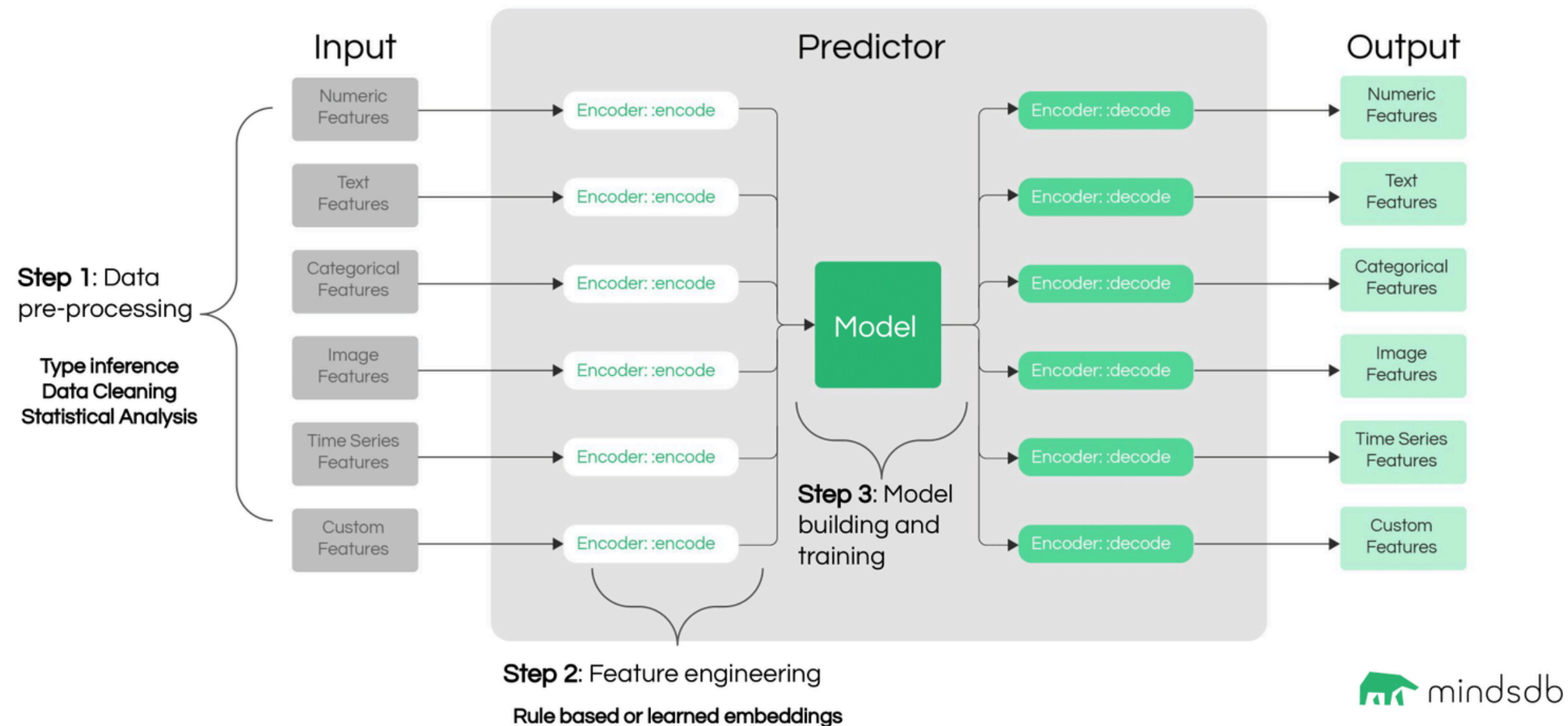
```
SELECT income, debt
FROM mindsdb.debt_model
WHERE income = 90000;
```

```
+-----+-----+
|income|debt |
+-----+-----+
|90000 |27820|
+-----+-----+
```

# Architecture et fonctionnalités clés

- Facilité d'utilisation de l'AutoML avec Lightwood dans MindsDB

Lightwood, le moteur d'IA par défaut de MindsDB, rend la création de modèles d'apprentissage automatique accessible à tous, même sans expertise approfondie en science des données. Il automatise et simplifie le processus en trois étapes clés :





# Architecture et fonctionnalités clés

- **Qu'est-ce que le Moteur de Requêtes Fédéré ?**

Le Moteur de Requêtes Fédéré de MindsDB est au cœur de sa capacité à unifier les données d'entreprise.

- **Connexion aisée** : MindsDB se connecte à plus de 200 sources de données via une simple commande SQL CREATE DATABASE.
- **Requête unifiée** : Les sources disparates sont interrogées comme une base de données unique grâce à un dialecte SQL complet ou même en langage naturel.
- **Traduction intelligente** : Le moteur traduit chaque requête dans le dialecte spécifique de la source sous-jacente.
- **Zéro stockage** : Les données restent dans leur source d'origine, MindsDB extrait les données directement à chaque requête.
- **Agrégation fluide** : Les résultats des sources sont combinés et présentés de manière unifiée.

# Architecture et fonctionnalités clés

- **Pourquoi les Requêtes Fédérées sont-elles Cruciales pour la BI ?**

La BI exige une vue globale et à jour des données d'entreprise. Pourtant, ces données sont souvent fragmentées dans des silos (bases de données, fichiers, ...), rendant l'analyse complexe et lente.

- **Zéro ETL** : Les données sont analysées sans être déplacées ni copiées, restant dans leurs sources d'origine.
- **Accès unifié** : Une interface unique (SQL ou langage naturel) permet d'interroger toutes les sources comme une seule base, simplifiant l'accès pour les outils BI et les utilisateurs.
- **Données en temps réel** : Les requêtes directes à la source offrent des informations toujours à jour, essentielles pour des décisions rapides et éclairées.
- **Analyse accélérée** : En supprimant la complexité de l'intégration, les analystes posent des questions complexes et obtiennent des réponses instantanées, sans attendre les rapports ETL.

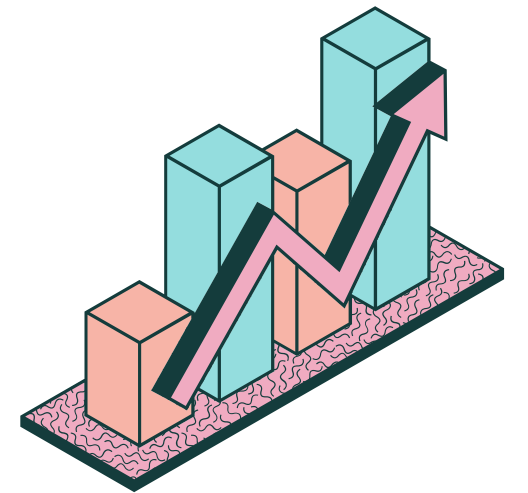
# **Cas d'utilisation en Business Intelligence**

# Cas d'utilisation en Business Intelligence

MindsDB intègre des prédictions intelligentes directement dans les flux BI, rendant l'analyse plus puissante.

## 1. Prédiction des Ventes :

- **Objectif** : Anticiper les ventes futures pour améliorer la planification et les stocks.
- **MindsDB** : Crée des modèles de séries temporelles via `CREATE MODEL`, en spécifiant la table des ventes et la colonne cible (sales).
- **Intégration BI** : Les prédictions sont accessibles via l'AI Table, jointes aux données réelles. Les analystes visualisent les prévisions dans des tableaux de bord BI pour optimiser la planification, les stocks et le marketing.



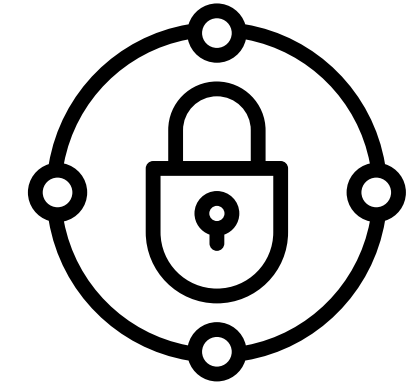
# Cas d'utilisation en Business Intelligence



## 2. Analyse de Sentiment des Clients :

- **Objectif** : Évaluer les avis clients (positif, neutre, négatif) pour améliorer produits et services.
- **MindsDB** : Utilise un modèle d'analyse textuelle (avec OpenAI) via CREATE MODEL et une prompt\_template pour traiter les commentaires.
- **Intégration BI** : Les sentiments prédits, accessibles via l'AI Table, sont joints aux données clients. Les outils BI affichent la répartition des sentiments, révélant les problèmes récurrents et l'impact des améliorations.

# Cas d'utilisation en Business Intelligence



## 3. Détection de Fraudes :

- **Objectif** : Identifier automatiquement des comportements suspects au sein des bases de données transactionnelles.
- **Comment ?**
  - Le modèle apprend à reconnaître les comportements "normaux" pour ensuite signaler les écarts significatifs.
- **Intégration dans la BI** :
  - Les résultats du modèle (score de risque, indicateur binaire) sont exposés via une AI Table.
  - Les outils BI (Power BI, Tableau...) peuvent interroger cette AI Table comme une source normale de données.

# **Démonstration Pratique**

# Installation de MindsDB

Il suffit d'installer Docker, puis de lancer ces deux lignes de commande :

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

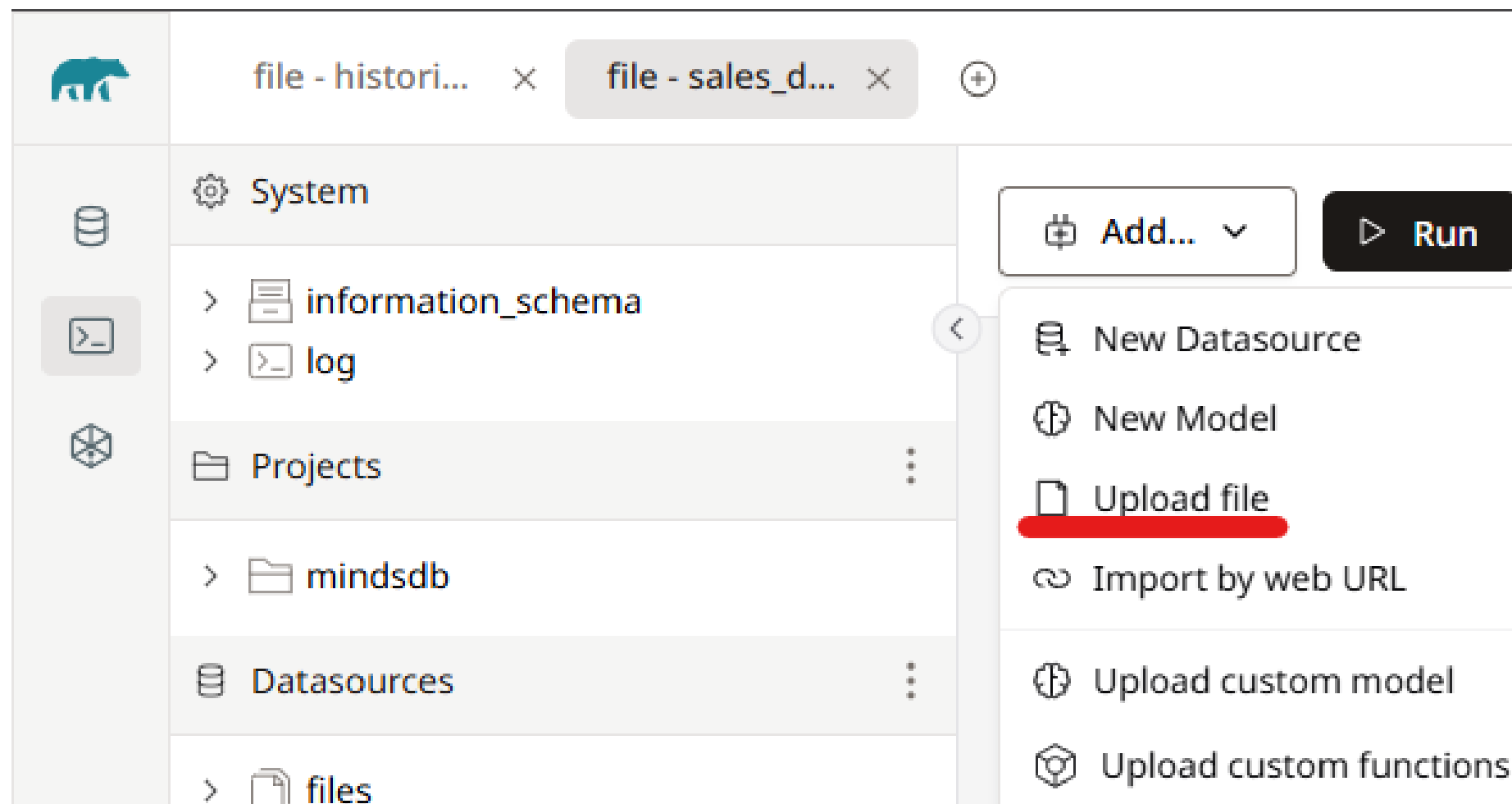
Cela va ouvrir une interface web accessible à l'adresse **<http://localhost:47334>** où l'on peut interagir facilement avec MindsDB.



# Importation des Données

Une fois MindsDB lancé, la première étape est d'importer nos données. Cela se fait très simplement en glissant-déposant un fichier .csv via l'interface.

Chaque fichier importé devient une table dans une base appelée files.



## ← 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 →

# Importation des Données

Après l'importation, MindsDB nous redirige vers un éditeur SQL intégré, ce qui permet d'interroger et de visualiser les données très rapidement.

```
< 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

# Comprendre les Données

Le dataset contient plusieurs colonnes, mais pour notre projet, nous allons nous concentrer sur les plus pertinentes :

- *Order Date* : la date à laquelle le produit a été commandé.
- *Ship Date* : la date d'expédition du produit.
- *Ship Mode* : la méthode de livraison choisie (Standard, Express, etc.).
- *Segment* : le type de client (particulier, entreprise, etc.).
- *Category* : la catégorie du produit (mobilier, technologie, etc.).

Ce sont ces variables que nous allons utiliser pour entraîner notre modèle et prédire le montant des ventes (***Sales***).

# Comprendre les Données

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 ▾

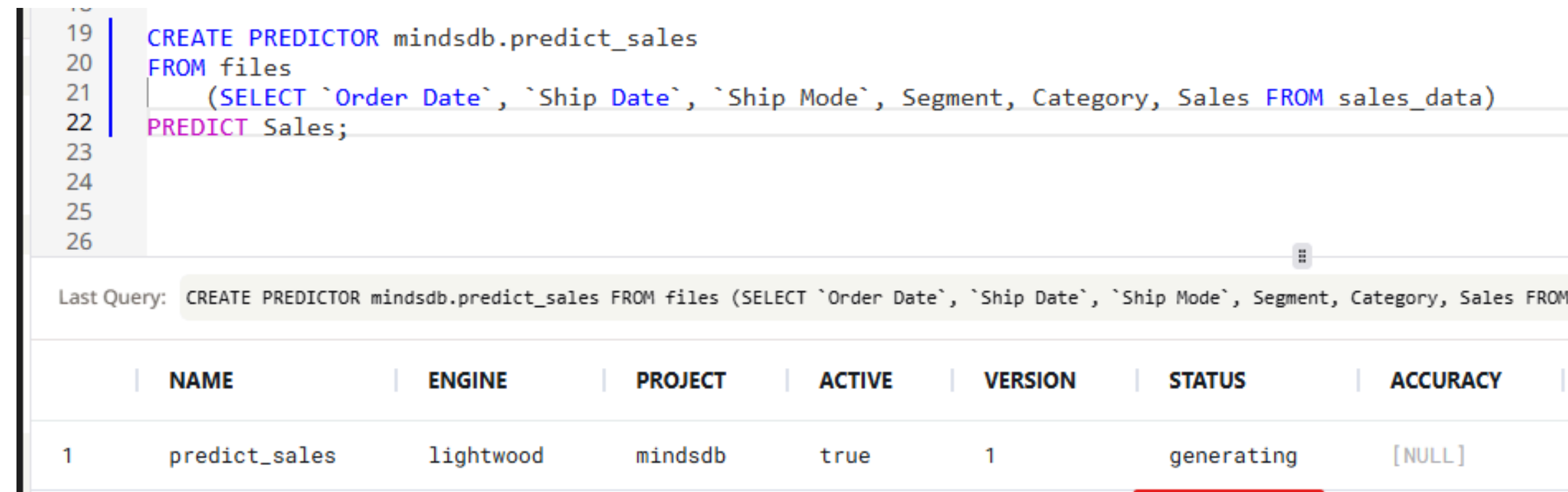
1 to 50 of 9,800

# Entraîner un Modèle de Prédiction

Une fois les données prêtes, on peut maintenant créer notre modèle. Dans MindsDB, cela se fait en écrivant une requête SQL spéciale, comme ceci :

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

Avec cette commande, MindsDB va automatiquement entraîner un modèle de machine learning pour prédire la colonne sales à partir des autres champs. C'est extrêmement simple et rapide.



The screenshot shows the MindsDB interface. At the top, a SQL query is entered in a text area:

```
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
```

Below the query, a status bar indicates the last query executed: "Last Query: CREATE PREDICTOR mindsdb.predict\_sales FROM files (SELECT `Order Date`, `Ship Date`, `Ship Mode`, Segment, Category, Sales FROM sales\_data) PREDICT Sales;".

At the bottom, a table displays the status of the created predictor:

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

# Faire une Prédiction

Maintenant que notre modèle est entraîné, on peut lui poser une question précise, comme dans un cas réel. Prenons un exemple :

Un client passe une commande le 28 mars 2019, qui est expédiée le 31 mars 2019.

Le produit est livré en First Class, pour un client du segment Consumer, et la catégorie est Furniture.

On souhaite savoir : Combien ce produit va-t-il coûter ?

Dans MindsDB, on écrit la requête suivante :

```
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';

# Faire une Prédiction

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

Montant prédit : 292,72 \$

Confiance du modèle : 80 %

Cela signifie que, selon les données historiques et les tendances apprises par le modèle, une commande avec ces caractéristiques a de fortes chances de coûter autour de 292 dollars.

# **Perspectives et Conclusion**



# Points forts de MindsDB


MindsDB présente de nombreux avantages pour les équipes non-expertes en machine learning :

- Intégration naturelle avec SQL : On peut créer et interroger des modèles comme si c'était des tables SQL.
- Prise en main rapide : Même sans connaissance en IA, on peut faire des prédictions en quelques lignes.
- Déploiement simplifié : Disponible via Docker ou le Cloud, sans configuration lourde.
- Capitalisation sur les compétences existantes : Les analystes et data engineers peuvent tirer parti de leurs connaissances en SQL, sans devoir apprendre Python ou R.


# Nouveautés récentes (2024–2025)


1. Interface chat no-code « MindsDB Chat UI » : tu parles à tes données en langage naturel, et l'outil orchestre automatiquement SQL et recherche sémantique.
2. Knowledge Bases intégrées : fusion des données structurées et textes non-structurés (support tickets, documents...) via SQL, avec recherche sémantique optimisée.
3. Modèle Context Protocol (MCP) : enable federation sur plusieurs sources de données, unifiant les accès avec un protocole standard.
4. Mise à jour d'agents et intégrations : gestion avancée des agents (avec support de "writer"), nouveaux connecteurs comme Salesforce, meilleures capacités Docker/A2A.


# Nouveautés récentes (2024–2025)





BETA











Clear Chat

Hello

Hello! I am a Data-Mind, created by MindsDB to help answer data questions. I can access information from the `call_summaries` and `prospects_details` tables. For example, I could answer questions like:

"What are the most common topics discussed in call summaries?"  
"How many prospects are in a specific industry?"  
"What is the average deal size for prospects in a particular region?"  
How can I help you today?

How many prospects are in a specific industry?

There are 257 prospects in the `prospects_details` table. Here's a breakdown of the top 5 use cases and the number of prospects interested in each:

use_case	prospect_count
Evaluating AI solutions for business automation	59
Interested in AI search over my data	54
Interested to check out your solution as it was recommended by my colleague	52
Interested in setting up a customer support chatbot	50
Looking into predictive analytics for customer behavior	42

Ask a question...

Thoughts

Agents

Connected Agents

Add new agent

☐  agent\_mindsdb\_minds\_docs ⋮ ▾

☐  agent\_mindsdb\_minds ⋮ ▾

☐  kb\_mindsdb\_docs\_agent ⋮ ▾

☐  test\_agent ⋮ ▾

☐  crewai\_agent\_multiple\_tables ⋮ ▾

☐  crewai\_agent\_single\_table ⋮ ▾

☐  ai\_agent ⋮ ▾

☐  my\_agent\_kb ⋮ ▾

☐  my\_agent\_1422 ⋮ ▾

☐  my\_agent\_14222 ⋮ ▾

☐  langchain\_agent ⋮ ▾

☐  my\_agent ⋮ ▾

☒  my\_agent\_salesmanager ⋮ ▾

# Conclusion

Pour conclure, MindsDB est un excellent exemple de technologie qui rend l'intelligence artificielle accessible, simple et directement exploitable par les équipes métiers.

Que ce soit pour des prédictions simples ou des cas d'usage plus complexes, il offre une expérience fluide, compatible avec les outils existants, et évolue très rapidement.

C'est une solution qui mérite vraiment d'être explorée dans les environnements Data modernes.

**Thank you  
very much!**