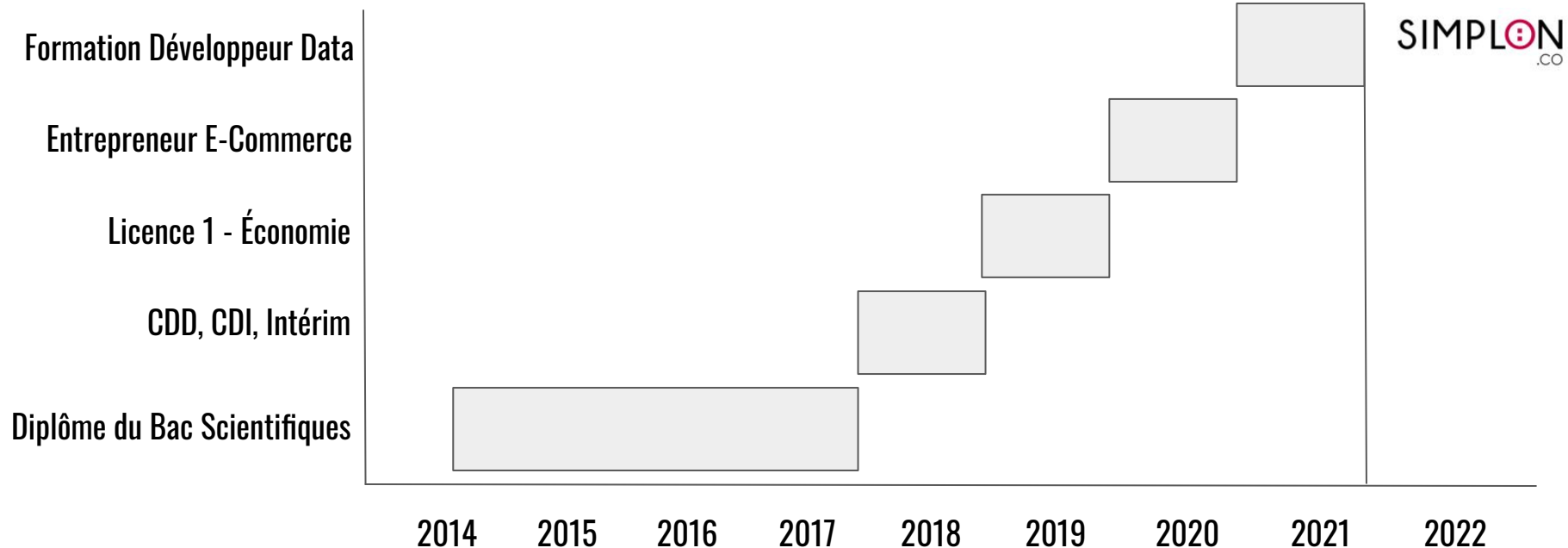




# **BOURSE : COMPRENDRE LES ACTIONS**

JORDAN NAMOR

# PRÉSENTATION



# SIMPLON

- Favoriser les reconversions professionnelles
- Former plus de 13 000 personnes aux métiers du numérique
- Proposer un maximum de débouchés
- Apprendre en France et à l'international

**13 021**

Simplonien.nes dans le monde  
dont 42 % d'apprenantes

**111**

Fabriques dans le monde

**65 %**

De sorties positives après la  
formation

## CONTEXTE

- Le capital d'une entreprise est divisé en plusieurs parts appelées des actions. En bourse, l'évolution du cours des actions est négocié sans devenir propriétaire des actifs.
- Étant un lieu de confrontation entre l'offre et la demande, dès qu'un échange est réalisé, le prix de l'action change : le prix de l'action représente donc le prix pratiqué lors du dernier échange.
- Plus de 6 milliards d'actions s'échangent chaque jour, ce qui correspond à plus de 5000 milliards de dollars en moyenne.

# OBJECTIFS

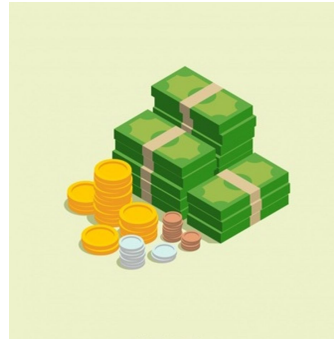
Le but est d'identifier et de comprendre **les différents facteurs** qui **peuvent** influencer sur le prix d'une action.

- Quelles sont les entreprises qui ont les résultats financiers les plus élevés ?
- Ceux-ci peuvent-ils avoir une influence sur sa valorisation boursière ? Le secteur d'activité ?
- Existe-il une corrélation entre ces différents facteurs ?
- Le nombre d'employé représente t-il un facteur important ?

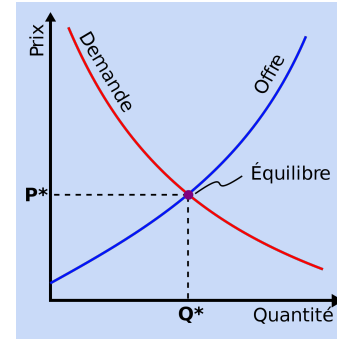
# INDICATEURS



RÉSULTATS  
BOURSIERS



RÉSULTATS  
FINANCIERS



RÉSULTATS  
INVESTISSEURS

# PLANIFICATION - COMPÉTENCES

**Lexiques**

- À faire
- En cours
- Terminé

+ Ajouter une carte

**A1. Développement d'une base de données**

- C1. Identifier la modélisation à effectuer
- C1. Concevoir et structurer physiquement une base de données relationnelle ou non
- C2. Collecter et structurer les données
- C3. Intégrer les données dans la base de données
- C4. Optimiser la base de données

+ Ajouter une carte

**A2. Exploitation d'une base de données**

- C5. Interroger et traiter des données
- C6. Visualiser les données collectées
- C7. Réaliser des visualisations pertinentes

+ Ajouter une carte

**A3. Gestion de projet**

- C8. Analyser et formaliser la demande ou le besoin en développement et en exploitation de base de données.
- C9. Autocontrôler le travail
- C10. Suivre, adapter et rendre compte de la réalisation du projet à partir du planning projet validé
- C10. Effectuer un rapport de 30 pages

+ Ajouter une carte

**Tâches**

- Concaténer l'ensemble des données travaillées pour être en conformité avec les tables
- Créer les tables directement via le fichier Notebook
- Automatiser l'insertion des données en prenant en compte les clés étrangères
- Effectuer la liaison des données via des requêtes SQL
- Nettoyer et organiser le code
- Étendre les axes d'analyse en utilisant le Machine Learning et l'intelligence artificielle

+ Ajouter une carte

# TECHNOLOGIES EXPLOITÉES

## CONCEVOIR

- Python
- MySQL
- Sources de données :
  - Google Finance
  - Wikipedia

## EXPLOITER

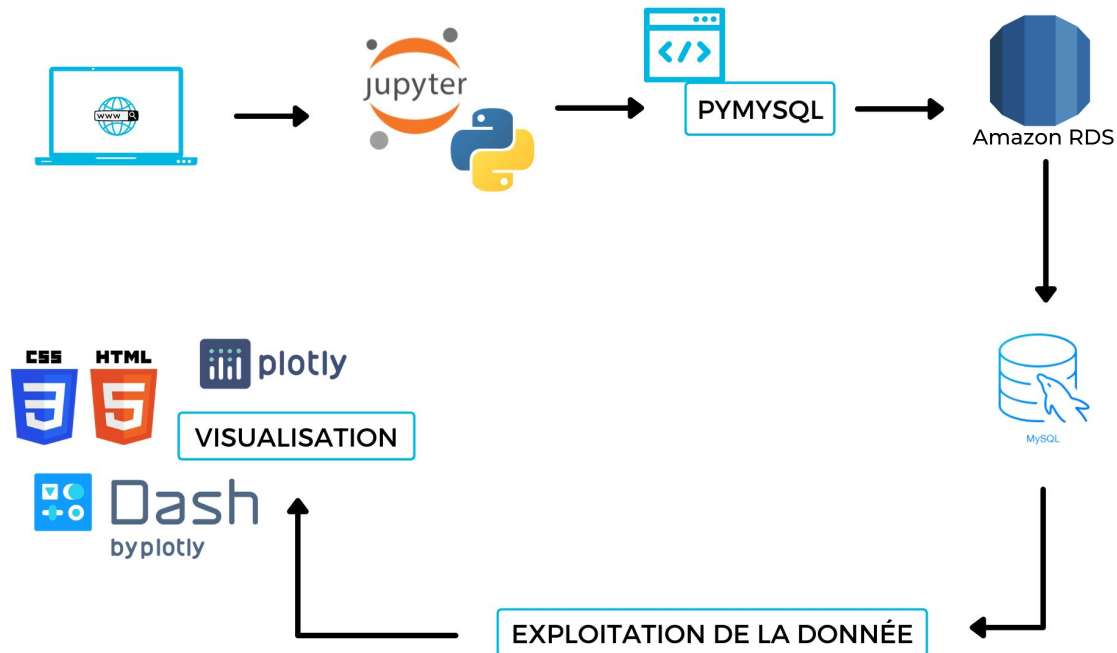
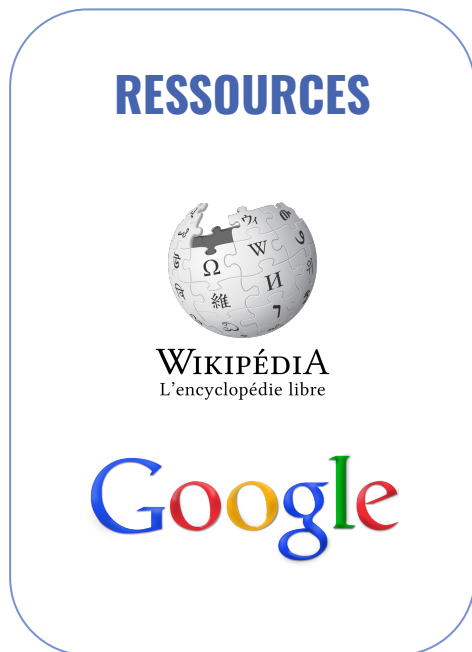
- Interroger la donnée :
  - Python
  - SQL
- IDE :
  - VSCode
  - Jupyter Notebook
- Visuels :
  - Plotly
  - Dash/HTML/CSS

## GÉRER

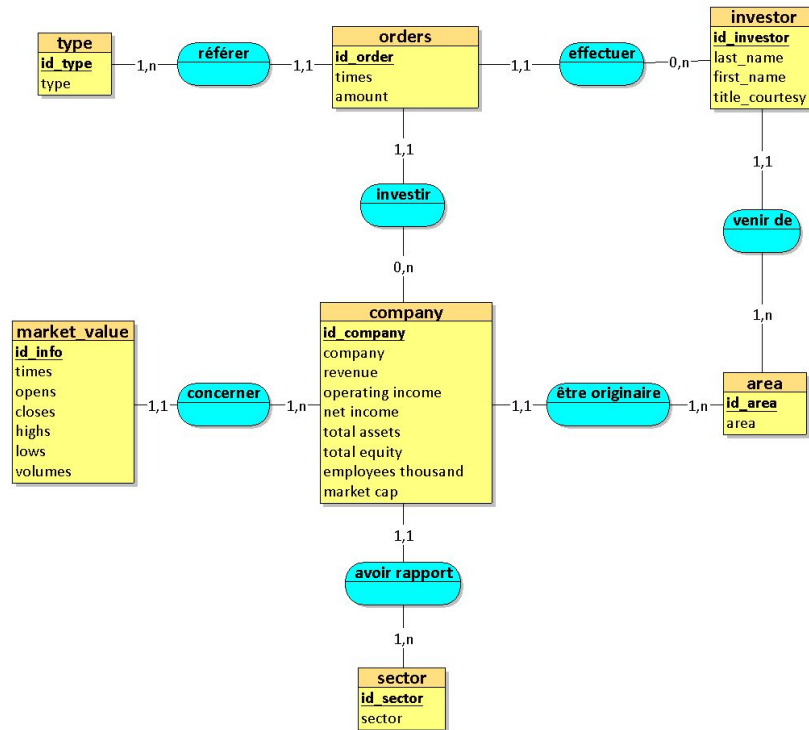
- Trello
- Github/Google Drive
- Stack Overflow
- Documentations :
  - Pandas
  - Dash
  - Plotly



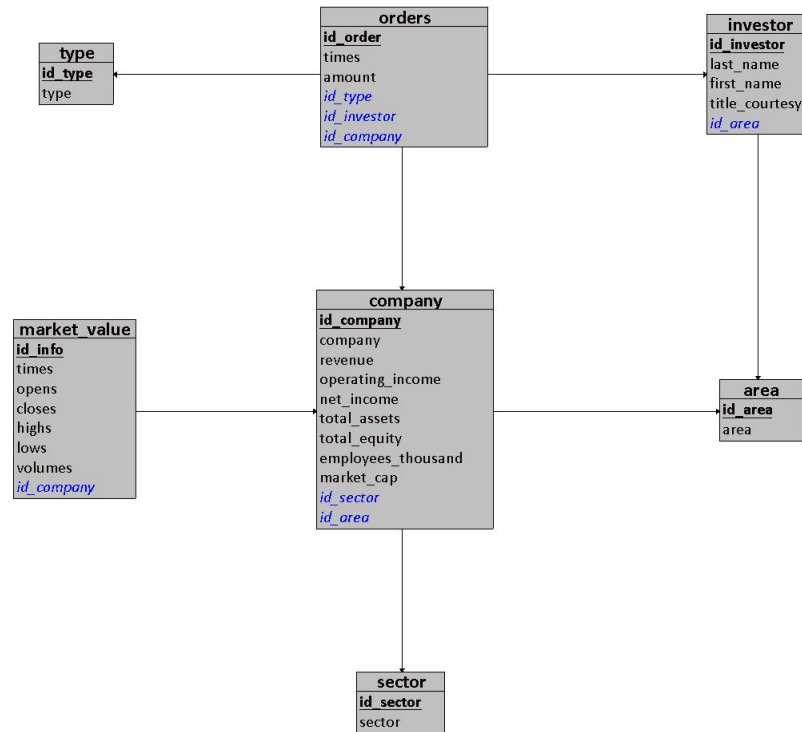
# MODÈLE ORGANISATIONNEL DES DONNÉES



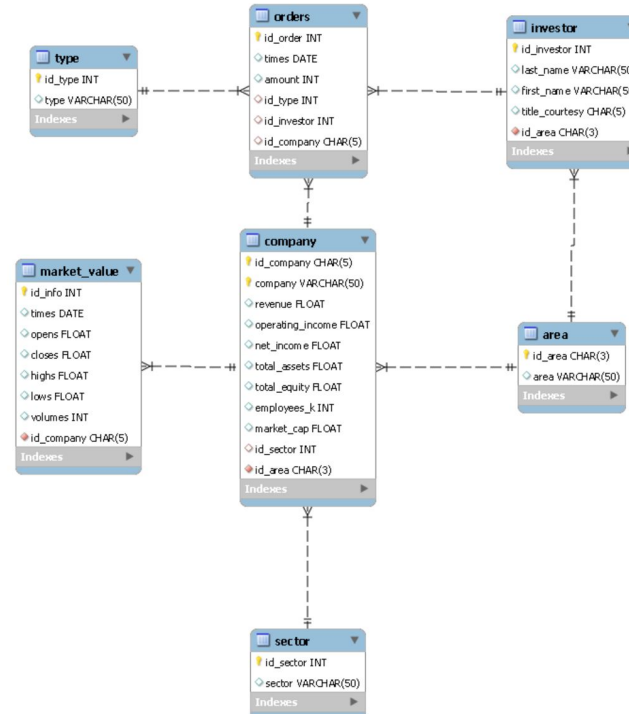
# MODÈLE CONCEPTUEL DES DONNÉES



# MODÈLE RELATIONNEL DES DONNÉES



# MODÈLE PHYSIQUE DES DONNÉES



# DATA PRÉPARATION

## Fichier brut avant traitement

	Company	Link	Revenue	Operating income	Net income	Total assets	Total equity	Number of employees
0	3M	<a href="https://en.wikipedia.org/wiki/3M">https://en.wikipedia.org/wiki/3M</a>	US\$32.18 billion (2020)[4]	US\$7.161 billion (2020)[4]	US\$5.38 billion (2020)[4]	US\$47.3 billion (2020)[4]	US\$12.931 billion (2020)[4]	94,987 (2020)[4]
1	American Express	<a href="https://en.wikipedia.org/wiki/American_Express">https://en.wikipedia.org/wiki/American_Express</a>	US\$36.09 billion (2020)[1]	US\$4.3 billion (2020)[1]	US\$3.14 billion (2020)[1]	US\$191.37 billion (2020)[1]	US\$22.98 billion (2020)[1]	63,700 (December 2020) [1]
2	Amgen	<a href="https://en.wikipedia.org/wiki/Amgen">https://en.wikipedia.org/wiki/Amgen</a>	US\$25.424 billion (2020)	US\$9.674 billion (2019)	US\$7.842 billion (2019)	US\$59.707 billion (2019)	US\$9.673 billion (2019)	22,000 (February 2020)
3	Apple Inc.	<a href="https://en.wikipedia.org/wiki/Apple_Inc.">https://en.wikipedia.org/wiki/Apple_Inc.</a>	US\$274.515 billion[4] (2020)	US\$66.288 billion[4] (2020)	US\$57.411 billion[4] (2020)	US\$323.888 billion[4] (2020)	US\$65.339 billion[4] (2020)	147,000[5] (2020)

## Fichier prêt pour importation

	company	id_company	revenue	operating_income	net_income	total_assets	total_equity	employees_k	id_sector	market_cap	id_area
0	3M	MMM	32.1800	7.161000	5.38000	47.300000	12.931000	94987	14	115.311	USA
1	American Express	AXP	36.0900	4.300000	3.14000	191.370000	22.980000	63700	11	135.173	USA
2	Amgen	AMGN	25.4240	9.674000	7.84200	59.707000	9.673000	22000	13	135.305	USA
3	Apple Inc.	AAPL	274.5150	66.288000	57.41100	323.888000	65.339000	147000	17	2429.768	USA

# CRÉATION DES TABLES

```
# Connection à l'instance de la database
def connect_to_instance_db():
    try:
        conn = pymysql.connect(
            host='db-company.cwm601wtiqo3.eu-west-3.rds.amazonaws.com',
            user='admin',
            password=ACCESS_DB,
            charset='utf8mb4')
    except pymysql.err.OperationalError as e:
        raise e
    else:
        print("Connection Successful!")
    return conn

connection = connect_to_instance_db()
cursor = connection.cursor()

# Création de la database
def create_db(cursor):
    queries = (
        "DROP DATABASE IF EXISTS db_company",
        "CREATE DATABASE db_company",
        "USE db_company"
    )

    for query in queries:
        cursor.execute(query)
        connection.commit()
```

```
# Création des tables
def create_table(cursor):
    queries = (
        "DROP TABLE IF EXISTS type",
        '''CREATE TABLE type(
            id_type INT,
            type VARCHAR(50),
            PRIMARY KEY(id_type))''',

        "DROP TABLE IF EXISTS sector",
        '''CREATE TABLE sector(
            id_sector INT,
            sector VARCHAR(50),
            PRIMARY KEY(id_sector))''',

        "DROP TABLE IF EXISTS orders",
        '''CREATE TABLE orders(
            id_order INT,
            times DATE,
            amount INT,
            capital INT,
            id_type INT,
            id_investor INT,
            id_company CHAR(5),
            PRIMARY KEY(id_order),
            FOREIGN KEY(id_type) REFERENCES type(id_type),
            FOREIGN KEY(id_investor) REFERENCES investor(id_investor),
            FOREIGN KEY(id_company) REFERENCES company(id_company))'''
    )

    for query in queries:
        cursor.execute(query)
        connection.commit()

create_db(cursor)
create_table(cursor)
```

# INSERTION DES DONNÉES LIGNE PAR LIGNE

```
# Importation des données
df_type = pd.read_csv('data/type.csv')
df_sector = pd.read_csv('data/sector.csv')
df_area = pd.read_csv('data/area.csv')
df_investor = pd.read_csv('data/investor.csv')
df_company = pd.read_csv('data/company.csv')
df_company['employees_k'] = (df_company['employees_k'] * 0.001)
df_market_value_short_2 = pd.read_csv('data/market_value_short_2.csv')
df_orders = pd.read_csv('data/orders.csv')
```

```
# Insertion du DataFrame ligne par ligne
def insert_into_table(data, table):
    # Création d'une liste de colonnes pour l'insertion
    cols = ','.join(str(x) for x in data.columns)

    for i, row in data.iterrows():
        query = f"INSERT IGNORE INTO {table}({cols}) VALUES(" + "%s,"*(len(row)-1) + "%s)"
        cursor.execute(query, tuple(row))
        print(i)

    connection.commit()
```



```
insert_into_table(df_type, 'type')
insert_into_table(df_sector, 'sector')
insert_into_table(df_area, 'area')
insert_into_table(df_investor, 'investor')
insert_into_table(df_company, 'company')
insert_into_table(df_orders, 'orders')
insert_into_table(df_market_value_short_2, 'market_value')
```

# SAUVEGARDE JOURNALIÈRE DE LA BDD

The screenshot displays the AWS Management Console interface for Amazon RDS. The left-hand navigation pane lists various RDS features, with 'Automated backups' highlighted in orange. The main content area is divided into two sections. The top section provides details for a specific database instance, including its storage size (20 GB), IOPS, and time zone. The bottom section, titled 'System snapshots (8)', contains a table listing all automated backups for the instance. Each row in the table shows the snapshot name, the database source, the creation time, the status (all are 'available'), the progress (all are 'Completed'), and the VPC.

**Amazon RDS**

- Dashboard
- Databases
- Query Editor
- Performance Insights
- Snapshots
- Automated backups**
- Reserved instances
- Subnet groups
- Parameter groups
- Option groups
- Events
- Event subscriptions
- Recommendations 1
- Certificate update

Storage: 20  
IOPS: -  
Timezone: -  
Amazon Resource Name (ARN): arn:aws:rds:eu-west-3:431139978922:db:db-company

**System snapshots (8)**

Filter system snapshots

Name	DB source	Creation time	Status	Progress	VPC
<a href="#">rds:db-company-2021-08-11-10-02</a>	db-company	Wed Aug 11 2021 11:02:31 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-12-10-02</a>	db-company	Thu Aug 12 2021 11:02:55 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-13-10-02</a>	db-company	Fri Aug 13 2021 11:02:42 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-14-10-02</a>	db-company	Sat Aug 14 2021 11:02:31 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-15-10-02</a>	db-company	Sun Aug 15 2021 11:02:32 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-16-10-02</a>	db-company	Mon Aug 16 2021 11:02:47 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-17-10-02</a>	db-company	Tue Aug 17 2021 11:02:19 GMT+0100	available	Completed	vpc-d6d31dbe
<a href="#">rds:db-company-2021-08-18-10-02</a>	db-company	Wed Aug 18 2021 11:02:53 GMT+0100	available	Completed	vpc-d6d31dbe



# SÉCURISATION - CRÉATION D'UN UTILISATEUR LIMITÉ

```
Anaconda Prompt (anaconda3) - mysql -h db-company.cwm601wtiqo3.eu-west-3.rds.amazonaws.com -u admin -p

(base) C:\Users\namor>mysql -h db-company.cwm601wtiqo3.eu-west-3.rds.amazonaws.com -u admin -p
Enter password: *****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 3281
Server version: 8.0.23 Source distribution

Copyright (c) 2000, 2021, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> CREATE USER 'jordan_new'@ '%' IDENTIFIED BY 'welcome123' ;
Query OK, 0 rows affected (0.21 sec)

mysql> GRANT SELECT ON *.* TO 'jordan_new'@ '%' WITH GRANT OPTION ;
Query OK, 0 rows affected (0.31 sec)

mysql> SHOW GRANTS FOR 'jordan_new' ;
+-----+
| Grants for jordan_new@% |
+-----+
| GRANT SELECT ON *.* TO `jordan_new`@`%` WITH GRANT OPTION |
+-----+
1 row in set (0.20 sec)

mysql>
```

## REQUÊTE : EXEMPLE 1

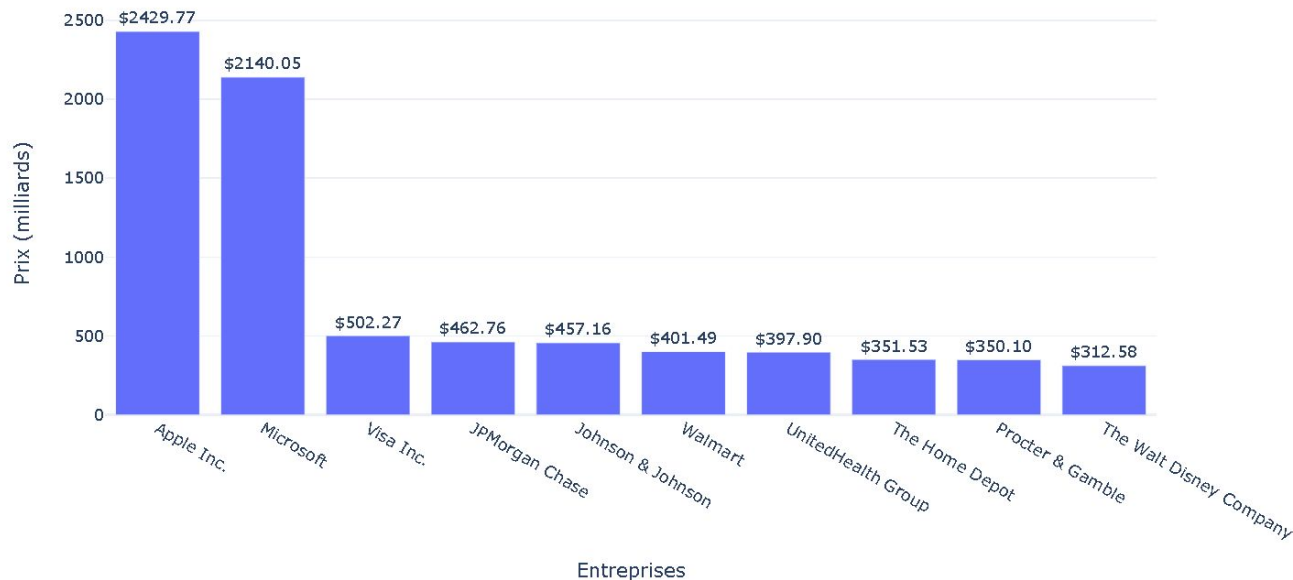
```
SELECT company, market_cap  
FROM company  
ORDER BY market_cap DESC  
LIMIT 10;
```



	company	market_cap
▶	Apple Inc.	2429.77
	Microsoft	2140.05
	Visa Inc.	502.269
	JPMorgan Chase	462.757
	Johnson & Johnson	457.157
	Walmart	401.491
	UnitedHealth Group	397.903
	The Home Depot	351.534
	Procter & Gamble	350.097
	The Walt Disney Company	312.585

# REQUÊTE : EXEMPLE 1 - RÉSULTAT FINAL

Top 10 des entreprises ayant les capitalisations boursières les plus élevées



## REQUÊTE : EXEMPLE 2

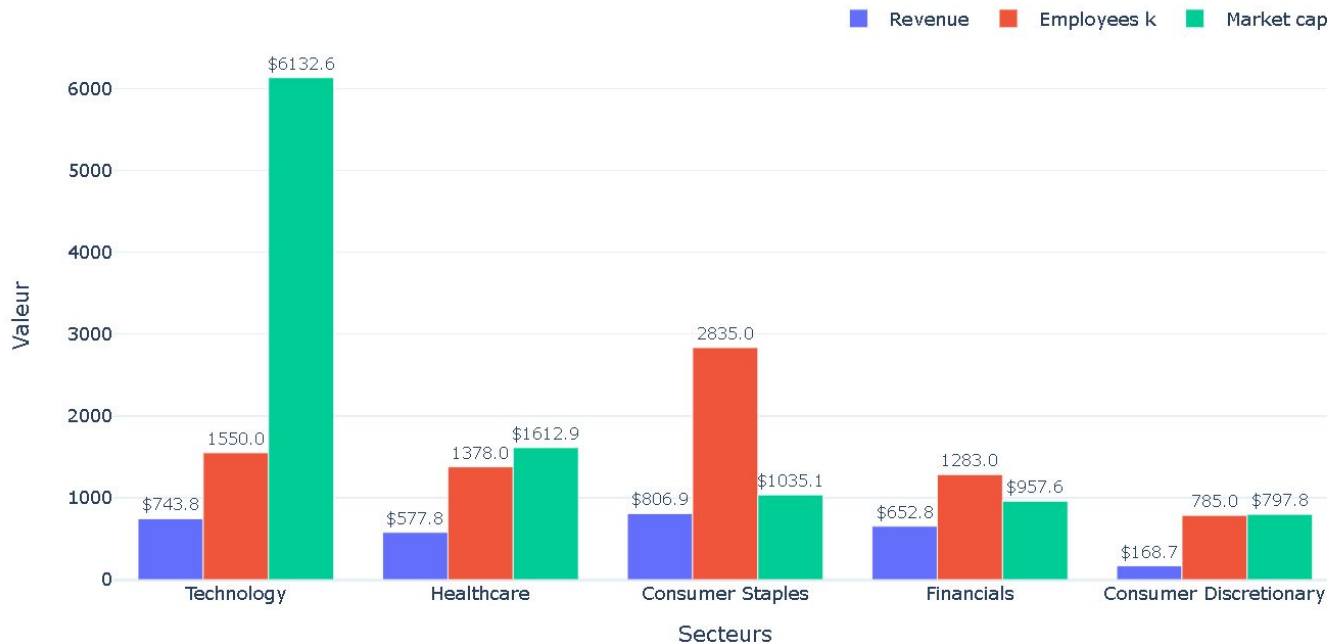
```
SELECT SUM(revenue) as revenue, SUM(employees_k) as employees_k, SUM(market_cap) as market_cap, C.id_sector, S.sector
FROM company C
      LEFT JOIN sector S ON C.id_sector = S.id_sector
GROUP BY id_sector
ORDER BY market_cap DESC
LIMIT 5;
```



	revenue	employees_k	market_cap	id_sector	sector
►	743.7720165252686	1550	6132.575043005403	17	Technology
	577.7749853134155	1378	1612.9180251802318	13	Healthcare
	806.9200096130371	2835	1035.137981414795	8	Consumer Staples
	652.7910034656525	1283	957.6419948041439	11	Financials
	168.67599868774414	785	797.7669982910156	6	Consumer Discretionary

## REQUÊTE : EXEMPLE 2 - RÉSULTAT FINAL

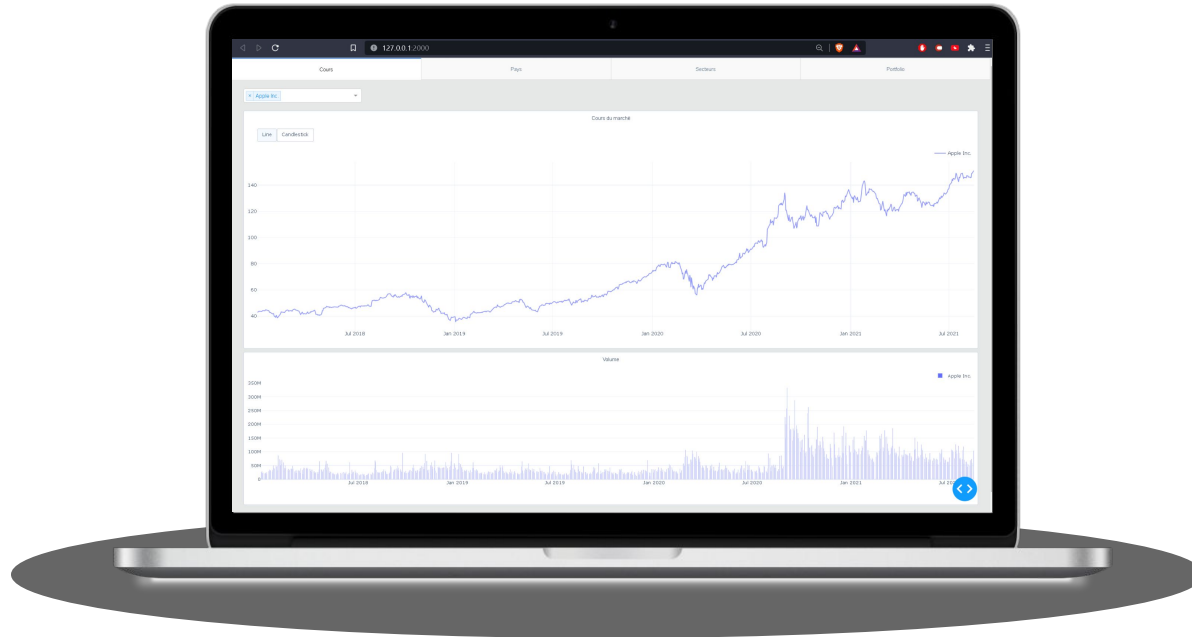
Résultats financiers des 5 meilleurs secteurs



# DÉVELOPPEMENT D'UNE APPLICATION WEB AVEC DASH

```
app.py x # style.css
Database > apps > app.py > ...
683         )
684     ]
685 )
686
687
688 app.layout = html.Div(
689     className="row",
690     children=[
691         html.Div(
692             className="row tabs_div",
693             children=[
694                 dcc.Tabs(
695                     id="tabs",
696                     style={"height": "20", "verticalAlign": "middle"},
697                     children=[
698                         dcc.Tab(label="Cours", value="trading_tab"),
699                         dcc.Tab(label="Pays", value="entreprise_tab"),
700                         dcc.Tab(label="Secteurs", value="entreprise_tab2"),
701                         dcc.Tab(label="Portfolio", value="portfolio_tab")
702                     ],
703                     value="trading_tab"
704                 )
705             ],
706         ),
707         html.Div(id="tab_content", className="row"),
708     ]
709 )
710
711
712 @app.callback(
```

# DÉMO



# CONCLUSION

L'action d'une entreprise peut être influencée par :

- ses résultats financiers (chiffre d'affaires, sa rentabilité, les anticipations sur ses bénéfices, ses actifs, ...)
- l'appréhension du secteur dans lequel elle évolue
- le pays d'origine



## **DIFFICULTÉS RENCONTRÉES**

- **Collecter et nettoyer la donnée**
- **Apprendre des technologies nouvelles (HTML, CSS, Dash, Plotly)**
- **Créer un tableau de bord en ligne et le dynamiser**

## **ALLER PLUS LOIN ...**

- **Actualiser et automatiser la donnée**
- **Utiliser des méthodes en machine learning et en intelligence artificielle pour analyser les différents actifs financiers**
- **Récupérer les différents articles qui parlent de sujets financiers et effectuer de la NLP**

**UN GRAND MERCI POUR VOTRE ATTENTION !**

