

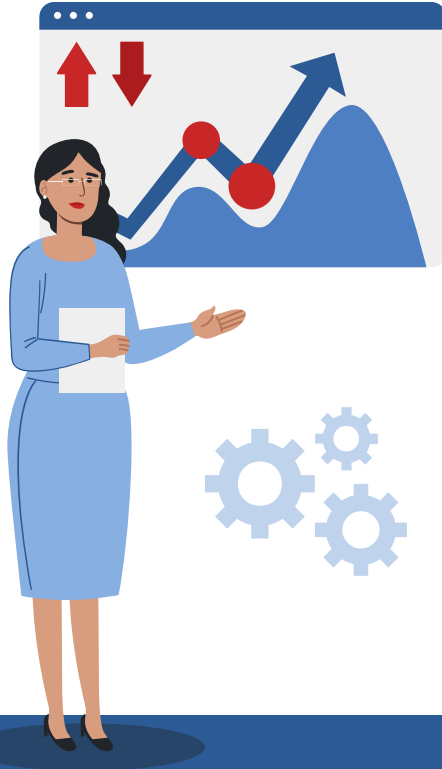
NSE STOCK MARKET ANALYSIES



Projet PI classe **1BA3**

PLAN

1. Introduction et contexte
2. Business understanding
3. Méthode de gestion et de planification du projet
4. Analyse et spécification des besoins
5. Etude conceptuelle
6. Integration Talend
7. Tableau de bord POWER BI
8. Machine learning : analyse predictive





01

INTRODUCTION

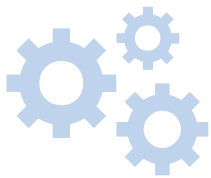
INTRODUCTION

Notre projet se concentre sur l'exploitation de la base de données NYSE, une source d'informations financières complète sur le marché boursier.

Notre objectif est d'extraire des informations exploitables en analysant les tendances, les secteurs et les entreprises à partir de cette base de données.

En utilisant les capacités avancées de la Business Intelligence, nous convertirons ces données brutes en connaissances stratégiques.





02

Business understanding



Contexte et cadre general

Notre projet de Business Intelligence se concentre sur l'exploitation de la base de données NYSE afin de fournir des informations exploitables et pertinentes pour les décideurs financiers, améliorant ainsi leur compréhension et leur prise de décision dans le domaine boursier.



Critique de l'existant

NYSE offre des informations complètes sur les entreprises cotées en bourse, avec des avantages tels que la facilité de manipulation et la mise à jour régulière. Elle présente des limitations telles que l'absence de données sur toutes les entreprises, des erreurs potentielles et l'absence de normalisation des données.

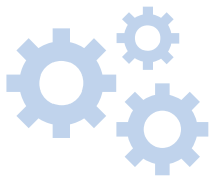
Problématique

Quels secteurs d'activité ont connu la croissance la plus rapide au cours des cinq années indiquées et quels sont les facteurs clés de cette croissance, en utilisant la base de données NYSE ?



Solutions proposées

Nous proposons d'explorer les données financières des entreprises dans chaque secteur, d'examiner les tendances économiques, utilisons des solutions d'analyse et de visualisation via des tableaux de bord, fournir une modélisation prédictive pour identifier les facteurs de croissance.

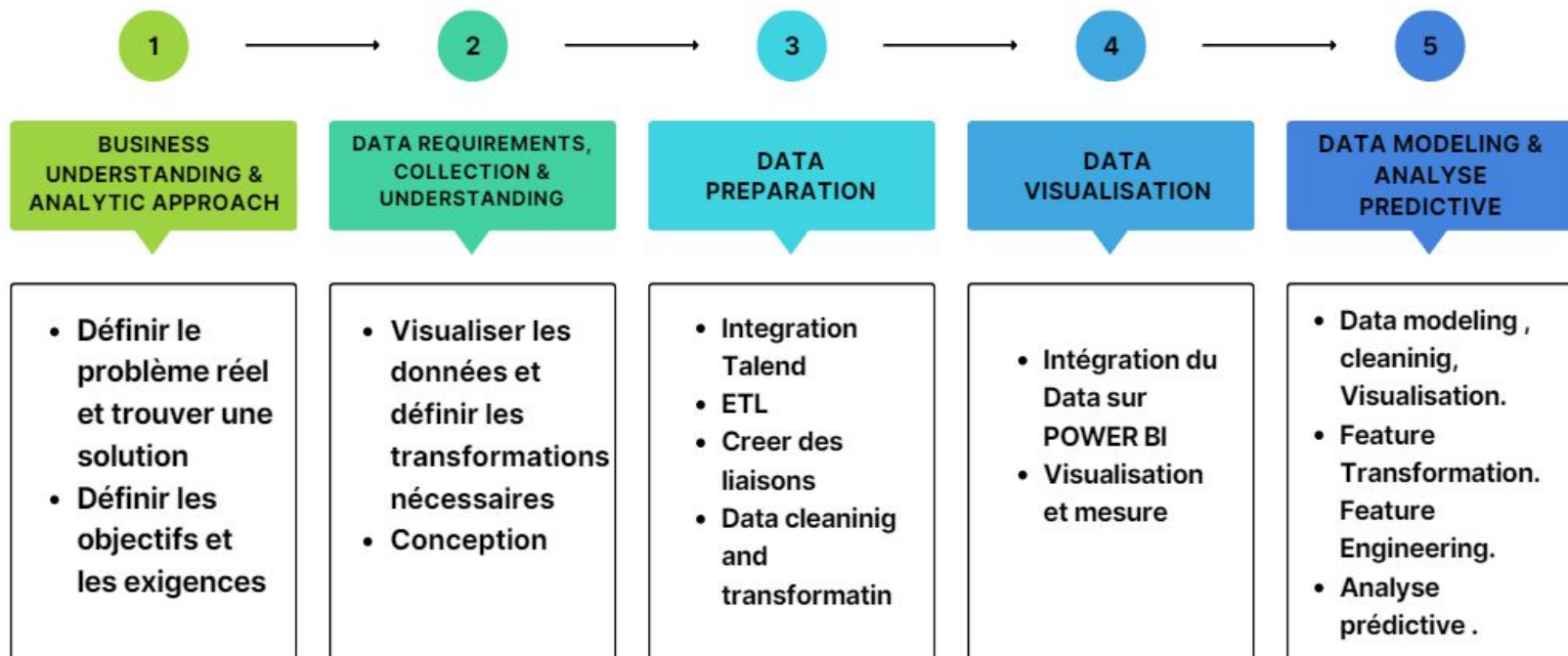


03

Méthode de gestion et de planification du projet



Project Pipeline



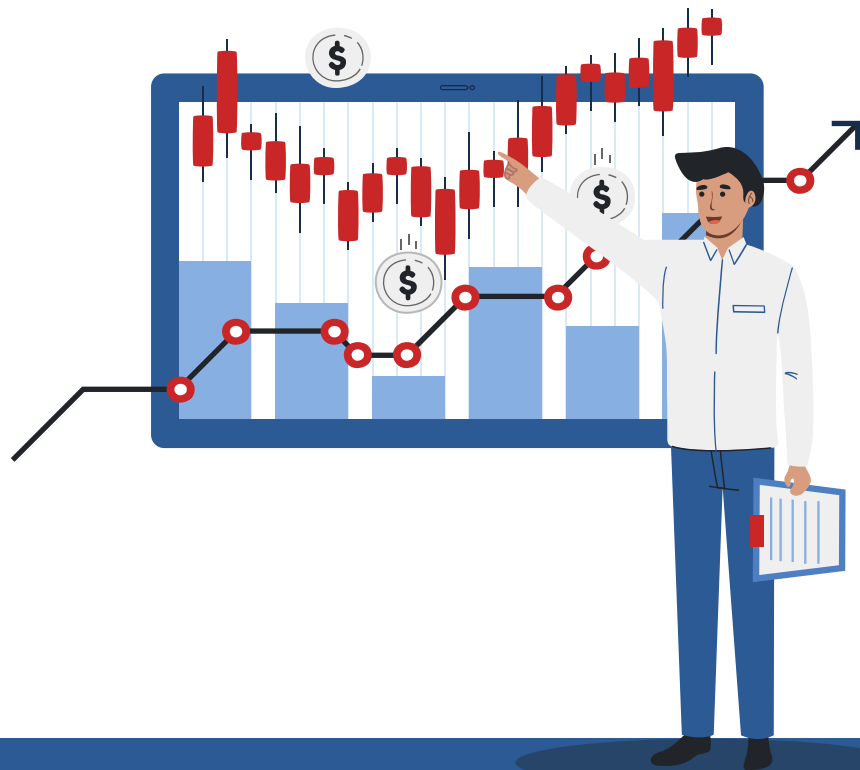
	TASK	Month 1				Month 2				Month 3			
		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12
1	Collecter et nettoyer les données												
1.1	Identifier les sources de données												
1.1.1	Collecter les données												
1.2	Nettoyer les données												
1.3	Vérifier la qualité des données												
2	Analyser les tendances du marché												
2.1	Analyser les tendances du marché												
2.2	Segmenter les données												
2.3	Identifier les entreprises sous/surévaluées												
3	Analyser les données fondamentales des entreprises												
3.1	Collecter les données fondamentales												
3.2	Analyser les ratios financiers												
4	Prévoir les prix futurs des actions												
4.1	Entraîner un modèle de prévision												
4.2	Tester le modèle												
4.3	Optimiser le modèle												





04

Analyse et spécification des besoins



Besoins fonctionnels

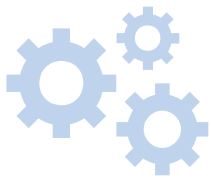
- **Intégration des données**
- **Manipulation des données**
- **Dashboarding**
- **Analyse et prediction**



Besoins non-fonctionnels

- La fiabilité
- La facilité d'utilisation
- L'efficacité
- La maintenabilité



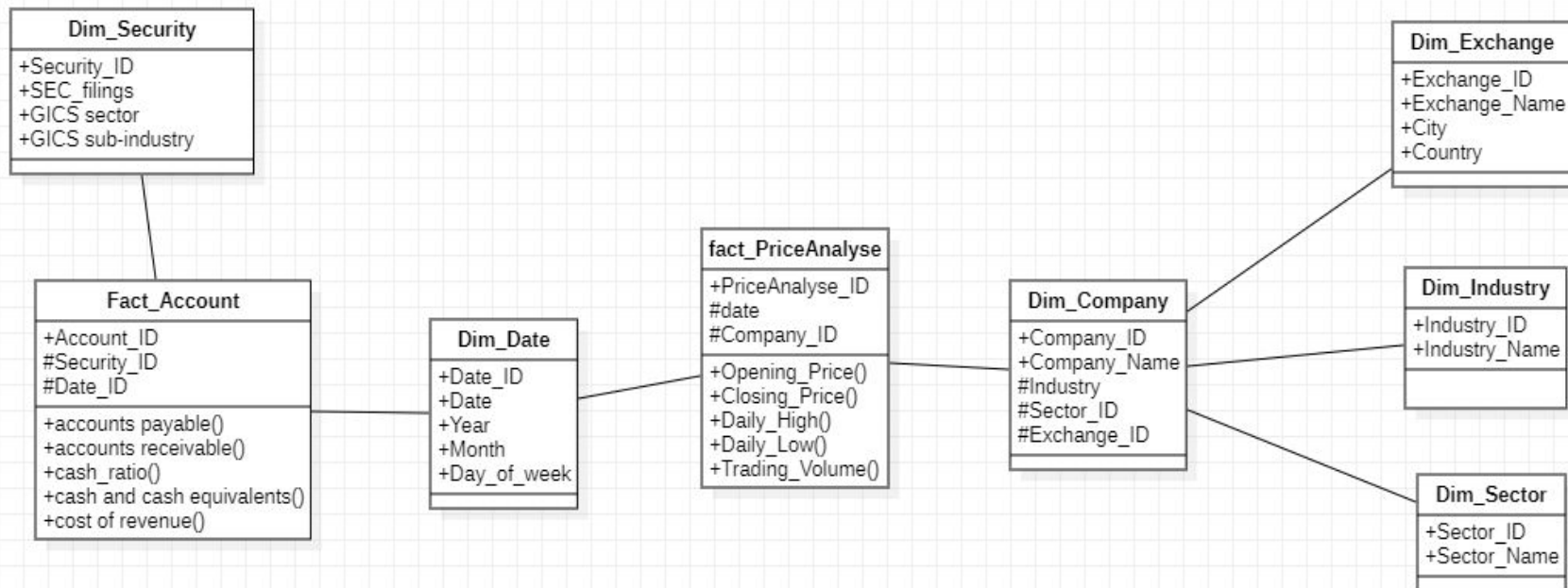


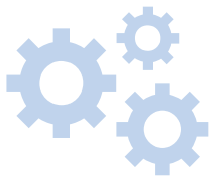
05

Etude conceptuelle



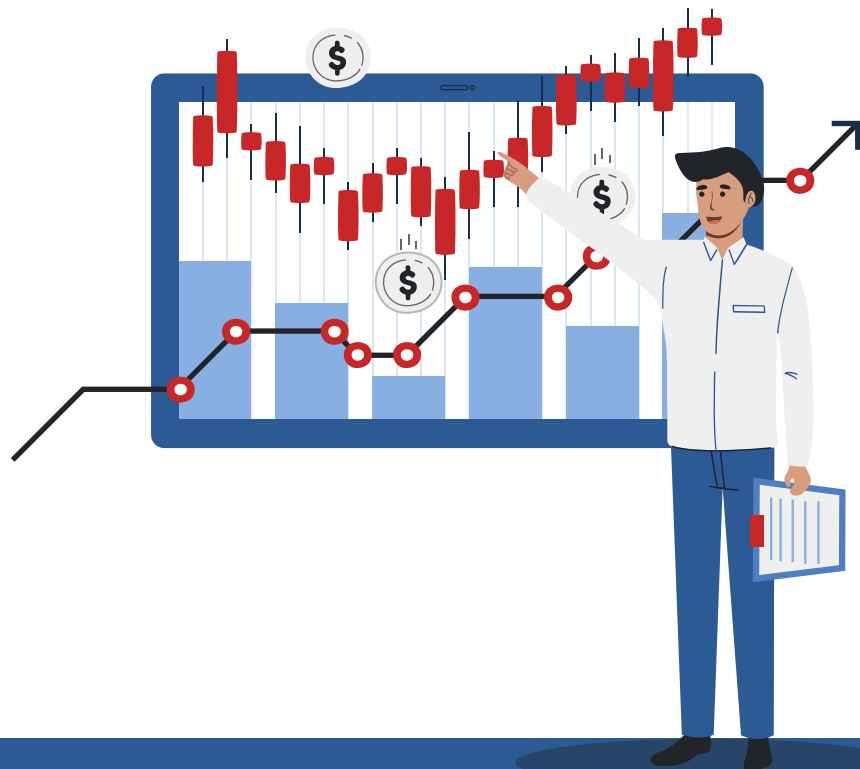
Etude conceptuelle





06

Integration Talent

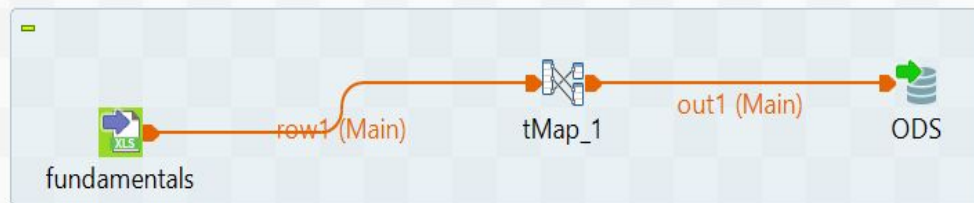


LOCAL: PI

- Jobs
 - Dimensions
 - ODS
 - ODS_fundamentals 0.1
 - ODS_Prices 0.1
 - ODS_Securities 0.1
 - Tables_Fait
- Contextes
- Code
- Modèles SQL
- Métadonnées
- Documentation
- Corbeille

Job ODS (OPERATIONAL DATA STORE) ODS fundamentals

Job ODS_fundamentals 0.1 x



jointure t-map

row1		Var		out1	
Column				Expression	Column
ID				row1.ID	ID
Ticker_Symbol				row1.Ticker_Symbol	Ticker_Symbol
Period_Ending				row1.Period_Ending	Period_Ending
Accounts_Payable				row1.Accounts_Payable	Accounts_Payable
Accounts_Receivable				row1.Accounts_Receivable	Accounts_Receivable
Add_I_income_expense_items				row1.Add_I_income_expense_items	Add_I_income_expense...
After_Tax_ROE				row1.After_Tax_ROE	After_Tax_ROE
Capital_Expenditures				row1.Capital_Expenditures	Capital_Expenditures
Capital_Surplus				row1.Capital_Surplus	Capital_Surplus
Cash_Ratio				row1.Cash_Ratio	Cash_Ratio
Cash_and_Cash_Equivalents				row1.Cash_and_Cash_Equivalents	Cash_and_Cash_Equiva...
Changes_in_Inventories				row1.Changes_in_Inventories	Changes_in_Inventories
Common_Stocks				row1.Common_Stocks	Common_Stocks
Cost_of_Revenue				row1.Cost_of_Revenue	Cost_of_Revenue
Current_Ratio				row1.Current_Ratio	Current_Ratio
Deferred_Asset_Charges				row1.Deferred_Asset_Charges	Deferred_Asset_Charges
Deferred_Liability_Charges				row1.Deferred_Liability_Charges	Deferred_Liability_Char...
Depreciation				row1.Depreciation	Depreciation
Earnings_Before_Interest_and_Tax				row1.Earnings_Before_Interest_and_Tax	Earnings_Before_Intere...
Earnings_Before_Tax				row1.Earnings_Before_Tax	Earnings_Before_Tax
Effect_of_Exchange_Rate				row1.Effect_of_Exchange_Rate	Effect_of_Exchange_Rate
Equity_Earnings_Loss_Unconsolidated_Subsiary				row1.Equity_Earnings_Loss_Unconsolid...	Equity_Earnings_Loss_...
Fixed_Assets				row1.Fixed_Assets	Fixed_Assets
Goodwill				row1.Goodwill	Goodwill
Gross_Margin				row1.Gross_Margin	Gross_Margin
Gross_Profit				row1.Gross_Profit	Gross_Profit
Income_Tax				row1.Income_Tax	Income_Tax
Intangible_Assets					

INPUT

OUTPUT

Éditeur de schéma Éditeur d'expression

row1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
Other_Current_Assets	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Other_Current_Liabilities	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		10	0		
Other_Equity	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Other_Financing_Activi...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Other_Investing_Activit...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Other_Liabilities	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Other_Operating_Activi...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Other_Operating_Items	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Pre_Tax_Margin	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Pre_Tax_ROE	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Profit_Margin	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Quick_Ratio	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Research_and_Develop...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Retained_Earnings	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Sale_and_Purchase_of_...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Sales_General_and_Ad...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Short_Term_Debt_Cur...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Short_Term_Investments	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Total_Assets	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Total_Current_Assets	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Total_Current_Liabilities	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Total_Equity	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Total_Liabilities	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Total_Liabilities__Equity	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Total_Revenue	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Treasury_Stock	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		

out1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
ID	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
Ticker_Symbol	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		10	0		
Period_Ending	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	"dd-MM-yyyy"		0		
Accounts_Payable	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Accounts_Receivable	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Add_I_income_expense...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
After_Tax_ROE	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Capital_Expenditures	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Capital_Surplus	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Cash_Ratio	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Cash_and_Cash_Equival...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Changes_in_Inventories	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Common_Stocks	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Cost_of_Revenue	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		12	0		
Current_Ratio	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Deferred_Asset_Charges	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		10	0		
Deferred_Liability_Char...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Depreciation	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Earnings_Before_Intere...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Earnings_Before_Tax	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Effect_of_Exchange_Rate	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Equity_Earnings_Loss_...	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Fixed_Assets	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Goodwill	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		
Gross_Margin	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>			0		
Gross_Profit	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		11	0		

ODS-fundamentales

phpMyAdmin

Récentes Préférences

dim_pi

faits

information_schema

mysql

ods_pi

Nouvelle table

ods_fundamentals

Colonnes

1 > >>

Nouvelle colonne

Accounts_Payable

Accounts_Receivable

Add_I_income_expense_items

After_Tax_ROE (bigint(20))

Capital_Expenditure

Capital_Surplus (bigint(11))

Cash_and_Cash_Equivalents

Cash_Ratio (bigint(20))

Changes_in_Inventories

Common_Stocks (bigint(11))

Cost_of_Revenue (bigint(12))

Server : 127.0.0.1 » Base de données : ods_pi » Table : ods_fundamentals

Parcourir Structure SQL Rechercher Insérer Exporter Importer Privilèges Opérations Suivi

#	Nom	Type	Interclassement	Attributs	Null	Valeur par défaut	Commentaires	Extra	Action
<input type="checkbox"/>	1 ID	int(2)			Non	Aucun(e)			Modifier
<input type="checkbox"/>	2 Ticker_Symbol	varchar(10)	latin1_swedish_ci		Oui	NULL			Modifier
<input type="checkbox"/>	3 Period_Ending	date			Oui	NULL			Modifier
<input type="checkbox"/>	4 Accounts_Payable	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	5 Accounts_Receivable	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	6 Add_I_income_expense_items	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	7 After_Tax_ROE	bigint(20)			Oui	NULL			Modifier
<input type="checkbox"/>	8 Capital_Expenditures	bigint(12)			Oui	NULL			Modifier
<input type="checkbox"/>	9 Capital_Surplus	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	10 Cash_Ratio	bigint(20)			Oui	NULL			Modifier
<input type="checkbox"/>	11 Cash_and_Cash_Equivalents	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	12 Changes_in_Inventories	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	13 Common_Stocks	bigint(11)			Oui	NULL			Modifier
<input type="checkbox"/>	14 Cost_of_Revenue	bigint(12)			Oui	NULL			Modifier
<input type="checkbox"/>	15 Current_Ratio	bigint(20)			Oui	NULL			Modifier
<input type="checkbox"/>	16 Deferred_Asset_Charges	bigint(10)			Oui	NULL			Modifier
<input type="checkbox"/>	17 Deferred_Liability_Charges	bigint(11)			Oui	NULL			Modifier

Console de requêtes SQL

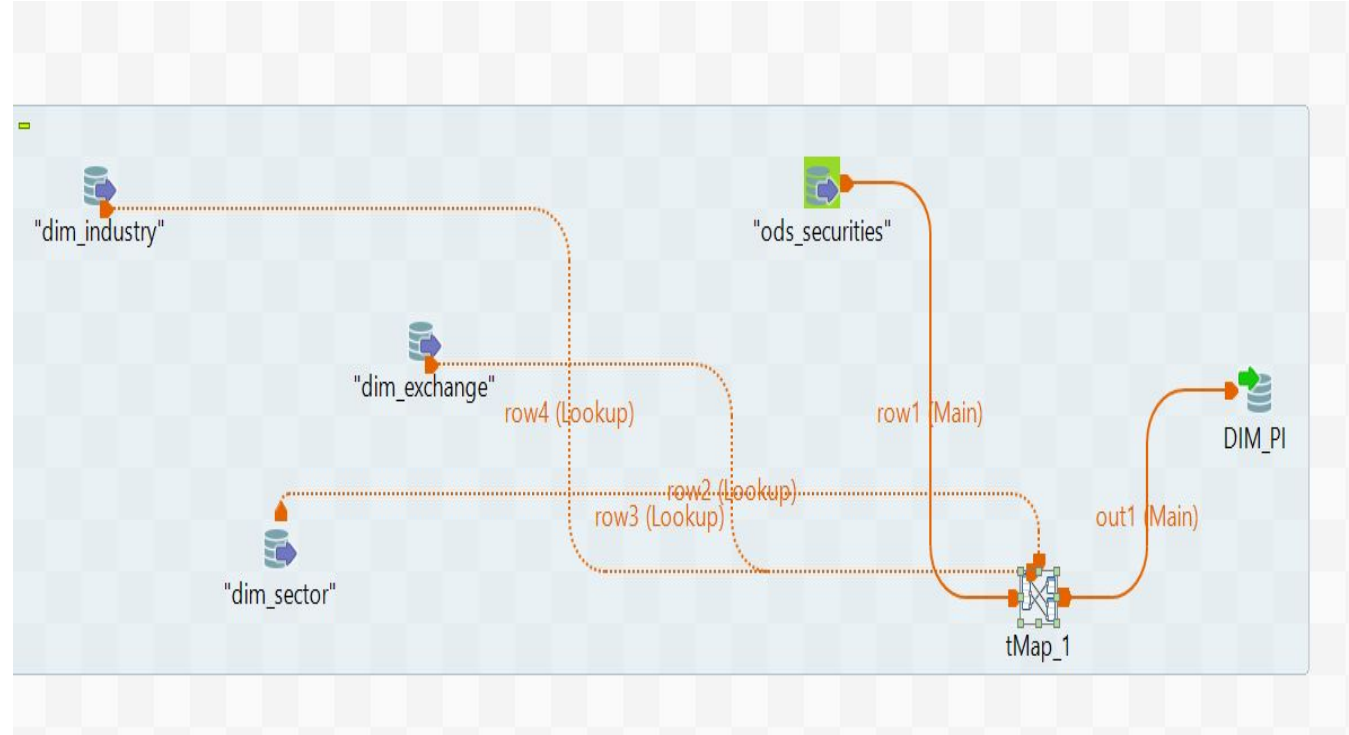
JOBS DW (Data warehouse)

Dimension_Company

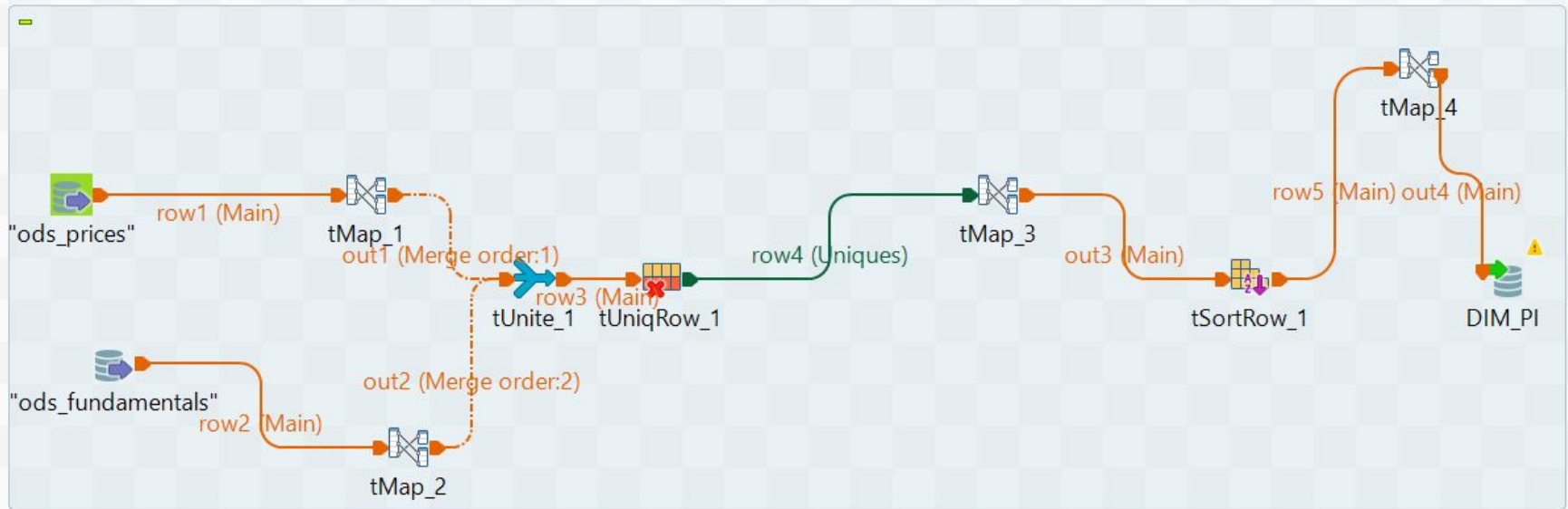
Référentiel x

LOCAL: PI

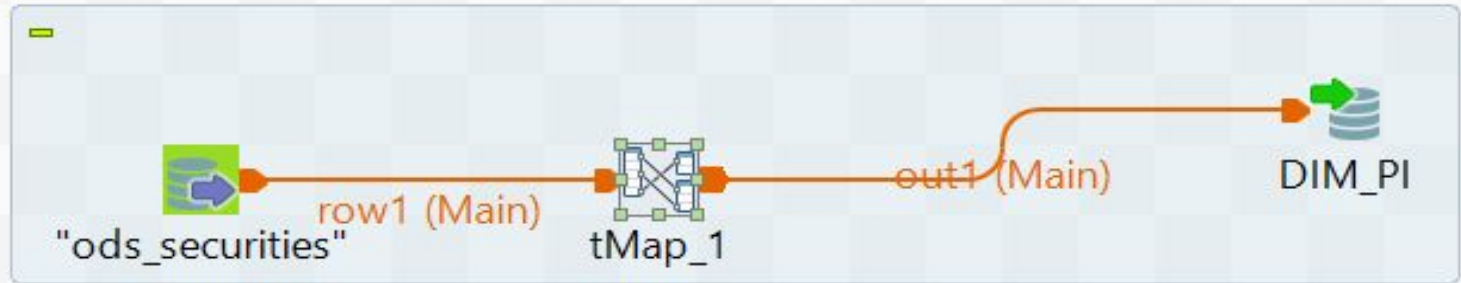
- Jobs
 - Dimensions
 - Dimension_Company 0.1
 - Dimension_Date 0.1
 - Dimension_Exchange 0.1
 - Dimension_Industry 0.1
 - Dimension_Sector 0.1
 - Dimension_Security 0.1
 - ODS
 - Tables_Fait
 - Contextes
 - Code
 - Modèles SQL



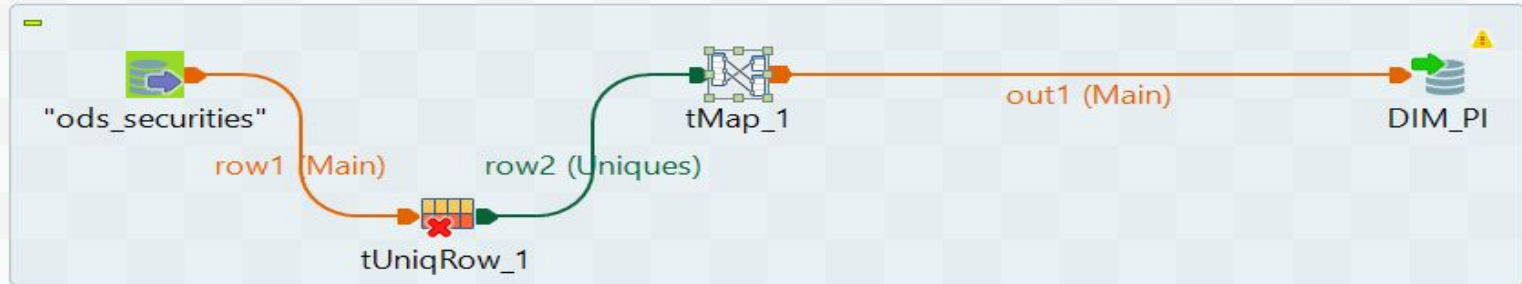
Dimension_Date



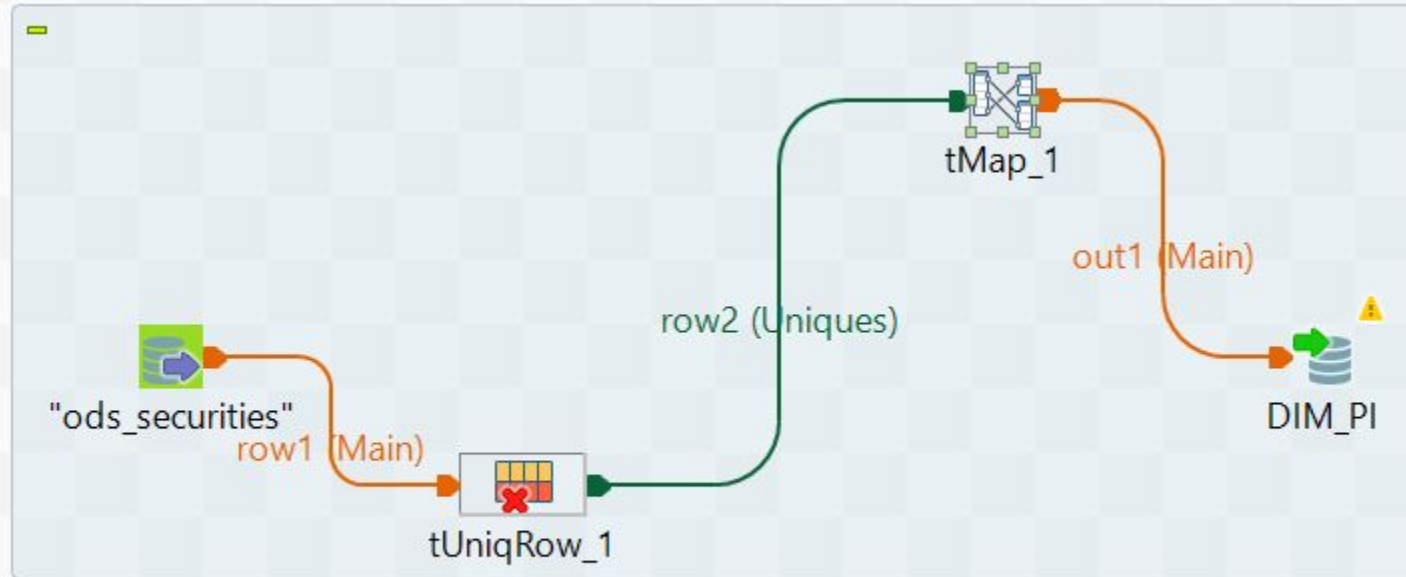
Dimension_Echange



Dimension_Industry



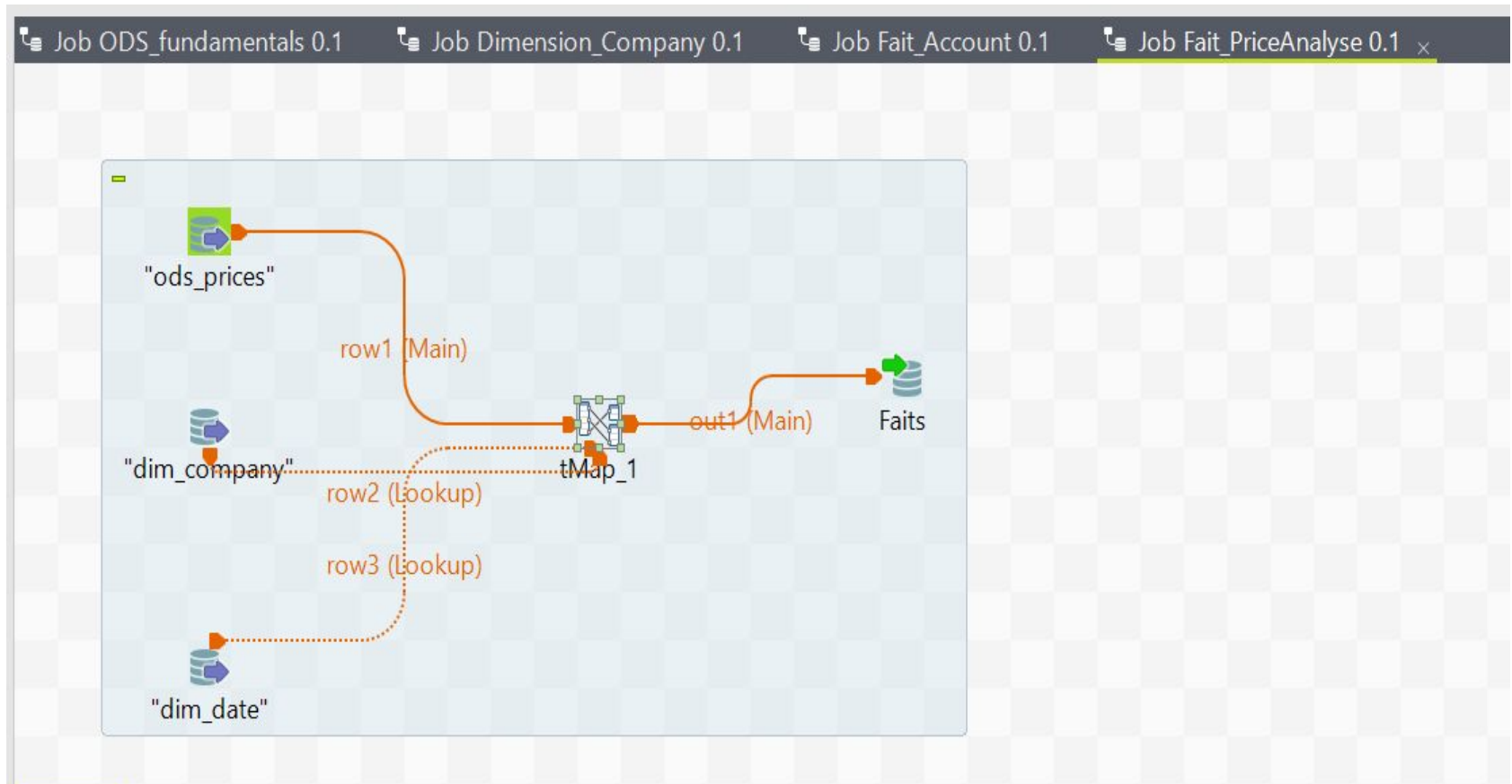
Dimension_Sector



Dimension_Security



Table fails (Fait-price analyse)



jointure t-map

row1

Column

ID

date

symbol

open

close

low

high

volume

row2

Clé d'expr.

Column

Company_ID

Sector_ID

Exchange_ID

Industry_ID

Company_Name

row3

Clé d'expr.

Column

Var

out1

Expression

Numeric.sequence("s1",1,1)

row2.Company_ID

row3.Date_ID

row1.open

row1.close

row1.low

row1.high

row1.volume

Column

PriceAnalyse_ID

Company_ID

Date_ID

Opening_Price

Closing_Price

Daily_Low

Daily_High

Trading_Volume

Éditeur de schéma

Éditeur d'expression

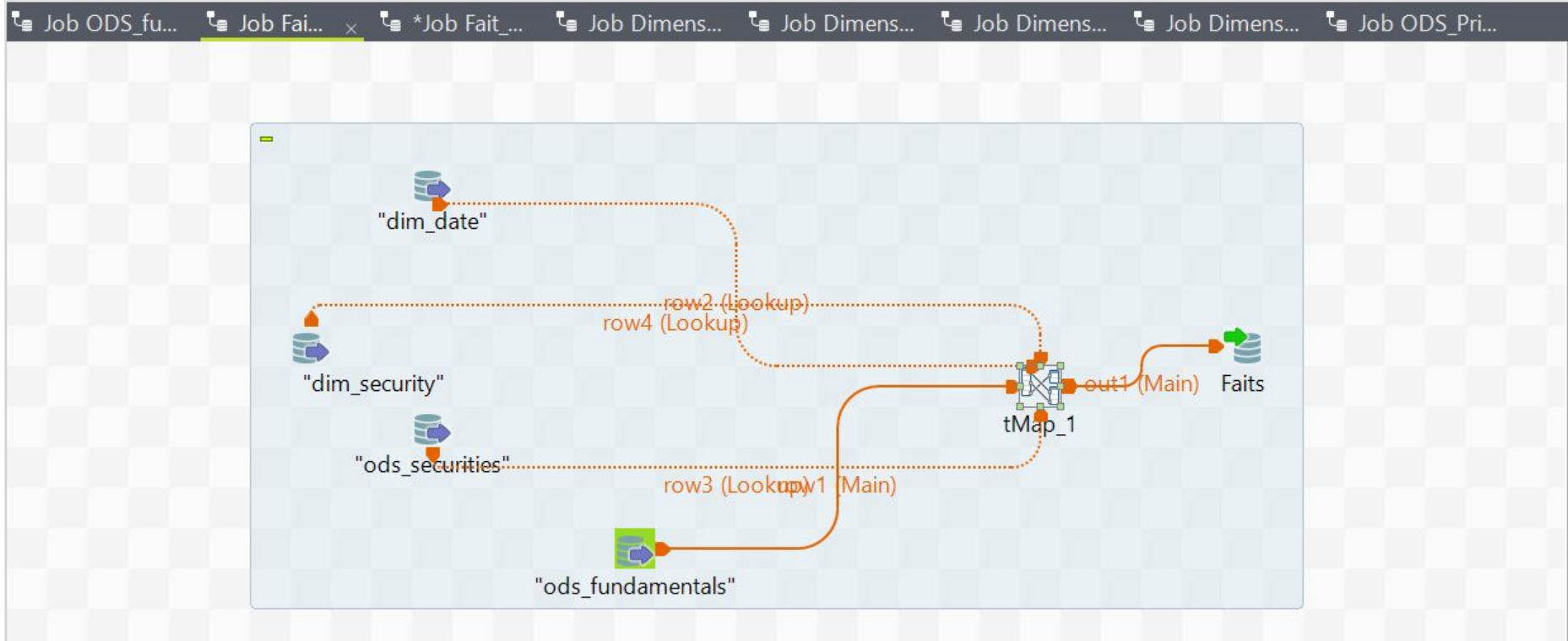
row1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
date	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	"dd-MM-yyyy"	10	0		
symbol	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		10	0		
open	<input type="checkbox"/>	Float	<input checked="" type="checkbox"/>		100	7		

out1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
PriceAnalyse_ID	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>					
Company_ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
Date_ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
Opening_Price	<input type="checkbox"/>	Float	<input checked="" type="checkbox"/>		100	7		

Table faits (Fait-Account)



jointure t-map

row1

Column

- ID
- Ticker_Symbol
- Period_Ending
- Accounts_Payable
- Accounts_Receivable
- Add_I_income_expense_items
- After_Tax_ROE
- Capital_Expenditures
- Capital_Surplus
- Cash_Ratio
- Cash_and_Cash_Equivalents
- Changes_in_Inventories
- Common_Stocks
- Cost_of_Revenue
- Current_Ratio
- Deferred_Asset_Charges
- Deferred_Liability_Charges
- Depreciation
- Earnings Before Interest and Tax

Var

out1

Expression	Column
Numeric.sequence("s1",1,1)	Account_ID
row4.Security_ID	Security_ID
row2.Date_ID	Date_ID
row1.Accounts_Payable	Accounts_Payable
row1.Accounts_Receivable	Accounts_Receivable
row1.Cash_Ratio	Cash_Ratio
row1.Cash_and_Cash_Equivalents	Cash_and_Cash_Equivalents
row1.Cost_of_Revenue	Cost_of_Revenue

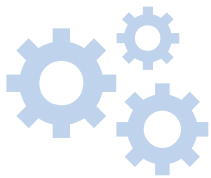
Éditeur de schéma Éditeur d'expression

row1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
Ticker_Symbol	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		10	0		
Period_Ending	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	"dd-MM-yyyy"	10	0		
Accounts_Payable	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		19	0		

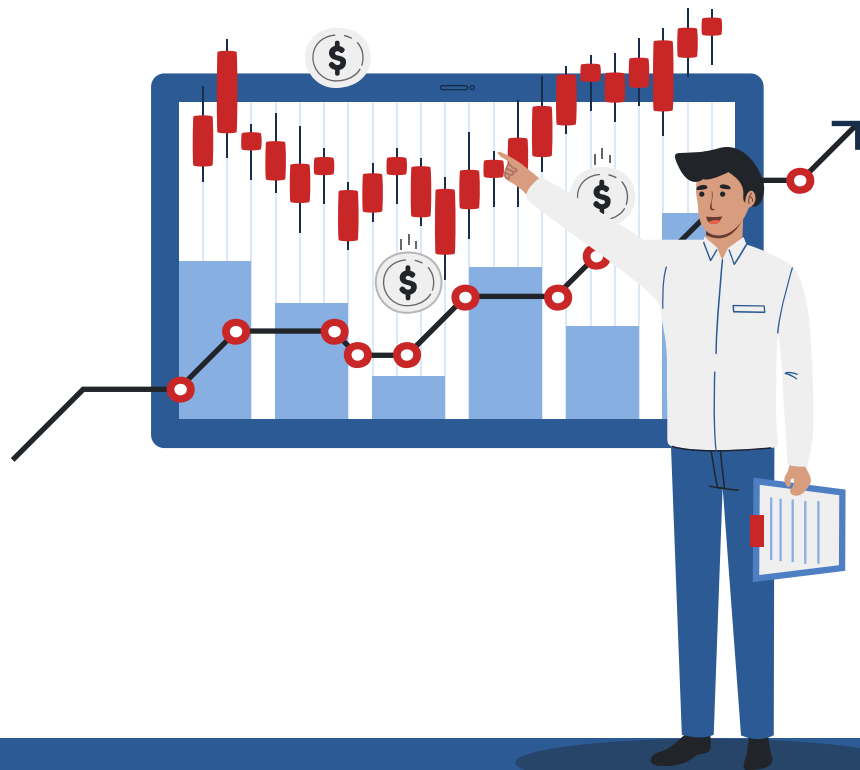
out1

Colonne	C...	Type	<input checked="" type="checkbox"/> N.	Modèle de dat...	Longueur	Précision	Par d...	Comme...
Account_ID	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>					
Security_ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
Date_ID	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
Accounts_Payable	<input type="checkbox"/>	Long	<input checked="" type="checkbox"/>		19	0		



07

Visualisation (POWER BI)



SECTOR TRADING

Industry_Name

Tout

30/06/2003

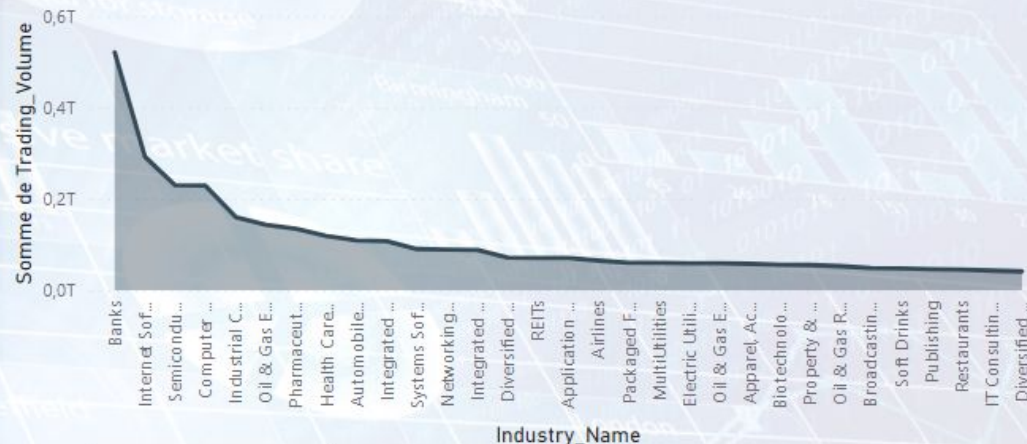
01/01/2017



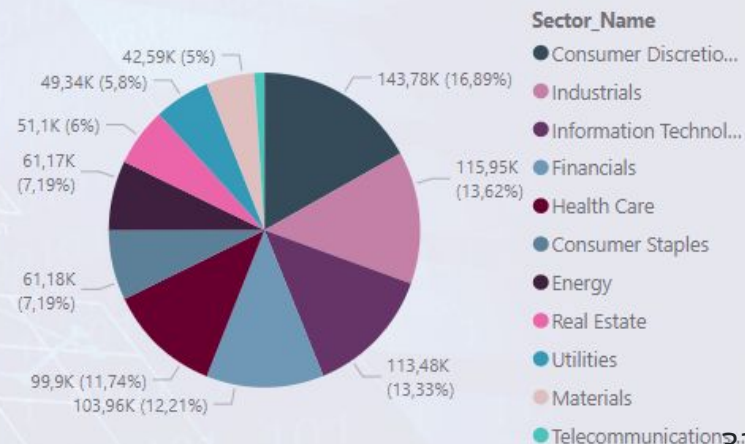
Company_Name	Company_ID	Somme de Closing_Price	Somme de Daily_High	Somme de Daily_Low	Somme de Opening_Price	Somme de PriceAnalyse_ID	Somme de Trading_Volume
AIV	45	53 666,02	54 177,53	53 089,64	53 628,95	740135017	2 587 225 000,00
ZION	504	44 172,44	44 727,19	43 568,31	44 152,82	740132268	5 686 304 100,00
ZBH	503	146 023,74	147 227,42	144 704,98	145 939,09	740130506	2 660 524 400,00
YUM	502	117 162,17	118 159,56	116 096,90	117 106,39	740128744	9 098 410 900,00
YHOO	501	48 088,49	48 648,78	47 531,01	48 098,93	740126982	35 185 494 184,00
XRX	497	17 606,90	17 808,81	17 396,62	17 607,96	740124212	19 602 500 600,00
XRAY	139	78 511,35	79 155,01	77 777,20	78 469,72	740122450	1 878 871 100,00
XOM	180	148 019,42	149 022,59	146 860,70	147 945,92	740120688	28 754 969 904,00
Total		60 318 105,99	60 902 385,46	59 689 281,65	60 300 976,10	362325624480	4 609 600 446 002 000



Somme de Trading_Volume par Industry_Name



Nombre de PriceAnalyse_ID par Sector_Name



COMPANY TRADING

Company Name

Tout

Month

Tout

15T

Somme de Cash_and_Cas...

8T

Somme de Accounts_Paya...

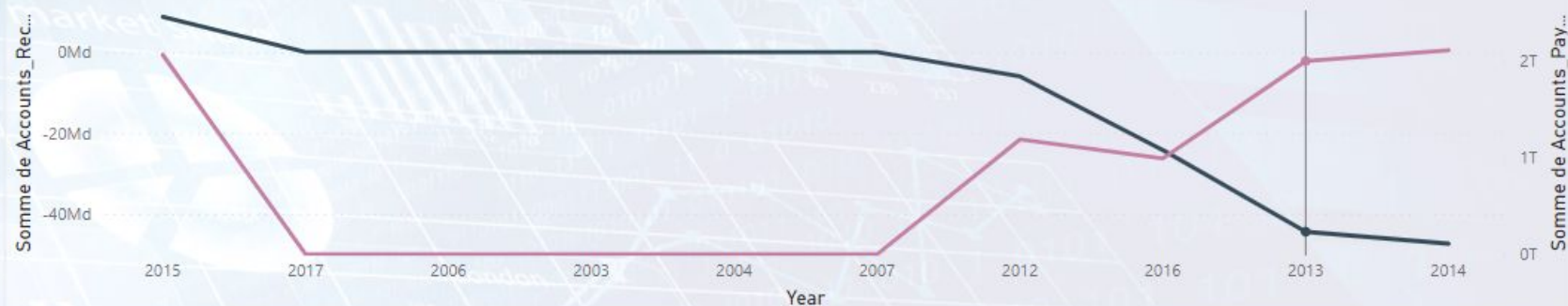
Nombre de Account_ID	Security_ID	Somme de Accounts_Payable	Somme de Accounts_Rec...
3	228	42264000000	1720
4	8	12391815000	-191
4	34	2001346000	-259
4	48	52090000000	3387
4	60	5805500000	-36
4	68	26640000000	152
4	88	8921976000	-759
4	94	1698770000	-166
4	106	2251217000	-153
4	111	143500000	-91
1781		8323357023000	-113155

Country



Somme de Accounts_Receivable et Somme de Accounts_Payable par Year

● Somme de Accounts_Receivable ● Somme de Accounts_Payable



CASH & REVENU ANALYSIS

Company Name

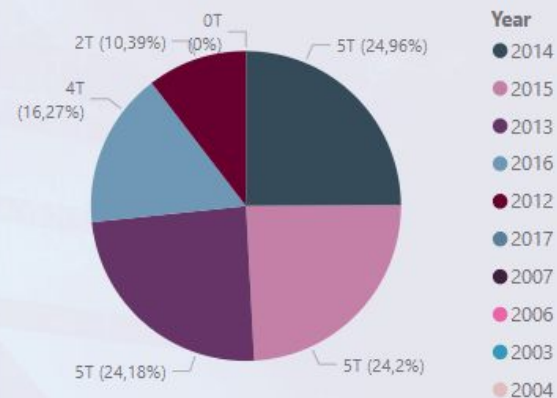
Tout

30/06/2003

01/01/2017

Account_ID	Cost_of_Revenue	Somme de Cash_and_Cash_Equivalents	Somme de Cash_Ratio	Somme de Cost_of_Revenue
69	0	1325500000		
70	0	1682000000		
71	0	1892500000		
72	0	1937600000		
105	0	430400000		
106	0	469600000		
107	0	550600000		
108	0	563800000		
133	0	291000000		
134	0	477000000		
135	0	374000000		
136	0	704000000		
Total		15176612360000	110346	2199913819

Somme de Cost_of_Revenue par Year



Somme de Cash_and_Cash_Equivalents et Somme de Cash_Ratio par Year



DAILY MARKET TRENDS

Company Name

Tout

30/06/2003

01/01/2017



Company_Name	Company_ID	Year	Somme de Daily_High	Somme de Daily_Low	Somme de Opening_Price	Somme de Closing_Price	Nombre de PriceAnalyse_ID	Somme de
MMM	1	2010	21 244,71	20 866,11	21 078,17	21 075,40	252	
MMM	1	2011	22 034,32	21 627,69	21 837,15	21 832,84	252	
MMM	1	2012	22 403,11	22 126,65	22 260,86	22 278,54	250	
MMM	1	2013	28 800,19	28 482,07	28 615,92	28 659,84	252	
MMM	1	2014	36 176,61	35 751,06	35 961,59	35 975,55	252	
MMM	1	2015	39 665,99	39 136,70	39 396,35	39 408,51	252	
MMM	1	2016	42 673,28	42 196,94	42 421,22	42 465,13	252	
ART	2	2010	12 871,07	12 685,52	12 702,00	12 784,46	252	
Total			60 902 385,46	59 689 281,65	60 300 976,10	60 318 105,99	851264	4 66



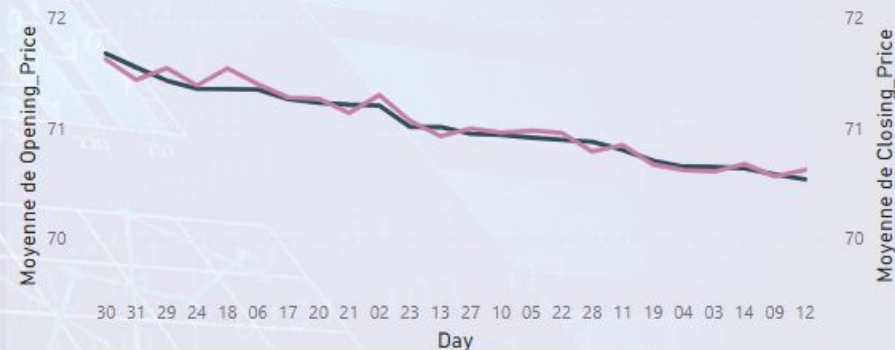
Moyenne de Daily_Low et Moyenne de Daily_High par Day

● Moyenne de Daily_Low ● Moyenne de Daily_High



Moyenne de Opening_Price et Moyenne de Closing_Price par Day

● Moyenne de Opening_Price ● Moyenne de Closing_Price



FORECAST

Sector_Name

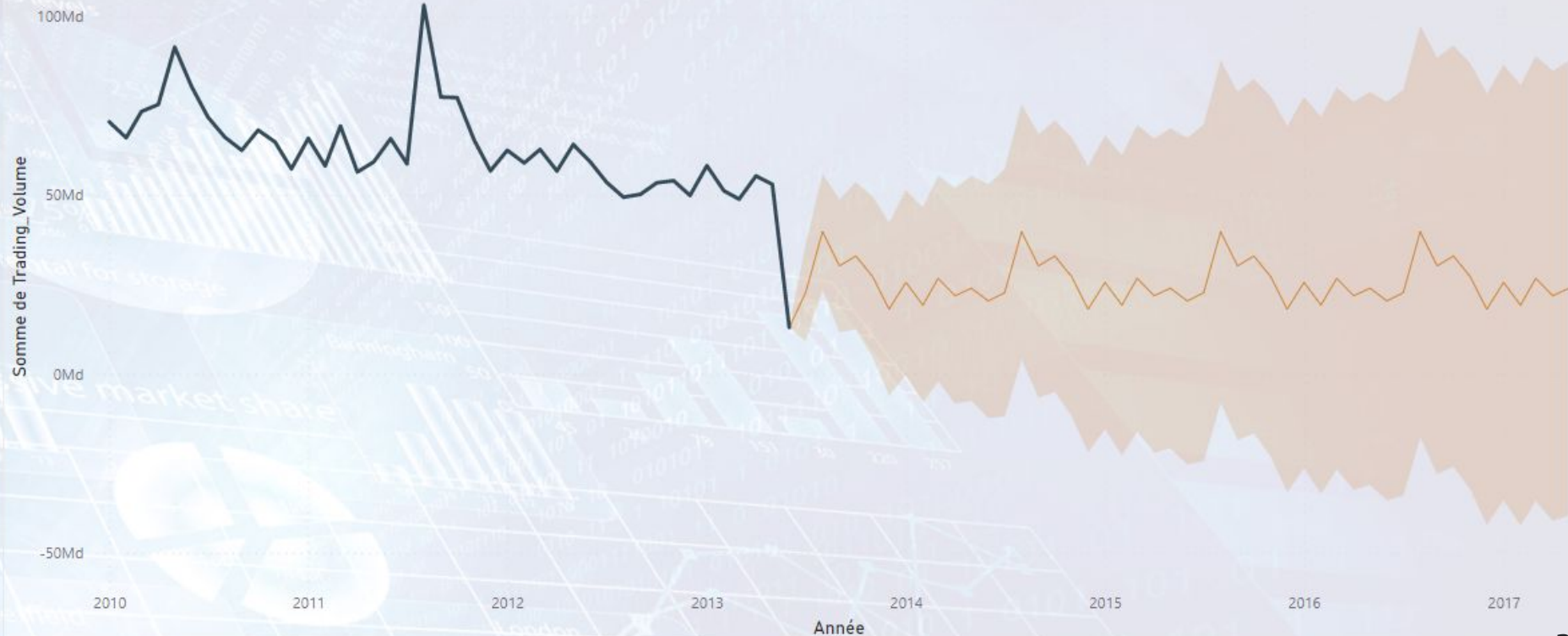
Tout

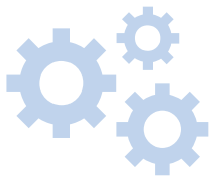
30/06/2003

09/06/2013



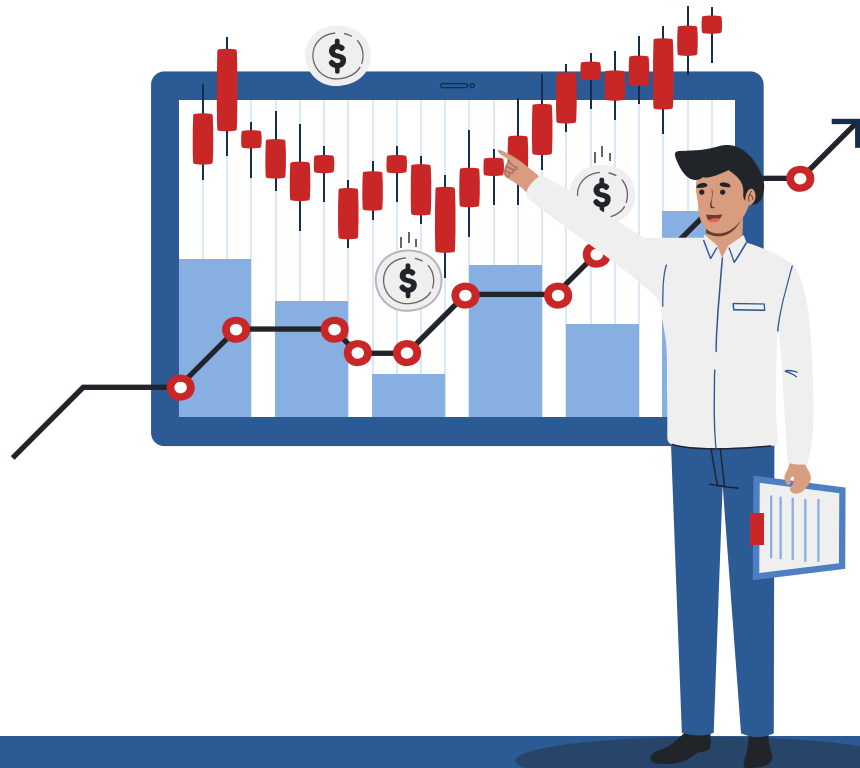
Somme de Trading_Volume par Année et Mois





08

Machine learning



Stock Market Dataset (New York Stock Exchange)

Phase 1 : Data Understanding

Importation des bibliotheques et de Data

```
Entrée [ ]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
data1 = pd.read_csv('C:/Users/Desktop/fact_priceanalyse.csv')
data = pd.read_csv('C:/Users/Desktop/prices-split-adjusted.csv')
```

Afficher les 5 premier lignes des deux Data

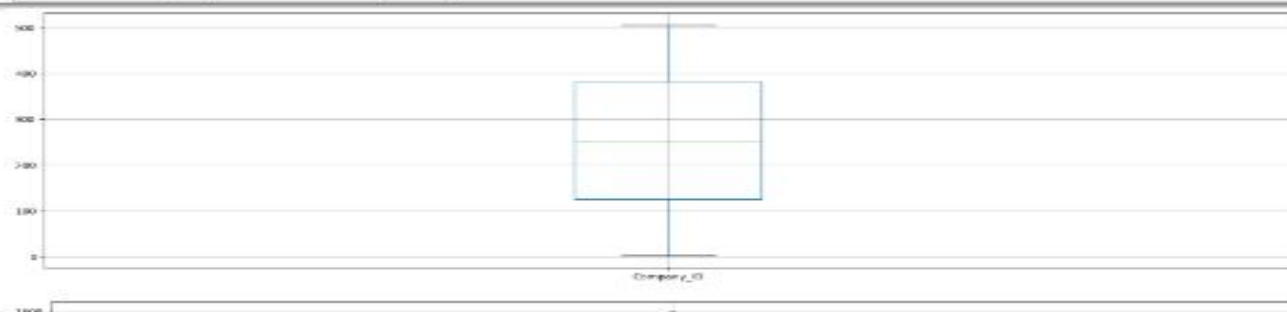
Entrée [77]: data.head()

```
Out[77]:
```

	date	symbol	open	close	low	high	volume
0	2016-01-05	WLTW	123.430000	125.839998	122.309998	126.250000	2163600.0
1	2016-01-06	WLTW	125.239998	119.980003	119.940002	125.540001	2388400.0
2	2016-01-07	WLTW	116.379997	114.949997	114.930000	119.739998	2489500.0
3	2016-01-08	WLTW	115.480003	116.620003	113.500000	117.440002	2006300.0
4	2016-01-11	WLTW	117.010002	114.970001	114.089996	117.330002	1408600.0

Entrée [78]: data1.head()

```
Entrée [188]: for column in dataC.columns:
    fig, ax = plt.subplots(figsize=(20,6))
    dataC.boxplot(column = column , ax=ax)
```



Après avoir observé le boxplot, nous avons remarqué la présence de valeurs aberrantes dans la plupart des colonnes. Cependant, après une analyse plus approfondie, nous avons réalisé que ces valeurs ne sont pas réellement aberrantes, mais plutôt significatives dans le contexte réel. Par conséquent, nous avons pris la décision de ne pas les traiter.

```
Entrée [189]: plt.figure(figsize=(15,8))
plt.title('Opening Price of the Stock Historically')
plt.plot(df['Date_ID'],df['Opening_Price'])
plt.xlabel('Year', Fontsize=20)
plt.ylabel('Opening Price Historically ($)', Fontsize=20)
plt.show()
```



Fichier 2

▶ `data.head()`



	Account_ID	Security_ID	Date_ID	Accounts_Payable	Accounts_Receivable	Cash_Ratio	Cash_and_Cash_Equivalents	Cost_of_Revenue
0	1	459	761	3068000000	-222000000	53	1330000000	10499000000
1	2	459	1035	4975000000	-93000000	75	2175000000	11019000000
2	3	459	1308	4668000000	-160000000	60	1768000000	15620000000
3	4	459	1585	5102000000	352000000	51	1085000000	11096000000
4	5	8	759	2409453000	-89482000	23	598111000	3106967000

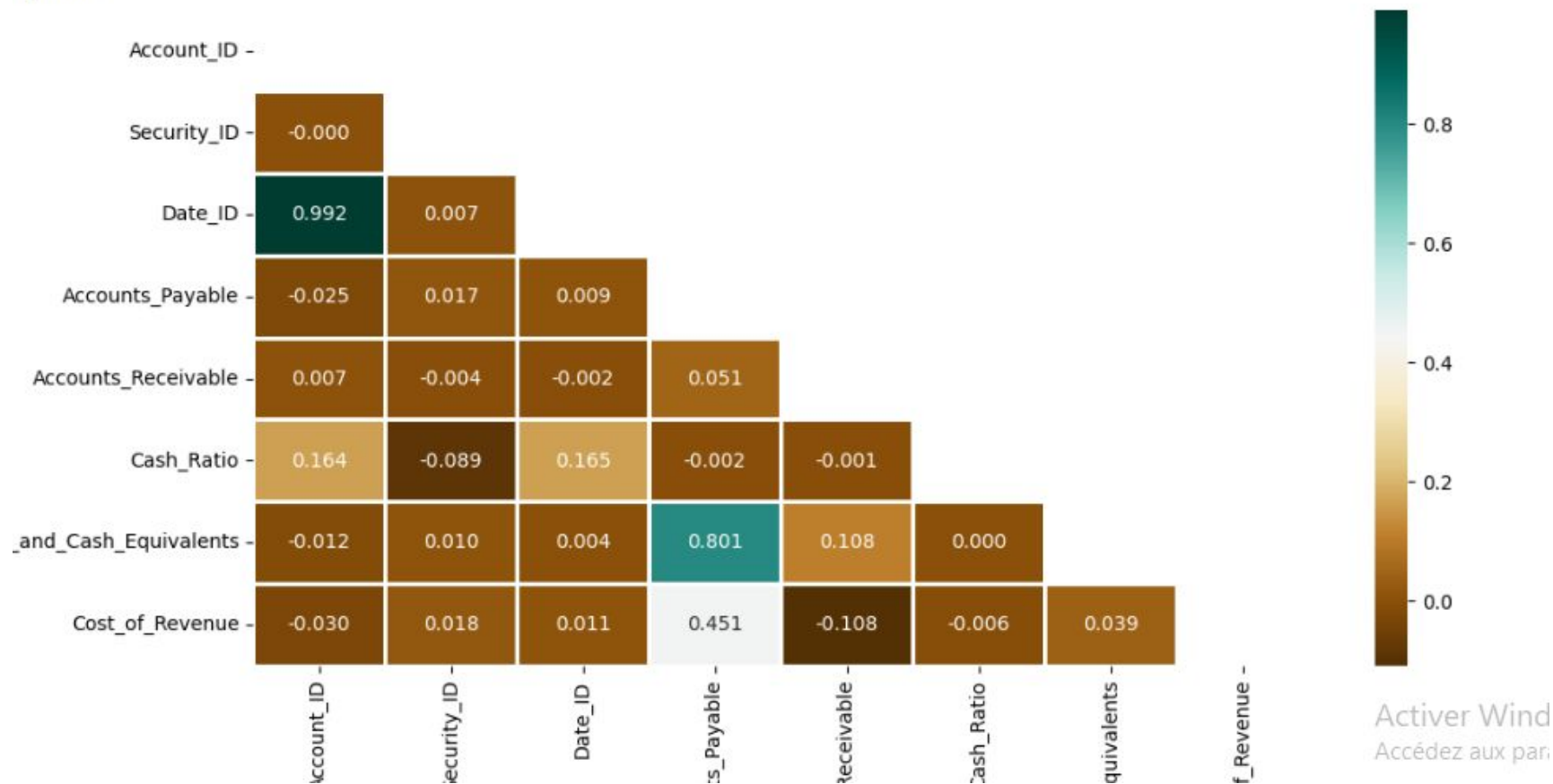
▶ `data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 853046 entries, 0 to 853045
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Account_ID            853046 non-null object
1   Security_ID           853046 non-null object
2   Date_ID               853046 non-null object
3   Accounts_Payable      853046 non-null object
4   Accounts_Receivable   853046 non-null object
5   Cash_Ratio            852747 non-null object
```

Activer Window
Accédez aux param

Fichier 2

unplot: >



MERCI DE VOTRE ATTENTION!

