

- Conflit Israël Palestine - Analyse Multidimensionnelle

Département : Informatique
Filière : Science des données



Réalisé par :

Ismaili Taha

Ennair Aya

Encadré par :

Pr. ZAHIR Jihad

Année universitaire : 2023-2024



Plan

1. Modèle multidimensionnel
2. Source et transformation
3. Mise à jour et Prédiction
4. Outils et Démonstration

1

Introduction

Questions décisionnelles

? Questions décisionnelles





Questions décisionnelles

1. Décès : Comment le nombre de décès liés au conflit a-t-il évolué, et comment ces chiffres se répartissent-ils entre civils et militaires ?
2. Prisonniers : Quelles sont les tendances et les profils des prisonniers , en tenant compte des différents statuts et catégories ?
3. Réfugiés Palestiniens : Quelle est l'évolution du nombre de réfugiés palestiniens depuis 1952, et comment se répartissent-ils par pays d'accueil ?
4. Colonie : Quelles sont les tendances démographiques par colonie au fil des années ?
5. Démographie d'Israël : Comment la composition démographique d'Israël a-t-elle évolué, en termes de flux migratoires et de nationalités représentées ?

2

Modèle Multidimensionnel

Matrice de BUS

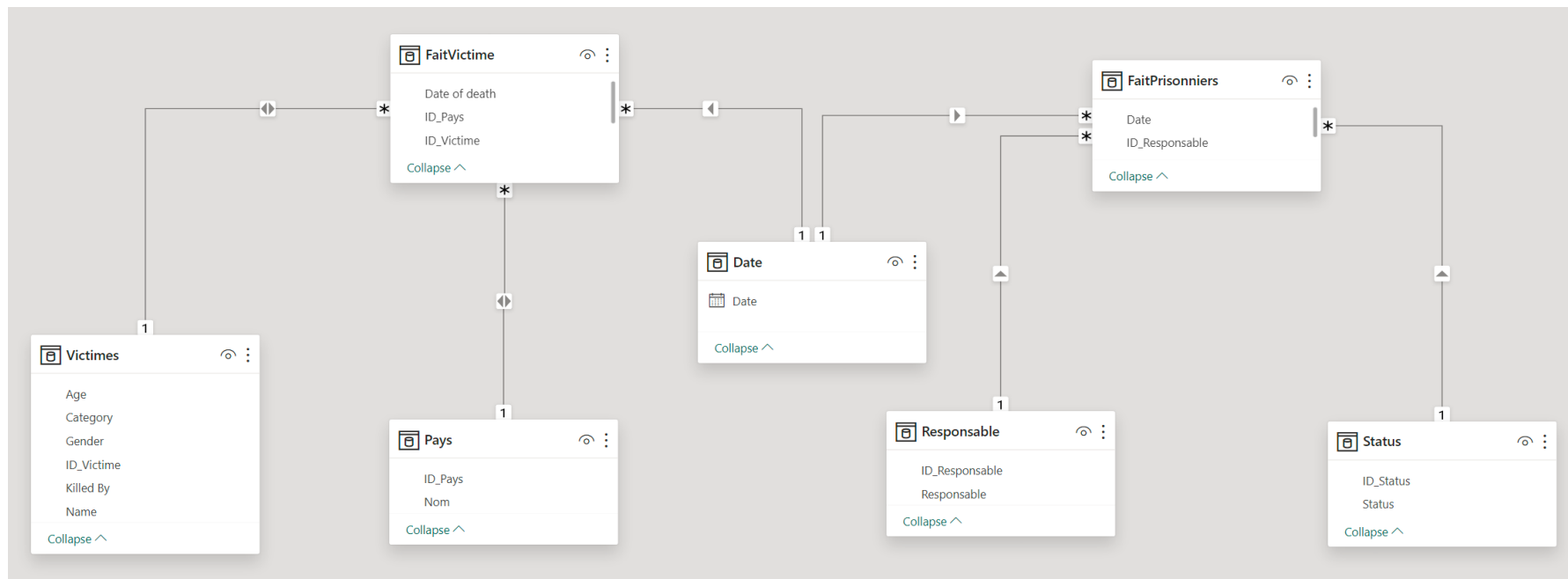
	Victime	Pays	Responsable	Status	Colonie	Continent	Pays origine	Pays d'asile	Religion	Année	Date
Nombre de Victime	<div><div>Granularité plus grossière → Les tendances annuelles</div><div>Granularité Fine → Jour, Mois et l'année</div></div>										
Nombre de Prisonniers											
Nombre d'Habitant											
Nombre de Population											
Nombre d'Immigrant											
Nombre de Réfugiés											
Pourcentage de Religion											

Matrice de BUS

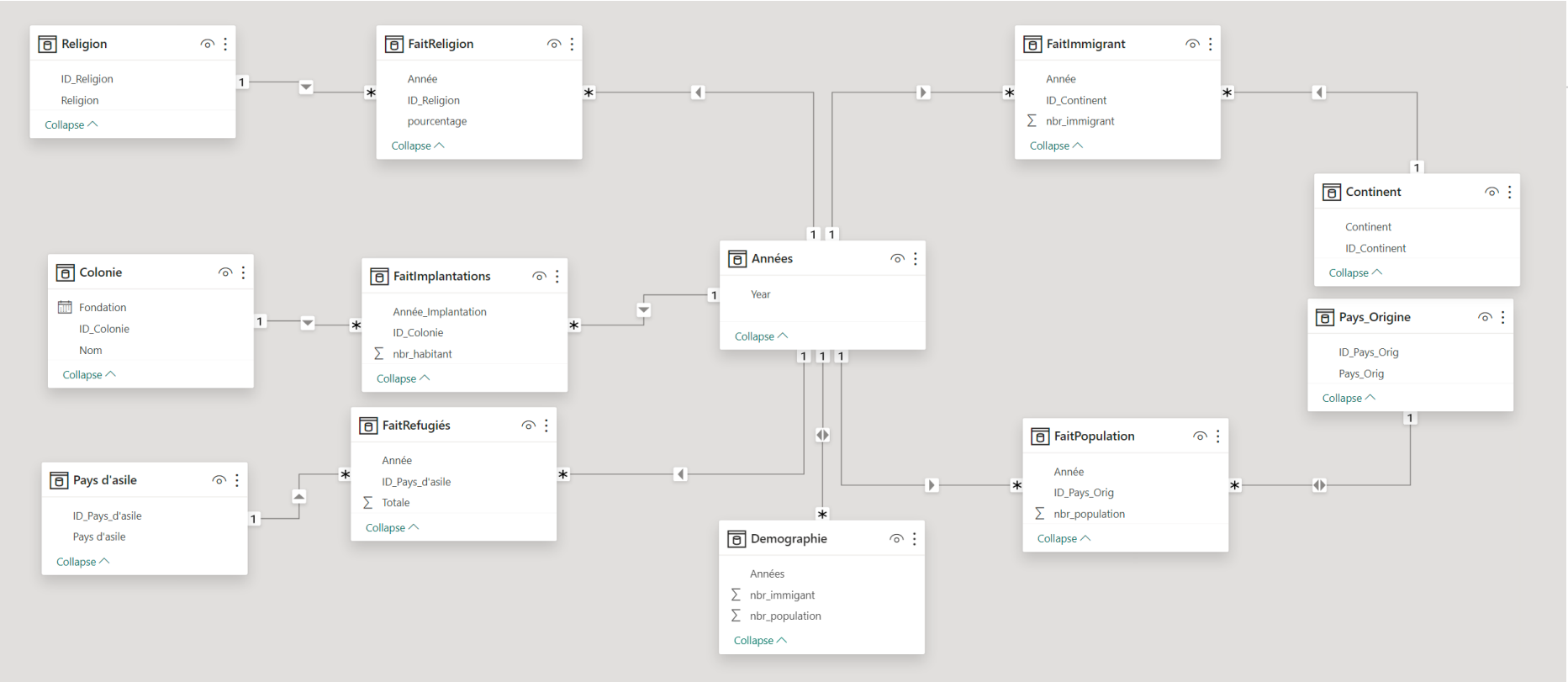
	<i>Victime</i>	<i>Pays</i>	<i>Responsable</i>	<i>Status</i>	<i>Colonie</i>	<i>Continent</i>	<i>Pays origine</i>	<i>Pays d'asile</i>	<i>Religion</i>	<i>Année</i>	<i>Date</i>
Nombre de Victime	×	×									×
Nombre de Prisonniers			×	×							×
Nombre d'Habitant					×					×	
Nombre de Population							×			×	
Nombre d'Immigrant						×				×	
Nombre de Réfugiés								×		×	
Pourcentage de Religion									×	×	



Le modèle logique



Le modèle logique

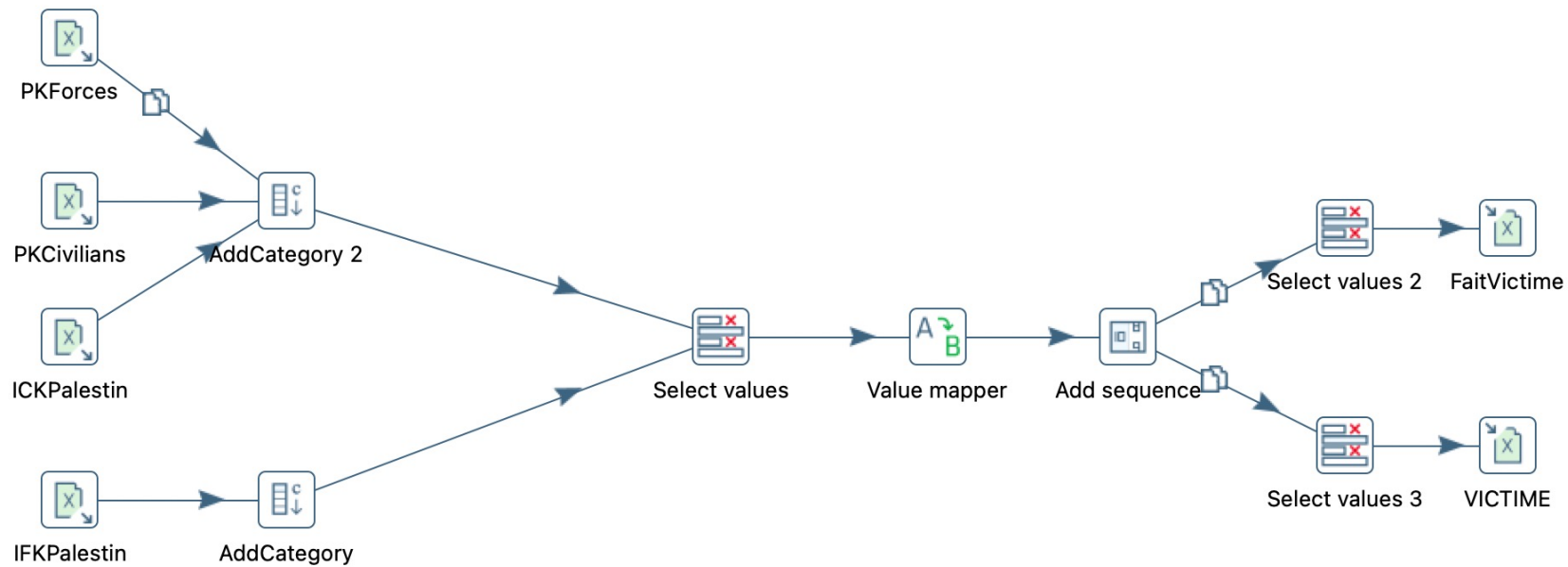


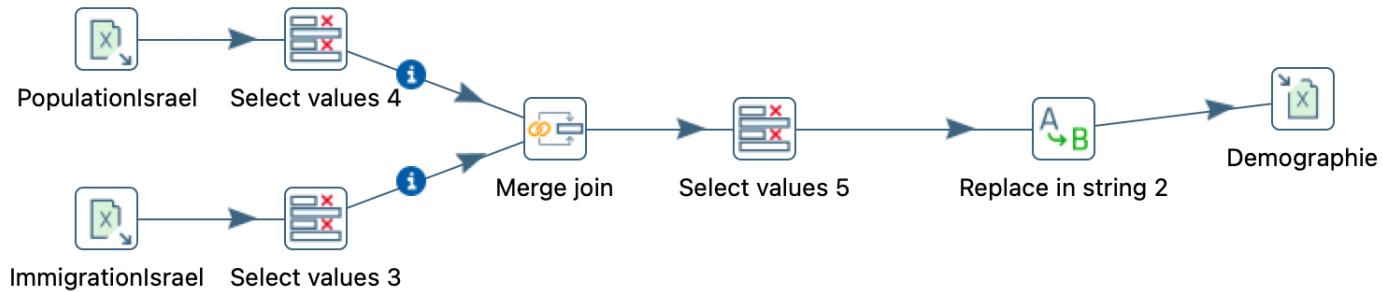
3

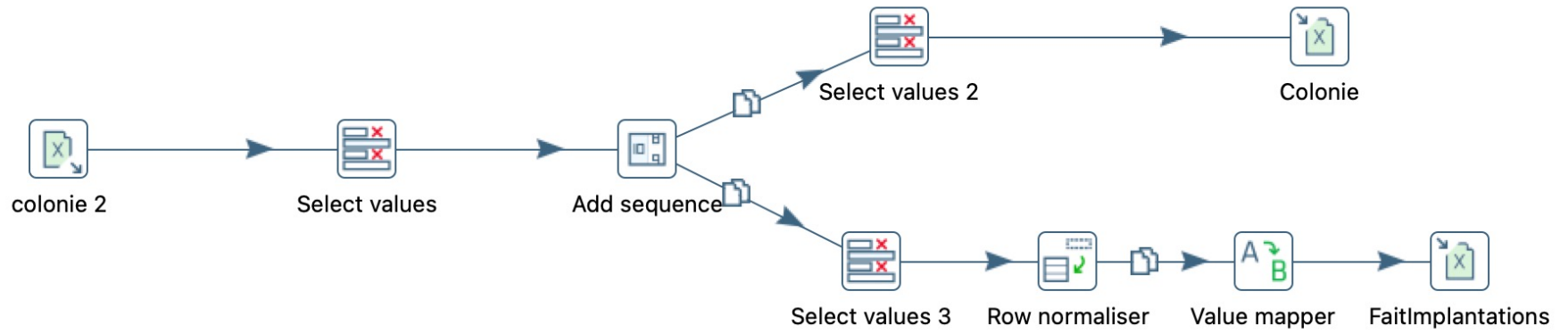
Transformation des Données

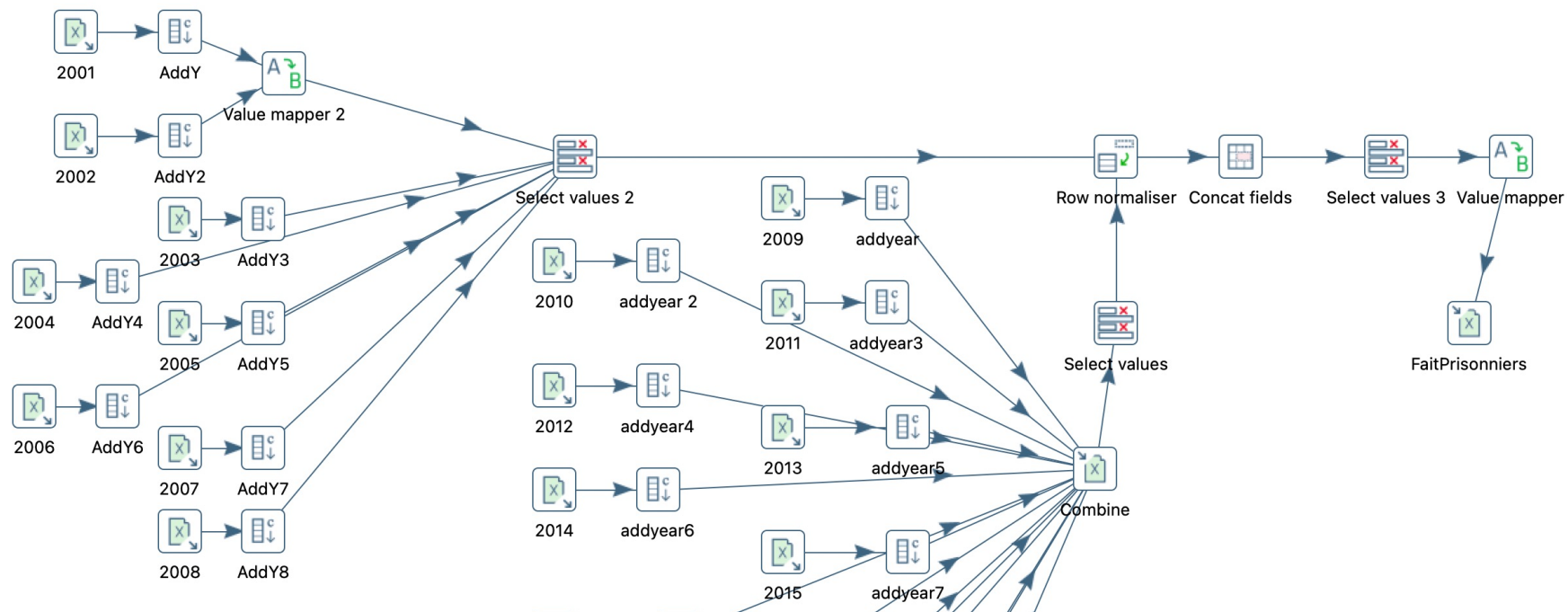
Description des sources de données

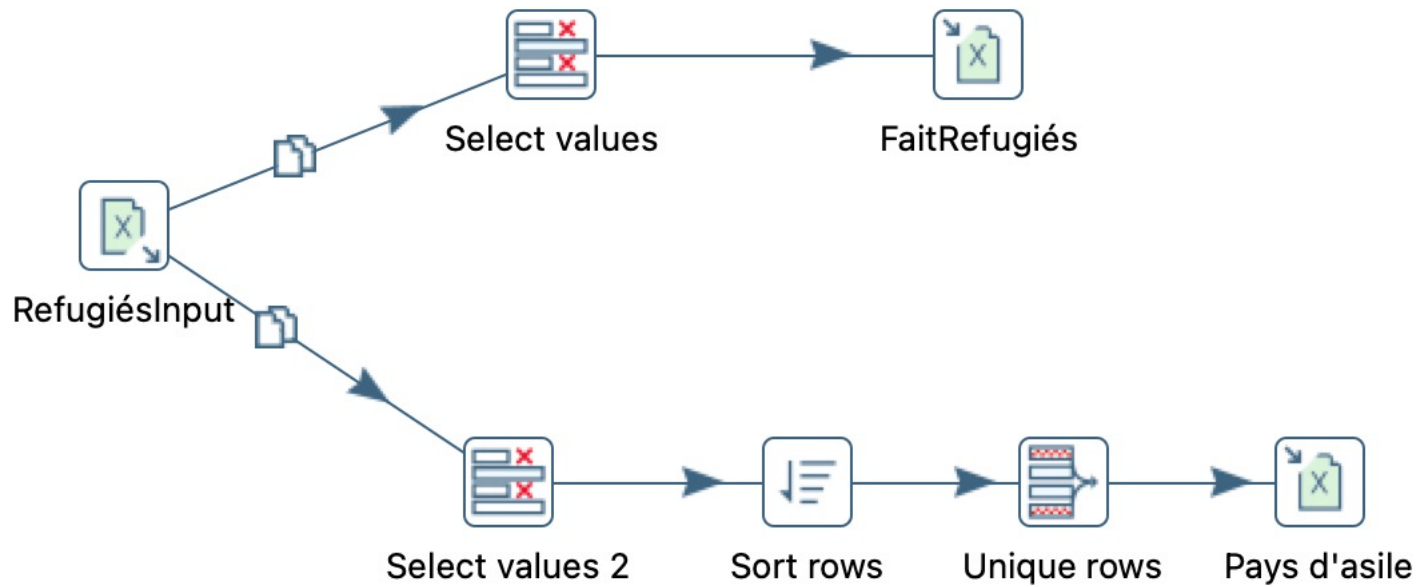
Intitulé du fichier	Source	Lien hypertexte	F de m.à.j dans S	Méthode d'acquisition
2001-2023	B'TSELEM	Prisonniers_source	1 ans	Scrapping
ImmigrationIsrael	Perspective Monde	immegration_source	1 ans	Copier coller
PopulationIsrael	Perspective Monde	Population_source	1 ans	Copier coller
PKForces-PKCivilians IFKPalestin- ICKPalestin	B'TSELEM	Victimes_source	1 mois	Téléchargement
ContinentIsrael	wikipedia	ContIsrael_source	n/a	Copier coller
Colonies	wikipedia	Colonies_source	n/a	Scrapping
PaysIsrael	wikipedia	PaysIsrael_source	n/a	Copier coller
ReligionIsrael	wikipedia	ReligionIsreal_source	n/a	Copier coller
Refugie	UNHCR	Refugiés_source	1 ans	Téléchargement







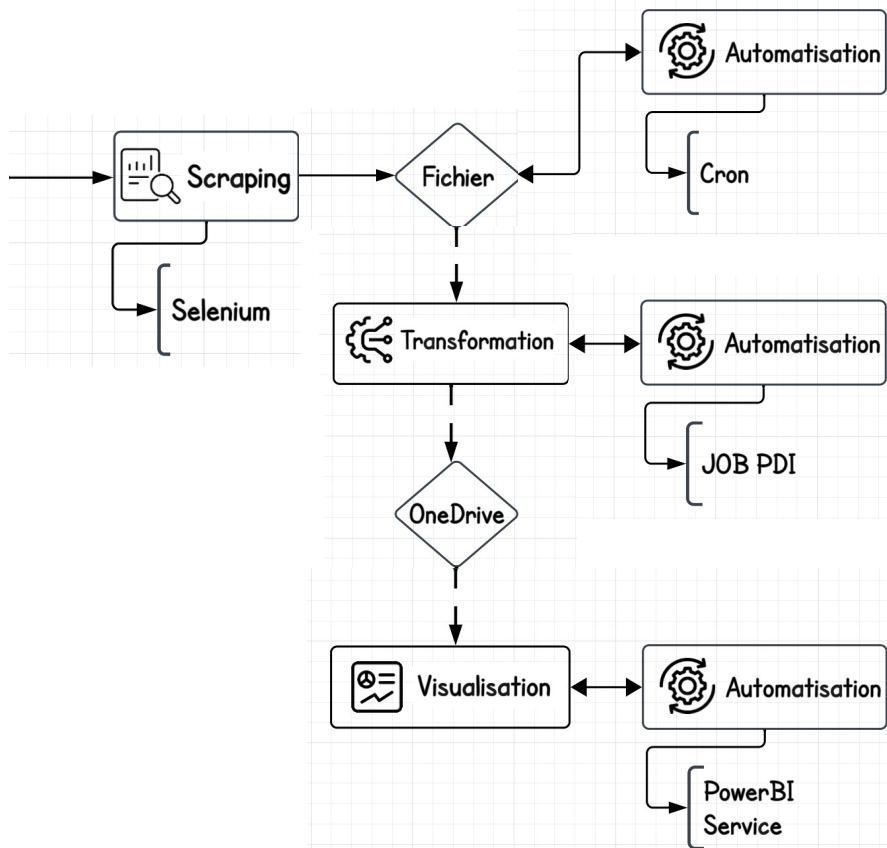




3

Mise à jour et Prédiction

Source de données



```

# Chemin du dossier de téléchargement
download_folder = "/Users/pro/Desktop/DataSource"
if not os.path.exists(download_folder):
    os.makedirs(download_folder)
# Options de Chrome
chrome_options = Options()
prefs = {"download.default_directory" : download_folder}
chrome_options.add_experimental_option("prefs", prefs)
service = Service(ChromeDriverManager().install())
driver = webdriver.Chrome(service=service, options=chrome_options)
driver.get("https://statistics.btselem.org/en/all-fatalities/by-date-of-incident?section=overall&tab=overview")

try:
    download_button = WebDriverWait(driver, 10).until(
        EC.element_to_be_clickable((By.CSS_SELECTOR, "#overall > div > div > div > div:nth-child(1) > div > div.v-ca
    )

    download_button.click()
    print("Bouton cliqué")

    # Attendre que le fichier téléchargé apparaisse dans le dossier de téléchargement
    timeout = 30
    start_time = time.time()
    previous_file_count = len([filename for filename in os.listdir(download_folder) if filename.endswith('.xlsx')])

    while True:
        current_file_count = len([filename for filename in os.listdir(download_folder) if filename.endswith('.xlsx')])

        if current_file_count > previous_file_count:
            print("Nouveau fichier téléchargé.")
            break

        if time.time() - start_time > timeout:
            print("Timeout atteint. Aucun fichier téléchargé.")
            break
        time.sleep(1) # Attendre une seconde

    # Identifier le fichier le plus récent dans le dossier de téléchargement
    latest_file = max([os.path.join(download_folder, f) for f in os.listdir(download_folder)], key=os.path.getctime)

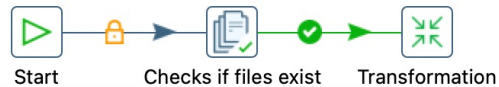
    # Définir le nouveau nom pour le fichier téléchargé
    new_file_name = os.path.join(download_folder, "PKForces.xlsx")

    # Remplacer le fichier s'il existe déjà
    if os.path.exists(new_file_name):
        os.remove(new_file_name)

    # Renommer le fichier téléchargé
    os.rename(latest_file, new_file_name)
    print("Fichier renommé et enregistré sous:", new_file_name)

finally:
    driver.quit()

```



Start

Job entry name:

Repeat: ☒

Type:

Interval in seconds:

Interval in minutes:

Time of day:

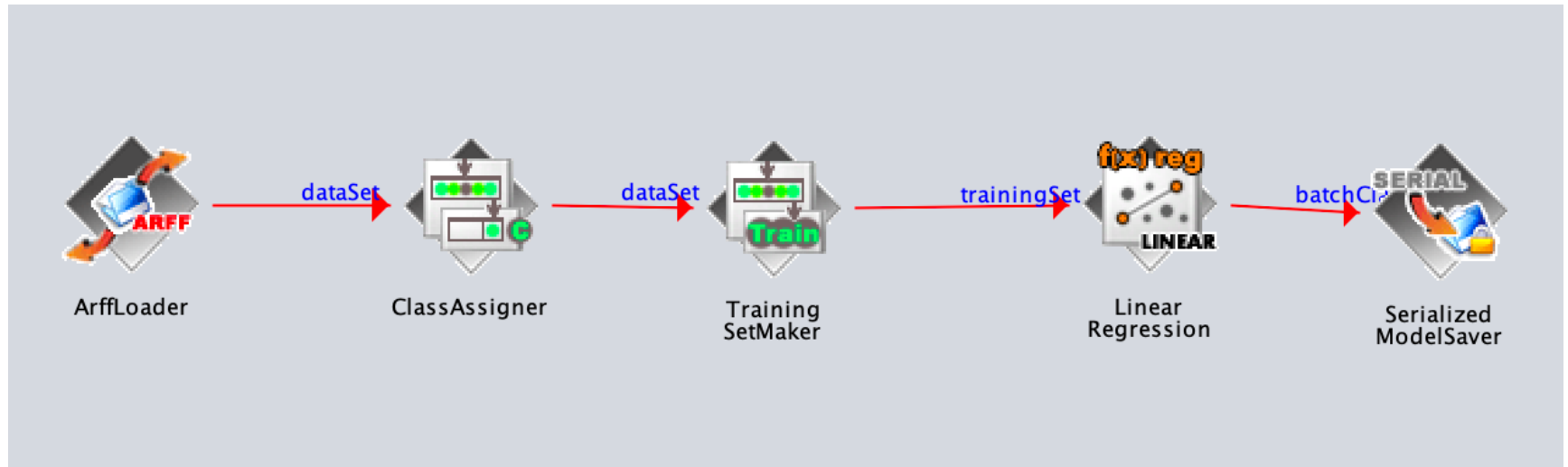
Day of week:

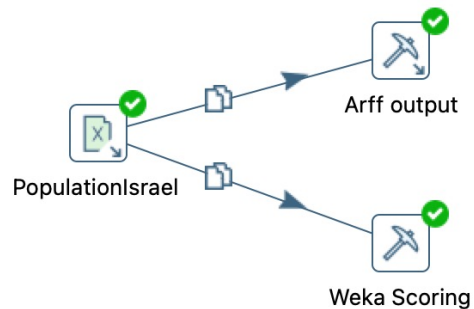
Day of month:

```
pro — vi — crontab — 92x24
0 10 * * 1 /opt/homebrew/anaconda3/bin/python /Users/pro/Desktop/DataSource/scrapingData.py
>> /Users/pro/Desktop/ProjetBI/cron_log.txt 2>&1
```



Le modèle prédictif





Execution Results

<div> <div>Logging</div> <div>Execution History</div> <div>Step Metrics</div> <div>Performance Graph</div> <div>Metrics</div> <div>Preview data</div> </div>				
<div> <div>First rows</div> <div>Last rows</div> <div>Off</div> </div>				
#	Années	Valeurs	Variations	Valeurs_predicted
1	1960	2114020.0	<null>	5326566.984126985
2	1961	2185000.0	3.36	5690575.9658376165
3	1962	2293000...	4.94	5690575.9658376165
4	1963	2379000.0	3.75	5690575.9658376165
5	1964	2475000.0	4.04	5690575.9658376165
6	1965	2563000...	3.56	5690575.9658376165
7	1966	2629000...	2.58	5393029.21897071
8	1967	2745000.0	4.41	5690575.9658376165
9	1968	2803000...	2.11	5455632.288744874
10	1969	2877000.0	2.64	5455632.288744874
11	1970	2974000.0	3.37	5455632.288744874
12	1971	3069000...	3.19	5455632.288744874
13	1972	3148000.0	2.57	5455632.288744874
14	1973	3278000.0	4.13	5455632.288744874

2

Outils et Demonstration



Outils Utilisés



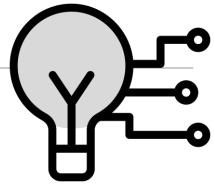
Pentaho Data
Integration



PowerBI



Weka



Conclusion



**Merci
pour votre
attention!**