Analysis and report on the human consequences of the conflict between Palestine and Israel

BENALI Ayoub and ANOUZLA Yassine
Supervised by Pr. ZAHIR Jihad

MARRAKECH
جامعة القاضي عياض
UNIVERSITÉ CADI AYYAD

2023/2024

Considered decision-making questions

- What is the historical context of the conflict between Palestine and Israel ?

- How has the conflict evolved over time, and what are the key events that have shaped its course?

- How have displacement, refugee movements, and internal migrations affected communities on both sides?

- What is the overall number of fatalities attributed to the conflict, and how has it evolved over different phases of the conflict?

- What types of injuries are most prevalent among palestine and israel in the conflict zones

- How has the conflict influenced the overall human rights framework in the region?

- How have human rights been impacted by the conflict, both in terms of violations and protections?

# Proposed multidimensional model: BUS Matrix

The BUS matrix shows the frequency of how many times a dimension or a fact has been used.

Facts	Year	Gender	Citizenship	Country	Month	Occupation	Country ISO	Event Location	Facility	Minor	Dimension Usage Count
Fatalities	1	1	1	1	1	1		1		1	8
Immigration	1		1	1			1				4
Settlements	1										1
Events	1										1
Demography	1			1			1				3
Prisoners	1					1			1		3
Refugees	1		1	1			1				4
Destruction	1										1
Fact Usage Count	8	1	3	4	2	1	3	1	1	1	

Description of data sources

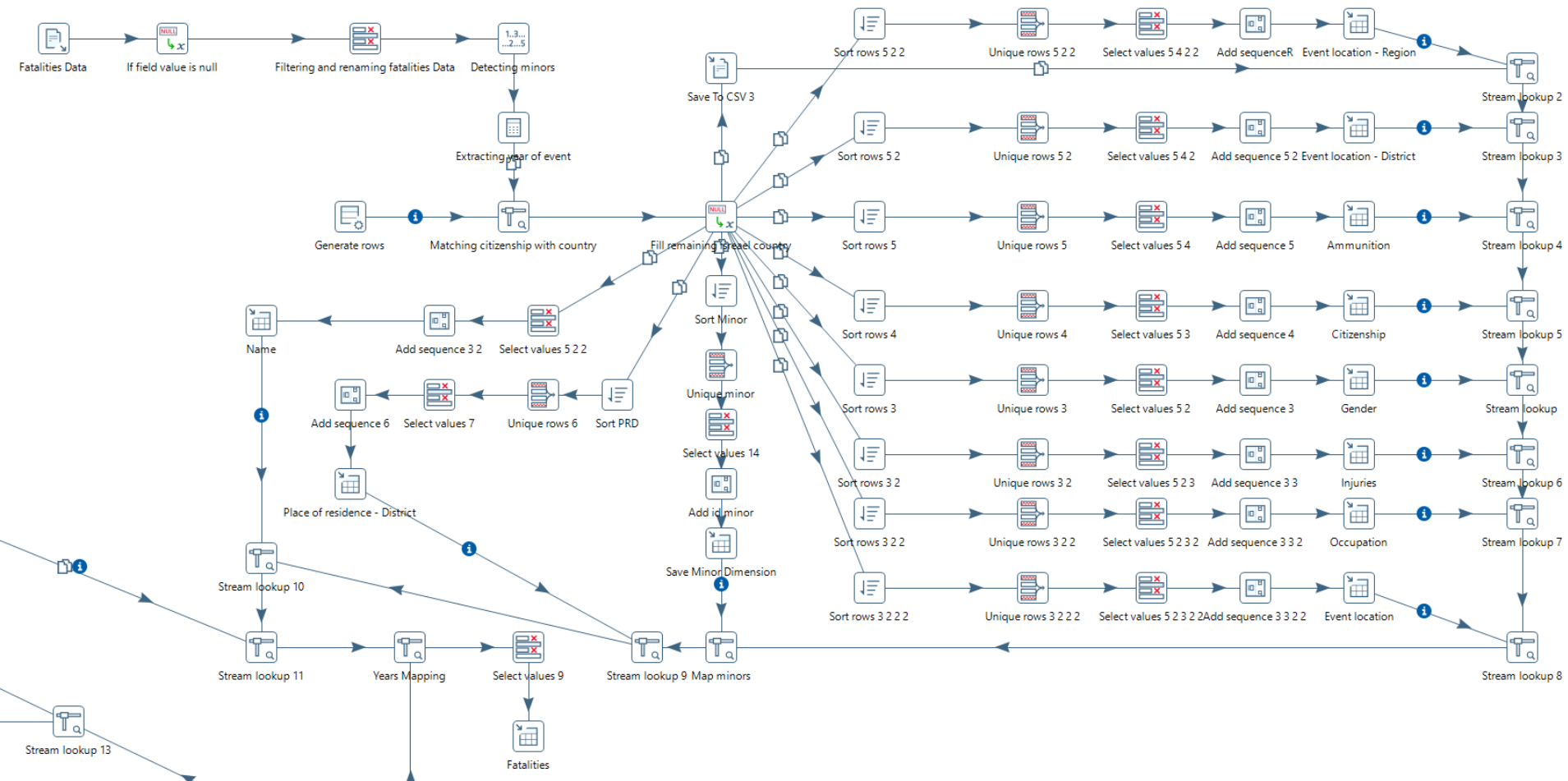| Intitulé du fichier	Source	Lien hypertexte	Fréquence de mise à jour dans la source	Méthode d'acquisition
Fatalities_data	**B'Tselem**	https://statistics.btselem.org/	6 months	Scrapping
prisonners_data	**B'Tselem**	https://www.btselem.org/statistics/detainees_and_prisoners	3 month	Scrapping
refugees_data	**UNHCR**	https://api.unhcr.org/	Each year	Scrapping
settlements_data	**PCBS => Palestinian Central Bureau of Statistics**	https://www.pcbs.gov.ps/Portals/_Rainbow/Documents/SETT 5E	Each year	Scraping
demographics_data	**worldbank**	https://api.worldbank.org/v2/en/indicator/SP.POP.TOTL?downloadformat=csv	Each year	Scraping
live_statistics	**PCBS**	https://www.pcbs.gov.ps	Each day	Scraping

Prop

5
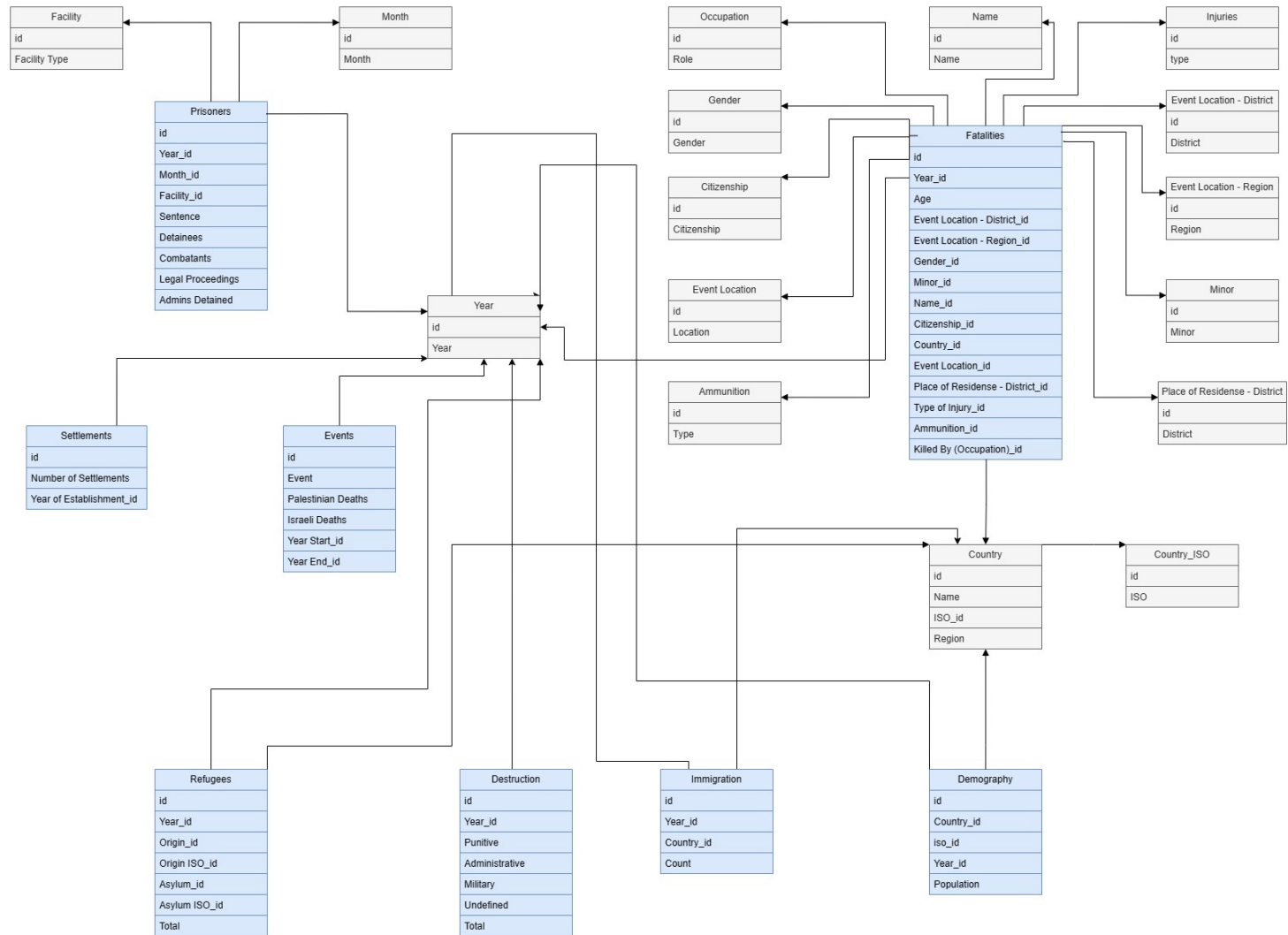
Transformations defined on the data

We used Pentaho Data Integration during the phase of transforming raw data into a format and into many useful informations that are pertinent to our report.

Fatalities Data
If field value is null
Filtering and renaming fatalities Data
Detecting minors
Extracting year of event
Generate rows
Matching citizenship with country
Fill remaining "israel country"
Save To CSV 3

Sort rows 5 2 2
Unique rows 5 2 2
Select values 5 4 2 2
Add sequenceR
Event location - Region
Stream lookup 2

Sort rows 5 2
Unique rows 5 2
Select values 5 4 2
Add sequence 5 2
Event location - District
Stream lookup 3

Sort rows 5
Unique rows 5
Select values 5 4
Add sequence 5
Ammunition
Stream lookup 4

Sort rows 4
Unique rows 4
Select values 5 3
Add sequence 4
Citizenship
Stream lookup 5

Sort rows 3
Unique rows 3
Select values 5 2
Add sequence 3
Gender
Stream lookup

Sort rows 3 2
Unique rows 3 2
Select values 5 2 3
Add sequence 3 3
Injuries
Stream lookup 6

Sort rows 3 2 2
Unique rows 3 2 2
Select values 5 2 3 2
Add sequence 3 3 2
Occupation
Stream lookup 7

Sort rows 3 2 2 2
Unique rows 3 2 2 2
Select values 5 2 3 2 2
Add sequence 3 3 2 2
Event location
Stream lookup 8

Name
Add sequence 3 2
Select values 5 2 2
Add sequence 6
Select values 7
Unique rows 6
Sort PRD
Place of residence - District

Sort Minor
Unique minor
Select values 14
Add id minor
Save Minor Dimension

Stream lookup 10
Stream lookup 11
Years Mapping
Select values 9
Stream lookup 9
Map minors
Fatalities

Stream lookup 13

Update Strategy

In our update strategy we created a backend server with python that handles consuming the APIs and Scraing the data from the given websites.

EXTRACT   **TRANSFORM**   **LOAD**

Update Strategy

The update strategy has 2 main players

The Pentaho jobs scheduler

The Python Backend Server

The predictive model
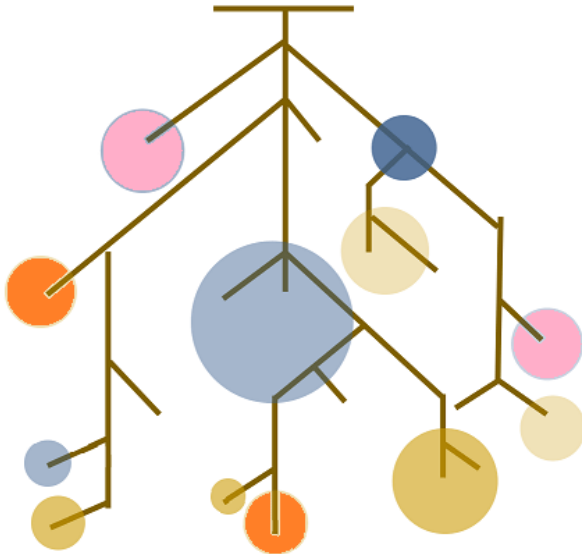
Random Forest Classifier

- High accuracy

- Robustness

- Versatility

- Scalability

The predictive model

The dataset we used is the prisoner's dataset in our project

Prisoners Data							
Facility	Sentence	Detainees	Combatants	Legal Proceedings	Admins Detained	Month	Year
Categorical	Numerical	Numerical	Numerical	Numerical	Numerical	Categorical	Numerical
Total Number of rows : 421							

We applied our classification model on the Facility categorical column that has if the prisoners was held by the IDF (Israel defense force) or the IPS (Israel Police Service)

The predictive model

In our first iteration of the training, the model surprisingly performed very well! on the default hyperparameters and the amount of the data we gave it

Metric	Result
Correctly Classified Instances	95.2381 %
Incorrectly Classified Instances	4.7619 %
Root mean squared error	0.1412

Tools used

➢ **pentaho** : Pentaho is business intelligence software that is used on our project as tool for (Extract, transform, load) process .Also pentaho is used as Backend server to run a scraping tool with schedule timing.

➢ : PostgreSQL , also known as Postgres, is a free and open-source relational database management system used on our project as a tool to store the scraped data on the format of relational database .

➢ : Python is a high-level, general-purpose programming language used on our project to run the script that can scrape data from the given links or url of the websites .

Tools used

➢ : Waikato Environment for Knowledge Analysis is a collection of machine learning and data analysis used on our project as a tools of a predictive model using the Random Forest Classifier technique.

➢ : Microsoft Power BI is an interactive data visualization software used on our project as a tool of creating the main report that can provide some useful informations to more understand the human consequences of the conflict between Palestine and Israel.

Demonstration

Conclusion

The successful completion of the project on the Israeli-Palestinian conflict was a enriching experience, successfully implementing multidimensional models, ETL tools such as PDI, as well as the Power BI visualization platform. The project provided a profound understanding of the conflict, simultaneously expanding our expertise in data analysis. In summary, this project not only yielded tangible results but also contributed significantly to our professional development and a better understanding of this complex subject.

Analysis and report on the human consequences of the conflict between Palestine and Israel

BENALI Ayoub and ANOUZLA Yassine
Supervised by Pr. ZAHIR Jihad

MARRAKECH
جامعة القاضي عياض
UNIVERSITÉ CADI AYYAD

2023/2024