Corentin MAPELLA & Glodie TANDU

**Projet EDD**

**Service des renseignements et surveillance du Territoire (DGSI) :** Prévenir du terrorisme

2ème Année de Master MIAGE - 2018



# Information du projet :

Ce document comporte le cahier des charges du projet relatif à YOUNG PEOPLE SURVEY. Il inscrit les objectifs et les besoins attendus de l’application ainsi qu’une présentation des fonctionnalités, les spécifications et le détail des développements de celui-ci.



|  |  |
| --- | --- |
| **Nom du projet :** | **PREVENIR DU TERRORISME** |
| **Groupe :** | Corentin MAPELLA – Glodie TANDU |
| **Version du document :** | 0.4 |
| **Statut :** | Ajout de l’UML |

# Information groupe :

|  |  |
| --- | --- |
| **NOM** | **RÔLE** |
| Corentin MAPELLA | Rédacteur du cahier des charges  Analyse des données  Utilisateur de Power BI |
| Glodie TANDU | Approbateur du cahier des charges  Analyse des données  Utilisateur de Power BI |

# Révision du document :

|  |  |  |
| --- | --- | --- |
| **VERSION** | **NOM** | **DESCRIPTION** |
| 0.1 | gtandu – cmapella | Création du document |
| 0.2 | gtandu – cmapella | Rédaction de la présentation du sujet |
| 0.3 | gtandu – cmapella | Ajout du contexte et de l’objectif |
| 0.4 | gtandu – cmapella | Ajout de l’UML |

# Introduction

Dans le cadre de notre Master 2 MIAGE Apprentissage, un projet d'entrepôt de données doit être réalisé. Ce projet consiste à collecter et stocker des données pour en faire une analyse afin d’aider une entreprise à prendre des décisions : Nous parlerons **d’analyse décisionnelle**.

Nous allons utiliser un Data Warehouse comme outil d’aide à la décision. Pour rappel, le Data Warehouse, ou entrepôt de données, est une **base de données dédiée au stockage de l'ensemble des données** utilisées dans le cadre de la prise de décision et de l'analyse décisionnelle.

Le Data Warehouse est exclusivement réservé à cet usage. Il est alimenté en données depuis les bases de production grâce notamment aux outils d'ETL Extract Transform Load.

|  |  |  |
| --- | --- | --- |
| **Étude de besoins spécifiques** | **Modélisation** | **Intégration des données** |
| Contexte  Présentation du sujet  Objectifs (Indicateurs à mesurer)  Données utilisées (Sources)  Indicateurs de risques | Modèle  Dimensions (Axes)  Faits (Mesures)  Granularité  Volumétrie | ETL (Extract Transform Load)  Base de données |

## Contexte

Dans une époque où le mot « **terrorisme** » est sur les lèvres de chacun et dans laquelle les craintes sont de plus en plus nombreuses, nous avons souhaité comprendre et analyser toutes les attaques référencées.

Tout le monde comprend ce mot mais personne ne sait l’expliquer :

Le terrorisme est l'emploi de **la terreur à des fins politiques, religieuses ou idéologiques.**

Les multiples définitions varient sur :

* L'usage de la **violence** (certaines comprennent des groupes n'utilisant pas la violence mais ayant un discours radical), les techniques utilisées, la nature du sujet (mettant à part le terrorisme d'État)
* L'usage de la **peur**
* Le niveau d'organisation
* L’**idéologie**
* Etc.

Dans nombre de définitions intervient aussi **le critère de la victime** du terrorisme (civile, désarmée, innocente, attaque contre un État démocratique qui aurait permis au terroriste de s'exprimer légalement, etc.).

Un grand nombre d'organisations politiques ou criminelles ont cependant recouru au terrorisme pour faire avancer leur cause ou en retirer des profits. Des partis de gauche comme de droite, des groupes nationalistes, religieux ou révolutionnaires, voire des États, ont commis des actes de terrorisme.

Une constante du terrorisme est **l'usage indiscriminé de la violence meurtrière à l'égard de civils** dans le but de promouvoir un groupe, une cause ou un individu, ou encore de pratiquer l'extorsion à large échelle (mafias, cartels de la drogue, etc.).

## Présentation du sujet

Blabla, nous sommes … Nous voulons … Nous avons …

Données :

<https://www.kaggle.com/ananthreddy/global-terrorism-database/data>

## Objectifs

## Données utilisées

Afin de répondre aux besoins de l’agence des renseignements, nous avons utilisé des données récoltées sur Kaggle, au lien suivant :

[**https://www.kaggle.com/ananthreddy/global-terrorism-database/data**](https://www.kaggle.com/ananthreddy/global-terrorism-database/data)

Les données sont fournies **en format .CSV**.

Elles sont composées de **plus de 170 350 lignes** et **plus de 70 colonnes**.

La taille du fichier est égale à **150 Mo**.

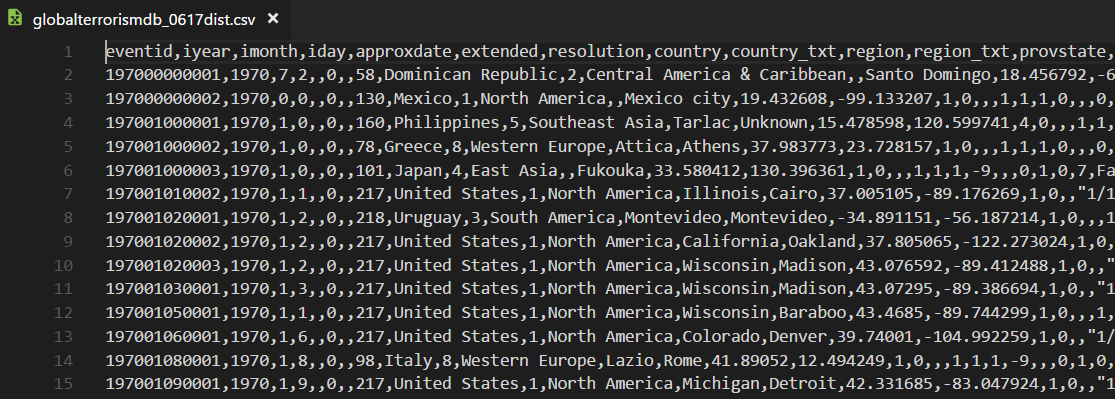
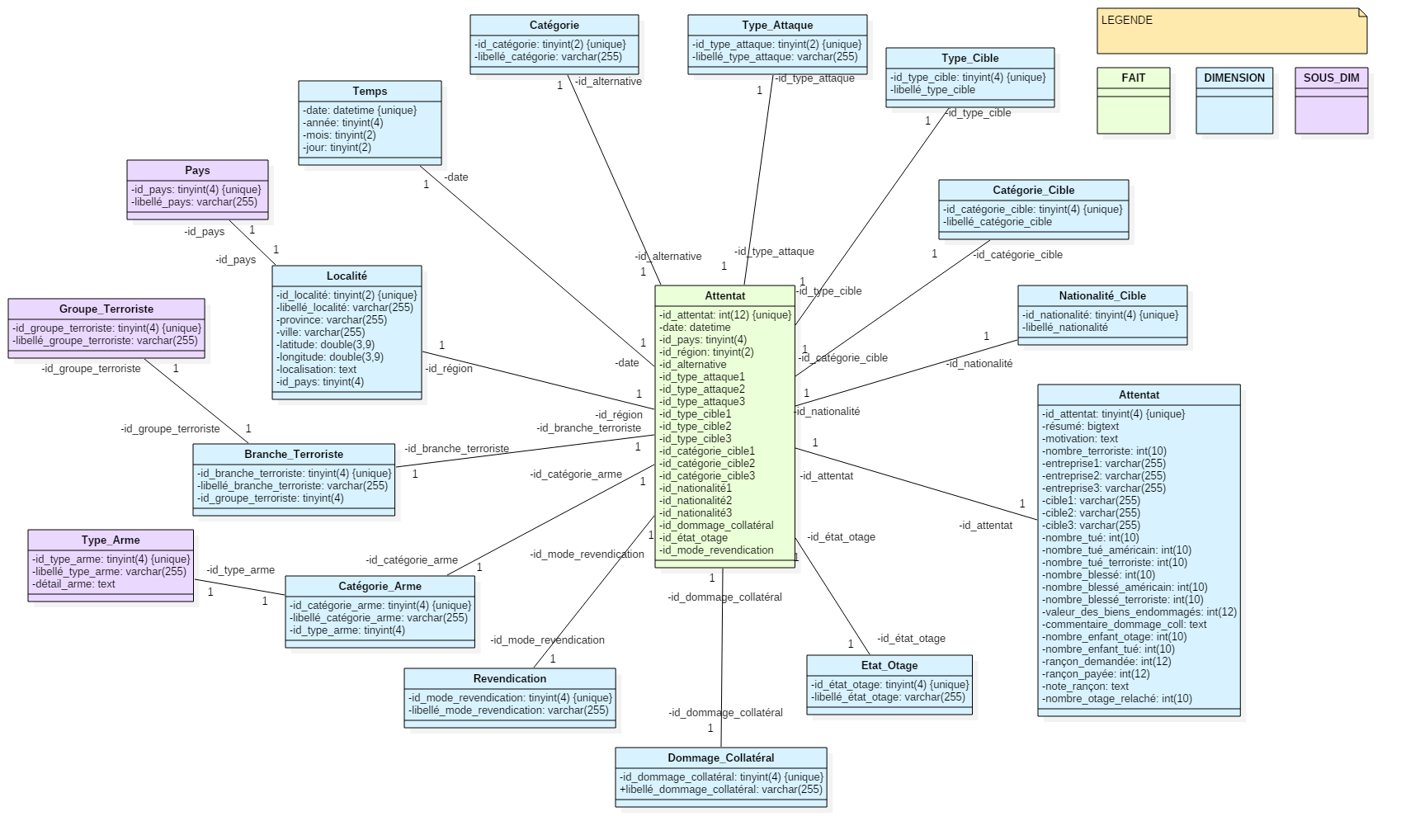


Figure - Jeu de données pour le terrorisme

# Modélisation

## Proposition du modèle



# Intégration

## Outil d’intégration ETL

## Base de données