Corentin MAPELLA & Glodie TANDU

**Projet EDD**

**Young People Survey** : Quels sont les goûts, intérêts et les peurs des jeunes ?

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 :** | **YOUNG PEOPLE SURVEY** |
| **Groupe :** | Corentin MAPELLA – Glodie TANDU |
| **Version du document :** | 0.3 |
| **Statut :** | Contextualisation du projet |

# 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 |

# 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** |
| Périmètre  Objectifs (Indicateurs à mesurer)  Besoins décisionnels  Données utilisées (Sources)  Indications des risques | Modèle  Dimensions (Axes)  Faits (Mesures)  Granularité  Volumétrie | ETL (Extract Transform Load)  Base de données |

## Périmètre

En 2013, des étudiants en Statistiques à l’Université de FSEV UK en Slovaquie ont demandé à 1000 jeunes de participer à une enquête.

Sous forme électronique et papier, toutes les informations récoltées lors des enquêtes ont été **retranscrites en Anglais** et mises à disposition de tout le monde :

* Le fichier de données comprend **1010 lignes** et **150 colonnes** (11 thèmes et 139 questions).
* Certaines données sont **manquantes mais ne seront pas comblées**.
* La population étudiée regroupe exclusivement des jeunes entre **15 ans et 30 ans de nationalité Slovaque**.

Les thèmes abordés sont les suivants :

* Musiques (19 questions)
* Films (12 questions)
* Hobbits et intérêts (32 questions)
* Phobies (10 questions)
* Habitudes de santé (3 questions)
* Traits de personnalités, opinions (57 questions)
* Habitudes de dépenses (7 questions)
* Démographique (10 questions)

## Objectifs

## Données utilisées

# Modèles

## Proposition du modèle

## Correction du modèle

## Définition des dimensions et des faits

# Intégration

## Outil d’intégration ETL