**Rapport**

**1. Introduction**

Ce rapport présente le processus d'exportation des données d'accidents de New York à partir d'un site web dédié aux statistiques d'accidents, suivi de l'alimentation de ces données dans une base de données SQL Server en utilisant SQL Server Integration Services (SSIS). L'objectif principal était de centraliser et d'analyser les données d'accidents pour permettre une meilleure compréhension des tendances et des facteurs associés.

**2. Outils**



**2. Extraction des Données depuis le Site Web**

Les données d'accidents ont été extraites à partir du site web [**https://data.cityofnewyork.us**](https://data.cityofnewyork.us). Le site fournit des données structurées sous forme de fichiers Excel, comprenant des détails sur les incidents tels que la date, l'emplacement, le type d'accident, les conditions météorologiques, etc.

**3. Alimentation de la Base de Données SQL Server avec SSIS**

Pour alimenter la base de données SQL Server, SSIS a été utilisé en raison de sa capacité à gérer efficacement les flux de données et les transformations. Les étapes suivantes ont été réalisées :

- ***Connexions*** : Configuration des connexions sources pour le fichier Excel et de destination pour la base de données SQL Server.

- ***Flux de Données*** : Conception d'un flux de données dans SSIS pour transférer les données de l'Excel vers la base de données SQL Server. Des transformations telles que la conversion de types de données et la suppression des doublons ont été appliquées pendant ce processus.

A screenshot of a computer

Description automatically generated

**4. Validation et Contrôle Qualité**

Des procédures de validation ont été mises en place pour garantir l'intégrité et la qualité des données transférées. Cela inclut la vérification des contraintes des contrôles de qualité spécifiques aux données (par exemple, valeurs aberrantes, données manquantes).

A screenshot of a computer

Description automatically generated

**5. Résultats et Performances**

Le processus d'alimentation a été couronné de succès, avec un volume important de données d'accidents transférées dans la base de données SQL Server. En termes de performances, le temps nécessaire pour compléter le processus était raisonnable compte tenu du volume de données traitées. Des tests de charge ont été effectués pour évaluer les performances du système dans des conditions de charge élevée.

A screenshot of a computer

Description automatically generated