Une image contenant texte, Police, Graphique, logo

Description générée automatiquement

**Rapport SID - 2**

SAE Intégration de données dans un DW

Sujet choisi : Météo

Adèle DALLE

Abinaya SUBRAMANIAM

John-Etan UZAN

Celia ZUK

BUT SD2 EMS FI

2024-2025

Table des matières

[Londres - ANGLETERRE 3](#_Toc182643118)

[Paris - FRANCE 3](#_Toc182643119)

[Rome - ITALIE 6](#_Toc182643120)

[Berne - SUISSE 8](#_Toc182643121)

Nous avons choisi le thème de la météo en nous concentrant plus particulièrement sur Paris, Londres, Rome et Berne. Nous avons sélectionné quatre bases de données (une par ville) sur les données météorologiques annuelles de l’année 2020. Nos différentes bases de données proviennent du site [www.kaggle.com](http://www.kaggle.com), meteo.data.gouv.fr. Pour analyser les variables et réaliser un tableau de bord nous utiliserons principalement le logiciel Excel. Nous pensons aussi réaliser des courbes pour comparer les différentes données présentes dans chaque base sélectionnée.

# Londres - ANGLETERRE

Londres est la capitale du Royaume-Unis et est située au sud de l’Angleterre. Pour étudier sa météo en 2020, nous allons utiliser neuf variables :

A screenshot of a computer

Description automatically generated

Ces variables ont été relevées quotidiennement de 1979 à 2020. Pour ce projet, nous nous concentrerons davantage sur les données de 2020, cela nous apportera une meilleure précision de la météo Londonienne lors de cette année.

# Paris - FRANCE

Pour analyser et interpréter les données météorologiques de la France, nous avons choisi une base de données issu de la capitale, c’est-à-dire Paris

Voici les variables que nous allons utiliser :

Une image contenant texte, capture d’écran, nombre, Police

Description générée automatiquement

Contenu des données :

Notre base de données contient en tout 17 variables étudiées sur de nombreuses dates. Cette base de données prend également en compte des données de plusieurs postes différents, ce qui nous permet ainsi d’avoir une comparaison davantage précise.

Notre base de données est riche en informations car elle contient divers aspects météorologiques comme le nombre de jours avec des précipitations dépassant un certains seuils, la latitude, la longitude, ou même l’altitude.

Périodicité :

Ces données sont représentées mensuellement. On pourra y trouver les année et les mois des informations.

Moyen d’accès aux données :

Ces données ont été récupéré sur « meteo.data.gouv.fr »

Droit de ré-utilisation des données :

Ces données proviennent d’une source publique. Elles peuvent ainsi être ré-utilisées.

# Rome - ITALIE

Concernant les données de l’Italie, nous avons décidé d’utiliser les données climatiques quotidiennes de la capitale, Rome. Ce jeu de données inclut les relevés météorologiques de Rome de chaque jour de l’année 2020, allant du **1er janvier au 31 décembre** (année bissextile avec 366 jours). Ces données permettent une analyse détaillée des variations climatiques dans cette ville au climat méditerranéen.

Les données utilisées sont des données libres de droit, issues de la plateforme publique **historique-meteo.net**, et sont accessible et peuvent être téléchargées sous format CSV directement via le site web.

Pour les variables, nous avons choisi d’en sélectionner seulement 10, la date, les températures minimales et maximales en °C, les précipitations en mm, le taux d’humidité en %, la vitesse du vent en km/h, la couverture nuageuse en %, le nombre d’heures ensoleillé, la quantité de neige en mm et l’indice de chaleur maximal.

Tableau des variables :

Une image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement

Concernant le nettoyage de la base de données sur Power BI, nous avons supprimé les colonnes qui n’étaient pas utiles afin de garder que les 10 variables sélectionnées et d’obtenir une base de données propre avec 10 colonnes et 366 lignes. Nous avons également vérifié s’il n’y avait pas de valeur manquante.

Base de données :

Une image contenant texte, nombre, capture d’écran

Description générée automatiquement

Par la suite, nous devrons réaliser des représentations graphiques, afin de visualiser les données, comme par exemple, faire des graphiques pour représenter les températures minimales et maximales pour pouvoir repérer les piques de chaleur et de froid. Nous pourrons aussi faire des graphiques pour représenter la répartition des jours de pluie de l’année 2020 et la répartition de la vitesse du vent ou des niveaux d’humidité. Enfin, nous aurons aussi la possibilité de représenter un graphique pour représenter le nombre d’heures d’ensoleillement pour pouvoir repérer les différences entre les différentes saisons.

Ces représentations, nous permettrons de comprendre les tendances climatiques de Rome en 2020.

# Berne - SUISSE

Pour la Suisse, nous avons sélectionné les données météorologiques de la ville de Berne, la capitale. Ce jeu de données couvre la période du **1er janvier au 31 décembre 2020**, avec des mesures effectuées quotidiennement donc chaque observation représente un jour, ce qui donne un total de **366 enregistrements** pour l’année 2020, en tenant compte de l’année bissextile. Les informations proviennent de la plateforme **Meteostat**, accessible via le lien suivant : <https://meteostat.net/fr/place/ch/bern?s=06631&t=2020-01-01/2024-05-24> . Ces données sont ouvertes, librement accessibles et réutilisables.

Voici le tableau qui récapitule les variables de notre jeu de données.



Après cela, nous avons importé la base de données sur Power BI afin d’analyser la distribution de ces variables.