



Présentation du projet de Data visualisation

PROJET P8 : FAITES UNE
ÉTUDE SUR L'EAU POTABLE



PARCOURS DATA ANALYST_V2

ADELINE LE RAY

Sommaire



Contexte du projet



Conception du tableau de bord



Pré-traitement des données



Blueprint



Mockup



Tableau de bord

Contexte du projet

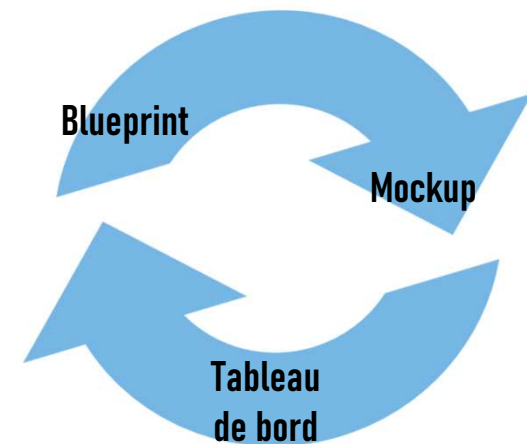
- **L'accès à l'eau potable dans le monde**
- **Drinking Water For All (DWFA)** est une ONG qui intervient dans 3 domaines d'expertises :
 - la création de services d'accès à l'eau potable
 - la modernisation des services
 - le consulting à propos des politiques d'accès à l'eau
- **Mission :** Réaliser un tableau de bord permettant :
 - d'identifier les pays qui rencontrent des difficultés d'accès à l'eau potable
 - de choisir le pays à cibler en fonction du domaine d'expertise choisi par le bailleur de fond.

Conception du tableau de bord

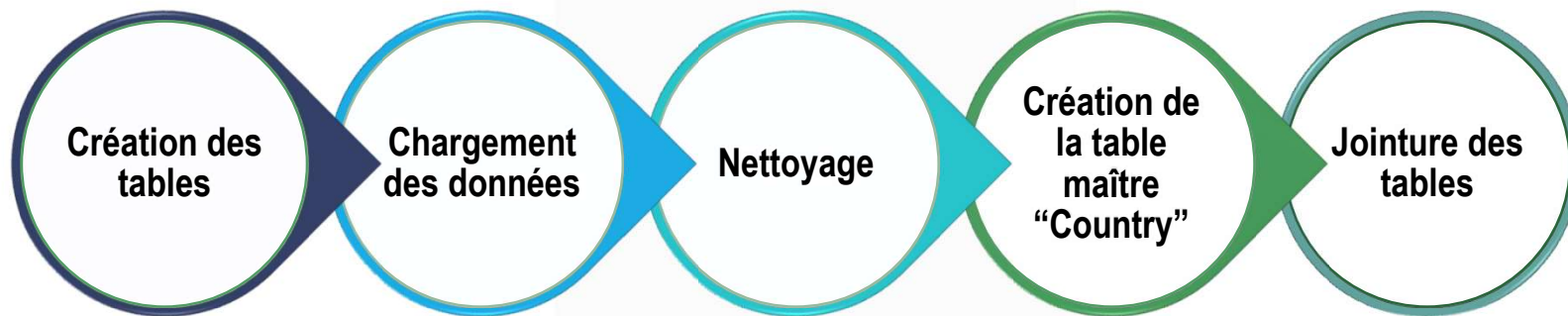
Les étapes clés de la conception du tableau de bord sont :

- **Pré-traitement des données**
- **Blueprint** : clarification les objectifs de la visualisation (besoin utilisateurs) et de définition des éléments clés
- **Mockup** : réalisation d'un prototype des différentes vues du tableau de bord sous forme de schéma
- **Tableau de bord** : réalisation du tableau de bord dans l'outil de data viz

Démarche itérative



Pré-traitement des données avec



Plus de détails en annexe

Blueprint

■ Domaine 1 : Création de services d'accès à l'eau :

- Part de la population ayant accès aux services basiques d'eau potable faible
- Taux de population urbaine élevé
- Taux de mortalité due à l'eau insalubre élevé

■ Domaine 2 : Modernisation de services d'accès à l'eau déjà existants :

- Part de la population ayant accès aux services basiques d'eau potable élevé
- Part de la population ayant accès aux infrastructures de qualité faible
- Ecart entre taux d'accès aux services basiques et infrastructures de qualité
- Taux de mortalité due à l'eau insalubre élevé

■ Domaine 3 : Consulting auprès d'administrations/gouvernements à propos des politiques d'accès à l'eau :

- Part de la population ayant accès aux services basiques d'eau potable élevé
- Taux de mortalité due à l'eau insalubre faible
- Stabilité politique élevé

2. Blueprint – Vue mondiale

Besoin / Question utilisateurs	Interactions utilisateurs	Mesures spécifiques à utiliser	Traitement	Visualisation
Quelle est la population à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Année (liste choix unique) Stabilité politique (curseur) 	<ul style="list-style-type: none"> Population Année Granularité : total 	Agrégation par la somme	Chiffre
Quelle est la part de la population ayant accès aux services basiques à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Année (liste choix unique) Stabilité politique (curseur) 	<ul style="list-style-type: none"> Taux de services basiques Année Granularité: total 	Agrégation par la moyenne	Chiffre
Quelle est la part de population ayant accès aux infrastructures de qualité à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Année (liste choix unique) Stabilité politique (curseur) 	<ul style="list-style-type: none"> Taux de services « safely managed » Année Granularité: total 	Agrégation par la moyenne	Chiffre
Quel est le taux de mortalité moyen à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Stabilité politique (curseur) 	<ul style="list-style-type: none"> Taux de mortalité Année : 2016 Granularité: total 	Agrégation par la moyenne	Chiffre
Quel est le nombre de morts dues à l'eau insalubre à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Stabilité politique (curseur) 	<ul style="list-style-type: none"> Nombre de morts dues à l'eau insalubre Année : 2016 Granularité: total 	Agrégation par la somme	Chiffre
Quelle est la stabilité politique moyenne à l'échelle du monde ?	Avoir la possibilité de filtrer sur : <ul style="list-style-type: none"> Année (liste choix unique) Stabilité politique (curseur) 	<ul style="list-style-type: none"> Stabilité politique Année Granularité: total 	Agrégation par la moyenne	Chiffre

Mockup

- Prototype de visualisation réalisé à partir du blueprint
- Permet de structurer la présentation du tableau de bord

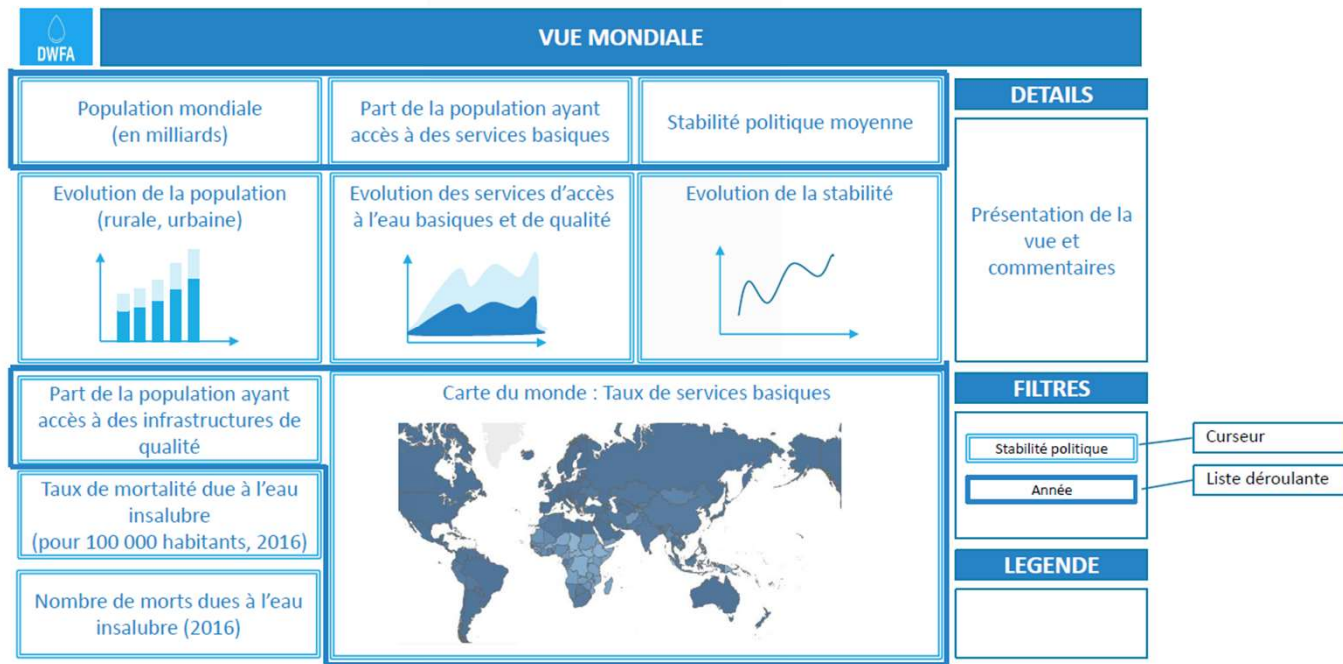


Tableau de bord

- Outil de visualisation utilisé : Tableau Desktop version 2022.4 (licence étudiant)
- Liaison directe BDD MySQL et Tableau Desktop via MySQL server



- Retour sur Tableau :
 - Interface intuitive
 - Nombreuses aides pour les alternatives de graphiques
 - Apprentissage de l'outil nécessaire : beaucoup de « trucs et astuces » pour réaliser des visualisations pertinentes et les rendre dynamiques.
- Tableau de bord sur Tableau Public :

https://public.tableau.com/shared/ZP8PSK492?:display_count=n&:origin=viz_share_link

Merci

DES QUESTIONS ?

Annexes



- Pré-traitement des données avec MySQL Workbench

Pré-traitements des données avec



Création
des tables

Chargement
des
données

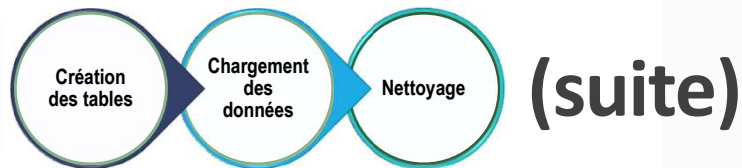
Nettoyage

```
1  #Création de la table
2  • DROP TABLE IF EXISTS dw_services;
3  • CREATE TABLE dw_services(
4      YEAR YEAR NOT NULL COMMENT '2000->2017',
5      COUNTRY VARCHAR(100) NOT NULL,
6      GRANULARITY VARCHAR(10) NOT NULL COMMENT 'Total, Urban, Rural',
7      BASIC VARCHAR(20) COMMENT 'Population using at least basic drinking-water services (%)',
8      SAFELY_MANAGED VARCHAR(20) COMMENT 'Population using safely managed drinking-water services (%)',
9      PRIMARY KEY(COUNTRY, YEAR, GRANULARITY)
10 );
11 #Importation des données
12 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/BasicAndSafelyManagedDrinkingWaterServices.csv'
13 INTO TABLE dw_services
14 FIELDS TERMINATED BY ','
15 OPTIONALLY ENCLOSED BY '"'
16 LINES TERMINATED BY '\r\n'
17 IGNORE 1 ROWS;
```

```
19 #Création de la table
20 • DROP TABLE IF EXISTS region;
21 • CREATE TABLE region(
22     CONTINENT VARCHAR(100) NOT NULL,
23     COUNTRY VARCHAR(100) NOT NULL,
24     PRIMARY KEY(COUNTRY)
25 );
26 #Importation des données
27 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/RegionCountry.csv'
28 INTO TABLE region
29 FIELDS TERMINATED BY ','
30 OPTIONALLY ENCLOSED BY '"'
31 LINES TERMINATED BY '\r\n'
32 IGNORE 1 ROWS;
```

```
35 #Création de la table
36 • DROP TABLE IF EXISTS population;
37 • CREATE TABLE population(
38     COUNTRY VARCHAR(100) NOT NULL,
39     GRANULARITY VARCHAR(100) NOT NULL,
40     YEAR YEAR NOT NULL,
41     POPULATION FLOAT,
42     PRIMARY KEY(COUNTRY, YEAR, GRANULARITY)
43 );
44 #Importation des données
45 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/Population.csv'
46 INTO TABLE population
47 FIELDS TERMINATED BY ','
48 OPTIONALLY ENCLOSED BY '"'
49 LINES TERMINATED BY '\r\n'
50 IGNORE 1 ROWS;
51 #Correction du nom 'Republic of North Macedonia' en 'North Macedonia' pour cohérence entre les tables
52 • UPDATE population
53 SET COUNTRY = REPLACE(COUNTRY, 'North Macedonia', 'Republic of North Macedonia')
54 WHERE COUNTRY LIKE 'North Macedonia';
55
56 #Création de la table
57 • DROP TABLE IF EXISTS mortality;
58 • CREATE TABLE mortality(
59     YEAR YEAR NOT NULL,
60     COUNTRY VARCHAR(100) NOT NULL,
61     GRANULARITY VARCHAR(10) NOT NULL,
62     MORTALITY_RATE_UNSAFE_WATER VARCHAR(20) COMMENT 'Mortality rate attributed to exposure to unsafe WASH services',
63     DEATHS VARCHAR(20) COMMENT 'Number of deaths attributed to exposure to unsafe WASH services',
64     PRIMARY KEY(COUNTRY, YEAR, GRANULARITY)
65 );
66 #Importation des données
67 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/MortalityRateAttributedToWater.csv'
68 INTO TABLE mortality
69 FIELDS TERMINATED BY ','
70 OPTIONALLY ENCLOSED BY '"'
71 LINES TERMINATED BY '\r\n'
72 IGNORE 1 ROWS;
```

Pré-traitements des données avec



```
76 #Création de la table
77 • DROP TABLE IF EXISTS political_stability;
78 • CREATE TABLE political_stability(
79     COUNTRY VARCHAR(100) NOT NULL,
80     YEAR YEAR NOT NULL,
81     STABILITY FLOAT,
82     GRANULARITY VARCHAR(10) NOT NULL,
83     PRIMARY KEY(COUNTRY,YEAR)
84 );
85 #Importation des données
86 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/PoliticalStability.csv'
87 INTO TABLE political_stability
88 FIELDS TERMINATED BY ','
89 OPTIONALLY ENCLOSED BY '"'
90 LINES TERMINATED BY '\r\n'
91 IGNORE 1 ROWS;
92 #Correction du nom 'Republic of North Macedonia' en 'North Macedonia' pour cohérence entre les tables
93 • UPDATE political_stability
94 SET COUNTRY = REPLACE(COUNTRY, 'North Macedonia','Republic of North Macedonia')
95 WHERE COUNTRY LIKE 'North Macedonia';
```

```
98 #Création de la table
99 • DROP TABLE IF EXISTS surface_area;
100 • CREATE TABLE surface_area(
101     #YEAR 2020,
102     COUNTRY VARCHAR(100) NOT NULL,
103     SURFACE VARCHAR(20) COMMENT "Superficie pays en 2020",
104     PRIMARY KEY(COUNTRY)
105 );
106 #Importation des données
107 • LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/DAN-P8-donnees/P8_surface_area.csv'
108 INTO TABLE surface_area
109 FIELDS TERMINATED BY ','
110 OPTIONALLY ENCLOSED BY '"'
111 LINES TERMINATED BY '\r\n'
112 IGNORE 1 ROWS;
113 #Correction des noms de pays pour cohérence entre les tables
114 • UPDATE surface_area
115 SET COUNTRY = REPLACE(COUNTRY, 'Bahamas, The','Bahamas')
116 WHERE COUNTRY LIKE 'Bahamas, The';
117 • UPDATE surface_area
118 SET COUNTRY = REPLACE(COUNTRY, 'North Macedonia','Republic of North Macedonia')
119 WHERE COUNTRY LIKE 'North Macedonia';
120 • UPDATE surface_area
121 SET COUNTRY = REPLACE(COUNTRY, 'Bolivia','BOLIVIA (PLURINATIONAL STATE OF)')
122 WHERE COUNTRY LIKE 'Bolivia';
```

Etc. ...

Pré-traitements des données avec



Création
de la table
maître
"Country"

Jointure
des tables

Table maître « Country »

```
1 • DROP TABLE IF EXISTS country;
2
3 • CREATE TABLE country AS (
4     SELECT UPPER(COUNTRY) as COUNTRY FROM population
5     UNION DISTINCT
6     SELECT UPPER(COUNTRY) as COUNTRY FROM region
7     UNION DISTINCT
8     SELECT UPPER(COUNTRY) as COUNTRY FROM political_stability
9     UNION DISTINCT
10    SELECT UPPER(COUNTRY) as COUNTRY FROM mortality
11    UNION DISTINCT
12    SELECT UPPER(COUNTRY) as COUNTRY FROM surface_area
13 );
```

Jointures

```
1 • DROP TABLE IF EXISTS WASH_data;
2
3 • CREATE TABLE WASH_data AS (
4     SELECT c.COUNTRY, CONTINENT, p.YEAR, p.GRANULARITY, POPULATION, CAST(SURFACE AS FLOAT) AS SURF_AREA,
5           CAST(MORTALITY_RATE_UNSAFE_WATER AS FLOAT) AS MORTALITY_RATE_UNSAFE_WATER,
6           CAST(DEATHS AS FLOAT) AS DEATHS, STABILITY, CAST(BASIC AS FLOAT) AS BASIC, CAST(SAFELY_MANAGED AS FLOAT) AS SAFELY_MANAGED
7     FROM COUNTRY c
8     LEFT OUTER JOIN population p ON c.COUNTRY = UPPER(p.COUNTRY)
9     LEFT OUTER JOIN region r ON c.COUNTRY=UPPER(r.COUNTRY)
10    LEFT OUTER JOIN mortality m ON c.COUNTRY=UPPER(m.COUNTRY) AND p.YEAR = m.YEAR AND p.GRANULARITY = m.GRANULARITY
11    LEFT OUTER JOIN political_stability s ON c.COUNTRY=UPPER(s.COUNTRY) AND p.YEAR = s.YEAR AND p.GRANULARITY = s.GRANULARITY
12    LEFT OUTER JOIN dw_services dw ON c.COUNTRY=UPPER(dw.COUNTRY) AND p.YEAR = dw.YEAR AND p.GRANULARITY = dw.GRANULARITY
```