



# Étude sur l'eau potable

## Compte-rendu de lancement

# Sommaire

## Introduction – démarche d'élaboration du dashboard

### 1 – Contexte de l'étude

1.1 – Audience & blueprint

1.2 – Objectifs & mock-up

### 2 – Design & alimentation de la base de données

2.1 – Sources & format des données

2.2 – Base de données : structure & implémentation

2.3 – Chargement, nettoyage & ré-export des données

2.4 – Choix de l'outil de visualisation

### 3 – Analyse & interprétation des résultats

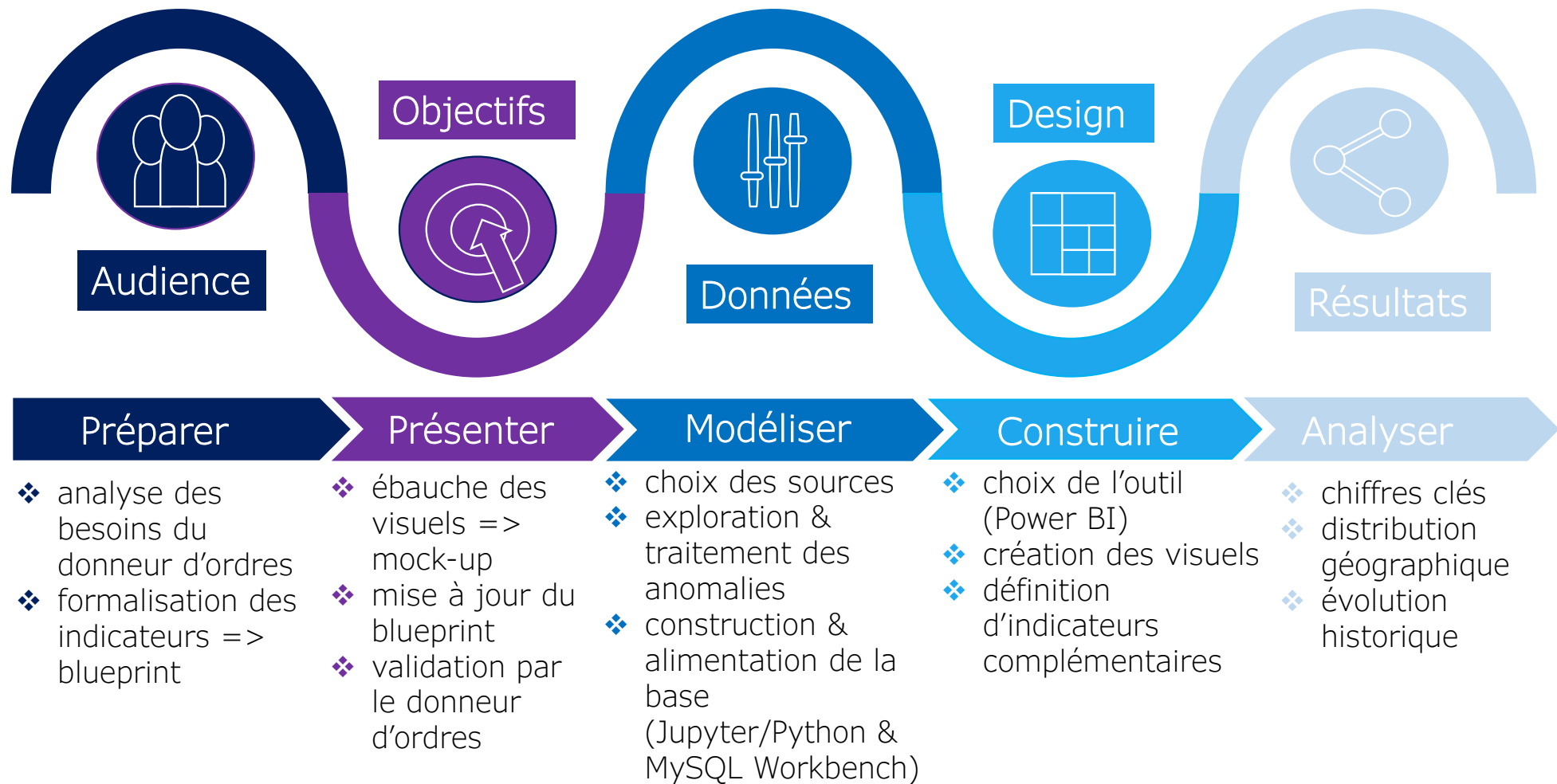
3.1 – Vue mondiale

3.2 – Vue régionale

3.3 – Vue nationale

## Conclusion & prochaines étapes

# Introduction - démarche d'élaboration du dashboard



# 1 – Contexte de l'étude

## 1.1 – Audience & blueprint

- ❖ Public : bailleur de fonds
- ❖ 3 domaines d'expertise à valoriser:
  - Création de services d'accès à l'eau potable
  - Modernisation de services d'accès à l'eau potable existants
  - Consulting auprès d'administrations/gouvernements sur les politiques d'accès à l'eau potable
- ❖ Identification d'indicateurs
  - Pertinents à ces 3 domaines
  - Sur 3 niveaux mondial, régional et national

Besoin	Utilisateurs	Mesures spécifiques à utiliser	Visualisation	Onglet
Domaine 1 : création de services	Évolution dans le temps de la situation mondiale	Indicateur : répartition de la population (rurale/urbaine) vs accès assainissement & eau potable Filtres : monde Agrégation : somme	Stacked bar chart with line plot Dynamique (slider années)	Onglet monde
	Évolution dans le temps de la situation régionale	Indicateur : urbanisation de la population Filtres : région Agrégation : total	Stacked area chart Dynamique	Onglet région
	Vue synthétique de la situation actuelle du pays (année la plus récente disponible) & comparaison avec la situation mondiale moyenne	Indicateur : taux d'urbanisation Filtres : année Agrégation : pays	Donut chart Dynamique	Onglet pays
Domaine 2 : modernisation des services	Évolution dans le temps de la situation mondiale	Indicateur : population totale & rurale, et accès à l'eau potable & assainissement Filtres : monde Agrégation : somme	Line chart Fixe	Onglet monde
	Évolution dans le temps de la situation régionale	Indicateur : Accès eau potable & assainissement Filtres : région Agrégation : total	Line chart Dynamique	Onglet région
	Vue synthétique de la situation actuelle du pays (année la plus récente disponible)	Indicateur : accès eau potable, assainissement, mortalité, populations totale, rurale & urbaine, Filtres : année Agrégation : pays	Chiffres Dynamique	Onglet pays

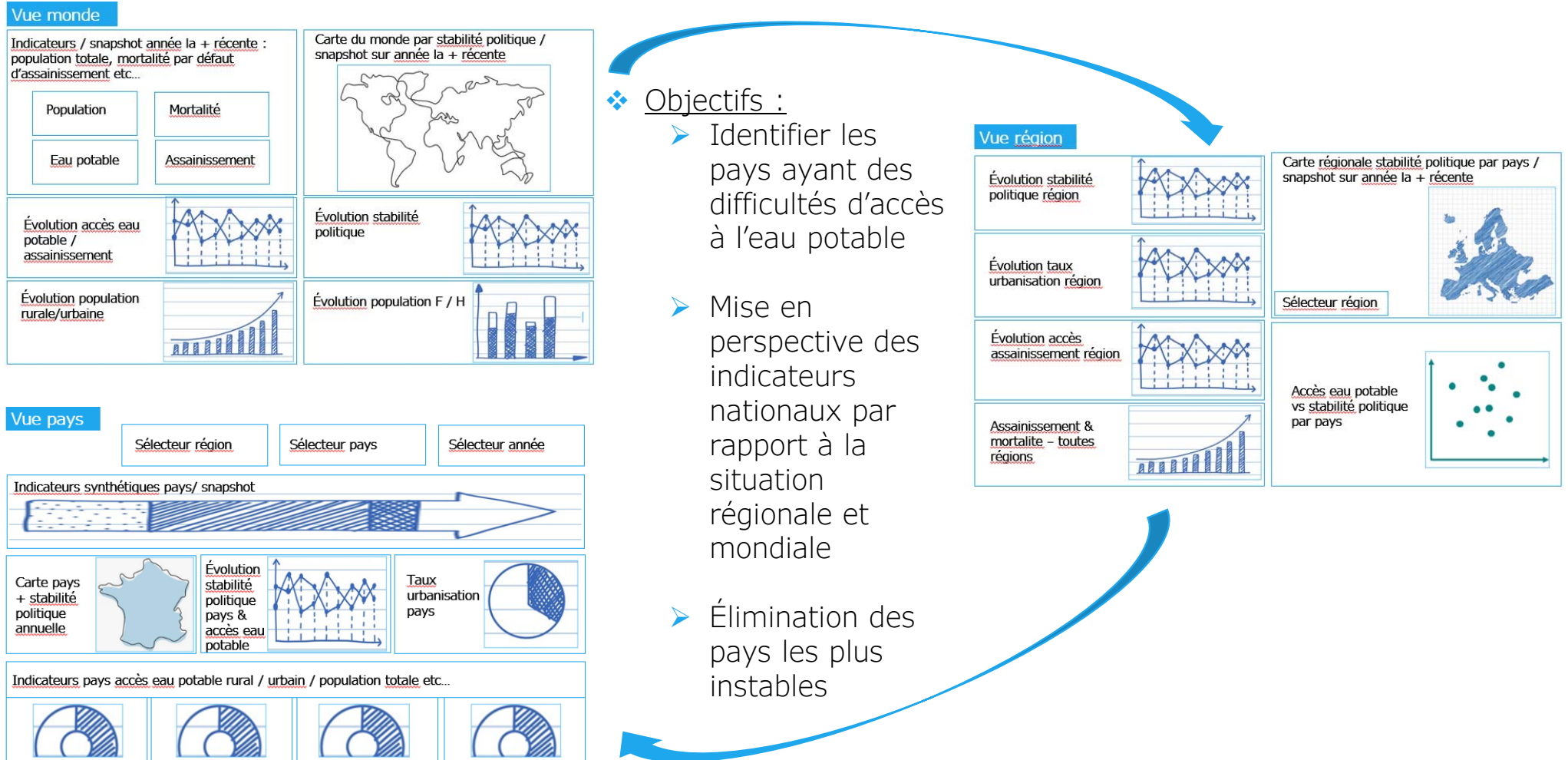
# 1 – Contexte de l'étude

## 1.1 – Audience & blueprint

Domaine 3 : consulting	Vue synthétique de la situation actuelle du pays (année la plus récente disponible)	<u>Indicateur</u> : stabilité politique vs accès a l'eau potable <u>Filtres</u> : année 2016 <u>Agrégation</u> : région & pays	Scatter plot Dynamique	Onglet région	
	Évolution dans le temps de la stabilité politique et son impact sur les infrastructures via l'accès a l'eau potable et aux services d'assainissement	<u>Indicateur</u> : stabilité politique <u>Filtres</u> : année <u>Agrégation</u> : pays	Carte pays Dynamique	Onglet pays	
	Vue synthétique de la situation actuelle du pays (année la plus récente disponible)	<u>Indicateur</u> : Évolution stabilité politique et accès eau potable & assainissement <u>Filtres</u> : année <u>Agrégation</u> : pays	Line chart Dynamique	Onglet pays	
	Vue synthétique de la situation actuelle du pays (année la plus récente disponible)	<u>Indicateur</u> : accès eau potable pays vs moyenne mondiale <u>Filtres</u> : année, pays <u>Agrégation</u> : total	Gauge Dynamique	Onglet pays	
Autres indicateurs		Vue synthétique de la situation actuelle mondiale (année la plus récente disponible)	Indicateurs statistiques synthétiques <u>Indicateur</u> : population totale, mortalité totale, accès eau potable & satiation totaux <u>Filtres</u> : Année 2016, monde <u>Agrégation</u> : sommes	Chiffres Fixe	Onglet monde
		Évolution dans le temps de la situation mondiale	<u>Indicateur</u> : stabilité politique mondiale <u>Filtres</u> : monde <u>Agrégation</u> : moyenne	Line chart Fixe	Onglet monde
		Évolution dans le temps de la situation mondiale	<u>Indicateur</u> : répartition F / H de la population totale <u>Filtres</u> : monde <u>Agrégation</u> : somme	Groupe bar chart Fixe	Onglet monde
		Élimination des pays les plus instables	stabilité politique > -2	Filtre Dynamique	Tous onglets

# 1 – Contexte de l'étude

## 1.2 – Objectifs & mock-up



## 2 – Design & alimentation de la base de données

### 2.1 – Sources & format des données

- ❖ Sources : sites internet FAO, WHO & World Bank
- ❖ Format : fichiers Excel
- ❖ Problèmes : incomplets (granularité / années manquantes) et partiellement pivotés

➤ *Extrait du fichier d'accès à l'eau potable:*

	A	B	C	D	E
1	Year	Country	Granularity	Population using at least basic drinking-water services (%)	Population using safely managed drinking-water services (%)
2	2000	Afghanistan	Rural	21.61913	
3	2000	Afghanistan	Total	27.7719	
4	2000	Afghanistan	Urban	49.48745	
5	2000	Albania	Rural	81.78472	
6	2000	Albania	Total	87.86662	49.29324
7	2000	Albania	Urban	96.35529	

➤ *Extrait du fichier de mortalité:*

	A	B	C	D	E
1	Year	Country	Granularity	Mortality rate attributed to exposure to unsafe WASH services	WASH deaths
2	2016	Afghanistan	Female	15.31193	
3	2016	Afghanistan	Male	12.61297	
4	2016	Afghanistan	Total	13.92067	4824.353
5	2016	Albania	Female	0.12552	
6	2016	Albania	Male	0.2065	
7	2016	Albania	Total	0.16641	4.86975

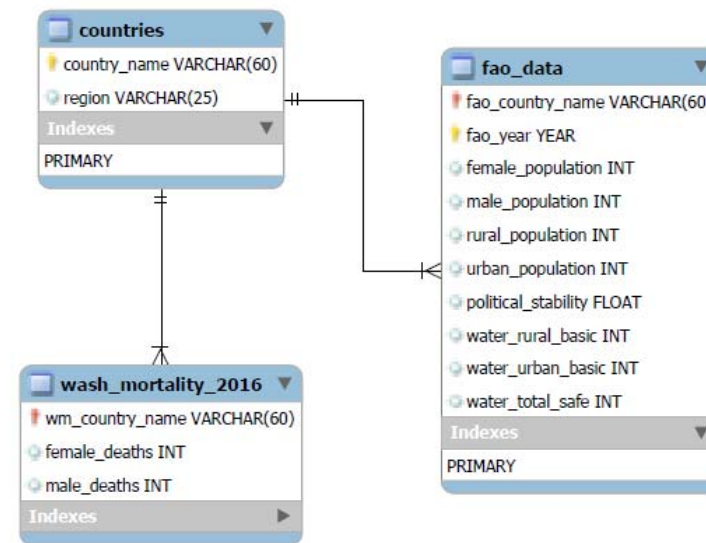
## 2 – Design & alimentation de la base de données

### 2.2 – Base de données : structure & implémentation

❖ Outil retenu : MySQL

❖ Démarche :

- Création du schéma dans MySQL Workbench
- Utilisation de la fonction “Forward Engineer” pour la génération du code et la création de la base et des tables
- Mise en conformité avec la 3NF





## 2 – Design & alimentation de la base de données

### 2.2 – Base de données : structure & implémentation

The image displays two screenshots of the MySQL Workbench interface, illustrating the database structure and the results of a query.

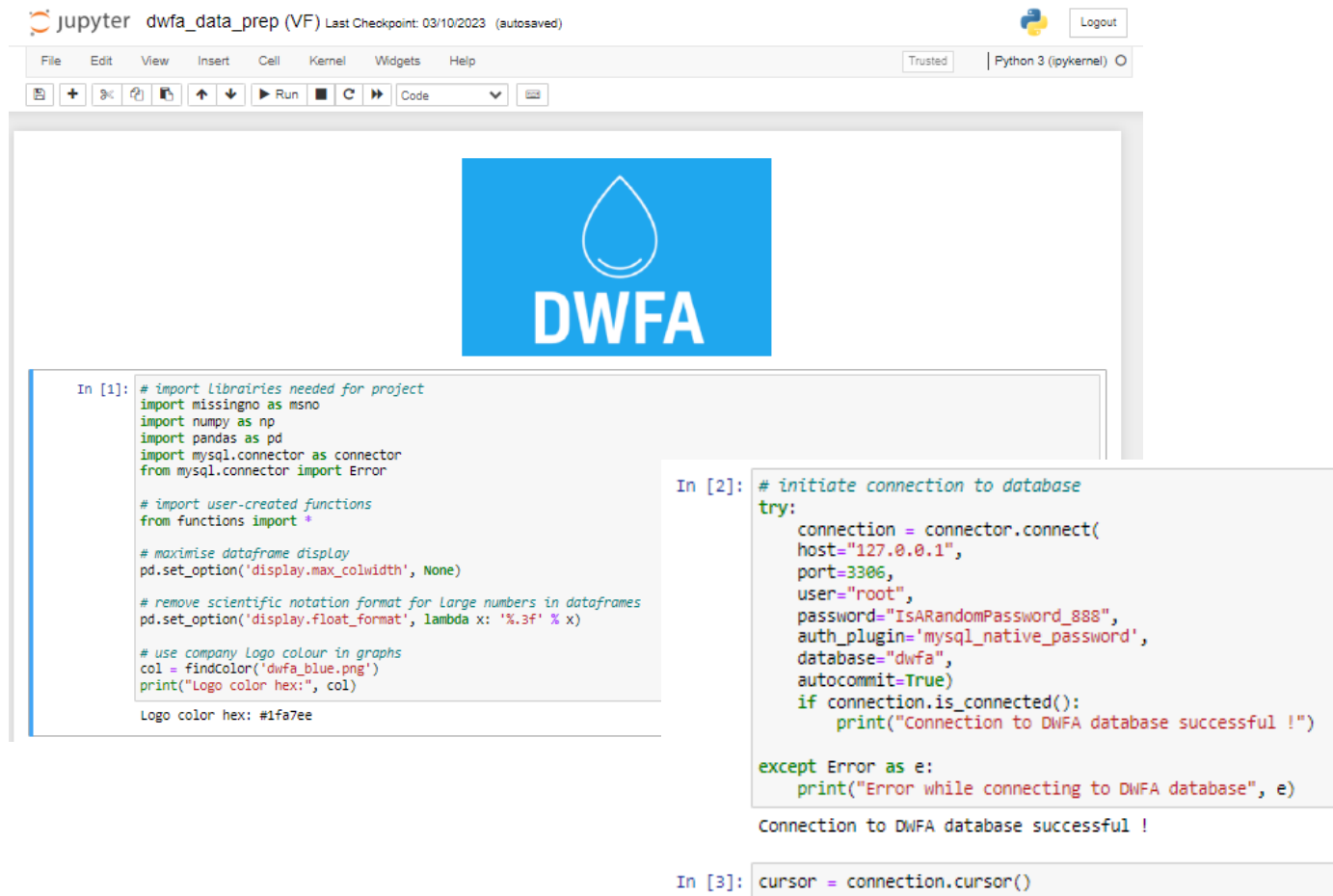
**Left Screenshot:** The 'SCHEMAS' pane on the left shows the database structure. The 'dwfa' schema is expanded, showing tables: 'countries', 'fao\_data', and 'wash\_mortality\_2016'. The 'countries' table is highlighted, showing its columns: 'country\_name', 'region', 'female\_population', 'male\_population', 'rural\_population', 'urban\_population', 'political\_stability', 'water\_rural\_basic', 'water\_urban\_basic', and 'water\_total\_safe'. The 'Result Grid' shows the query results for 'SELECT \* FROM countries;'. The table has two columns: 'country\_name' and 'region'. The data rows are: Afghanistan (Eastern Mediterranean), Albania (Europe), Algeria (Africa), Andorra (Europe), Angola (Africa), Antigua and Barbuda (Americas), and Argentina (Americas).

**Right Screenshot:** This screenshot shows the same MySQL Workbench interface, but with the 'Result Grid' expanded to show the full query results. The table has two columns: 'country\_name' and 'region'. The data rows are: Afghanistan (Eastern Mediterranean), Albania (Europe), Algeria (Africa), Andorra (Europe), Angola (Africa), Antigua and Barbuda (Americas), and Argentina (Americas).

## 2 – Design & alimentation de la base de données

### 2.3 – Chargement, nettoyage & ré-export des données

- ❖ Données Excel importées en Python sur Jupyter Notebook
- ❖ Retraitements : unicité des clés, gestion des valeurs manquantes & redondantes, dé-pivotage des données, suppression des champs calculés, conformité avec la 3NF
- ❖ Création de nouveaux fichiers .csv & export
- ❖ Chargement des fichiers .csv dans la base MySQL avec la librairie MySQL Connector Python



```
jupyter dwfa_data_prep (VF) Last Checkpoint: 03/10/2023 (autosaved)
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) Logout

In [1]: # import libraries needed for project
import missingno as msno
import numpy as np
import pandas as pd
import mysql.connector as connector
from mysql.connector import Error

# import user-created functions
from functions import *

# maximise dataframe display
pd.set_option('display.max_colwidth', None)

# remove scientific notation format for large numbers in dataframes
pd.set_option('display.float_format', lambda x: '%.3f' % x)

# use company logo colour in graphs
col = findColor('dwfa_blue.png')
print("Logo color hex:", col)

Logo color hex: #1fa7ee

In [2]: # initiate connection to database
try:
    connection = connector.connect(
        host="127.0.0.1",
        port=3306,
        user="root",
        password="IsARandomPassword_888",
        auth_plugin='mysql_native_password',
        database="dwfa",
        autocommit=True)
    if connection.is_connected():
        print("Connection to DWFA database successful !")
except Error as e:
    print("Error while connecting to DWFA database", e)

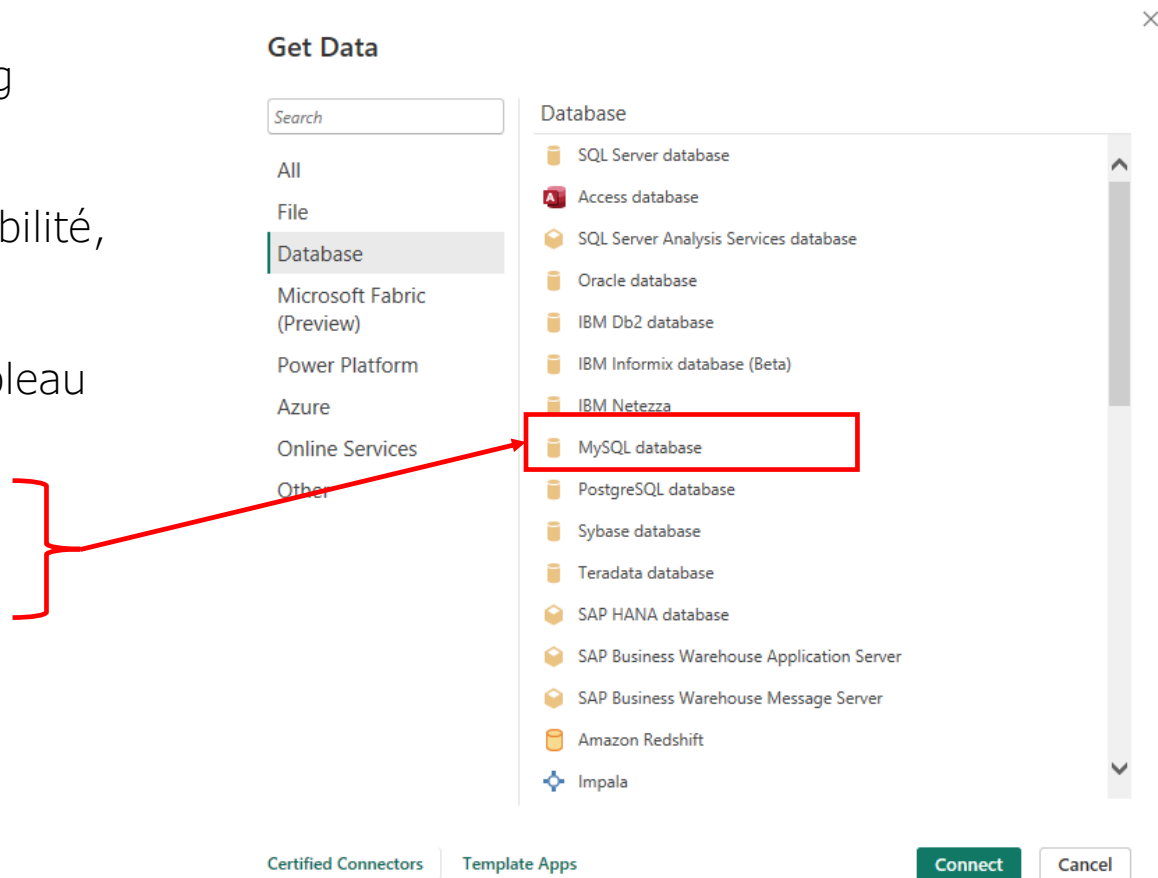
Connection to DWFA database successful !

In [3]: cursor = connection.cursor()
```

## 2 – Design & alimentation de la base de données

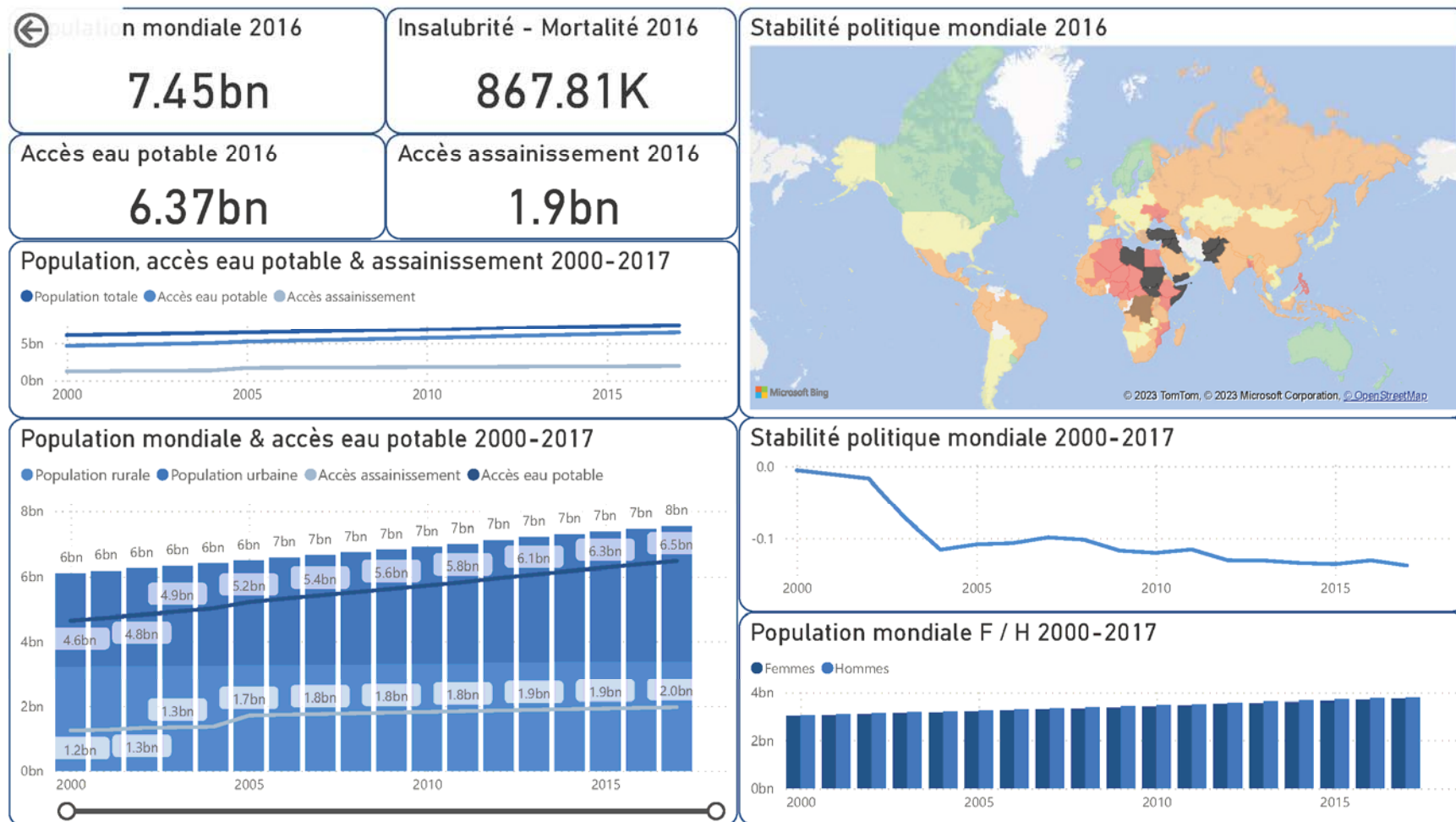
### 2.4 – Choix de l'outil de visualisation

- ❖ Besoin d'un outil de dashboarding spécialisé
- ❖ Critères: fonctionnalité, prix, stabilité, popularité, ergonomie
- ❖ Marché diversifié dominé par Tableau et Power BI
  - Power BI retenu
  - Connexion directe à la base MySQL



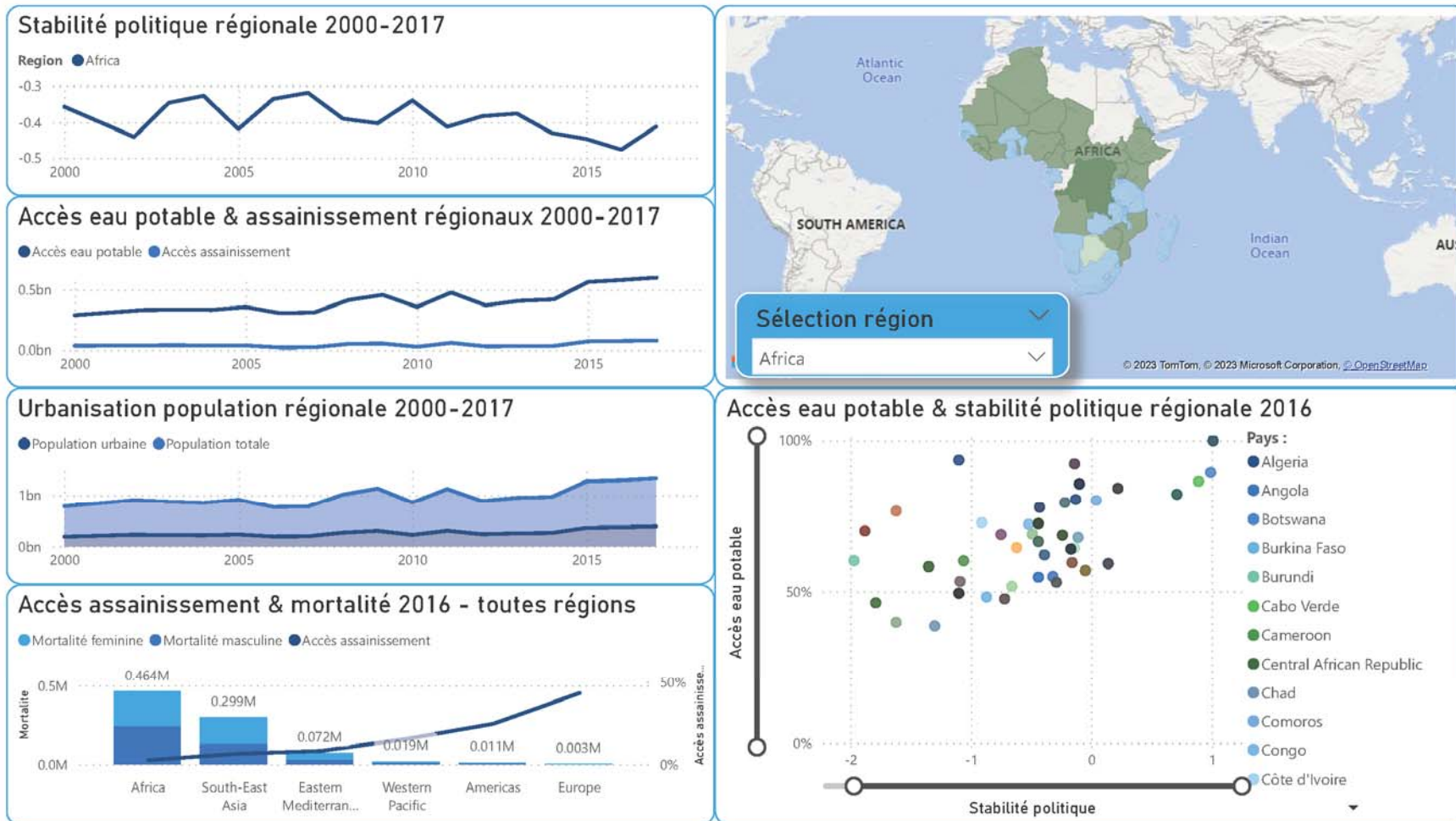
### 3 – Analyse & interprétation des résultats

#### 3.1 – Vue mondiale



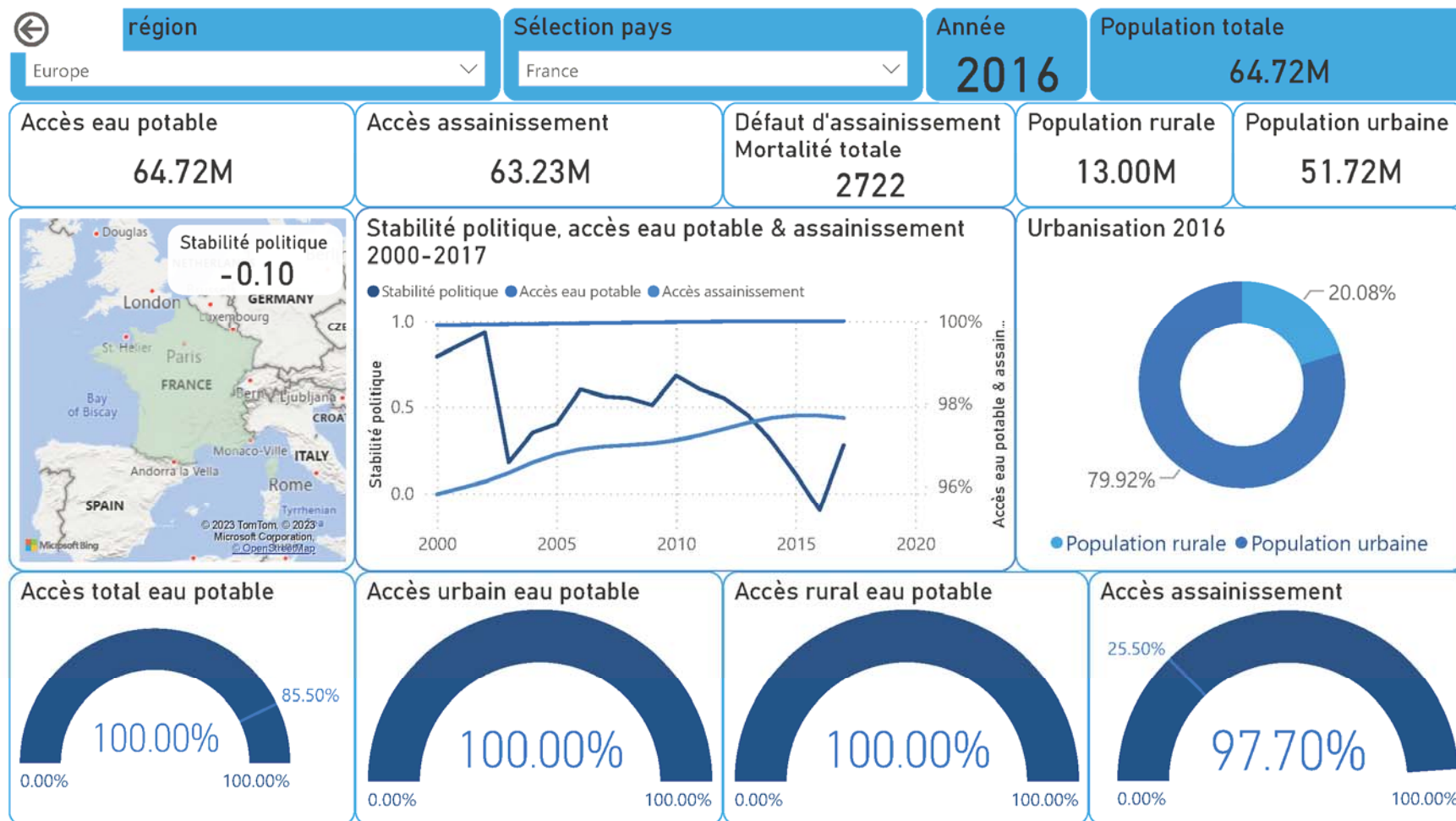
## 3 – Analyse & interprétation des résultats

### 3.2 – Vue régionale



### 3 – Analyse & interprétation des résultats

#### 3.3 – Vue nationale



# Conclusion & prochaines étapes

- ❖ Choix des pays à finaliser en fonction du retour du donneur d'ordres :
  - Stabilité politique
  - Urbanisation
  - Potentiel de création ou développement de services d'accès & traitement de l'eau
- ❖ Mise à jour automatique de la base à mettre en place :
  - Exhaustivité des données à revoir
  - Enrichissement de la base avec des indicateurs de développement économique & financier