



## Description

This database contains the records of a water survey for the Integrated Project.

## Data model

The database **md\_water\_services** includes more than 60 000 unique records and 43 columns. The data model, including the column names, data types and feature descriptions, is included in *Table 1* ordered by table.

*Table 1: The data model for the md\_water\_services database.*

employee table		
Column name	Description	Data type
<b>assigned_employee_id</b>	Unique ID assigned to each employee.	INT
<b>employee_name</b>	Name of the employee.	VARCHAR(255)
<b>phone_number</b>	Contact number of the employee.	VARCHAR(15)
<b>email</b>	Email address of the employee.	VARCHAR(255)
<b>address</b>	Residential address of the employee.	VARCHAR(255)
<b>town_name</b>	Name of the town where the employee resides.	VARCHAR(255)
<b>province_name</b>	Name of the province where the employee resides.	VARCHAR(255)
<b>position</b>	Position or job title of the employee.	VARCHAR(255)

global_water_access table		
Column name	Description	Data type
<b>name</b>	The country or area name.	VARCHAR(255)
<b>region</b>	Geographical region.	VARCHAR(255)
<b>year</b>	Year of the data record.	INT
<b>pop_n</b>	The national population size estimate in thousands.	FLOAT
<b>pop_u</b>	The urban population share estimate in percentage points (%).	FLOAT

<b>wat_bas_n</b>	The estimated national share of people with at least basic service (%).	FLOAT
<b>wat_lim_n</b>	The estimated national share of people with limited service (%).	FLOAT
<b>wat_unimp_n</b>	The estimated national share of people with unimproved service (%).	FLOAT
<b>wat_sur_n</b>	The estimated national share of people with surface service (%).	FLOAT
<b>wat_bas_r</b>	The estimated rural share of people with at least basic service (%).	FLOAT
<b>wat_lim_r</b>	The estimated rural share of people with limited service (%).	FLOAT
<b>wat_unimp_r</b>	The estimated rural share of people with unimproved service (%).	FLOAT
<b>wat_sur_r</b>	The estimated rural share of people with surface service (%).	FLOAT
<b>wat_bas_u</b>	The estimated urban share of people with at least basic service (%).	FLOAT
<b>wat_lim_u</b>	The estimated urban share of people with limited service (%).	FLOAT
<b>wat_unimp_u</b>	The estimated urban share of people with unimproved service (%).	FLOAT
<b>wat_sur_u</b>	The estimated urban share of people with surface service (%).	FLOAT

location table		
Column name	Description	Data type
<b>location_id</b>	Unique ID assigned to each location.	VARCHAR(255)
<b>address</b>	Address of the location.	VARCHAR(255)
<b>province_name</b>	Name of the province where the location is situated.	VARCHAR(255)
<b>town_name</b>	Name of the town where the location is situated.	VARCHAR(255)

<b>location_type</b>	Type or category of the location.	VARCHAR(255)
----------------------	-----------------------------------	--------------

visits table		
Column name	Description	Data type
<b>record_id</b>	Unique ID assigned to each visit.	INT
<b>location_id</b>	ID of the location visited.	VARCHAR(255)
<b>source_id</b>	ID of the water source visited.	VARCHAR(510)
<b>time_of_record</b>	Date and time of the visit.	DATETIME
<b>visit_count</b>	Number of visits made to this location.	INT
<b>time_in_queue</b>	Time spent by people waiting for water in a queue at the location.	INT
<b>assigned_employee_id</b>	ID of the employee who visited the location.	INT

water_quality table		
Column name	Description	Data type
<b>record_id</b>	Unique ID assigned to each record.	INT
<b>subjective_quality_score</b>	Score representing the subjective quality of the water source.	INT
<b>visit_count</b>	Number of visits made for data collection.	INT

water_source table		
Column name	Description	Data type
<b>source_id</b>	Unique ID assigned to each water source.	INT
<b>type_of_water_source</b>	Type or category of the water source. Can be: tap_in_home, tap_in_home_broken, well, shared_tap, river.	INT
<b>Number_of_people_served</b>	Number of people served by this water source.	INT

well_pollution table		
Column name	Description	Data type
source_id	ID of the water source being tested for pollution.	VARCHAR(258)
date	Date of the pollution test.	DATETIME
description	Description of the pollution test.	VARCHAR(255)
pollutant_ppm	Result of the pollution test in parts per million.	FLOAT
biological	Biological contamination level.	FLOAT
results	Result of the pollution test.	VARCHAR(255)