

1. Завантажте дані:

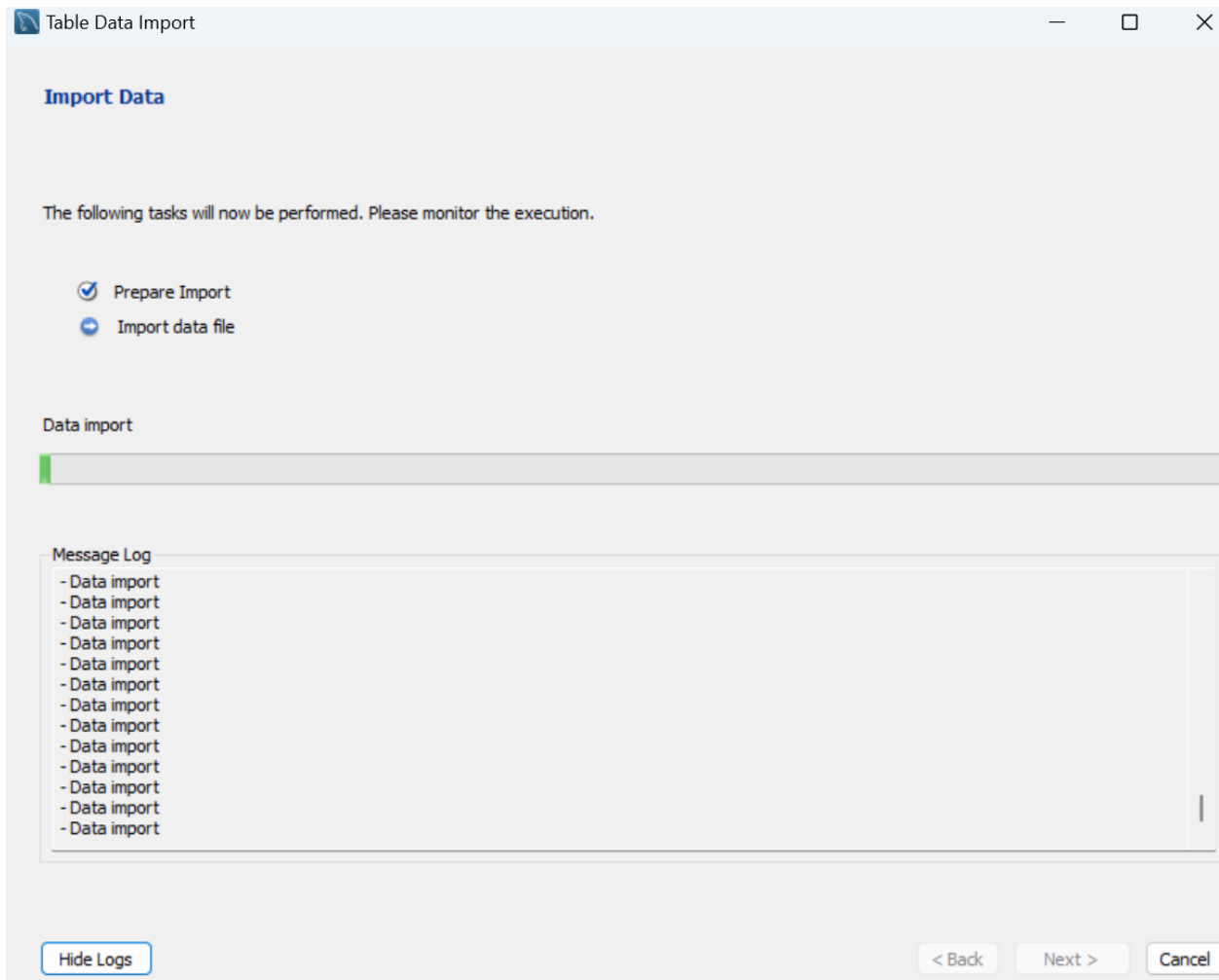
- Створіть схему pandemic у базі даних за допомогою SQL-команди.
- Оберіть її як схему за замовчуванням за допомогою SQL-команди.
- Імпортуйте [дані](#) за допомогою Import wizard так.

Як бачите, атрибути Entity та Code постійно повторюються. Позбудьтеся цього за допомогою нормалізації даних.

2. Нормалізуйте таблицю infectious_cases до 3ї нормальної форми. Збережіть у цій же схемі дві таблиці з нормалізованими даними.

Виконайте запит SELECT COUNT(*) FROM infectious_cases , щоб ментор міг зрозуміти, скільки записів ви завантажили у базу даних із файла.

```
17 • CREATE schema pandemic;  
18 • use pandemic;
```



Import Data

The following tasks will now be performed. Please monitor the execution.

- ✓ Prepare Import
- ✓ Import data file

Finished performing tasks. Click [Next >] to continue.

Message Log

- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Data import
- Import finished
Import data file finished
Finished performing tasks.

Hide Logs

< Back

Next >

Cancel

Таблиця після завантаження даних через wizard

22 • `select count(1) from infectious_cases;`
23

Result Grid | Filter Rows: | Export:

	count(1)
▶	10521

21 • `select * from infectious_cases`

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Entity	Code	Year	Number_yaws	polio_cases	cases_guinea_worm	Number_rabies	Number_malaria	Number_hiv	Number_tuberculosis	Number_smallpox	Number_cholera
Afghanistan	AFG	1987	NULL	4396	0	NULL	NULL	NULL	NULL	NULL	NULL
Afghanistan	AFG	1988	NULL	2149	0	NULL	NULL	NULL	NULL	NULL	NULL
Afghanistan	AFG	1989	NULL	385	0	NULL	NULL	NULL	NULL	NULL	NULL
Afghanistan	AFG	1990	NULL	336	0	2.145772	1219845.8	87.665596	16303.112	NULL	NULL
Afghanistan	AFG	1991	NULL	14	0	2.3209748	1132781.5	103.77359	17582.254	NULL	NULL
Afghanistan	AFG	1995	NULL	0	0	3.126361	725758.3	143.54527	22163.428	NULL	19903
Afghanistan	AFG	1997	NULL	21	0	3.446122	605853.06	160.59525	23140.863	NULL	4170
...

infectious_cases 4 x

Нормалізовані таблиці

- countries

Filter objects

hw3db

librarymanagement

mydb

pandemic

Tables

countries

Columns

id

country

country_code

Indexes

Foreign Keys

Triggers

infectious_cases

Views

Administration Schemas

Information

Table: countries

Columns:

id

country

country_code

int AI PK

varchar(255)

varchar(10)

26

•

CREATE TABLE IF NOT EXISTS countries (

27

id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,

28

country VARCHAR(255),

29

country_code VARCHAR(10)

30

);

31

32

•

INSERT INTO countries (country, country_code)

33

SELECT DISTINCT Entity, Code FROM infectious_cases;

34

35

•

select * from countries

Result Grid

Filter Rows:

Edit

Export

	id	country	country_code
▶	1	Afghanistan	AFG
	2	Africa	NULL
	3	African Region (WHO)	NULL
	4	Albania	ALB
	5	Algeria	DZA
	6	American Samoa	ASM
	7	Americas	NULL
	8	Andorra	AND

- записи про захворювання (infectious_cases_new)

1

•

CREATE TABLE IF NOT EXISTS infectious_cases_new AS

2

SELECT (SELECT c.id FROM countries c WHERE c.country = ic.Entity) AS country_id,

3

ic.Year,

4

ic.Number_yaws,

5

ic.polio_cases,

6

ic.cases_guinea_worm,

7

ic.Number_rabies,

8

ic.Number_malaria,

9

ic.Number_hiv,

10

ic.Number_tuberculosis,

11

ic.Number_smallpox,

12

ic.Number_cholera_cases

13

FROM infectious_cases ic;

Result Grid

Filter Rows:

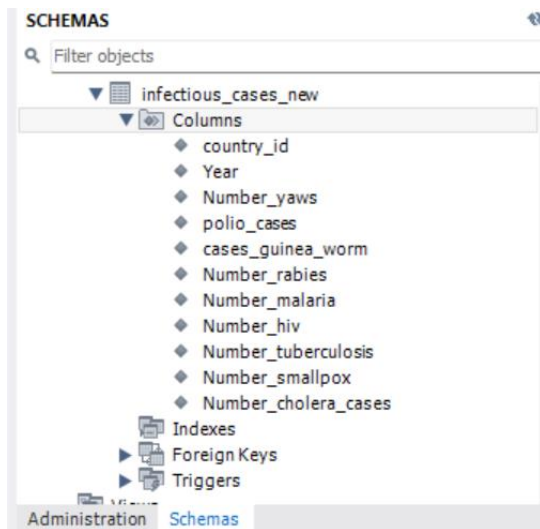
Export

Wrap Cell Content:

Fetch rows:

	country_id	Year	Number_yaws	polio_cases	cases_guinea_worm	Number_rabies	Number_malaria	Number_hiv	Number_tuberculosis	Number_smallpox
▶	1	1980	NULL	6160	0	NULL	NULL	NULL	NULL	NULL
	1	1981	NULL	5859	0	NULL	NULL	NULL	NULL	NULL
	1	1982	NULL	9730	0	NULL	NULL	NULL	NULL	NULL
	1	1983	NULL	13937	0	NULL	NULL	NULL	NULL	NULL
	1	1984	NULL	3864	0	NULL	NULL	NULL	NULL	NULL
	1	1985	NULL	13867	0	NULL	NULL	NULL	NULL	NULL

infectious_cases_new 1 x



3. Проаналізуйте дані:

- Для кожної унікальної комбінації Entity та Code або їх id порахуйте середнє, мінімальне, максимальне значення та суму для атрибута Number_rabies. Врахуйте, що атрибут Number_rabies може містити порожні значення — вам попередньо необхідно їх відфільтрувати.
- Результат відсортуйте за порахованим середнім значенням у порядку спадання.
- Оберіть тільки 10 рядків для виведення на екран.

```

8 • SELECT c.country,
9         ROUND(AVG(inf.Number_rabies), 4) AS average,
10        MIN(inf.Number_rabies) AS min_number,
11        MAX(inf.Number_rabies) AS max_number,
12        ROUND(SUM(inf.Number_rabies), 4) AS total_cases_sum
13 FROM    infectious_cases_new inf
14 JOIN    countries c ON inf.country_id = c.id
15 WHERE   inf.Number_rabies IS NOT NULL
16 GROUP BY 1
17 ORDER BY average DESC
18 LIMIT 10;

```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Fetch rows:

	country	average	min_number	max_number	total_cases_sum
▶	World	20192.3704	14075.508	24744.658	605771.111
	Lower Middle Income (WB)	15193.9594	10202.528	19182.795	455818.7815
	South Asia (WB)	11729.8888	7271.2754	15361.878	351896.6626
	South-East Asia Region (WHO)	11424.3269	6806.0127	15641.958	342729.8068
	G20	10189.046	6339.0796	13164.881	305671.3801
	India	8599.1733	5425.8726	11121.142	257975.1984
	Sub-Saharan Africa (WB)	4748.2681	4356.8184	5017.977	142448.0427
	African Region (WHO)	4649.7024	4283.073	4917.784	139491.0713
	Low Income (WB)	3568.1054	3014.617	3873.533	107043.1635
	East Asia & Pacific (WB)	3427.4611	2888.8305	4581.8107	102822.824

Result 2

Result 2 x

4. Побудуйте колонку різниці в роках.

Для оригінальної або нормованої таблиці для колонки Year побудуйте з використанням вбудованих SQL-функцій:

- атрибут, що створює дату першого січня відповідного року (наприклад, якщо атрибут містить значення '1996', то значення нового атрибута має бути '1996-01-01').
- атрибут, що дорівнює поточній даті,
- атрибут, що дорівнює різниці в роках двох вищезгаданих колонок.

Перераховувати всі інші атрибути, такі як Number_malaria, не потрібно.

```
43 • SELECT
44     Year AS original_year,
45     MAKEDATE(year, 1) AS original_year_start,
46     CURRENT_DATE AS today,
47     TIMESTAMPDIFF(YEAR, MAKEDATE(year, 1), CURRENT_DATE) AS years_since_record_date
48 FROM infectious_cases_new
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

	original_year	original_year_start	today	years_since_record_date
▶	1980	1980-01-01	2025-07-16	45
	1981	1981-01-01	2025-07-16	44
	1982	1982-01-01	2025-07-16	43
	1983	1983-01-01	2025-07-16	42
	1984	1984-01-01	2025-07-16	41
	1985	1985-01-01	2025-07-16	40
	1986	1986-01-01	2025-07-16	39

Result 3 x

5. Побудуйте власну функцію.

- Створіть і використайте функцію, що будує такий же атрибут, як і в попередньому завданні: функція має приймати на вхід значення року, а повертати різницю в роках між поточною датою та датою, створеною з атрибута року (1996 рік → '1996-01-01').

```
55 • DROP FUNCTION IF EXISTS GetYearsDifference;
56 DELIMITER //
57
58 • CREATE FUNCTION GetYearsDifference(start_year INT) -- YEAR datatype ranges between 1901 and 2155 so better to have INT to avoid errors
59 RETURNS INT
60 DETERMINISTIC
61 BEGIN
62     DECLARE start_date DATE;
63     SET start_date = MAKEDATE(start_year, 1);
64     RETURN TIMESTAMPDIFF(YEAR, start_date, CURRENT_DATE);
65 END //
66
67 DELIMITER ;
68
69 • SELECT GetYearsDifference(2020); -- function test
```

Output

Action Output

#	Time	Action	Message	Duration /
✓ 1	20:13:43	CREATE FUNCTION GetYearsDifference(start_year INT) -- YEAR datatype range...	0 row(s) affected	0.016 sec

Перевіряємо виконання функції на таблиці `infectious_cases_new` – заміняємо відповідну частину запиту з попереднього завдання на функцію:

```
71 • SELECT
72     Year AS original_year,
73     CURRENT_DATE AS today,
74     GetYearsDifference(Year) AS years_since_record_date
75 FROM infectious_cases_new;
```

Result Grid

	original_year	today	years_since_record_date
▶	1980	2025-07-16	45
	1981	2025-07-16	44
	1982	2025-07-16	43
	1983	2025-07-16	42
	1984	2025-07-16	41
	1985	2025-07-16	40
	1986	2025-07-16	39
	1987	2025-07-16	38
	1988	2025-07-16	37
	1989	2025-07-16	36
	1990	2025-07-16	35
	1991	2025-07-16	34

Result 7

Output

Action Output

#	Time	Action	Message	Duration / F
✓ 1	20:10:48	SELECT Year AS original_year, CURRENT_DATE AS today, GetYears...	1000 row(s) returned	0.015 sec /