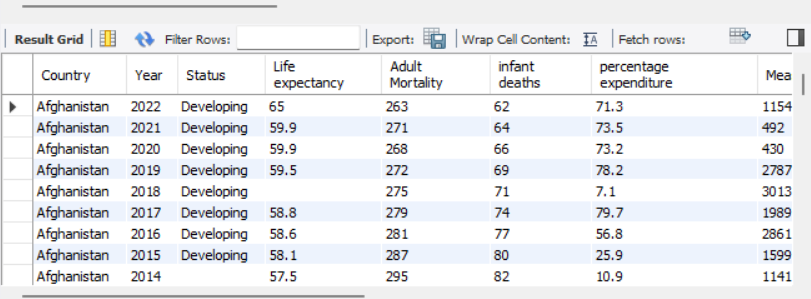
**WORLD LIFE EXPECTANCY**

**Data Cleaning**

**Initial Data Exploration**

SELECT \*

FROM worldlifeexpectancy;



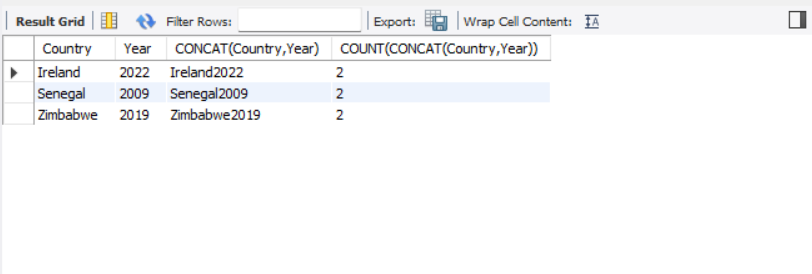
**Identifying Duplicates**

SELECT Country, Year , CONCAT(Country,Year) , COUNT(CONCAT(Country,Year))

FROM worldlifeexpectancy

GROUP BY Country, Year , CONCAT(Country,Year)

HAVING COUNT(CONCAT(Country,Year)) > 1;



SELECT \*

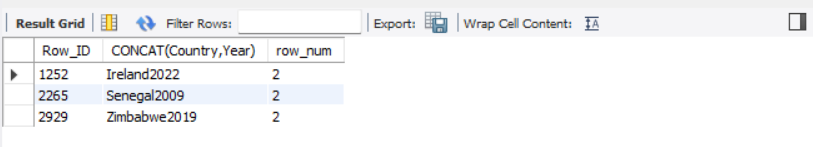
FROM

(SELECT Row\_ID , CONCAT(Country,Year) ,

ROW\_NUMBER() OVER(PARTITION BY CONCAT(Country,Year) ORDER BY CONCAT(Country,Year)) AS row\_num

FROM worldlifeexpectancy)

WHERE row\_num >1;



**Removing Duplicates**

DELETE FROM worldlifeexpectancy

WHERE Row\_ID IN (

SELECT Row\_ID

FROM

(SELECT Row\_ID , CONCAT(Country,Year) ,

ROW\_NUMBER() OVER(PARTITION BY CONCAT(Country,Year) ORDER BY CONCAT(Country,Year)) AS row\_num

FROM worldlifeexpectancy) AS rowww

WHERE row\_num >1

);

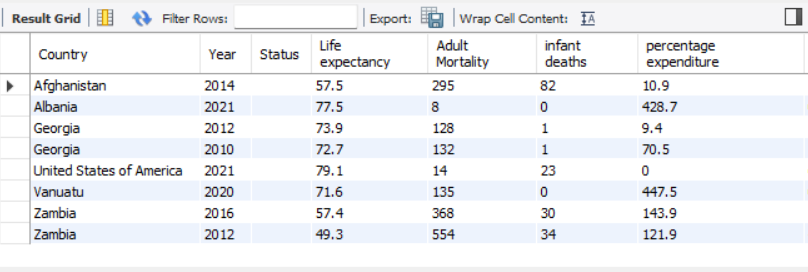


**Checking for Missing Status Values**

SELECT \*

FROM worldlifeexpectancy

WHERE status = '';



**Filling Missing Status Values**

UPDATE worldlifeexpectancy t1

JOIN worldlifeexpectancy t2

ON t1.Country = t2.Country

SET t1.Status = 'Developing'

WHERE t1.Status = ''

AND t2.Status <> ''

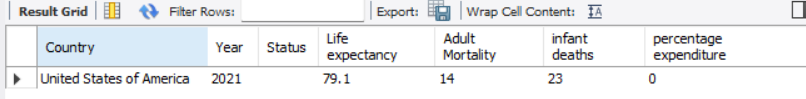
AND t2.Status = 'Developing';



SELECT \*

FROM worldlifeexpectancy

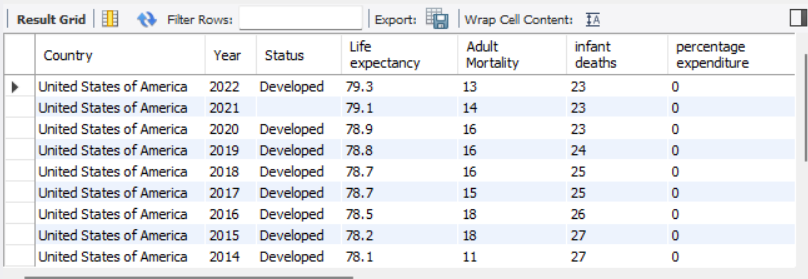
WHERE Status = '';



SELECT \*

FROM worldlifeexpectancy

WHERE Country = 'United States of America' ;



UPDATE worldlifeexpectancy t1

JOIN worldlifeexpectancy t2

ON t1.Country = t2.Country

SET t1.Status = 'Developed'

WHERE t1.Status = ''

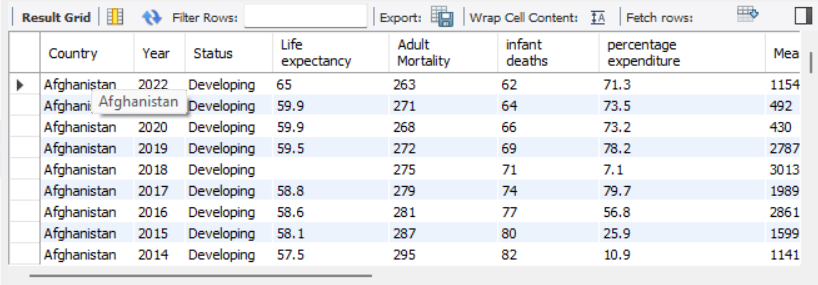
AND t2.Status <> ''

AND t2.Status = 'Developed';



SELECT \*

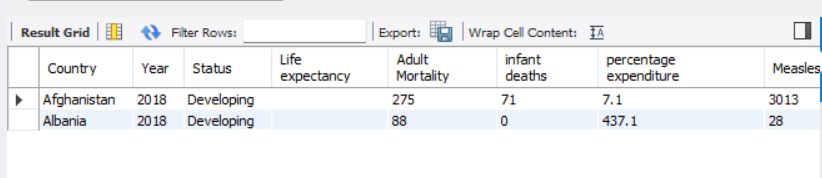
FROM worldlifeexpectancy;



SELECT \*

FROM worldlifeexpectancy

WHERE `Life expectancy` = '';



SELECT t1.Country, t1.Year, t1.`Life expectancy`, t2.Country, t2.Year, t2.`Life expectancy`, t3.Country, t3.Year, t3.`Life expectancy`,

ROUND((t2.`Life expectancy` + t3.`Life expectancy`)/2,1)

FROM worldlifeexpectancy t1

JOIN worldlifeexpectancy t2

ON t1.Country = t2.Country

AND t1.Year = t2.Year - 1

JOIN worldlifeexpectancy t3

ON t1.Country = t3.Country

AND t1.Year = t3.Year + 1

WHERE t1. `Life expectancy` = '';

