



Module 4: Project 1

By- Rohit Raman,

Course - ITC 6000,

Professor- Anis Sharafoddini,

Date: 05/12/2022

Project Title:

HIV Analysis.

Introduction of HIV analysis:

The propose analysis is done for the region Latin America and Caribbean.

Starting from the dataset, we have downloaded it from the UNICEF official website. There are seven dataset that we downloaded for the analysis. Following to this we will explain the name of datasets with its descriptions:



1. **HIV_Adolescent_ART_Coverage_2022:** It gives the information about the percentage of adolescent living with HIV and receiving **Antiretrovial treatment (ART)**
2. **HIV_ANC_Testing_2022:** Conveys the information about the **number of pregnant women** diagnosed with **HIV** for the year 2015-2021.
3. **HIV_Discrimination_2022:** This dataset includes data on discrimination against people who is living with HIV.
4. **HIV_Early_Infant_Diagnosis_2022:** It has a report about the **infants suffering** from **HIV** who got infection from mother during pregnancy.

5. **HIV_Orphans_2022:** This dataset has information about the **infants** who lost their parents due to **HIV**.
6. **HIV_Paediatric_ART_Coverage_2022:** This data has information about the coverage of **antiretroviral treatment (ART)** among **children** aged 0-14 living with **HIV**.
7. **HIV_PMTCT_Coverage_2022:** It has an information about **the coverage of prevention of mother to child transmission (PMTCT)**.
8. **HIV_Prevention_Adolescents_2022:** This dataset conveys the information about the **HIV prevention efforts among adolescents with comprehensive knowledge of HIV precautions measures**.

SQL Analysis:

In this section we will explain and demonstrate the analysis which we have done using SQL.

1. **Total Records For pregnant woman in HIV_ANC_Testing dataset.**

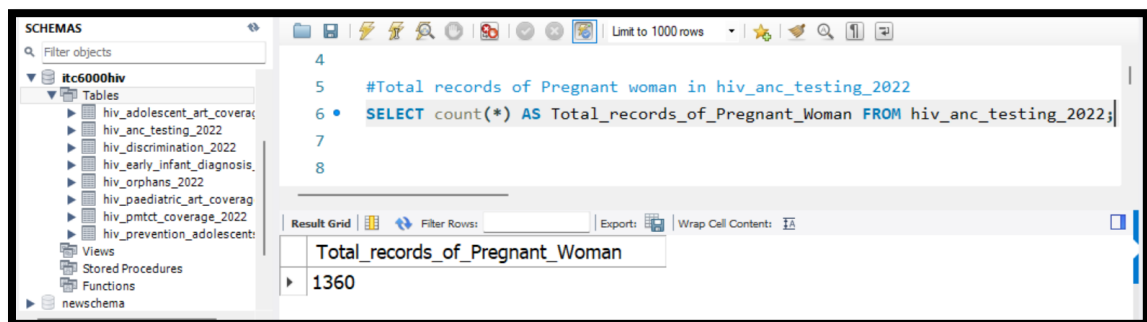
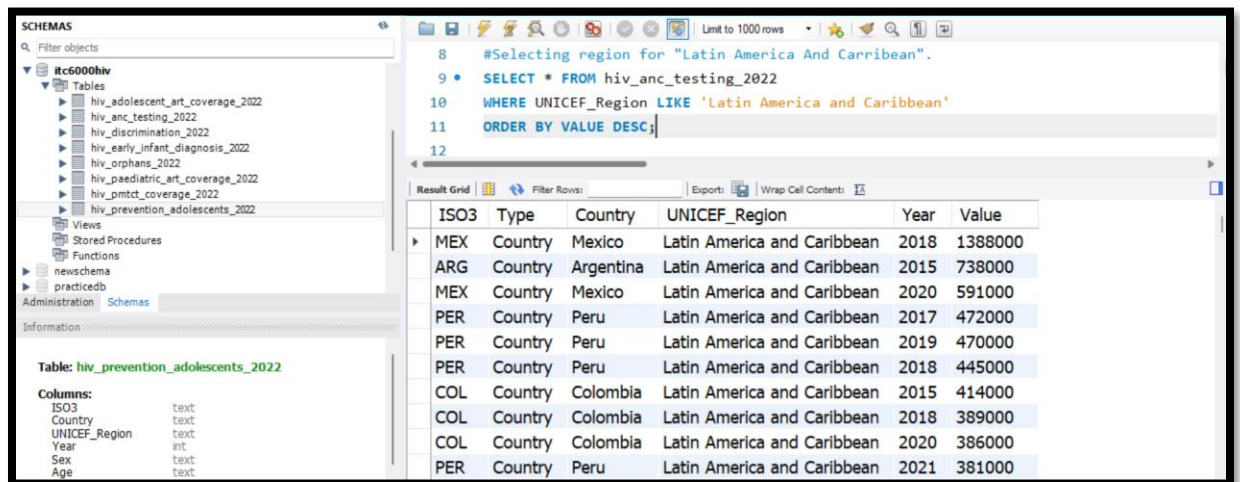


Figure1: Total records of pregnant woman in the dataset.

Observation:

There are around 1360 records present of pregnant woman, it includes every region

2. Selecting the assigned region for the Analysis:



The screenshot shows a database query interface. On the left, a tree view displays the database schema, including tables like `hiv_adolescent_art_coverage_2022`, `hiv_anc_testing_2022`, `hiv_discrimination_2022`, `hiv_early_infant_diagnosis_2022`, `hiv_orphans_2022`, `hiv_paediatric_art_coverage_2022`, `hiv_pmtct_coverage_2022`, and `hiv_prevention_adolescents_2022`. The main query editor shows the following SQL query:

```

8 #Selecting region for "Latin America And Caribbean".
9 * SELECT * FROM hiv_anc_testing_2022
10 WHERE UNICEF_Region LIKE 'Latin America and Caribbean'
11 ORDER BY VALUE DESC;
12

```

The results are displayed in a table with the following columns: ISO3, Type, Country, UNICEF_Region, Year, and Value. The data is sorted by Value in descending order.

| ISO3 | Type | Country | UNICEF_Region | Year | Value |
|------|---------|-----------|-----------------------------|------|---------|
| MEX | Country | Mexico | Latin America and Caribbean | 2018 | 1388000 |
| ARG | Country | Argentina | Latin America and Caribbean | 2015 | 738000 |
| MEX | Country | Mexico | Latin America and Caribbean | 2020 | 591000 |
| PER | Country | Peru | Latin America and Caribbean | 2017 | 472000 |
| PER | Country | Peru | Latin America and Caribbean | 2019 | 470000 |
| PER | Country | Peru | Latin America and Caribbean | 2018 | 445000 |
| COL | Country | Colombia | Latin America and Caribbean | 2015 | 414000 |
| COL | Country | Colombia | Latin America and Caribbean | 2018 | 389000 |
| COL | Country | Colombia | Latin America and Caribbean | 2020 | 386000 |
| PER | Country | Peru | Latin America and Caribbean | 2021 | 381000 |

Figure 2: Query for selecting the region for Latin America and Caribbean.

Observation:

Since our analysis is projected on the Latin America and Caribbean, so we have selected the above-mentioned region.

3. Total number of women suffered from HIV in Latin America and Caribbean till date.

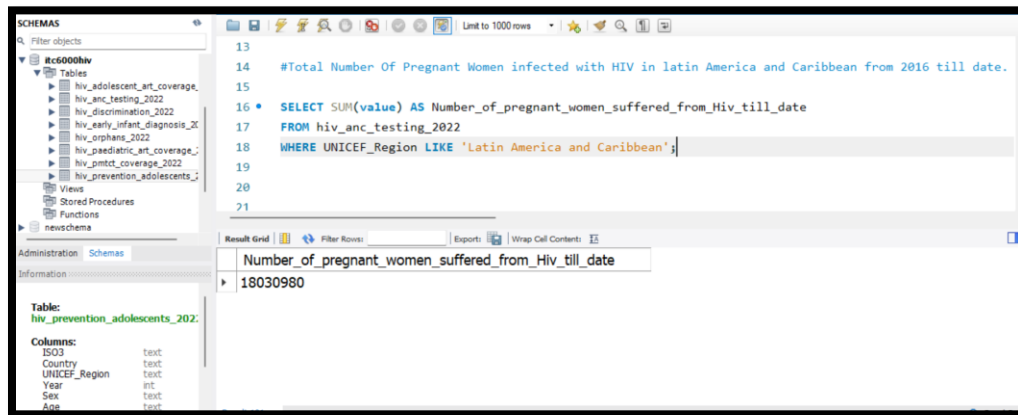


Figure 3: Total number pregnant women suffered from HIV in Latin America and Caribbean.

Observation:

Around **1,80,30,980** pregnant women have been contaminated with HIV till date in Latin America and Caribbean.

4. Pregnant women suffering from HIV in the year 2015 for country Argentina

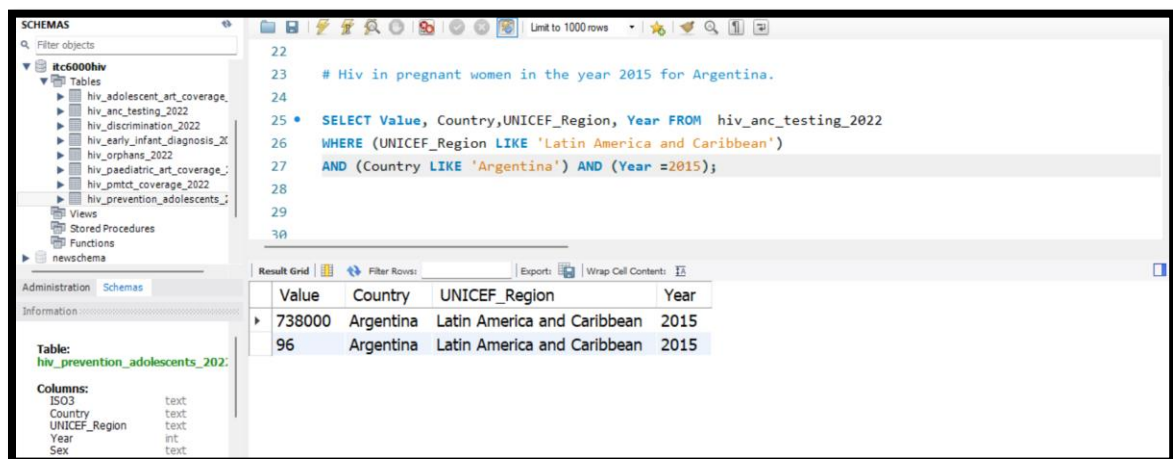


Figure 4: Pregnant women suffering from HIV in Argentina.

Observation:

Figure 4 conveys the information on **women** suffering from HIV in the year 2015 in **Argentina**.

5. Total number of women who got HIV in Argentina for the year 2015.

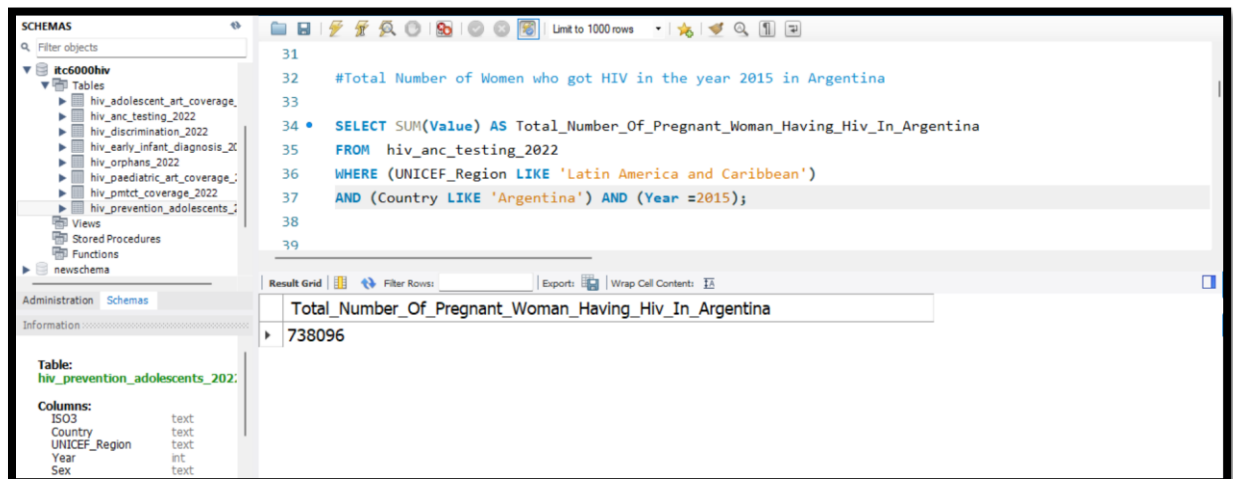


Figure 5: Total Number of Women who got HIV in the Argentina for the year 2015.

Observation.

There are **738096** pregnant women who got **HIV** in **Argentina** in the year **2016**.

6. Country that witnessed maximum number of HIV case pregnant women in one year for Latin America and Caribbean region.

47 #Country in Latin America and caribbean that witnessed maximum HIV case in Pregnant women in one year.

48

49 * SELECT * FROM hiv_anc_testing_2022

50 WHERE UNICEF_Region LIKE 'latin America and Caribbean'

51 ORDER BY Value DESC;

| ISO3 | Type | Country | UNICEF_Region | Year | Value |
|------|---------|-----------|-----------------------------|------|---------|
| MEX | Country | Mexico | Latin America and Caribbean | 2018 | 1388000 |
| ARG | Country | Argentina | Latin America and Caribbean | 2015 | 738000 |
| MEX | Country | Mexico | Latin America and Caribbean | 2020 | 591000 |
| PER | Country | Peru | Latin America and Caribbean | 2017 | 472000 |
| PER | Country | Peru | Latin America and Caribbean | 2019 | 470000 |
| PER | Country | Peru | Latin America and Caribbean | 2018 | 445000 |
| COL | Country | Colombia | Latin America and Caribbean | 2015 | 414000 |
| COL | Country | Colombia | Latin America and Caribbean | 2018 | 389000 |
| COL | Country | Colombia | Latin America and Caribbean | 2020 | 386000 |
| PER | Country | Peru | Latin America and Caribbean | 2021 | 381000 |

Figure 6: Country with maximum number of HIV case in pregnant women in one year.

Observation:

Mexico contributed the greatest number of **HIV** cases among **pregnant woman** in one year that is for the **year 2018** with a value more than **13.5 lakhs**.

NESTED QUERY DEMONSTRATION.

7. Mexican pregnant women contribution in HIV for the year 2018 in terms of percentage.

62 * SELECT ROUND(SUM(value)/(SELECT SUM(value) FROM hiv_anc_testing_2022

63 WHERE UNICEF_Region LIKE 'latin America and Caribbean' AND year =2018) *100,2)

64 AS Percentage_of_HIV_contamination_contributed_by_Mexican_women_in_2018

65 FROM hiv_anc_testing_2022

66 WHERE (UNICEF_Region LIKE 'latin America and Caribbean') AND (year =2018) AND (country like 'Mexico');

67

68

| Percentage_of_HIV_contamination_contributed_by_Mexican_women_in_2018 | |
|--|--|
| 32.98 | |

Result 161 x

Output

| # | Time | Action | Message | Duration / Fetch |
|-----|----------|--|-------------------|-----------------------|
| 130 | 19:32:01 | SELECT round(sum(value)/(select sum(value) from hiv_anc_testing_2022 WH... | 1 row(s) returned | 0.000 sec / 0.000 sec |
| 131 | 19:33:10 | SELECT round(sum(value)/(SELECT sum(value) FROM hiv_anc_testing_2022 ... | 1 row(s) returned | 0.000 sec / 0.000 sec |
| 132 | 19:34:20 | SELECT ROUND(SUM(value)/(SELECT SUM(value) FROM hiv_anc_testing_2... | 1 row(s) returned | 0.000 sec / 0.000 sec |

Figure 7: Mexican pregnant women contribution in HIV for the year 2018.

Observation:

Form the figure 7 it can be said **maximum** number of women contributed **the highest number of HIV** cases in the year 2018 are form **Mexico**. In term of **percentage**, it has value of around **32.98** which means **32.98 percent** of **Mexican pregnant women** contributed in **HIV** in the Latin America and Caribbean region for the year 2018.

8. Percentage of Infants who got HIV from mother in Latin America and Caribbean Region.

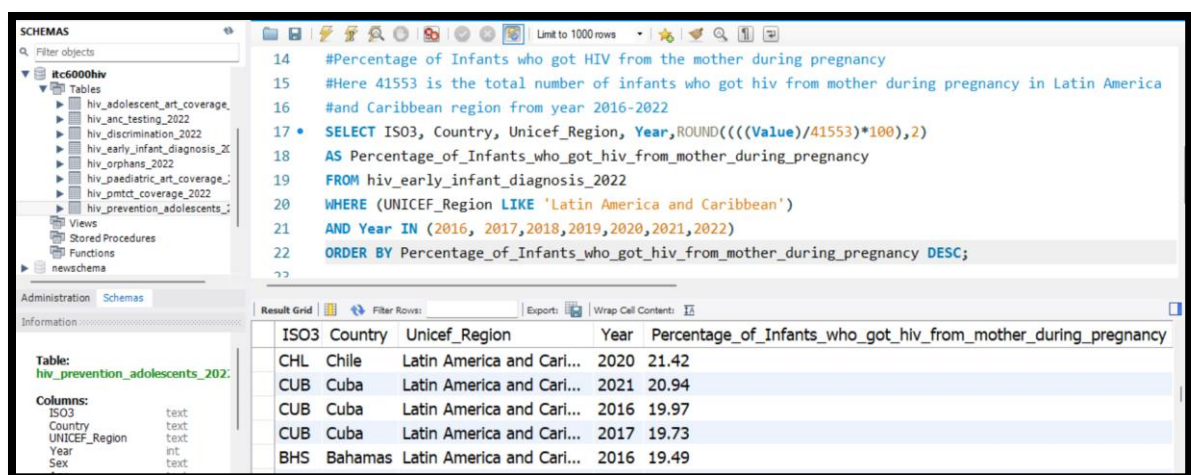


Figure 8: Percentage of Infants who got HIV from the mother during pregnancy.

Observation:

Chile, Cuba and Bahamas have the **highest number of transmissions** of HIV from mother to children. We can see for the year **2020** “**Chile**” has the **highest number of transmissions of HIV** from mother to infants with value of **21.42**, **Cuba** stood **1st** for the years **2021, 2016 and 2017** with value of almost **20 percent**. For **2016** it was **Bahamas** with maximum number of transmissions having values of **19.49**.

9. Total HIV case in infants in Cuba from 2016 till 2022.

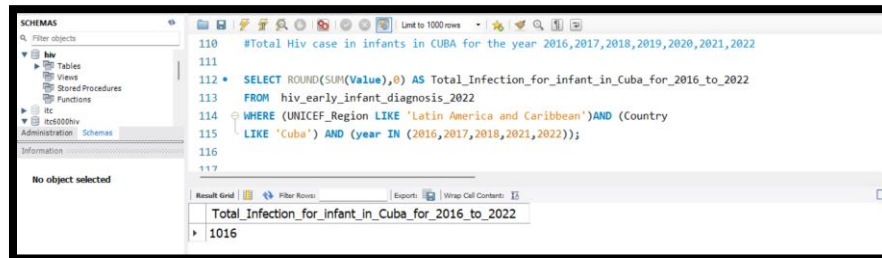


Figure 9: Total HIV case in infants in Cuba from 2016 till 2022.

Observation:

The Total HIV case in **infants** from the year **2016 till 2022** is **1016**.

10. Percentage of infants with gender female getting education on HIV.

```

1 #Percentage of Infants age from 15-24 who is receiving knowledge on HIV transmission
2 * SELECT ISO3, Country, UNICEF_Region,year, Sex, Age, DISAGG_CATEGORY,
3 value AS percentage_of_female_having_knowledge_HIV
4 FROM hiv_prevention_adolescents_2022
5 WHERE (UNICEF_Region LIKE 'Latin America and Caribbean') AND (age like '15-24')
6 AND (sex LIKE 'female' AND DISAGG_CATEGORY LIKE 'Education')
7 GROUP BY Country
8 ORDER BY percentage_of_female_having_knowledge_HIV DESC;

```

| ISO3 | Country | UNICEF_Region | ye | Sex | Age | DISAGG_C | percentage_of_female_having_knowledge_HIV |
|------|-------------|---------------------|-----|--------|--------|-----------|---|
| GUY | Guyana | Latin America an... | 2.. | Female | 15-... | Education | 77 |
| CUB | Cuba | Latin America an... | 2.. | Female | 15-... | Education | 60 |
| ARG | Argentina | Latin America an... | 2.. | Female | 15-... | Education | 57 |
| URY | Uruguay | Latin America an... | 2.. | Female | 15-... | Education | 54 |
| BRB | Barbados | Latin America an... | 2.. | Female | 15-... | Education | 54 |
| DOM | Dominica... | Latin America an... | 2.. | Female | 15-... | Education | 52 |

Figure 10. Percentage of female infants getting knowledge on HIV.

Observation:

Female Infants from **Guyana** stood at the table top when it comes to getting knowledge on **HIV** and constitute around **77 percent** in the table. Second come from **Cuba** with a value of **60%** and **Argentina** stood at third with value of **57 percent**.

Note: Here the value is already given in percentage, so no need to do any mathematical manipulation on the variable.

11. Percentage Female adolescent receiving ART (Antiretroviral Treatment) Coverage.

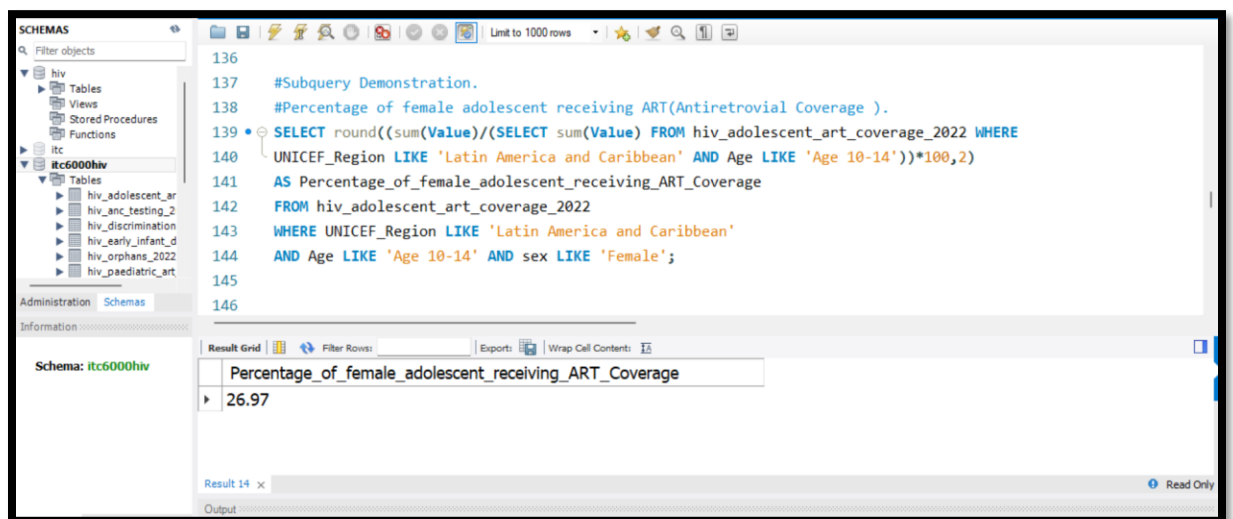


Figure 11: Percentage of adolescent female receiving ART coverage.

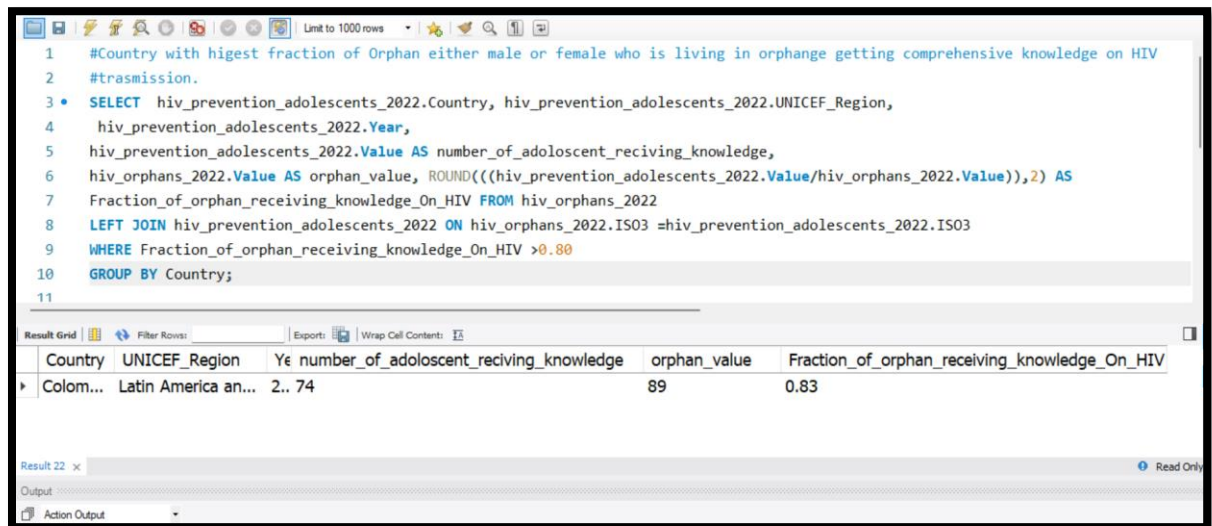
Observation:

Around **26.97 percent** of female from **Latin America and Caribbean** region receiving Art coverage.

JOINT DEMONSTRATION.

LEFT JOIN

12. Country with the Highest fraction of orphan who is living in orphan and getting comprehensive knowledge on HIV.



The screenshot shows a SQL query editor with a query that finds the country with the highest fraction of orphans receiving comprehensive HIV knowledge. The query uses a LEFT JOIN between the 'hiv_prevention_adolescents_2022' and 'hiv_orphans_2022' tables. The results table shows that Colombia, in the Latin America and Caribbean region, has the highest fraction of orphans receiving knowledge, with a value of 0.83.

```

1  #Country with highest fraction of Orphan either male or female who is living in orphanage getting comprehensive knowledge on HIV
2  #transmission.
3  • SELECT hiv_prevention_adolescents_2022.Country, hiv_prevention_adolescents_2022.UNICEF_Region,
4     hiv_prevention_adolescents_2022.Year,
5     hiv_prevention_adolescents_2022.Value AS number_of_adolescent_reciving_knowledge,
6     hiv_orphans_2022.Value AS orphan_value, ROUND(((hiv_prevention_adolescents_2022.Value/hiv_orphans_2022.Value)),2) AS
7     Fraction_of_orphan_receiving_knowledge_On_HIV FROM hiv_orphans_2022
8  LEFT JOIN hiv_prevention_adolescents_2022 ON hiv_orphans_2022.IS03 =hiv_prevention_adolescents_2022.IS03
9  WHERE Fraction_of_orphan_receiving_knowledge_On_HIV >0.80
10 GROUP BY Country;
11

```

| Country | UNICEF_Region | Ye | number_of_adolescent_reciving_knowledge | orphan_value | Fraction_of_orphan_receiving_knowledge_On_HIV |
|----------|---------------------|-----|---|--------------|---|
| Colom... | Latin America an... | 2.. | 74 | 89 | 0.83 |

Figure 12: Fraction of orphan getting comprehensive knowledge on HIV

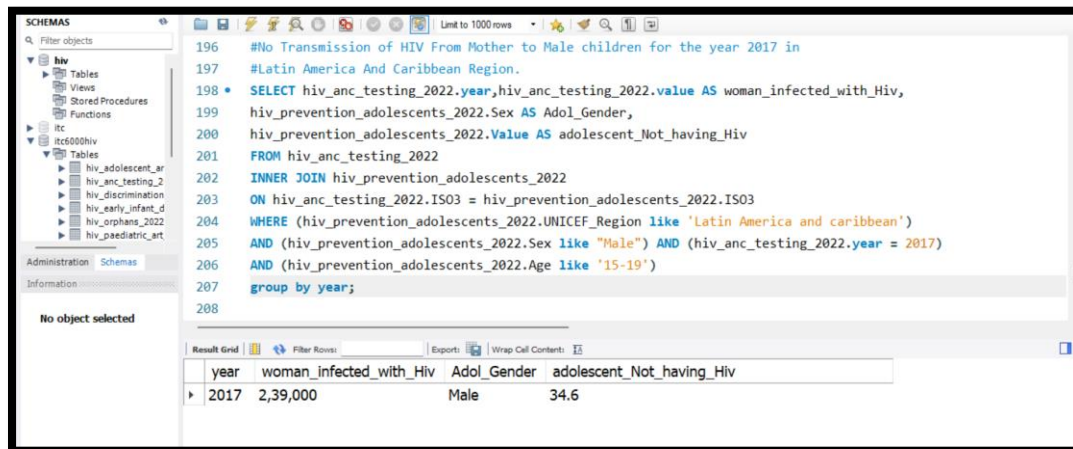
Transmission using Left Join.

Observation:

To get the highest fraction value of **orphan** receiving **comprehensive knowledge on HIV transmission** in Latin America and Caribbean region, we have used **left join** to join two table i.e., **Orphan table** and **HIV prevention adolescent** table. Here we got to know **Colombia** is the country with **high value in** fraction of orphan getting knowledge on HIV transmission. Around **83 percent** of **orphans** in **Colombia** who is staying in orphan getting education on HIV

INNER JOIN DEMONSTRATION.

13. Percentage of adolescent who got no transmission of HIV from mother.



```

196 #No Transmission of HIV From Mother to Male children for the year 2017 in
197 #Latin America And Caribbean Region.
198 * SELECT hiv_anc_testing_2022.year, hiv_anc_testing_2022.value AS woman_infected_with_Hiv,
199       hiv_prevention_adolescents_2022.Sex AS Adol_Gender,
200       hiv_prevention_adolescents_2022.Value AS adolescent_Not_having_Hiv
201 FROM hiv_anc_testing_2022
202 INNER JOIN hiv_prevention_adolescents_2022
203 ON hiv_anc_testing_2022.IS03 = hiv_prevention_adolescents_2022.IS03
204 WHERE (hiv_prevention_adolescents_2022.UNICEF_Region like 'Latin America and caribbean')
205 AND (hiv_prevention_adolescents_2022.Sex like 'Male') AND (hiv_anc_testing_2022.year = 2017)
206 AND (hiv_prevention_adolescents_2022.Age like '15-19')
207 group by year;
208

```

| year | woman_infected_with_Hiv | Adol_Gender | adolescent_Not_having_Hiv |
|------|-------------------------|-------------|---------------------------|
| 2017 | 2,39,000 | Male | 34.6 |

Figure 13: Percentage of adolescent who got no transmission of HIV from mother using Inner Join.

Observation:

Through **Inner join**, we have merged two table i.e.; **Hiv_Anc_Testing_2022** and **Hiv_prevention_adoscents_2022** and we got to know the **34.6 percent** of adolescent children staying with their mother did receive any infection from their mother.

However, in that year **2,39,000 women** suffering from **HIV**.

RIGHT JOIN DEMOSTRATION.

14. Percentage of women in Latin America suffering from HIV and percentage of it receiving coverage for the prevention of mother to child transmission in all the year.

220 #Percentage of women who is suffering from HIV and percentage of it receiving coverage
 221 #for the prevention of mother to child tranmission in all the year.
 222
 223 • SELECT hiv_anc_testing_2022.UNICEF_Region, hiv_anc_testing_2022.Year,
 224 hiv_anc_testing_2022.value AS Percentage_Of_Women_diagnosed_with_hiv,
 225 hiv_pmtct_coverage_2022.value AS Percentage_of_women_receiving_PMCTC
 226 FROM hiv_anc_testing_2022
 227 RIGHT JOIN hiv_pmtct_coverage_2022
 228 ON hiv_anc_testing_2022.IS03 =hiv_pmtct_coverage_2022.IS03
 229 WHERE hiv_anc_testing_2022.UNICEF_Region LIKE 'Latin America and Caribbean'
 230 GROUP BY year;
 231
 232

| UNICEF_Region | Year | Percentage_Of_Women_diagnosed_with_hiv | Percentage_of_women_receiving_PMCTC |
|-----------------------------|------|--|-------------------------------------|
| Latin America and Caribbean | 2017 | 96 | 45.2 |
| Latin America and Caribbean | 2015 | 89.3 | 45.2 |
| Latin America and Caribbean | 2019 | 1 | 79.1 |

Figure 14: Percent of woman suffering from HIV and percent of it receiving coverage using Right Join.

Observation:

From figure 5 we have demonstrated the **percentage of women** suffering from HIV and **percentage of it receiving coverage** of **prevention of mother to child transmission** for the year **2017,2015 and 2019**. We have used the **right join** for joining two tables.

INNER JOIN

15. Top 3 Country with highest percentage of discrimination in orphanage.

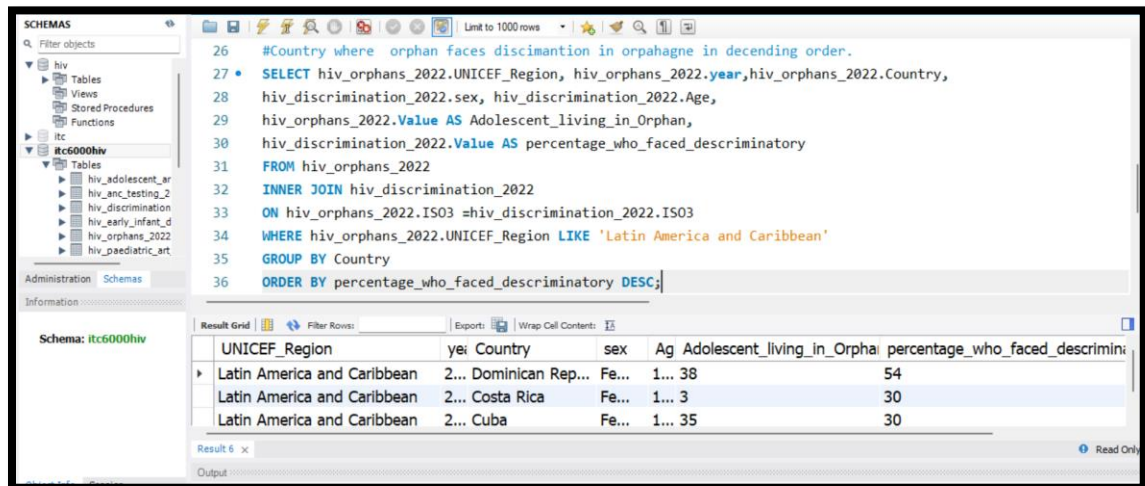


Figure 15: Discrimination percentage in orphanage.

Observation:

In the analysis of discrimination, we have analyzed that orphans who are staying in **Dominican Republic** face a **high rate** of discrimination. **In 2018** among **38 adolescent** who is staying in **orphanage**; **54 percentage** of them **faced discrimination**. Cost Rica stood at 2nd position with 30 percentage of discrimination in orphanage and same is the case with Cuba.

KPI:

1. HIV in early infants in children from 2010 to 2030(expected).

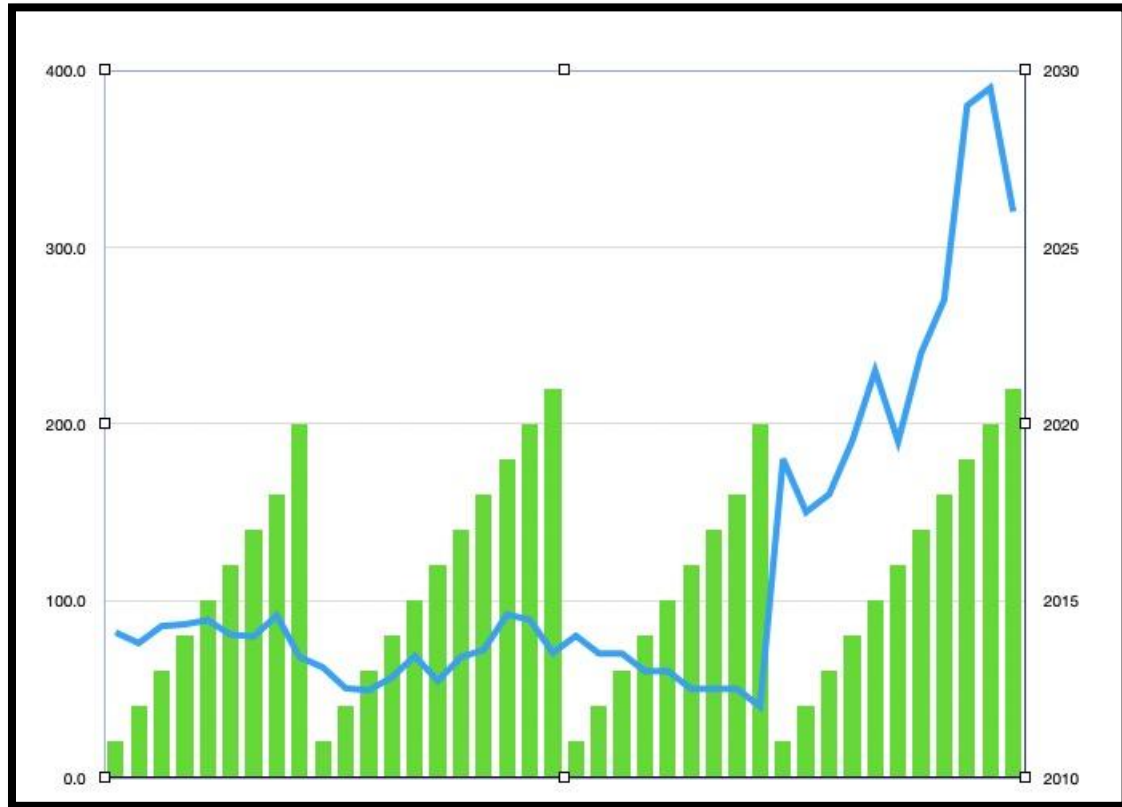


Figure 16: HIV transmission predication among Infants.

Observation:

From the KPI analysis, it can be said that the Transmission of HIV among infants are bound to the increase in the region of Latin America and Costa Rica Region.

2.Mother to child transmission in Mexico from 2007 to 2024.

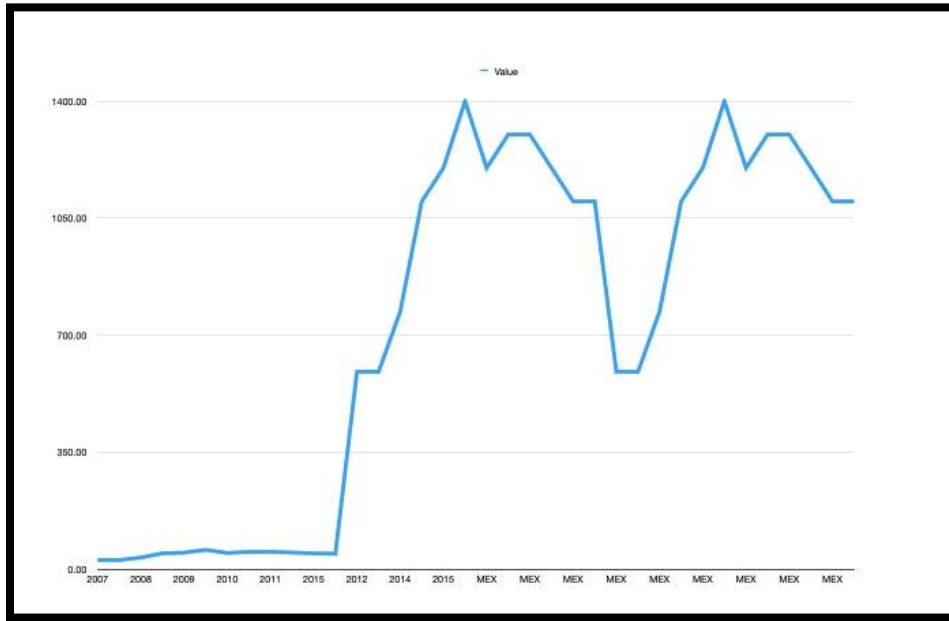


Figure 17: HIV transmission analysis from mother to child from 2007 to 2015.

Observation:

The transmission of HIV from mother to child in Mexico is in irregular pattern, Initially from 2005 till 2015 there was significant rise in the transmission, after which Mexico witnessed a decline in transmission in the year 2019, later after 2019, there was again increase in the cases of transmission from mother to child.

SUMMARY:

Through HIV analysis we got an opportunity to learn a lot on SQL and about the real condition of HIV patients. There are several analyses which we have done by using **basic SQL search query** like **total number of pregnant women suffered from HIV** in Latin America and Caribbean region till date. Furthermore, we did **query** on Country in **Latin**

America and Caribbean that **witnessed maximum number of HIV** case in **pregnant women**. We have used “**nested query**” to do analysis on **percent** of **women** who is **receiving Antiretroviral treatment**. Additionally, we have performed some of **Join statement** where we got exposure of **Table join**, we used **left join** to know the what **fraction** of **orphan** living in orphanage getting **comprehensive knowledge** on HIV using left join. We used **Inner join** to know what percentage of **children** got **no HIV transmission** from their **mother**. Lastly, we used **Right join** to know **what percent** of **women** suffering from HIV and what **percent** of it **receiving coverage for the treatment**.

REFERENCES:

<https://databank.worldbank.org/source/world-development-indicators>

<https://data.unicef.org/resources/dataset/hiv-aids-statistical-tables/>

<https://data.unicef.org/topic/hivaids/global-regional-trends/>

<https://data.unicef.org/topic/gender/gender-and-hiv-aids/>

<https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics>