

**Subject :** DASHBOARD AND DATA ANALYSIS (EXL1002 – 93951**)**

**Topic :** Malaria in Colombia 2015 - 2023

**Group Member Name :** Gressy Cruz De La Rosa, Fatima Kong De Puga, Juan Diego Wilches Vega, Ayotomiwa Salawu, Vikram Jeet Singh

**Date :** December 11, 2023

**EXECUTIVE SUMMARY**

This report analyzes malaria cases in Colombia - a country in South America - focusing on the years between 2015 and 2023. Colombia is one of the countries located in the South America Amazon, where most cases of Malaria on the continent have emanated (Recht et al., 2017). It has been reported that after a decline between 2005 and 2014, the cases of Malaria began to spike in 2015 following the conflict in Venezuela (a neighboring country), which has led most of its citizens to migrate to Colombia and other neighboring countries. This phenomenon impacts the number of cases of Malaria in Colombia, which rocketed after the significant migration began.

**DATA DECISIONS**

Following a thorough review of the available data, the team decided to determine the disease's age group distribution and classify the disease's cases by sex, ethnicity, and healthcare coverage. This will enable us to determine the age group most affected, the sex with the highest disease cases, and the ethnic group most affected by the disease to plan target-specific interventions and allocate resources efficiently.

**DATA ACQUISITION**

Data was obtained from the National Institute of Health of Colombia, which provided details about individuals who had suffered from the disease. Codes were uniquely assigned to each individual, and all identifiers were eliminated to keep individual identities anonymous. Variables collected in the data include the date health authorities were notified about the case, year of occurrence, age of the individual, ethnicity, migrant status, pregnancy status for females, and the department of occurrence (synonymous with province).

**DATA TIMING**

This research explored reported Malaria cases in Colombia between 2015 and 2023 by sex, ethnicity, age distribution, and health coverage. The data was thoroughly screened for duplicates, cleaned, and made analysis ready.

**DATA ANALYSIS**

The data analysis focused on aggregating the malaria cases recorded in Colombia by sex, ethnicity, migrant status, pregnancy, and hospitalization. The years covered in the analysis were 2015 to 2023 to determine the pattern of occurrence of the disease among people in Colombia.

**Cases by sex**

|  |  |  |  |
| --- | --- | --- | --- |
| **Row Labels** | **F** | **M** | **Grand Total** |
| **2015** | **9838** | **14317** | **24155** |
| **2016** | **13975** | **19741** | **33716** |
| **2017** | **8938** | **12961** | **21899** |
| **2018** | **11886** | **17817** | **29703** |
| **2019** | **14822** | **21618** | **36440** |
| **2020** | **15702** | **23353** | **39055** |
| **2021** | **14742** | **20742** | **35484** |
| **2022** | **17775** | **25668** | **43443** |
| **2023** | **270** | **418** | **688** |
| **Grand Total** | **107948** | **156635** | **264583** |

Table 1 shows the total occurrence of Malaria by sex between 2015 and 2023.

A total of 264,583 cases of Malaria were reported, with more cases reported in males than females. Notably, there was a decline in the number of cases recorded in 2017, followed by a rise in cases in 2018. A year-by-year analysis also revealed more yearly cases in males than females, while 2022 recorded the highest number of cases of Malaria in Colombia between 2015 and 2023.

**Cases by Social Security**

Chart 1: Visual representation of cases by social security.

The stacked column chart above shows the yearly distribution of cases among individuals in different social security classes. S refers to subsidized people, P shows the retired population, N for not covered, I for unknown status, E for a particular category that includes military, police officers, and public teachers, and C for Contributory, the working Colombians.

The S social security class reported the highest cases, while E had the lowest. It is also interesting to note that the E Social Security class had no reported case of Malaria in the data collected between 2021 and 2023.

**Cases by Ethnicity**

Chart 2: Visual representation of cases by ethnicity.

The ethnic groups were coded using numbers 1 to 6, each representing a particular ethnic group. 1 indicates Indigenous People, 2 Romani, 3 Raizal, which refers to an ethnic group that lives on the Colombian Caribbean islands, 4 Palenque, another unique ethnic group that only inhabits the northern swamps of Colombia, 5 for Afro, and 6 for others.

It was expected that most cases were suffered by the population in the “Others” category, as it is most inhabitants in Colombia, with a total of 119.912 cases. Moreover, Indigenous and Afro communities collected 93.976 and 49.852 cases, respectively. In Colombia, Indigenous and Afro communities historically have been affected by state abandonment (Rocha, 2020), increasing their risk of suffering from diseases like Malaria.

**Cases by migrant status**

Chart 3: Visual representation of data by migration status.

The migration status was encoded using numbers 1 and 2, where 1 represents a migrant, and 2 represents a citizen of Colombia. From the bar chart above, the citizens (coded as '2') have reported more malaria cases yearly, with the highest number recorded in 2022. The rise of migrant cases in Colombia corresponds with the exodus of Venezuelan people due to the political and social situation in their country, making Colombia the primary receptor of Venezuelans (Arena, n.d.).

**Cases by Pregnancy Status**

Chart 4: a visual representation of analysis of case by pregnancy status.

For females, the immunological changes that occur during pregnancy put pregnant women living in malaria-prone regions at high risk of developing Malaria (Rogerson, 2017). Pregnant women residing in malaria-endemic areas face an elevated risk of contracting Malaria due to immunological alterations that accompany pregnancy (Rogerson, 2017). The depicted graph illustrates, in line 1, the recorded instances of pregnant women and, in line 2, the occurrences of pregnant women who experienced malaria infection. This prevalence escalated annually from 1% to 2% between 2015 and 2019. Furthermore, a notable surge of 5% in malaria cases among pregnant women occurred from 2020 to 2022. 2023, there will be a dramatic drop, probably due to a lack of registration and because the year has yet to end.

**Cases by Hospitalization**

Chart 5: visual representation of cases by hospitalization.

The clustered column chart compares the number of cases by hospitalization. The data was grouped into two and coded using numbers 1 and 2, where 1 represented malaria cases that led to hospitalization and two represented cases that didn't. From the chart above, the highest number of cases that led to hospitalization and those that did not was recorded in 2022.

**Cases by Age Group**

Chart 6: Classification of cases by age group.

The histogram above shows that the highest cases of Malaria reported in Colombia between 2015 and 2023 were in the age group 11 – 20, closely followed by 0 to 10 years, while the least number of cases was recorded in the elderly aged 81 and above. This result confirms that children and young adults are at higher risk of Malaria, as has been reported by other authors (Cardona-Arias, 2019).

**DATA TRANSFORMATION**

This Project is based on a structural database, as data was recognized by renaming, moving, or combining columns and techniques. The technique used was aggregation, for accurate analysis and reporting. Some benefits of this data transformation technique are that we could have better organization, improved data quality, and simplified data management.

**DATA AUDIT**

The data obtained was audited prior to analysis because we needed to ascertain that it would meet our objectives. This was done by checking to ensure that data for all years under review (2015 – 2023) were available, after which data was cleaned, trimmed, and made into a proper format fit for our analysis. Pivot tables were created for all variables analyzed to give insights into the details of the data available and spot if there is presence of any abnormal value. The cleaned data was analyzed and charts were drawn.

**DATA DESIGN**

For the data design of the present Project, after gathering the datasets for each year, all of them were joined together in a single dataset. All the variable names were translated from Spanish to English, and several variables were deleted for different reasons. For example, it was not relevant for our purposes if the individual was in prison or a former guerrilla. Also, the Metadata sheet was created to explain the meaning of the codified values for each variable.

**ASSUMPTION**

Given the data extracted from the National Institute of Health of Colombia about the cases of Malaria between 2015 and 2023, despite the efforts made by The Ministry of Health, the cases have continued to increase year by year and seem to continue growing in numbers. Although it is thought that the migration of people from Venezuela to Colombia was the main reason for the rise in the number of cases of Malaria in the country, our data shows that the proportion of cases of migrants is deficient compared to the number of cases among Colombians. A possible reason for this is that Colombians, besides the migrant population, might act as an enormous reservoir where the mosquitoes get infected. Also, migrants may have increased the number of cases as they mostly traveled on foot through Colombia, passing by endemic regions of Malaria, where they could get infected.

On the other hand, groups like Indigenous people and Afro communities also showed a steadier increment in the number of cases of Malaria compared to the “Other” population. This could be explained by the fact that migrants were mainly located in big cities, where the proportion of Indigenous and Afro communities is lower. At the same time, major Indigenous and Afro communities are in places with difficult access and poor economic and social conditions that couldn’t represent an attractive place to settle as a migrant.

Finally, the age distribution of malaria cases in Colombia shows a similar behavior than in other countries where Malaria is also present. This means that young adults of working age are at higher risk of Malaria, as well as people with nocturnal habits and five years old and younger children (Cardona-Arias, 2019). Young adults spend more time outdoors and have a more active nightlife than older adults are more suspected to get infected with Malaria.

**REFERENCES**

Arena, P. M., Corugedo, E. F., & Yepez, J. G. y J. F. (n.d.). Los migrantes venezolanos brindan oportunidades económicas a América Latina. IMF. Retrieved December 9, 2023, from https://www.imf.org/es/News/Articles/2022/12/06/cf-venezuelas-migrants-bring-economic-opportunity-to-latin-america

Cardona-Arias, J. A., Salas-Zapata, W. A., & Carmona-Fonseca, J. (2019). Determinación y determinantes sociales de la malaria: Revisión sistemática, 1980-2018. Revista Panamericana de Salud Pública, 43, e39. https://doi.org/10.26633/RPSP.2019.39

Recht, J., Siqueira, A. M., Monteiro, W. M., Herrera, S. M., Herrera, S., & Lacerda, M. V. G. (2017). Malaria in Brazil, Colombia, Peru and Venezuela: current challenges in malaria control and elimination. *Malaria Journal 2017 16:1*, *16*(1), 1–18. https://doi.org/10.1186/S12936-017-1925-6

Rogerson, S. J. (2017). Management of Malaria in pregnancy. *The Indian Journal of Medical Research*, *146*(3), 328. https://doi.org/10.4103/IJMR.IJMR\_1304\_17

Rocha, J. O. (2020, August 19). Las etnias y la pandemia en Colombia [Text]. Universidad de los Andes - Colombia - Sitio oficial. https://uniandes.edu.co/es/noticias/economia-y-negocios/la-etnias-y-pandemia-en-colombia