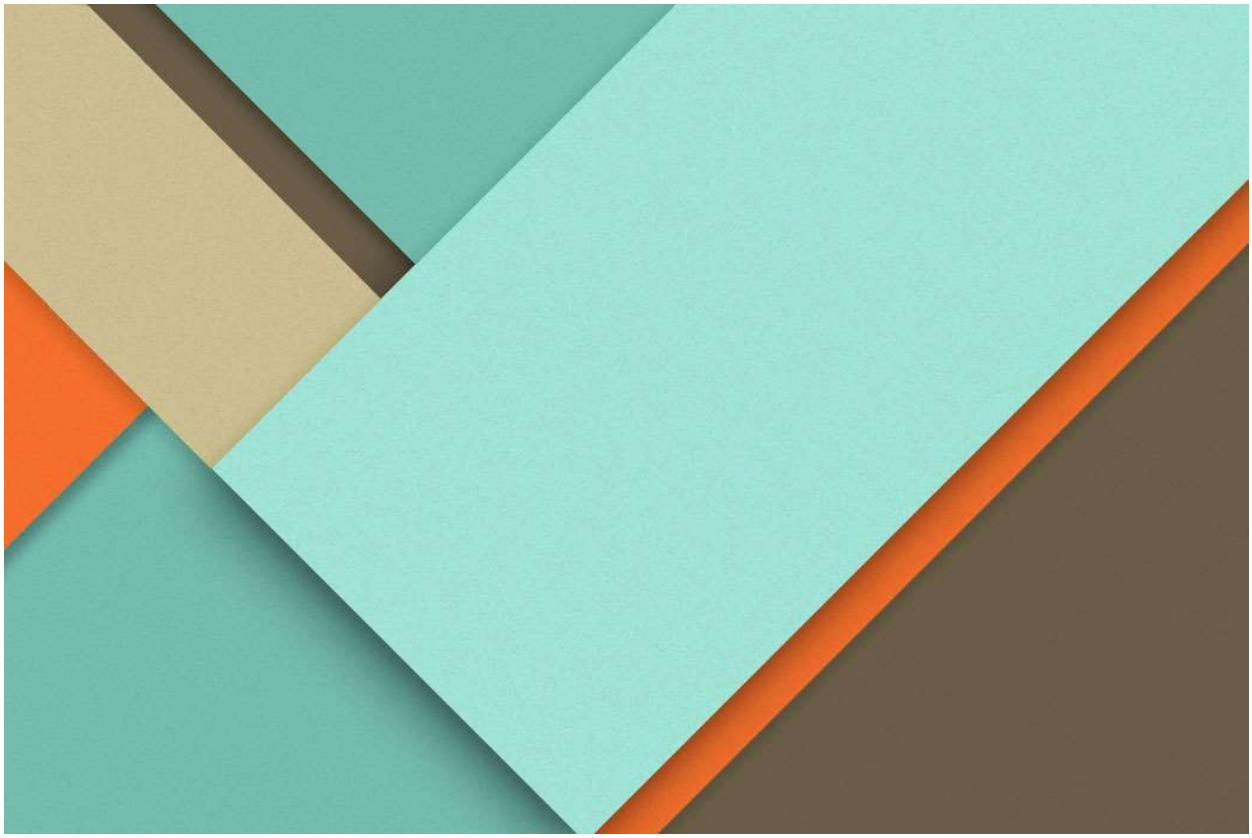


Analysis Report

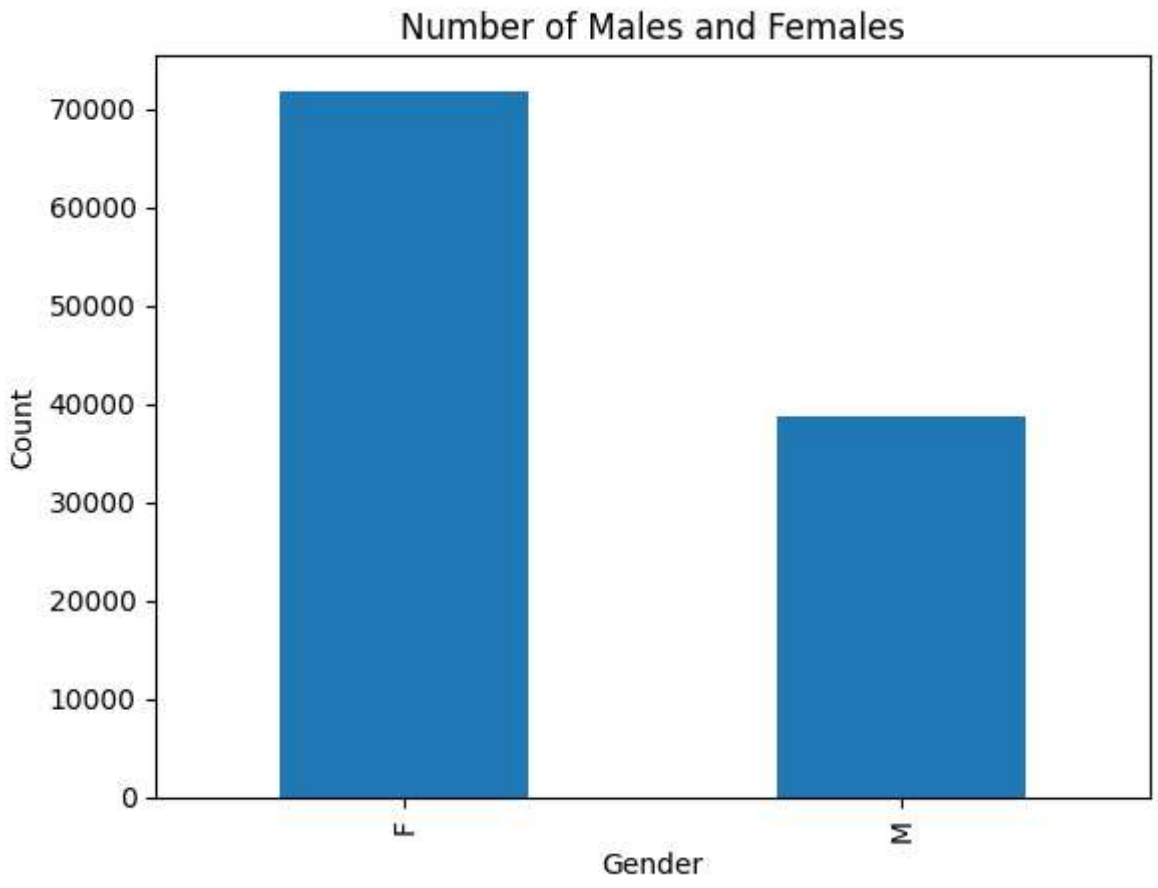
03.06.2023

by: Japanjot Singh

Overview

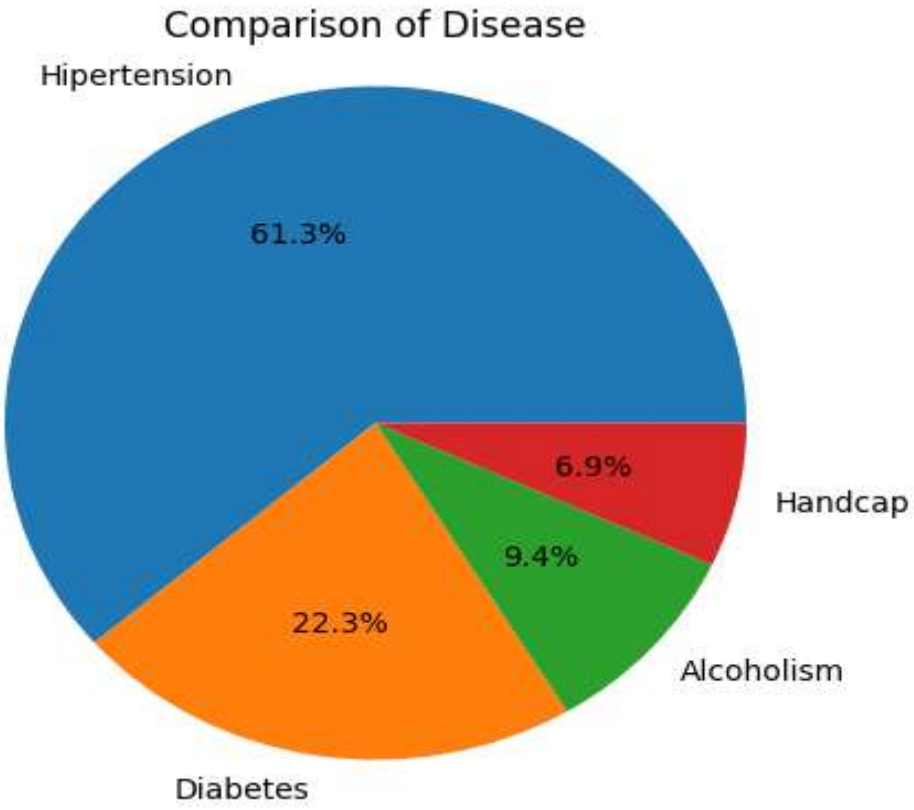
The Clinical Data Analysis Dataset is a comprehensive collection of patient information aimed at exploring various aspects of clinical data, including patient demographics and diseases. This dataset serves as a valuable resource for researchers, healthcare professionals, and data analysts interested in understanding patterns, trends, and relationships within clinical data.

Analysis and Findings



The bar graph presented illustrates a significant disparity in the number of female and male patients within the dataset. The data reveals that the female patient population is approximately double the size of the male patient population, with females accounting for approximately 70,000 individuals compared to approximately 38,000 males. This stark contrast in patient numbers emphasizes the need for gender-specific analysis and considerations when exploring clinical data.

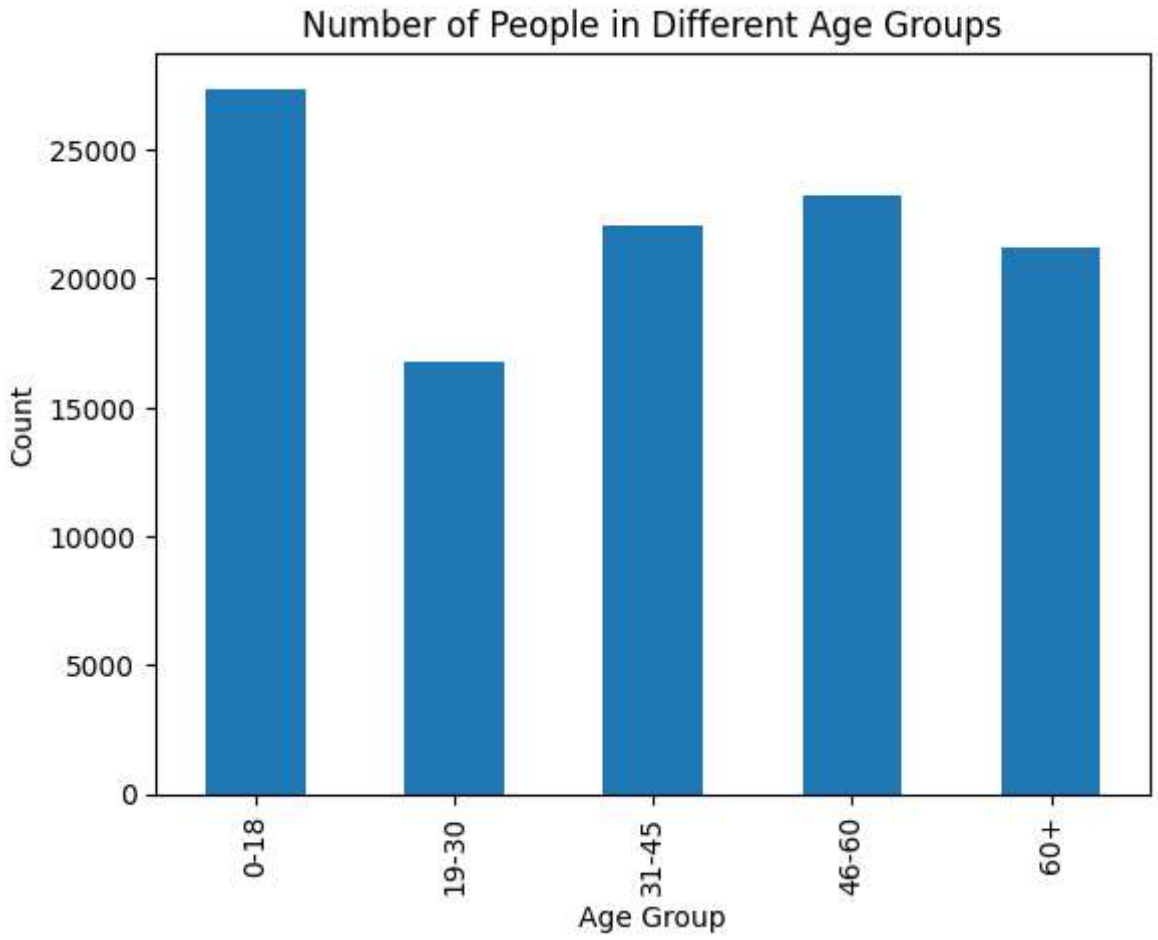
Understanding such demographic variations can provide valuable insights into potential gender-related differences in disease prevalence, treatment outcomes, and overall healthcare planning and delivery.



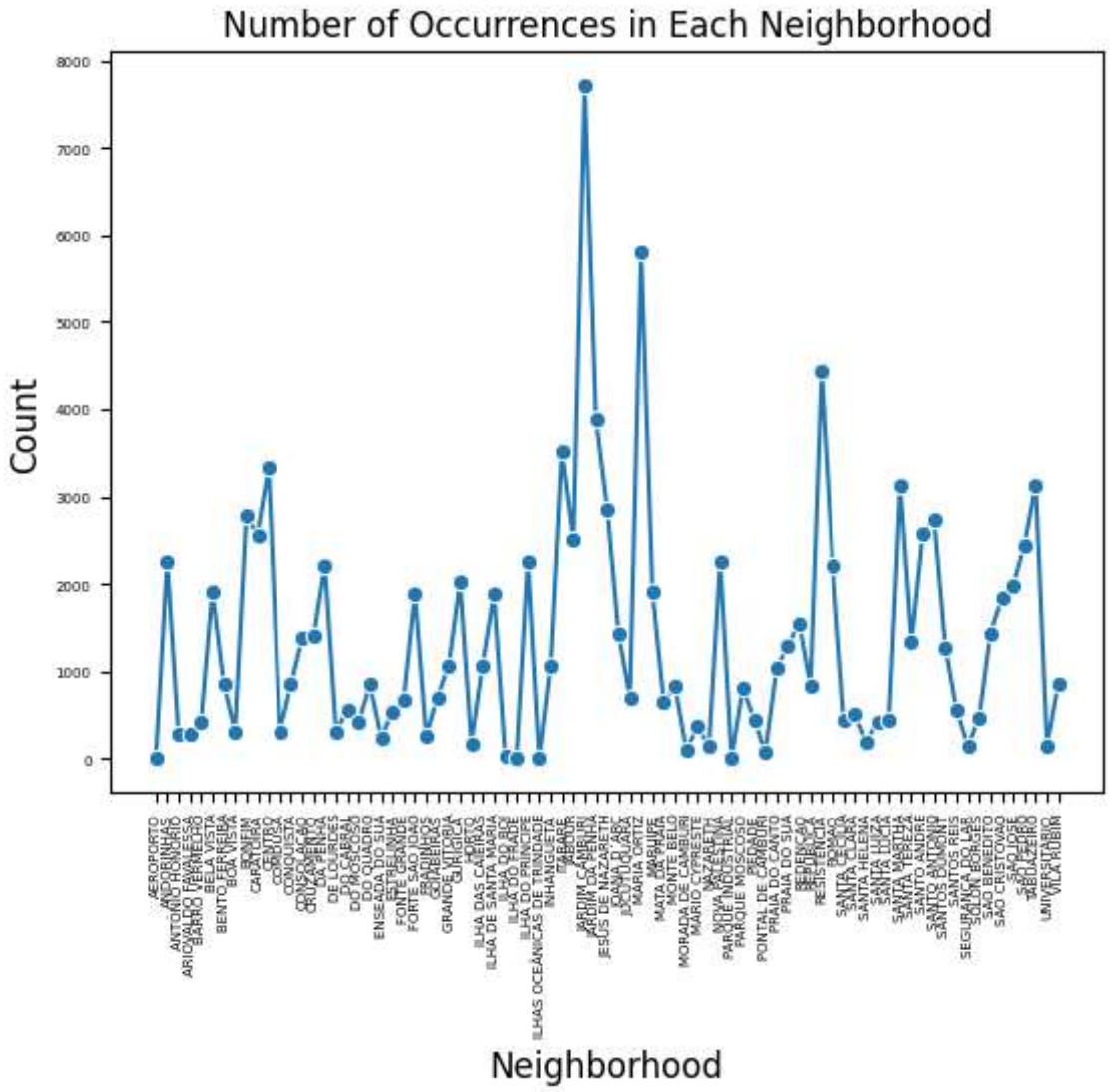
The pie chart above illustrates the distribution of diseases within the Clinical Data Analysis Dataset. The analysis reveals that hypertension has the highest prevalence among patients, followed by diabetes as the second most common condition. Alcoholism ranks third in terms of disease frequency, while the category of handicap exhibits the lowest occurrence.

The pie chart provides a visual representation of the relative proportions of these diseases within the dataset. Hypertension appears to have the largest slice, indicating a higher prevalence compared to the other diseases. Diabetes is depicted with the second largest portion, signifying its substantial presence within the patient population. Alcoholism is depicted by a moderately sized slice, suggesting a notable but comparatively lower occurrence. Lastly, the category of handicap exhibits the smallest slice, indicating the least prevalence among the listed diseases.

This information is valuable for researchers, healthcare professionals, and policymakers aiming to gain insights into the distribution and relative frequencies of these diseases. By understanding the prevalence of specific conditions, healthcare providers can allocate appropriate resources and design targeted interventions to address the healthcare needs of patients affected by these diseases.



The graph depicts the distribution of clinic visits among patients across different age groups. Notably, patients aged 0-18 display the highest attendance rate, followed by individuals in the 46-60 and 31-45 age brackets. Patients aged 60 and above demonstrate a similar visitation pattern to those in the 19-30 age group, while the latter exhibits the lowest rate of clinic visits.



In the aforementioned line graph, the occurrence of patients from different neighborhoods is visualized, providing insights into their geographical distribution. The graph effectively presents the count of patients associated with each neighborhood, allowing for an analysis of patient origin and geographical patterns. The neighborhoods are represented on the x-axis, while the y-axis denotes the corresponding count. The font size for the x-axis and y-axis labels has been reduced to enhance clarity, and the x-axis labels are rotated for improved readability. Additionally, the graph incorporates a larger figure size, gridlines for better visualization, and an optimized layout. These adjustments contribute to a professional and concise representation of the data.

For source code of this analysis is available at link given below

https://github.com/sjapanjots/Analysis-report-3