

Exploring Demographic and Health Factors: A Comprehensive Analysis of a Multivariate Dataset

We propose to conduct an analysis utilizing a comprehensive dataset encompassing various demographic and health-related variables.

The dataset includes information on individuals' state of residence, **gender**, general health status, physical and mental health days, healthcare utilization such as last checkup time, lifestyle factors like physical activity and **sleep hours**, as well as **medical history** including occurrences of **heart attacks**, **strokes**, asthma, **skin cancer**, COPD, **depressive disorders**, **kidney disease**, arthritis, and **diabetes**, among others. Additionally, the dataset contains data on sensory impairments, difficulties with daily activities, **smoking** and **e-cigarette** usage, chest scans, **race/ethnicity**, **age**, height, weight, **BMI**, **alcohol consumption**, **HIV testing**, flu vaccination, pneumococcal vaccination, tetanus vaccination, high-risk status, and **COVID-19** positivity.

Our analysis aims to investigate various **null hypotheses**, including the relationship between demographic factors and health outcomes, the impact of lifestyle choices on disease prevalence, and disparities in healthcare access and utilization among different population groups. We will employ appropriate **statistical tests** and **regression analyses** to test these hypotheses, providing valuable insights into factors influencing health outcomes and informing public health interventions and policy decisions.

These are our hypotheses questions:

1. Is there a significant difference in **general health** between **males** and **females**?
2. Is there a relationship between the **number of physical health days** and **mental health days**?
3. Does the time since the **last checkup** significantly affect **physical health**?
4. Are there **differences** in **physical activity levels** between individuals of **different ages**?

5. Is there an **association** between **sleep hours** and the **number of removed teeth**?
6. Is there a significant **difference** in the **prevalence** of **heart attacks** between **different races/ethnicities**?
7. Does **smoking** status impact the likelihood of having had **asthma**?
8. Is there a relationship between **BMI** and the presence of **diabetes**?
9. Do individuals with a history of **skin cancer** have a higher likelihood of also having **COPD**?
10. Is there a difference in **HIV testing** rates between **deaf** or **hard of hearing individuals** and those **without hearing difficulties**?
11. Is there a significant difference in **physical health days** between individuals who have had a **heart attack** and those who have not?
12. Do individuals with a history of **angina** experience more **mental health days** compared to those without?
13. Is there an association between the occurrence of **strokes** and the presence of **arthritis**?
14. Does the frequency of **asthma** episodes differ between individuals who are **deaf** or **hard of hearing** and those with **no hearing difficulties**?
15. Is there a relationship between **skin cancer** history and **difficulties in walking**?
16. Do individuals with **COPD** report more **difficulty in dressing and bathing** compared to those without?
17. Is there a difference in the **prevalence of depressive disorders** among **smokers** and **non-smokers**?
18. Are individuals with **kidney disease** more likely to have **diabetes** compared to those **without kidney issues**?
19. Does **age** impact the likelihood of having **skin cancer**?
20. Is there a significant difference in **BMI** between individuals who have had **skin cancer** and those who have not?
21. Do **alcohol drinkers** report more **physical health days** compared to **non-drinkers**?
22. Is there an association between **HIV testing** and the **presence of heart conditions** such as **angina** or **heart attacks**?

23. Does receiving the **flu vaccine** in the **last 12 months** affect the likelihood of having had **asthma**?
24. Are individuals who have received the **pneumococcal vaccine** more likely to have had a **stroke**?
25. Is there a relationship between **tetanus vaccination** in the **last 10 years** and difficulties in **concentrating**?
26. Does **high-risk status** impact the **frequency of physical activities**?
27. Are there **differences** in **COVID-19 positivity rates** between **different racial/ethnic groups**?
28. Is there an association between **smoking status** and the **frequency of chest scans**?
29. Do individuals with a **history of heart attacks** report **more difficulty in errands** compared to those without?
30. Is there a **difference in sleep hours** between **individuals who have had a chest scan** and those who have not?

These questions aim to explore various relationships and associations within the dataset, providing insights into health outcomes and potential factors influencing them.

Team members and contributions:

<i>Mohamed Ibrahim</i>	<i>22-101058</i>	<i>questions assigned from 1-6.</i>
<i>Mohamed Mostafa</i>	<i>22-101203</i>	<i>questions assigned from 7-12.</i>
<i>Mohamed Karim</i>	<i>22-101273</i>	<i>questions assigned from 13-18.</i>
<i>Seif El Islam</i>	<i>22-101172</i>	<i>questions assigned from 19-25.</i>
<i>Abd Allah Moussa</i>	<i>22-101114</i>	<i>questions assigned from 26-30.</i>