

Appendix

Tables

Table 1: Summary of Predictive Performance (AUC) Across Chronic Disease Flags.

This table reports the number of positive cases and the Area Under the Receiver Operating Characteristic Curve (AUC) for each modeled disease condition using the XGBoost classifier. AUC values quantify the model's ability to discriminate between cases and controls. Higher AUC indicates better predictive performance. Predictive accuracy was moderate for common conditions like Hypertension (AUC = 0.678) and Type 2 Diabetes (AUC = 0.729), while rarer conditions such as Chronic Hepatitis and Asthma exhibited lower AUCs, reflecting limited training data and class imbalance.

| Disease Flag | Positive Cases | AUC |
|--------------------------------|----------------|----------|
| Has Hypertension | 3315 | 0.677725 |
| Has Type 2 Diabetes | 1940 | 0.728739 |
| Has Hypertensive Heart Disease | 1477 | 0.620277 |
| Has Type 1 Diabetes | 82 | 0.619078 |
| Has Asthma | 32 | 0.529488 |
| Has Epilepsy | 32 | 0.587534 |
| Has Chronic Hepatitis | 32 | 0.451427 |

Table 2: Ch-Statistic Summary Summary of Statistical Significance and SHAP-Based Feature Contributions Across Disease and Sociodemographic Groups. This comprehensive table aggregates two complementary analyses: **Tab1:** Chi-Square Association Tests: For each chronic disease, we report the Chi-squared test statistic (chi2), degrees of freedom (dof), and associated p-value (p value) assessing the independence between the disease and each socio demographic group (tribe, religion, income, education). Statistically significant associations (e.g., $p < 0.05$) suggest meaningful disparities in disease prevalence across these subgroups. For example, hypertension and Type 2 diabetes show highly significant variation by tribe and religion. **Tab2:** SHAP Feature Importance: For each disease-group pair, we provide the mean absolute SHAP values that quantify the magnitude of each feature's contribution to the model's

prediction. Higher values indicate greater influence on disease risk prediction. Key features consistently influencing model outputs include BMI, religion affiliations, and types of cooking oil (e.g., palm oil, vegetable oil), underscoring their relevance in shaping health outcomes across Nigerian populations. This dual-view framework enables both statistical validation and model interpretability for uncovering actionable health disparities.

| Tab1: Ch-Statistic Summary <i>Summary of Statistical Significance</i> | | | | |
|--|-------------------------------------|----------|----------|-----|
| Group | Disease | chi2 | P value | dof |
| Tribe | Has Hypertension | 1601.446 | 0 | 342 |
| Tribe | Has Hypertensive Heart Disease | 801.3073 | 7.03E-39 | 342 |
| Tribe | Has Acute Myocardial Infarction | 16.23648 | 1 | 342 |
| Tribe | Has Type 2 Diabetes | 2032.685 | 0 | 342 |
| Tribe | Has Type 1 Diabetes | 308.0206 | 0.906454 | 342 |
| Tribe | Has Alcoholic Liver Disease | 54.91963 | 1 | 342 |
| Tribe | Has Fibrosis And Cirrhosis Of Liver | 118.2206 | 1 | 342 |
| Tribe | Has Chronic Kidney Disease | 39.85492 | 1 | 342 |
| Tribe | Has Asthma | 1522.934 | 0 | 342 |
| Tribe | Has Depressive Episode | 75.10267 | 1 | 342 |
| Tribe | Has Polyarthrosis | 190.5664 | 1 | 342 |
| Religion | Has Hypertension | 412.0295 | 7.24E-86 | 6 |
| Religion | Has Hypertensive Heart Disease | 6.191814 | 0.40205 | 6 |

| | | | | |
|----------|-------------------------------------|----------|----------|---|
| Religion | Has Acute Myocardial Infarction | 0.010674 | 1 | 6 |
| Religion | Has Type 2 Diabetes | 306.7057 | 2.99E-63 | 6 |
| Religion | Has Type 1 Diabetes | 24.17069 | 0.000486 | 6 |
| Religion | Has Alcoholic Liver Disease | 3.818785 | 0.701184 | 6 |
| Religion | Has Fibrosis And Cirrhosis Of Liver | 0.215856 | 0.999807 | 6 |
| Religion | Has Chronic Kidney Disease | 0.954624 | 0.987274 | 6 |
| Religion | Has Asthma | 4.599536 | 0.5961 | 6 |
| Religion | Has Depressive Episode | 1.829442 | 0.934697 | 6 |
| Religion | Has Polyarthrosis | 1.829442 | 0.934697 | 6 |
| Income | Has Hypertension | 210.8594 | 9.24E-43 | 6 |
| Income | Has Hypertensive Heart Disease | 77.44815 | 1.20E-14 | 6 |
| Income | Has Acute Myocardial Infarction | 3.966233 | 0.681246 | 6 |
| Income | Has Type 2 Diabetes | 126.2489 | 7.92E-25 | 6 |
| Income | Has Type 1 Diabetes | 11.61915 | 0.071025 | 6 |
| Income | Has Alcoholic Liver Disease | 6.455257 | 0.374167 | 6 |
| Income | Has Fibrosis And Cirrhosis Of Liver | 9.75578 | 0.135321 | 6 |

| Income | Has Chronic Kidney Disease | 2.236306 | 0.896734 | 6 |
|--|-------------------------------------|----------|---------------|----|
| Income | Has Asthma | 3.846727 | 0.697409 | 6 |
| Income | Has Depressive Episode | 13.64126 | 0.033911 | 6 |
| Income | Has Polyarthrosis | 0.748842 | 0.993376 | 6 |
| Education | Has Hypertension | 182.2573 | 7.96E-34 | 10 |
| Education | Has Hypertensive Heart Disease | 32.25457 | 0.000363 | 10 |
| Education | Has Acute Myocardial Infarction | 3.463613 | 0.968319 | 10 |
| Education | Has Type 2 Diabetes | 76.64426 | 2.28E-12 | 10 |
| Education | Has Type 1 Diabetes | 12.65154 | 0.243812 | 10 |
| Education | Has Alcoholic Liver Disease | 12.78697 | 0.235827 | 10 |
| Education | Has Fibrosis And Cirrhosis Of Liver | 5.05722 | 0.887323 | 10 |
| Education | Has Chronic Kidney Disease | 5.504093 | 0.855067 | 10 |
| Education | Has Asthma | 22.39077 | 0.013233 | 10 |
| Education | Has Depressive Episode | 13.24154 | 0.210479 | 10 |
| Education | Has Polyarthrosis | 17.50324 | 0.063944 | 10 |
| Tab2: SHAP-Based Feature Contributions Across Disease and Sociodemographic Groups | | | | |
| Disease Flag | Group | Feature | Mean Abs SHAP | |

| | | | | |
|---------------------|-------|--------------------------------|----------|--|
| Has Type 1 Diabetes | Tribe | BMI | 41.92658 | |
| Has Type 1 Diabetes | Tribe | Religion Christian | 0.71624 | |
| Has Type 1 Diabetes | Tribe | Oil Type Palm Oil | 0.679871 | |
| Has Type 1 Diabetes | Tribe | Religion Muslim | 0.281206 | |
| Has Type 1 Diabetes | Tribe | Oil Type Vegetable Oil | 0.21771 | |
| Has Type 1 Diabetes | Tribe | Oil Type Ground Oil | 0.092032 | |
| Has Type 1 Diabetes | Tribe | Oil Type I Don't Know/Not Sure | 0.006335 | |
| Has Type 1 Diabetes | Tribe | Tribe Yoruba Ondo | 0.004505 | |
| Has Type 1 Diabetes | Tribe | Tribe Yoruba Kwara | 0.004505 | |

| | | | | |
|---------------------|----------|-------------------------------|----------|--|
| Has Type 1 Diabetes | Tribe | Tribe Yoruba Kogi | 0.004505 | |
| Has Type 1 Diabetes | Religion | BMI | 52.12078 | |
| Has Type 1 Diabetes | Religion | Oil Type Palm Oil | 0.553007 | |
| Has Type 1 Diabetes | Religion | Oil Type Vegetable Oil | 0.374386 | |
| Has Type 1 Diabetes | Religion | Religion Others | 0.166667 | |
| Has Type 1 Diabetes | Religion | Religion Atheist | 0.166667 | |
| Has Type 1 Diabetes | Religion | Religion Christian | 0.166667 | |
| Has Type 1 Diabetes | Religion | Religion Prefer not to Answer | 0.166667 | |
| Has Type 1 Diabetes | Religion | Religion Traditional | 0.166667 | |

| | | | | |
|---------------------|----------|--------------------------------|----------|--|
| Has Type 1 Diabetes | Religion | Religion MusLIM | 0.166667 | |
| Has Type 1 Diabetes | Religion | Tribe Ukwani (Kwale) Delta | 0.09731 | |
| Has Asthma | Tribe | BMI | 42.70558 | |
| Has Asthma | Tribe | Religion Christian | 0.700473 | |
| Has Asthma | Tribe | Oil Type Palm Oil | 0.678465 | |
| Has Asthma | Tribe | Religion Muslism | 0.297617 | |
| Has Asthma | Tribe | Oil Type Vegetable Oil | 0.211385 | |
| Has Asthma | Tribe | Oil Type Groundnut Oil | 0.098884 | |
| Has Asthma | Tribe | Oil Type I Don't Know/Not Sure | 0.00728 | |
| Has Asthma | Tribe | Tribe 265 | 0.004425 | |
| Has Asthma | Tribe | Tribe 492 | 0.004425 | |
| Has Asthma | Tribe | Tribe Ukwani (Kwale) Delta | 0.004425 | |

| | | | | |
|--------------------------------|----------|-------------------------------|----------|--|
| Has Asthma | Religion | BMI | 51.90283 | |
| Has Asthma | Religion | Oil Type Palm Oil | 0.493374 | |
| Has Asthma | Religion | Oil Type Vegetable Oil | 0.300547 | |
| Has Asthma | Religion | Oil Type Groundnut Oil | 0.183254 | |
| Has Asthma | Religion | Tribe Hausa Kano | 0.178792 | |
| Has Asthma | Religion | Religion Prefer Not to Answer | 0.166667 | |
| Has Asthma | Religion | Religion Christian | 0.166667 | |
| Has Asthma | Religion | Religion Others | 0.166667 | |
| Has Asthma | Religion | Religion Atheist | 0.166667 | |
| Has Asthma | Religion | Religion Muslim | 0.166667 | |
| Has Hypertensive Heart Disease | Tribe | BMI | 42.59139 | |

| | | | | |
|--------------------------------------|-------|---------------------------|----------|--|
| Has Hypertensive Heart Disease | Tribe | Religion Christian | 0.725496 | |
| Has Hypertensive Heart Disease | Tribe | Oil Type Palm Oil | 0.698084 | |
| Has Hypertensive Heart Disease | Tribe | Religion Muslim | 0.272444 | |
| Has Hypertensive Heart Disease | Tribe | Oil Type Vegetable Oil | 0.216807 | |
| Has Hypertensive Heart Disease | Tribe | Oil Tyoe Groundnut Oil | 0.077747 | |
| Has Hypertensive Heart Disease | Tribe | Tribe Yoruba Lagos | 0.004566 | |
| Has Hypertensive Heart Disease | Tribe | Tribe Yoruba Ondo | 0.004566 | |

| | | | | |
|--------------------------------|----------|-------------------------------|----------|--|
| Has Hypertensive Heart Disease | Tribe | Tribe Yoruba Ogun | 0.004566 | |
| Has Hypertensive Heart Disease | Tribe | Tribe Yandang Adamawa | 0.004566 | |
| Has Hypertensive Heart Disease | Religion | BMI | 57.83145 | |
| Has Hypertensive Heart Disease | Religion | Oil Type Palm Oil | 0.552008 | |
| Has Hypertensive Heart Disease | Religion | Oil Type Vegetable Oil | 0.235916 | |
| Has Hypertensive Heart Disease | Religion | Religion Prefer Not to Answer | 0.2 | |
| Has Hypertensive Heart Disease | Religion | Religion Traditional | 0.2 | |

| | | | | |
|--------------------------------|----------|------------------------|----------|--|
| Has Hypertensive Heart Disease | Religion | Religion Christian | 0.2 | |
| Has Hypertensive Heart Disease | Religion | Religion Others | 0.2 | |
| Has Hypertensive Heart Disease | Religion | Religion Muslim | 0.2 | |
| Has Hypertensive Heart Disease | Religion | Oil Type Groundnut Oil | 0.187339 | |
| Has Hypertensive Heart Disease | Religion | Tribe Hausa Kano | 0.149395 | |
| Has Hypertension | Tribe | BMI | 25.81747 | |
| Has Hypertension | Tribe | Religion Christian | 0.743904 | |
| Has Hypertension | Tribe | Oil Type Palm Oil | 0.693254 | |

| | | | | |
|------------------|----------|--------------------------------------|----------|--|
| Has Hypertension | Tribe | Religion Muslim | 0.254653 | |
| Has Hypertension | Tribe | Oil Type Vegetable Oil | 0.217414 | |
| Has Hypertension | Tribe | Oil Type Groundnut Oil | 0.06697 | |
| Has Hypertension | Tribe | Oil Type I Don't Know/Not Sure | 0.010909 | |
| Has Hypertension | Tribe | Oil Type Margarine or Cooking Butter | 0.009136 | |
| Has Hypertension | Tribe | Tribe Ukwane (Kwale) Delta | 0.004464 | |
| Has Hypertension | Tribe | Tribe Uncinda Kaduna | 0.004464 | |
| Has Hypertension | Religion | BMI | 25.99914 | |
| Has Hypertension | Religion | Oil Type Palm Oil | 0.632212 | |

| | | | | |
|------------------|----------|-------------------------------|----------|--|
| Has Hypertension | Religion | Oil Type Vegetable Oil | 0.274079 | |
| Has Hypertension | Religion | Religion Prefer Not to Answer | 0.2 | |
| Has Hypertension | Religion | Religion Others | 0.2 | |
| Has Hypertension | Religion | Religion Muslim | 0.2 | |
| Has Hypertension | Religion | Religion Christian | 0.2 | |
| Has Hypertension | Religion | Religion Traditional | 0.2 | |
| Has Hypertension | Religion | Tribe Hausa Sokoto | 0.118618 | |
| Has Hypertension | Religion | Tribe Hausa Kaduna | 0.095228 | |
| Has Epilepsy | Tribe | BMI | 35.1768 | |
| Has Epilepsy | Tribe | Religion Christian | 0.743026 | |

| | | | | |
|--------------|----------|-------------------------------|----------|--|
| Has Epilepsy | Tribe | Oil Type Palm Oil | 0.669148 | |
| Has Epilepsy | Tribe | Religion Muslim | 0.254868 | |
| Has Epilepsy | Tribe | Oil Type Vegetable Oil | 0.223961 | |
| Has Epilepsy | Tribe | Oil Type Groundnut Oil | 0.094427 | |
| Has Epilepsy | Tribe | Oil Type Prefer Not to Answer | 0.007173 | |
| Has Epilepsy | Tribe | Tribe Yoruba Oyo | 0.004425 | |
| Has Epilepsy | Tribe | Tribe 144 | 0.004425 | |
| Has Epilepsy | Tribe | Tribe Yoruba Lagos | 0.004425 | |
| Has Epilepsy | Religion | BMI | 47.93493 | |
| Has Epilepsy | Religion | Oil Type Palm Oil | 0.555548 | |
| Has Epilepsy | Religion | Oil Type Vegetable Oil | 0.351248 | |
| Has Epilepsy | Religion | Tribe Izon (Ijaw) Bayelsa | 0.18849 | |

| | | | | |
|---------------------|----------|-------------------------------|-----------|--|
| Has Epilepsy | Religion | Religion Others | 0.166667 | |
| Has Epilepsy | Religion | Religion Christian | 0.166667 | |
| Has Epilepsy | Religion | Religion Atheist | 0.166667 | |
| Has Epilepsy | Religion | Religion Muslim | 0.166667 | |
| Has Epilepsy | Religion | Religion Prefer Not to Answer | 0.166667 | |
| Has Epilepsy | Religion | Religion Traditional | 0.166667 | |
| Has Type 2 Diabetes | Tribe | BMI | 25.457336 | |
| Has Type 2 Diabetes | Tribe | Religion Christian | 0.703667 | |
| Has Type 2 Diabetes | Tribe | Oil Type Palm Oil | 0.673111 | |
| Has Type 2 Diabetes | Tribe | Religion Muslim | 0.294905 | |
| Has Type 2 Diabetes | Tribe | Oil Type Vegetable Oil | 0.221578 | |

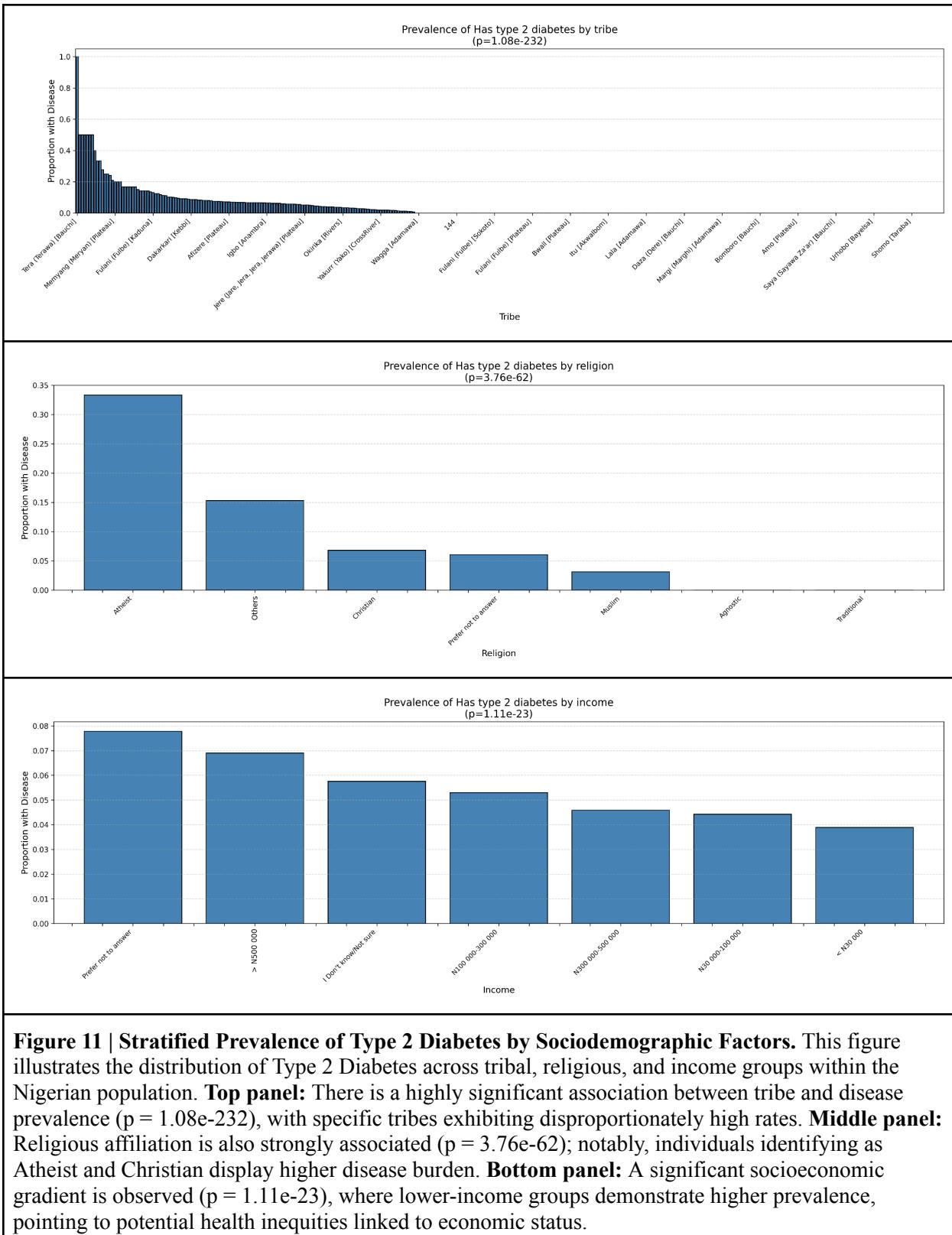
| | | | | |
|---------------------|----------|--------------------------------|----------|--|
| Has Type 2 Diabetes | Tribe | Oil Type Groundnut Oil | 0.095933 | |
| Has Type 2 Diabetes | Tribe | Oil Type I Don't Know/Not Sure | 0.005779 | |
| Has Type 2 Diabetes | Tribe | Tribe 66 | 0.004292 | |
| Has Type 2 Diabetes | Tribe | Tribe Anioma Delta | 0.004292 | |
| Has Type 2 Diabetes | Tribe | Tribe Anioma Edo | 0.004292 | |
| Has Type 2 Diabetes | Religion | BMI | 26.12487 | |
| Has Type 2 Diabetes | Religion | Oil Type Palm Oil | 0.460193 | |
| Has Type 2 Diabetes | Religion | Oil Type Vegetable Oil | 0.424128 | |
| Has Type 2 Diabetes | Religion | Tribe Hausa Kano | 0.183898 | |

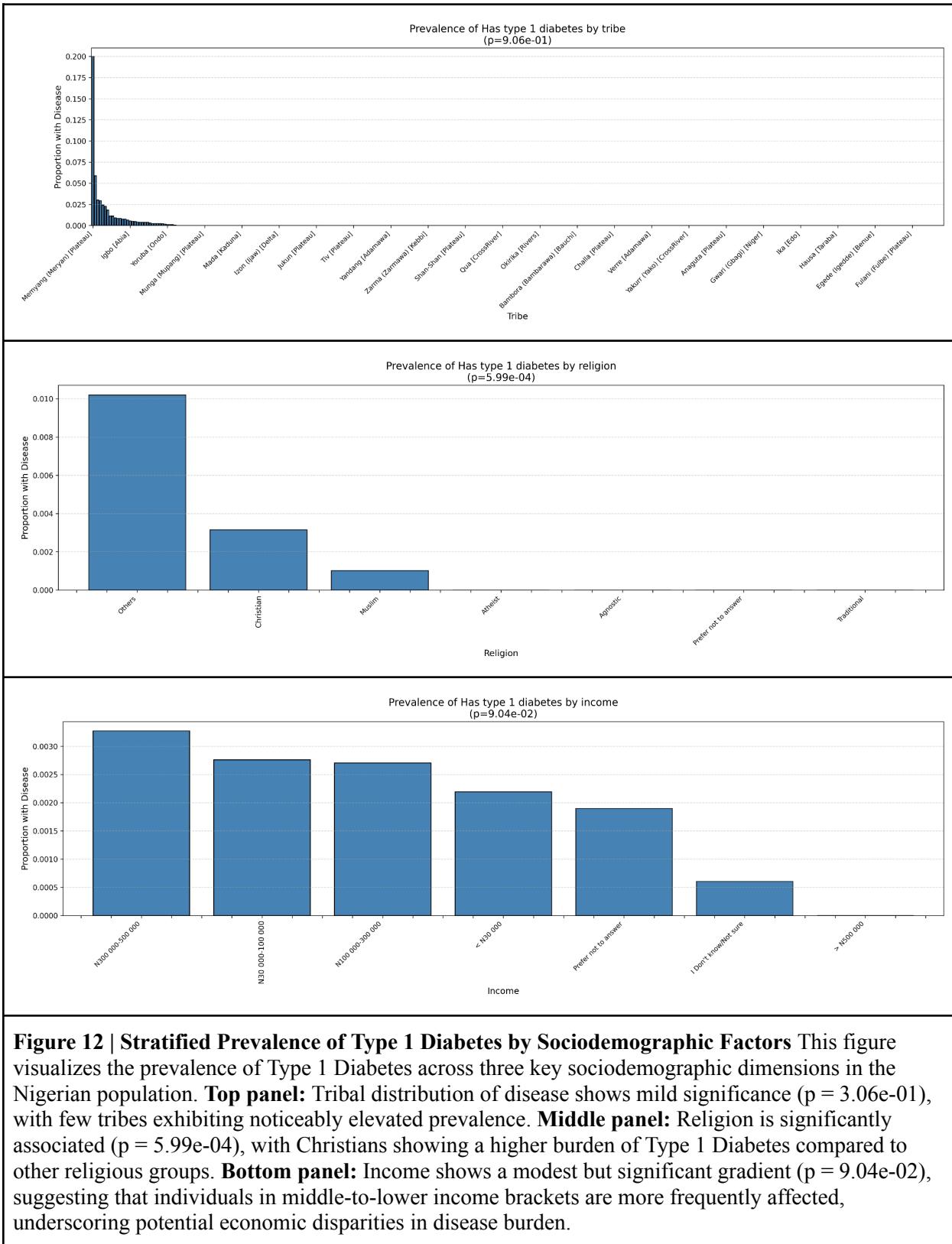
| | | | | |
|-----------------------|----------|------------------------------|----------|--|
| Has Type 2 Diabetes | Religion | Tribe Izon (Ijaw) Bayelsa | 0.18084 | |
| Has Type 2 Diabetes | Religion | Religion Muslim | 0.166667 | |
| Has Type 2 Diabetes | Religion | Religion Christian | 0.166667 | |
| Has Type 2 Diabetes | Religion | Religion Others | 0.166667 | |
| Has Type 2 Diabetes | Religion | Religion Atheist | 0.166667 | |
| Has Type 2 Diabetes | Religion | Religion Traditional | 0.166667 | |
| Has Chronic Hepatitis | Tribe | BMI | 25.20762 | |
| Has Chronic Hepatitis | Tribe | Religion Christian | 0.73037 | |
| Has Chronic Hepatitis | Tribe | Oil Type Palm Oil | 0.681849 | |

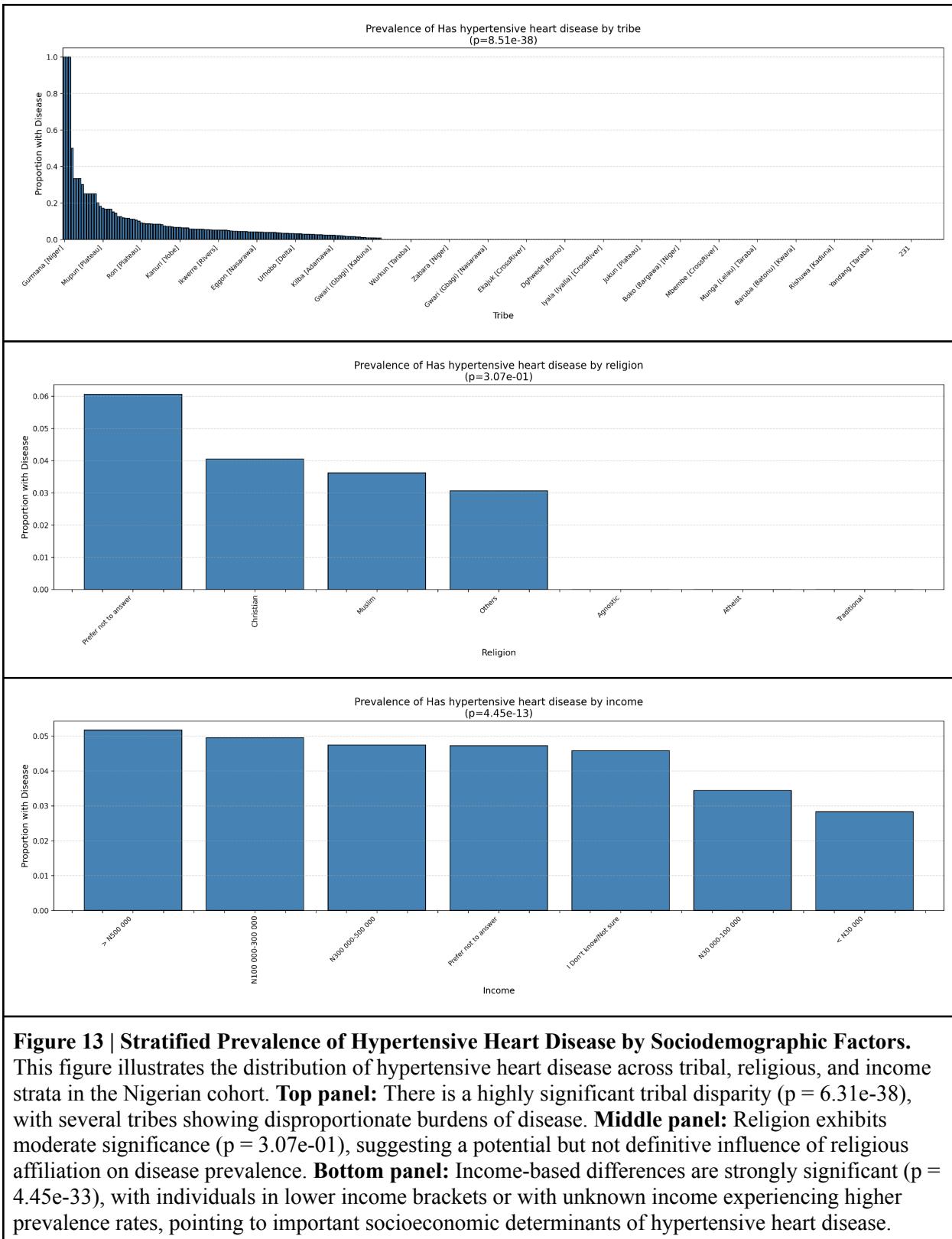
| | | | | |
|-----------------------|----------|---------------------------|----------|--|
| Has Chronic Hepatitis | Tribe | Religion Muslim | 0.268221 | |
| Has Chronic Hepatitis | Tribe | Oil Type Vegetable Oil | 0.20999 | |
| Has Chronic Hepatitis | Tribe | Oil Type Groundnut Oil | 0.092322 | |
| Has Chronic Hepatitis | Tribe | Oil Type I Don't/Not Sure | 0.009402 | |
| Has Chronic Hepatitis | Tribe | Tribe Yoruba Osun | 0.004444 | |
| Has Chronic Hepatitis | Tribe | Tribe Yoruba Kwara | 0.004444 | |
| Has Chronic Hepatitis | Tribe | Tribe Yoruba Kogi | 0.004444 | |
| Has Chronic Hepatitis | Religion | BMI | 27.44751 | |
| Has Chronic Hepatitis | Religion | Oil Type Palm Oil | 0.440001 | |

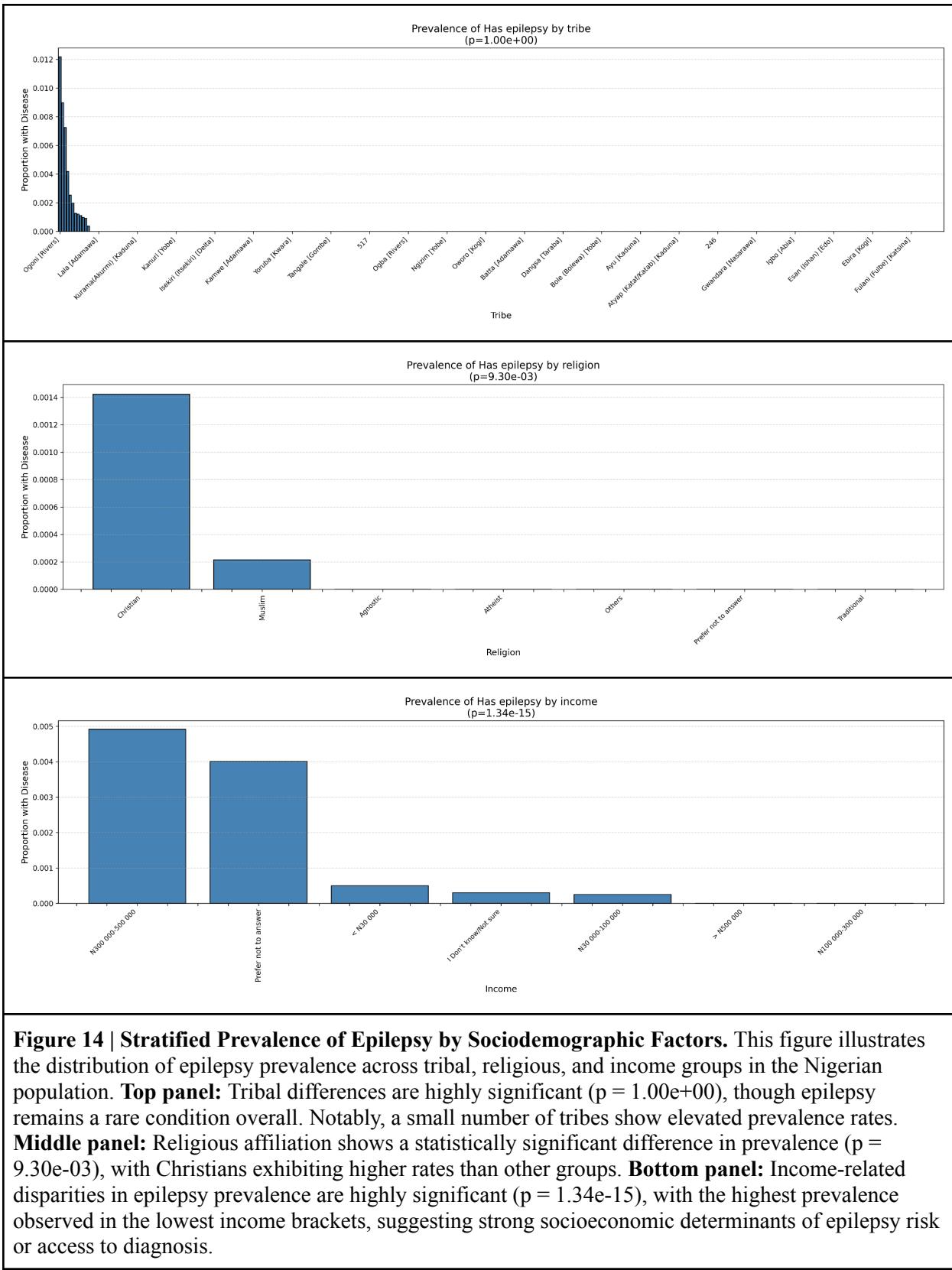
| | | | | |
|-----------------------|----------|-------------------------------|----------|--|
| Has Chronic Hepatitis | Religion | Oil Type Vegetable Oil | 0.362885 | |
| Has Chronic Hepatitis | Religion | Oil Type Groundnut Oil | 0.193135 | |
| Has Chronic Hepatitis | Religion | Tribe Hausa Kano | 0.182986 | |
| Has Chronic Hepatitis | Religion | Religion Atheist | 0.166667 | |
| Has Chronic Hepatitis | Religion | Religion Muslim | 0.166667 | |
| Has Chronic Hepatitis | Religion | Religion Prefer Not to Answer | 0.166667 | |
| Has Chronic Hepatitis | Religion | Religion Others | 0.166667 | |
| Has Chronic Hepatitis | Religion | Religion Christian | 0.166667 | |

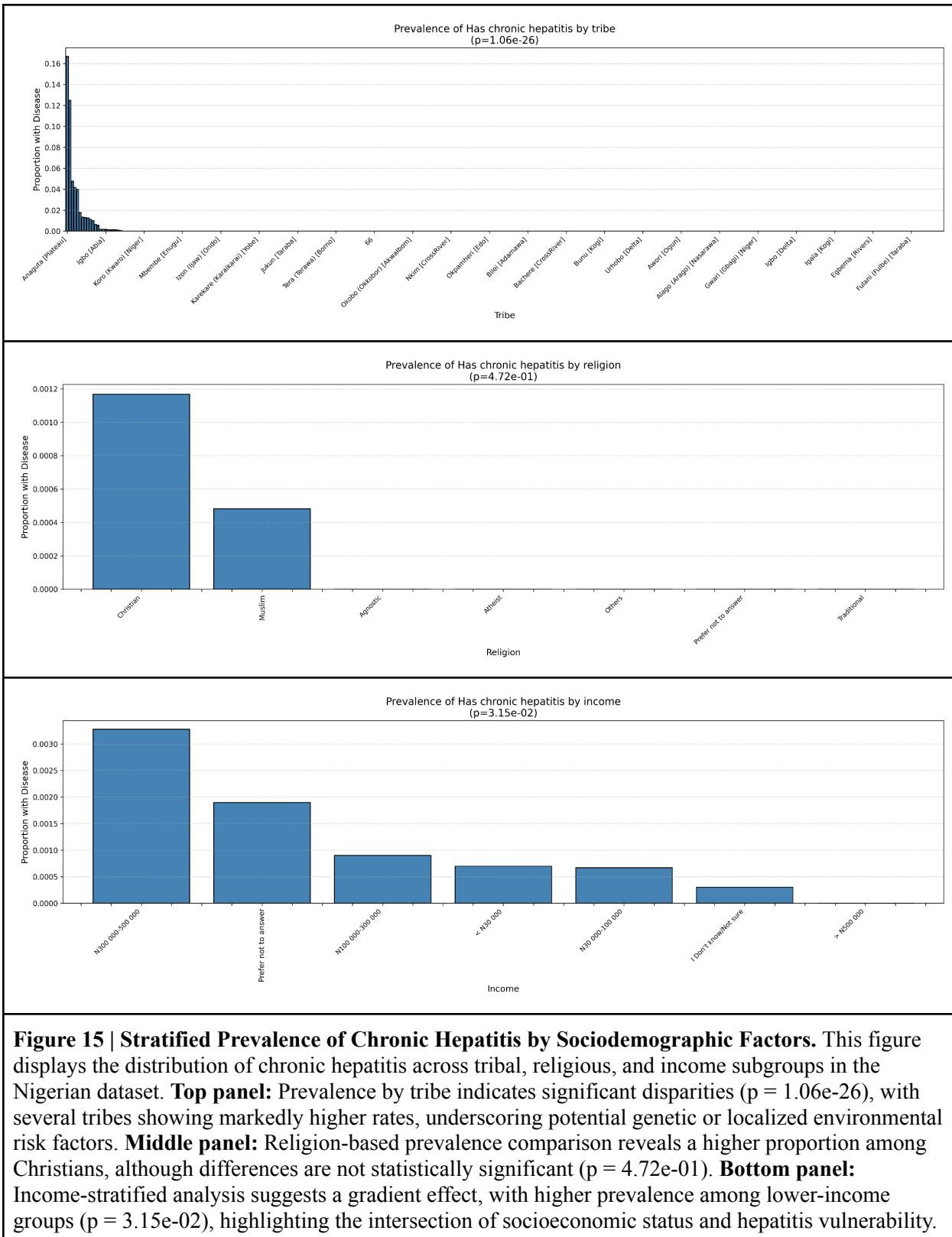
Figures

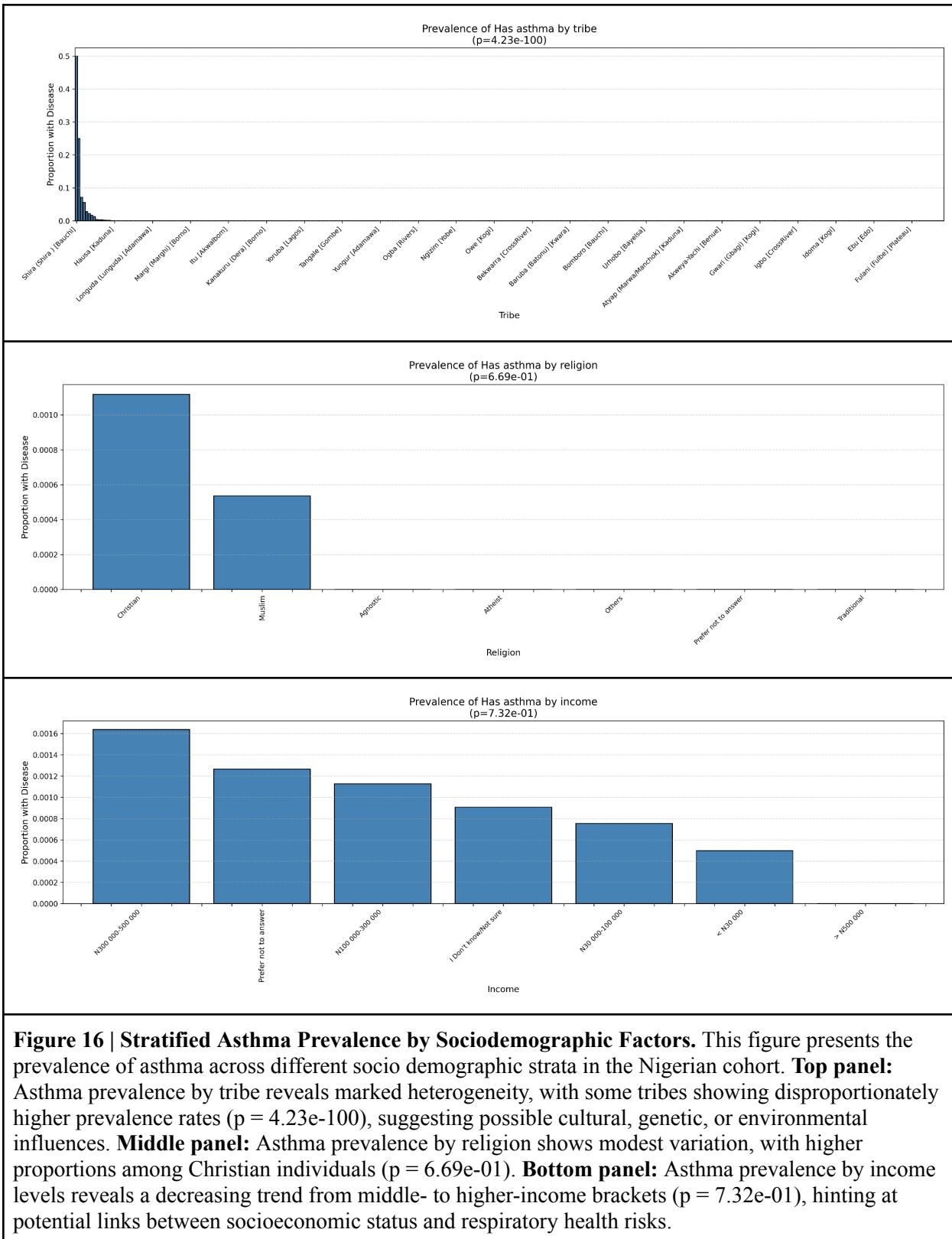












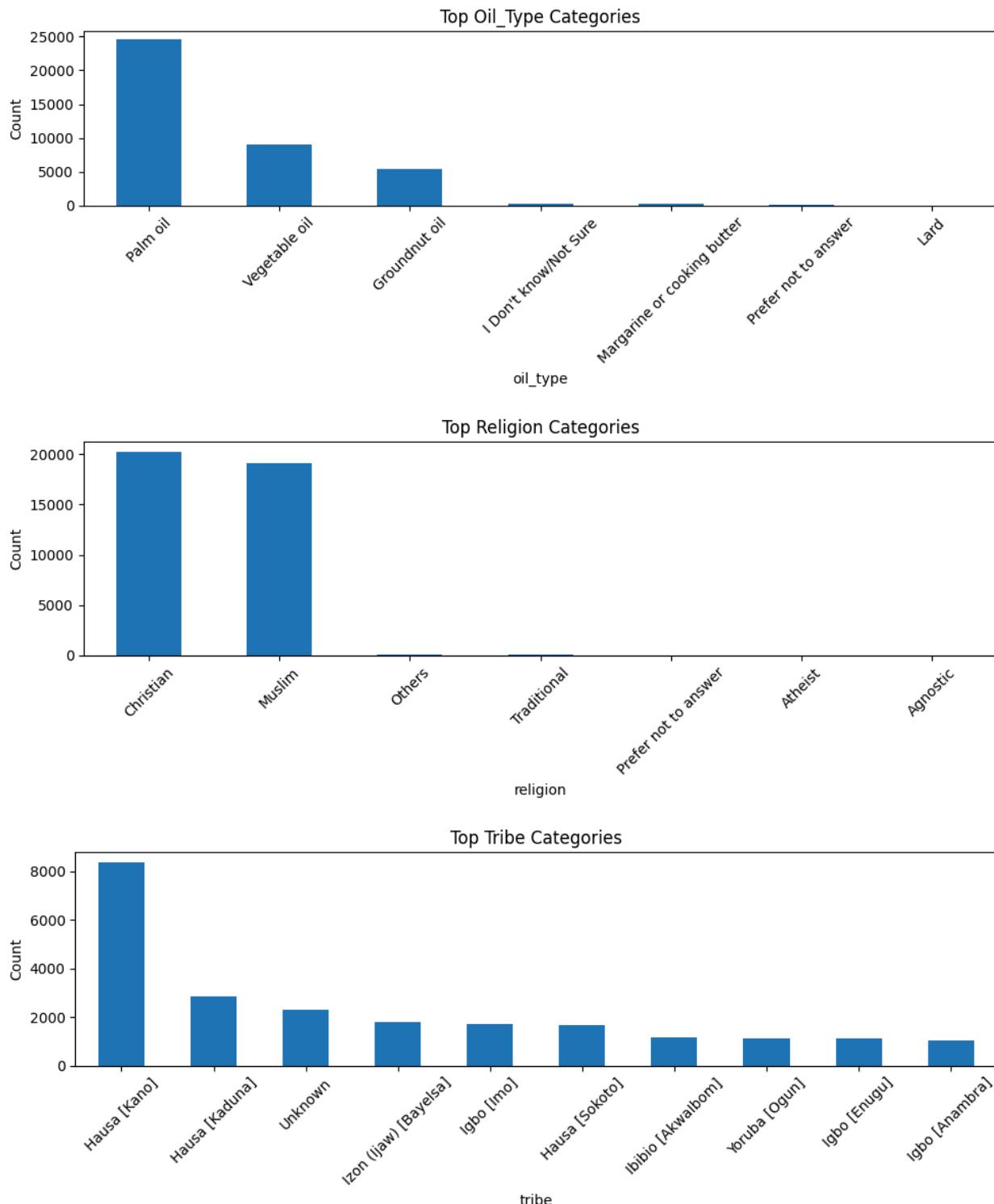
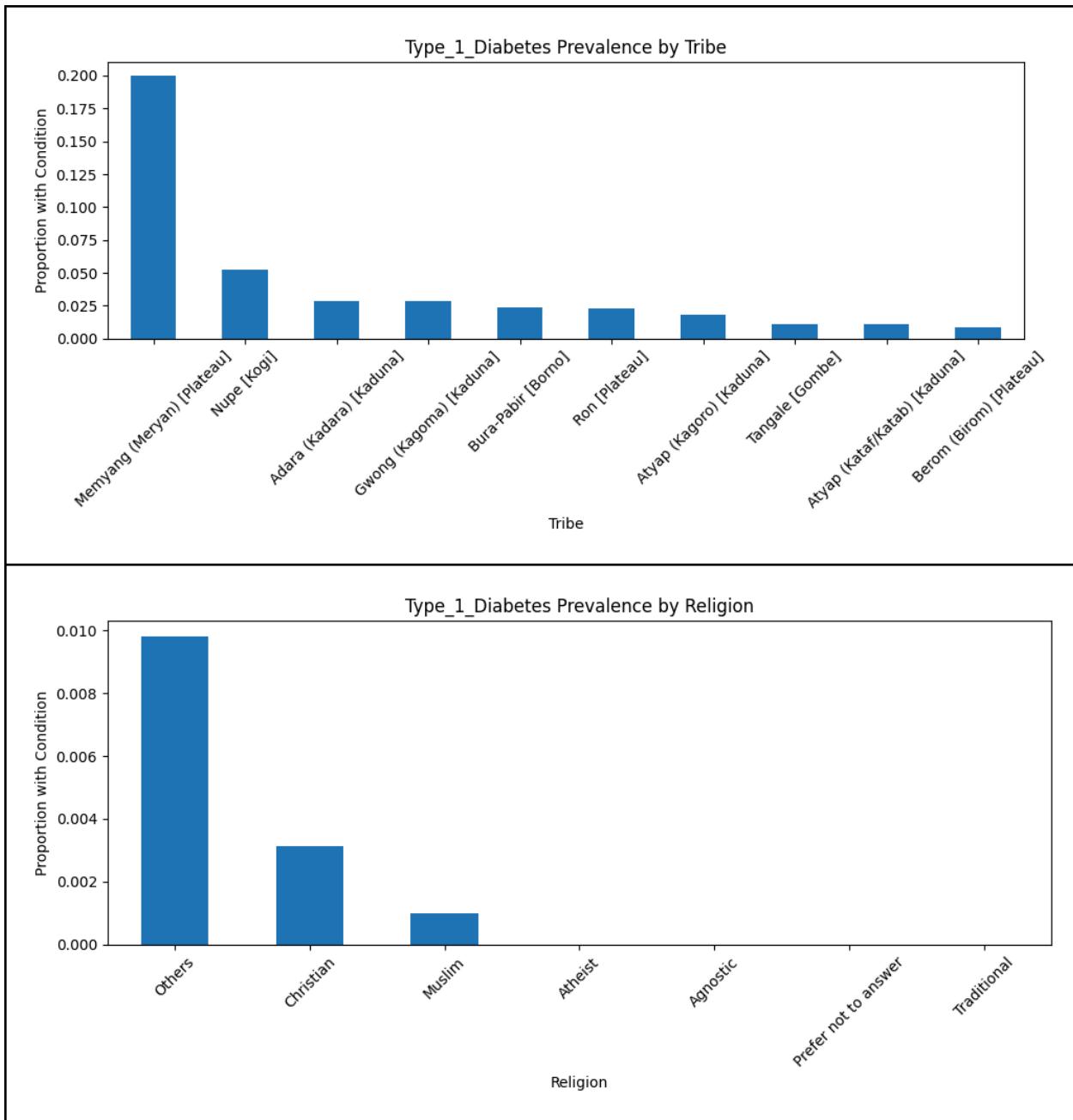
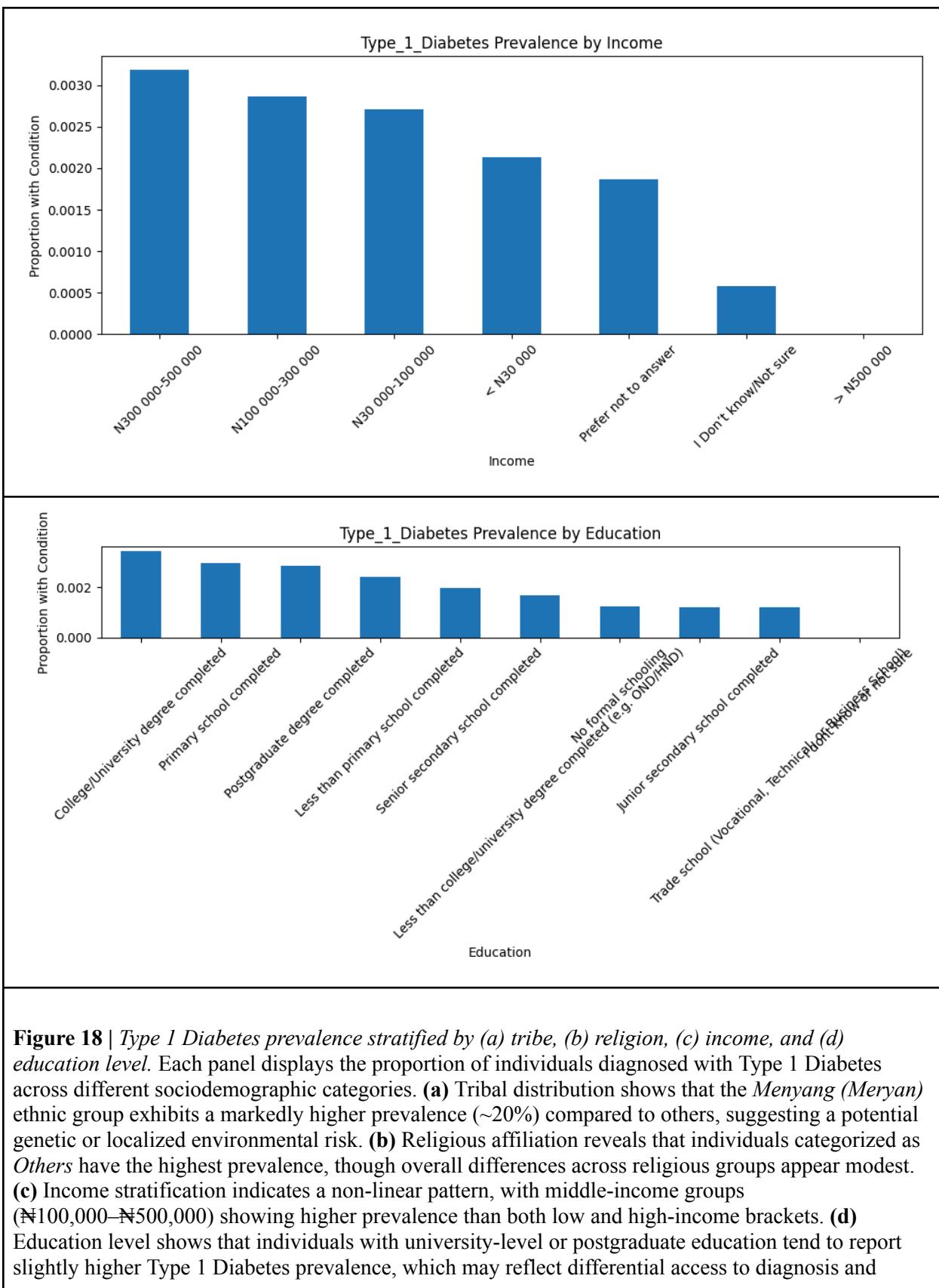


Figure 17 | Distribution of Key Sociodemographic and Dietary Variables. Bar plots showing the top categories for (top) oil type, (middle) religion, and (bottom) tribe reported by participants. Palm oil is the predominant cooking oil, followed by vegetable and groundnut oil. Christianity and Islam are the two major religions represented, with other categories comprising a minor fraction. The Hausa (Kano) tribe is

the most frequently reported tribal affiliation, followed by Hausa (Kaduna), Izon (Ijaw), and various Igbo and Yoruba subgroups.





healthcare awareness rather than causal mechanisms.

These stratified plots highlight potential sociodemographic disparities in disease burden and motivate further investigation into underlying genetic, cultural, and environmental determinants.

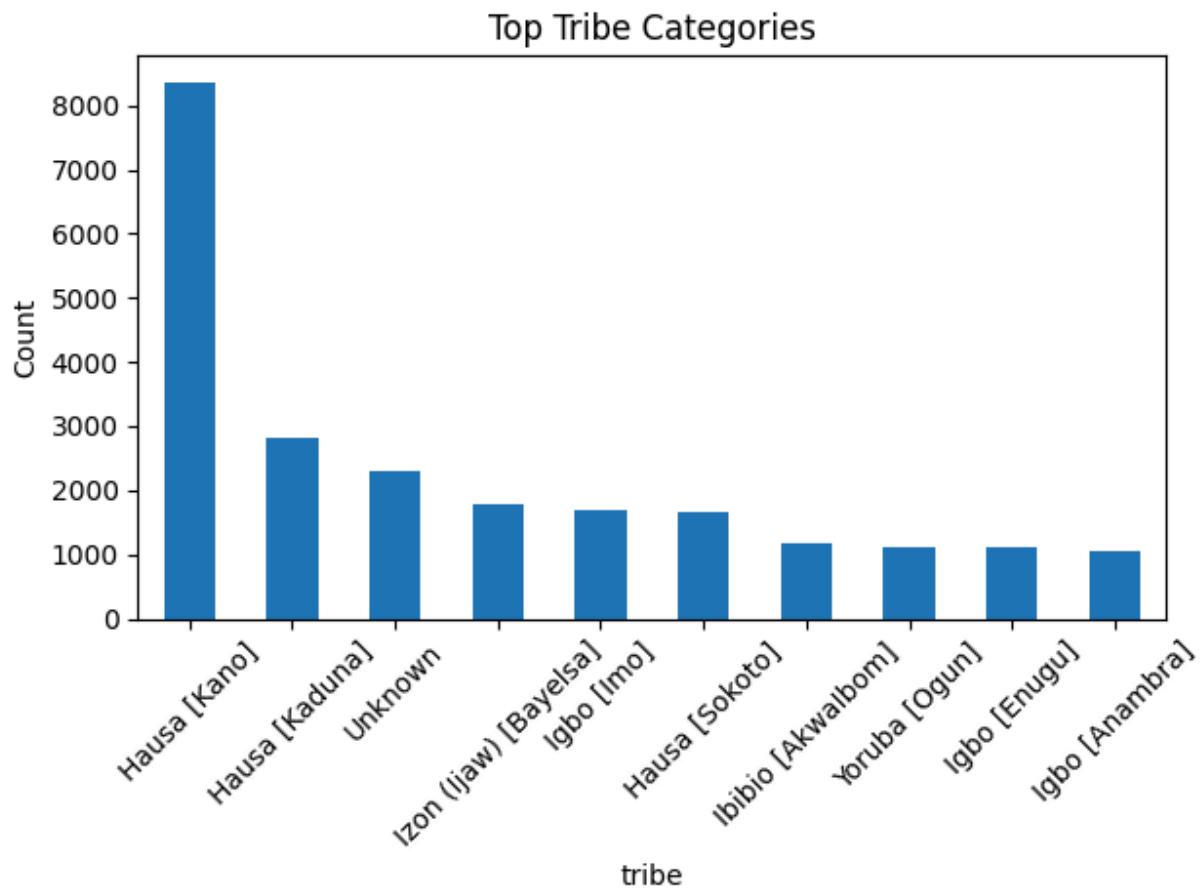


Figure 19 | Distribution of Top Tribe Categories in the Dataset. Bar chart showing the frequency of the top 10 most represented tribes in the study population. The Hausa (Kano) group is the most prevalent, followed by Hausa (Kaduna), Unknown, and Izon (Ijaw) [Bayelsa]. This distribution highlights the demographic composition and potential cultural diversity relevant to downstream sociogenomic analyses.

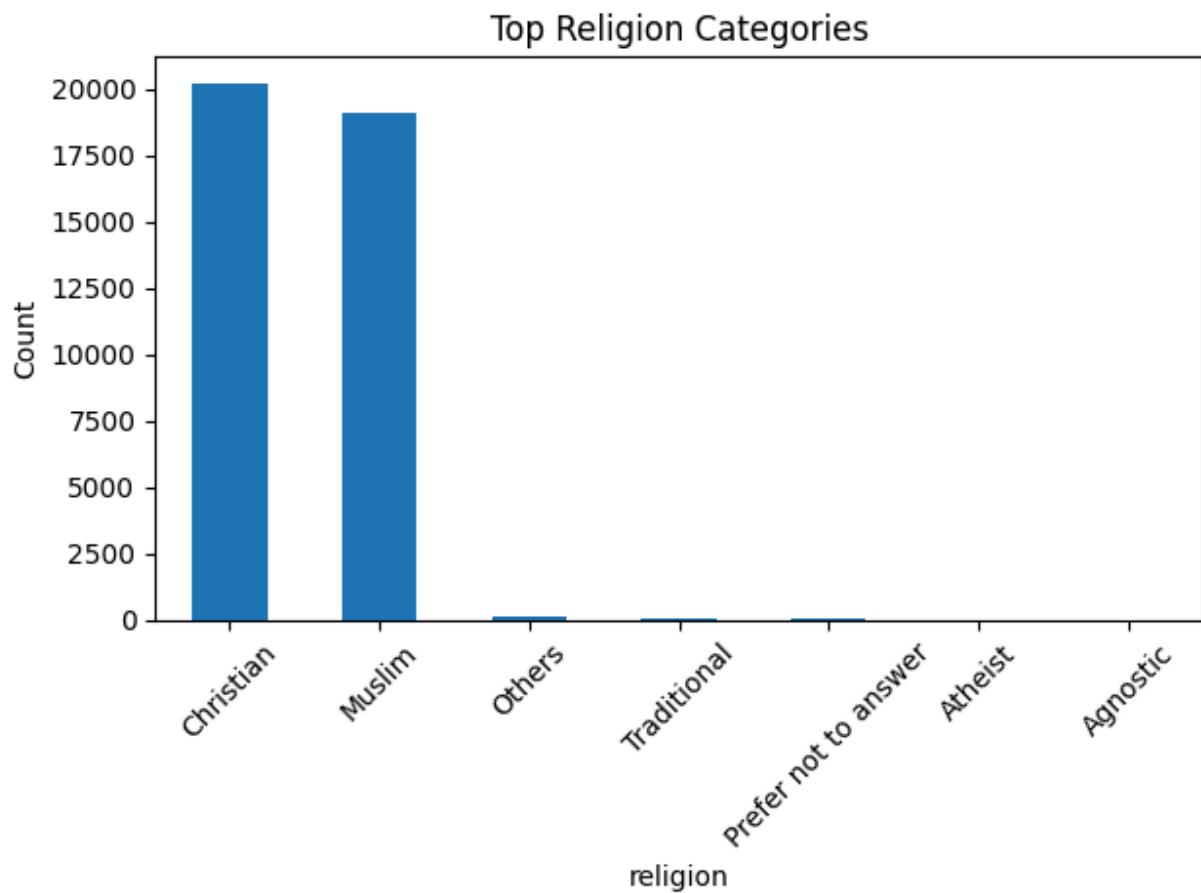


Figure 20 | Distribution of Top Tribe Categories in the Dataset. Bar chart showing the frequency of the top 10 most represented tribes in the study population. The Hausa (Kano) group is the most prevalent, followed by Hausa (Kaduna), Unknown, and Izon (Ijaw) [Bayelsa]. This distribution highlights the demographic composition and potential cultural diversity relevant to downstream socio genomic analyses.

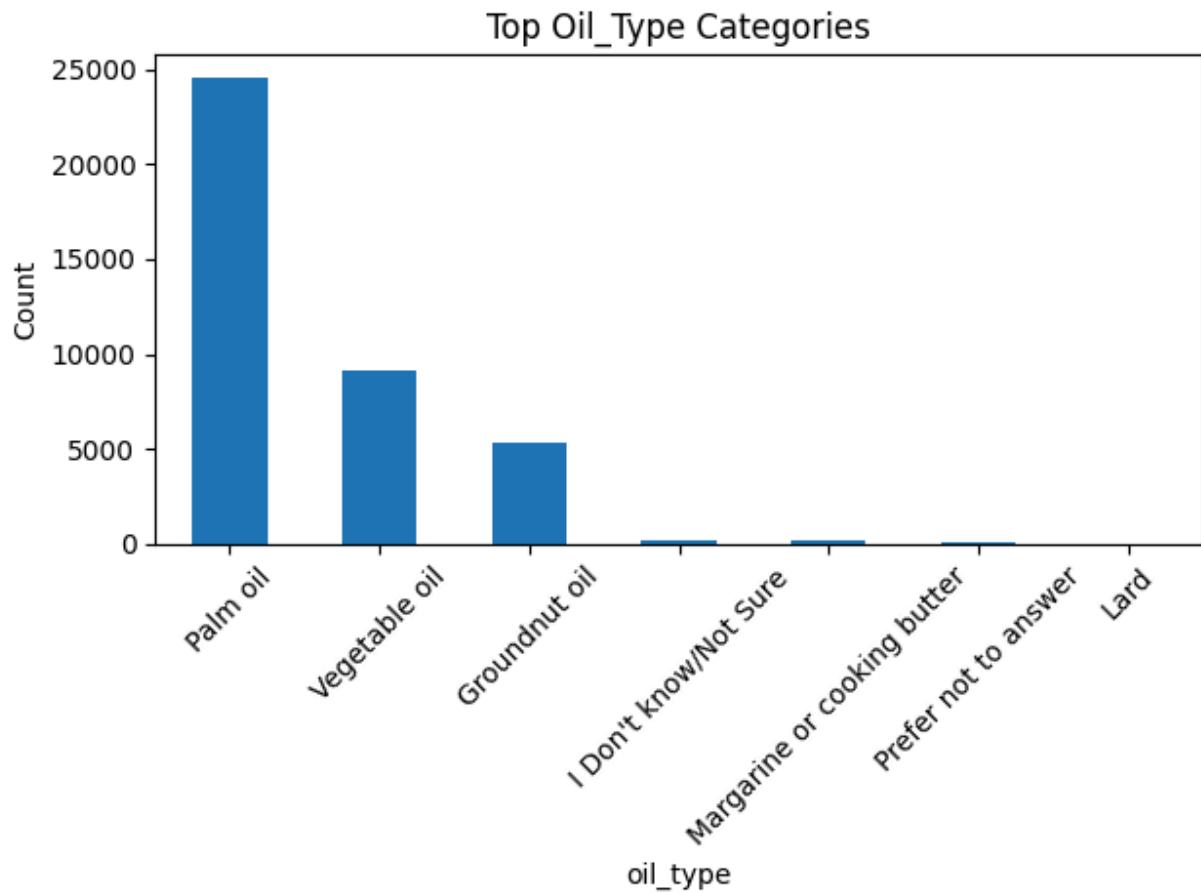
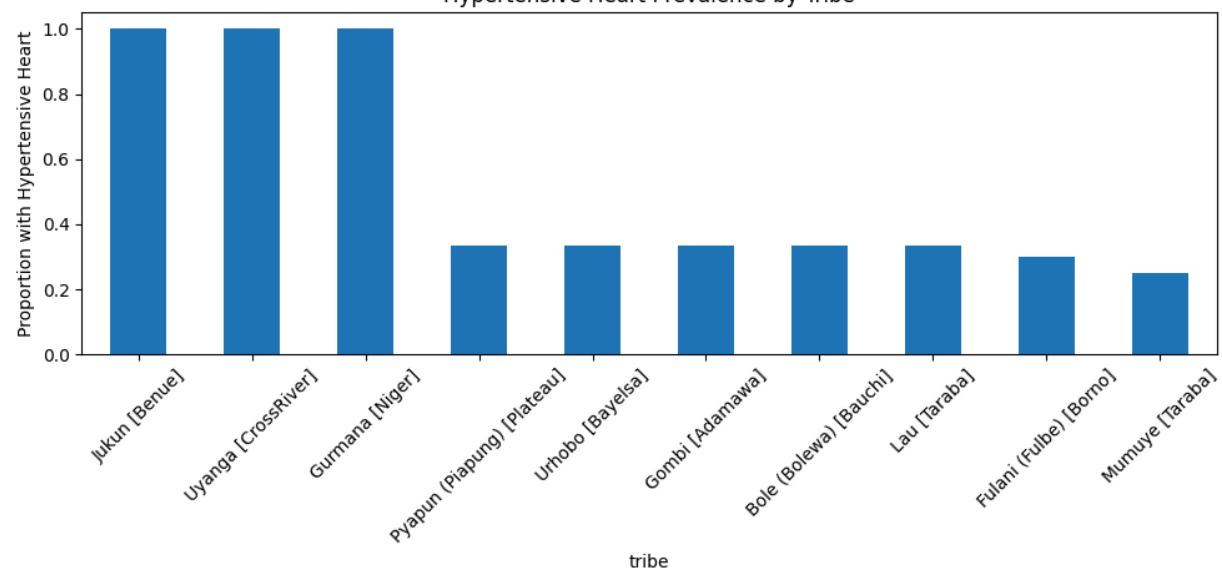
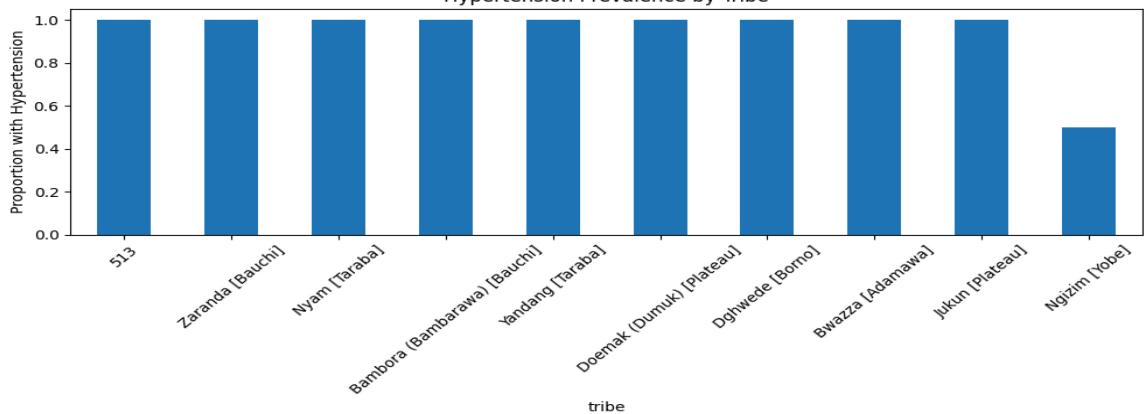


Figure 21 | Distribution of Cooking Oil Type Preferences in the Study Population. Bar plot illustrating the frequencies of self-reported primary cooking oil types used by participants. Palm oil is the most commonly reported oil type, followed by vegetable oil and groundnut oil. Other categories, including margarine, lard, or respondents unsure or unwilling to answer, make up a minimal proportion of the sample.

Hypertensive Heart Prevalence by Tribe



Hypertension Prevalence by Tribe



Type 2 Diabetes Prevalence by Tribe

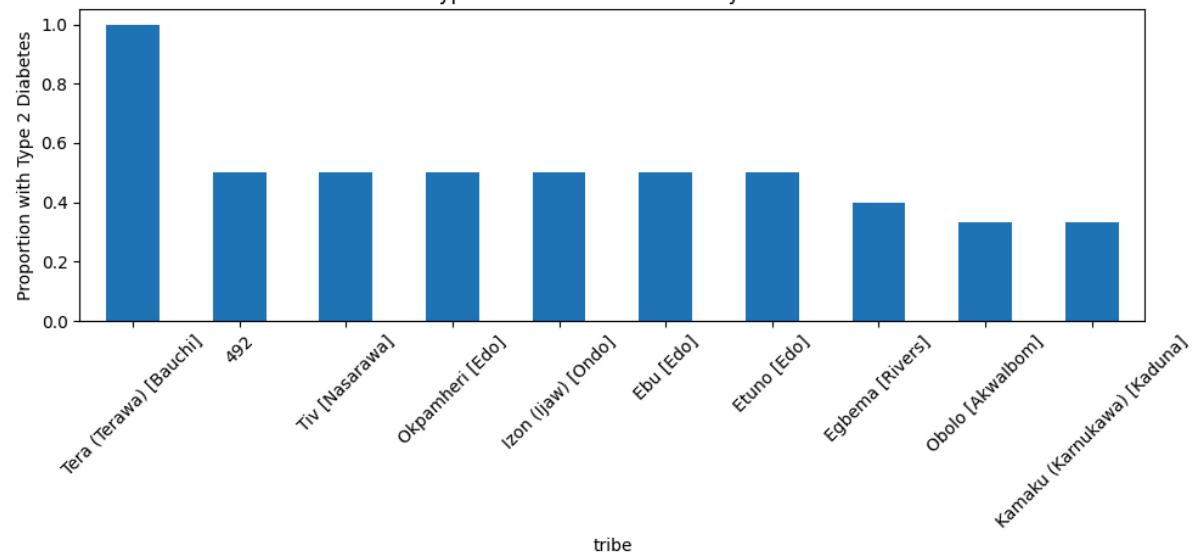
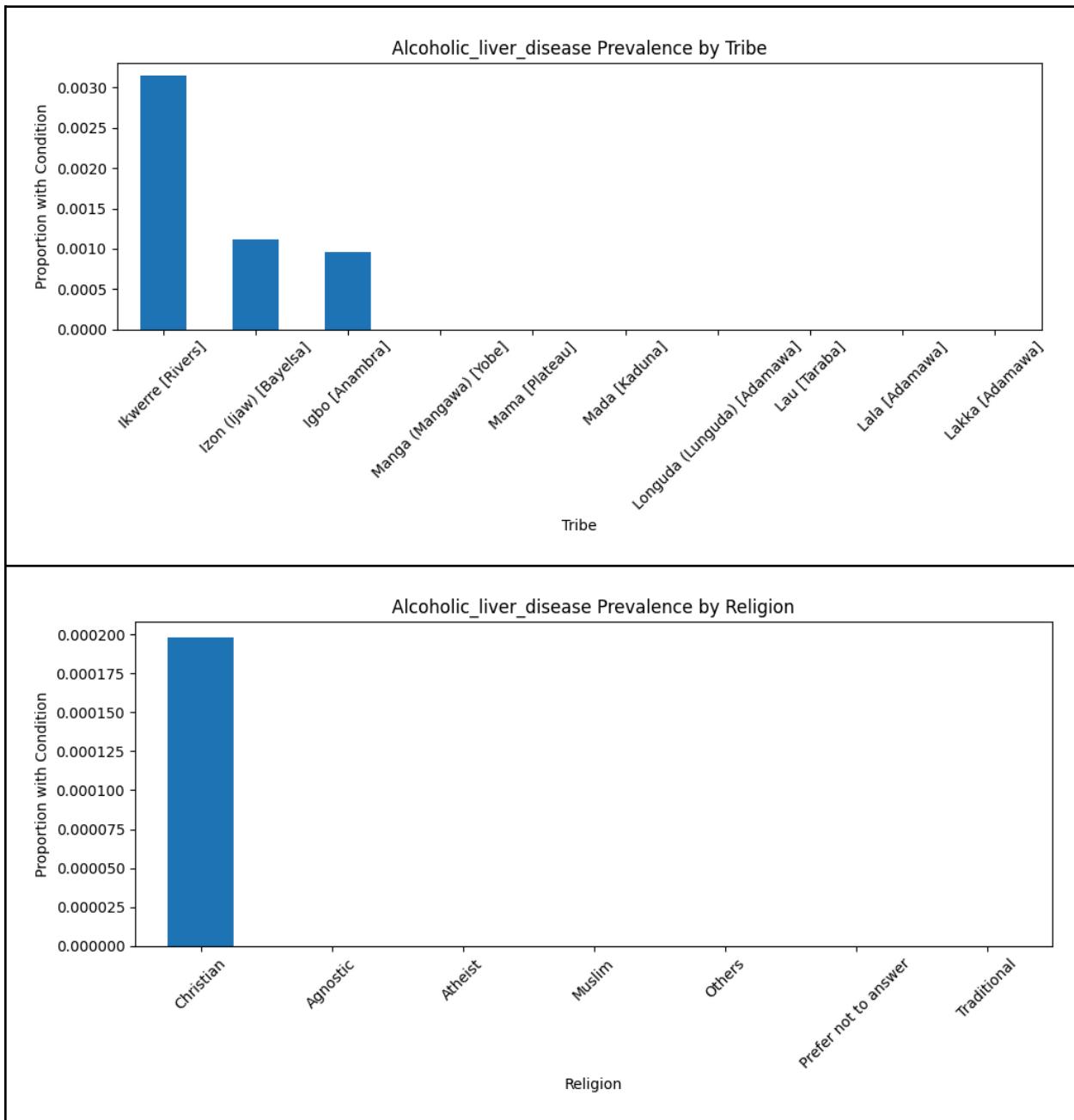


Figure 22 | Disease Prevalence across Nigerian Tribes. (Top) Proportion of individuals with hypertensive heart disease by tribe, with near-complete prevalence observed in certain smaller ethnic groups such as Jukun (Benu) and Gurmana (Niger). (Middle) Proportion of individuals with general hypertension by tribe, showing similarly high prevalence in specific tribes such as Zaranda (Bauchi) and Swazza (Adamawa). (Bottom) Proportion of individuals with Type 2 Diabetes by tribe, with highest recorded prevalence among Tera (Bauchi), Tiv (Nasarawa), and Okpamheri (Edo).



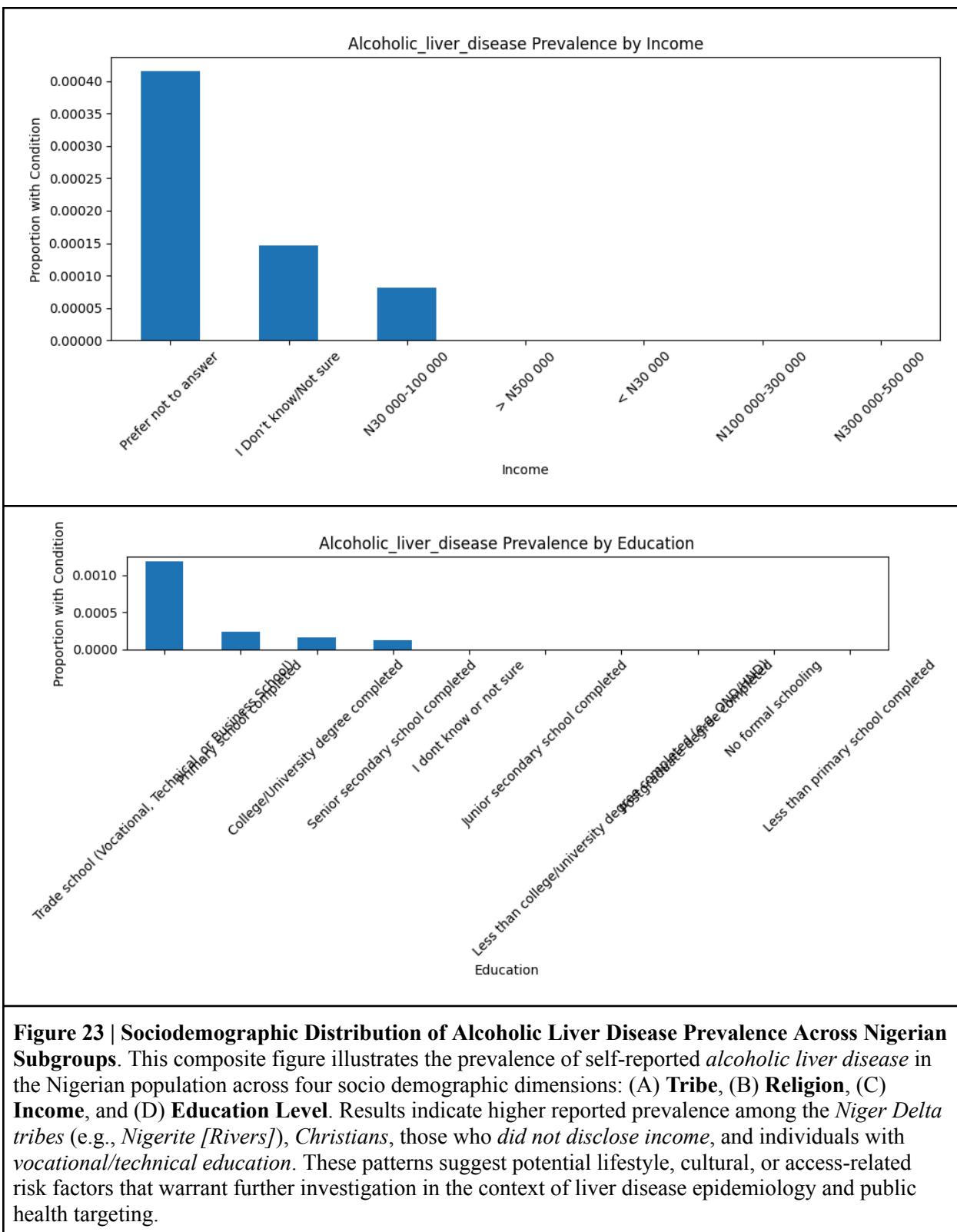


Figure 23 | Sociodemographic Distribution of Alcoholic Liver Disease Prevalence Across Nigerian Subgroups. This composite figure illustrates the prevalence of self-reported *alcoholic liver disease* in the Nigerian population across four socio demographic dimensions: (A) **Tribe**, (B) **Religion**, (C) **Income**, and (D) **Education Level**. Results indicate higher reported prevalence among the *Niger Delta tribes* (e.g., *Nigerite [Rivers]*), *Christians*, those who *did not disclose income*, and individuals with *vocational/technical education*. These patterns suggest potential lifestyle, cultural, or access-related risk factors that warrant further investigation in the context of liver disease epidemiology and public health targeting.

