

AN ANALYTICAL STUDY ON PROBABLE FACTORS LEADING TO ORAL CANCER IN ASSAM

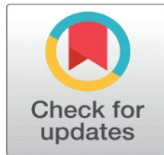
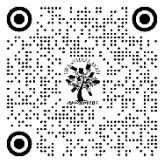
Dr. Parbin Sultana ¹, Jayanta Sarma Kakoty ², Dr. Sahnaj Rahmatulla Laskar ³, Ferdowsi Kawchar ⁴

¹ Professor, School of Technology & Management, University of Science & Technology Meghalaya

² Research Scholar, Department of Mathematics, University of Science & Technology Meghalaya

³ Assistant Professor, Department of Business Administration, University of Science & Technology Meghalaya

⁴ Research Scholar, Department of Business Administration, University of Science & Technology Meghalaya



DOI

[10.29121/shodhkosh.v5.i1.2024.3380](https://doi.org/10.29121/shodhkosh.v5.i1.2024.3380)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2024 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

According to National Centre for Disease Informatics and Research (NCDIR) (2021) [1], oral cancer is among the top three cancers in India and it is among top five cancers in Assam. Different studies revealed that the incidence of oral cancer is more with the use of tobacco related products and the maximum number of cases are in Cachar District. Besides the use of tobacco, consumption of alcohol and metabolic risks are also the factors of the causes of oral cancer. Also some demographic factors are responsible for the causes of oral cancer. In this study, a Factorial Analysis model with ANOVA have been applied to understand the probable factors leading to oral cancer in Assam. In this study, chi-square test has been applied to study the association between Gender and some probable factors like locations, behavioural factors and metabolic risks factors on account of the occurrence of oral cancer. It is concluded that the development of oral cancer among the gender of the patients are not significantly different. It is found that Esophagus and Tongue cancers are significantly difference. The patients with overweight, raised blood pressure and raised blood glucose are significantly difference.

Keywords: Demographic Factors, Factorial Analysis, ANOVA, Chi-Square

1. INTRODUCTION

Cancer is a disease caused by an uncontrolled division of abnormal cells in the parts of the body. Cancer has become more than just a global health problem and is also a cause of deep suffering to the individual patients, their families and to the community at large. The GLOBOCAN 2021 estimates that there were 19.3 million new cases of cancer and almost 10 million deaths from cancer in 2020. According to World Health Organisation (WHO), cancer is found to be second leading cause of death throughout the global [1]

The number of patients registered due to cancer in 2016 in Cachar District, Dibrugarh District and Kamrup Urban given in Table 1.

Table 1: Patients registered due to cancer in 2016

Gender	Cachar District	Dibrugarh District	Kamrup District	Total
	Number of patient registered	Number of patients registered	Number of patients registered	
Males	4663	2535	6223	13421
Females	3943	2238	4790	10971
Total	8606	4773	11013	24392

Source: ICMR – National Center for Disease Informatics and Research (2021)

According to NCDIR (National Centre for Disease Informatics and Research)[2], latest report oral cancer is among the top three cancers in India and it is among the top five cancers in Assam. The rate of growth of oral cancer is more in Kamrup District followed by Cachar and Dibrugarh Districts. According to ICMR report [1], the probable factors of oral cancer are classified into four categories as follows:

Category-I: Sites of oral cancer

The sites of oral cancer in the given region are

- 1) **Hypopharyngeal cancer:** It is a disease in which malignant cells form in the tissues of hypopharynx. Signs and symptoms due to this cancer include a sore throat and ear pain.
- 2) **Mouth cancer:** It occurs in any of the parts like lips, gums, inner lining of the cheeks, roof of the mouth and floor of the mouth.
- 3) **Esophageal cancer:** It occurs in esophagus – a long hollow tube that runs from throat to the stomach.
- 4) **Tongue cancer:** It is a form of cancer that begins in the thin, flat squamous cells that line the surface of the tongue.

Category-II: Demographic factors

In demographic factor, the proportion of developing oral cancer among males and females in the age group (0 – 74) [1] has been studied.

Category-III: Behavioural factors

The maximum number of people in the age group 15 and above consume tobacco and alcohol. So, according to ICMR [2], behavioral factors are classified as – Tobacco and Alcohol consumptions.

Category-IV: Metabolic Risks factors

Metabolic syndrome is not a disease in itself. Instead, it's a group of risks factors- high blood pressure, high blood sugar, unhealthy cholesterol levels and abdominal fats. Occurrence of these factors in a person are influenced by his / her lifestyle. The metabolic factors generally support the growth of any disease. According to ICMR [1] report, it is found that 80% of cancer patients have the following metabolic factors-

- Overweight / Obese BMI > 25
- Raised blood pressure
- Raised blood glucose (random)

The study is trying to analyse the different probable factors leading to Oral Cancer in Assam. In this analysis, a combination of Factorial Analysis and ANOVA, is applied to study the significance of different probable factors leading to Oral Cancer in Assam. Chi-square test is applied to study the association between the gender of the patients and the given probable factors.

2. OBJECTIVES OF THE STUDY

The study reveals the following objectives:

- To analyse the principle sites of oral cancer in Assam.
- To study the demographic factors which influence the occurrence of oral cancer.
- To study the behavioural and metabolic risk factors which influences the occurrences of oral cancer.

3. REVIEW OF LITERATURE

Different literatures on related topics have been reviewed and some of them are present as follows.

- 1) Areia Anibal and Mexia João T. (2016) [3] used ANOVA to study the influence of year, sex, country and location on the average cancer death age. The data used was from World Health Organization (WHO). The location considered were: Kidney, leukaemia, melanoma of skin and oesophagus. By this study it was observed that the average cancer death age is significant with respect to year, sex, contry and location.
- 2) Koc zelina et al (2014) [4] applied one-way ANOVA, t-test and Turkey Honest Significance Difference test to evaluate the impact of nurses healthy lifestyle behaviours on utilization from breast cancer. It was found that Healthy Lifestyle Behaviours Scale total score was higher in nurses with sufficient level of breast cancer knowledge, who perform regular BSE (Breast self examination) and who attended training on breast cancer.
- 3) Prasad Dr S. N. and Diwakar Ajay Kumar (2022) [5] applied ANOVA to anayse the significant difference between the various cancer in India in different years. By the analysis of cancer data from 2019 to 2021, it is interpreted that there is no significant difference between the mean effect of various cancers. By analysing Cancer data, it has been concluded that the various cancer cases which is maximum effected by Cancer is Uttar Pradesh and the prediction of cancer cases follows a linear pattern decreasing year by year.
- 4) Bharathi A., Natarajan Dr. A. M. (2010) [6] computed the importance ranking of each gene using an Analysis of variance (ANOVA). For the purpose of finding the smallest gene subsets for accurate cancer classification, both ANOVA and Cross Validation (CV) are highly effective ranking schemes, where Support Vector Machines (SVM) is sufficiently good classifiers. Finally it is obtained very good accuracy compared to t-test method.
- 5) Rana Rakesh, Singhal Richa (2015) [7] analysed whether there was an association between smoking and lung disease by using chi-square test for independence of attributes and finally it was concluded that at 5% significance level, it is concluded that there is no association between smoking and lung cancer.
- 6) Adimu Patience I et al. (2018) [8] applied chi-square test to study whether there was an association between the topological location of cancer and gender of the patients in North eastern states of Nigeria. At 5% level of significance, it was concluded that there is a significant association between the topological location of cancer and the gender of patients.
- 7) Etxeberria Jaione et al (2015) [9] considered an ANOVA-type P-spline model to predict the number of mortality cases in forth coming years in regions within a country.
- 8) Glenn II Paul D. and Griesinger N. L. Glenn (2021) [10] performed statistical power analyses of ANOVA to investigate the effects of violating assumptions such as normality using mice intestinal polyp data.

4. RESEARCH METHODOLOGY

The data used for the study is secondary data and is based on 2016 registered patients collected from three different regions – Guwahati, Dibrugarh and Cachar recorded in ICMR – National Center for Disease Informatics and Research published in 2021. In this study, the factors are classified into four categories as mentioned earlier.

Analysis of variance (ANOVA) is applied to test the significant of the means proportion of the development of oral cancer

- 1) among three different regions.
- 2) among the sites.

- 3) due to the metabolic risk factors.
- 4) among the gender of the patients with the given metabolic risk factors.

If the test is significant and the number of groups/ samples/ treatments > 2 , then it is required to find out which pair(s) of the factor means differ significantly. For this, Turkey's Honest Significant Difference (HSD) test statistic $T = Q \times \sqrt{\frac{MSE}{n}}$, where, $n_i = n_j = n$ and Q is the Turkey's critical value at $(k, N - k)$ degrees of freedom, MSE is the mean square error. n_i and n_j are the respective sample sizes. In this study, N = Total number of observations (patients) under study and k = Number of factors under study.

If the difference between the two means i.e., $|\bar{x}_i - \bar{x}_j| > T$, there is significant difference between the two means $\bar{x}_i$ and $\bar{x}_j$, otherwise there is no significant difference between the two means.

Chi-square (χ^2) test for the independence of attributes is applied to test

- 1) the association between the gender of the patients and locations (regions) on account of the development of oral cancer.
- 2) the association between the gender of the patients and behavioural factors on account of the development of oral cancer.
- 3) the association between the gender of the patients and metabolic risk factors on account of the development of oral cancer.

5. ANALYSIS

The analyses of all the categories are employed which are described as follows.

5.1. ANALYSIS OF CATEGORY-I:

The proportions of developing oral cancer in the age group (0 – 74) in the three given regions are computed in the table 2.

Table 2: Proportions of developing oral cancer

Sites of oral cancer	Cachar District			Dibrugarh District			Kamrup District			Total
	Male	Female	Total	Male	Female	Total	Male	Female	Total	
Hypopharynx	0.0156	0.0000	0.0156	0.0137	0.0000	0.0137	0.0283	0.0000	0.0283	0.0576
Mouth	0.0114	0.0069	0.0183	0.0081	0.0000	0.0081	0.0152	0.0000	0.0152	0.0416
Esophagus	0.0189	0.0097	0.0286	0.0185	0.0094	0.0279	0.0371	0.0227	0.0598	0.1163
Tongue	0.0119	0.0000	0.0119	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0119
Total	0.0578	0.0166	0.0744	0.0403	0.0094	0.0497	0.0806	0.0227	0.1033	0.2274

Source: ICMR – National Center for Disease Informatics and Research (2021)

Null hypothesis H_{01} : There is no significant difference of the means proportion of the development of oral cancer among three different regions, i.e.,

$$\mu_{\text{Cachar}} = \mu_{\text{Dibrugarh}} = \mu_{\text{Kamrup}}$$

Alternative hypothesis H_{11} : There is significant difference of the means proportion of the development of oral cancer among three different regions. In other words, at least two regions are significantly different.

Null hypothesis H_{02} : There is no significant difference of the means proportion of the development among the sites of oral cancer. That is,

$$\mu_{\text{Hypopharynx}} = \mu_{\text{Mouth}} = \mu_{\text{Esophagus}} = \mu_{\text{Tongue}}$$

Alternative hypothesis H_{12} : There is significant difference of the means proportion of the development among the sites of oral cancer. In other words, at least two means are significantly different.

To test the hypothesis, ANOVA is applied that are shown in Table 3 and Table 4.

Table 3: ANOVA: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance		
Cachar District	4	0.0744	0.0186	5.13E-05		
Dibrugarh District	4	0.0497	0.012425	0.000138		
Kamrup District	4	0.1033	0.025825	0.000647		
Hypopharynx	3	0.0576	0.0192	6.3E-05		
Mouth	3	0.0416	0.013867	2.73E-05		
Esophagus	3	0.1163	0.038767	0.000332		
Tongue	3	0.0119	0.003967	4.72E-05		

Table 4: ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Centres	0.00036	2	0.00018	1.864196	0.234601	5.143253
Oral sites	0.001929	3	0.000643	6.663075	0.024467	4.757063
Error	0.000579	6	9.65E-05			
Total	0.002868	11				

Interpretations:

Test statistic value of the region variations $F = 1.864196 < F_{\text{crit}} = 5.143253$ or $P\text{-value} = 0.234601 > \alpha = 0.05$.

Therefore, the null hypothesis H_{01} is not rejected at $\alpha = 0.05$.

Hence, it is concluded that There is no significant difference of the means proportion of the development of oral cancer among three different regions.

Again, test statistic value of oral sites $F = 6.663075 > F_{\text{crit}} = 4.757063$ or $P\text{-value} = 0.024467 < \alpha = 0.05$.

Therefore, we may reject the null hypothesis H_{02} at $\alpha = 0.05$

Hence, it is concluded that there is significant difference of the means proportion of the development among the sites of the oral cancer. In other words, at least two means are significantly different.

To test significant difference among the means proportion of the development of a pair of sites of oral cancer, Turkey's HSD test is applied.

Turkey's HSD Test:

Degrees of freedom for Turkey's critical value Q is $(4, 12 - 4) = (4, 8)$

Therefore, $Q_{0.05} = 4.53$

$$\text{Turkey's HSD test statistic } T = 4.53 \times \sqrt{\frac{0.0000965}{3}} = 0.0257$$

The results are interpreted by the table 5.

Table 5: Comparison Analysis by Turkey's HSD test

Pairs	$ \bar{x}_i - \bar{x}_j $	Comparison	Conclusion
Hypopharynx and Mouth	0.005333	< T	Not significant
Hypopharynx and Esophagus	0.019567	< T	Not significant
Hypopharynx and Tongue	0.015233	< T	Not significant
Mouth and Esophagus	0.0248903	< T	Not significant
Mouth and Tongue	0.0099	< T	Not significant
Esophagus and Tongue	0.0348	> T	Significant

Interpretations:

The pairs of sites of the oral cancer - Hypopharynx and Mouth, Hypopharynx and Esophagus, Hypopharynx and Tongue, Mouth and Esophagus, and Mouth and Tongue are not significant as their respective mean differences are < T. But the Esophagus and Tongue are significantly different as their respective mean difference > T.

5.2. ANALYSIS OF CATEGORY-II

The proportions of developing oral cancer among the gender of the patients in the age group (0 – 74) in the three given regions are computed in the table 6.

Table 6: Proportions of developing oral cancer among the gender of the patients.

	Cachar District	Dibrugarh District	Kamrup District
Males	0.0578	0.0403	0.0806
Females	0.0166	0.0094	0.0227

Source: ICMR – National Center for Disease Informatics and Research (2021)

Null hypothesis H_{03} : There is no association between the gender of the patients and locations (regions) on account of the development of oral cancer.

H_{13} : There is association between gender of the patients and locations (regions) on account of the development of oral cancer.

To test the hypothesis, data are analysed by chi-square test shown in table 7.

Table 7: Chi-square test for the association between the gender of the patients and the locations (regions)

Location→		Cachar District	Dibrugarh District	Kamrup Urban	Total	(O - E) ² E	Degrees of freedom	χ ² crit	P-value		α
Gender↓											
Male	Observed	0.0578	0.0403	0.0806	00.1787	5.131E-05	2	5.9915	0.99988		0.05
	Expected	0.058466	0.0390562	0.081177							
Female	Observed	0.0166	0.0094	0.0227	0.0487	0.00018828					
	Expected	0.015934	0.0106438	0.022123							
					χ ² =	0.00023959					

Interpretations:

The test statistic value $\chi^2 = 0.00023959 < \chi^2_{crit} = 5.9915$ or P-value = $0.99988 > \alpha = 0.05$. So, the null hypothesis H_{03} is not rejected. Hence, it is concluded that there is no association between the gender of the patients and the locations (regions) on account of the development of oral cancer.

5.3. ANALYSIS OF CATEGORY-III

The proportion of male and female of age 15 and above who used tobacco and consumed alcohol are shown in the table 8.

Table 8: The proportion of male and female who used tobacco and consumed alcohol.

		Behavioural factors	
		Tobacco	Alcohol
Gender	Male	0.629	0.251
	Female	0.329	0.073

Source: ICMR – National Center for Disease Informatics and Research (2021)

Null hypothesis H_{04} : There is no association between the gender of the patients and the behavioural factors on account of the development of oral cancer.

Alternative hypothesis H_{14} : There is association between the gender of the patients and behavioural factors on account of the development of oral cancer.

To test the hypothesis, data are analysed by chi-square test that shown in table 9.

Table 9: Chi-square test for independence of the gender of the patients and behavioural factors.

Behavioural factors→ Risks		Tobacco	Alcohol	Total	$\frac{(O - E)^2}{E}$	Degrees of freedom	χ^2 crit	P-value	α
Genders↓									
Male	Observed frequency	0.629	0.251	0.88	0.004921	1	3.841	0.9003	0.05
	Expected frequency	0.657598	0.222402						
Female	Observed frequency	0.329	0.073	0.402	0.010772				
	Expected frequency	0.300402	0.101598						
				$\chi^2 =$	0.015693				

Interpretations:

$\chi^2 = 0.015693 < \chi^2_{crit} = 3.841$ or p-value = $0.9003 > \alpha = 0.05$. So, the null hypothesis H_{04} is not rejected. Hence, it is concluded that there is no association between the gender of the patients and behavioural risk factors on account of the development of oral cancer.

5.4. ANALYSIS OF CATEGORY-IV

The proportion of male and female of age 15 and above with the given metabolic risk factors are shown in the table 10.

Table 10: The proportion of male and female with given metabolic risks factors

	Male	Female
Overweight/Obese BMI >25 (%)	0.162	0.152
Raised blood pressure (%)	0.203	0.191
Raised blood glucose (random) (%)	0.160	0.128

Source: ICMR – National Center for Disease Informatics and Research (2021)

Null hypothesis H_{05} : There is no association between the gender of the patients and the metabolic risk factors on account of the development of oral cancer.

Alternative hypothesis H_{15} : There is association between the gender of the patients and metabolic risk factors on account of the development of oral cancer.

To test the hypothesis, data are analysed by Chi-square test that shown in table 11

Table 11: Chi-square test for the association between the gender of the patients and the metabolic risk factors

Gender→		Male	Female	Total	(O - E) ² / E	Degrees of freedom	χ ² crit	p-value	α
Metabolic Risk factors↓									
Overweight/Obese BMI >25 (%)	Observed frequency	0.162	0.152	0.314	0.001098	2	5.991465	0.99867	0.05
	Expected frequency	0.171245	0.142755						
Raised blood pressure (%)	Observed frequency	0.203	0.191	0.394	0.001443				
	Expected frequency	0.214875	0.179125						
Raised blood glucose (random) (%)	Observed frequency	0.2	0.128	0.288	0.000121				
	Expected frequency	0.157066	0.130934						
				χ ² =	0.002662				

Interpretations:

$\chi^2 = 0.002662 < \chi^2_{crit} = 5.991465$ or $p\text{-value} = 0.99867 > \alpha = 0.05$. So, the null hypothesis H_{05} is not rejected. Hence, it is concluded that there is no association between the gender of the patients and metabolic risk factors on account of the development of oral cancer.

Null hypothesis H_{06} : There is no significant difference between the means proportion of the development of oral cancer due to the given metabolic risk factors. That is,

$$\mu_{\text{Overweight}} = \mu_{\text{Raised blood pressure}} = \mu_{\text{Raised blood glucose}}$$

Alternative hypothesis H_{16} : There is significant difference between the means proportion of the development of oral cancer due to the given metabolic risk factors. That is, at least two means are significantly different.

Null hypothesis H_{07} : There is no significant difference of the proportion of the development of oral cancer among the gender of the patients with given metabolic risk factors.

That is, $\mu_{\text{Male}} = \mu_{\text{Female}}$

Alternative hypothesis H_{17} : There is significant difference of the proportion of the development of oral cancer among the gender of the patients with given metabolic risk factors.

That is, $\mu_{\text{Male}} \neq \mu_{\text{Female}}$

To test the hypothesis, ANOVA is applied that shown in the table 12 and table 13.

Table 12: ANOVA: Two-factor analysis

SUMMARY	Count	Sum	Average	Variance		
Overweight/Obese BMI >25 (%)	2	0.314	0.157	5E-05		
Raised blood pressure (%)	2	0.394	0.197	7.2E-05		
Raised blood glucose (random) (%)	2	0.288	0.144	0.000512		
Male	3	0.525	0.175	0.000589		
Female	3	0.471	0.157	0.001011		
Table 13: ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Metabolic risk factors	0.003052	2	0.001526	20.6216	0.04625	19
Gender of the patients	0.000486	1	0.000486	6.567568	0.124465	18.51282
Error	0.000148	2	7.4E-05			
Total	0.003686	5				

Interpretations:

Test statistic value for metabolic risks variation $F = 20.62162 > F \text{ crit} = 19$ or $P\text{-value} = 0.04625 < \alpha = 0.05$.

Therefore, the null hypothesis H_{06} is rejected at $\alpha = 0.05$.

Hence, it is concluded that there is significant difference between the means proportion of the development of oral cancer due to the given metabolic risk factors.

Again, test statistic value for gender variation $F = 6.567568 < F \text{ crit} = 18.51282$.

Therefore, the null hypothesis H_{07} is not rejected at $\alpha = 0.05$.

Hence, we conclude that There is no significant difference of the proportion of the development of oral cancer among the gender of the patients with given metabolic risk factors.

Turkey's HSD Test:

Degrees of freedom for Turkey's critical value Q is $(3, 6 - 3) = (3, 3)$

Therefore, $Q_{0.05} = 5.91$

Turkey's HSD test statistic $T = 5.91 \times \sqrt{\frac{7.4 \times 10^{-5}}{2}} = 0.03595$

The results are interpreted in table 14.

Table 14: Comparison Analysis by Turkey's HSD test

Pairs	$ \bar{x}_i - \bar{x}_j $	Comparison	Conclusion
Overweight and Raised blood pressure	0.04	$> T$	Significant
Overweight and Raised blood glucose.	0.013	$< T$	Not significant
Raised blood pressure and Raised blood glucose	0.053	$> T$	Significant

Interpretations:

The pairs of factors Overweight & Raised Blood pressure; Raised blood pressure and Raised blood glucose are significant as their mean differences $> T$. However, the pair of factors Overweight and Raised blood glucose are not significantly different as their mean difference $< T$.

6. CONCLUSIONS AND DISCUSSIONS

According to the study, it may be concluded that the mean proportions of the development of oral cancer among Cachar district, Dibrugarh district and Kamrup District are not significantly different, whereas the mean proportions of the development of oral cancer among the sites are significantly different as the mean proportions of the development of Oesophagus and Tongue cancers are significant. From this study, it is found that there is no association of the gender of the patients with the locations(regions); behavioural factors and the metabolic risk factors.. Also, development of oral cancer among the gender of the patients due to overweight and raised blood pressure; raised blood pressure and raised glucose are significantly different. However, the development of oral cancer among the gender of the patients due to overweight and raised blood glucose is not significant.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Hyuna Sung, Jacques Ferlay "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries". <https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21660>
- "Assam-National Centre for Disease Informatics and Research"(2021) https://ncdirindia.org/All_Reports/NorthEast2021/resources/NE_chapter3.pdf
- Areia Anibal and Mexia João T. (2016), "ANOVA like analysis of cancer death age" <https://doi.org/10.1063/1.4951825>.
- Koc Zelina et. Al (2014). "Evaluation of Impact of Nurses Healthy Lifestyle Behaviours on Utilization from Breast Cancer Early Diagnosis Methods". <https://ncbi.nlm.nih.gov/pmc/articles/pmc5351542>

- Prasad Dr. S. N., Diwakar Ajay Kumar (2022). "Statistical Analysis Cancer Data". <https://ijsr.net/archive/V11i8/SR22819190317.pdf>
- Bharathi A., Natarajan Dr. A. M. (2010). "Cancer Classification of Bioinformatics data using ANOVA". <https://ijcte.org/papers/169-G652.pdf>.
- Rana Rakesh, Singhal Richa (2015). "Chi-square Test and its Application in Hypothesis Testing". https://j-pcs.org/temp/JPractCardiovascSci1169-5772018_160200.pdf
- Adimu Patience I. et al (2018). "Statistical data analysis of cancer incidences in insurgency affected states in Nigeria". <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5998707/>
- Etxeberria Jaione et al (2015). "ON PREDICTING CANCER MORTALITY USING ANOVA-TYPEP-SPLINE MODELS". <https://ine.pt/revstat/pdf/rs150102.pdf>
- Paul D. Glenn II and Griesinger N. L. Glenn (2021). "STATISTICAL POWER ANALYSIS OF ANOVA WITH APPLICATION TO COUNT DATA FROM INTESTINAL POLYP MEDICINE IN APE MIN + MICE". <https://coset.tsu.edu/statistical-power-analysis-of-anova-with-application-to-count-data-from-intestinal-polyp-incidence-in-apemin-mice-paul-d-glenn-it-and-n-l-glenn-griesinger/>