

Visual analysis of leading Cancer sites using SPSS Software

V. Manikanta^{1*}, N.G Yethiraj²

¹ SRP-Technical (CIIL), Bharathiar University, Coimbatore, Tamilnadu, India

² Department of Computer Science, Maharani's Science College for Women, Bangalore, India

*Corresponding Author: ciil.manikanta@gmail.com

DOI: <https://doi.org/10.26438/ijcse/v7si9.3846> | Available online at: www.ijcseonline.org

Abstract— Many diseases are arising depending upon the environment, culture, life style, etc. In many Diseases, Multiple types of cancer are increased in various region based on mouth, breast, cervix uteri, ovary, gall bladder, NHL, thyroid, Brain, Stomach, etc. Human of each organ is depending of the cell functionality. Each cell functionality is works our daily foods of sweet, sour, salty, etc. In our country many medical research centers are involved many research areas of Tuberculosis, Cancer, Diabetes, etc. Leading cancer sites data is one of the big tasks to region wise, gender wise, age wise, etc. The study of leading sites of cancer in various region wise, gender wise, age wise are easy to find the avoid the cancer to protect with any immediate solution. Visual analysis of data is easily exploring the details of Cancer leading site and analysis describes to how it is important to save our country for immediate treatment procedure.

Keywords— Visualization, Cancer, data mining, SPSS.

I. INTRODUCTION

Variation of cells are damage the immune system, tumours, impairment results and uncontrollable condition to growth arises the Cancer Disease. Cancer disease are visible state is easy to find some solution or not. But, some of the cancer diseases are not visible and identifying the cancer is also takes time. Many of the cells spread easily with more human parts like nodes of Lymph. Cancer diseases are preventable when some of the new technology with expertise doctor and spending money like Cricketer Yuvaraj Singh. Scientists and Doctors are trying to research the cancer disease for solution to resolve the diseases. Unpreventable Cancer diseases are dividing the cell to irritating control of human life. Human is focus on health issues rather than any other works and always concentration 'my health'. Cancer is one disease but due to mental preparation also another disease. Each and every cancer sites are increasing with alcohol, cigarette, quality of foods, smoking, physical exercise, etc. More and more people are died with cancer throughout the world multiple types cancer. Cells are not working every time same and varies with age. Cancer are occurs due to the age factor because cells are weak and activities of physical is less. Physical activity is very important in every aspect of diseases except some of the disease. Genetic of cells are contribute to developing cell growth with uncontrollable and cells are dividing with genes alteration with proteins production. Cells are developing and dividing due to carrying the instruction of protein. Cancer cells are generated by birth, smoking, chemical dependant, mutation process, etc.

II. SPSS

SPSS is to find knowledge enhancement functionality to produce efficient and effective manner of data with visualization. Data of huge size content is to perform easiest way of understandable solution in several steps using SPSS. SPSS is available many tools such as graph builder, Analysis data in various techniques, Define variable structure, Transformation, Valid output results, etc. SPSS is not expressing just like software and it is huge application structure oriented analysis data with complex problem to solve easily compared to any other software. Innovation of data with many aspects of medical science, business activity and industrial management field is very important tool is required for further instruction to manipulate data. In this case, SPSS is important role of software. Below tables are sample examples of dataset.

Table 1. Type of Cancer in different region from Female group

Cancer type	Place	Count
Breast	Mumbai	5620
Cervix Uteri	Mumbai	2480
Ovary	Mumbai	936
Mouth	Mumbai	815
Gall Bladder	Mumbai	723
NHL	Mumbai	647

Thyroid	Mumbai	607	Brain, NS	Chennai	103
Lung	Mumbai	524	Hypopharynx	Chennai	406
Oesophagus	Mumbai	510	Rectum	Chennai	302
Tongue	Mumbai	452	Corpus Uteri	Chennai	342
Myeloid Leukaemia	Mumbai	405	Breast	Thiruvanthapuram	5354
Stomach	Mumbai	333	Cervix Uteri	Thiruvanthapuram	1743
Brain, NS	Mumbai	253	Ovary	Thiruvanthapuram	950
Hypopharynx	Mumbai	192	Mouth	Thiruvanthapuram	889
Rectum	Mumbai	311	Gall Bladder	Thiruvanthapuram	76
Corpus Uteri	Mumbai	343	NHL	Thiruvanthapuram	516
Breast	Bangalore	2052	Thyroid	Thiruvanthapuram	2095
Cervix Uteri	Bangalore	3585	Lung	Thiruvanthapuram	521
Ovary	Bangalore	778	Oesophagus	Thiruvanthapuram	293
Mouth	Bangalore	1354	Tongue	Thiruvanthapuram	548
Gall Bladder	Bangalore	67	Myeloid Leukaemia	Thiruvanthapuram	721
NHL	Bangalore	254	Stomach	Thiruvanthapuram	287
Thyroid	Bangalore	449	Brain, NS	Thiruvanthapuram	391
Lung	Bangalore	224	Hypopharynx	Thiruvanthapuram	123
Oesophagus	Bangalore	666	Rectum	Thiruvanthapuram	393
Tongue	Bangalore	163	Corpus Uteri	Thiruvanthapuram	624
Myeloid Leukaemia	Bangalore	337	Breast	Dibrugarh	336
Stomach	Bangalore	280	Cervix Uteri	Dibrugarh	266
Brain, NS	Bangalore	254	Ovary	Dibrugarh	200
Hypopharynx	Bangalore	169	Mouth	Dibrugarh	116
Rectum	Bangalore	192	Gall Bladder	Dibrugarh	175
Corpus Uteri	Bangalore	215	NHL	Dibrugarh	33
Breast	Chennai	3921	Thyroid	Dibrugarh	23
Cervix Uteri	Chennai	4462	Lung	Dibrugarh	36
Ovary	Chennai	921	Oesophagus	Dibrugarh	237
Mouth	Chennai	914	Tongue	Dibrugarh	51
Gall Bladder	Chennai	114	Myeloid Leukaemia	Dibrugarh	18
NHL	Chennai	301	Stomach	Dibrugarh	133
Thyroid	Chennai	467	Brain, NS	Dibrugarh	27
Lung	Chennai	409	Hypopharynx	Dibrugarh	75
Oesophagus	Chennai	559	Rectum	Dibrugarh	60
Tongue	Chennai	305	Corpus Uteri	Dibrugarh	32
Myeloid Leukaemia	Chennai	478	Breast	Guwahati	674
Stomach	Chennai	646	Cervix Uteri	Guwahati	764

Ovary	Guwahati	228
Mouth	Guwahati	251
Gall Bladder	Guwahati	544
NHL	Guwahati	60
Thyroid	Guwahati	60
Lung	Guwahati	140
Oesophagus	Guwahati	487
Tongue	Guwahati	135
Myeloid Leukaemia	Guwahati	25
Stomach	Guwahati	201
Brain, NS	Guwahati	46
Hypopharynx	Guwahati	127
Rectum	Guwahati	86
Corpus Uteri	Guwahati	71
Breast	Chandigarh	341
Cervix Uteri	Chandigarh	385
Ovary	Chandigarh	166
Mouth	Chandigarh	18
Gall Bladder	Chandigarh	138
NHL	Chandigarh	45
Thyroid	Chandigarh	14
Lung	Chandigarh	47
Oesophagus	Chandigarh	106
Tongue	Chandigarh	38
Myeloid Leukaemia	Chandigarh	55
Stomach	Chandigarh	18
Brain, NS	Chandigarh	94
Hypopharynx	Chandigarh	11
Rectum	Chandigarh	23
Corpus Uteri	Chandigarh	80
Myeloid Leukaemia	Mumbai	880
Hypopharynx	Mumbai	869
Larynx	Mumbai	841
Stomach	Mumbai	783
Lymphoid Leukaemia	Mumbai	662
Brain, NS	Mumbai	540
Liver	Mumbai	318
Thyroid	Mumbai	329
Tonsil	Mumbai	251
Rectum	Mumbai	605
Gall Bladder	Mumbai	530
Oth. Oropharynx	Mumbai	271
Mouth	Bangalore	731
Lung	Bangalore	790
Tongue	Bangalore	754
NHL	Bangalore	520
Oesophagus	Bangalore	887
Myeloid Leukaemia	Bangalore	442
Hypopharynx	Bangalore	946
Larynx	Bangalore	447
Stomach	Bangalore	620
Lymphoid Leukaemia	Bangalore	390
Brain, NS	Bangalore	398
Liver	Bangalore	281
Thyroid	Bangalore	172
Tonsil	Bangalore	228
Rectum	Bangalore	265
Gall Bladder	Bangalore	63
Oth. Oropharynx	Bangalore	213
Mouth	Chennai	1351
Lung	Chennai	1404
Tongue	Chennai	1061
NHL	Chennai	655
Oesophagus	Chennai	892
Myeloid Leukaemia	Chennai	715
Hypopharynx	Chennai	765
Larynx	Chennai	599
Stomach	Chennai	1446

Table 2. Type of Cancer in different region from Male group

Cancer type	Place	Count
Mouth	Mumbai	2776
Lung	Mumbai	1808
Tongue	Mumbai	1603
NHL	Mumbai	1381
Oesophagus	Mumbai	1045

Lymphoid Leukaemia	Chennai	450	Tonsil	Dibrugarh	98
Brain, NS	Chennai	142	Rectum	Dibrugarh	80
Liver	Chennai	479	Gall Bladder	Dibrugarh	62
Thyroid	Chennai	180	Oth. Oropharynx	Dibrugarh	54
Tonsil	Chennai	210	Mouth	Guwahati	545
Rectum	Chennai	462	Lung	Guwahati	503
Gall Bladder	Chennai	155	Tongue	Guwahati	472
Oth. Oropharynx	Chennai	273	NHL	Guwahati	127
Mouth	Thiruvanthapuram	1593	Oesophagus	Guwahati	1020
Lung	Thiruvanthapuram	2847	Myeloid Leukaemia	Guwahati	36
Tongue	Thiruvanthapuram	1214	Hypopharynx	Guwahati	807
NHL	Thiruvanthapuram	954	Larynx	Guwahati	382
Oesophagus	Thiruvanthapuram	798	Stomach	Guwahati	439
Myeloid Leukaemia	Thiruvanthapuram	948	Lymphoid Leukaemia	Guwahati	26
Hypopharynx	Thiruvanthapuram	490	Brain, NS	Guwahati	106
Larynx	Thiruvanthapuram	766	Liver	Guwahati	85
Stomach	Thiruvanthapuram	914	Thyroid	Guwahati	36
Lymphoid Leukaemia	Thiruvanthapuram	739	Tonsil	Guwahati	334
Brain, NS	Thiruvanthapuram	573	Rectum	Guwahati	141
Liver	Thiruvanthapuram	452	Gall Bladder	Guwahati	258
Thyroid	Thiruvanthapuram	635	Oth. Oropharynx	Guwahati	205
Tonsil	Thiruvanthapuram	130	Mouth	Chandigarh	94
Rectum	Thiruvanthapuram	595	Lung	Chandigarh	243
Gall Bladder	Thiruvanthapuram	95	Tongue	Chandigarh	192
Oth. Oropharynx	Thiruvanthapuram	451	NHL	Chandigarh	114
Mouth	Dibrugarh	257	Oesophagus	Chandigarh	164
Lung	Dibrugarh	103	Myeloid Leukaemia	Chandigarh	113
Tongue	Dibrugarh	164	Hypopharynx	Chandigarh	52
NHL	Dibrugarh	83	Larynx	Chandigarh	140
Oesophagus	Dibrugarh	380	Stomach	Chandigarh	61
Myeloid Leukaemia	Dibrugarh	47	Lymphoid Leukaemia	Chandigarh	155
Hypopharynx	Dibrugarh	441	Brain, NS	Chandigarh	200
Larynx	Dibrugarh	106	Liver	Chandigarh	69
Stomach	Dibrugarh	197	Thyroid	Chandigarh	2
Lymphoid Leukaemia	Dibrugarh	27	Tonsil	Chandigarh	58
Brain, NS	Dibrugarh	44	Rectum	Chandigarh	41
Liver	Dibrugarh	28	Gall Bladder	Chandigarh	101
Thyroid	Dibrugarh	17	Oth. Oropharynx	Chandigarh	30

Table 3. Age wise in various regions from Male group

Place	0-14	15-34	35-64	65+
Mumbai	1331	3001	13904	4344
Bangalore	762	1178	6510	2823
Chennai	475	1717	9716	3823
Thiruvanthapuram	1098	1627	10982	5512
Dibrugarh	67	194	1807	827
Guwahati	114	378	4266	2045
Chandigarh	289	286	1459	609

Table 4. Age wise in various regions from Female group

Place	0-14	15-34	35-64	65+
Mumbai	622	2121	13163	2622
Bangalore	477	1262	9175	2211
Chennai	313	1665	12692	2829
Thiruvanthapuram	776	2232	12046	3755
Dibrugarh	46	227	1690	313
Guwahati	87	406	3414	772
Chandigarh	133	214	1398	347

III. STUDIES TO CANCER DISEASE VISUALIZATION

SPSS visualizations are explored the data effectively with various parameters. Our dataset consists of region wise, gender wise and age wise data (Table I to IV). Each region has described the number of people who suffering from type of the cancer and age of the person is carried out.

Table 5: Age wise Descriptive

DESCRIPTIVE				
Age			Statistic	Std. Error
0-14	Mean		470.71	108.039
	95% Confidence Interval for Mean	Lower Bound	237.31	
		Upper Bound	704.12	
	5% Trimmed Mean		446.52	
	Median		394	
	Variance		163412.681	
	Std. Deviation		404.243	
	Minimum		46	
	Maximum		1331	
	Range		1285	

	Interquartile Range		658	
	Skewness		880	597
	Kurtosis		34	1.154
15-34	Mean		1179.14	244.651
	95% Confidence Interval for Mean	Lower Bound	650.61	
		Upper Bound	1707.68	
	5% Trimmed Mean		1132.66	
	Median		1220	
	Variance		837955.67	
	Std. Deviation		915.399	
	Minimum		194	
	Maximum		3001	
	Range		2807	
	Interquartile Range		1547	
	Skewness		0.468	0.597
	Kurtosis		-0.835	1.154
35-64	Mean		7301.57	1300.933
	95% Confidence Interval for Mean	Lower Bound	4491.08	
		Upper Bound	10112.07	
	5% Trimmed Mean		7262.75	
	Median		7842.5	
	Variance		23693986.26	
	Std. Deviation		4867.647	
	Minimum		1398	
	Maximum		13904	
	Range		12506	
	Interquartile Range		10430	
	Skewness		-0.009	0.597
	Kurtosis		-1.823	1.154
65+	Mean		2345.14	437.018
	95% Confidence Interval for Mean	Lower Bound	1401.02	
		Upper Bound	3289.26	
	5% Trimmed Mean		2282.1	

Median	2416.5	
Variance	2673790.44	
Std. Deviation	1635.173	
Minimum	313	
Maximum	5512	
Range	5199	
Interquartile Range	3041	
Skewness	0.349	0.597
Kurtosis	0.778	1.154

In Graphs, Graph Builder to choose Boxplot and classification analysis using classify and tree method.

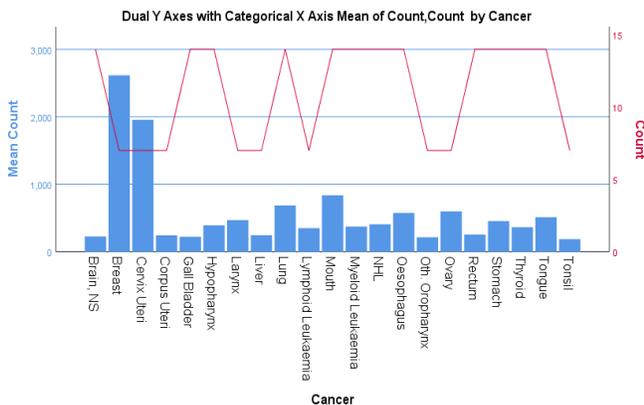


Figure 1: Cancer types of mean value

Breast Cancer, Cervix Uteri are increased to other type of cancer in graph representation of various regions. Tonsil, Oropharynx, Gall Bladder are less type of cancer in graph representation of various regions.

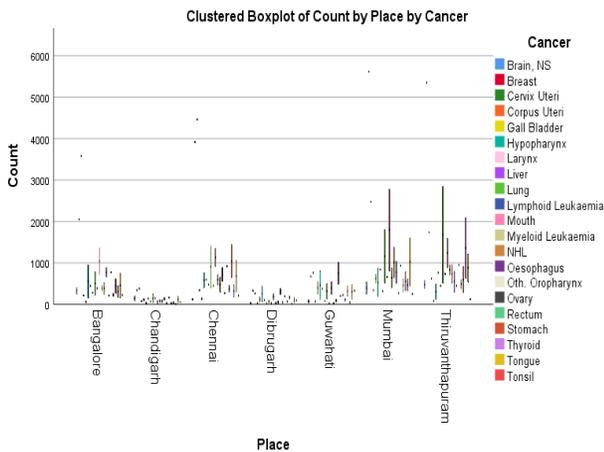


Figure 2: Region wise of Cancer types

Breast Cancer is increased in Thiruvanthapuram, Mumbai and Chennai. Cervix Uteri is increased in Chennai and Bangalore.



Figure 3: Age wise of various region

Cancer patients of age 35 to 64 people are large compare to any other age group in various regions.

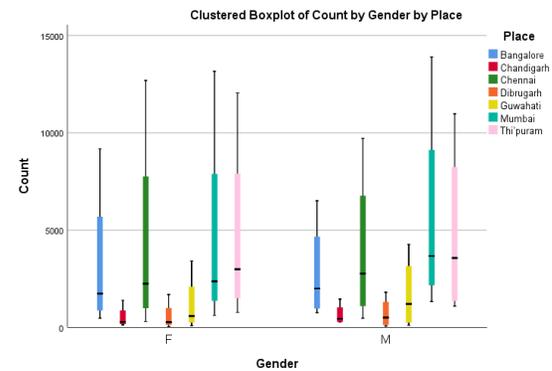


Figure 4: Gender wise distribution of various regions

In Male, Cancer patients are increased in Mumbai, Thiruvanthapuram and more compared to Female. In Female, Chennai and Bangalore places cancer patients are increased to more than Male Gender.

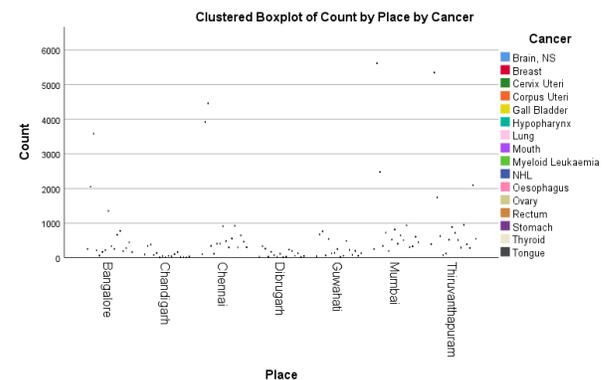


Figure 5: Region wise of cancer types from Female

In Female, Breast, Cervix Uteri, Thyroid cancer patients increased in various regions.

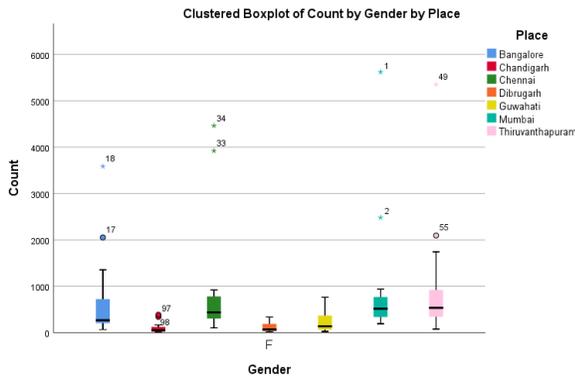


Figure 6: Region wise count from Female Group

From Female group Thiruvanthapuram, Mumbai, Chennai are more number of cancer patients.

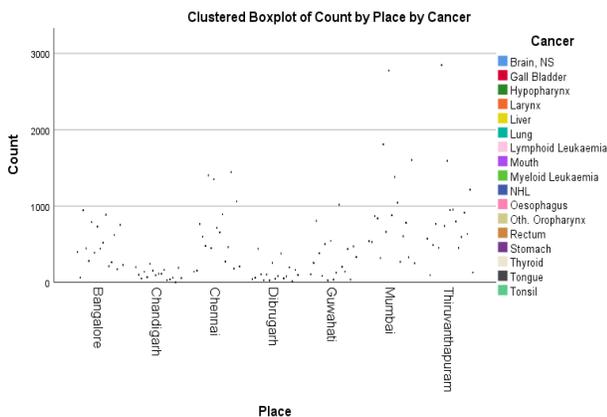


Figure 7: Region wise cancer types from Male Group

In Male group, Lung, Mouth, Tongue, Stomach type of cancer increased in various locations.

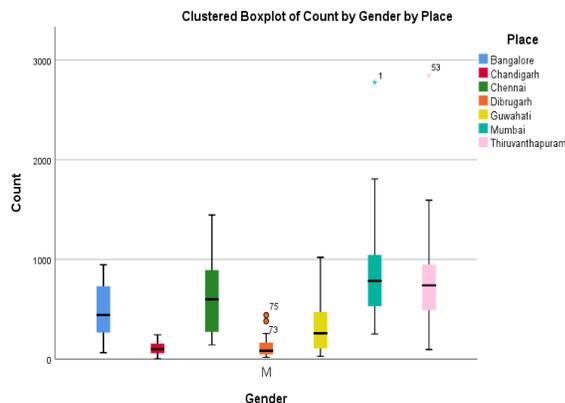


Figure 8: Region wise count from Male Group

In Male group, Places of cancer patients are increased in Mumbai, Thiruvanthapuram, Chennai, Bangalore, etc.

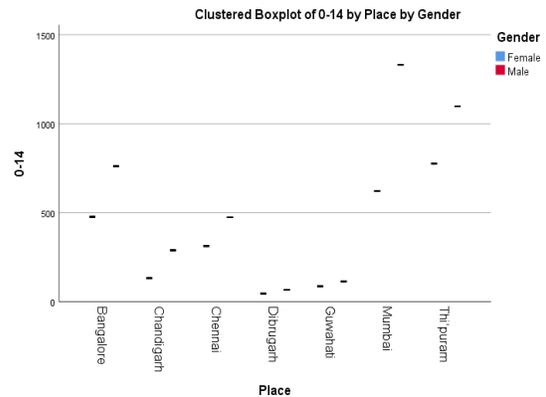


Figure 9: Gender wise of various regions in 0-14 age groups

In the childhood cancer patients are more in Mumbai (1953 cancer patients), Thiruvanthapuram (1874 cancer patients), Bangalore (1239 cancer patients), etc.

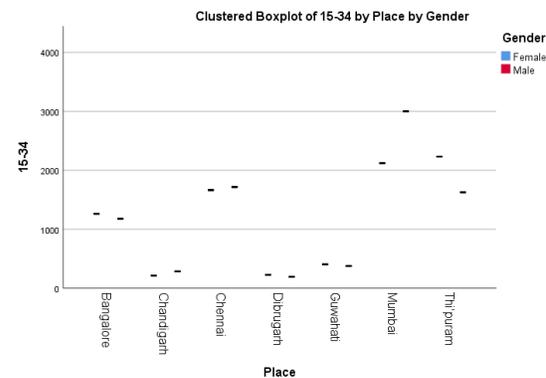


Figure 10: Gender wise of various regions in 15-34 age groups

More number of cancer patients in Thiruvanthapuram, Mumbai, and Chennai regions.

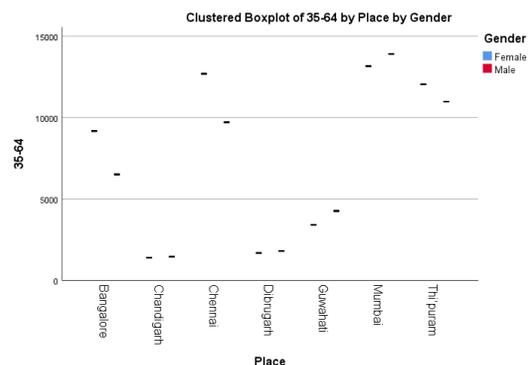


Figure 11: Gender wise of various regions in 35-64 age groups

More number of cancer patients in Thiruvanthapuram, Mumbai, and Chennai regions. In Guwahti and Chandigarh is varied in 15-34 years and 35-64 age groups.

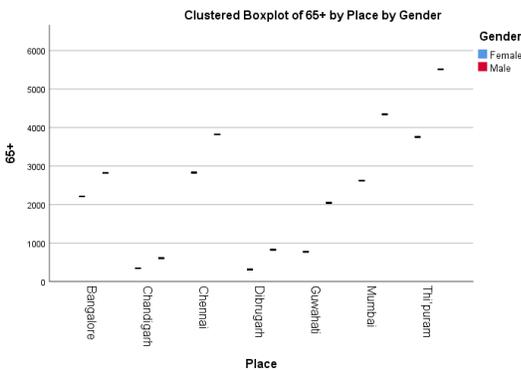


Figure 12: Gender wise of various regions in 65+ age groups

Number of cancer patients is increased in Thiruvanthapuram, Mumbai, Chennai, Bangalore, etc.

CHAID Decision tree in SPSS is support decision making values in the gender wise list of Cancer patients in different type of cancer are mentioned in the below figure. Total Percentage between male and female are 51.5 and 48.5. The decision tree describes the type of cancer in male and female with their percentage.

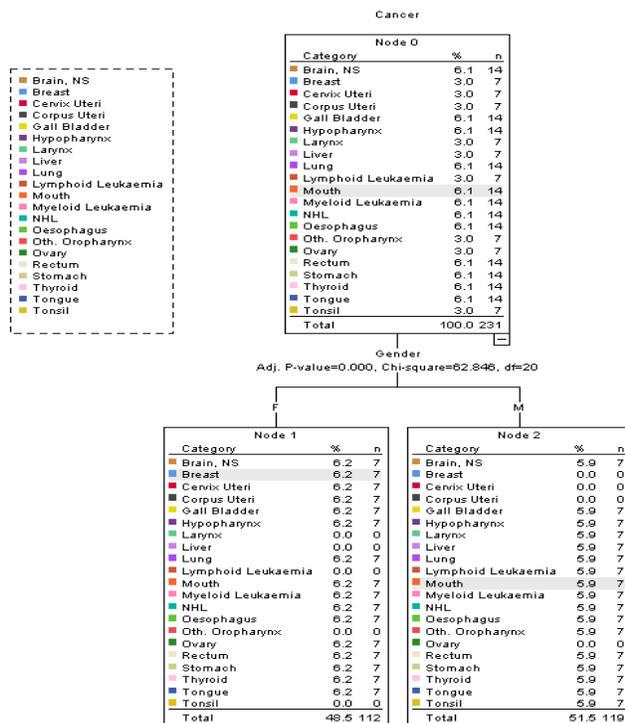


Figure 13: Gender wise comparison with cancer type

Classification of cancer patients in various regions using range values of parameters. The parameters are 1-500, 501-1000, 1001-2000, 2001-3000, 3001-4000, 4001-5000, 6000+ using Recode into Different Variables from SPSS and values are 1,2,3,4,5,6,7 respectively. From this decision graph all cancer type patients occurs in all regions. But Breast cancer, Cervix Uteri, Lung and Mouth cancer are more number of patients in regions.

Cancer patients are increased in various regions with respect to their cancer type in the following.

Breast Cancer-> Mumbai, Thiruvanthapuram, Bangalore

Cervix Uteri Cancer -> Chennai, Bangalore, Mumbai

Lung Cancer -> Thiruvanthapuram, Mumbai

Mouth Cancer -> Mumbai

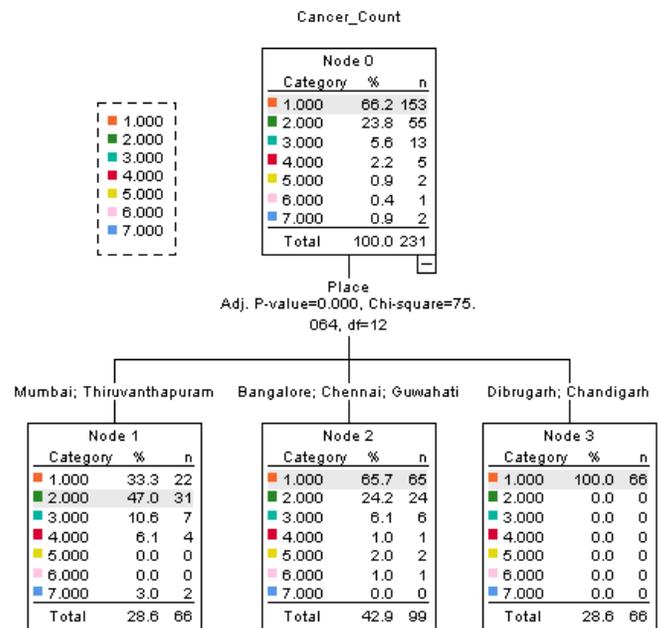


Figure 14: Range wise analysis of various regions in types of cancer

IV. RESULT AND CONCLUSION

SPSS is representing the graph builder to generate graph based on user requirement analysis and it is clearly influences the value to get inferential results with datasets. Graph builder creates the graph of Dual axes, Box Plot, Scatter Plot, etc are easily represented the dataset with effective manner. Datasets consists of groups based on gender wise, region wise, and age wise of multiple type of

cancer in our country. From these datasets easily explored the graphical representation in multiple view analysis are explored. SPSS classification is one of the decision support system applications to express the representation in visualization. Cancer leading sites are increased in our world based several factors of life style. According to visualization statistical data is easy to find the solution in various parameters of food based criteria, avoid tobacco, immediate treatment to recognize the cancer type, region people adaptation, population strength, cancer test based on age factor, cancer type of identification in region wise, gender wise calculation, childhood cancer type, age based cancer type occurrence, etc. From these all phenomenon, diseases are mainly involved in our biological cell activity. How to protect our health based on maintenance of cell activity is crucial thing any of the country. Research and Development projects are very important to avoid the cancer disease of multiple types. Cancer is not like one disease and it consists of many types. Today in the world, how it is different types of cancer occurred is one of the risk task and several types of cancer. Research is required for many people of clinical trials are required to find the cause or treatment of the cancer.

ACKNOWLEDGMENT

The authors thank Director, Bharathiar University, and Coimbatore for his encouragement in undertaking this study for R and D Development.

REFERENCES

- [1] S. Krishnaveni and Dr. M. Hemalatha, "A perspective analysis of Traffic Accident Using Data Mining Techniques", International Journal of Computer Application.
- [2] Jusoh Shaidah and Alfawareh Hejab M., "Techniques Applications and Challenging Issue in Text Mining uses, Applications", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 2, November 2012 .
- [3] Aarti Sharma, Rahul Sharma, Vivek Kr. Sharma, Vishal Shrivastava, "Application of Data Mining – Computer Science and Information Technologies", Vol. 5, Issue 2, Pg. no. 2023-2025. 2014.
- [4] A Multimodal SVM Approach for Fused Biometric Recognition Geethu S Kumar Jyothirmati Devi Department of Computer Science and Engineering College of Engineering , Chengannur, Kerala.
- [5] D. Barbara, J. Couto, S. Jajodia, L. Popyack, and N. Wu, "ADAM: Detecting intrusions by data mining," in Proc. 2nd Annu. IEEE Workshop Inf. Assur.Secur., New York, Jun. 2001, pp. 11-16.
- [6] Mohammad Khubeb Siddiqui and Shams Naahid, "Analysis of KDD CUP 99 dataset using Clustering based Data Mining", International Journal of Database Theory and Application, Vol. 6, No.5, 2013, pp. 23-34
- [7] Sarfraz Ahmed, "Global cancer statistics," CA Cancer Journal for Clinicians, vol. 61, no. 2, pp. 69-90, 2011.
- [8] Xingquan Zhu, Ian Davidson, "Knowledge Discovery and Data Mining: Challenges and Realities", ISBN 978- 1-59904- 252, Hershey, New York, 2007.
- [9] Lior Rokach and Oded Maimon, "Data Mining with Decision Trees: Theory and Applications (Series in Machine Perception and Artificial

Intelligence)", ISBN: 981-2771- 719, World Scientific Publishing Company, , 2008.

Authors Profile

Mr. Manikanta V MS in Computer Science and Technology from University of Mysore. He is currently working as Senior Resource Person (Technical) in Central Institute of Indian Languages, Mysore. He is expertised in Microsoft Technology, ArcGIS, MOSS, PHP, MYSQL, OPEN SOURCE Softwares, Data Mining, Matlab, .Net, VB, JAVA, ANDROID Development, BIG Data Analysis, etc. He is published 6 research articles and 4 International conference papers.

Dr N G Yethiraj was working as Assistant Professor in Department of Computer Science, Maharani's Arts and Science College, Bangalore.