**PROJECT REPORT ON**



**Impact Of Lifestyle on Human Health**

A project report submitted in partial fulfilment of requirements for the degree

of M. Sc (Statistics)with specialization in Industrial Statistics

Presented by

Miss. Jamadar Chetana Sunilsing



**Department of statistics**

**school of mathematical sciences**

**kavayitri bahinabai chaudhari north maharashtra**

**university,jalgaon-425001**

**(Academic Year:2022-23)**

**INDEX **



**Contents**

1. **INTRODUCTION**
2. **Purpose of STDY**
3. **Determination of sample size**
4. **Wald-wolfowitz (or Runs) test for randomness**

**5)Chi-Square Test of Independence**

**6)Chi-Square Test of Fisher exact**

**7) General linear model**

**8)Graphical Repesentation**

**9)Conclusion.**

**INTRODUCTION** 





Life style is a way used by people, groups & nations & is formed in specific geographical, economic, political, cultural & religious text. It includes day behaviours & functions of individuals in job, activities, fun & diet.

According to WHO, 60% of related factors to individual health & quality of life are correlated to lifestyle. Millions of people follow unhealthy lifestyle. Hence, they encounter illness, disability, & even death.

Today, wide changes have occurred in life of all people. Malnutrition, unhealthy dies, smoking, alcohol consuming, drug abuse, stress and so on, are the presentations of unhealthy life style that they are used as dominant form of lifestyle.

Therefore, according to the existing studies, it can be said that: lifestyle has a significant influence on physical and mental health of human being significant influence on physical and mental health of human being.

**Variable of lifestyle on health are as follows:-**

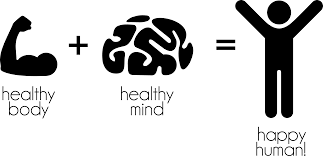
**1)Eating a Fruit/Vegetable:-**



Vegetables and fruits are important part of a healthy Lifestyle, and variety is as important quantity.

A diet rich in vegetables and fruits can lower blood pressure, reduce the risk of heart disease and stroke, prevent some types of cancer, lower risk of eye and digestive problems, and have a positive effect upon blood sugar, which can help keep appetite in check.

**2)Physically healthy:-**

****

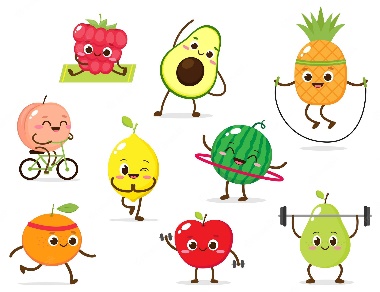
Regular physical activity is one of the most important things you can do for your health. Being physically active can improve your [brain health](https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm#brain-health), help [manage weight](https://www.cdc.gov/healthyweight/physical_activity/index.html), reduce the [risk of disease](https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm#reducing-disease), strengthen [bones and muscles](https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm#bones-muscles), and improve your ability to do [everyday activities](https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm#prevent).

**3)Sleep:-**



Sleep is essential to every process in the body, affecting our physical and mental functioning the next day, our ability to fight disease and develop immunity, and our metabolism and chronic disease risk.

**4)Exercise:-**



Regular exercise keeps your muscles, bones and joints strong and functioning well, but also helps prevent falls and fractures. Try doing exercises that strengthen your bones and muscles twice a week. Exercise burns calories and helps to maintain a healthy weight which in turn reduces your risk of many health problems.

**4)Drinking water:-**



Drinking water can prevent dehydration, a condition that can cause unclear thinking, result in mood change, cause your body to overheat, and lead to constipation and kidney stones.

Etc.

**Purpose of Study**

The main aim of survey is to point out health & lifestyle issues appropriately to improve the health status of a person.

➢ To manage the circumstances that hinder in achieving a healthy lifestyle.

➢ To access & evaluate the health & lifestyle of people in particular area.

➢TO study the effect of life style in routine.

**Data collection**

➢Collect as data online questionnaire due to COVID-19 situation.

**Data type**

➢Secondary data.

**Population size**

➢300.

**TOOLS**

➢Graphical representation, Hypothesis testing, RBD, etc, by using SAS software

**Keyword**

**Determination of sample size**

**• Sample size:-**

The sample size for a study needs to be estimated at the time the study is proposed; too large a sample is unnecessary and unethical, and too small a sample is unscientific and also unethical. The necessary sample size can be calculated, using statistical software, based on certain assumptions. If no assumptions can be made, then an arbitrary sample size is set for a pilot study.

•**Margin of error:-(MOE)**

The margin of error is a statistic expressing error in the results of survey. The margin of error helps to determine the accuracy of a value, by expressing in a range.

Formula:-(1) Sample size (2) MOE

1)n =N/(1+N\*e^2)

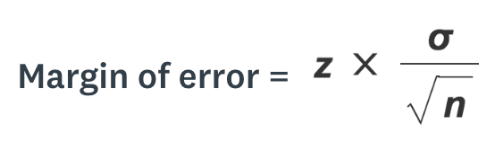
• where,

N=Population size

n=Sample size

e=MOE

2)



**n = sample size**

**• σ = population standard deviation**

**• z = Confidence level**

•**calculating by sample size and MOE in SAS: -**

**Command: -**

**/\*Determination of sample size\*/**

**Data first;**

**N=300;**

**e=0.1;/\*margin of error\*/**

**n=N/(1+N\*e\*\*2);**

**Output: -**

|  |  |  |
| --- | --- | --- |
|  | N | e |
| 1 | 75 | 0.1 |

**/\*To find the Margin of Error\*/**

**data second;**

**sum\_xi=300;**

**n=75;**

**mean=sum\_xi/n;**

**std=0.9867;**

**z\_value=1.65;/\*the confidence interval z\_value \*/**

**MOE=z\_value\*std/sqrt(n);/\*margin of error\*/**

**run;**

**proc print;**

**Output: -**

| **Obs** | **sum\_xi** | **n** | **mean** | **std** | **z\_value** | **MOE** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | 300 | 75 | 4 | 0.9867 | 1.65 | 0.18799 |

## **Where , the MOE 0.18799**

## ****wald****-Wolfowitz (or runs) test for

## randomness

**A randomness test (or test for randomness), in data evaluation, is a test used to analyze the distribution of a set of data to see if it can described as random.**

The Wald-Wolfowitz test, also known as the Runs test for randomness, is used to test the hypothesis that a series of numbers is random. A run is a set of sequential values that are either all above or below the mean. To simplify computations, the data are first centered about their mean. To carry out the test, the total number of runs is computed along with the number of positive and negative values. A positive run is then a sequence of values greater than zero, and a negative run is a sequence of values less than zero. We can then test if the number of positive and negative runs are distributed equally in time.

The test statistic is asymptotically normally distributed, so this program computes Z, the large sample test statistic, as follows:

Z = (R – E(R)) / sqrt(V(R))

where R is number of runs. The expected value and variance of R are:

E(R) = (2nm / (n + m)) + 1

V(R) = (2nm (2nm – n – m)) / ((n + m)2 (n + m – 1))

where n is the number of positive values and m is the number of negative values.

**The of hypothesis in randomness data**

H0: Sample is random

H1: Sample is not random

To test above hypothesis

1= user is male

2= user is female

**Command: -Test of randomness**

**data MYDATA.Gender;**

**drop i;**

**do i=1 to 75;**

**d=rannor (12);**

**output;**

**end;**

**run;**

**proc standard data=MYDATA.Gender out=two mean=0;**

**var d;**

**run;**

**data runcount;**

**set two nobs=nobs;**

**if d=0 then delete;**

**if d>0 then n+1;**

**if d<0 then m+1;**

**retain runs 0 numpos 0 numneg 0;**

**previous=lag(d);**

**if \_n\_=1 then do;**

**runs=1;**

**prevpos=.;**

**currpos=.;**

**prevneg=.;**

**currneg=.;**

**end;**

**else do;**

**prevpos= (previous > 0);**

**currpos= (d > 0);**

**prevneg= (previous < 0);**

**currneg= (d < 0);**

**if \_n\_=2 and (currpos and prevpos) then numpos+1;**

**else if \_n\_=2 and (currpos and prevneg) then numneg+1;**

**else if \_n\_=2 and (currneg and prevpos) then numpos+1;**

**else if \_n\_=2 and (currneg and prevneg) then numneg+1;**

**if currpos and prevneg then do;**

**runs+1;**

**numpos+1;**

**end;**

**if currneg and prevpos then do;**

**runs+1;**

**numneg+1;**

**end;**

**end;**

**run;**

**data runcount;**

**set runcount end=last;**

**if last;**

**run;**

**data waldwolf;**

**label z='Wald-Wolfowitz Z'**

**pvalue='Pr > |Z|';**

**set runcount;**

**mu = ((2\*n\*m) / (n + m)) + 1;**

**sigmasq = ((2\*n\*m) \* (2\*n\*m-(n+m))) / ( ((n+m)\*\*2) \* (n+m-1) );**

**sigma=sqrt(sigmasq);**

**drop sigmasq;**

**if N GE 75 then Z = (Runs - mu) / sigma;**

**else if Runs-mu LT 0 then Z = (Runs-mu+0.5)/sigma;**

**else Z = (Runs-mu-0.5)/sigma;**

**pvalue=2\*(1-probnorm(abs(Z)));**

**run;**

**title 'Wald-Wolfowitz Test for Randomness';**

**title2 'H0: The data are random';**

**proc print data=waldwolf label noobs;**

**var z pvalue;**

**format pvalue pvalue.;**

**run;**

**Output: -**

**Wald-Wolfowitz Test for Randomness**

**H0: The data are random**

| **Wald-Wolfowitz Z** | **Pr > |Z|** |
| --- | --- |
| 1.77915 | 0.0752 |

**Conclusion: -**

As you would expect of values from a random number generator, the test accept Ho the null hypothesis that the data is random (p=0.0752).

**A) Chi-Square Test of Independence**

A [Chi-Square Test of Independence](https://www.statology.org/chi-square-test-of-independence/) is used to determine whether or not there is a significant association between two [categorical variables](https://www.statology.org/categorical-vs-quantitative/).

The chi-square test of independence is an [inferential](https://www.scribbr.com/statistics/inferential-statistics/) statistical test, meaning that it allows you to draw conclusions about a [**population**](https://www.scribbr.com/methodology/population-vs-sample/) based on a **sample**. Specifically, it allows you to conclude whether two variables are related in the population.

Like all hypothesis tests, the chi-square test of independence evaluates a [null and alternative hypothesis](https://www.scribbr.com/statistics/null-and-alternative-hypotheses/).

**Command:-**

A)To check whether the two variable i.e.’exercise routine’ and ‘sickness’are independent or not.

/\*chi-test\*/

data first;

input routine $ sick $ total;

datalines;

d one 21

d three 8

d mo 0

d ne 10

s one 10

s three 4

s mo 2

s ne 2

we one 4

we three 5

we mo 2

we ne 1

ne one 2

ne three 2

ne mo 2

ne ne 0

;

run;

proc print;

/\*perform Chi-Square Goodness of Fit test\*/

proc freq data=first;

tables routine\*sick / chisq;

weight total;

**Output: -**

**The FREQ Procedure**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Frequency**  **Percent**  **Row Pct**  **Col Pct** | | | **Table of routine by sick** | | | | | | | --- | --- | --- | --- | --- | --- | | **routine** | **sick** | | | | | | **mo** | **ne** | **one** | **three** | **Total** | | **d** | 0  0.00  0.00  0.00 | 10  13.33  25.64  76.92 | 21  28.00  53.85  56.76 | 8  10.67  20.51  42.11 | 39  52.00 | | **ne** | 2  2.67  33.33  33.33 | 0  0.00  0.00  0.00 | 2  2.67  33.33  5.41 | 2  2.67  33.33  10.53 | 6  8.00 | | **s** | 2  2.67  11.11  33.33 | 2  2.67  11.11  15.38 | 10  13.33  55.56  27.03 | 4  5.33  22.22  21.05 | 18  24.00 | | **we** | 2  2.67  16.67  33.33 | 1  1.33  8.33  7.69 | 4  5.33  33.33  10.81 | 5  6.67  41.67  26.32 | 12  16.00 | | **Total** | 6  8.00 | 13  17.33 | 37  49.33 | 19  25.33 | 75  100.00 | |

**Statistics for Table of routine by sick**

| **Statistic** | **DF** | **Value** | **Prob** |
| --- | --- | --- | --- |
| **WARNING: 69% of the cells have expected counts less than 5. Chi-Square may not be a valid test.** | | | |
| **Chi-Square** | 9 | 15.9116 | 0.0688 |
| **Likelihood Ratio Chi-Square** | 9 | 17.5418 | 0.0409 |
| **Mantel-Haenszel Chi-Square** | 1 | 0.0007 | 0.9784 |
| **Phi Coefficient** |  | 0.4606 |  |
| **Contingency Coefficient** |  | 0.4184 |  |
| **Cramer's V** |  | 0.2659 |  |

**Sample Size = 75**

**Conclusion**

There are two values of interest in the output:

* Chi-Square Test Statistic: 15.9116
* Corresponding p-value: 0.0688

Recall that the Chi-Square Test of Independence uses the following null and alternative hypotheses:

* **H0:**The two variables are independent.
* **HA:**The two variables are not independent.

Since the [p-value](https://www.statology.org/p-values-statistical-significance/) (0.0688) of the test is not less than 0.05, we fail to reject the null hypothesis.

This means we do not have sufficient evidence to say that there is an association between exercise routine and sick in year preference.

In other words, exercise routine and sick in year preference are independent.

**B) CHI-SQUARE TEST OF FISHER EXACT**

**Fisher’s Exact Test**is used to determine whether or not there is a significant association between two categorical variables.

Fisher’s Exact Test uses the following null and alternative [hypotheses](https://www.statology.org/hypothesis-testing/):

* **H0: (null hypothesis)**The two variables are independent.
* **H1: (alternative hypothesis)**The two variables are not independent.

**Command:-**

data second;

input healthy $ eating\_fruit $ total;

datalines;

yes daily 23

yes weekly 25

yes monthly 2

yes sometimes 11

no daily 1

no weekly 7

no monthly 4

no sometimes 2

;

run;

proc print;

proc freq data=second order=data;

weight total;

tables healthy\*eating\_fruit/ fisher;

run;

**Output: -**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Frequency**  **Percent**  **Row Pct**  **Col Pct** | | | **Table of healthy by eating\_fruit** | | | | | | | --- | --- | --- | --- | --- | --- | | **healthy** | **eating\_fruit** | | | | | | **daily** | **weekly** | **monthly** | **sometime** | **Total** | | **yes** | 23  30.67  37.70  95.83 | 25  33.33  40.98  78.13 | 2  2.67  3.28  33.33 | 11  14.67  18.03  84.62 | 61  81.33 | | **no** | 1  1.33  7.14  4.17 | 7  9.33  50.00  21.88 | 4  5.33  28.57  66.67 | 2  2.67  14.29  15.38 | 14  18.67 | | **Total** | 24  32.00 | 32  42.67 | 6  8.00 | 13  17.33 | 75  100.00 | |

**Statistics for Table of healthy by eating\_fruit**

| **Statistic** | **DF** | **Value** | **Prob** |
| --- | --- | --- | --- |
| **WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test.** | | | |
| **Chi-Square** | 3 | 12.7382 | 0.0052 |
| **Likelihood Ratio Chi-Square** | 3 | 11.4680 | 0.0094 |
| **Mantel-Haenszel Chi-Square** | 1 | 2.4285 | 0.1191 |
| **Phi Coefficient** |  | 0.4121 |  |
| **Contingency Coefficient** |  | 0.3810 |  |
| **Cramer's V** |  | 0.4121 |  |

| **Fisher's Exact Test** | |
| --- | --- |
| **Table Probability (P)** | 0.0002 |
| **Pr <= P** | 0.0073 |

**Sample Size = 75**

**Conclusion**

The null hypothesis for Fisher’s Exact Test is that the two variables are independent. our null hypothesis is that Healhty and eating food preference are independent, which is a two-sided test.

Thus, we’ll look at the **two-sided p-value** in the final table of the output, which turns out to be 0.0073

Since this p-value is less than 0.05, we reject the null hypothesis.

This means we do have sufficient evidence to say that there is a significant association between healthy and eating fruit.

**General linear model**

To check the relationship betaween time spend on tv and sleep by different age group is independent

**Command:-**

data time;

input time\_spendtv $ time\_sleep $ Age;

datalines;

3-5\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 22

2-3\_hours 7-12\_hours 19

2-3\_hours 7-12\_hours 21

2-3\_hours 4-7\_hours 31

2-3\_hours 4-7\_hours 17

2-3\_hours 7-12\_hours 26

Morethan\_5\_hours 7-12\_hours 21

3-5\_hours 7-12\_hours 22

2-3\_hours 4-7\_hours 20

2-3\_hours 4-7\_hours 20

3-5\_hours 4-7\_hours 21

2-3\_hours 4-7\_hours 21

2-3\_hours 4-7\_hours 20

2-3\_hours 7-12\_hours 21

2-3\_hours 4-7\_hours 22

2-3\_hours 7-12\_hours 36

2-3\_hours 7-12\_hours 32

2-3\_hours 7-12\_hours 19

More\_than\_5\_hours 7-12\_hours 21

3-5\_hours 7-12\_hours 18

More\_than\_5\_hours 7-12\_hours 21

3-5\_hours 7-12\_hours 20

2-3\_hours 2-4\_hours 21

More\_than\_5\_hours 2-4\_hours 18.9

2-3\_hours 7-12\_hours 20

2-3\_hours 4-7\_hours 20

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 21

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 17

3-5\_hours 4-7\_hours 21

3-5\_hours 7-12\_hours 21

More\_than\_5\_hours 7-12\_hours 24

More\_than\_5\_hours 7-12\_hours 28

3-5\_hours 7-12\_hours 15

2-3\_hours 4-7\_hours 17

2-3\_hours 7-12\_hours 21

3-5\_hours 7-12\_hours 17

2-3\_hours 4-7\_hours 21

More\_than\_5\_hours 7-12\_hours 20

More\_than\_5\_hours 7-12\_hours 20

3-5\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 15

2-3\_hours 7-12\_hours 20

3-5\_hours 7-12\_hours 20

More\_than\_5\_hours 7-12-hours 20

3-5\_hours 7-12\_hours 20

3-5\_hours 7-12\_hours 22

More\_than\_5\_hours 7-12\_hours 23

2-3\_hours 4-7\_hours 23

More\_than\_5\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 22

3-5\_hours 7-12\_hours 21

3-5\_hours 7-12\_hours 20

More\_than\_5\_hours 4-7\_hours 21

2-3\_hours 7-12\_hours 20

More\_than\_5\_hours 4-7\_hours 21

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 20

2-3\_hours 4-7\_hours 20

2-3\_hours 4-7\_hours 16

2-3\_hours 4-7\_hours 21

4-7\_hours 4-7\_hours 27

2-3\_hours 4-7\_hours 27

More\_than\_5\_hours 7-12\_hours 21

2-3\_hours 7-12\_hours 20

2-3\_hours 7-12\_hours 21

More\_than\_5\_hours 7-12\_hours 28

2-3\_hours 4-7\_hours 17

2-3\_hours 4-7\_hours 20

2-3\_hours 4-7\_hours 17

3-5\_hours 7-12\_hours 22

;

run;

proc print;

proc glm data=time;

class time\_spendtv time\_sleep;

model Age =time\_spendtv time\_sleep;

run;

**Output: -**

**The GLM Procedure**

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **time\_spendtv** | 5 | 2-3\_hour 3-5\_hour 4-7\_hour More\_tha Morethan |
| **time\_sleep** | 4 | 2-4\_hour 4-7\_hour 7-12-hou 7-12\_hou |

|  |  |
| --- | --- |
| **Number of Observations Read** | 75 |
| **Number of Observations Used** | 73 |

**The GLM Procedure**

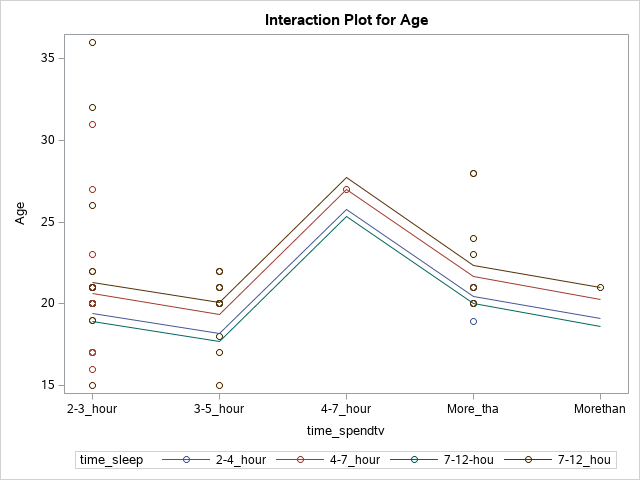
**Dependent Variable: Age**

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **Model** | 7 | 80.9905616 | 11.5700802 | 0.91 | 0.5011 |
| **Error** | 65 | 822.3042329 | 12.6508344 |  |  |
| **Corrected Total** | 72 | 903.2947945 |  |  |  |

| **R-Square** | **Coeff Var** | **Root MSE** | **Age Mean** |
| --- | --- | --- | --- |
| 0.089661 | 16.90517 | 3.556801 | 21.03973 |

| **Source** | **DF** | **Type I SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **time\_spendtv** | 4 | 64.68556375 | 16.17139094 | 1.28 | 0.2876 |
| **time\_sleep** | 3 | 16.30499790 | 5.43499930 | 0.43 | 0.7325 |

| **Source** | **DF** | **Type III SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **time\_spendtv** | 4 | 75.81913448 | 18.95478362 | 1.50 | 0.2130 |
| **time\_sleep** | 3 | 16.30499790 | 5.43499930 | 0.43 | 0.7325 |



**Conclusion**

the ho= relation is indepednce

and h1=relationship does not independent

conclusion p\_value 0.5011 therefore there are relationship between the variable and they are indepedent

Graph:-

From graph we can observe that as they ageses incrasing the watching tv they will impact on your sleeping.

**Graphical representation**

**1)Bubble chart shows you have often a green-vegetables and pulses**

**Command:-**

ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgplot data=MYDATA.VEGETABLES;

bubble x=name y=total size=total/ group=name bradiusmin=7 bradiusmax=14;

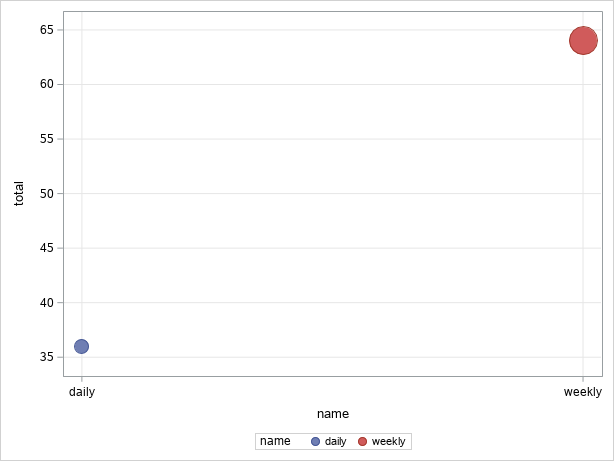
xaxis grid;

yaxis grid;

run;

ods graphics / reset;

**Output: -**



**Conclusion: -**

The blue bubble are shows 36% often a daily vegetable and red bubble are shows 64% of often a vegetable weekly.

**2)pie chart shows a junk food and living healthy or not**

**Command**

/\* Define Pie template \*/

proc template;

define statgraph SASStudio.Pie;

begingraph;

layout region;

piechart category='eating junk food'n / group='healthy / unhealthy'n

groupgap=2% datalabellocation=inside;

endlayout;

endgraph;

end;

run;

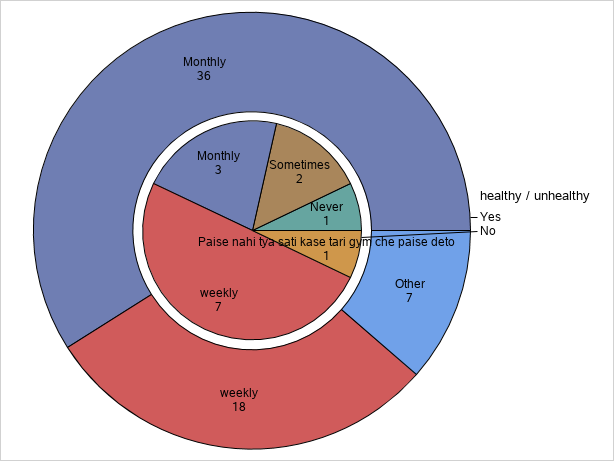
ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgrender template=SASStudio.Pie data=MYDATA.'EATING JUNK FOOD HEALTHY'n;

run;

ods graphics / reset;

**Output**



**Conclusion**

Eating junk food monthly 36 responses are healthy and 3 responses are unhealthy, weekly 18 response are healthy and 7 responses unhealthy, sometimes 2 responses unhealthy and other 7 responses healthy.

**3)the bar chart shows health check-up and healhty or not**

**Command**

ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgplot data=MYDATA.'HEALTH CHECKUP'n;

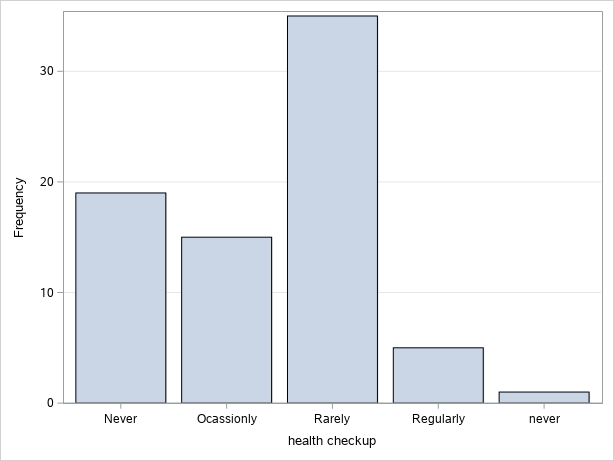
vbar 'health checkup'n /;

yaxis grid;

run;

ods graphics / reset;

**Output**



**Conclusion: -** The people often health check-up people rarely 35%

and 15% 2% people never check the health.

**4)The bar chart shows particularly take time out for your intreast/Hobbies**

**Command**

ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgplot data=MYDATA.HOBBIES;

title height=14pt "Hobbies";

vbar Hobbies / fillattrs= (color=CXCAD5E5 transparency=0.25) datalabel

fillType=gradient dataskin=gloss;

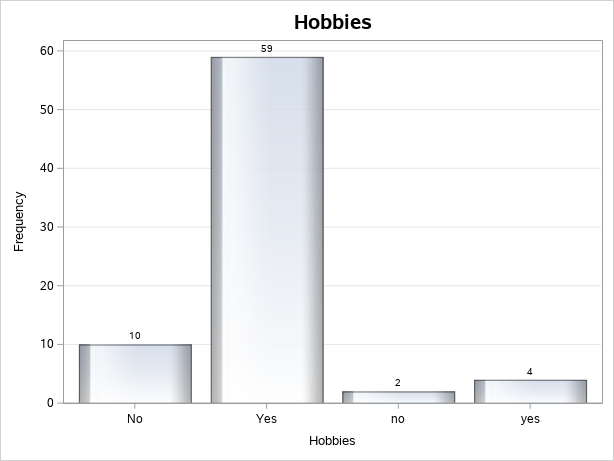
yaxis grid;

run;

ods graphics / reset;

title;

**Output**



**Conclusion**

The bar Chart shows total 63% people intreastead in hobbies and 12% not intreasted in hobbies

**5)The pie chart shows glass of water drink in day**

**Command**

/\* Define Pie template \*/

proc template;

define statgraph SASStudio.Pie;

begingraph;

entrytitle "Water drink" / textattrs=(size=14);

layout region;

piechart category='WATER DRINK'n / datalabelattrs=(size=10)

fillattrs=(transparency=0.25) dataskin=crisp;

endlayout;

endgraph;

end;

run;

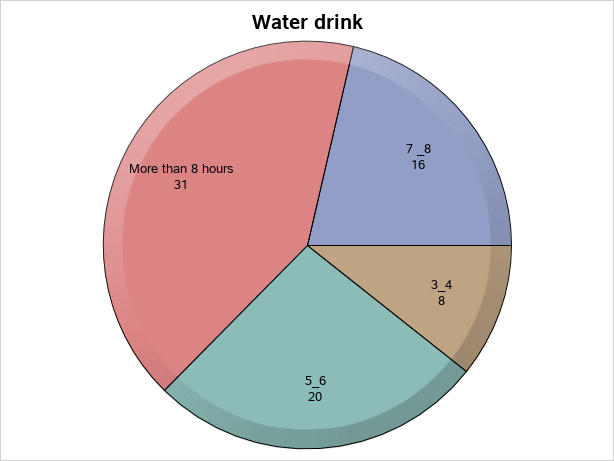
ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgrender template=SASStudio.Pie data=MYDATA.'WATER DRINK'n;

run;

ods graphics / reset;

**Output**



**Conclusion: -**

Many people drink water 31% more than 8 glass in a day ,5\_6 glass drink water in a people 20% in a day,7\_8 glass drink water in a people 16% in day, 3\_4 glass drink water in a people 8% in day.

**6) bar chart shows Sick vs dieses**

**Command: -**

ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgplot data=MYDATA.'SICK VS DIESES'n;

title height=14pt "Sick vs Dieses";

vbar sickness / group=Dieses groupdisplay=cluster

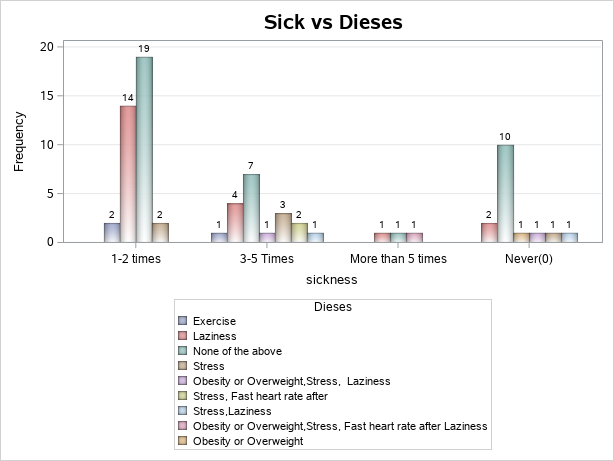
fillattrs=(transparency=0.25) datalabel fillType=gradient dataskin=pressed;

yaxis grid;

run;

ods graphics / reset;

title;



**Conclusion: -**

The graph shows 1\_2 time get sick in year the 19 % of people not any health issue

14% laziness and 2% of health issues are exercise and obesity or overweight.

**Conclusion**

**1)Chi-Square Test**

**A) Chi-Square Test of Independence**

We conclude that do not have sufficient evidence to say that there is an association between exercise routine and sick in year preference.

**B) Chi-Square Test of Fisher Exact**

We conclude that do have sufficient evidence to say that there is a significant association between healthy and eating fruit.

**2)General linear model**

We conclude that relationship betaween time spend on tv and sleep

**THANK YOU**