Part 1 Introduction

1.1 Abstract

Nowadays, many dual-income families especially grassroots families live at a fast pace, work at high pressure and often overlook good eating habits. Grassroots families may not have any knowledge about eating healthy and nutrition in food. This sector of the population, as a result of the factors listed above, see their daily diet as an obligation and so ignore proper nutritional and healthy requirements which leads to the development of bad eating habits and unhealthy lifestyle. This could also have a significant influence on their kids since they are the carers. As a consequence of such behaviour children's health can be adversely affected to a great degree.

The eating and cooking habits of grassroots families who currently reside in subdivided units in To Kai Wan, Kowloon City, were the subject of the study including how they eat, how their diet is made, their eating habits, etc. in an effort to determine the typical diet of children and the most prevalent unhealthy eating habits in order to develop an educational and interventional program for them. As well as examining the connection between these children's eating habits and their health. The end result is a list of the most common eating habits and an understanding of how much these children, who live in subdivided apartments, are currently misunderstood about their diet. Based on this, make a number of suggestions to improve the health of this group of people from the personal, catering, government, and relevant charitable organizations. This will serve as a guide for charities on how to better serve and support these individuals while also improving their health.

1.2 Introduction of People Service Center

People Service Center was founded in 2002 to address the growing gap between rich and poor in society, provide assistance to low-income families and the elderly, advocate for social change, and work toward creating a society that is fair and equal. By bringing together social forces, encouraging residents to lend a hand to one another, and advocating for the development of social policies, they hope to provide families living in poverty with care, change, and hope. The center we visited is To Kwa Wan Center which mainly provides services to families living in subdivided flats and the elderly.

1.3 Research Purpose and significance

We want to get the most up-to-date information about the needy, find out if families and children living in subdivided apartments get enough food, and find out if any significant factors encourage them to eat healthily. During the epidemic and other related work, we hope that this will also serve as a guide for individuals, the government, and society on how to improve people's health, promote nutritious and healthy eating and pay attention to the dietary health of grassroots citizens.

1.4 Research time and progress

09/2022	Expert consultation, design questionnaire
10/2022	Questionnaire survey, qualitative interview through phone calls and home visiting
11/2022	Data entry and analysis
11-12/2022	Writing a research report

1.5 Respondents Chosen

Family carers with at least one 3 to 12 years old kid and currently live in subdivided flats in To Kwa Wan, Kowloon City.

1.6 Feature Engineering

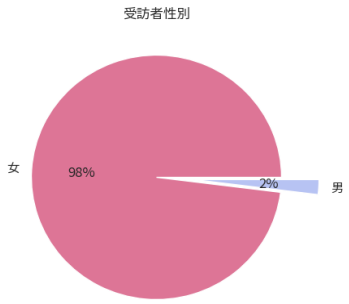
We created a new attribute called “child’s nutrition score”, which is used to check if the child fulfils the criteria of the food pyramid. If the child can accomplish the criteria, then he or she can get 1 mark. Here are the daily criteria per day: (Full mark = 6)

- 3 to 4 bowls of cereal
- 2 servings of vegetables
- 2 servings of fruit
- 3 to 5 taels of meat, fish, eggs or substitutes
- 2 servings of milk or alternatives
- 6 to 8 drinks

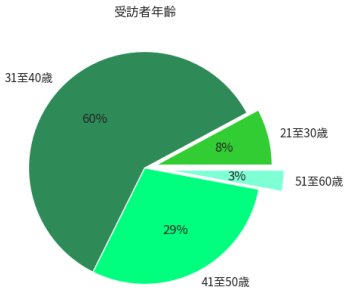
Part 2 Data Analysis

2.1 Demographic Data

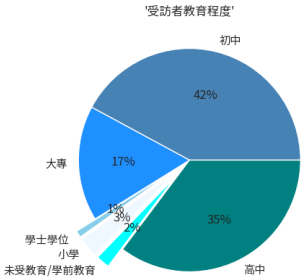
We spoke with 102 family caregivers. In the middle, 100(98%) of them are female while 4(2%) of them are male(Graph 1.1). The majority (60%) were between the ages of 31 and 40, while only 3% were between the ages of 51 and 60. Based on class marks from the group data, the average age is 38.2 years old (Graph 1.2). Only a small percentage of respondents have a bachelor's degree or have never attended college, but the majority of respondents graduated from junior high school(Graph 1.3). The majority of families only have one child between the ages of 3 and 12; only two families have three children between the ages of 3 and 12 (Graph 1.4). Only 1% of the families interviewed had more than six members, and many of them (40%) were families of four(Graph 1.5). Almost all of them do not share rooms with others(Graph 1.6).



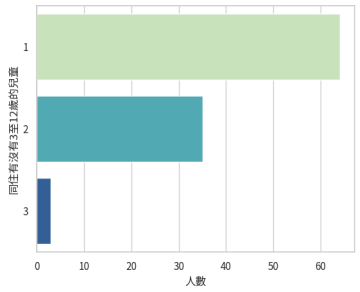
Graph 1.1 The gender



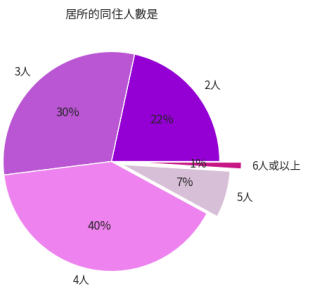
Graph 1.2 The age



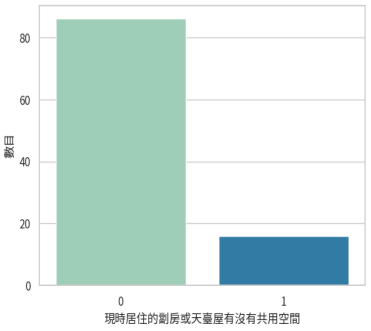
Graph 1.3 The highest education



Graph 1.4 The number of 3 to 12 years old children



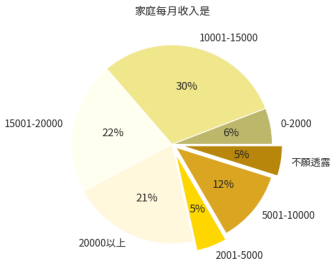
Graph 1.5 The number of family members



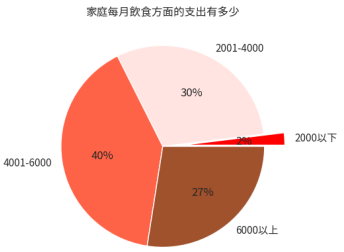
Graph 1.6 Existence of common area

2.2 Economic Data

30% of the interview families were between the monthly family income of HK$10001 and HK$15000, while only 5% were between the monthly family income of HK$0 and HK$2000. Only 2% of the families spend less than HK$2000 for food in a month, and many of them spend around HK$4001 to HK$6000 as monthly food expenses(Graph 2.2).



Graph 2.1 Monthly family income



Graph 2.2 Monthly family food expenses

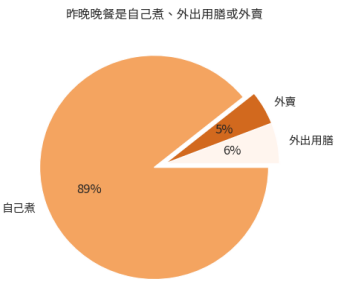
2.3 Eating Habits Data

The majority of families (89%) prepare their final meal at home; only 5% and 6% of the families, respectively, takeaway or ate at a restaurant (Graph 3.1).

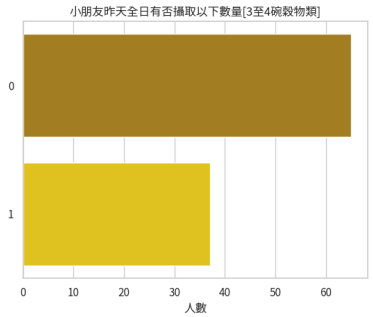
According to the Department of Health's Healthy Eating Food Pyramid for children, 64%, 56%, and 69% of children do not consume the recommended amounts of grains, fruits, and milk (Graphs 3.2, 3.4, and 3.6). 58%, 85%, and 54% of children consume the recommended amounts of vegetables, meat, and fluid(Graphs 3.3, 3.5, 3.7).

The most common reason for not eating a healthy diet is partiality toward vegetables, fruits, meat, fish, eggs and milk and alternatives. Different reasons like children can't eat anymore, don't cook excessively and don't have a lot of chance to plan are additionally normal reasons(Graph 3.8 - 3.13).

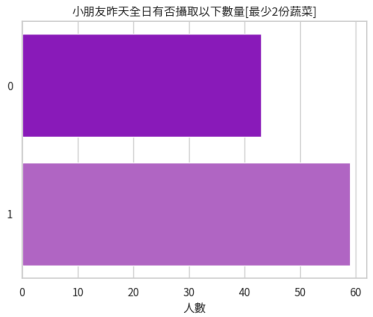
The majority of families spend the majority of the week preparing their final meal at home; Only 5% of the families ate twice a week at a restaurant. They would occasionally order takeout and eat at home. The average number of having dinner at home and eating at a restaurant and buying takeaways are 6.1 times, 0.4 times and 0.5 times respectively(Graph 3.14 - 3.16).



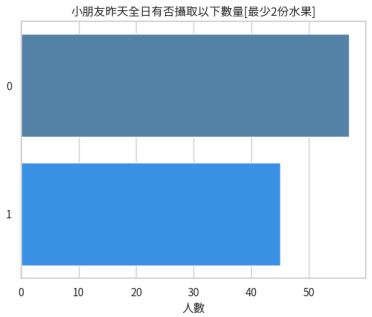
Graph 3.1 Eating habits for last dinner



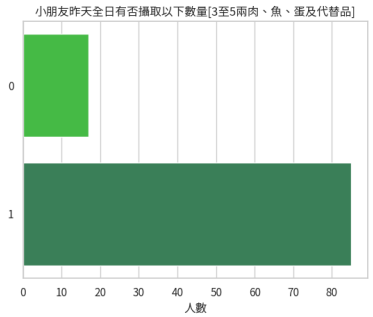
Graph 3.2 Whether the Children Ate at Least Three to Four Bowls of Grains Yesterday



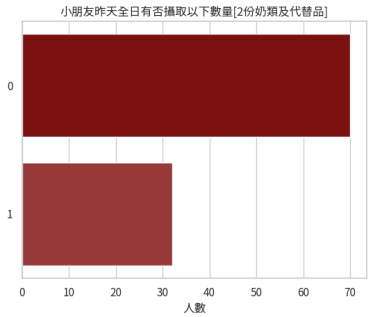
Graph 3.3 Whether the Children Ate at Least Two Servings of Vegetables Yesterday



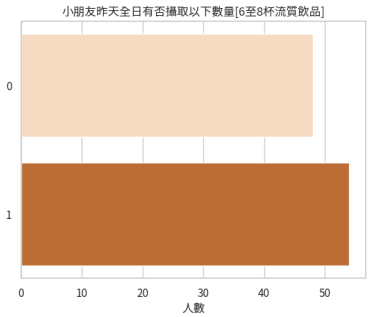
Graph 3.4 Whether the Children Ate at Least Two Servings of Fruits Yesterday



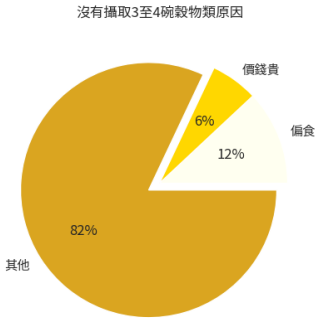
Graph 3.5 Whether the Children Ate at Least Three to Five Taels of Meat, Fish, Egg and Alternatives Yesterday



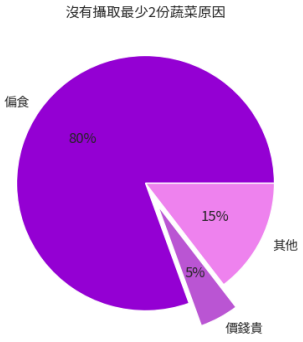
Graph 3.6 Whether the Children Ate at Least Two Servings of Milk and Alternatives Yesterday



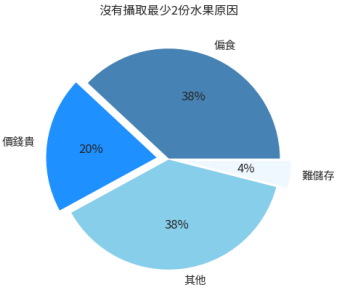
Graph 3.7 Whether the Children Drank at Least Six to Eight Bottles of Fluid Yesterday



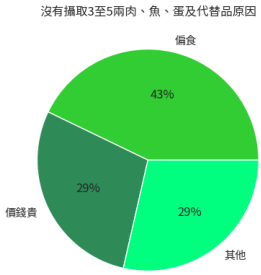
Graph 3.8 Reasons for not eating Three to Four bowls of grains



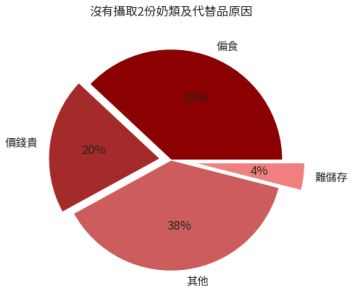
Graph 3.9 Reasons for not eating Two servings of vegetables



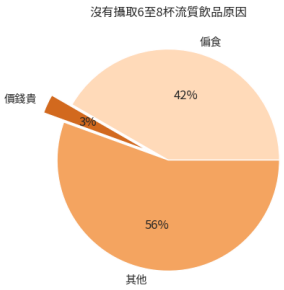
Graph 3.10 Reasons for not eating Two servings of fruit



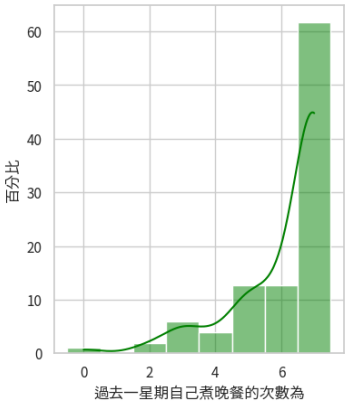
Graph 3.11 Reasons for not eating Three to Five taels of meat, fish, egg and alternatives



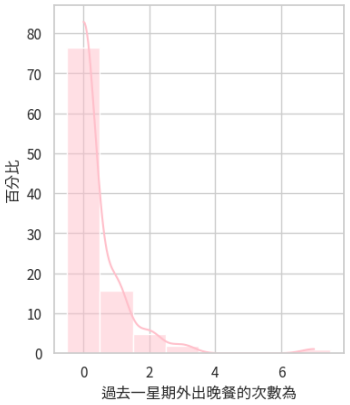
Graph 3.12 Reasons for not eating Two servings of milk and alternatives



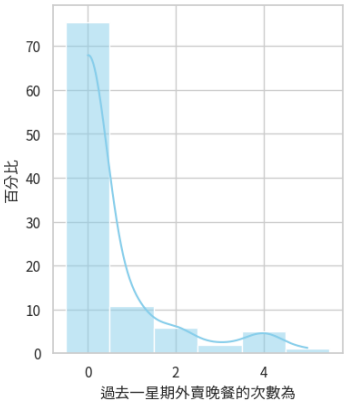
Graph 3.13 Reasons for not eating Six to Eight glasses of fluid



Graph 3.14 The number of cooking at home in a week



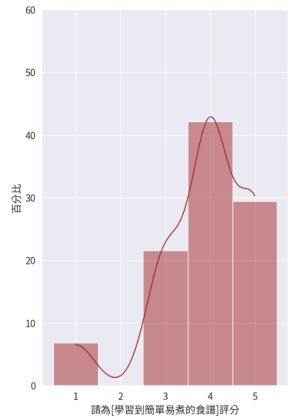
Graph 3.15 The number of eating at a restaurant in a week



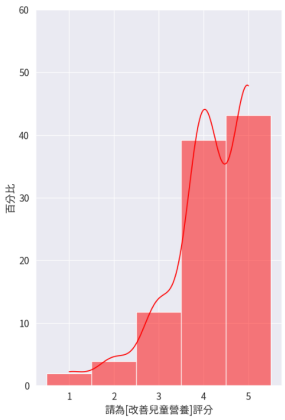
Graph 3.16 The number of buying takeaways in a week

2.4 Ways to Increase the Willingness to Cooking at home Data

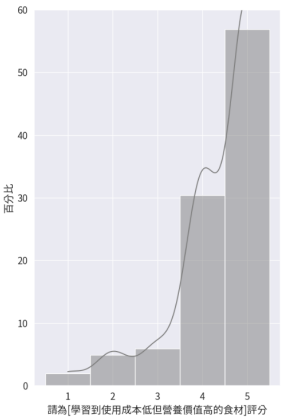
More family carers agree that the desire to cook at home can be stoked by learning about inexpensive but nutritious ingredients.



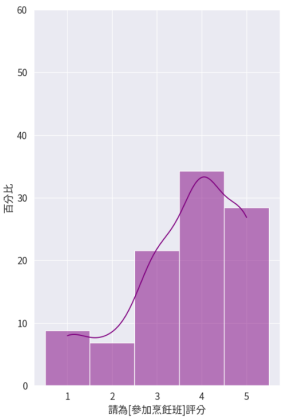
Graph 4.1 Scale of "Learning Easy Recipes" can increase the willingness to cook at home



Graph 4.2 Scale of "Improving Child Nutrition" can increase the willingness to cook at home



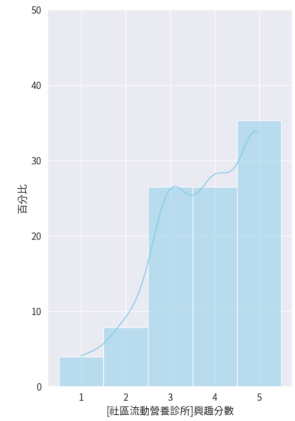
Graph 4.3 Scale of "Learning about Inexpensive but Nutritious Ingredients" can increase the willingness to cook at home



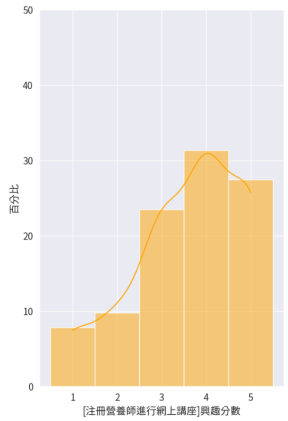
Graph 4.4 Scale of "Attending Cooking Class" can increase the willingness to cook at home

2.5 Willingness on Attending Activities Data

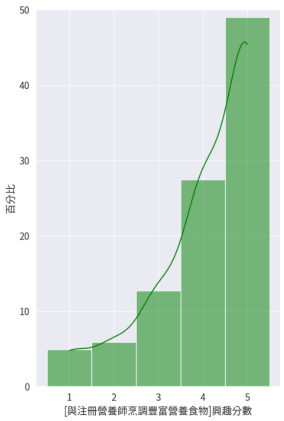
More children and families are willing to visit the game booth.



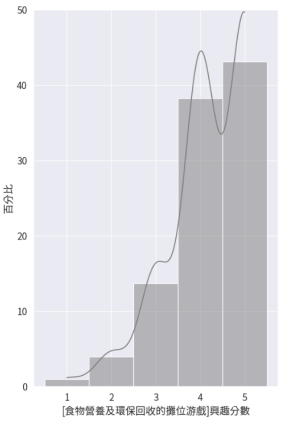
Graph 5.1 Willingness on attending the "Mobile Nutrition Clinic"



Graph 5.2 Willingness on attending the "Presented online by a registered nutritionist"



Graph 5.3 Willingness on attending the "Cooking class with a registered nutritionist"



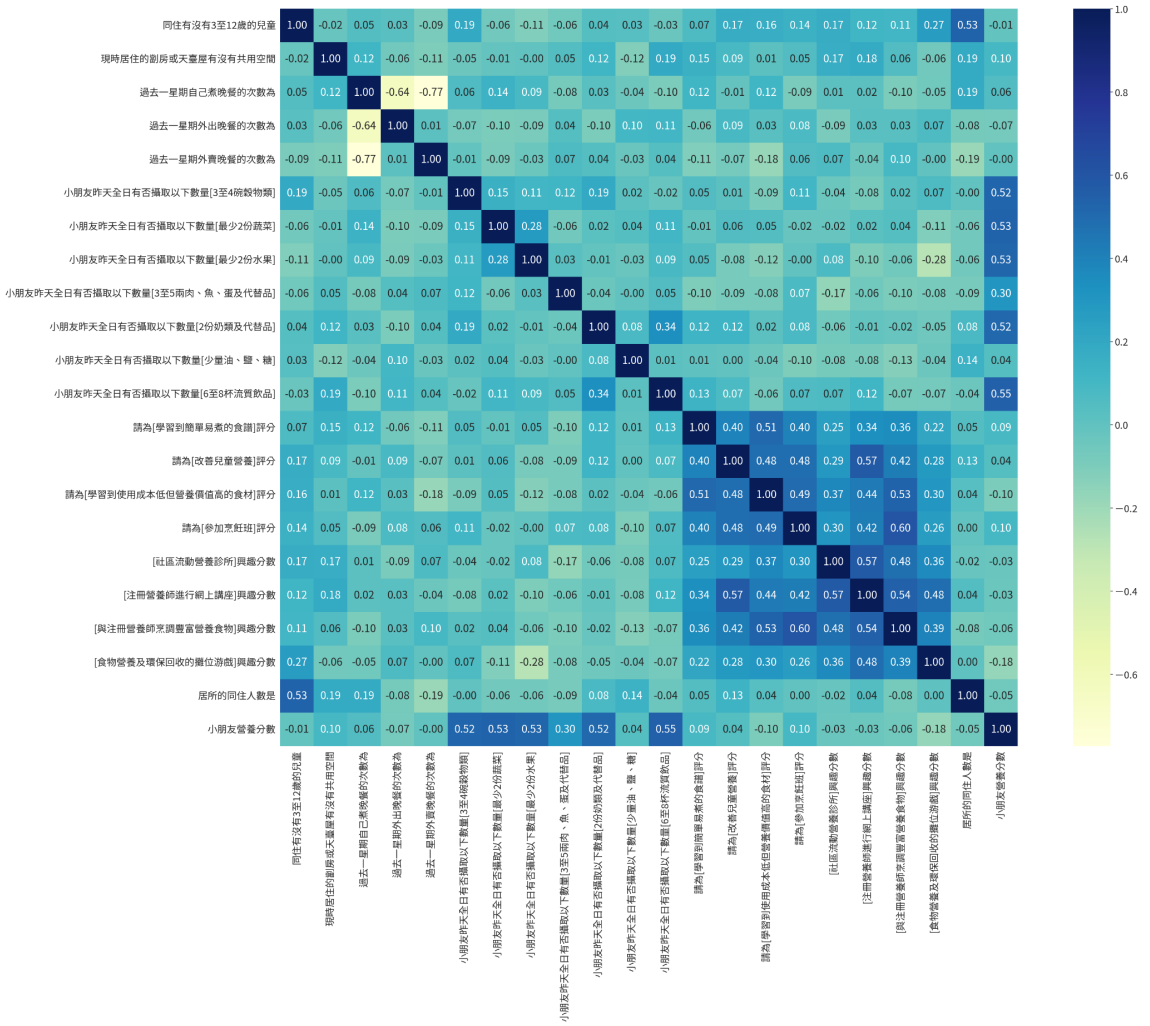
Graph 5.4 Willingness on attending the "Game booth related to food nutrition and recycle"

2.6 Relationships between Data

2.6.1 Heat Map

The heat map is a data visualization technique that shows the magnitude of a phenomenon as colour in two questions from the questionnaire.

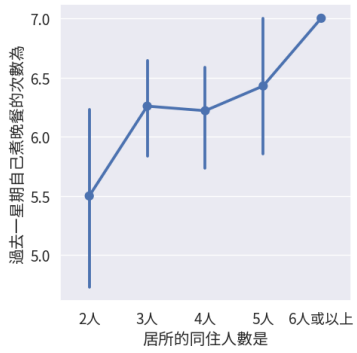
The plots and further analyses of the data from two questions with magnitudes greater than 0.2 are provided below.



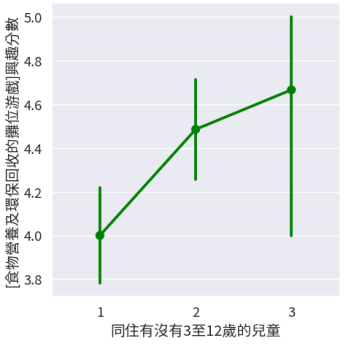
Graph 6.1 Heat Map

2.6.2 Further Investigations on the Relationship of Two Assigned Data

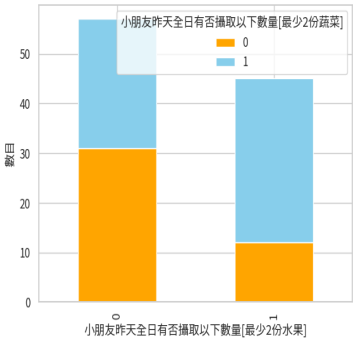
In a week, more family members may cook at home more frequently(Graph 6.2). The food nutrition and recycling-related game booth are more likely to attract families with children ranging in age from 3 to 12 years old(Graph 6.3). The children who ate at least two servings of vegetables are more likely to also eat at least two servings of fruits yesterday(Graph 6.4).



Graph 6.2 Relationship between "The Frequency of Cooking at Home in Last Week" and "The Number of Family Members"

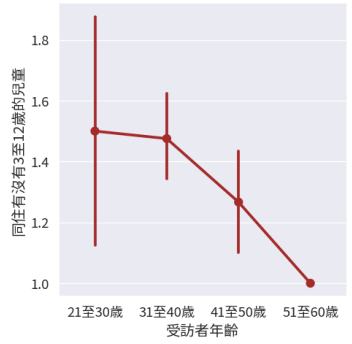


Graph 6.3 Relationship between "Scale of Willingness of Attending the Food Nutrition and Recycle Related Game Booth" and "The Number of children aged between 3 to 12 years old"

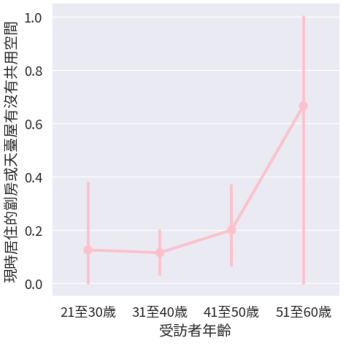


Graph 6.4 Relationship between "Whether the Children Ate at Least Two Servings of Vegetables Yesterday" and "Whether the Children Ate at Least Two Servings of Fruits Yesterday"

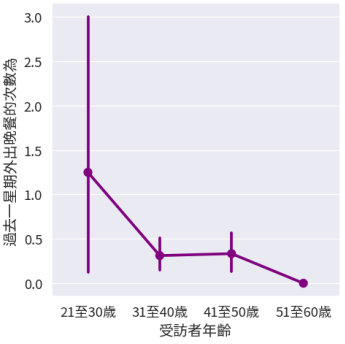
The younger respondents are more likely to have a larger number of children in their families, ranging in age from 3 to 12 years old(Graph 6.5). The older respondents are more likely to live in subdivided apartments that share a common area with other families(Graph 6.6). The younger respondents are more likely to dine out for dinner(Graph 6.7). According to Graph 6.6, older respondents are more likely to agree that learning a simple recipe could encourage them to cook at home.



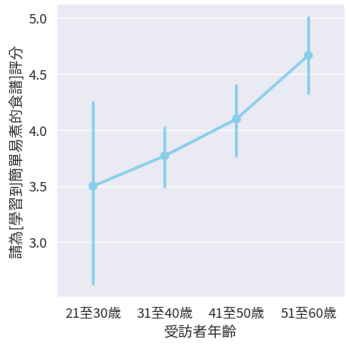
Graph 6.5 Relationship between "The Number of children aged between 3 to 12 years old" and "The Age"



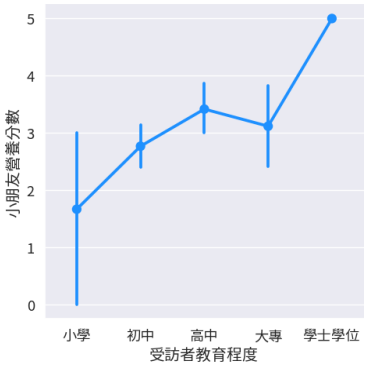
Graph 6.6 Relationship between "Existence of common area" and "The Age"



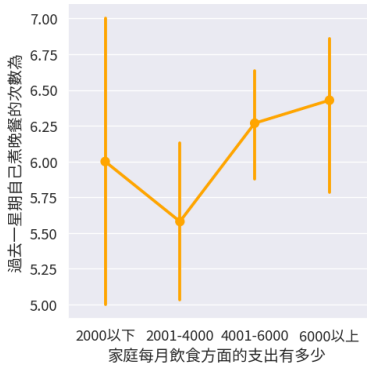
Graph 6.7 Relationship between "The number of cooking at home in a week" and "The Age"



Graph 6.8 Relationship between "Learning Easy Recipes" and "The Age"



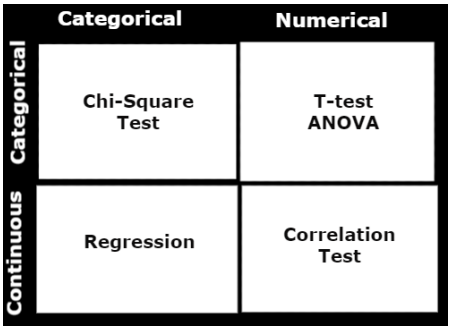
Graph 6.9 Relationship between "Nutrition Score of Children" and "The Highest Education"



Graph 6.10 Relationship between "The number of cooking at home in a week" and "Monthly Family Food Expenses"

Part 3 Hypothesis Testing

Hypothesis testing is a statistical calculation that tests an assumption about a population parameter. The purpose of the analysis and the nature of the data used determine the methodology.



Graph 7.1 Statistical Tests towards different types of data

We will use the Chi-Square Test to test our hypotheses because the majority of our questionnaire data are categorical. The most important aspects of the child's nutrition score are the primary focus of this report.

Chi-Square Test

$$\chi^2 = \sum \frac{(O - E)^2}{E} \sim \chi^2_{(r-1)(c-1)}$$

Graph 7.2 The test statistic of Chi-Square

A Chi-Square ($\chi 2$) can be used to measure the difference between the actual observation of a set of attributes and the expected frequencies of the outcomes.

Test of independence

The test of independence is used to decide if is there any relationship that exists between two attributes of our data. It is useful to analyze our survey results of 2 categorical variables.
Assumptions:
- The null hypothesis H0 = The two categorical variables have no relationship
- The alternative hypothesis H1 = There is a relationship between two categorical variables
- $\alpha = 0.05$

We will use the Chi-Square ($\chi 2$) and α to determine whether the null hypothesis will be rejected or not. To get the expected frequencies of the outcomes we assume our test data are independent first.

Assumed independent:

$$\mathbf{E_{ij}} = (R_i \cap C_j) = P(R_i) \times P(C_j) \times \text{Grand Total}$$
$$= \frac{R_i \text{ Total}}{\text{Grand Total}} \times \frac{C_j \text{ Total}}{\text{Grand Total}} \times \text{Grand Total}$$
$$= \frac{\mathbf{R_i \text{ Total} \times C_j \text{ Total}}}{\mathbf{\text{Grand Total}}}$$

Graph 7.3 Assumed the test data are independent

3.1 Food expenses versus the number of family members (Example test)

H0 = No relationship between Food expenses and the number of family members.
H1 = There is a relationship between Food expenses and the number of family members.

居所的同住人數是 家庭每月飲食方面的支出有多少	2人	3人	4人	5人	6人或以上	Total
2000以下	0	1	0	0	1	2
2001-4000	13	11	6	1	0	31
4001-6000	9	12	19	1	0	41
6000以上	0	7	16	5	0	28
Total	22	31	41	7	1	102

Graph 7.4 Data preprocessing for the independence test 3.1

```
chiTest(data['家庭每月飲食方面的支出有多少'],data['居所的同住人數是'])
✓ 0.1s
The p-value approach to hypothesis testing in the decision rule
chisquare-score is: 76.61283553950395 and p value is: 1.8189560968551177e-11
Null Hypothesis is rejected.
```

Graph 7.5 The independence test 3.1 result

Therefore, the null hypothesis was rejected, there is a relationship between Food expenses and the number of family members.

3.2 Nutrition score of the child versus the Respondent's education level

H0 = No strong relationship between the Nutrition score of the child and the Respondent's education level.
H1 = There is a strong relationship between the Nutrition score of the child and the Respondent's education level

小朋友營養分數 受訪者教育程度	0	1	2	3	4	5	6	Total
初中	2	3	13	15	6	3	1	43
大專	0	3	3	5	2	3	1	17
學士學位	0	0	0	0	0	1	0	1
小學	1	0	1	1	0	0	0	3
未受教育/學前教育	0	0	1	0	0	1	0	2
高中	0	3	8	6	11	6	2	36
Total	3	9	26	27	19	14	4	102

Graph 7.6 Data preprocessing for the independence test 3.2

```
chiTest(data['受訪者教育程度'],data['小朋友營養分數'])
✓ 0.2s
The p-value approach to hypothesis testing in the decision rule
chisquare-score is: 34.12491400435096 and p value is: 0.2758663725165147
Failed to reject the null hypothesis.
```

Graph 7.7 The independence test 3.2 result

Therefore, there is no strong relationship between the Nutrition score of the child versus the Respondent's education level.

3.3 Nutrition score of the child versus the Family's monthly income

H0 = No strong relationship between the Nutrition score of the child and the Family's monthly income.

H1 = There is a strong relationship between the Nutrition score of the child and the Family's monthly income.

小朋友營養分數 家庭每月收入是	0	1	2	3	4	5	6	Total
0-2000	0	0	1	3	2	0	0	6
10001-15000	0	2	6	9	9	5	0	31
15001-20000	1	2	6	5	2	5	1	22
20000以上	1	4	7	4	2	1	2	21
2001-5000	1	1	2	0	0	0	1	5
5001-10000	0	0	4	3	3	2	0	12
不願透露	0	0	0	3	1	1	0	5
Total	3	9	26	27	19	14	4	102

Graph 7.8 Data preprocessing for the independence test 3.3

```
chiTest(data['家庭每月收入是'],data['小朋友營養分數'])
✓ 0.8s
The p-value approach to hypothesis testing in the decision rule
chisquare-score is: 38.092028437153346 and p value is: 0.3743985170547956 5
Failed to reject the null hypothesis.
```

Graph 7.9 The independence test 3.3 result

Therefore, there is no relationship between the Nutrition score of the child versus the Family's monthly income.

3.4 Nutrition score of the child versus the Family food monthly expenses

H0 = No strong relationship between the Nutrition score of the child and the Family food monthly expenses.

H1 = There is a strong relationship between the Nutrition score of the child and the Family food monthly expenses.

小朋友營養分數 家庭每月飲食方面的支出有多少	0	1	2	3	4	5	6	Total
2000以下	0	0	0	1	0	0	1	2
2001-4000	1	5	9	6	6	3	1	31
4001-6000	0	2	12	13	6	6	2	41
6000以上	2	2	5	7	7	5	0	28
Total	3	9	26	27	19	14	4	102

Graph 7.10 Data preprocessing for the independence test 3.4

```
chiTest(data['家庭每月飲食方面的支出有多少'],data['小朋友營養分數'])
✓ 0.6s
The p-value approach to hypothesis testing in the decision rule
chisquare-score is: 23.230007694973708 and p value is: 0.1819054730527253
Failed to reject the null hypothesis.
```

Graph 7.11 The independence test 3.4 result

Therefore, there is no strong relationship between the Nutrition score of the child versus the Family food monthly expenses.

3.5 Conclusion of Chi-Square Test

In our observations, there is no element that has a direct impact on the child's diet.

Part 4 Clustering

Clustering is, by definition, one of the machine learning-based data mining methods. It is able to group our data according to similar properties. In this report, we classify the families using clustering and investigate the families' preferences for the various events, which can assist the centre in effectively maintaining the various events. We will use those demographic data for the clustering in this report.

4.1 Data Preprocessing

One Hot Encode → Convert each categorical data into a new categorical column and assign a binary value of 1 or 0

Scaling → Rescale the numerical data that prevent models from getting biased by particular attributes

	家庭每月飲食方面的支出有多少_2000以下	家庭每月飲食方面的支出有多少_2001-4000	家庭每月飲食方面的支出有多少_4001-6000	家庭每月飲食方面的支出有多少_6000以上	受訪者性別_女	受訪者性別_男	受訪者年齡_21至30歲	受訪者年齡_31至40歲	受訪者年齡_41至50歲	受訪者年齡_51至60歲	...	家庭每月收入是_15001-20000	家庭每月收入是_20000以上	家庭每月收入是_2001-5000	家庭每月收入是_5001-10000	家庭每月收入是_不願透露	同住有沒有3至12歲的兒童	現時居住的劏房或天臺屋有沒有共用空間	過去一星期自己煮晚餐的次數為	過去一星期外出晚餐的次數為	過去一星期外賣晚餐的次數為
0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	...	0.0	0.0	0.0	0.0	1.0	0.5	0.0	0.428571	0.000000	0.8
1	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	...	1.0	0.0	0.0	0.0	0.0	0.5	1.0	1.000000	0.000000	0.0
2	0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	1.0	0.0	0.0	1.000000	0.000000	0.0
3	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.714286	0.142857	0.2
4	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	...	0.0	1.0	0.0	0.0	0.0	0.5	0.0	0.714286	0.285714	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

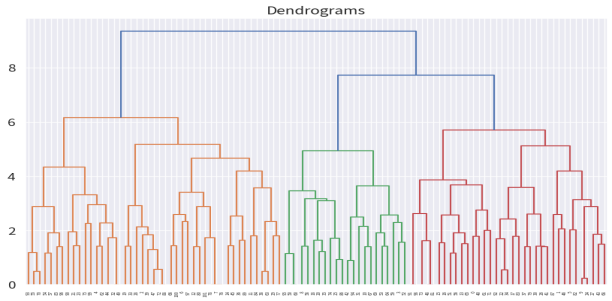
Graph 8.1 The processed demographic data

4.2 Agglomerative Hierarchical clustering

This is the most common type of unsupervised learning used to group objects in clusters based on their similarity. It is a “bottom-up” approach, which means each observation starts in its cluster, and pairs of clusters are merged with the nearest distance.

4.3 Dendrogram

A Dendrogram is a tree diagram that shows the hierarchical relationships between different attributes of data.



Graph 8.2 Sample of Dendrogram Tree Diagram

In our case, we can observe that **3 clusters** will be efficient to classify.

4.4 Implementation

HierCluster	受訪者性別	受訪者年齡	受訪者教育程度	居所的同住人數是	家庭每月收入是	同住有沒有3至12歲的兒童	家庭每月飲食方面的支出有多少
1	女	31至40歲	初中	4人	20000以上	2	4001-6000
2	女	31至40歲	初中	3人	10001-15000	1	4001-6000
3	女	31至40歲	初中	2人	10001-15000	1	2001-4000

Graph 8.3 The Demographic Data by Clustering

```
HierCluster
1    45
2    35
3    22
Name: HierCluster, dtype: int64
```

Graph 8.4 The Number of Data in Different Clusters

HierCluster	[社區流動營養診所]興趣分數	[注冊營養師進行網上講座]興趣分數	[與注冊營養師烹調豐富營養食物]興趣分數	[食物營養及環保回收的攤位游戲]興趣分數
1	3.777778	3.733333	4.000000	4.333333
2	3.885714	3.371429	4.142857	4.028571
3	3.772727	3.727273	4.227273	4.136364

Graph 8.5 The Willingness to Attend the Event by Mean

HierCluster	請為[學習到簡單易煮的食譜]評分	請為[改善兒童營養]評分	請為[學習到使用成本低但營養價值高的食材]評分	請為[參加烹飪班]評分
1	3.933333	4.288889	4.444444	3.666667
2	3.828571	4.142857	4.285714	3.628571
3	3.818182	4.000000	4.272727	3.727273

Graph 8.6 The Ways to Increase the Willingness to Cook at Home Mean

4.5 Conclusion of Clustering

In most cases, there are three groups: families with 2, 3, or 4 members. Four-person families are more likely to join the game booth. Furthermore, families with three or fewer members are more open to cooking with the nutritionist. On the other hand, all of the groups believe that cooking is most tempting by learning how to use inexpensive but nutritious ingredients.

Part 5 Main Findings and Suggestions

5.1.1 A partial eclipse and unbalanced diet

The children in the survey have very different diets than the types and amounts of food that the Department of Health recommends for children. 64% of children do not consume three to four bowls of grains daily. The daily diet still includes a greater proportion of animal products, including grains, meat, poultry, fish, and eggs. The average amount of food consumed at each meal is inconsequential and small. The main reason is that kids don't eat much. The partial eclipse is more common, and the majority of kids don't like vegetables and fruits.

5.1.2 Low milk intake, less fruit consumption

Milk consumption and daily milk consumption by children are low. Only about 30% of children consume two servings of milk per day. The food pyramid's recommended daily intake of fruits is only met by less than half of children. Additionally, fruits are very rarely consumed, and vegetables are the primary food. Only 40% of kids are able to eat at least two servings of fruits each day. However, at least two servings of vegetables can be eaten by nearly 60% of children. Additionally, children were more likely to consume two servings of vegetables when they consumed at least two servings of fruit per day.

5.1.3 Insufficient drinking water, educational level related to nutrition score

Most children fail to drink six to eight glasses of water a day, and many children do not have the habit of actively drinking water. Most of them think of drinking water only after being urged by their parents after returning home. Children whose caregivers had earned more education had better nutritional scores, and their diets were generally well-balanced.

5.1.4 Simplified daily dining places, cooking at home accounts for a large proportion

The majority of families surveyed have primary caregivers who are not employed and are able to cook for the children at home. Impacted by the plague, certain individuals will decide to purchase takeaways and afterwards return home to eat. About 0.5 times per week is the average number of takeout orders for dinner. Only a small percentage of people eat out for dinner, with an average of 0.4 per week.

5.1.5 Children's nutrition score has no obvious relationship with all the factors involved in the survey

Children's nutritional scores did not clearly relate to any of the survey's factors, according to Chi-square hypothesis tests. The chart's visual analysis reveals a weakly positive correlation between children's nutrition scores and their education. The most probable justification for this error is the little example size.

5.2 Suggestions

5.2.1 Personal - Optimize food consumption structure

People should eat more traditional coarse grains, such as millet, corn, oats, and buckwheat, in moderation, avoid being picky eaters and eat a wider variety of foods. Eggs and pork, both of which are high in fat, should be limited in consumption. Instead, they can appropriately increase their intake of poultry and fish products. Soy products can provide protein and unsaturated fat to those who are concerned about the fat in meat and fish. Individuals can intentionally pick soy food or soy items for example soy milk when they eat out. It is

recommended to insist on drinking milk at least once per day. People who live in better economic circumstances and have access to refrigerated conditions can purchase milk or milk powder as a substitute. Last but not least, for a healthy and balanced diet, individuals should consume two servings of fruits and vegetables per day.

5.2.2 Personal - Established correct eating habits

Parents should be persistent in selecting nutritious foods. They should also not allow children to eat high-salt, high-fat, and high-calorie foods for convenience or to please them, such as by avoiding unhealthy snacks. Because the children's growth model is concerned with the family's living habits, parents must insist on instilling good habits in their children and acting as role models.

5.2.3 Personal - Learn more about healthy eating

Scientific research should be used to guide daily diet. Learning from authoritative books on healthy eating, such as "Dietary Guidelines for Chinese Residents," is one method. People frequently use the Internet to inquire about new knowledge and achievements in dietary health by visiting the official websites of some nutrition and health education institutions.

5.2.4 Organization - Host booth game day

Based on the clustering results, it is suggested that the first group (a family of four members) hold a booth game day based on their interests. They can participate as participants or as organizers to gain experience in various roles. As participants, parents and children are taught the benefits of a balanced diet in order to cultivate healthy lifestyle habits; while organizers can use this opportunity as volunteer work and learn from the process of preparing material for the game booth. Finally, parents can spend their family day with their children in the booth game day and learn through play.

5.2.5 Organization - Host cooking classes

As a result of the clustering, the majority of groups are more interested in taking cooking classes with registered dietitians. Hosting cooking classes for these families is an excellent way to promote a healthy diet. Cooking together can also be a great way for the family to spend time together. The entire family can enjoy the results, which is heartwarming. Sometimes the recipes contain ingredients that children dislike, but they usually believe that the dishes they make with their hands are delicious. Parents can use this to encourage their children to try new foods and improve their eating habits. At the same time, the family can learn how to utilize cheaper but higher nutrition value food. It also has the highest factor score in increasing the respondents' desire to cook. Every class can help children gain a better understanding of food and nutrition. Children learn to cherish and be grateful while experiencing and feeling.

5.2.6 Society - Create a supportive environment

Environmental factors have a significant impact on people's health. To promote health, we must begin with daily habits and living environments in order to create a safe, satisfying, comfortable, and interesting living environment in order to help people live healthier and happier lives. To create a healthy food culture, schools can develop a healthy eating policy that prohibits the sale, and promotion of drinks and snacks that do not meet health guidelines. Also, they can encourage tuckshops to provide students with fruits and other

healthy snacks at discounted prices. As a result, schoolchildren were more likely to eat healthily.

5.2.7 Society - Promote a balanced diet

People will be more effective in maintaining their health and their environment if they have access to information, health education, and the development of personal life skills. The Department of Health, for example, can actively promote a balanced diet through its website, leaflets, radio and television advertisements, and so on. It can also take the lead in encouraging citizens to consume at least two servings of fruit and three servings of vegetables per day, as well as suggestions on how to incorporate these points into daily life.

5.3 Limitation

5.3,1 Limitation - Limited Responses

The questionnaire frequently fails to capture respondents who are busy, lazy, or indifferent. Although they are rarely identified, these are the people who make up a significant portion of the respondents to be gathered in data collection. Meanwhile, the limitations of the questions set (i.e. too less questions, not specific enough) bring the problem of insufficient data. As a result, with a different method of analysis, relationships that aren't obvious or even contradictions result from insufficient data. The data may not be adequately explained by the clustering.

5.3.2 Unreliability

The information gathered through the questionnaire cannot be considered reliable or valid. There is little that can be done to connect a response if the subject misinterprets a question or gives an incomplete or indefinite response. Furthermore, human errors would occur due to the limited amount of time. For example, the interviewer may incorrectly mark answers and manipulate entries by asking leading questions.

Part 6 Conclusion

Personal living habits, health awareness, economic conditions, and the objective environment are all factors that influence eating habits. Although the impact of the respondents' diet on their current health is not obvious, there are still hidden dangers of many chronic diseases developing in the future. As a result, gradually improving their eating habits and eating conditions through intervention measures is a systematic and long-term endeavor that necessitates the collaboration of individuals, relevant organizations, and society. And it is only through perseverance that we will be able to continuously improve the dietary health awareness and health level of subdivided families.

Appendix

Dendrogram:

```
import scipy.cluster.hierarchy as shc
from sklearn.preprocessing import MinMaxScaler
X=data[['家庭每月飲食方面的支出有多少','受訪者性別', '受訪者年齡', '受訪者教育程度', '居所的同住人數是', '家庭每月收入是']]
X=pd.get_dummies(X)
Y =data[['同住有沒有3至12歲的兒童', '現時居住的劏房或天臺屋有沒有共用空間',
       '過去一星期自己煮晚餐的次數為', '過去一星期外出晚餐的次數為', '過去一星期外賣晚餐的次數為',]]
G = pd.concat([X,Y],axis=1)
MinMax = MinMaxScaler()
G_std = pd.DataFrame(MinMax.fit_transform(G),columns = G.columns)
plt.figure(figsize=(10, 7))
plt.title("Dendrograms")
dend = shc.dendrogram(shc.linkage(G_std, method='ward'))
```

Clustering:

```
from sklearn.cluster import AgglomerativeClustering
cluster = AgglomerativeClustering(n_clusters=3, affinity='euclidean', linkage='ward')
data['HierCluster'] = cluster.fit_predict(G_std)
data['HierCluster'] = data['HierCluster'] + 1
```

Chi-Square Test

```
def chiTest(df1,df2):
    data_crosstab = pd.crosstab(df1,
                                df2,
                                margins=True, margins_name="Total")
    alpha = 0.05
    chi_square = 0
    rows = df1.unique()
    columns = df2.unique()
    for i in columns:
        for j in rows:
            O = data_crosstab[i][j]
            E = data_crosstab[i]['Total'] * data_crosstab['Total'][j] / data_crosstab['Total']['Total']
            chi_square += (O-E)**2/E

    print("The p-value approach to hypothesis testing in the decision rule")
    p_value = 1 - stats.chi2.cdf(chi_square, (len(rows)-1)*(len(columns)-1))
    conclusion = "Failed to reject the null hypothesis."
    if p_value <= alpha:
        conclusion = "Null Hypothesis is rejected."

    print("chisquare-score is:", chi_square, "and p value is:", p_value)
    print(conclusion)
```