STAT3613 Marketing Engineering

Project

To launch a new off-campus restaurant near The University of Hong Kong

Name	UID	Contribution
Ding Ding	3035334670	33.3%
Ke Daqi	3035234600	33.3%
Zhang Xinyi	3035234571	33.3%

1. Introduction

Every student at The University of Hong Kong (HKU) hopes for an ideal restaurant that they can run for after class. After the renovation of HKU campus, this semester there are several new dining places. However, it is rumored that students are not satisfied with them in terms of flavor, price or cost performance, etc. This motivates us to explore the opportunity of launching an off-campus restaurant, which hopefully can meet the requirements of the majority of our students. The objective is to discuss the possible and relative importance of the attributes of a restaurant targeting HKU students, identify the optimal setup of the new restaurant for each attributes that maximize the gross profit of the restaurant. By studying the characteristics of the sampled customers, we find out our target customer cluster and makes an estimation of total gross profit accordingly. Thus, the best hypothetical new restaurant and it's estimated total gross profit can serve as the baseline or starting point for further marketing research.

2. Method

2.1 Target population and sample size

Corresponding to the objective of this project, the project mainly targets at HKU students. According to the official figure, the total number enrollment of students in all programmes at HKU is 28744 by November 2017. Initially, the survey aims to obtain 300 responses which would be around 1 percent of the whole population. However, due to time limit and low response rate, 53 and 20 students respond to our first and second questionnaires respectively.

2.2 Sampling method

Convenient sample method will be adopted for practical reasons. Although simple random sample and stratified sample are more likely to provide unbiased and representative results, they are time-consuming. Given that this is only a hypothetical "new restaurant", it is not worthwhile spending the excessive resources, especially when reasonably good results can be achieved by convenient sample method.

2.3 Data source and data collection method

Since the target population of this project is HKU students, using questionnaires to collect data is feasible and reliable. It also gives us the flexibility to design variables needed for analysis. To reduce workload and increase the sample size, we designed relatively short questionnaires and asked students to fill them in. The questionnaires were distributed through WeChat HKU student group, WhatsApp HKU student group, and Facebook HKU student group to reach the majority of the target population. The first questionnaire was distributed on 27 November, 2018 and the deadline was set at 12:00 pm on 29 November, 2019. After we finish analyzing the results of the first questionnaire, the attributes of the hypothetical restaurant were identified. Based on the result, the second questionnaire was distributed on 29 November, 2018, and the deadline was set at 12:00pm on 1 December, 2018. The first-hand data from collected questionnaires are used in our analysis.

2.4 Sample questionnaires

2.4.1 Questionnaire 1

Link: https://goo.gl/forms/n8q457j4AGLN4Hfk1

Would you like to have meals in the following restaurants? (Waiting time= waiting time in the queue + waiting time to get your meal after ordering) *

	Yes	No
Asian food, $45/meal, 10-min waiting time	○	○
Western food, $45/meal, 10-min waiting time	○	○
Asian food, $75/meal, 10-min waiting time	○	○
Western food, $75/meal, 10-min waiting time	○	○
Asian food, $45/meal, 20-min waiting time	○	○
Western food, $45/meal, 20-min waiting time	○	○
Asian food, $75/meal, 20-min waiting time	○	○
Western food, $75/meal, 20-min waiting time	○	○

Do the restaurants have the following attributes? If yes, tick the box.

	Delicious food	Healthy & nutritious food	Pleasant environment	Short waiting time	Short distance from HKU
Yu Mai (魚米)	☐	☐	☐	☐	☐
Ninoen Japanese Shop (日の苑)	☐	☐	☐	☐	☐
Bafang Dumpling (八方雲集)	☐	☐	☐	☐	☐
King Prawn Restaurant (大頭蝦餐廳)	☐	☐	☐	☐	☐
Yoshinoya（吉野家）	☐	☐	☐	☐	☐
Mexim' MX (美心)	☐	☐	☐	☐	☐
Oppa Chicken	☐	☐	☐	☐	☐
McDonald's	☐	☐	☐	☐	☐

2.4.2 Questionnaire 2

Link: https://goo.gl/forms/E5xuRnqedVnu8v072

Based on the result of the first questionnaire (please refer to Section 3 for detailed analysis), the attributes of hypothetical restaurant are identified.

Suppose a new off-campus restaurant is launched. It sells Asian food which is nutritious (i.e. less salty, less oily, and sufficient meat and vegetables). The distance is 5-min walk from HKU and the waiting time (=waiting time in the queue + waiting time to get your meal after ordering) is 10 minutes. Will you have meals in the restaurant at the following prices? *

	Yes	No
20	○	○
25	○	○
30	○	○
35	○	○
40	○	○
45	○	○
50	○	○
55	○	○
60	○	○
65	○	○
70	○	○

Rate the importance of the attributes of a restaurant (1=least important, ..., 5=most important) *

	1 (least important)	2	3	4	5 (most important)
Nutrition (healthy or not)	○	○	○	○	○
Taste (delicious or not)	○	○	○	○	○
Environment	○	○	○	○	○
Waiting time	○	○	○	○	○
Distance from HKU	○	○	○	○	○

Gender *

- ○ Male
- ○ Female

Age *

Your answer

Faculty *

- ○ Architecture
- ○ Arts
- ○ Business and Economics
- ○ Dentistry
- ○ Education
- ○ Law
- ○ Medicine
- ○ Science
- ○ Social Science
- ○ Other: _____

Total monthly expense *

Your answer

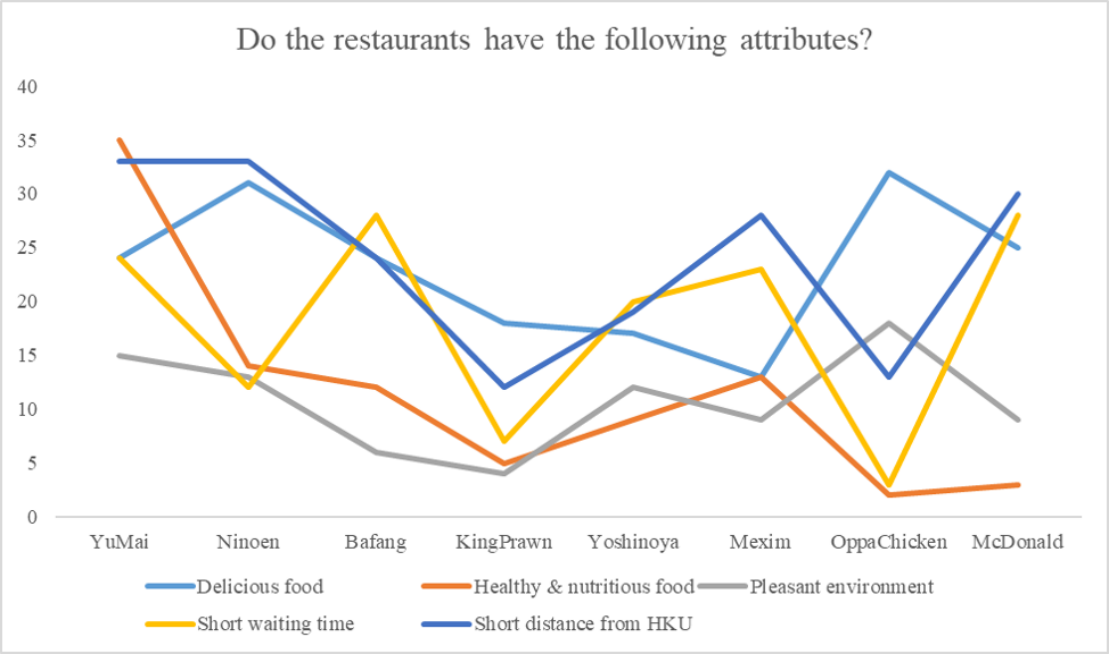
Are you an undergraduate or a postgraduate student? *

- ○ Undergraduate
- ○ Postgraduate

3. Results

3.1 Results for Questionnaire 1

3.1.1. Correspondence analysis



In this question, we presented eight existing off-campus restaurants that are quite popular among HKU students to the respondents. Also, we presented five attributes including delicious food, healthy & nutritious food, pleasant environment, short waiting time and short distance from HKU. We asked respondents to click "Yes" if they think the presented restaurant has certain attributes. Through applying correspondence analysis to the data, we can find the association between the five attributes and eight existing off-campus restaurants. By dimension reduction, we would be able to find out important attributes that can account for the positioning of existing restaurants.

After data collection, we have calculated the simple frequencies for each restaurant across the entire set of attributes as follows.

	YuMai	Ninoen	Bafang	KingPrawn	Yoshinoya	Mexim	OppaChicken	McDonald
Delicious food	24	31	24	18	17	13	32	25
Healthy & nutritious food	35	14	12	5	9	13	2	3
Pleasant environment	15	13	6	4	12	9	18	9

Short waiting time	24	12	28	7	20	23	3	28
Short distance from HKU	33	33	24	12	19	28	13	30

```
Principal inertias (eigenvalues):

 dim    value      %   cum%   scree plot
 1      0.080683  57.8  57.8  **************
 2      0.043047  30.9  88.7  ********
 3      0.009786   7.0  95.7  **
 4      0.006004   4.3 100.0  *
        -------- -----
 Total: 0.139520 100.0
```

Two dimensions can explain 88.7% of the variation and increasing to a 3-dimensional solution only provides 7% increment. So we can use correspondent analysis to reduce the dimension to 2.

For attributes:

```
Rows:
    name   mass  qlt  inr    k=1 cor ctr    k=2 cor ctr
1 | Dlcs |  263  904  229 |  332 903 358 |  -11   1   1 |
2 | Hlth |  133  992  319 | -386 444 245 |  429 548 567 |
3 | Plsn |  123  724  163 |  337 616 173 |  141 108  57 |
4 | Shrt |  207  950  242 | -285 497 208 | -272 453 355 |
-------------------------------------------------------
5 | SHKU |  274  321   46 |  -67 190  15 |  -56 131  20 |
```

As for dimension 1, delicious food, health & nutritious food and possibly short waiting time are primary contributors. 60.3% of dimension 1 is accounted for. Dimension 1 is between delicious food and healthy & nutritious food. Lots of delicious food, such as fried chicken at McDonald, are salty, oily and probably with fewer vegetables, which become a contrast with healthy & nutritious food. As for dimension 2, the primary contributors are healthy & nutritious food and short waiting time. They contribute to 92.2% of dimension 2. Dimension 2 is between healthy & nutritious food and short waiting time. People who are keen on healthy & nutritious food may be willing to spend more time waiting.

Based on the above R output, the quality values range from a high 99.2% for healthy & nutritious food to a low of 32.1% for short distance from HKU. Only short distance from HKU has a value of total contribution below 50%. Therefore, the model fits quite well.

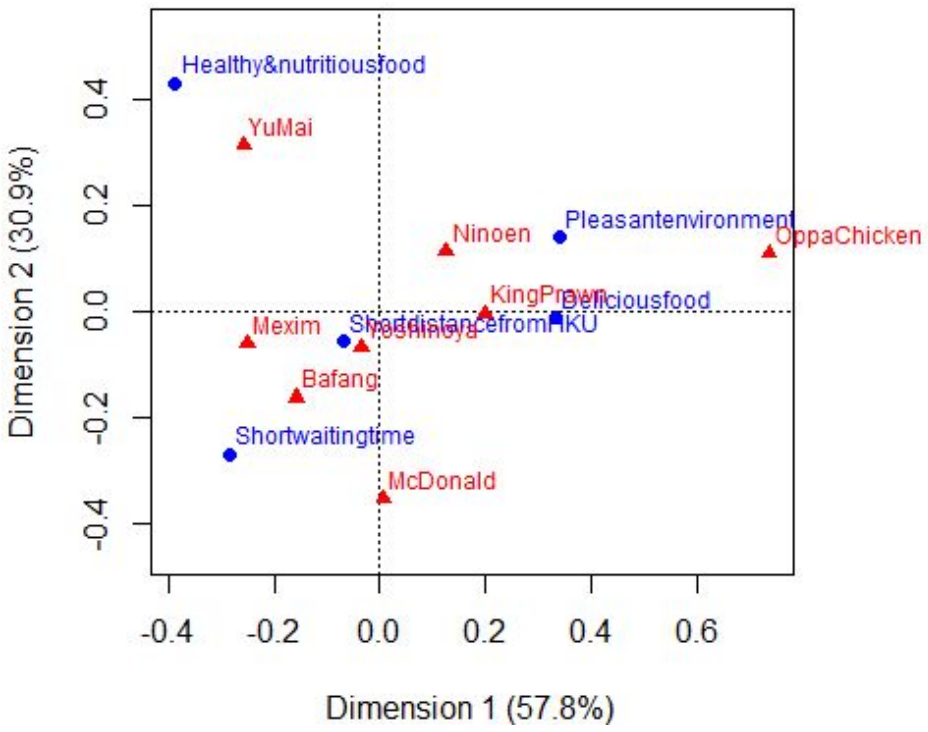
For restaurants:

```
Columns:
    name   mass  qlt  inr    k=1 cor ctr    k=2 cor ctr
1 | YuMa |  187  992  221 | -256 397 152 |  313 595 427 |
2 | Ninn |  147  531   56 |  124 291  28 |  113 240  44 |
3 | Bfng |  134  746   67 | -158 359  42 | -164 387  84 |
4 | KngP |   66  424   44 |  198 424  32 |   -5   0   0 |
5 | Yshn |  110  180   27 |  -34  35   2 |  -70 145  12 |
6 | Mexm |  123  827   70 | -249 779  95 |  -62  47  11 |
7 | OppC |   97  981  391 |  734 960 650 |  108  21  26 |
8 | McDn |  136  989  123 |    7   0   0 | -354 989 396 |
```

As for dimension 1, Oppa Chicken and Yu Mai contribute quite a lot. Thus, we can interpret dimension 1 as Oppa Chicken vs. Yu Mai. As for dimension 2, McDonald and Yu Mai contribute quite a lot. Thus, we can interpret dimension 2 as McDonald vs. Yu Mai.

The fit values range from 18% to 98.9%. Only two restaurants (King Prawn and Yoshinoya) have values of total contribution below 50%.

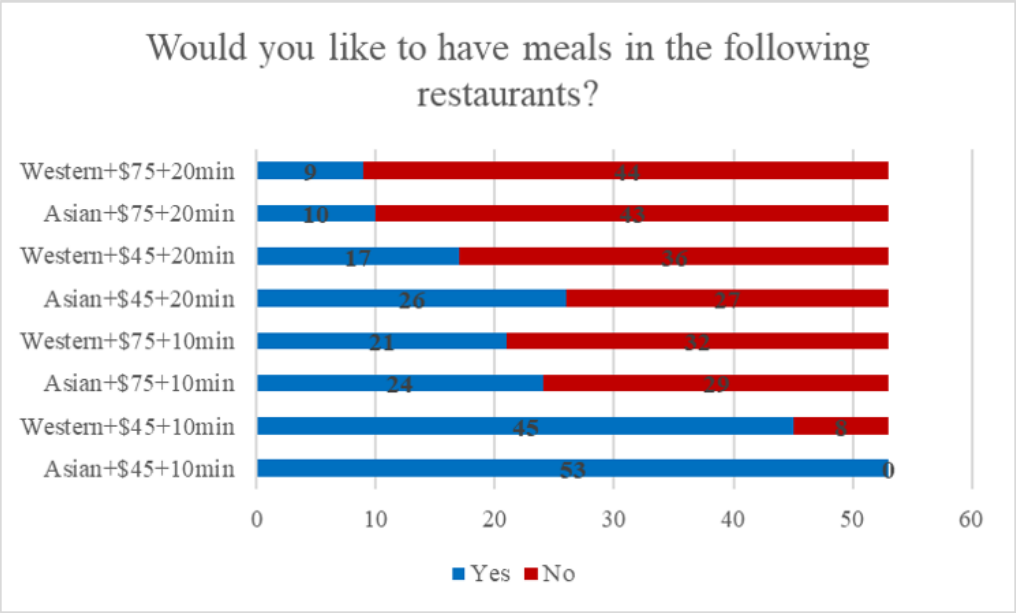
```
$rows
                             Dim1        Dim2
Deliciousfood           0.33152006 -0.01067056
Healthy&nutritiousfood -0.38589093  0.42872250
Pleasantenvironment     0.33744807  0.14149683
Shortwaitingtime       -0.28471278 -0.27167159
ShortdistancefromHKU   -0.06692195 -0.05564666

$columns
                   Dim1        Dim2
YuMai       -0.256099207  0.31329418
Ninoen       0.124295318  0.11297937
Bafang      -0.158334501 -0.16432282
KingPrawn    0.198347488 -0.00544126
Yoshinoya   -0.034454317 -0.06983414
Mexim       -0.249384897 -0.06155932
OppaChicken  0.734487614  0.10805637
McDonald     0.006957503 -0.35429761
```



Positioning of the firms is as follows:

1. Group 1: Yu Mai
2. Group 2: Ninoen, King Prawn, Oppa Chicken
3. Group 3: Mexim, Bafang, Yoshinoya
4. Group 4: McDonald

3.1.2 Logit choice model



In the question, we have used full profile method, listing all eight combinations of three attributes (each attribute have two levels) to be presented to the respondents. Respondents are required to choose "Yes" if they would like to have meals in a certain restaurant. The results are shown in the figure above. In order to study the relative importance of three different attributes (food style, price & waiting time*) as well as to construct a hypothetical restaurant with the highest utility for further analysis, we apply logit choice model to the data.

* Waiting time = waiting time in the queue + waiting time to get your meal after ordering

The frequency table of various profiles is as follows:

	Foodstyle	Price	Waitingtime	n	c	nc
1	asian	$45	10min	53	53	0
2	western	$45	10min	53	45	8
3	asian	$75	10min	53	24	29
4	western	$75	10min	53	21	32
5	asian	$45	20min	53	26	27
6	western	$45	20min	53	17	36
7	asian	$75	20min	53	10	43
8	western	$75	20min	53	9	44

For example, profile 1 (Asian + $45 per meal + 10 minutes' waiting time) was selected by 100% of the subjects. In total, there were 424 restaurants and 205 were selected. Number of events (chosen) is equal to 205 and number of nonevent (not chosen) is equal to 219.

```
Anova(logit,type="III")

## Analysis of Deviance Table (Type III tests)
##
## Response: cbind(c, nc)
##             LR Chisq Df Pr(>Chisq)
## Foodstyle      5.799  1    0.01603 *
## Price         69.791  1    < 2e-16 ***
## Waitingtime   76.144  1    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on p-value, we can conclude that all three attributes are significant at 5%.

```
contrasts(restaurant$Foodstyle)

##         western
## asian         0
## western       1

contrasts(restaurant$Price)

##     $75
## $45   0
## $75   1

contrasts(restaurant$Waitingtime)

##       20min
## 10min     0
## 20min     1
```

```
Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)        2.1408     0.2714   7.889 3.05e-15 ***
Foodstylewestern  -0.5550     0.2321  -2.391   0.0168 *
Price$75          -1.9177     0.2504  -7.658 1.89e-14 ***
Waitingtime20min  -1.9948     0.2506  -7.961 1.70e-15 ***
```

	Utility	Range	Importance

Western	-0.5550	0.5550	12.4%
Asian	0		
$75	-1.9177	1.9177	42.9%
$45	0		
20min	-1.9948	1.9948	44.7%
10min	0		
		4.4675	

Based on the above summary of logit choice model, we conclude that waiting time (10min/20min) and price ($45/$75) are the two most important factors for our students when choosing restaurants, with waiting time being slightly more important than the price. After calculating the part-worth utility of each level of each factor, we find out that the most prefered profile is a restaurant with asian food style, $45 and 10 minutes' waiting time (Utility (Asian, $45, 10min) = 2.1408), and a restaurant with western food style, $45 adn 10 minutes' waiting time (Utility (Western, $45, 10min) = 1.5858) is the second preferred profile.

Below are utility and probability of choosing for each of the eight profiles:

	Foodstyle	Price	Waitingtime	n	c	nc	utility	p
1	asian	$45	10min	53	53	0	2.1407519	0.89480141
2	western	$45	10min	53	45	8	1.5857503	0.83001736
3	asian	$75	10min	53	24	29	0.2230421	0.55553051
4	western	$75	10min	53	21	32	-0.3319595	0.41776394
5	asian	$45	20min	53	26	27	0.1459421	0.53642091
6	western	$45	20min	53	17	36	-0.4090594	0.39913768
7	asian	$75	20min	53	10	43	-1.7717676	0.14532265
8	western	$75	20min	53	9	44	-2.3267692	0.08893008

Based on the logit model result of our Question 1, Asian food style, relatively low price and short time would be three important factors that we need to take into consideration when constructing a new off-campus restaurant near HKU.

In a nutshell, according to our logit model for Question 1 and correspondence analysis for Question 2, we discover that they echo each other's result in some way in terms of the importance of the attributes.

Therefore, we want to propose a new off-campus restaurant which offers healthy & nutritious Asian food with relatively low prices, short distance from HKU and short waiting time.

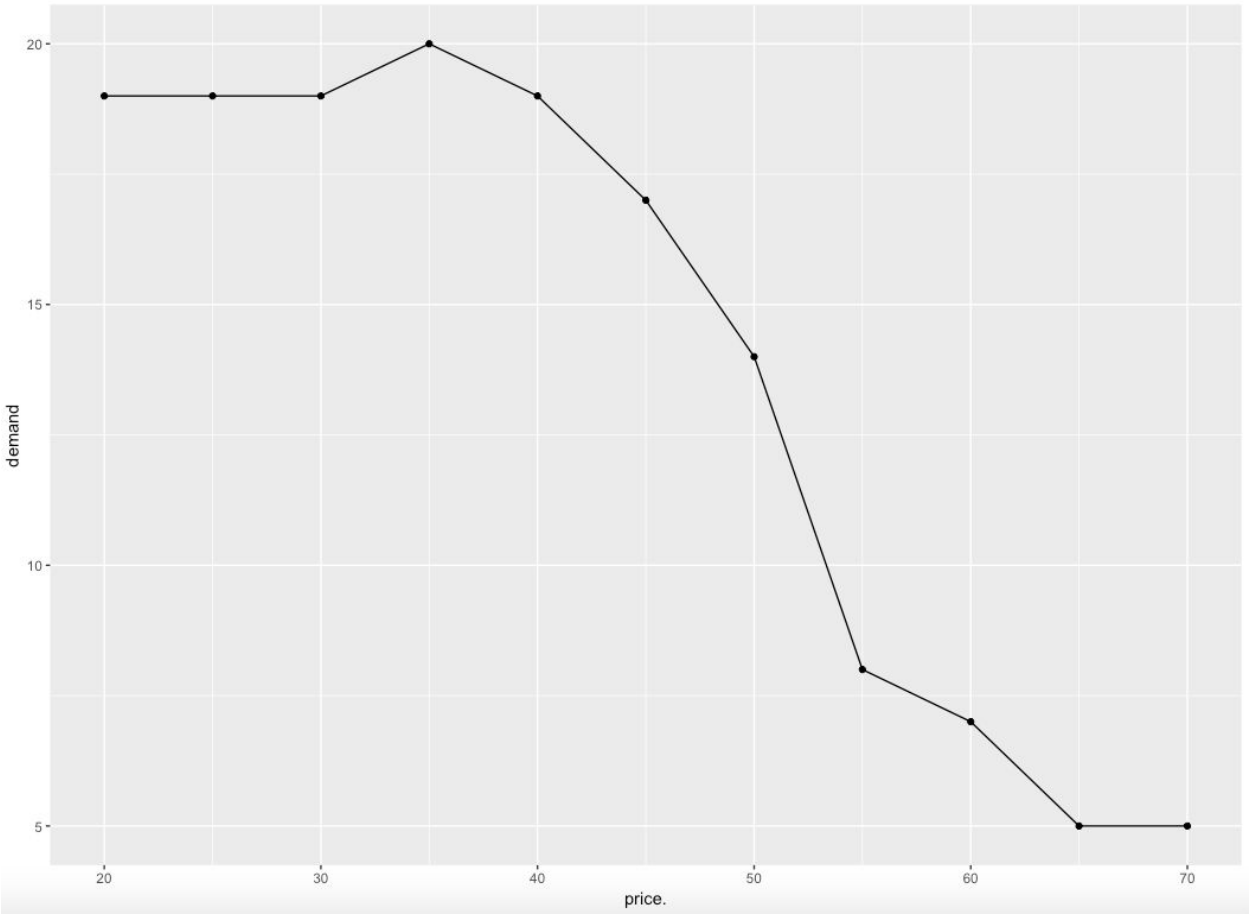
Our second step is to research the profitability of this restaurant and customer profiles. Therefore, we constructed a hypothetical restaurant with the aforementioned attributes in our Questionnaire 2 for further analysis.

3.2 Results for Questionnaire 2

3.2.1 Market response model

X	price	demand
1	20	19
2	25	19
3	30	19
4	35	20
5	40	19
6	45	17
7	50	14
8	55	8
9	60	7
10	65	5
11	70	5

In this question, students are asked to choose whether or not they will have meals in this hypothetical restaurant at certain prices. Since the result of questionnaire 1 shows that students prefer $45 to $75, the range of the choices of the price is set to be from $20 to $70. The result of this questions shows that 19 students are willing to pay at the price of $20, $25, $30, and $40, 20 students are willing to pay at the price of $35, 17 students are willing to pay at $45, etc. Generally speaking, the demand drops as the price goes up. However, we note an interesting phenomenon that when the price is too low, the demand may decrease. It is reasonable, since it may arouse people’s concern about the food quality.



To begin with, we try to find the relationship between demand and the price.

First, we try constant elasticity price model.

```
Formula: demand ~ a * price^-b

Parameters:
  Estimate Std. Error t value Pr(>|t|)
a 214.9422   160.8918   1.336  0.21436
b   0.7435     0.2126   3.497  0.00676 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.105 on 9 degrees of freedom

Number of iterations to convergence: 9
Achieved convergence tolerance: 1.073e-06
```

The parameter a is not significant, and R^2 is 0.6. Thus, we would like to try other models.

Then, we try power series model.

```
Formula: demand ~ a + b * price + c * price^2

Parameters:
   Estimate Std. Error t value Pr(>|t|)
a 16.453613   5.117845   3.215   0.0123 *
b  0.311469   0.245862   1.267   0.2408
c -0.007319   0.002699  -2.711   0.0266 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.977 on 8 degrees of freedom

Number of iterations to convergence: 1
Achieved convergence tolerance: 1.186e-08
```

The parameters a and c are significant at 5%, while b is not significant. Hence, we would like to improve this model by removing the “b*price” in the formula.

```
Formula: demand ~ a + c * price^2

Parameters:
    Estimate Std. Error t value Pr(>|t|)
a 22.7815497  1.1512440  19.789 9.97e-09 ***
c -0.0039399  0.0004276  -9.215 7.04e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.042 on 9 degrees of freedom

Number of iterations to convergence: 1
Achieved convergence tolerance: 4.718e-10
```

After the adjustment, both parameters are significant, and R^2 is 0.9. The model fits well to a certain degree.

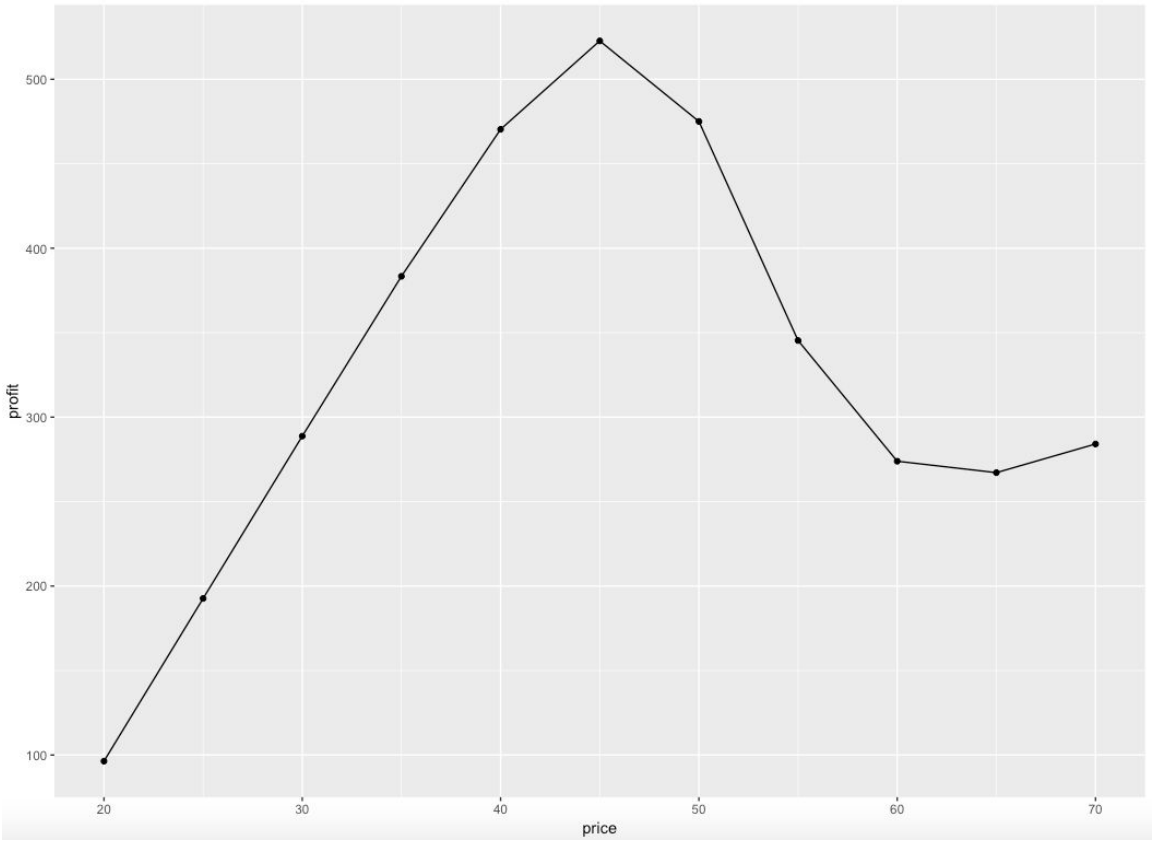
The last model we try is logistic model. Since from the plot of demand vs price, there is an S-shape.

```
Formula: demand ~ a/(1 + exp(-b - c * price)) + d

Parameters:
  Estimate Std. Error t value Pr(>|t|)
a 14.16005    0.58783  24.089 5.41e-08 ***
b 15.39972    2.28023   6.754 0.000264 ***
c -0.30010    0.04487  -6.689 0.000280 ***
d  5.11371    0.46183  11.073 1.09e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6185 on 7 degrees of freedom

Number of iterations to convergence: 5
Achieved convergence tolerance: 1.497e-06
```

All 4 parameters are very significant, and R^2 of this model is 0.99, which is higher than the power series model we tried previously.

Hence, we choose logistic model to describe the relationship between demand and price. Then we move to the next step: optimization.

Suppose the cost is $15 per meal, then the gross profit will be equal to demand*(price-cost). This cost is assumed to be constant for every meal. Besides, fixed costs for an restaurant is not considered, so that net profit will not be calculated in this project.



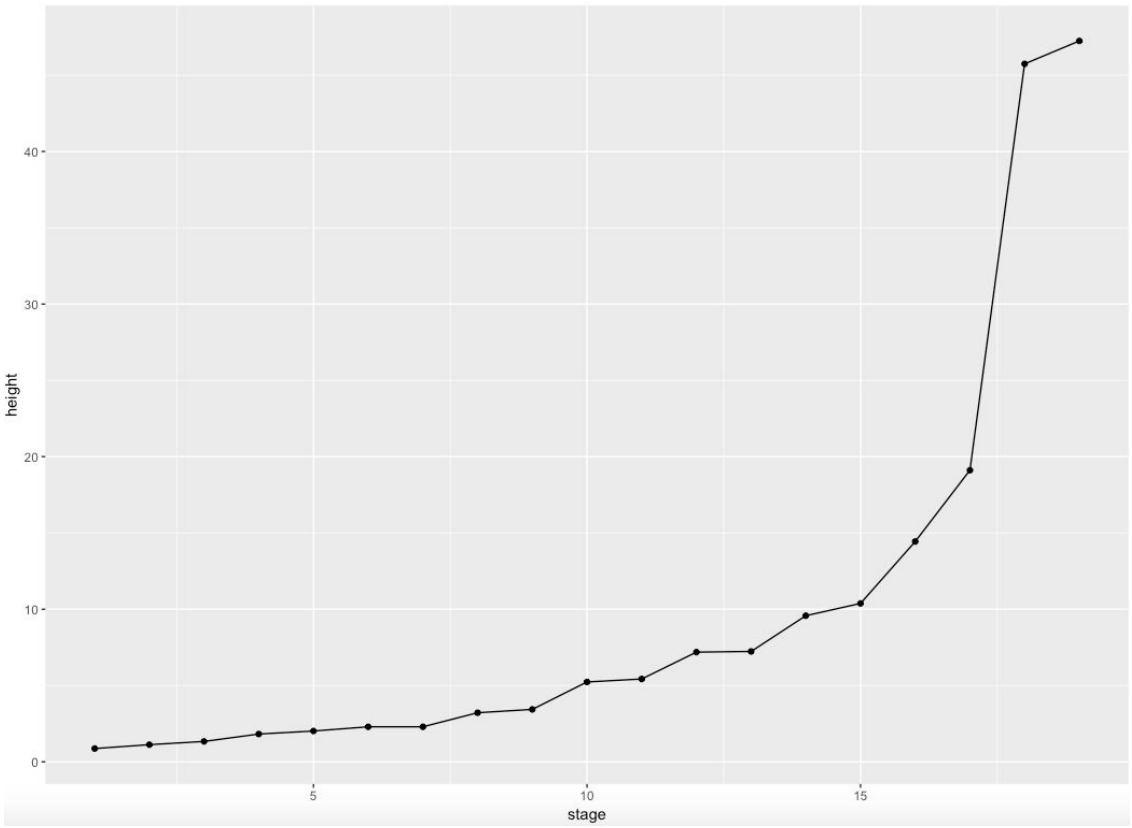

```
> op<-optimize(f=profit,fit=fit3,interval=c(20,70),maximum =TRUE)
> op
$maximum
[1] 45.66626
```

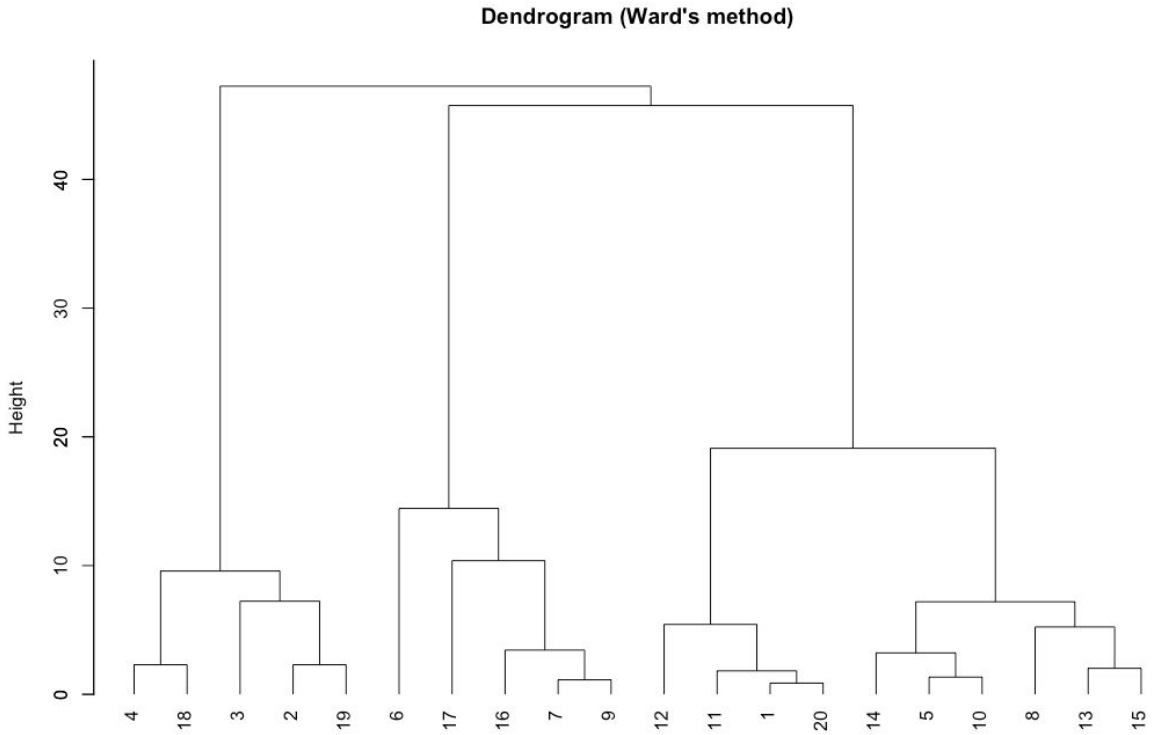
We find that when price=$45.67, the gross profit will be maximized. Rounding the number to be an integer, we may set the price to be $46. If we assume this 17/20 response rate can be applied to the university and people who find the price acceptable will visit once per month. The estimated total gross profit would be 17/20* 28744*1*（45.67-15）=HKD 749,341.

3.2.2 Cluster analysis

id	nutrition	taste	environment	wt	ds	gender	age	faculty	expense	status
1	4	5	3	5	5	Female	21	Science	4000	Undergraduate
2	5	3	2	5	5	Female	22	Science	5000	Undergraduate
3	2	3	1	4	5	Female	20	Business and Economics	3000	Undergraduate
4	5	2	1	4	3	Female	20	Science	4500	Undergraduate
5	4	4	4	4	4	Female	20	Science	3000	Undergraduate
6	1	3	3	2	2	Male	21	Business and Economics	4500	Undergraduate
7	2	5	1	4	3	Female	20	Science	3500	Undergraduate
8	2	5	4	3	3	Female	20	Engineering	6000	Undergraduate
9	1	5	2	4	3	Female	24	Education	4000	Postgraduate
10	3	5	4	4	4	Female	20	Business and Economics	3500	Undergraduate
11	4	5	3	4	5	Male	22	Business and Economics	2500	Undergraduate
12	5	4	5	5	5	Male	20	Business and Economics	5000	Undergraduate
13	3	5	3	5	3	Male	24	Science	20000	Undergraduate
14	5	5	4	3	4	Female	20	Business and Economics	2500	Undergraduate
15	3	4	3	4	3	Male	21	Business and Economics	5000	Undergraduate
16	1	5	2	3	4	Female	20	Social Science	7000	Undergraduate
17	4	5	1	2	3	Female	21	Business and Economics	4000	Undergraduate
18	5	2	1	3	4	Female	21	Business and Economics	3500	Undergraduate
19	5	3	2	4	4	Female	21	Science	4500	Undergraduate
20	4	4	3	5	5	Female	20	Business and Economics	5000	Undergraduate

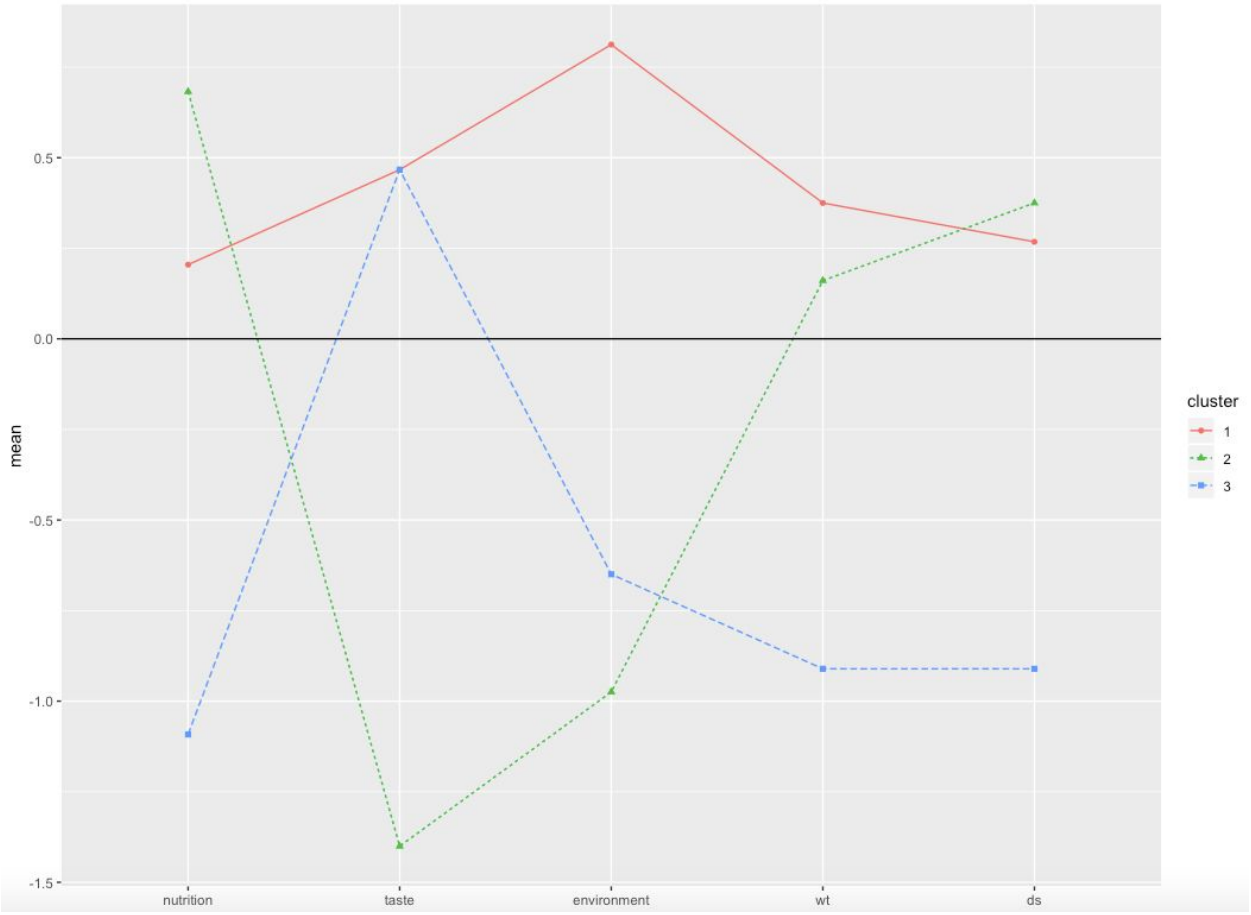
To begin with, Ward's method is applied based on the standardized value of students' rate for the importance of nutrition, taste, environment, waiting time and distance from HKU.





Based on the plot and the dendrogram, 3 groups are suggested, since in the plot of between-group sum of squares, there is a sharp jump from 3-cluster solution to 2-cluster solution, which indicates that there are critical differences between these three clusters, and if we adapt 2-cluster solution, the variance within the group may be too significant. It is also easy to identify three clusters in the dendrogram. Then we need to find the characteristics of students in these three clusters.

cluster	nutrition	taste	environment	wt	ds
1	0.20	0.47	0.81	0.38	0.27
2	0.68	-1.40	-0.97	0.16	0.38
3	-1.09	0.47	-0.65	-0.91	-0.91



Students in cluster 1 tend to pay attention to every attribute of the restaurant, so they are “considerate customers”. Students in cluster 2 tend to emphasize on nutrition, waiting time and the distance from HKU, and from their perspective, taste and environment do not play an important role, so they can be called “pragmatic customers”. Students in cluster 3 tend to pay attention to taste only, so they may be foodies.

The cluster size for cluster 1, 2, and 3 are 10, 5 and 5 respectively.

After that, we can have a look at the relationship between the clustering of customers and the variable gender, faculty, status, age and total monthly expense respectively.

```
        1   2   3
Female 0.6 1.0 0.8
Male   0.4 0.0 0.2
```

	1	2	3
Postgraduate	0.0	0.0	0.2
Undergraduate	1.0	1.0	0.8

	1	2	3
Business and Economics	0.6	0.4	0.4
Education	0.0	0.0	0.2
Engineering	0.1	0.0	0.0
Science	0.3	0.6	0.2
Social Science	0.0	0.0	0.2

Age

	cluster	x
1	1	21
2	2	21
3	3	21

Total monthly expense

	cluster	x
1	1	5650
2	2	4100
3	3	4600

Cluster 1: 60% are female and all are undergraduates. Most are from Faculty of Business and Economics, and the remaining are from Faculty of Engineering and Faculty of Science. On average, the total monthly expense is $5650, which is the highest among three clusters.
Cluster 2: All are female and undergraduates. The majority are from Faculty of Science and the remaining are from Faculty of Business and Economics. On average, the total monthly expense is $4100, which is the lowest among three clusters.
Cluster 3: 80% are female and 80% are undergraduates. Students spread from different faculties. On average, the total monthly expense is $4600.
With respect to age, all clusters have an average age of 21, so there is no significant difference among them in terms of age.
For our hypothetical restaurant, we set short waiting time, short distance from HKU and nutrition as our selling points. Hence, our main customers may be like students in cluster 2. Since they are likely to be “pragmatic customers”, we can launch some campaign or marketing strategies to attract this specific segment.

4. Limitation

We are aware that there are some limitations this study.

4.1 Limited time and energy

In this project, there are limited resources and time. Hence, the sample size is smaller than we expected in the proposal, and the response rate is relatively low.

4.2 Cost estimation

In this study, we do not take net profit into account, and only gross profit is considered. Cost for each meal is assumed to be constant (i.e. $15) and other costs are not taken into consideration. This is a strong assumption which is likely to differ from reality. Costs are likely to be different for different levels of nutrition, taste of food, and for different decorations, etc. Furthermore, considering the high rental fees in Hong Kong, the rental fees may account for a large portion of the actual costs and could be one critical factor when considering launching a new restaurant.

4.3 Availability of restaurant location

In this project, we do not consider whether our proposed restaurant location is available for renting or not.

4.4 Choice of existing off-campus restaurants

Eight existing off-campus restaurants are randomly chosen based on our subjective recognition of their popularity among HKU students. It is possible that our respondents might not have come to those restaurants before, thus giving rise to biase in our results.

5. Conclusion

As mentioned in the introduction part, we are conducting preliminary exploratory market research of launching an off-campus restaurant. One special characteristic is that instead of constructing a general picture, our objective is to use minimum surveying cost to focus on the fiction that we are interested most. To achieve this goal, we conducted two rounds of surveys and we design the second questionnaire according to the result of the first round so that we could dig deeper towards the direction that interests us the most.

The goal for the first round of survey is to find the features that mainly affect the consumer's choice of restaurant, and their relative importance.

According to the result of the correspondence analysis, the raised 5 possible attributes (Delicious food, Healthy & Nutritious, Environment, Waiting time, Distance) can be reduced to two dimensions.The first dimension is mainly contributed by taste vs. health and nutrition while the second dimension mostly reflects the convenience (short waiting time) vs. quality (healthy & nutritious).

The finding from correspondence analysis happens to be consistent with our designed attributes for logit choice model (food style, price & waiting time) to some degree. The preference between the Asian & Western food style might reflect how consumers are thinking about them in the taste vs. health dimension.

It is generally believed that western foods tend to sacrifice healthiness in order to achieve better taste. It can be deduced that consumers prefer the food style that maintains a better balance between taste vs. health. The attribute waiting time can reflect the customer's' concern of convenience vs food quality (healthy & nutrition).

Logit choice model was conducted to evaluate the relative importance of the three selected attributes. We found that the waiting time (10min/20min) and price ($45/$75) are the two most important factors for our students when choosing restaurants, with waiting time being slightly more important than the price. The profile (Asian, $45, 10 min) that has the highest utility value of 2.1409 was selected as the tentative combination for the new restaurant. Then we launched the second round survey to estimate the performance of this tentative combination and the reliability of the prediction.

Since we found that our target customers are quite price-sensitive, the choice between $45 and $75 might not be adequate to capture their price preference. We provide a wider range of price choices and smaller steps to find the optimized gross profit. (We targeted gross profit instead of net profit because the cost is hard to estimate and it will not affect the procedure of analysis.) The rate of accepting to pay this price to have meals at this restaurant is 17/20=85%. By assuming this 85% response rate can be applied to the university and people who find the price acceptable will visit once per month, the optimized gross profit would be $749,341 at a price = $45.67. This is a general estimation.

Then we conducted the cluster analysis to study the characteristics of the respondents (sampled customers). They can be divided into 3 clusters. Students in cluster 1 tend to pay attention to every attribute of the restaurant, so they are “considerate customers.” Students in cluster 2 tend to emphasize on nutrition, waiting time and the distance from HKU, and in their eyes, taste and environment do not play an important role, so they can be called “pragmatic customers”(They emphasize more on dimension 2). Students in cluster 3 tend to pay attention to taste only, so they may be foodies (dimension 1). The cluster size for cluster 1, 2, and 3 are 10, 5 and 5 respectively. Students in cluster 2 will be our targets for the hypothetical restaurant, and they account for 25% of the total number of students. Hence, to make a very conservative estimation, assume only the 25% of students are interested in this restaurant, and they will visit the restaurant twice per month, then the total monthly gross profit can be estimated as 25%*85%*28744*2*$46=$561,945.2. We find this estimation is somewhat close to the previous estimation based on the generalized acceptance rate of 85%. The results from marketing response model and clustering analysis are consistent. We could further assign different estimated visiting frequency for each cluster of customers to estimate the total number of customers visiting from other angles. This requires further studies.

The result of the combination of (Asian, $45.67, 10 mins) and its corresponding predicted monthly gross profit and the predicted number of visiting could serve as a baseline and a starting point for the management team. Firstly we will ask them whether it is acceptable or achievable to have the monthly gross profit of a certain amount of money for a restaurant near HKU. If the maximum possible gross profit is too low, the management team can abandon this project right away with minimum researching cost. If the estimated gross profit is acceptable, they could continue the research by replicating the (Asian, $45.67, 10 mins) attributes like finding a suitable place, estimate number of waitress needed to hire. If all attributes cannot be met at the same time, they can refer to the relative importance to decide which attribute to compromise. They could also conduct a third round of survey to ask consumers' preference among a few actual combinations of profiles.

The advantage of this keep narrowing investigating method is that we can focus on the most interested section and neglect the others. However, it is sensitive to the result of the previous surveys since it decides the direction of later research. We suggest using this method only in exploratory situations so the benefit from efficiency will be more likely to weigh out the risks.

Project report of 15 – 50 pages including:

(iv) Introduction section
Background and motivation.
Precise objectives.
(v) Methods section
Target population with description and reasons.
Sampling design including a sample questionnaire, the sampling method, sample size and response rate.
Data collection procedures including the mode of data collection such as face-to-face interview and the time and place of data collection.
Statistical methods.
(vi) Results section
Provide summary statistics with means, frequencies and standard deviations in tables, graphs or both forms for each variables as well as correlation between relevant variables.

You may also describe them to review the basic information about the data.
Data analysis with description of the variables and models used and how they are related to each of the objective.
Analysis results with relevant tables, graphs and interpretation. The use of graphs is always suggested to facilitate presentation.
(vii) Limitations section
Sources of errors and possible remedies.
Sample size, omitted important factors, representativeness of the sample, response rate, etc.
Model assumptions.
(viii) Conclusion section
Address to the objectives, state your decision of the business and purpose any necessary marketing policies.
Support any conclusion or assertion you made by quoting your analyses results or figures.