International Journal of Multidisciplinary Approach and Studies
ISSN NO:: 2348 – 537X

Analysis of Relationships between Internet Usage Reasons By Log-Linear Models

Dr. Seda Bağdatli Kalkan

Assistant Professor, Department of Statistics, Istanbul Commerce University, Istanbul, Turkey.

ABSTRACT

In consequence of the rapid advancements in the information technologies globewide, the internet usage started to occupy a dominant place in our lives. Thanks to the internet, it is now viable to access any given type of information any time and share the information with the rest of users. When the internet first came to the scene its main usage objective was accessing academic knowledge whereas in modern age internet can be used for a myriad of purposes. In parallel with the pervasive use of smart phones and tablet computers, internet usage climbed the peak and opened a variety of purposes in internet usage.

In present study, by applying data collected from 2017-dated Household Information Technologies Usage Survey conducted by Turkish Statistical Institute (TÜİK), it was aimed to specify the top three reasons for internet usage among participants. Interrelations concerning the three-identified variables were analyzed by applying Log-Linear models. In the next stage Multiple Correspondence Analysis was harnessed to visualize the relations across the categories of variables and collected findings were then interpreted.

KEY WORDS: *Log-Linear Analysis, Multiple Correspondence Analysis, Internet Usage*

1. INTRODUCTION

In consequence of the rapid advancements in the information technologies globewide, computers, smart phones and tablet computers started to play an increasingly more powerful role in our lives. One of the reasons behind the rising popularity of these tools is the presence internet access. Internet is a complex network system that can synchronously connect millions of people overall the world and via synchronicity the very same system can eliminate the vitality of place and transform the entire world into a global village (Balcı, Arsal Gölcü, & Eray Öcalan, 2013). Mostly identified as pool or library; any information shared is open to the access of all users having an connection (Tüysüz, Balaban, & Atalar, 2012). Internet is a cluster of resources established within two or more local networks or a wider web in which it can be possible for millions of subnetworks operating overall the world to communicate with one another within the specifics of a joint protocol and allow them to share one another's resources (Balay, Kaya, & Çevik, 2014). During the very first years internet access was possible, information sharing was basically on academic level but recently present usage of social media, internet shopping, banking transactions via internet allowed to reach any kind of information via internet. In any given space in the world with an internet connection, it is now viable to access and share obtained information.

In line with the popularity and ease of internet usage, reasons behind internet usage have either multiplied or transformed. Personal reasons behind internet usage are affected by a variety of socio-economic factors. In order to identify these factors and monitor general tendency among citizens, in different countries periodical surveys to analyze usage of information technologies are conducted by respective national statistical institutes. In this study, data collected from 2017-dated Household Information Technologies Usage Survey conducted by Turkish Statistical Institute (TÜİK) were applied. It was first aimed to specify the top-three reasons behind internet usage. These reasons were respectively ordered as connecting the social media, discovering health-related information and seeking information on goods and services. Interrelations across the three-identified variables were analyzed by applying Log-Linear Models. In the next stage Multiple Correspondence Analysis was harnessed to visualize the relations across the categories of variables and collected findings were then interpreted.

2. MATERIAL AND METHODS

In this study data collected from 2017-dated Household Information Technologies Usage Survey conducted by TÜİK were applied. This survey has been repeatedly conducted since 2004 to the end of compiling data on the pervasiveness of information and communication technologies and its usage in houses and house members. This survey is the preliminary source of data providing information on the usage of particular technologies (TÜİK, 2017). The survey consists of 6 parts categorized as the presence of information and communication technologies in houses, computer usage, internet usage, e-commerce, e-state applications and information security. In addition, these data were collected and shared with the general public in 2 different categories as per-person and per-house levels. In 2017- dated survey, data from 30407 people and 12781 houses were compiled.

2.1. Variables

In this research, person-based data were gathered to detect personal reasons behind internet usage by applying Household Information Technologies Usage Survey. To reveal internet usage reasons of people, 14 questions were asked and answers were selected as Yes-No. In Table 1 percentages on the internet usage reasons of people are illustrated.

Table 1. Percentages on the Internet Usage Reasons of People

Variables	%
1-E-mail send/receive	44,7%
2- Phone call/video call on the Internet (via Skype or Facetime etc.)	61,3%
3- Profile setting, message transfer or photograph or context sharing on social media (via Facebook, Twitter, Instagram etc.)	**82,6%**
4-Uploading contexts such as personal texts, photograph, music, video, software to an internet web site for sharing purposes	59,7%
5-Surfing through online-news websites and reading online magazines and newspapers	60,2%

6-Discovering health-related information (on injuries, diseases, nutrition etc.)	**69,4%**
7- Seeking information on goods and services	**64,7%**
8-Making a Web site or a blog	17,3%
9-Participating to online polls about a social or political issue	7,1%
10-Job seeking or job application	8,7%
11-Participating to a professional group (Linkedln, Xing and similar career websites)	2,9%
12-Using services related to travel and accommodation (Hotel reservation, ticket sale etc.)	14%
13-Selling goods and services (gittigidiyor.com, sahibinden.com etc.)	17,8%
14-Internet banking	33,6%

Table 1 shows that the most popular reasons for internet usage are respectively ordered as social media connection, discovering health-related information and seeking information on goods and services. Based on these findings, it was resolved to analyze the interrelations among the top-3 factors guiding individuals' internet usage.

2.2. Log-Linear Models

Log-Linear models are the models that identify the level that frequencies in the cells of crosstabs depend on the categories of variables and these models also analyze the interrelations between categorical variables and joint relations (Von Eye & Niedermeier, Statistical Analysis of Longitudinal Categorical Data in the Social and Behavioral Sciences, 2008). In such models there is no differentiation between dependent and independent variables. Thus, it is infeasible to analyze cause-effect relations. In cases where there is no differentiation between dependent and independent variables, it is more logical to use Logistic regression analysis. In situations where chi-square analysis can be applied but fail to be comprehensive, Log-Linear model is a favorable method that can analyze multi-dimensional tables by utilizing all the models. The main objective of these models is modeling the cell frequencies for the entire range of combinations surfacing in relation to the categories of variables. Log-linear models are described as the generalized version of bidirectional crosstabs in which conditioned relationship between two or more variables is analyzed by taking the natural logarithm of the cell frequencies within crosstabs (Erdugan & Türkan, 2017). Log-Linear models serve three purposes: distribution of the compound forged by variables, to detect if variables are interdependent or not and to test the relation between variables without taking cause-effect relationship into account (Von Eye & Mun, Log-Lineer Modeling, 2013).

Observation values in multidimensional crosstabs are measured in accordance with specific probability models. In general; Multinomial, Poisson and Product Multinomial sampling models are used in crosstabs. In Poisson sampling model there is not a first value that is taken in response to the total sum of observation. Within a preset length of fixed interval, Poisson process is observed in every single cell of the tables. In this model n -representing the sampling size- is random variable. t provides time-dependent examples like traffic accidents and birth ratios. In Multinomial sampling model there is one constant n sampling size and it is suggested that observations are multinomially distributed from multiple categories. Product

Multinomial sampling model is the kind of model in which it is hypothesized that samplings collected from varied populations are independent and each sampling is multinomially distributed. In these models, sizes of samplings in rows are fixed. For each single type of sampling model, estimates of expected frequencies can be predicted by using the same method (Arı, 2016).

Logarithmic linear models illustrate the logarithm of every-cell frequency within the crosstabs as the linear combination of all the probable interaction between the variables displayed in the table (Yurt Öncel & Erdugan, 2015). Three directional crosstabs are employed to demonstrate relations across three categorical variables. A model in which the relation among A, B, C variables possessed with *i*, *j* and *k* number of categories can be expressed as below:

$$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \quad (i=1,2,\ldots,I,\ j=1,2,\ldots,J,\ k=1,2,\ldots,K)$$

This is a saturated model which is also a third-degree hierarchical model. In this model there are certain parameters that involve average main effect (μ), main effects ($\lambda_i^A, \lambda_j^B, \lambda_k^C$), all expected dual relations ($\lambda_{ij}^{AB}, \lambda_{ik}^{AC}, \lambda_{jk}^{BC}$) and triple interactions (λ_{ijk}^{ABC}) (Bülbül S. , 2006). A relation structure fitting for the data is not necessarily identical with the equation above in all instances. Less complex models that involve the subset of parameters in this model can be the most applicable model for data. For three-directional crosstabs it is viable to form 9 Log-Linear model samples. These models can be categorized under 5 groups which are respectively; Mutual independence, Joint independence, Conditional independence, Homogeneous association and Saturated Models (Agresti, Categorical Data Analysis, 2013). Mutual independence is the model that involves one single-directional relationship term only, it is also known as corresponding independence model. Joint independence model is the model in which there is only one interaction term. Conditional independence model is the model in which two interaction terms are presented. Homogeneous association model is the model in which all three variables are in interdependent conditionality. This model states that between any given two variables the ratio of conditioned odds would be even in any given level of the third variable (Bülbül S., 2006). Saturated model *is* the model that involves the main effect and interaction parameters (Agresti, Categorical Data Analysis, 2013). In Table 2, models that can be formed in three-directional crosstabs are at display;

Table 2: Log-Linear Models that can be formed in Three-directional Crosstabs

Model	Display
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C$	(A, B, C)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$	(AB, C)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC}$	(AC, B)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC}$	(BC, A)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC}$	(AB, BC)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$	(AB, AC)

$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	(AC, BC)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	(AB, AC, BC)
$lnf_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$	(ABC)

Models displayed in Table 2 are hierarchal log-linear models. Each and every model is represented by a symbol standing for A, B and C variables. Here; ABC stands for the three-directional interaction across A, B and C whilst AB stands for bidirectional relations with AC and BC. The first model displays mutual independence model in Table 2; subsequent three models represent Joint independence model in which there is only one interaction term; Model 5-7 displays conditional independence model in which there is one interaction term; Model 8 symbolizes homogeneous association model that integrates all dual interaction terms and the last model exhibits all dual interrelations and saturated model that integrates interaction of all the three variables (Filiz, 2007).

In Log-Linear models, parameter estimates are conducted through Maximum Likelihood method. Maximum Likelihood method is the selected method because compared to many alternatives such as weighted least squares, minimum chi-square, and minimum discriminate information, it is the most effective prediction method for random variables of which probability distribution is known (Vermunt, Log-Linear Models for Events Histories, 1997). In order to determine Maximum Likelihood function it should be assumed that observed frequencies in crosstabs fits to a specified sampling distribution. In Log-Linear analysis, it is argued that observations fit with either Poisson or Multinomial Distribution (Von Eye & Mun, Log-Lineer Modeling, 2013). In specific events it is unfeasible to reach direct estimate of Log-Linear models. In that case iterative methods viz. Newton-Ralphson and Iterative Proportional Homogeneity Algorithm are harnessed (Agresti, Categorical Data Analysis, 2013).

To the end of selecting the best homogeneity model for the observed data, it is required to investigate which effects are significant or insignificant. Since in Multidimensional Tables there would be a large number of models to test for selecting the best homogeneous model for available data, it is recommended to test the hypotheses via a systematic method (Zeren Yıldırım, 2003). Modeling process is initiated with a saturated model that involves all potential relations across independent variables. To detect the most homogeneous model it is essential to hierarchically analyze the models and to eliminate the relations that contribute to fitness of the model in a minimum degree. To that end, first, backward-stepwise elimination is applied (Bülbül S., 2006). Hence testing process of the hypotheses is commenced with the first hypothesis which includes parameters with the highest degree. In order to step into the next hypothesis, it is essential that all hypotheses manifesting the insignificance of parameters have been accepted till that stage. By combining the parameters in various forms, different hypotheses are tested and the most homogeneous model is then determined (Zeren Yıldırım, 2003). Chi-Square Fitness Test (Likelihood Ratio - G^2 and Pearson Chi-Square - X^2statistics), Model residuals analysis and Partial Relations Terms Test are harnessed in searching which model is the most homogenous model for data on hand.

2.3. Multiple Correspondence Analysis

Multiple Correspondence Analysis (*Homogeneity Analysis by Alternating Least Squares HOMALS*) is presented as the generalized version of simple correspondance analysis. In relevant literature Multiple Correspondence Analysis is also termed as Digitization Method, Dual Scaling, Multiple Homogeneity Analysis and Scalogram Analysis as well. Multiple Correspondence Analysis is a method that can reveal similarities, differences and interrelations of the rows and columns in crosstabs and this method can also graphically display its changes in a few dimensional plots (Suner & Çelikoğlu, 2008).

In Multiple Correspondence Analysis G indicator matrix is used. Row number of the G matrix is equivalent with observation sum and its column number is equivalent with the total of the category numbers in all variables. Members of this matrix are composed of 0s and 1s. In each row, 1 is put next to the related category of relevant variable and 0 is put next to the other category or categories. Via correspondence analysis to apply with the support of G matrix and homogeneity analysis that would be performed with the support of $G'G$ matrix a.k.a Burt matrix are equivalents (Alpar, 2011). Burt Matrix is a quadratic and symmetrical matrix. Members on its principal diagonal point are single marginal frequencies whereas members outside the scope of diagonal point are dual marginal frequencies. This matrix integrates dual crosstabs for all variables in the analysis and since it is a quadratic matrix, it can be exposed to eigenvalue differentiation process. Eigenvalues obtained at the end of this process are principal inertia values (Giray, 2011). Application of all the processes that are present in simple homogeneity analysis into Burt matrix is the equation of Multiple Correspondence Analysis process (Van de Geer, 1993).

The main objective of this analysis is to offer an interpretation for the graphic picture revealed at the end of Multiple Correspondence Analysis. In the interpretation of graphic; the points to consider are as shortly explained below:

- Distance of category point to the origin refers to the significance of category.
- If the direction of one point is opposite to the direction of other points, there is negative correlation in between. If the direction is the same, correlation is then positive.
- If the angle between the lines showing the distance of two category points to the origin is small, or in a different saying if the points are close to one another, correlation between them is high and if the angle is big, correlation is lower (Giray, 2011).

There are similarities between Multiple Correspondence Analysis and Principal Components Analysis. In the two analysis methods, the objective is to degrade dimension of data matrix and to show it in a more comprehensible way in a multidimensional plot. In the two analysis methods, the key difference lies in the type of data matrix. In Principal components analysis it is required that data are composed of variables measured with a constant or intermittent scale that can provide the multi-variable normal distribution hypothesis; but in Multiple Correspondence Analysis data are categorical and there is no requirement for any distribution hypothesis (Suner & Çelikoğlu, 2008). Multiple Correspondence Analysis has also similarities with Multidimensional Scaling. Yet, since it shows the relationship between categories in the same plot it differs from Multidimensional Scaling (Etikan, Uysal, Sanisoğlu, & Dirican, 2000). In relevant literature there is a myriad of studies in which Multiple Correspondence Analysis and Log-Linear models are synchronously used. Relations that surface with Log-Linear models can also be tested via Multiple Correspondence Analysis. It can thus be suggested that these two analysis methods are complementary in nature.

3. FINDINGS

In order to determine the best fitting Log-linear model to analyze the reasons behind internet usage among people, all models were tested independently starting with the model presenting all of the main effects and interaction terms. G^2 and X^2 test results of these models were then examined. Consequently it was decided that saturated model in which all main effects and interaction terms were included was the best fitting model to answer the purpose of this research. This model is expressed such;

*ln(f_{ijk})= Constant + Social Media + Health + Goods and Services + (Social Media * Health)+ (Social Media * Goods and Services) + (Health * Goods and Services) + (Social Media * Health * Goods and Services).*

Parameter estimates of this Model are as displayed in Table 3.

Table 3: Parameter Estimates of the Best Fitting Log-Linear Model

Parameter Estimates						
					95% Confidence Interval	
Parameter	Estimate	Std. Error	Z	Sig.	Lower Bound	Upper Bound
Constant	6,606	,037	179,641	,000	6,534	6,678
[Social Media = 1]	1,350	,041	**32,705**	,000	1,269	1,431
[Social Media = 2]	0	.	.	.	.	.
[Health = 1]	-,067	,053	-1,269	,205	-,171	,037
[Health = 2]	0	.	.	.	.	.
[Goods and Services = 1]	-,546	,061	-8,988	,000	-,665	-,427
[Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 1] * [Health = 1]	-,265	,060	-4,387	,000	-,383	-,146
[Social Media = 1] * [Health = 2]	0	.	.	.	.	.
[Social Media = 2] * [Health = 1]	0	.	.	.	.	.
[Social Media = 2] * [Health = 2]	0	.	.	.	.	.

[Social Media = 1] * [Goods and Services = 1]	-,123	,069	-1,796	,073	-,258	,011
[Social Media = 1] * [Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 2] * [Goods and Services = 1]	0	.	.	.	.	.
[Social Media = 2] * [Goods and Services = 2]	0	.	.	.	.	.
[Health = 1] * [Goods and Services = 1]	1,141	,077	14,811	,000	,990	1,291
[Health = 1] * [Goods and Services = 2]	0	.	.	.	.	.
[Health = 2] * [Goods and Services = 1]	0	.	.	.	.	.
[Health = 2] * [Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 1] * [Health = 1] * [Goods and Services = 1]	,945	,087	10,855	,000	,774	1,115
[Social Media = 1] * [Health = 1] * [Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 1] * [Health = 2] * [Goods and Services = 1]	0	.	.	.	.	.
[Social Media = 1] * [Health = 2] * [Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 2] * [Health = 1] * [Goods and Services = 1]	0	.	.	.	.	.
[Social Media = 2] * [Health = 1] * [Goods and Services = 2]	0	.	.	.	.	.
[Social Media = 2] * [Health = 2] * [Goods and Services = 1]	0	.	.	.	.	.
[Social Media = 2] * [Health = 2] * [Goods and Services = 2]	0	.	.	.	.	.

Log-Linear models are analyzed in two stages. In the first stage main effects are analyzed, subsequently interaction terms are examined. In both cases, the findings are interpreted via examining standardized parameter estimates (Z). Z values provide insights about the categories in which interrelations are more effective.

In this Model an analysis of main effects shows that the highest-standardized parameter estimate value (Z=32,705) belongs to those claiming to use the internet to connect the social media. Thus it is evident that the most significant factor determinant on the frequencies in contingency table is this factor. It was also identified that those responding to the question of using the internet to discover health-related information as yes or no were not statistically significant independently (p=0,205); yet as the interaction terms in the model were evaluated it surfaced that this variable interrelated with the other variables; thereby contributing to the overall model. Thus it is not advised to eliminate this variable from the model. An analysis of this table exhibits that some estimate values failed to be computed. These values can only be computed by benefiting from the limit that total parameter equates with zero. For instance; parameter estimate of those claiming to use the internet to connect social media was measured as 1,350. In that case, parameter estimate of those not using the internet to connect social media could be computed as -1,350.

As the interaction terms in the model are analyzed it is viable to determine the interrelations of variables across the categories. The table shows that there is no correlation between those claiming to use the internet to connect social media and those using the internet to seek information on goods and services (p=0,073). This finding indicates that none of the categories of both variables are in an interactive relation. As dual interaction terms are evaluated it is evidenced that the highest interaction can be measured between those using the internet to discover health-related information and those using the internet to seek information on goods and services (Z=14,811). The second highest relationship is between those using the internet to connect social media and those using the internet to discover health-related information. Since this is a saturated model it also involves triple interaction term which significantly contributes to the model. It is thus evident that there is a strong interaction among those claiming to use the internet to connect social media, those claiming to use the internet to discover health-related information and those arguing to use the internet for seeking information on goods and services.

In Log-Linear models it is also viable to provide certain comments on odds ratios. Odds ratios can also be analyzed in two stages as main effects and interaction effects. Odds ratio is computed by taking *e*-based antilogarithms of parameter estimates and then be interpreted. Odds ratio belonging to social media variable is computed as 3,86 ($e^{1,350}$). Based on this outcome it can be concluded that for a randomly selected subject, the likelihood of using the internet to connect social media is 3,86 times above the likelihood of not using the internet for that purpose. Odds ratio of health variable is computed as 0,94 ($e^{-0,067}$). Based on this outcome it can be concluded that for a randomly selected subject, the likelihood of using the internet to discover health-related information is 0,94 times above the likelihood of not using the internet for that purpose. Finally odds ratio of goods and services variable is computed as 0,58 ($e^{-0,546}$). Based on this outcome it can be concluded that for a randomly selected subject, the likelihood of using the internet to seek information on goods and services is 0,58 times above the likelihood of not using the internet for that purpose.

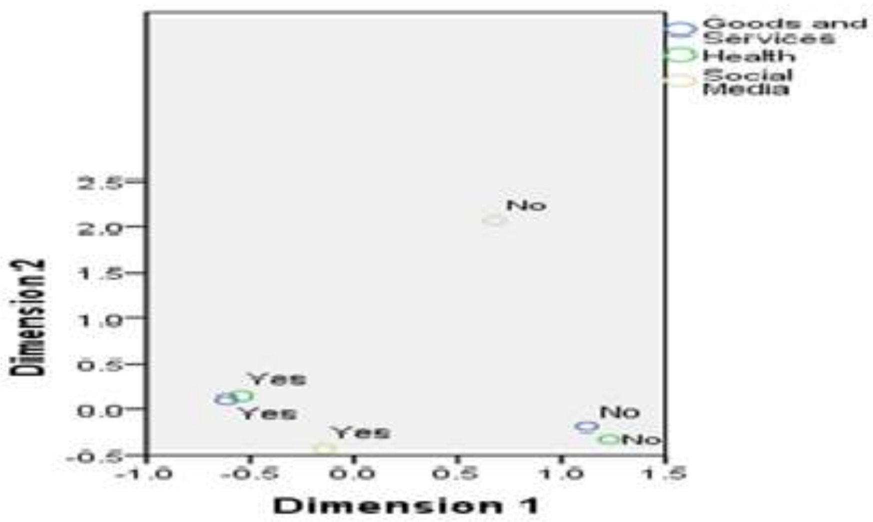
An analysis of odds ratios of interaction effects determines that odds ratio belonging to social media variable and health variable interaction can be computed as 0,77 ($e^{-0,265}$). Therefore likelihood of those using the internet to connect social media and discover health-related information is 0,77 times above the likelihood of those not using the internet to connect social media but using the internet to discover health-related information only.

Odds ratio of the interaction between social media variable and goods & services variable is computed as 0,88 ($e^{-0,123}$). It can thus be argued that the likelihood of those using the internet to connect social media and seek information on goods and services is 0,88 above those using the internet not to connect social media but to seek information on goods and services.

Odds ratio of the interaction between health variable and goods & services variable is computed as 3,13 ($e^{1,141}$). It can thus be argued that the likelihood of those using the internet to discover health-related information and seek information on goods and services is 3,13 above those using the internet not to discover health-related information but to seek information on goods and services.

Subsequent to this stage, Multiple Correspondence Analysis was administered to better visualize interrelations of the variables across the categories. The graphic designed in accordance with the findings of Multiple Correspondence Analysis can be monitored in Figure 1.

Figure 1: Joint Plot of Category Points



In this graphic designed in line with the findings of Multiple Correspondence Analysis, in parallel with the results of Log-linear analysis, it is evidenced that there is a correspondence between those claiming to use the internet to connect social media, for discovering health-related information and for seeking information on goods and services; in a different saying those dimensions are in an interactive relationship. As seen in the graph above, there is another interaction between those using the internet not to seek information on goods and services and not to discover health-related information likewise.

4. CONCLUSIONS

It is quite a natural outcome that individuals' reasons behind internet usage experienced a shift in parallel with the advancing technology. To illustrate; internet banking and e-commerce activities are some of the applications that have recently come into the scene. Internet usage reasons are subject to be affected by the geographical setting and socio-economic factors influential on the person. Yet considering the everyday existence of internet in our lives, the main purpose behind internet usage is quite evident. As seen in this study too, internet is used most frequently for three reasons essentially, which are alternatively listed as connecting the social media, discovering health-related information and seeking information on goods and services. Considering that people connect the social media to discover information on others' lives, it can be argued that in Turkey the generic reason behind internet usage is to attain information.

The purpose of this research is to analyze the relations among the main three reasons behind internet usage which are connecting the social media, discovering health-related information and seeking information on goods and services. As the specifically-designated log-linear model for this objective is analyzed, it is concluded that connecting the social media is the variable that contributes to the model at the highest level. It can thus be claimed that people mostly use the internet to connect social media. The principal objective of the model is to unveil the relations of variables across the categories. As the interaction terms are examined to find the answer for research question, the highest interaction is observed between those claiming to use the internet to discover health-related information and those using it to seek information on goods and services. It is thus concluded that those claiming to use the internet to discover health-related information also use the internet to seek information on goods and services. The second highest interaction is measured between those claiming to use the internet to connect social media and to discover health-related information. It is thus concluded that those claiming to use the internet to connect social media also use the internet to discover health-related information. All in all, since triple interaction term also significantly contributes to the model, it is exhibited that there is a strong interaction among those claiming to use the internet to connect the social media, to discover health-related information and to seek information on goods and services. This finding is in parallel with the results of Multiple Correspondence analysis shown in Figure 1.

REFERENCES

i. Agresti, A. (2013). *Categorical Data Analysis.* USA: WILEY.

ii. Alpar, R. (2011). *Çok Değişkenli İstatistiksel Yöntemler.* Ankara: Detay Yayıncılık.

iii. Arı, E. (2016). Trafik Kaza Verilerinin Log Lineer Modeller İle İncelenmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 17-37.

iv. Balay, R., Kaya, A., & Çevik, M. N. (2014). ÖğretmenleriInternete Yönelik Tutumları ve Eğitsel İnternet Kullanım Öz-Yeterlilik İnanç Düzeyleri. *Karadeniz Uluslararası Bilimsel Dergi*, s. 16-31.

v. Balcı, Ş., Arsal Gölcü, A., & Eray Öcalan, M. (2013). Üniversite Öğrencileri Arasında İnternet Kullanım Örüntüleri. *Selçuk İletişim*, s. 5-22.

vi. Bülbül, S. (2006). Üç Boyutlu Çapraz Tablolarda Logaritmik Doğrusal Analiz: Çocuk İşgücü Değişkenleri Arasındaki Etkileşimler. *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 41-70.

vii. Erdugan, F., & Türkan, A. H. (2017). Üç Yönlü Kontenjans Tablolarında Log-Lineer Model İle İş Kazası Verilerinin İncelenmesi. *Karaelmas Fen ve Mühendislik Dergisi*, 462-468.

viii. Etikan, İ., Uysal, M., Sanisoğlu, Y., & Dirican, B. (2000). Uygunluk Analizi İle Kanser Vakalarının Çözümlenmesi. *5. Ulusal Biyoistatistik Kongresi*, (s. 118-127). Eskişehir.

ix. Filiz, Z. (2007). Üç Yönlü Log-Lineer Modeller İle Üniversite Öğrencilerinin Sigara Alkol ve Nargile İçme Nedenlerini Etkileyen Faktörlerin Belirlenmesi. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 225-250.

x. Giray, S. (2011). Doğrusal Olmayan Kanonik Korelasyon Analizi ve Yaşam Memnuniyeti Üzerine Bir Uygulama. *Doktora Tezi*. Marmara Üniversitesi Sosyal Bilimler Enstitüsü.

xi. Powers, D. A., & Xie, Y. (2000). *Statistical Methods for Categorical Data Analysis*. USA: Academic Press.

xii. Suner, A., & Çelikoğlu, C. (2008). Uygunluk Analizinin BenzerÇok Değişkenli Analiz Yöntemleri İle Karşılaştırılması. *İstatistikçiler Dergisi*, 9-15.

xiii. TÜİK. (2017, Ağustos). *Hanehalkı Bilişim Teknolojileri Kullanım Araştırması.* Temmuz 2018 tarihinde www.tuik.gov.tr: www.tuik.gov.tr adresinden alındı

xiv. Tüysüz, C., Balaban, F., & Atalar, T. (2012). Eğitim Fakültesi Öğrencilerinin İnternet Kullanım Amaçlarının Belirlenmesi. *Batman Üniversitesi Yaşam bilimleri Dergisi*, s. 251-260.

xv. Van de Geer, J. (1993). *Multivariate Analysis of Categorical Data: Theory*. California: Sage Publication.

xvi. Vermunt, J. (1997). *Log-Lineer Models for Events Histories*. California: Sage.

xvii. Von Eye, A., & Mun, E.-Y. (2013). *Log-Lineer Modeling*. USA: WILEY.

xviii. Von Eye, A., & Niedermeier, K. (2008). *Statistical Analysis of Longitudinal Categorical Data In The Social and Behavioral Sciences*. USA: Taylor&Francis Group.

xix. Yurt Öncel, S., & Erdugan, F. (2015). Kontenjans Tablolalarının Analizinde Log-Lineer Modellerin Kullanımı ve Sigara Bağımlılığı Üzerine Bir Uygulama. *Sakarya Üniversitesi Fen Bilimleri Dergisi*, 221-235.

xx. Zeren Yıldırım, E. (2003). Gizli Sınıf Analizi ve Türkiye'de Üniversite Mezunu Gençlerin Avrupa Birliği'ne Bakış Açıları Üzerine Bir Uygulama. Marmara Üniversitesi Sosyal Bilimler Enstitüsü, Yayınlanmamış Doktora Tezi.