

İnternet kullanıcılarının demografik özellikleri ile önde gelen web arama motorları tercihleri arasındaki ilişki üzerine bir araştırma

Evren AYRANCI*

Özet

Günümüzde milyarlarca insan interneti farklı amaçlar için kullanmaktadır. Web arama motorları internet kullanıcıları arasında, ihtiyaç duyulan bilgiyi bulmak için çok popülerdir. Bu makale, internet kullanıcılarının bazı demografik özellikleri ile önde gelen web arama motorları tercihleri arasında bir ilişkinin olup olmadığını anlamak için yapılan bir araştırmayı içermektedir. Elde edilen sonuçlara göre demografik özelliklerin çoğu önde gelen web arama motorları tercihleri ile bir takım bağlantılara sahiptir.

Anahtar Kelimeler: İnternet, arama motoru, seçim, demografik özellikler.

A research on the relationship between demographic properties and prominent web search engine choices of internet users

Abstract

Today, billions of people use internet for different purposes. Web search engines are very popular among internet users to find the information needed. This article includes a research conducted, in order to understand whether there is a relationship between some demographic properties of internet users and prominent web search engine choices. According to results obtained, most of the demographic properties have some connections with the choices of prominent web search engines.

Keywords: *Internet, search engine, choice, demographic properties.*

Introduction

Internet is used frequently in every-day life. It is seen as the biggest source of information, fun and the best medium for communication. In this article, internet's role of being a source of information is considered.

* Evren AYRANCI, İstanbul Aydın Üniversitesi, İktisadi ve İdari Bilimler Fakültesi. evrenayranci@adolubil.edu.tr

According to the statistics, currently there are about 6.8 billions of people on Earth and about 1.7 billion of people use Internet (Internet World Stats, 2009). From the aspect of information storage, it may be concluded that internet has the biggest accumulation of information *ever* in the history of mankind as it includes more than five million terabytes of data (McGuigan, 2009).

Prominent Web Search Engines and Directories

To cope up with internet's enormous amount of data, people have to use web search engines or internet directories*. The earliest web search engine was "Archie" which was started up in 1990 (Wall, 2009). Archie was mostly used to search the web pages of universities in US and therefore it was a popular web search engine among scholars at the beginning of 1990s.

Prominent web search engines and directories that are currently known by average internet users started to appear after the year 1993. Table 1 includes prominent web search engines as well as directories (Wall, 2009):

Table 1. Prominent Web Search Engines, Internet Directories and Their Years of Foundation

Name of the Web Search Engine or Internet Directory	Year of Foundation
Excite	1993
Altavista	1994
Infoseek	1994
Lycos	1994
Yahoo	1994
Webcrawler	1994
Inktomi	1996
Google	1998
MSN Search	1998
AllTheWeb	1999
Live Search (Formerly MSN Search)	2006
Bing (Formerly Live Search)	2009

The web search engines or directories in Table 1 are all for general purposes, which means they can be used to obtain any kind of information. As it was mentioned, the internet is huge and people need very specific information. Some entrepreneurs realized this situation and formed thematic web search engines or directories. These thematic solutions are focusing on very specific types of information and some examples can be found in Table 2.

* Internet directories include lists of topics that can be clicked on, in order to see the detailed information of each topic.

Table 2. Some Examples to Thematic Web Search Engines and Directories

Type of Information to Search for	Name of the Thematic Web Search Engine or Directory
Scientific	Proquest Ojose.com Scirus
Games	GamePublic GameSpot GameSpider
Music	Musicrobot.com Mp3va.com Songza
Prices of products or services	Pricewatch Pricegrabber
Business	Business.com Nexis Globalspec
Jobs	Careerbuilder.com Hotjobs.com Monster.com
Legal	Quicklaw Westlaw
Medical	Searchmedica Healia WebMD
News	Newslookup Topix.net
Real Property	Hotpads.com Rightmove Zillow.com

In the last few years, it is seen that local or national web search engines and directories have flourished. Table 3 includes some examples for these in Turkey.

Table 3. Some Examples to Turkish Web Search Engines and Directories

Name of the Web Search Engine or Internet Directory
Arasam.net
Arabul.com
Ara.com.tr
Turkcearama.com
Hemenbul.com
iyinet.com
interturknet.com
Bilgi.com
Dizin.com
Turkarama.com
Arrama.com
Azbuz.com
Adanzye.com

Research Methodology

Purpose of the Research

In this research, the purpose is:

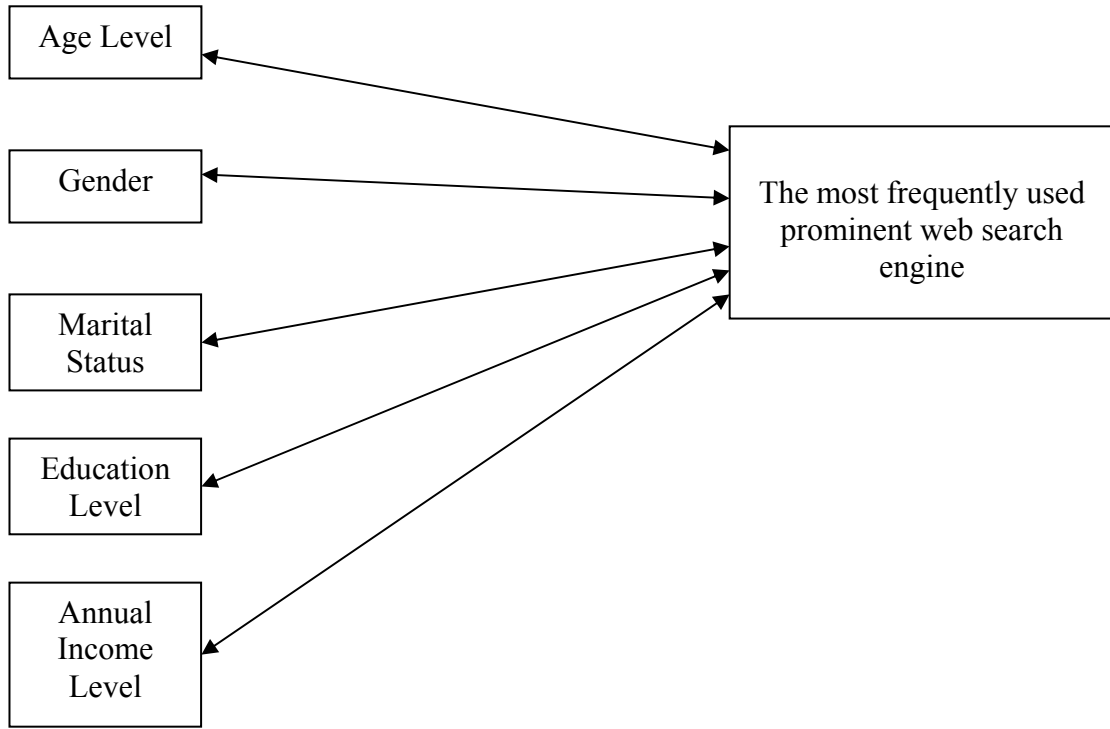
- To understand the relationships between some demographical properties and web search engine preference.

As it is seen, the major focus of this research is on web search engine preference. This preference is very important for the web search engines as rival web search engines provide similar services today. Their main function is to perform web searches. In order to survive, a web search engine must retain the current customers first of all. The best approach to do this is to protect and enhance the brand loyalty and/or to become a powerful choice.

Research Model and Hypotheses

In this research, the relationships between some demographic properties and the most frequently used prominent web search engine (the most preferred major web search engine) are analyzed. Research is performed under the assumption that α is 5%.

Internet Kullanıcılarının Demografik Özellikleri İle Önde Gelen Web Arama Motorları Tercihleri Arasındaki İlişki Üzerine Bir Araştırma



For the design above, five hypotheses are to be tested:

H_0 : There is NO relationship between age level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between age level and the most frequently used prominent web search engine.

H_0 : There is NO relationship between gender and the most frequently used prominent web search engine.

H_1 : There IS a relationship between gender and the most frequently used prominent web search engine.

H_0 : There is NO relationship between marital status and the most frequently used prominent web search engine.

H_1 : There IS a relationship between marital status and the most frequently used prominent web search engine.

H_0 : There is NO relationship between education level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between education level and the most frequently used prominent web search engine.

H_0 : There is NO relationship between annual income level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between annual income level and the most frequently used prominent web search engine.

Sampling method

Convenience sampling method was used. A list of 5000 e-mails that belong to Turkish Internet users was obtained. E-mails were sent to these 5000 people and they were invited to an internet site that included the online survey. Therefore, the method to obtain data was online survey. The internet site that included the survey was www.aramatercihi.com and this survey site was online for two weeks between 12 and 25 January 2009. Out of 5000, 372 people provided answers to the survey. According to this, the response rate was 7,44%.

Results Based on Descriptive Statistics

Table 4 summarizes the findings about the demographical properties of the survey-takers. It is seen that most of the attendees are males and young, single, undergraduate people are common. More than half of the attendees earn 20000 TL or less annually.

Table 4. Demographic Properties of the Attendees

	Gender				
	<i>male</i>		<i>female</i>		
frequency	264		108		
percentage	70,90%		29,10%		
	Age				
	<i>less than 15</i>	<i>15-24</i>	<i>25-34</i>	<i>35-49</i>	<i>50 and more</i>
frequency	89	122	94	52	15
percentage	23,92%	32,79%	25,27%	13,98%	4,00%
	Marital status				
	<i>married</i>		<i>single</i>		
frequency	137		235		
percentage	36,90%		63,10%		
	Education level				
	<i>high school</i>	<i>undergraduate</i>	<i>postgraduate</i>		
frequency	118	176	78		
percentage	31,70%	47,30%	20,90%		
	Annual income				
	<i>10000 TL and less</i>	<i>10001-20000 TL</i>	<i>20001-30000 TL</i>	<i>300001-40000 TL</i>	<i>400001 TL and more</i>
frequency	128	95	86	44	19
percentage	34,40%	25,60%	23,00%	11,80%	5,20%

Internet usage patterns of the attendees are also to be analyzed. According to Table 5, almost 2/3 of survey-takers use internet more than five hours a week. The main purpose of using internet is to have fun. This seems to be normal as most of the attendees are young people. It is interesting to see that 27% of the attendees connect to internet mostly from internet cafes and another 38% from offices.

Table 5. Internet Usage Patterns

	Weekly internet usage			
	<i>less than 1 hour</i>	<i>1 hour-5 hours</i>	<i>more than 5 hours</i>	
frequency	11	124	237	
percentage	3,00%	33,30%	63,70%	
	Why use internet			
	<i>information</i>	<i>fun</i>	<i>business</i>	<i>other</i>
frequency	85	144	110	33
percentage	22,80%	38,70%	29,50%	8,80%
	From where mostly			
	<i>home</i>	<i>office</i>	<i>internet cafe</i>	<i>other</i>
frequency	115	142	101	14
percentage	30,90%	38,10%	27,10%	3,70%

Finally, Table 6 shows the other extent of the research - the most frequently used prominent web search engine.

Table 6. Most Frequently Used Prominent Web Search Engine

	Which prominent web search engine do you use most frequently?				
	Google	Altavista	Yahoo	Live Search	Others
frequency	113	97	106	11	45
percentage	30,38%	26,08%	28,49%	2,96%	12,10%

It should be noted that internal consistency analysis must be performed upon survey questions and the analysis, mostly used to check internal consistency is Cronbach's Alpha. However internal consistency check can be made on continuous scales like Likert-type scale. As the survey used in this research doesn't have any question to be measured with Likert-type scale, Cronbach's Alpha analysis can not be performed.

Results Based on Inferential Statistics

As told earlier, there are six hypotheses to be tested. The data to be used is mostly qualitative and the relationships between variables are examined. Therefore it seems best to use Chi-square analysis to test the hypotheses.

Research Hypotheses Testing

1. PART Demographic properties – search engine preference

H_0 : There is NO relationship between age level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between age level and the most frequently used prominent web search engine.

Table 7 includes the observed variables table (contingency table) for age levels and the most frequently used prominent web search engine.

Table 7. Contingency Table for Age Level-The Most Frequently Used Prominent Web Search Engine

Observed Values (O_{ij} values)						
	less than 15	15 - 24	25 - 34	35 - 49	50 and more	TOTAL
Google	32	35	21	23	2	113
Altavista	26	32	25	11	3	97
Yahoo	21	41	29	9	6	106
Live Search	2	1	5	2	1	11
Other	8	13	14	7	3	45
TOTAL	89	122	94	52	15	372

The expected values calculated are in Table 8.

Table 8. Expected Values Table for Age Level-The Most Frequently Used Prominent Web Search Engine

Expected Values (E_{ij} values)						
	less than 15	15 - 24	25 - 34	35 - 49	50 and more	TOTAL
Google	27,03	37,06	28,55	15,80	4,56	113
Altavista	23,21	31,81	24,51	13,56	3,91	97
Yahoo	25,36	34,76	26,78	14,83	4,27	106
Live Search	2,63	3,61	2,78	1,54	0,44	11
Other	10,77	14,76	11,37	6,29	1,81	45
TOTAL	89	122	94	52	15	372

Degrees of freedom is 16, α is 5%. In this case critical value is found to be 26,3 and Chi-square value is calculated to be 20,89. According to this, we accept the null hypothesis.

There is NO relationship between age level and the most frequently used prominent web search engine.

2. PART Demographic properties – search engine preference

H_0 : There is NO relationship between gender and the most frequently used prominent web search engine.

H_1 : There IS a relationship between gender and the most frequently used prominent web search engine.

The relevant data are found in Table 9.

Table 9. Contingency Table for Gender-The Most Frequently Used Prominent Web Search Engine

Observed Values (O_{ij} values)			
	Male	Female	TOTAL
Google	91	22	113
Altavista	56	41	97
Yahoo	78	28	106
Live Search	3	8	11
Others	36	9	45
TOTAL	264	108	372

Table 10 includes expected values derived from Table 9.

Table 10. Expected Values Table for Gender-The Most Frequently Used Prominent Web Search Engine

Expected Values (E_{ij} values)			
	Male	Female	TOTAL
Google	80,19	32,81	113
Altavista	68,84	28,16	97
Yahoo	75,23	30,77	106
Live Search	7,80	3,20	11
Others	31,94	13,06	45
TOTAL	264	108	372

Degrees of freedom is 4, α is 5%. In this case critical value is found to be 9,49 and Chi-square value is calculated to be 25,55. In this case, the null hypothesis is rejected.

There is a relationship between gender and the most frequently used prominent web search engine.

3. PART Demographic properties – search engine preference

H_0 : There is NO relationship between marital status and the most frequently used prominent web search engine.

H_1 : There IS a relationship between marital status and the most frequently used prominent web search engine.

Table 11. Contingency Table for Marital Status - The Most Frequently Used Prominent Web Search Engine

Observed Values (O_{ij} values)			
	Married	Single	TOTAL
Google	38	75	113
Altavista	13	84	97
Yahoo	37	69	106
Live Search	5	6	11
Others	44	1	45
TOTAL	137	235	372

Expected values are found in Table 12.

Table 12. Expected Values Table for Marital Status - The Most Frequently Used Prominent Web Search Engine

Expected Values (E_{ij} values)			
	Married	Single	TOTAL
Google	41,62	71,38	113
Altavista	35,72	61,28	97
Yahoo	39,04	66,96	106
Live Search	4,05	6,95	11
Others	16,57	28,43	45
TOTAL	137	235	372

Degrees of freedom is 4, α is 5%. In this case critical value is found to be 9,49 and Chi-square value is calculated to be 95,77. The null hypothesis is rejected. So, there is a relationship between marital status and the most frequently used prominent web search engine

4. PART Demographic properties – search engine preference

H_0 : There is NO relationship between education level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between education level and the most frequently used prominent web search engine.

Table 13. Contingency Table for Education Level - The Most Frequently Used Prominent Web Search Engine

Observed Values (O_{ij} values)				
	High school	Undergraduate	Postgraduate	TOTAL
Google	51	40	22	113
Altavista	23	48	26	97
Yahoo	15	65	26	106
Live Search	2	7	2	11
Others	27	16	2	45
TOTAL	118	176	78	372

Expected values are in Table 14.

Table 14. Expected Values Table for Education Level - The Most Frequently Used Prominent Web Search Engine

Expected Values (E_{ij} values)				
	High school	Undergraduate	Postgraduate	TOTAL
Google	35,84	53,46	23,70	113
Altavista	30,77	45,89	20,34	97
Yahoo	33,62	50,15	22,23	106
Live Search	3,49	5,20	2,31	11
Others	14,27	21,29	9,44	45
TOTAL	118	176	78	372

Degrees of freedom is 8, α is 5%. In this case critical value is found to be 15,51 and Chi-square value is calculated to be 48,74. The null hypothesis is rejected and therefore there is a relationship between education level and the most frequently used prominent web search engine.

5. PART Demographic properties – search engine preference

H_0 : There is NO relationship between annual income level and the most frequently used prominent web search engine.

H_1 : There IS a relationship between annual income level and the most frequently used prominent web search engine.

Table 15. Contingency Table for Annual Income Level - The Most Frequently Used Prominent Web Search Engine

Observed Values (O_{ij} values)						
	10000 TL and less	10001-20000 TL	20001-30000 TL	30001-40000 TL	40001 TL and more	TOTAL
Google	51	29	24	4	5	113
Altavista	22	23	11	25	8	97
Yahoo	46	38	19	8	3	106
Live Search	5	1	2	2	1	11
Other	4	4	30	5	2	45
TOTAL	128	95	86	44	19	372

Table 16 includes the expected values derived from Table 15.

Table 16. Expected Values Table for Annual Income Level - The Most Frequently Used Prominent Web Search Engine

Expected Values (E_{ij} values)						
	10000 TL and less	10001-20000 TL	20001-30000 TL	30001-40000 TL	40001 TL and more	TOTAL
Google	38,88	28,86	26,12	13,37	5,77	113
Altavista	33,38	24,77	22,43	11,47	4,95	97
Yahoo	36,47	27,07	24,50	12,54	5,42	106
Live Search	3,78	2,81	2,54	1,30	0,57	11
Other	15,48	11,49	10,40	5,33	2,30	45
TOTAL	128	95	86	44	19	372

Degrees of freedom is 16, α is 5%. In this case critical value is found to be 26,30 and Chi-square value is calculated to be 101,92. The null hypothesis is rejected, there is a relationship between annual income level and the most frequently used prominent web search engine.

Results

When the results are aggregated, it is clearly understood that almost every demographic property of the attendees (namely gender, marital status, education level and annual income level) have relationships with the choice of prominent web search engine, except the age level. It should be noted that all attendees are Turkish internet users.

These results can be summarized as follows: Most of Turkish internet users' demographic properties have relationships with their choices of prominent web search engines.

From this point, some advices can be made. As there are connections between the demographic properties and the choice of prominent web search engine:

- New Turkish search engines that perform general search but that are *tailored* for the demographic properties of Turkish internet users can be built. Table 3 includes some Turkish web search engines and directories but none of them appeal to different demographic properties of Turkish internet users. This “appealing” can be about the design (different designs of the internet site for different genders, education levels...) or the deepness of the search (according to education level of the user...).
- New Turkish search engines that perform search, *tailored* for the demographic properties of Turkish internet users can be built. This means thematic Turkish internet search engines can go online. There might be differences between the subjects searched by male and female Turkish internet users and some further research can be performed to understand mainly what these subjects are. New web search engines can emerge that only scan for the main subjects which males or females want. Some research can also be conducted to see whether there are differences between the subjects searched by Turkish internet users with different marital status, income levels and education levels. Again, new web search engines with the abilities to scan for the subjects that these people are interested, can be founded.
- Current Turkish web search engines can *become thematic*, that is they can be designed or can perform search for Turkish internet users with different demographic properties.

Bibliography

Internet World Stats, 2009. “World Internet Users and Population Stats”, <http://www.internetworldstats.com/stats.htm>, 23.12.2009.

McGuigan, B., 2009. “How Big is the Internet?”, <http://www.wisegeek.com/how-big-is-the-internet.htm>, 23.12.2009.

Wall, A., 2009. “History of Search Engines: From 1945 to Google 2007”, <http://www.searchenginehistory.com>, 23.12.2009.