Health Sciences Research
2016; 3(1): 10-16
Published online January 29, 2016 (http://www.aascit.org/journal/hsr)
ISSN: 2375-379X

American Association for Science and Technology

Keywords
Nutrition,
Students,
Data Analysis,
R

Received: November 25, 2015
Revised: December 22, 2015
Accepted: December 23, 2015

Nutrition Data Analysis Using R: Applications in Higher Education

Efstathios Stefos[1], Georgia Koulianidi[2]

[1]School of Physical Sciences and Mathematics, Faculty of Exact and Natural Sciences, Pontifical Catholic University of Ecuador, Proyecto Prometeo, Quito, Ecuador
[2]Post-graduate of Medical School, University of Athens, Athens, Greece

Email address
estefos@aegean.gr (E. Stefos), koulianidi@yahoo.gr (G. Koulianidi)

Citation
Efstathios Stefos, Georgia Koulianidi. Nutrition Data Analysis Using R: Applications in Higher Education. *Health Sciences Research*. Vol. 3, No. 1, 2016, pp. 10-16.

Abstract
During the academic year 2015-2016, students of the Pontifical Catholic University of Ecuador worked on a project that studied the relationship between food quality and school performance of Greek students aged 10-12 years from schools located in various places of Greece. The university students used R in order to study the data and the statistical analysis showed that students who consume more frequently low quality food products are the ones with the poorest school performance. After the completion of the project, questionnaires delivered to Ecuadorian students in order to explore their views towards the use of R. A Descriptive Analysis and a Hierarchical Clustering were realized and the results showed that the students reacted very positively to the use of R in teaching Statistics.

1. Introduction

The aim of this study is to investigate university students' views towards the use of software R in teaching Statistics. The investigation was realized in the School of Physical Sciences and Mathematics of the Pontifical Catholic University of Ecuador. The students worked on a project studying the effects of consumption of junk food, sweets and soft drinks on the school achievement of students aged 10-12 years [1]. The students used R, statistical techniques and quantitative research methods in order to extract the results. The project used real life data from a survey that was realized by Georgia Koulianidi in the Medical School of the University of Athens in Greece. Data from Greece was used because the interdisciplinarity and the exchange of knowledge, experience and information at international level is essential for the advancement of science and the achievement of the goals and objectives of academic education and research [2]. Also, the use of project method, real life data and new technology were chosen because they improve students' attitudes towards Statistics and make undergraduates to be more engaged in class activities [3].

2. Methodology

During the first semester of the academic year 2015-2016, a project was realized in the School of Physical Sciences and Mathematics of the Pontifical Catholic University of Ecuador. The project investigated the relationship between food quality and school performance of 577 Greek students aged 10-12 from schools located in various places of Greece [4]. The students of the Pontifical Catholic University of Ecuador investigated the relationship between the variables "School performance", "Do you eat

sweets?", "Do you eat junk food?" and "Do you drink soft drinks?" [5]. They used the software R in order to calculate the frequencies and percentages of the variables, the mean and standard deviation of each variable, the t-test, the contingency tables and investigate whether there is a correlation between students' school performance and the frequency of consumption of junk food, sweets and soft drinks [6]. The analysis with R showed that students who consume more frequently low quality food products, widely known as "junk food" are the ones with the poorest school performance. After the completion of the project, questionnaires delivered to university students in order to explore their views towards the use of R in teaching Statistics. A Descriptive Analysis and a Hierarchical Classification were realized in order to investigate students' views and the results are presented in the following paragraphs [7].

2.1. The Descriptive Analysis

In total of 75 university students that took place in the research, a 44.00% were men and a 56.00% were women (Table 1).

Table 1. *Sex.*

	n	%
Men	33	44.00%
Women	42	56.00%
Total	75	100.00%

To the item "I enjoyed working with R in the course of Statistics" 12.12% of the men and 4.76% of the woman answered "Strongly disagree". 9.09% of the men and 4.76% of the women answered "Disagree". 24.24% of the men and 7.14% of the women answered "Neither agree nor disagree". 33.33% of the men and 61.90% of the women answered "Agree" while 21.21% of the men and 21.43% of the women answered "Strongly agree" (Table 2).

Table 2. *I enjoyed working with R in the course of Statistics.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	4	12.12%	2	4.76%	6	8.00%
Disagree	3	9.09%	2	4.76%	5	6.67%
Neither agree nor disagree	8	24.24%	3	7.14%	11	14.67%
Agree	11	33.33%	26	61.90%	37	49.33%
Strongly agree	7	21.21%	9	21.43%	16	21.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "I could understand the use of R" 3.03% of the men answered "Strongly disagree" while no woman answered "Strongly disagree". 33.33% of the men and 19.05% of the women answered "Disagree". 6.06% of the men and 9.52% of the women answered "Neither agree nor disagree". 42.42% of the men and 45.24% of the women answered "Agree" while 15.15% of the men and 26.19% of the women answered "Strongly agree" (Table 3).

Table 3. *I could understand the use of R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	0	0.00%	1	1.33%
Disagree	11	33.33%	8	19.05%	19	25.33%
Neither agree nor disagree	2	6.06%	4	9.52%	6	8.00%
Agree	14	42.42%	19	45.24%	33	44.00%
Strongly agree	5	15.15%	11	26.19%	16	21.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "R is important" 3.03% of the men and 4.76% of the women answered "Strongly disagree". 9.09% of the men answered "Disagree" while no woman answered "Disagree". 12.12% of the men and 11.90% of the women answered "Neither agree nor disagree". 48.48% of the men and 50.00% of the women answered "Agree" while 27.27% of the men and 33.33% of the women answered "Strongly agree" (Table 4).

Table 4. *R is important.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	3	9.09%	0	0.00%	3	4.00%
Neither agree nor disagree	4	12.12%	5	11.90%	9	12.00%
Agree	16	48.48%	21	50.00%	37	49.33%
Strongly agree	9	27.27%	14	33.33%	23	30.67%
Total	33	100.00%	42	100.00%	75	100.00%

To the item “R is easy to use” 6.06% of the men and 4.76% of the women answered “Strongly disagree”. 27.27% of the men and 19.05% of the women answered “Disagree”. 12.12% of the men and 16.67% of the women answered “Neither agree nor disagree”. 30.30% of the men and 26.19% of the women answered “Agree” while 24.24% of the men and 33.33% of the women answered “Strongly agree” (Table 5).

Table 5. *R is easy to use.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	2	6.06%	2	4.76%	4	5.33%
Disagree	9	27.27%	8	19.05%	17	22.67%
Neither agree nor disagree	4	12.12%	7	16.67%	11	14.67%
Agree	10	30.30%	11	26.19%	21	28.00%
Strongly agree	8	24.24%	14	33.33%	22	29.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item “I am interested in using R” 3.03% of the men and 4.76% of the women answered “Strongly disagree”. 3.03% of the men and 9.52% of the women answered “Disagree”. 21.21% of the men and 14.29% of the women answered “Neither agree nor disagree”. 63.64% of the men and 28.57% of the women answered “Agree” while 9.09% of the men and 42.86% of the women answered “Strongly agree” (Table 6).

Table 6. *I am interested in using R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	1	3.03%	4	9.52%	5	6.67%
Neither agree nor disagree	7	21.21%	6	14.29%	13	17.33%
Agree	21	63.64%	12	28.57%	33	44.00%
Strongly agree	3	9.09%	18	42.86%	21	28.00%
Total	33	100.00%	42	100.00%	75	100.00%

To the item “I felt comfortable working with R” 3.03% of the men and 4.76% of the women answered “Strongly disagree”. 12.12% of the men and 7.14% of the women answered “Disagree”. 33.33% of the men and 14.29% of the women answered “Neither agree nor disagree”. 30.30% of the men and 42.86% of the women answered “Agree” while 21.21% of the men and 30.95% of the women answered “Strongly agree” (Table 7).

Table 7. *I felt comfortable working with R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	4	12.12%	3	7.14%	7	9.33%
Neither agree nor disagree	11	33.33%	6	14.29%	17	22.67%
Agree	10	30.30%	18	42.86%	28	37.33%
Strongly agree	7	21.21%	13	30.95%	20	26.67%
Total	33	100.00%	42	100.00%	75	100.00%

To the item “I was able to work with R” 3.03% of the men and 2.38% of the women answered “Strongly disagree”. 27.27% of the men and 2.38% of the women answered “Disagree”. 6.06% of the men and 9.52% of the women answered “Neither agree nor disagree”. 45.45% of the men and 50.00% of the women answered “Agree” while 18.18% of the men and 35.71% of the women answered “Strongly agree” (Table 8).

Table 8. *I was able to work with R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	1	2.38%	2	2.67%
Disagree	9	27.27%	1	2.38%	10	13.33%
Neither agree nor disagree	2	6.06%	4	9.52%	6	8.00%
Agree	15	45.45%	21	50.00%	36	48.00%
Strongly agree	6	18.18%	15	35.71%	21	28.00%
Total	33	100.00%	42	100.00%	75	100.00%

To the item “R is useful to all scientists” 3.03% of the men and 4.76% of the women answered “Strongly disagree”. 6.06% of the men and 2.38% of the women answered “Disagree”. 18.18% of the men and 11.90% of the women answered “Neither

agree nor disagree". 21.21% of the men and 45.24% of the women answered "Agree" while 51.52% of the men and 35.71% of the women answered "Strongly agree" (Table 9).

Table 9. *R is useful to all scientists.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	2	6.06%	1	2.38%	3	4.00%
Neither agree nor disagree	6	18.18%	5	11.90%	11	14.67%
Agree	7	21.21%	19	45.24%	26	34.67%
Strongly agree	17	51.52%	15	35.71%	32	42.67%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "I faced difficulties in learning how to use R" 6.06% of the men and 7.14% of the women answered "Strongly disagree". 24.24% of the men and 14.29% of the women answered "Disagree". 6.06% of the men and 14.29% of the women answered "Neither agree nor disagree". 54.55% of the men and 33.33% of the women answered "Agree" while 9.09% of the men and 30.95% of the women answered "Strongly agree" (Table 10).

Table 10. *I faced difficulties in learning how to use R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	2	6.06%	3	7.14%	5	6.67%
Disagree	8	24.24%	6	14.29%	14	18.67%
Neither agree nor disagree	2	6.06%	6	14.29%	8	10.67%
Agree	18	54.55%	14	33.33%	32	42.67%
Strongly agree	3	9.09%	13	30.95%	16	21.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "I understood what I was taught about R" 3.03% of the men and 4.76% of the women answered "Strongly disagree". 3.03% of the men and 9.52% of the women answered "Disagree". 12.12% of the men and 4.76% of the women answered "Neither agree nor disagree". 63.64% of the men and 40.48% of the women answered "Agree" while 18.18% of the men and 40.48% of the women answered "Strongly agree" (Table 11).

Table 11. *I understood what I was taught about R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	1	3.03%	4	9.52%	5	6.67%
Neither agree nor disagree	4	12.12%	2	4.76%	6	8.00%
Agree	21	63.64%	17	40.48%	38	50.67%
Strongly agree	6	18.18%	17	40.48%	23	30.67%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "It was a pleasure working with R" 3.03% of the men and 4.76% of the women answered "Strongly disagree". 12.12% of the men and 11.90% of the women answered "Disagree". 42.42% of the men and 11.90% of the women answered "Neither agree nor disagree". 24.24% of the men and 28.57% of the women answered "Agree" while 18.18% of the men and 42.86% of the women answered "Strongly agree" (Table 12).

Table 12. *It was a pleasure working with R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	1	3.03%	2	4.76%	3	4.00%
Disagree	4	12.12%	5	11.90%	9	12.00%
Neither agree nor disagree	14	42.42%	5	11.90%	19	25.33%
Agree	8	24.24%	12	28.57%	20	26.67%
Strongly agree	6	18.18%	18	42.86%	24	32.00%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "I made many mistakes when working with R" 15.15% of the men and 14.29% of the women answered "Strongly disagree". 24.24% of the men and 35.71% of the women answered "Disagree". 30.30% of the men and 16.67% of the women answered "Neither agree nor disagree". 30.30% of the men and 30.95% of the women answered "Agree". No man answered "Strongly agree" while 2.38% of the women answered "Strongly agree" (Table 13).

Table 13. *I made many mistakes when working with R.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	5	15.15%	6	14.29%	11	14.67%
Disagree	8	24.24%	15	35.71%	23	30.67%
Neither agree nor disagree	10	30.30%	7	16.67%	17	22.67%
Agree	10	30.30%	13	30.95%	23	30.67%
Strongly agree	0	0.00%	1	2.38%	1	1.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "R will be helpful to my profession" 6.06% of the men and 7.14% of the women answered "Strongly disagree". 9.09% of the men and 7.14% of the women answered "Disagree". 18.18% of the men and 14.29% of the women answered "Neither agree nor disagree". 51.52% of the men and 40.48% of the women answered "Agree" while 15.15% of the men and 30.95% of the women answered "Strongly agree" (Table 14).

Table 14. *R will be helpful to my profession.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	2	6.06%	3	7.14%	5	6.67%
Disagree	3	9.09%	3	7.14%	6	8.00%
Neither agree nor disagree	6	18.18%	6	14.29%	12	16.00%
Agree	17	51.52%	17	40.48%	34	45.33%
Strongly agree	5	15.15%	13	30.95%	18	24.00%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "It will be difficult for me to apply R in my future work" 9.09% of the men and 16.67% of the women answered "Strongly disagree". 36.36% of the men and 28.57% of the women answered "Disagree". 15.15% of the men and 19.05% of the women answered "Neither agree nor disagree". 36.36% of the men and 35.71% of the women answered "Agree". 3.03% of the men answered "Strongly agree" while no woman answered "Strongly agree" (Table 15).

Table 15. *It will be difficult for me to apply R in my future work.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	3	9.09%	7	16.67%	10	13.33%
Disagree	12	36.36%	12	28.57%	24	32.00%
Neither agree nor disagree	5	15.15%	8	19.05%	13	17.33%
Agree	12	36.36%	15	35.71%	27	36.00%
Strongly agree	1	3.03%	0	0.00%	1	1.33%
Total	33	100.00%	42	100.00%	75	100.00%

To the item "I am able to teach others what I learned about R in the course of Statistics" 6.06% of the men and 7.14% of the women answered "Strongly disagree". 36.36% of the men and 19.05% of the women answered "Disagree". 9.09% of the men and 2.38% of the women answered "Neither agree nor disagree". 42.42% of the men and 40.48% of the women answered "Agree" while 6.06% of the men and 30.95% of the women answered "Strongly agree" (Table 16).

Table 16. *I am able to teach others what I learned about R in the course of Statistics.*

	Men		Women		Total	
	n	%	n	%	n	%
Strongly disagree	2	6.06%	3	7.14%	5	6.67%
Disagree	12	36.36%	8	19.05%	20	26.67%
Neither agree nor disagree	3	9.09%	1	2.38%	4	5.33%
Agree	14	42.42%	17	40.48%	31	41.33%
Strongly agree	2	6.06%	13	30.95%	15	20.00%
Total	33	100.00%	42	100.00%	75	100.00%

2.2. The Hierarchical Classification

We used are the Hierarchical Clustering in order to show the clusters of the students according to their common characteristics. This method offers the advantage of representing the centroids of the clusters on the factor levels aiming to a detailed interpretation of the differences between the clusters [8]. The Hierarchical Classification led to the

formation of five clusters which are presented in Figure 1.

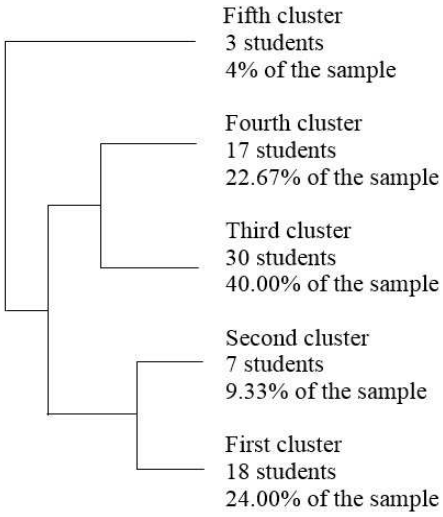

Figure 1. *The Classification Chart.*

First cluster (18 students, 24.00% of the sample)

The first cluster consists of students who answered the items "I am interested in using R", "R is useful to all scientists", "I made many mistakes when working with R", "I was able to work with R" and "R is important" with "neither agree nor disagree".

Second cluster (7 students, 9.33% of the sample)

The second cluster consists of students who answered the items "I was able to work with R", "R is easy to use", "I could understand the use of R" and "I am able to teach others what I learned about R in the course of Statistics" with "Disagree".

Third cluster (30 students, 40.00% of the sample)

The third cluster consists of students who answered the items "I was able to work with R", "It was a pleasure working with R", "I felt comfortable working with R", "R is easy to use" and "R is useful to all scientists" with "Agree".

Fourth cluster (17 students, 22.67% of the sample)

The fourth cluster consists of students who answered the items "I was able to work with R", "I felt comfortable working with R", "R is easy to use", "It was a pleasure working with R" and "I could understand the use of R" with "Strongly agree".

Fifth cluster (3 students, 4% of the sample)

The fifth cluster consists of students who answered the items "I understood what I was taught about R", "It was a pleasure working with R", "I felt comfortable working with R", "I am interested in using R" and "I faced difficulties in learning how to use R" with "Strongly disagree".

3. Conclusion

Students of the Pontifical Catholic University of Ecuador worked on a project with data from Greece in order to investigate their views towards the use of R in teaching Statistics [9]. The students reacted very positively to the use of R. 70.67% of the students answered "Agree" and "Strongly agree" to the item "I enjoyed working with R in the course of Statistics". 65.33% of the students answered "Agree" and "Strongly agree" to the item "I could understand the use of R". 80.00% of the students answered "Agree" and "Strongly agree" to the item "R is important". 72.00% of the students answered "Agree" and "Strongly agree" to the item "I am interested in using R". 64.00% of the students answered "Agree" and "Strongly agree" to the item "I felt comfortable working with R". 76.00% of the students answered "Agree" and "Strongly agree" to the item "I was able to work with R". 77.33% of the students answered "Agree" and "Strongly agree" to the item "R is useful to all scientists". 81.33% of the students answered "Agree" and "Strongly agree" to the item "I understood what I was taught about R" and 69.33% of the students answered "Agree" and "Strongly agree" to the item "R will be helpful to my profession" [10]. The Hierarchical Classification confirmed the above results. 62.67% of the students (third and fourth cluster) selected in the questionnaire the answers "Agree" and "Strongly agree" [11].

Acknowledgement

We would like to express our deepest gratitude to Prometeo Project of the Secretariat for Higher Education, Science, Technology and Innovation of the Republic of Ecuador that funded the work of Efstathios Stefos in the Pontifical Catholic University of Ecuador. We would also like to express our sincere gratitude to the Pontifical Catholic University of Ecuador and the Medical School of the University of Athens that supported the research of Efstathios Stefos and Georgia Koulianidi.

References

[1] Bellisle, F. (2004). Effects of diet on behaviour and cognition in children. *The British Journal of Nutrition,* Oct; 92 Suppl 2: S227-32.

[2] Stefos, E., Athanasiadis I., Gialamas, B. & Tsolakidis, C. (2011). The Use of New Technologies and the Project Method in Teaching Statistics: A Case Study in Higher Education. *HMS i JME*, Volume 3. 2010-2011. pp. 84-100.

[3] Stefos, E. & Papapostolou, I. (2013). *Research Methodology. Processes and suggestions.* p. 406. Rhodes: Evdimos Editions.

[4] Florence, M. D., Asbridge, M. & Veugelers, P. J. (2008). Diet quality and academic performance. *Journal of School Health*, Apr; 78(4): 209-215.

[5] Fu, M. L., Cheng, L., Tu S. H. & Pan, W. H. (2007). Association between unhealthful eating patterns and unfavorable overall school performance in children. *Journal of American Dietetic Association,* Nov; 107(11): 1935-43.

[6] Stefos, E. (2015). *Eating habits among High School students in Dodecanese. Impacts of the economic crisis in Greece.* p. 125. Rhodes: Evdimos Editions.

[7] Benzécri, P. (1992). *Correspondence Analysis Handbook*. New York: Dekker.

[8] Martin, O. (2008). *The analysis of quantitative data*, Transl. Athanasiadis, I. (pp. 86-88). Athens: Topos.

[9] Athanasiadis, I., Stefos, E., Gialamas, V. & Tsolakidis, K. (2010). New Technologies and the Project Method in Teaching Statistics: A case study in Higher Education. *Proceedings of the 27th Panhellenic Conference of Mathematical Education*. pp. 243-252. Halkida.

[10] Athanasiadis, I., Stefos, E., Gialamas, V. & Tsolakidis, C. (2011). Didactical interventions in the teaching of Statistics: Applications in Higher Education. *Euclides γ' Journal of Mathematics Education*, Vol.75. July-December 2011. pp. 35-56.

[11] Athanasiadis, I. (1995). *Correspondence Analysis and Hierarchical Classification*. New Technologies Editions, pp. 51-56.