

A PROJECT PROPOSAL ON

WHICH AUSTRALIAN HOUSEHOLDS SPEND THE MOST ON

ELECTRICITY

COMPUTER SCIENCE RESEARCH PROJECT

COMP7801

DONE BY

MR. SAISRIKAR PARUCHURI
44751575

PROJECT SUPERVISOR

DR. MARK PHILIP GRIFFIN



THE UNIVERSITY
OF QUEENSLAND
A U S T R A L I A



Create change

Table of Contents

Abstract.....	2
1. Introduction.....	2
1.1 Goals and Objectives	3
1.2 Significance of Project	3
2. Literature Review	4
2.1 Electricity Usage	4
2.2 Household/Family Structure	6
2.3 Employment	7
2.4 Location	8
2.5 HILDA Dataset.....	8
2.6 Statistical Methods and Software	9
2.6.1 Statistical Methods	9
2.6.2 Software	10
3. Methodology	10
3.1 Goals and Objectives	10
3.2 Initial Screening of Data	11
3.2.1 Household Membership (HF8).....	11
3.2.2 Total Children (G1)	11
3.2.4 Social Functioning (SCQ: A10)	11
3.2.5 Hours Per Week (Paid Employment SCQ: B19a)	12
3.2.6 Hours Per Week (Travelling to and from Place of Work SCQ: B19b).....	12
3.2.7 Annual Household Expenditure (SCQ:C5o)	12
3.4 Timeline using Gantt Chart	13
3.5 Risk Assessment.....	13
4. Occupational Safety and Health	15
5. Reference	15

Abstract

In the literature survey of electricity usage, all the references included in this paper (1-11) have explored the variation of electricity consumption by household characteristics like household size, age, IT appliances usage, working from home etc. But the references in this project (1-11) did not show which Australian households spend the most on electricity usage. The goal of this project is to find the Australian households that spend the most on electricity. This project will use data from the HILDA survey to explore the reasons for variation in these household expenses. The Household, Income and Labor Dynamics in Australia (HILDA) survey is a large, longitudinal survey that is the only survey of its kind that is nationally representative. This survey has followed approximately 17,000 people since 2001, with 15-time points of data released to date. This project uses advanced statistical methods like regression, market segmentation, spatial mapping, and scale development using factor analysis to analyze data from the HILDA survey.

1. Introduction

This project uses energy as an input to deliver the services that are needed for Australian households. Australian households represent a major group of consumers of energy resources like electricity ^[3]. Moreover, there is a need to focus on the electricity usage of the younger generation because future generations will be strongly affected by energy system changes ^[3]. Usage of electricity in households is an important political topic and an important topic in social science research ^[3]. For instance, electricity consumption seems to increase as the number of people in a household increases. House members have a crucial impact on their energy usage through behaviors such as their use of appliances ^[3]. Residential electricity usage can be explained by indirect and direct determinants ^[3]. Matthies et al. 2016 propose a system that can help explain the indirect sociodemographic and economic factors on electricity consumption by considering the direct behavioral and motivational components simultaneously.

The household members age is a relevant sociodemographic factor that is correlated with electricity consumption ^[3]. Several investigations have reported that when the number of household members in a household increases, electricity usage also increases ^[3]. Income has frequently been identified as a positive prediction of electricity consumption ^[3]. However, high

correlations between the resident's income, house characteristics, and several household members may lead to methodological problems such as multicollinearity ^[3].

To effectively target marketing campaigns, it is important to cluster the households into a manageable number of groupings so that each group can be presented differently ^[4]. Currently, the utility companies use demographic data that is house size, family size, location etc. as the basis for clustering ^[4]. The work shown makes use of electricity meter data to explore whether useful clusters can be achieved based on the household's behavior ^[4].

One stream of sustainable usage research is directed towards household consumption patterns related to environmental patterns ^[6]. Household expenditures are verified in a life cycle context, by identifying the environmental impact ^[6]. Mainly due to the availability of the data and quality, this project concentrates on urban households, and on purchased energy. This project evaluates panel data on household expenditure rather than aggregate income. In the following, we have explained the goals, objectives, and significance of the project.

1.1 Goals and Objectives

The aim of this project is to identify the Australian households that spend the most on electricity bills and explore the reasons for variation of the households in electricity consumption by using variables from the Household, Income and Labor Dynamics in Australia (HILDA) survey. This project considers household size, total children ever had, how often do you take care of your grandchildren, hours per week for paid employment, hours per week traveling to and from a place of employment, annual household expenditure on electricity bills, and where the household member works.

This project uses the above variables to explore the variation in electricity usage. Finally, with this research project, we will be able to find the Australian households that spend the most on electricity bills.

1.2 Significance of Project

This research project will provide information on the households that spend the most on electricity bills. The above data can be useful for end users as well as the

electricity providers. The end users can know the variation in household consumption as a function of an electricity usage on electric appliances, household size etc. Customers can also know whether they are getting a good electricity deal by comparing their electricity bills with other electricity providers. This research is useful for the energy providers so that they can know the usage of households and specifically target the households with good deals on electricity services and to improve the overall electricity network.

2. Literature Review

2.1 Electricity Usage

Extensive research has been done on various determinants of electricity usage. However, how specific socio-demographic, attitudinal determinants and behavioral influence residential electricity usage are still scarce [3]. Matthies et al. [3] used hierarchical regression analysis to systematically investigate the household engagement in electricity saving along with a wide range of other measures in some sample households. Special attention was given to households with young people and children by analyzing the influence of the number of teenage people on electricity usage in a path model [3]. The researchers suggest that the use of behavioral information provides a more detailed information of the condition of electricity usage [3]. Matthies et al. [3] aimed to explain the indirect influences of economic factors and socio-demographic variables on electricity usage by considering motivational components and behavioral simultaneously.

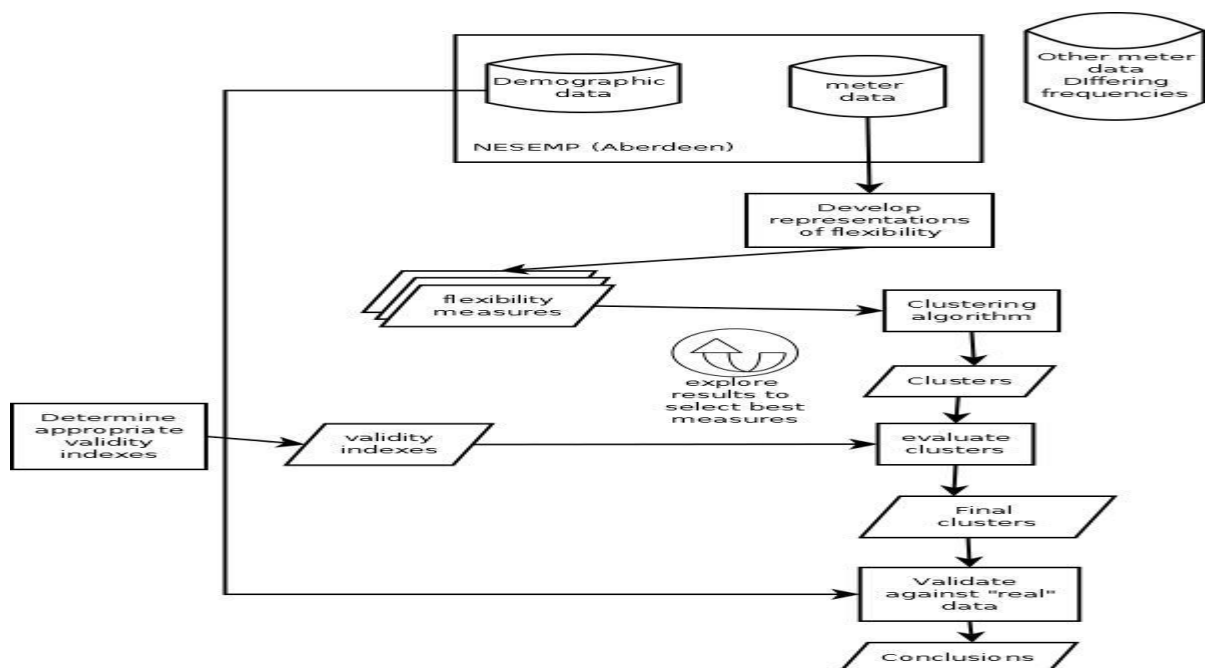
Matthies et al. [3] aim is to explain the indirect influences of economic factors and sociodemographic on electricity usage by find that past purchasing behaviors and present usage behaviors. They assume that other activities would indirectly provide information about these behaviors. For example, the time residents spend at home was expected to increase the use of appliances, such as more use of warm water and frequent cooking [3]. First, they examined the influence of indirect factors such as income, number of adolescents on electricity consumption and number of residents [3].

Second, they examined influences of income, adolescents, and house characteristics on electricity usage would be mediated by the behaviors and activities

[3]. Differences in electricity usage are primarily due to differences in purchasing and usage behaviors. Their research indicates that the adolescents in some households led to higher electricity usage because adolescents frequently use IT appliances that were positively correlated with high household electricity usage [3]. Finally, their research suggests that having adolescents in a household led to higher electricity usage because the adolescents spend the more time at home [3].

Generally knowing, how a household varies their regular consumption of electricity is useful for an organization to allow accurate targeting of behavior modifications with the aim of improving the quality of the electricity network. The variability of daily activities in a household is one possible way for the household to accept incentives to modify their behavior [4]. To evaluate the variability measures of a household they validate the number of clusters indexes [4]. These indices are varying with several clusters, quality of attributes and the number of attributes [4]. The Cluster Dispersion Indicator (CDI) and the Davis-Boulden Indicator (DBI) are taken into consideration for the household behavior variability [4].

To validate the household variability indices by the CDI and Davis-Boulden Indicator they have taken 180 households monitored over a year at an interval of 5 minutes. The time is taken from the peak electricity usage period that is 4pm to 8pm [4].



The authors have adopted the above approach to cluster the households according to the household's electricity usage ^[4].

2.2 Household/Family Structure

In today's culture, the electricity consumption is dependent on the household size, activities of the household, and age of the person (for example teenagers, kid's etc.) ^[3].

Generally knowing, the more time the people spend at home greater the electricity usage. Some people work from home at the time they will be using air-conditioners and other IT appliances, so the electricity bills will be higher for those households. Most of the households use IT appliances like refrigerators, air-conditioners, microwaves, heaters etc. for their daily needs. Matthies et al. ^[3] have considered the number of residents in a house and the time residents spend at home as indirect influences on electricity consumption.

M. Lenzen et al. ^[3] has discussed that socioeconomic-demographic factors generally have similar influences on energy requirements (that is age and household size). In some household's people maintain electric cars and they spend considerable electricity on charging their cars. Speidel et al. ^[10] say that in households 55% of electric vehicles are charged. If a household resident uses a vehicle for long distances, then the person must charge several times which results in higher electricity bills ^[10].

Lenzen et al. ^[8] describes strategies to reduce the household electricity usage for that he has calculated the Sydney household's electricity consumption for that he has considered household size, age, income, and degree of urbanity, lifestyle. All the above-mentioned household characteristics vary significantly from one household to another ^[8].

2.3 Employment

Employment is an important part of the Australian households and almost all the references in this project (1-11) has considered income to analyze the electricity usage. Matthies et al. ^[3] examined the impact of determinants such as income, number of residents, number of adolescents on electricity usage. In today's world because of technology several household residents work from home (For example Facebook declared the employees of Facebook can work from home).

If the household residents work from home depends on the climatic conditions household residents will be using air-conditioner or heater during the work, so the electricity bills are higher for those households. The authors have considered direct impacts like a number of IT appliances, hours of use of IT appliances, refrigerators etc. and indirect influences like a number of residents in a house and amount of time residents spend at home to determine electricity consumption ^[3].

Lenzen et al. ^[6] examined most of the energy in energy technology allowing for large differences across countries. Some residents in the household are away due to work and others spend the most on working and traveling to and from work, households spend less on electricity bills. Some start-up companies work in garage depending on the field (For instance, software company they must set up an environment. Also, if more employees are working in the garage then the electricity bills will be higher).

Pears et al. ^[1] predict the future energy services by evaluating energy usage on different aspects like household electricity usage, an industry which includes mining and agriculture, transport, lifestyle. Pears et al. ^[1] have discussed how the Australian household's consumption varies in different fields like households, mining, agriculture, transport, and lifestyle. Lenzen et al. ^[8] research is on the relationship between energy consumption, household members size, age, income, and degree of urbanity.

2.4 Location

Location is an important factor in analyzing the Australian household's electricity consumption. Based on the location electricity bills will be higher or lower (For instance, Queensland electricity bills are higher compared to New South Wales electricity bills ^[2]). South Australia has highest annual electricity bills compared to New South Wales, Victoria, and Queensland states in Australia ^[2]. In every location, it always depends on the electricity provider because in some locations electricity providers will charge higher and, in other locations, electricity providers will charge lower rates per unit.

Household members spend the most on traveling to another place due to work, school, college etc. The information confirms most charging is handled at home locations (55%) and business ^[10]. However, charging stations are used for 33% of charging events ^[10]. If the household resident travel for long distances in daily life, they must charge their electric car several times.

If the resident works from home they will be using IT appliances for a long time, the electricity will be higher for those households and in other households both the Mother and Father will be working during daytime and their kids or teenagers will be going for school or university, so they spend less time at home, so the electricity bills will be lower for those households.

2.5 HILDA Dataset

The HILDA dataset ^[12] is publicly available de-identified data. As such privacy and confidentiality of all survey participants will be respected. This project will use data from the HILDA survey to explore the reasons for variation in household expenses. The Household, Income and Labor Dynamics in Australia (HILDA) survey is a large, longitudinal survey that is the only survey of its kind that is nationally representative. This survey has followed approximately 17,000 people since 2001, with 15-time points of data released to date.

2.6 Statistical Methods and Software

2.6.1 Statistical Methods

Statistical methods are used for analyzing, summarizing, and interpreting data. Statistical methods are used in economics, agricultural sciences, and life sciences. Also, they have an important role in physical sciences for the measurement of errors (such as meteorological events) and to obtain appropriate results.

2.6.1.1 Regression

In statistical methods, regression ^[13] is used to estimate the relationship between variables. The focus is on the one or more independent variables and dependent variables. Regression analysis is used to understand how the typical values of the different variables change when the independent variables are varied.

2.6.1.2 Market Segmentation

Market segmentation ^[14] is the process of dividing the business market or a broad consumer (which consists of potential and existing customers) into sub-groups of consumers dependent on some type of the shared characteristics. In dividing the markets researchers typically look for some common characteristics (including common interests, similar lifestyles, shared needs and similar demographic profiles). The aim of the segmentation is to identify the segments that have high yields (that means the segments that are more likely to be profitable) so these segments can be selected for special attention.

2.6.1.3 Spatial Mapping

Spatial mapping ^[17] includes any of the formal techniques which study entities using their topological, geographic properties. Spatial analysis has a variety of techniques (but many of them are still in development) by using different analytical approaches. Spatial analysis is applied to the structure at the human scale in the analysis of geographic data.

2.6.1.4 Scale Development using Factor Analysis

Factor analysis ^[15] is used to define the variability among observed and correlated variables in terms of a potentially low number of variables that are unobserved (called factors). The variables that are observed are modeled as linear combinations of potential factors and error terms. The main aim of factor analysis is to find the independent variables that are latent.

2.6.2 Software

2.6.2.1 Rstudio for this Project

This project uses Rstudio ^[16], it is an integrated development environment for the R programming language. Mainly Rstudio is used for graphics and statistical computing.

3. Methodology

3.1 Goals and Objectives

The aim of this project is to identify the Australian households that spend the most on electricity bills and explore the reasons for variation of the households in electricity consumption by using the following variables from Household, Income and Labor Dynamics in Australia (HILDA) survey.

- Household Membership (HF8)
- Total children ever had (G1)
- How often do you take care of your grandchildren(G34a)?
- Social Functioning: Physical/emotional problems interfered with social activities (SCQ: A10)
- Hours per week – Paid Employment (SCQ: B19a)
- Hours per week – Travelling to and from a place of paid employment (SCQ: B19b)
- Annual household expenditure - electricity bills, gas bills and other heating fuel (SCQ: C5o)

3.2 Initial Screening of Data

3.2.1 Household Membership (HF8)

Household membership describes the household size, the members of the household and age. Also, it will decide whether the current household member is currently away or staying at home because of the household members and the household members activity at home estimates the consumption of electricity. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely the household membership variable is relevant to the electricity usage.

3.2.2 Total Children (G1)

This variable describes the couple or married people has how many children so that if they have children and depends on their age if they are very small (like kids, teenagers) there is a chance of spending more money on electricity bills. If they have few children, they spend less money on electricity bills. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely the total children variable is relevant to the electricity usage.

3.2.3 How Often Do You Take Care of Your Grandchildren (G34a)

Initial screening for this variable is to know how often household resident take care of their grandchildren whether once a week, fortnight, month, every day depends on time, if they often take care of their grandchildren they must spend the most on electricity bills. If they take care of their grandchildren like once a month, 6 months, 1 year then they will spend less on electricity bills. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely this variable is relevant to the electricity usage.

3.2.4 Social Functioning (SCQ: A10)

The social functioning variable discusses the people who actively participate in social activities. The household's members who participate

actively in social functions they spend less time in households, and other household members, if they are suffering from health problems many of their relatives might come to his house, so the electricity consumption varies. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely this variable is relevant to the electricity usage.

3.2.5 Hours Per Week (Paid Employment SCQ: B19a)

This variable is very important in the variation of electricity consumption. Hours per week will provide information about how many hours the household residents work. In some households, all the residents will work and spend less time at home or some people work from home and they will be using IT appliances like air-conditioners, laptops, tablets etc. so there is a huge electricity usage. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely this variable is relevant to the electricity usage.

3.2.6 Hours Per Week (Travelling to and from Place of Work SCQ: B19b)

The household members spend time on traveling to and from a place of work. If the work location is far from there household it takes a long time for traveling to and from work and if they work for more hour's outside electricity bills will be less in their households. Some people they will be currently away from the house because of the work. In few households, they work from home, so their activity will be more in their household. However, the above decision is an assumption to validate this variable this project uses regression analysis on a one-time basis on how likely this variable is relevant to the electricity usage.

3.2.7 Annual Household Expenditure (SCQ:C5o)

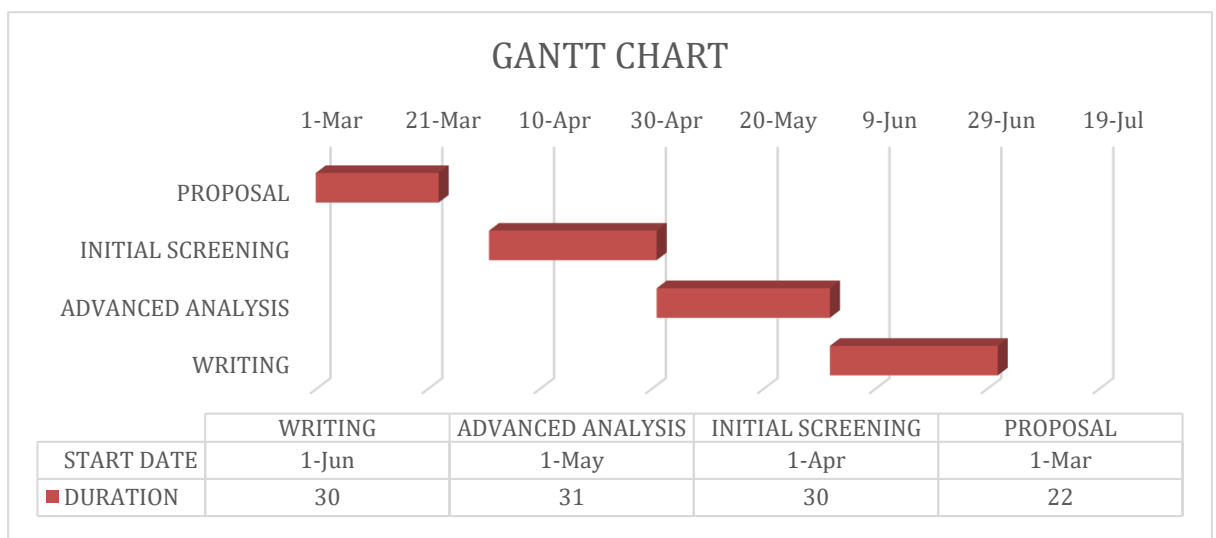
Annual household's expenditure variable provides information on household electricity bills so we can analyze the household that spend the most on electricity bills. However, the above decision is an assumption to validate

this variable this project uses regression analysis on a one-time basis on how likely this variable is relevant to the electricity usage.

3.3 Advanced Statistical Analysis

This project uses advanced regression, market segmentation using K means clustering, spatial mapping, and scale development using factor analysis statistical methods to analyze variables that are related to the electricity usage from Household, Income and Labor Dynamics in Australia (HILDA) survey.

3.4 Timeline using Gantt Chart



3.5 Risk Assessment

Risk Id	Risk	Likelihood	Impact	How I will Manage
R1	Managing workload with other course assignments	High	High	I will work every day minimum of 1 hour on each assignment to

				overcome the workload
R2	Part-time workload	High	Medium	If there is more study work, then I will try to reduce my job hours to work most on assignments
R3	My supervisor is not available during the holiday period	High	Low	Before the holiday period, project supervisor will allocate the work that I must do in the holiday period and we can communicate through emails during the holiday break
R4	My health conditions	Low	High	I will apply extension for my project submission
R5	During the holiday period, I am not available	Low	Medium	If I am not available during the holiday period I will be working on my project at my home location
R6	If I take a semester break	Low	Low	I will still be working on the project during semester break
R7	Short time for submission	Medium	High	I will try to figure out the important information in

				the project for completion
--	--	--	--	----------------------------

4. Occupational Safety and Health

I will be working on this research project to find the Australian households that spend most on electricity by using Household, Income and Labor Dynamics in Australia (HILDA) dataset which has de-identified data. While working on this project I will be completely safe as I am not dealing with any dangerous equipment. I am completely safe while working on this project.

5. Reference

1. A. K. Pears, "Imagining Australia's energy services futures," *Futures*, vol. 39, no. 2, pp. 253-271, 2007/03/01/ 2007.
2. C. BLUE. (2018, January 15). What is the average electricity bill. Available: <https://www.canstarblue.com.au/energy/electricity/average-electricity-bills/>
3. E. Matthies, M. Nachreiner, and H. Wallis, "Adolescents and electricity consumption; Investigating sociodemographic, economic, and behavioral influences on electricity consumption in households. (Special Section on the European Union: Markets and Regulators)," *Energy Policy*, vol. 94, pp. 224-234, 2016.
4. I. Dent, T. Craig, U. Aickelin, and T. Rodden, "An Approach for Assessing Clustering of Households by Electricity Usage," 2014.
5. K. Atalay, S. Whelan, and J. Yates, "Housing Wealth and Household Consumption: New Evidence from Australia and Canada," E. School Of, Ed., ed, 2013.
6. M. Lenzen, M. Wier, C. Cohen, H. Hayami, S. Pachauri, and R. Schaeffer, "A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan," *Energy*, vol. 31, no. 2, pp. 181-207, 2006/02/01/ 2006.
7. M. Lenzen, "Energy and greenhouse gas cost of living for Australia during 1993/94," *Energy*, vol. 23, no. 6, pp. 497-516, 1998/06/01/ 1998.
8. M. Lenzen, C. Dey, and B. Foran, "Energy requirements of Sydney households," *Ecological Economics*, vol. 49, no. 3, pp. 375-399, 2004/07/01/ 2004.
9. R. Kellogg, "Efficiency in energy production and consumption," S. Borenstein and J. Perloff, Eds., ed: ProQuest Dissertations Publishing, 2008.

10. S. Speidel and T. Bräunl, "Driving and charging patterns of electric vehicles for energy usage," *Renewable and Sustainable Energy Reviews*, vol. 40, pp. 97-110, 2014/12/01/ 2014.
11. Z. M. Chen and G. Q. Chen, "An overview of energy consumption of the globalized world economy," *Energy Policy*, vol. 39, no. 10, pp. 5920-5928, 2011/10/01/ 2011.
12. N. Watson and M. P. Wooden, "The HILDA survey: a case study in the design and development of a successful household panel survey," *Longitudinal and Life Course Studies*, vol. 3, no. 3, pp. 369-381, 2012.
13. M. J. Crawley, *Regression*. Chichester, UK: Chichester, UK: John Wiley & Sons, Ltd, 2012, pp. 449-497.
14. R. E. Frank, W. F. Massey, and Y. Wind, *Market segmentation* / Ronald E. Frank, William F. Massey, Yoram Wind. Englewood Cliffs, N.J.: Englewood Cliffs, N.J. : Prentice-Hall, 1972.
15. R. L. Gorsuch, *Factor analysis* / Richard L. Gorsuch, 2nd ed.. ed. Hillsdale, N.J.: Hillsdale, N.J. : Erlbaum Associates, 1983.
16. RStudio Team (2016). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.
17. R. P. Cabeen, M. E. Bastin, and D. H. Laidlaw, "A Comparative evaluation of voxel-based spatial mapping in diffusion tensor imaging," *NeuroImage*, vol. 146, pp. 100-112, 2017/02/01/ 2017.