Influence of demographic variables on uptake of domestic solar photovoltaic technology.

Sommerfeld, J., Buys, L., Mengersen, K., & Vine, D.

Sommerfeld, Jeffrey, Buys, Laurie, Mengersen, Kerrie, & Vine, Desley (2017) Influence of demographic variables on uptake of domestic solar photovoltaic technology. Renewable and Sustainable Energy Reviews, 67, pp. 315-323.. DOI: https://doi.org/10.1016/j.rser.2016.09.009

Title: Influence of demographic variables on uptake of domestic solar photovoltaic technology

Author names and affiliations: Jeff Sommerfeld, Laurie Buys, Kerrie Mengersen, Desley Vine

Jeff Sommerfeld, Queensland University of Technology, Brisbane, Australia
Prof Laurie Buys, Queensland University of Technology, Brisbane, Australia
Prof Kerrie Mengersen, Queensland University of Technology, Brisbane, Australia
Dr Desley Vine, Queensland University of Technology, Brisbane, Australia

Corresponding author: Jeff Sommerfeld j.sommerfeld@qut.edu.au

Present address: School of Design, Queensland University of Technology, 2 George Street, Brisbane, Australia

Abstract:
In Australia during the past decade there has been a significant transformation of the electricity demand and supply sector. In five years from 2008 to 2013 the number of Australians installing solar photovoltaic (PV) technology grew from 8000 to more than one million. Governments in Australia used a range of policy incentives such as feed-in tariffs (FiTs) to encourage the uptake of solar PV and this had a range of consequences. Solar PV technology has transformed the residential consumer electricity market providing some consumers with greater choice in demand and supply of their power. This study contributes to the growing understanding of the role that demographic factors play in household uptake of solar PV technology. Based on a review of relevant literature and a multi-phased statistical analysis of more than 2 million people in south-east Queensland over five years, the paper highlights the complex interplay between socio-economics and household uptake. The paper identifies key demographic variables and quantifies their relative influence, and provides new insights into the role of age in solar PV uptake. This more nuanced explanation of the socio-economic variables influencing solar PV uptake offers an opportunity to more effectively and efficiently shape future policies and incentives.

Key words: domestic consumers, consumer change, solar photovoltaic (PV), demographic variables, feed-in tariff (FiT); classification and regression tree (CART); boosted regression tree (BRT)

Abbreviations: Australian Bureau of Statistics (ABS); classification and regression tree (CART); boosted regression tree (BRT); feed-in tariff (FiT); greenhouse gas (GHG); Mandatory Renewable Energy Target (MRET); photovoltaic (PV); Photovoltaic Rebate Program (PVRP); Residential Monitoring to Decrease Energy Use and Carbon Emissions in Europe (REMODECE)

1. Introduction:

For most of the past century the dominant paradigm of the electricity demand and supply sector has been a provider technology-push versus consumer demand-pull which has defined traditional electricity market participants [1]. However, in recent years a demand-pull for greater environmental, economic and social sustainability from government and the electricity sector has altered its traditional linear demand and supply dichotomy. Since the 1990s in Australia, state and federal governments have progressively been devolving from centralized monopoly electricity markets, encouraging deregulation and removing often hidden subsidies [2, 3]. Coinciding with government deregulation of the electricity market has been the emergence of government policies and consumer preferences for energy from renewable sources that produce lower greenhouse gas (GHG) emissions delivering better environmental outcomes. Renewable energy, however, has been more costly than traditional sources of electricity. An unfortunate result of government policies encouraging electricity industry transformation towards renewables, such as solar photovoltaic (PV), is increased costs of electricity for consumers [2, 3]. In Australia, policies such as solar PV feed-in tariffs (FiTs) are added to the cost of electricity for all consumers and is a factor contributing to cost increases of more than 100 per cent in less than a decade [4].

The convergence of electricity sector deregulation and policies that promote solar has resulted in major market upheavals with significant economic and social impacts [3, 5, 6]. As renewable energy is continuing to be promoted as the future for global energy supply, it is important for regressive consequences of policy to be mitigated. Understanding consumer motivation and decision making regarding solar PV uptake is important to ensure negative consequences including equity issues are able to be mitigated in future solar policy initiatives.

The consumer decision to acquire a solar PV system is complex requiring information that the average consumer is unlikely to have in early stages of new technology [7]. Yet, research into the uptake of energy technology by consumers is considered to be narrowly focused and does not address the full range of external factors that influence decision making [8]. In particular, although the role of socio-economic factors has been suggested [2-6] the complement of relevant variables and their interplay have not yet been fully explored. The aim of this study is to provide a more comprehensive assessment of key demographic variables and their comparative influence on solar PV uptake. This is achieved by first evaluating policy drivers of consumer change, then conducting a population-based statistical analysis of demographic drivers of residential customer solar PV uptake. The target population is taken to be the greater Brisbane metropolitan region in Queensland, Australia, during the period from 2010 to 2014. A multi-phased analytic approach is adopted, comprising exploratory analyses to identify a suite of potential variables, followed by decision tree models to capture the relative importance of these variables and their potentially complex interactions. The need and importance of such research to track the influence of consumer behaviour and new trends in technology was a key finding of a review of the Residential Monitoring to Decrease Energy Use and Carbon Emissions in Europe (REMODECE) project [9].

1.1. Policy drivers of consumer change

Much of the electricity market transformation has been driven by government policy. Some of this transformation relates to policies on improved labelling and energy efficiency of appliances to increase the efficient use of energy, whilst generous government subsidies for domestic solar hot water systems and solar PV systems were a major factor encouraging consumer uptake of this technology [10]. In Australia, there have been three key periods in the evolution of solar PV technology, government policy and consumer behaviour since 2001. During the period from 2001 to 2008 policy focussed on solar hot water systems and early incentives for solar PV. From 2008 to 2012 government policies encouraged a rapid uptake of solar PV while from 2012 onwards these policies have been wound back or discontinued. As a result, the dynamics of the traditional push-pull paradigm of transmission and distribution of the electricity market has transformed with consumers becoming producers and contributing to a demand-pull for technology.

In 2001, the Commonwealth of Australia introduced the Mandatory Renewable Energy Target (MRET) scheme to encourage investment in renewable energy technologies [11]. During this period the Australian Government provided rebates to householders who acquired solar PV systems under the Photovoltaic Rebate Program (PVRP) which provided a fixed upfront incentive of about $5000 to reduce the capital cost of solar PV technology [6]. From 2007 most States and Territories commenced programs that offered the owners of small-scale solar PV installations generous FiTs for electricity generated [5] until these began to be reduced or concluded from 2012.

Table 1 summarises the growth of domestic solar hot water and solar photovoltaic (PV) systems in Australia since 2001. Commonwealth and State government policies and subsidies have been implemented at several stages of the solar energy production chain in Australia. During the period 2000-2008 the bulk of solar growth was in the form of solar thermal systems used for domestic water heating [10]. Since 2008 there has been a rapid uptake of small-scale solar PV systems on household rooftops. In five years from 2008 to 2013 the number of Australians installing solar photovoltaic (PV) technology grew from 8000 to more than one million [12]. In 2007, solar PV systems represented 9.6 MW of a 50,000MW power grid and in just four years this had increased by 100-fold to 1031 MW [4]. During the period from 2008 to 2012 FiTs were provided in most jurisdictions in Australia. Table 1 shows the uptake of solar PV during this period and the subsequent decrease in demand for solar PV once FiTs were reduced or concluded.

Table 1. Small Scale Solar Installations in Australia 2001 to 2014

Installation year	Solar PV systems	Solar water heaters
2001	118	10,075
2002	251	21,839
2003	664	28,653
2004	1,089	30,991
2005	1,406	33,964
2006	1,115	35,924
2007	3,480	50,977
2008	14,064	85,385
2009	62,916	194,695
2010	198,208	127,093
2011	360,745	105,050
2012	343,320	69,466
2013	196,429	55,189
2014	28,788	6,801
Grand total	1,212,593	856,102

Source: Clean Energy Regulator 2014

The transformation of the residential consumer electricity market in the past decade has resulted in some consumers having greater choice in demand and supply of their power by being both consumers and producers of electricity, sometimes coined prosumers [see for example, 13, 14]. This revolutionary change is a major paradigm shift in electricity demand and supply which will have ongoing policy and regulatory implications for some consumers. However, this dynamic shift has not been universally enjoyed with a growing number of consumers experiencing energy poverty [2, 5, 6, 15]. Much of the research to date has focused on either the reasons for adoption or non-adoption of renewable energy [16] and not on the widening of the social divide between consumers which has been found to be an unfortunate outcome of these policies [2, 3].

1.2. Evaluation of previous policy

Australian governments have adopted policies encouraging the rapid uptake of small rooftop solar PV systems through subsidies and FiTs. Although 11 per cent of the Australian population (about 2.6 million people) now use solar for their electricity, there remains almost 89 per cent of the Australian population who have not participated to date despite the financial incentives to uptake solar PV. An evaluation of the Australian Government solar PV incentive programs focusing on uptake of renewable energy, industry impact, emissions abatement and equity concluded that the programs were ineffective and costly and resulted in equity issues [2, 6]. In Germany, a significant adopter of domestic solar PV, Grosche and colleagues [15] identified similar equity concerns with solar PV policy that resulted in the transfer of income from lower socio-economic groups to higher socio-economic groups.

In Australia, infrastructure costs associated with maintaining network stability were found to be greater for small scale solar PV due to the multiple interfaces with the network for individual household PV units as opposed to the lower costs and economies of scale of the small number of interfaces required by large scale solar generation [4, 5]. Additional network costs were passed onto all consumers through higher network charges. The socially regressive aspects of these policies were that lower socio-economic groups who could not afford solar PV paid for network costs incurred by customers from higher socio-economic groups [2, 5, 6, 15]. In addition, home ownership was found to be the key factor influencing the ability of consumers to install solar PV. Critics of these policies identified that people who lived in rented accommodation or in apartments were unable to install solar PV with the result of a further transfer of income from lower socio-economic groups to higher socio-economic groups [3].

Reviews of past solar PV policies have identified that household financial resources influence solar PV uptake due to the upfront costs of acquiring a residential solar PV and that this may have excluded many low to medium income households from programs [3, 5, 6]. These reviews also identified that households with higher levels of education and in higher skilled occupations were more likely to find it easier to access information on residential solar PV systems highlighting other equity issues in the design of solar policy [3, 5, 6]. Despite the awareness of the general importance of socio-economic variables, there remains a knowledge gap about their individual and combined impact on the effectiveness of solar policies. Moreover, other demographic variables may also be important, such as size of family and age of residents, particularly for specific population groups. Importantly, there is a lack of quantitative analyses of these associations in large cohorts exposed to new policies over a substantial period of time. The purpose of this current study is to address this gap and contribute to the understanding of the role that demographic factors play in household uptake of solar PV technology

2. Method

2.1 Study area

Within Australia, Queensland has the largest number of solar photovoltaic installations of any state. The focal region for this study, the greater Brisbane area, was selected because it has almost one-half of the total Queensland population (1,920,205) [17] and data for solar energy programs have been collected for more than a decade. The study area comprises 117 postal areas (postcodes) which is a commonly used level of data collection and analysis used by multiple agencies in Australia. The solar PV and demographic information is publicly available at the postcode scale, since it is large enough that individual consumers cannot be identified. Although research at this scale does not capture the individual socio-economic profile and behaviour of a consumer, it was the best available socio-economic status (SES) measure as it captured the economic and social resources of the postal areas [6].

The data on the uptake of solar PV in south east Queensland (Table 2) showed there were 533 solar PV installations as at July 2008 when high solar FiTs ($0.44 per kWh) were legislated by the State Government. By July 2012 when $0.44 solar FiTs ended, this had increased to 157,849. Since July 2012 when both legislated and industry-based solar FiTs decreased to $0.06 or $0.08 per kWh, there has been a continuing growth in the numbers of solar PV installations.

Table 2. Domestic Solar PV South East Queensland 2008 to 2014

Installation year	Solar PV systems
As at July 2008	533
As at July 2009	5947
As at July 2010	27,100
As at July 2011	83,188
As at July 2012	157,849
As at July 2013	229,439
As at July 2014	264,807

Source: Energex 2014

A comparative tracking of small scale solar system uptake in Australia and domestic solar PV uptake in south east Queensland is illustrated in Figure 1, based on figures from Tables 1 and 2.

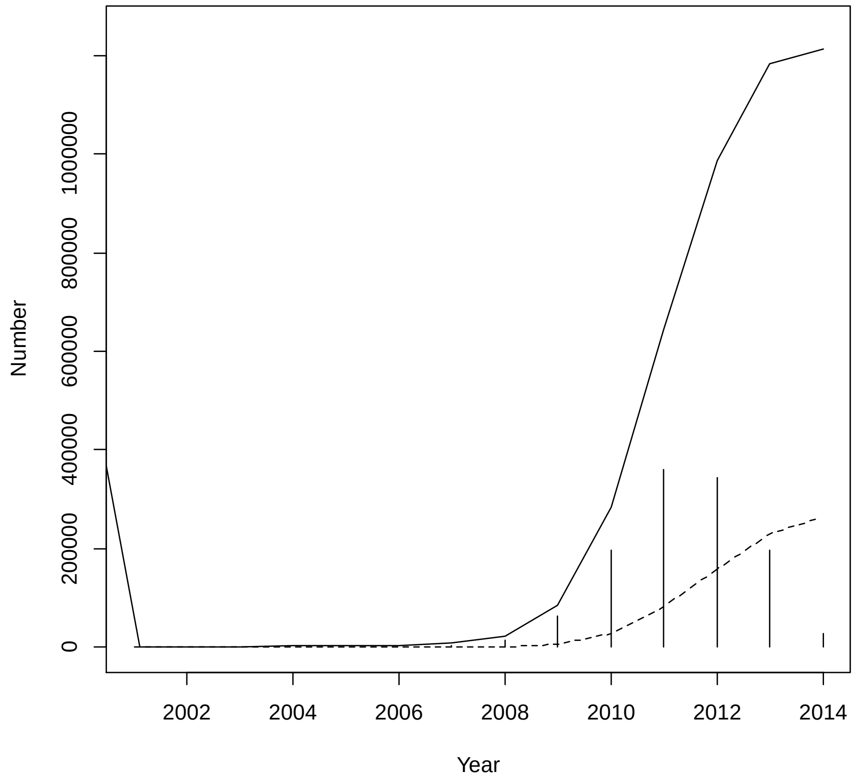

Figure 1. Small scale solar PV system installations in Australia and south east Queensland. The vertical lines indicate the annual number of installations and the solid line is the respective cumulative sum (Table 1). The dotted line is the cumulative sum of domestic solar PVs in the focal study region (Table 2).

2.2 Data acquisition

Data on solar installation were obtained from the Australian Government Clean Energy Regulator (AGCER) [18] and local electricity distributor Energex which was then correlated to postal areas using postcode information from the ABS. The AGCER data accumulated solar installations annually over a period of 12 years from 2001 to 2012. The Energex data provided information for the 117 postal areas from 2008 to 2014 which could be cross-validated with the AGCER data. The demographic data for this research were obtained from the Australian Bureau of Statistics (ABS) from the 2011 Census [17]. Table 3 shows the selected demographic information used for this study which includes: median weekly income; median mortgage payments; median weekly rent; population; dwelling type; ownership status; number of bedrooms and education status (university or tertiary).

Table 3. Explanatory variables

Socio-economic variables	Definition (ABS)
People	Total number of persons in the postal area
Families	Two or more persons, one of whom is at least 15 years of age, who are related by blood, marriage (registered or de facto), adoption, step or fostering, and who are usually resident in the same household.
Income	Gross income from all sources
Education	Number of persons with a university or tertiary qualification
Over 55 years	Persons aged over 55 years old
Over 65 years	Persons aged over 65 years
Owners	Own a dwelling outright or with a mortgage
Renters	Renting a dwelling
Mortgage	Housing loan repayments being paid on a monthly basis by a household to purchase the dwelling
Rent	Dollar amount of rent paid by households on a weekly basis for the dwelling
Private homes	Number of all private dwellings
Houses	House which stands alone in its own grounds separated from other dwellings by at least half a metre
Units	Includes flats, units and apartments - dwellings that do not have their own private grounds and usually share a common entrance foyer or stairwell
Duplexes	Semi-detached dwelling including terrace house and townhouses - dwellings that have their own private grounds and no other dwelling above or below them
Three bedrooms or more	Occupied private dwellings with three or more bedrooms

Source: ABS Census Directory http://www.abs.gov.au/ausstats/abs@.nsf/0/4B6D4A6E729E8275CA25720900078321?opendocument accessed 13 March 2015

2.3 Statistical analyses

A multi-phased analytic approach was adopted, comprising exploratory analyses at the first stage and statistical modelling at the second stage [19-21]. In the first stage, the exploratory analyses involved extraction of stratified sets of postcodes from the census data for each demographic variable of interest, and construction of tables with corresponding annual solar PV uptake figures. The stratified sets comprised 15 postcodes with the five largest, five middle and five smallest values of the respective variable; the middle values were those spanning the median. These summary statistics were then examined for consistency of trends and compared with previous policy analysis and socio-economic conclusions on uptake of solar PV.

In the second stage, the full dataset comprising the selected demographic variables and annual uptake data was analysed using two types of decision trees, namely classification and regression trees (CART) and boosted regression trees (BRT). The aim was to provide a simple to understand representation of the nuanced relationship between the set of demographic factors (explanatory variables) and the probability of an individual taking up solar PV (the response variable). Here, the response probability was calculated for individuals in a postal area as the proportion of households in the postal area that had adopted solar PV divided by the census population in that postal area.

Decision tree models are very well established statistical and machine learning techniques used for prediction and classification mechanisms that are used in a range of data mining and knowledge discovery [22]. They are often preferred over more traditional linear regression techniques since they can model complex relationships, accommodate correlated variables, easily include non-linear relationships and interactions between variables and the response, and allow predictor variables to be different types and scales of measurement [23, 24]. In the present study, other models such as linear regression and generalised additive models were considered but were discounted since the data exhibited these features.

Decision tree models aim to segment the data into a set of subgroups, where the responses within each subgroup are similar. The subgroups are formed by selecting a set of variables and a series of binary splits of these variables, until a specified stopping rule is reached. In the current study, the aim is to create subgroups of postal areas based on similar percentages of uptake of solar PV. The method first determines the variable and the splitting point that provides the best division of postal areas into two groups (higher and lower solar PV uptake), then within each group it identifies the next variable that splits into two subgroups, and so on until the stopping rule is reached. The result is usually depicted as a tree-like structure with the splitting variables as nodes in the tree, the binary splits as branches of the tree, and the subgroups of postal PVs as the terminal nodes. In the current study, the model provided the average solar PV uptake per person in the postal PVs in each of the terminal subgroups.

The CART and BRT decision trees provide complementary information. The CART is a single decision tree based on the full dataset and shows the set of most influential explanatory variables The BRT constructs many small trees based on subsets of the data and shows the relative influence of the variables [24]. It should be noted that decision trees do not provide a p-value as in traditional linear regression, but instead provide a measure of the relative importance of the variables in the model [25]. The overall fit of the decision tree model is optimised by multiple-fold cross-validation [22]. More details regarding the construction and interpretation of the trees are given in the Results section.

3. Results

3.1. Exploratory data analysis

Tables 4,5,6,7 and 8 show examples of the stratified subsets of data obtained in the first phase of analysis. Inspection of the tables in light of previous literature revealed the following results. Previous literature reported positive correlation between financial capacity and solar installations [4, 5]. However, the data in Table 4, for example, shows similar levels of solar PV across each of the groups regardless of income. Additionally, previous literature identified a relationship between education and access to information as key factors in higher uptakes of solar PV [16]. The exploratory assessment undertaken in Table 5 identified very low rates of installation of solar PV in postal areas with the highest numbers of university/tertiary educated persons, whereas the areas with the lowest levels of university/tertiary educated persons had more than double the installation rates of solar PV. However, the postal areas with the highest numbers of university/tertiary educated persons were also areas with high concentration of units and apartments.

Table 4. Exploratory Data Analysis – Solar installations and Median Weekly Income

	Rank	Postcode	Median weekly income	Solar 2010 %	Solar 2011 (%)	Solar 2012 (%)	Solar 2013 (%)	Solar 2014 %)

Top	1	4069	$2,347	3.26%	6.58%	13.10%	19.68%	22.70%
	2	4065	$2,286	2.51%	4.41%	8.15%	12.87%	14.63%
	3	4156	$2,265	4.11%	8.54%	16.97%	29.82%	35.62%
	4	4154	$2,261	2.58%	6.95%	15.76%	26.59%	31.78%
	5	4155	$2,230	6.29%	10.57%	22.00%	37.14%	43.43%
Middle	55	4165	$1,478	5.06%	11.08%	19.46%	28.17%	32.36%
	56	4101	$1,475	1.27%	2.13%	3.83%	5.65%	6.25%
	57	4173	$1,474	2.98%	6.91%	13.95%	20.39%	23.88%
	58	4179	$1,465	4.07%	7.35%	13.78%	19.81%	22.38%
	59	4159	$1,462	5.65%	10.95%	19.44%	28.03%	31.97%
Bottom	113	4303	$855	1.88%	4.91%	10.48%	14.45%	16.47%
	114	4205	$848	3.90%	10.47%	22.27%	27.55%	31.40%
	115	4183	$801	2.04%	5.34%	9.43%	14.37%	15.90%
	116	4507	$747	4.34%	12.42%	21.01%	27.36%	30.34%
	117	4184	$598	7.88%	13.34%	19.54%	25.11%	27.52%

Table 5. Exploratory Data Analysis – Solar installations and Tertiary Education

	Rank	Postcode	Education University/Tertiary	%	Solar 2010 %	Solar 2011 (%)	Solar 2012 (%)	Solar 2013 (%)	Solar 2014 %)
Top	1	4111	561	72.2%	2.07%	5.18%	12.18%	16.32%	18.91%
	2	4067	3,462	70.1%	1.11%	1.80%	3.17%	5.00%	5.67%
	3	4066	3,174	52.7%	1.22%	2.03%	3.50%	5.34%	5.98%
	4	4059	2,303	51.7%	1.30%	2.45%	4.56%	7.02%	7.93%
	5	4068	3,846	46.5%	1.46%	2.55%	4.98%	7.76%	9.08%
Middle	55	4119	251	15.6%	3.55%	8.99%	19.89%	30.42%	35.51%
	56	4055	905	15.3%	4.00%	8.55%	16.63%	25.44%	29.63%
	57	4123	667	14.8%	3.20%	7.91%	15.27%	23.01%	27.55%
	58	4054	516	14.7%	2.86%	6.22%	12.40%	18.56%	21.69%
	59	4130	338	14.3%	5.40%	10.62%	21.85%	33.09%	37.74%
Bottom	113	4303	74	5.4%	1.88%	4.91%	10.48%	14.45%	16.47%
	114	4508	321	5.3%	2.51%	6.39%	12.62%	19.04%	22.72%
	115	4114	601	5.3%	1.68%	3.90%	8.03%	11.18%	13.19%
	116	4132	363	4.9%	2.64%	7.08%	15.75%	21.58%	25.10%
	117	4184	70	4.5%	7.88%	13.34%	19.54%	25.11%	27.52%

Table 6. Exploratory Data Analysis – Solar installations and Home Ownership

	Rank	Postcode	Owners	%	Solar 2010 %	Solar 2011 (%)	Solar 2012 (%)	Solar 2013 (%)	Solar 2014 %)
Top	1	4155	288	90.8%	6.29%	10.57%	22.00%	37.14%	43.43%
	2	4156	786	88.9%	4.11%	8.54%	16.97%	29.82%	35.62%
	3	4037	2,037	85.9%	3.72%	7.68%	17.94%	28.24%	33.07%
	4	4035	6,452	84.1%	4.37%	8.82%	17.10%	26.55%	31.26%
	5	4069	9,048	83.6%	3.26%	6.58%	13.10%	19.68%	22.70%
Middle	55	4503	8,024	66.1%	2.29%	6.28%	14.02%	21.72%	26.34%
	56	4507	4,948	65.8%	4.34%	12.42%	21.01%	27.36%	30.34%
	57	4158	957	65.7%	3.14%	6.15%	11.56%	15.68%	17.90%
	58	4078	4,993	65.5%	2.99%	8.94%	17.64%	25.50%	29.20%
	59	4172	923	65.3%	2.75%	5.37%	12.50%	18.00%	20.42%
Bottom	113	4169	2,116	40.5%	0.71%	1.16%	2.19%	3.55%	4.24%
	114	4102	915	38.2%	1.79%	3.33%	6.77%	9.06%	10.46%
	115	4101	2,850	37.1%	1.27%	2.13%	3.83%	5.65%	6.25%
	116	4000	2,002	36.4%	0.33%	0.54%	0.76%	1.15%	1.43%
	117	4006	2,292	34.5%	0.18%	0.37%	0.76%	1.24%	1.43%

The summary figures in Table 6 indicate that home ownership is a factor in solar PV uptake which supports previous research and literature [6]. Tables 7 and 8 indicate a relationship between areas with the highest concentrations of persons who rented or lived in apartments and lower rates of solar PV installations. This would appear to further support a positive correlation between home ownership and PV installation.

Table 7. Exploratory Data Analysis – Solar installations and Renters

	Rank	Postcode	Renters	%	Solar 2010 %	Solar 2011 (%)	Solar 2012 (%)	Solar 2013 (%)	Solar 2014 %)
Top	1	4006	4,197	63.2%	0.18%	0.37%	0.76%	1.24%	1.43%
	2	4101	4,556	59.4%	1.27%	2.13%	3.83%	5.65%	6.25%
	3	4000	3,260	59.3%	0.33%	0.54%	0.76%	1.15%	1.43%
	4	4102	1,296	58.2%	1.79%	3.33%	6.77%	9.06%	10.46%
	5	4169	2,983	56.7%	0.71%	1.16%	2.19%	3.55%	4.24%
Middle	55	4163	1,918	33.2%	3.48%	7.60%	11.89%	17.33%	19.58%
	56	4078	2,515	33.0%	2.99%	8.94%	17.64%	25.50%	29.20%
	57	4133	1,626	33.0%	2.96%	8.86%	15.84%	22.18%	26.24%
	58	4111	112	32.7%	2.07%	5.18%	12.18%	16.32%	18.91%
	59	4018	1,163	32.0%	1.73%	5.26%	12.12%	18.42%	21.98%
Bottom	113	4125	320	13.6%	3.38%	8.86%	18.93%	27.22%	31.69%
	114	4037	306	12.9%	3.72%	7.68%	17.94%	28.24%	33.07%
	115	4035	985	12.8%	4.37%	8.82%	17.10%	26.55%	31.26%
	116	4156	83	9.4%	4.11%	8.54%	16.97%	29.82%	35.62%
	117	4155	25	7.9%	6.29%	10.57%	22.00%	37.14%	43.43%

Table 8. Exploratory Data Analysis – Solar installations and Units, Flats, Apartments

	Rank	Postcode	Units, Flats, Apartments	%	Solar 2010 %	Solar 2011 (%)	Solar 2012 (%)	Solar 2013 (%)	Solar 2014 %)
Top	1	4000	4,593	83.6%	0.33%	0.54%	0.76%	1.15%	1.43%
	2	4006	5,472	82.4%	0.18%	0.37%	0.76%	1.24%	1.43%
	3	4007	5,472	82.4%	1.05%	1.63%	2.78%	4.88%	5.75%
	4	4005	3,681	70.8%	0.64%	1.10%	2.21%	3.66%	4.04%
	5	4169	3,318	63.5%	0.71%	1.16%	2.19%	3.55%	4.24%
Middle	55	4157	277	4.2%	3.80%	8.30%	16.16%	23.52%	27.11%
	56	4305	864	4.2%	2.41%	5.91%	13.01%	19.83%	22.91%
	57	4021	146	4.0%	2.51%	6.03%	13.17%	18.71%	22.95%
	58	4173	125	4.0%	2.98%	6.91%	13.95%	20.39%	23.88%
	59	4054	175	3.6%	2.86%	6.22%	12.40%	18.56%	21.69%
Bottom	113	4164	4	0.1%	5.00%	10.25%	18.54%	29.02%	33.70%
	114	4117	0	0.0%	1.64%	6.54%	18.22%	23.60%	27.10%
	115	4130	0	0.0%	5.40%	10.62%	21.85%	33.09%	37.74%
	116	4154	0	0.0%	2.58%	6.95%	15.76%	26.59%	31.78%
	117	4155	0	0.0%	6.29%	10.57%	22.00%	37.14%	43.43%

A number of issues of significance were identified by the exploratory analysis of single and multiple variables. It re-confirmed links identified in previous literature such as the links between private home ownership and solar PV uptake. The analysis identified that tertiary education and the financial capacity of individuals did not appear to be as significant as identified by previous literature. New information revealed by the exploratory analysis was the number of bedrooms of homes and the type of dwelling were significant explanatory variables in solar PV uptake. The perceived differences between the published literature and the new information identified by the exploratory analysis justified the further examination of the information. This was undertaken using decision tree (CART and BRT) models, since these are specifically designed to identify interactions between variables [23].

3.2. CART analysis

The CART analyses were undertaken for each year of the study and generated decision trees similar to the one depicted in Figure 2. This figure shows the average (expected) probability of uptake of solar PV in a postal area, based on a set of the available demographic variables. This set, and the corresponding split points for each of the variables, are determined as part of the statistical analysis to be the most important in differentiating between solar uptake outcomes.

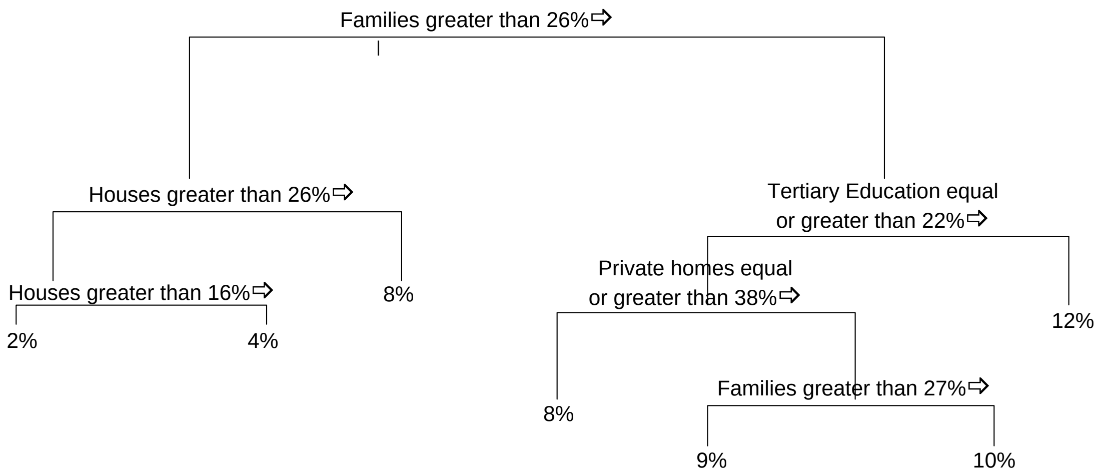

Figure 2 CART model showing predicted probability of solar PV uptake in the study area in 2014

The CART model shown in Figure 2 describes the inter-relationship between the explanatory variables in predicting the probability of solar PV uptake for a specific year, 2014. In the CART, the flow of the right branch is conditional on the node being *true* whilst the left is conditional on the node being *false*. In this example, the CART shows in 2014 *Families* (see Table 3 for definitions) was the most influential demographic variable in determining the percentage of households that had solar PV in the study region. Among those postcodes with more than 26% of *Families,* the next most influential variable in determining solar PV uptake was *Tertiary Education*. In the above example, if the percentage of *Families* in a postal area was greater than 26% and these postal areas had a level of *Tertiary Education* greater than 22% the predicted chance of an individual taking up solar PV in these postal areas was 12%. Using this example further, where the percentage of *Families* in a postal area is less than 26% the branch goes to the left and the next most significant feature in these postal areas is *Houses.* The average uptake of solar PV per person based on these two variables in these postal areas was 8%. The individuals with least chance of taking up solar PV were predicted to reside in postal areas with a small percentage (<26%) of families and a small percentage (<16%) of houses. This example demonstrates the importance of identifying the multiple contributions which may be required to effectively explain an outcome [23, 24].

Table 9 shows a summary of the important variables identified by the CART analyses for each year of the study. Here, the influence value indicates the level of the tree hierarchy at which the variable appeared. For example, the first variable used to split the data (the most important splitting variable), which appears at the top of the tree, is assigned an influence value of 1, and so on. The analyses showed that at the mid-point of the $0.44 solar FiT policy (July 2010), the most influential demographic feature in the uptake of solar PV was the percentage of families of two or more persons in the postal area. This was followed by the proportion of people aged over 55 years, then the proportion of persons who owned their own homes and privately owned dwellings.

However, by July 2012, when $0.44 solar FiTs ended in Queensland, the most influential explanatory variable impacting on solar PV uptake was the size of the dwellings (proportion of dwellings with three or more bedrooms), with the proportion of people aged over 55 years continuing to be a strong explanatory variable. Four of the five strongest explanatory variables in 2012 related to the type of dwelling or its ownership. By 2014, two years after the end of $0.44 solar FiTs, the primary explanatory variable was the proportion of families of two or more persons, with education emerging as an important explanatory variable.

Table 9 – Decision tree significance of explanatory variables as ranked by R program

Influence	July 2010	July 2011	July 2012	July 2013	July 2014
1	Families	Families	Three Bedrooms	Three Bedrooms	Families
2	Over 55	Over 55	Over 55	Families	Houses
3	Houses	Houses	Houses	Houses	Education
4	Owners	Private homes	Private homes	Education	Private homes
5	Private homes		Owners	Private homes	

These analyses highlight the inter-relationship between the explanatory variables indicating that combinations of issues can be significant in determining the influence of socio-economic factors on the uptake of the response variable, solar PV.

3.3 BRT analysis

The BRT analysis provided further insight into the relative influence of the explanatory variables on solar PV uptake. Figure 3 is an example of the BRT for the year 2014. Here, the length of the horizontal bar indicates the relative importance of the corresponding variable. It is clear that for this year, the percentage of bedrooms in the postal area was the most important variable (with a relative importance of approximately 50%), followed by the percentage of families, and this variable was around three times as important as the percentage of houses. The percentage of residents with tertiary education was only slightly important, after taking the three more dominant variables into account.

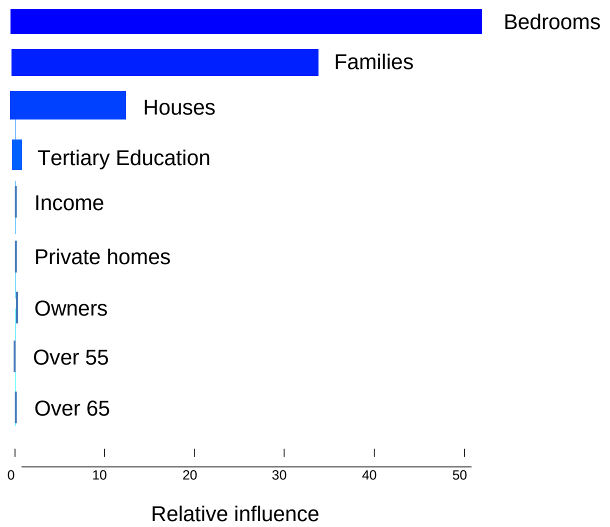

Figure 3 Boosted regression tree 2014

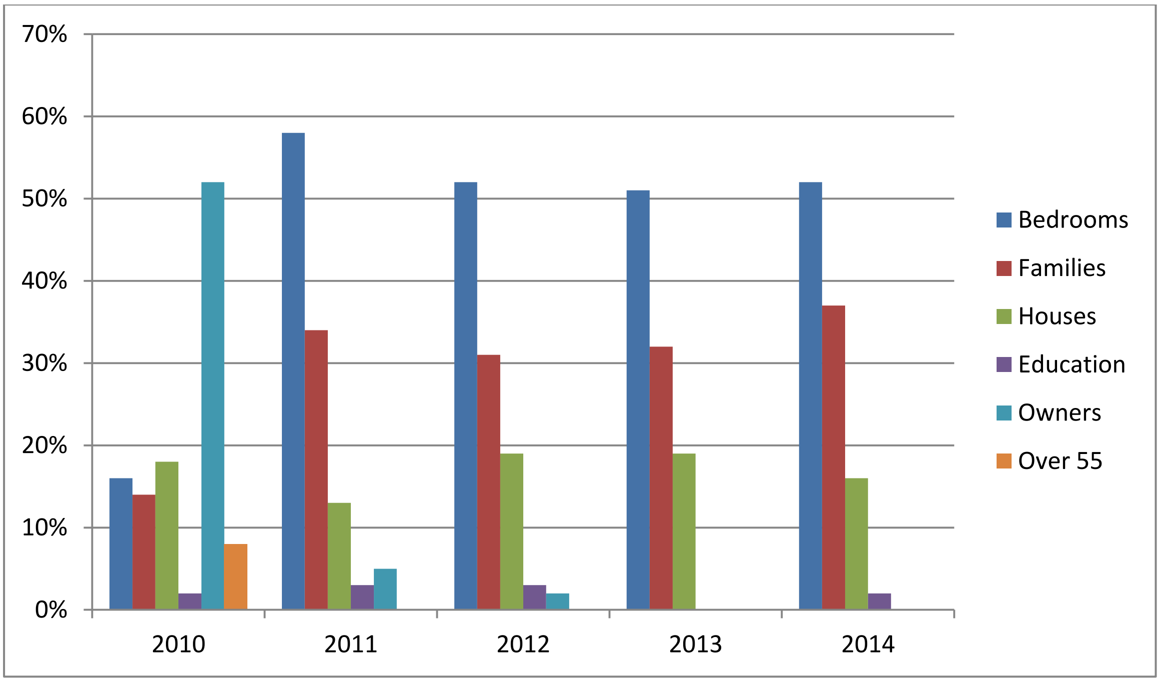

Figure 4. Summary of boosted regression tree relative influence 2010 to 2014

Figure 4 summarizes the BRT results for each year from 2010 to 2014. The examination of the relative influence of the explanatory variables as predictors shows that dwellings with three or more bedrooms, families of two or more person and houses reoccurred in most years. This generally correlated with the explanatory variables in the CART analysis.

3.3. Summary of the analyses

The multi-phased analysis undertaken in this study revealed a rich set of results. The exploratory data analysis which ranked each explanatory variable and compared the stratified subsample with the response variable (solar PV), identified differences with the published literature (Tables 4-8). The CART analysis identified explanatory variables that had not been mentioned in previous literature, in particular being aged over 55, the number of bedrooms and families with two or more persons. It also showed that having a tertiary education was less important, after taking other variables into account. The boosted regression tree analysis quantitatively confirmed the relative importance of the explanatory variables in the CART analysis.

4. Discussion and policy implications

Previous research on the factors that influence uptake of solar PV have asserted financial capacity and home ownership as important pre-requisites [2, 4, 6, 15], along with groups of people who were unable to install solar PV due to their living arrangements such as renting [2]. Acknowledging these assumptions about solar PV uptake, this research sought to introduce a wide range of socio-economic explanatory variables to examine the influence of specific socio-economic variables and whether these had linkages to other variables.

The convergent inferences arising from the multiple phases of the statistical analysis identified issues that have not been uncovered previously. Whilst home ownership was affirmed as a key explanatory variable in solar PV uptake, the significance of people aged over 55 years and families were also highlighted. In addition, having a tertiary education, which is commonly associated with better knowledge of technology, was found to be a less significant factor in solar PV uptake than in previous research. This study also re-examined previous research where financial capacity was significant in uptake of solar PV [2, 4, 6]. Income was not found to be a significant explanatory variable in any of the analyses: Table 3 showed that the five lowest socio-economic postcode areas had similar solar PV uptake rates to the top five postcode areas, and income was not selected as an important differentiator in the CART or BRT analyses (Table 9, Figures 2 and 3). Linkages between home ownership and financial capacity needs to be viewed cautiously based on this research, as persons aged over 55 may own their own homes but may be on low incomes such as government or private pensions.

During the years of Queensland government policy support for $0.44 solar FiTs (2010 to 2012), the decision tree analysis showed that being aged over 55 years was one of the top three (3) explanatory variables and this was further detected in the 2010 BRT. People aged over 55 on pensions or fixed incomes may be concerned about the impact of increasing electricity prices and installed solar PV and identified the $0.44 FiT as a cost effective means of managing electricity costs for persons on fixed incomes such as pensions. Many of the postal areas with larger numbers of people aged over 55 years are also in the lower SES areas when cross referenced with median weekly income. In postal areas with high levels of units and high levels of people aged over 55 years, solar PV uptake was low. This would tend to indicate that persons aged over 55 years who live in units or rent were more likely to be excluded from solar PV. In terms of equity, people aged over 55 years who rent would be the most vulnerable to electricity price increases.

In many areas with lower levels of university/tertiary education and lower incomes, solar PV uptake was more than double suburban profiles with high incomes and high levels of university/tertiary education (Table 5). Previous research by Caird and colleagues [16] indicated that knowledge was a critical feature in solar PV decisions. In this study, most of the postal areas with highest uptake of solar PV had the lowest levels of university/tertiary education. Conversely, in postcode areas with high levels of university/tertiary education, solar PV uptake was amongst the lowest. However, the postal areas with the highest numbers of university/tertiary educated persons were also areas with high concentration of units and apartments and any conclusions should take this into account.

4.1. Policy implications

This research reinforces some of the assumptions from previous literature on socio-economic variables that influence solar PV uptake, but identifies how other assumptions may not be as significant and how key socio-economic variables have been overlooked. Linkages between explanatory variables have been identified as being significant demonstrating that the model of solar PV uptake is a complex system with cause and effect that needs to be carefully examined. The analysis reinforced previous conclusions that owning a dwelling was one of the significant explanatory variables for solar PV installation and that the dwelling was most likely to be a house with three or more bedrooms occupied by a family of two or more persons. However, linkages with education and income were found to be less significant. New findings on the significance of being aged over 55 years highlight significant explanatory variables that may not have emerged in past research.

Although the policies that promoted $0.44 FiTs were available to the entire population, in reality this analysis identifies explanatory variables that may identify problems with the design of solar FiT policies. These policies may discriminate against people who do not own their own homes or those who live in dwellings that are not conducive to the installation of solar PV. The significance of people aged over 55 years may identify concerns by older people about electricity prices and the importance of policy measures that can assist these groups.

5. Conclusion

Rather than take for granted previous research on the demographic profile of consumers regarding socio-economic explanatory variables in solar PV uptake, the significance of this research is the re-examination of previous assumptions based on data covering almost two million people in one of the areas of greatest solar PV penetration in the world. This research scrutinized the results from previous research to ascertain the most significant explanatory variables in areas with high uptakes of solar PV and whether this changed over time and under differing policy settings. Additionally, this research sought to examine significance during and after solar PV policy interventions.

This research reaffirmed the significance of home ownership as being a significant socio-economic explanatory variable in solar PV uptake. However, it also showed the significance of the linkages between socio-economic explanatory variables in solar PV uptake. Although previous research identified education and knowledge as an important explanatory variable this was found to be less significant in the context of the larger suite of potential factors. Moreover, this study revealed the importance of other explanatory variables, such as being aged over 55 years, which had not been identified in previous research.

The research methodology used for this study and analysis shows the importance of a multi-phased analytic approach, which allows for cross referencing of the results of the different methods to identify socio-economic variables and how they may change over time and under different policy settings. It reinforces that much of the previous research and subsequent policy tends to concentrate on financial, regulatory and information drivers [16] and that the social context needs to be more fully explored. The use of multiple systems of analysis across large data collections used in this research provides for a greater contextual understanding of the phenomena of solar PV uptake. Discerning the highlighted explanatory variables further is an area for future research but one that will require clear understanding of the complexity of the decision to acquire solar PV.

In addition to elucidation of insights about the effect of socioeconomic influences on solar PV update, this paper also makes a methodological contribution to the field. Specifically, it introduces a complementary quantitative rigour to existing qualitative approaches that seek to understand social dimensions of this important problem. Moreover, the use of multiple models in an ensemble learning approach is of growing interest in the statistical and machine learning communities. This is the first time that decision tree methods, notably CART and BRT models, have been employed to investigate this complicated, and indeed statistically complex, problem.

Although the results from the data should not be generalised to other jurisdictions, the methodology used provides a guide to other researchers using statistical data to further explore socio-economic trends in uptake of solar PV. This paper makes a contribution to the field of energy policy by reviewing a range of external factors that influence decision making of consumers in the uptake of energy technology and also by highlighting some of the negative consequences including equity issues that need to be mitigated with such policies.

Supplementary material

R-software setup commands

```
library(rpart)
attach(datafile)
head(datafile)
summary(datafile)
mean(Income, na.rm=TRUE)   # mean
median(Income, na.rm=TRUE)   # median
sd(Income, na.rm=TRUE)  # standard deviation
datafile.p <- datafile/People
attach(datafile.p)
head(datafile.p)
summary(datafile.p)
```

Plot commands 2014 Solar data

```
j1=lm(datafile.p$Solar2014~datafile.p$Income+datafile.p$People+datafile.p$Familie
s+datafile.p$Privatehomes+datafile.p$Owners+datafile.p$Houses+datafile.p$Threebe
drooms+datafile.p$Education+datafile.p$Over55+datafile.p$Over65)
summary(j1)
predict.lm(j1,datafile.p)
plot(datafile.p$Julfour,predict.lm(j1,datafile.p))
lines(seq(0,.4,.01),seq(0,.4,.01))
```

Decision Tree 2014 Solar data

```
j2=rpart(datafile.p$Julfour~datafile.p$Income+datafile.p$People+datafile.p$Families
+datafile.p$Private+datafile.p$Owners+datafile.p$Houses+datafile.p$Threebedrooms
+datafile.p$Education+datafile.p$Over55+datafile.p$Over65, cp=0.01)
plot(j2)
text(j2, xpd=NA)
```

References

[1] Taylor M. Beyond technology-push and demand-pull: Lessons from California's solar policy. Energy Economics. 2008;30:2829-54.

[2] Byrnes L, Brown C, Foster J, Wagner LD. Australian renewable energy policy: Barriers and challenges. Renewable Energy. 2013;60:711-21.

[3] Grösche P, Schröder C. On the redistributive effects of Germany’s feed-in tariff. Empirical Economics. 2014;46:1339-83.

[4] Nelson T, Simshauser P, Nelson J. Queensland solar feed-in tariffs and the merit-order effect: economic benefit, or regressive taxation and wealth transfers? Economic Analysis and Policy. 2012;42:277-301.

[5] Nelson T, Simshauser P, Kelley S. Australian residential solar Feed-in Tariffs: industry stimulus or regressive form of taxation? Economic Analysis and Policy. 2011;41:113-29.

[6] Macintosh A, Wilkinson D. Searching for public benefits in solar subsidies: a case study on the Australian government’s residential photovoltaic rebate program. Energy Policy. 2011;39:3199-209.

[7] Guidolin M, Mortarino C. Cross-country diffusion of photovoltaic systems: modelling choices and forecasts for national adoption patterns. Technological forecasting & social change. 2010;77:279-96.

[8] Faiers A, Cook M, Neame C. Towards a contemporary approach for understanding consumer behaviour in the context of domestic energy use. Energy Policy. 2007;35:4381-90.

[9] de Almeida A, Fonseca P, Schlomann B, Feilberg N. Characterization of the household electricity consumption in the EU, potential energy savings and specific policy recommendations. Energy & Buildings. 2011;43:1884-94.

[10] Bahadori A, Nwaoha C. A review on solar energy utilisation in Australia. Renewable and Sustainable Energy Reviews. 2013;18:1-5.

[11] Ferrari D, Guthrie K, Ott S, Thomson R. Learning from interventions aimed at mainstreaming solar hot water in the Australian market. Energy Procedia. 2012;30:1401-10.

[12] Flannery TF, Sahajwalla V, Commission C. The Critical Decade: Australia's Future: Solar Energy: Climate Commission Secretariat, Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education; 2013.

[13] Grijalva S, Tariq MU. Prosumer-based smart grid architecture enables a flat, sustainable electricity industry. Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES: IEEE; 2011. p. 1-6.

[14] Vogt H, Weiss H, Spiess P, Karduck AP. Market-based prosumer participation in the smart grid. 4th IEEE International Conference on Digital Ecosystems and Technologies (DEST)2010. p. 592-7.

[15] Grösche P, Schröder C. Eliciting public support for greening the electricity mix using random parameter techniques. Energy Economics. 2011;33:363-70.

[16] Caird S, Roy R, Herring H. Improving the energy performance of UK households: Results from surveys of consumer adoption and use of low-and zero-carbon technologies. Energy Efficiency. 2008;1:149-66.

[17] Australian Bureau of Statistics. QuickStats Data from the 2011 Census. 2013.

[18] Australian Government Clean Energy Regulator. Small-scale installations by postode. 2014.

[19] Putler DS, Krider RE. Customer and business analytics: Applied data mining for business decision making using R. Boca Raton, FL: Chapman & Hall/CRC Press; 2012.

[20] Morgenthaler S. Exploratory data analysis. Wiley Interdisciplinary Reviews: Computational Statistics. 2009;1:33-44.

[21] The R Foundation for Statistical Computing. The R Project for Statistical Computing. Vienna: The R Foundation https://www.r-project.org/; 2014.

[22] de Ville B. Decision trees. Wiley Interdisciplinary Reviews: Computational Statistics. 2013;5:448-55.

[23] Hu W, O'Leary RA, Mengersen K, Choy SL. Bayesian classification and regression trees for predicting incidence of cryptosporidiosis. PloS one. 2011;6:e23903.

[24] McCluskey JW, Dzurllkanian ZD, Norhaya K. Boosted regression trees. Journal of Financial Management of Property and Construction. 2014;19:152-67.

[25] Morgan J. Classification and regression tree analysis: Technical Report No 1.: Boston University of Public Health; 2014.