

# Young Consumers' Behavioural Intentions to Engage in the Bike Pseudo-Sharing

Joey Gronovius, Lorenzo Lops, and Carmen Schlaugat

## I. INTRODUCTION

**M**ILLENNIALS have been shown to constitute a substantial segment of the collaborative consumer group in what has been labelled the “sharing economy” [1; 2; 3]. This realisation has prompted some researchers to highlight the potential for further investigation into young consumer's attitudes, intentions and behaviours towards collaborative consumption. As “collaborative” or “access-based consumption” tends to be facilitated by network technology and accessible via electronic devices, it comes as no surprise that younger consumers – often qualified as “tech savvy” – are more eager to engage in the use of these services than the average population [4]. In light of the explosion of bike sharing web enabled services among these young consumers, and particularly in a country where biking is one of the primary means of transportation like the Netherlands, it would be interesting to conduct additional research in this direction.

It is often argued that environmental considerations constitute important motives – especially for young consumers – to engage in sharing activities. This contention has been investigated by Hwang Griffiths [3], whose findings suggest that utilitarian and hedonic values have substantial influence on young consumers' attitudes towards collaborative consumption. However, symbolic values – such as those pertaining to issues of sustainability – have a negligible effect. According to Azjen [5], “attitude” towards a behaviour or action – among other factors – is a key determinant of behavioural intentions, which may further support the claim that young consumers are not highly influenced by environmental considerations in their decision to engage in sharing. The preceding finding is somewhat counterintuitive given that, from a business as well as a societal perspective, the proliferation of sharing platforms has the benefit of reducing waste associated with underutilized assets [4; 6]. As such, sustainability remains a topic relevant to sharing activities, and a salient issue on which various companies' advertisement seeks to capitalize.

Sharing activities and collaborative consumption have challenged conventional business models and challenged dominant economic logics, both in theory and in practice. For instance, the proliferation of platforms that allow sharing transportation with prior and subsequent customers has turned beliefs pertaining to the importance of ownership of resources upside down. Some scholars, namely Belk [1], go as far as to maintain that modern society is entering a “post-ownership economy” and that the effects could be as far reaching as those of the Industrial Revolution. To achieve more clarity, but also to illuminate certain cognitive and affective dimensions of

individual decision making, the concept of “psychological ownership” – developed in the field of organisational behaviour – has been deployed in literature dealing with resource sharing [6]. It refers to a state in which individuals feel as though they possess a certain target despite it does not belong to the individuals' legal possessions [7]. One interesting and particularly relevant semantic caveat discussed by Belk [4] in his follow up article pertains to the appropriation of the term “sharing” by many organisations (for example Airbnb or Zipcar) that do not really operate on the basis of actual sharing. These hybrid organisational structures – for which the most far-fetched form of sharing involves knowing that someone else has used the target in the past (and that someone will probably use it in the future) – would more formally qualify as pseudo-sharing, given the presence of “profit motives” and “expectations of reciprocity” as well as the absence of a “sense of community”. An example cited as a common instance of pseudo-sharing is “long-term rent and lease” [4], which is indeed the basis on which some subscription-based bike-“sharing” services operate. The focus of this investigation has hence been rephrased in terms of young consumers' intentions to engage in or extend their lease with Swapfiets, an emerging Dutch company. Swapfiets is a very interesting case study due to its market positioning: it has capitalized on the breadth of the bike sharing market in the Netherlands by offering a more convenient as well as personal service, emphasising some customisation and particularly sustainability. Moreover, it represents one of the few examples in the web enabled cycling market that operates on the basis of a medium to long-term lease contract. The latter point may allow for the conclusions of this study to be generalizable to other instances of long-term leasing. It appears as though this has not been studied so far. Belk [4] has identified, but not empirically investigated it as a form of pseudo-sharing. The next section will outline the research questions and objectives, as well as a preliminary statement on the aspired managerial implications. The third section introduces the hypotheses that will be tested, a visual representation of the proposed conceptual model and more extensive definitions of the applied concepts, namely, operationalization of responsiveness to both environmental issues and psychological ownership, taken from extant literature on resource sharing. Section four and five discuss the data and methods applied and the results of the analysis. The last section presents the implications of the study and concluding remarks.

## II. RESEARCH OBJECTIVES AND QUESTIONS

The influence of the aspects discussed in the introduction has been widely investigated by the research brought to bear and beyond. The present investigation will seek to apply existing knowledge on sharing activities to assess whether people's behavioural intentions to engage in bike pseudo-sharing can be satisfactorily modeled and studied, and whether these insights can be valuable to other leasing businesses. Aside from gaining insight into people's behavioral intentions to engage in pseudo-sharing, the study is additionally interested in directly investigating the supposed effect of feelings, or 'psychological' ownership. The latter will be modeled along the lines of Paundra et.al [6], whose findings suggest a negative relation between the latter and intent of selecting a shared car, whereby a low psychological ownership may lead to a higher preference for a shared car. This suggests that Swapfiets users should have a comparatively higher inclination towards feelings of ownership than users of other, more short-term bike sharing services, but a comparatively lower inclination than actual bicycle owners. This gives reason to believe that Swapfiets's customers value feelings of ownership more than the customers of other bike sharing services, and intuitively that it could attract customers that would have otherwise preferred purchasing a bike.

Its broader base of potential appeal perhaps offers a crude explanation for its widespread success. A comparative empirical study to assess this inter-firm difference is beyond the scope of this paper, which instead seeks to assess the relative influence of environmental considerations and psychological ownership on the intention to extend one's bike lease. A comprehensive consideration of convenience is excluded almost ad hoc, despite potentially being a key determinant.

The length of the contract has been effectively constructed as a proxy for convenience as it captures some dimensions of flexibility without focusing too highly on the service's peculiarities. This is primarily because advertisement of most bike sharing players already focuses on convenience and service. Likewise, it is not uncommon to see ads point to the benefits in terms of sustainability. It is the team's contention however, that Swapfiets' marketers emphasise issues of psychological ownership too distantly, if at all – and it is their niche. If the following analysis could reveal a significant influence of psychological ownership on intent to extend a lease, that could have important managerial implications especially for marketers. Stated more formally we wish to investigate the following questions:

- a) Does the length young consumers hold a Swapfiets contract for have an influence on their intentions to extend their (pseudo) bike-sharing lease?
- b) Do environmental considerations have an influence on young consumers' intentions to extend their bike-sharing lease?
- c) Does psychological ownership have a significant influence on young consumers' intentions to extend their bike-sharing lease?
- d) Is this effect stronger than that of environmental considerations?

- e) Does frequency of use have an effect have an influence on young consumers' intentions to extend their bike-sharing lease?

- f) Can an accurate model be constructed?

A last informal and consequential question that would be interesting to explore is whether the implications of the results could be applied to more effectively design customer retention campaigns, a somewhat typical churn application with revisited techniques. Some of the insights gained may well be of relevance for customer acquisition purposes.

## III. THEORETICAL BACKGROUND

In terms of conceptualisation, assuming that motivations to engage in a contract in the first place do not change substantially over time, the present study models the intent to extend the contract in a similar fashion to how previous literature (i.e. Paundra et. al) has modelled the intent to engage in pseudo-sharing. Intent may well change as a result of a shift in attitude, which indeed may be an effect of "time and place" on psychological ownership [8], but it is expected that neither of the other two determinants of behaviour – i.e. "subjective norm" and "perceived behavioural control" should change too much throughout [5; 9].

Contract length represents the amount of time an individual has been using a Swapfiets bike, and it is included in order to account for further potential motives for not extending one's contract. In particular, individuals who have been using a contract for an amount of time close to twelve months may be expected to redeem their contract because they no longer require a bike. In the analysis these individuals should be treated differently from those whose departure was more properly classed as a 'churn', such as relatively new customer deciding to purchase their own bike. Alternatively, the influence could be positive: the longer a customer has had the contract, the more they are assumed to be satisfied with the service or have a long-term commitment to it, and therefore the more likely they are to extend it. The exact direction of this effect on intentions to extend one's lease is not known however, given that a longer relationship with one's bike may equally lead individuals to develop stronger desires for actual ownership. It is hoped that our investigation will be able to clarify the existence and eventual significance and direction of this relation.

*A. H1: 'Contract Length' has an influence on 'Intent to extend Swapfiets Contract' .*

A general conceptualisation of psychological ownership states that it is generated by the fulfilment of four basic needs: the needs for efficacy, self-identity, place and stimulation [6; 7; 8]. When a product or service stimulates a person's self-identification; provides an effective daily reference; and when the consumer can exercise control over it, he is likely to develop a sense of psychological ownership that is independent, and often antecedent of actual ownership (ibidem). Therefore, and in line with prior research, psychological ownership has been hypothesised to have a negative relation with people's intentions to extend their lease [6]. Despite Swapfiets users

have been assumed to have relatively higher inclination towards feelings of ownership than users of most other short-term competitors, an excessively high inclination may lead them to prefer purchasing their own bike.

*B. H2a: ‘Psychological Ownership’ has an influence on ‘Intent to extend Swapfiets Contract’.*

For Swapfiets users in particular, in some way even merely ‘psychological’ ownership creates value. It would hence be strategically important to identify such a significant influence on intent to extend one’s contract. Specifically, it would be interesting to find a stronger influence than that of environmentalism. The motivating literature was discussed in the preceding sections.

*C. H2b: ‘Psychological ownership’ has a stronger effect on ‘intent to extend contract’ than ‘environmental considerations’.*

Environmentalism is conceptualised on the basis of consumers responsiveness to certain specific issues such as their subjective norm towards environmental conservation; their attitude towards and association between wasteful consumption and the consequences of these unsustainable practices on the environment [3; 5]; and on an explicit – though rough – measure of “perceived consumer effectiveness” [3], or respondents’ perceptions regarding the potential impact of their individual actions. The last two dimensions would intuitively seem to be related in their impact on intentions through environmentalism, but research has found no evidence to support this “moderating role of PCE for the path from attitude to [purchase] intention” [3]

*D. H3: ‘Environmental considerations’ have an influence on ‘Intent to extend Swapfiets Contract’.*

Frequency of use has been included to somewhat complement some of the possible explanations for churning customers proposed by length of contract by working via environmental considerations and psychological ownership. For instance, a very high frequency of use associated to a months-old customer may explain a churn by highlighting how purchasing a bike became more reasonable for that customer. On the other hand, a low frequency of use by a relatively new customer may explain a churn in terms of low psychological ownership and a consequent preference for pay-per-use sharing services. In addition, it may offer a rudimentary measure of the extent to which one’s bike represents a ‘daily point of reference’ and is hence included in the survey. Overall, the direction of this relation cannot be clearly anticipated: whether the influence is positive or negative is likely to depend on considerations like the individual situation of the consumer (some might rather get their own bike if they use it with relative frequency and/or are permanent residents). This is perhaps also why it may be interesting to further investigate it. The motivating literature was discussed in the preceding sections.

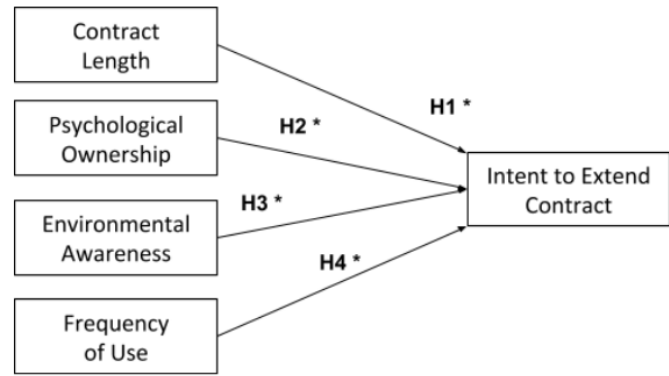


Fig. 1: Causal Scheme

*E. H4: ‘Frequency of Use’ has an influence on ‘Intent to extend Swapfiets Contract’.*

The motivating literature was discussed in the preceding sections.

## IV. DATA AND METHODS

### A. Sample

Due to the constraints of the assignment, the population was limited to include all higher education students residing in the Rotterdam metropolitan area. The team was unable to effectively expand the reach of the survey to other major Dutch cities, so restricting insights to the local population seemed more appropriate. The survey did not discriminate against students of international status. Several controls were included in the survey to allow for screening and the organization of respondents. Using publically available statistics, a rough estimation of the population was determined to be approximately 70,000 students [10; 11]. The minimum requirements for the findings to be statistically significant was defined to include a confidence level of 0.90 with a margin of error of 0.10. When considering these requirements together with the estimated population size, the respective target sample size was determined to be 68.

Convenience sampling was utilized to collect survey responses. The survey was distributed electronically using a version of Qualtrics provided by the university. Because the research group was based at the Erasmus University Rotterdam, the sample was restricted to those in the immediate vicinity of the campus. It should be acknowledged that this introduces many potential issues with the findings, such as the self-selecting nature of the respondents leading to the overestimation of targeted measures [12]. Therefore, the approach of this study was to explore whether there is significant information disproving the aforementioned hypotheses, as opposed to proving them.

### B. Manipulating the Control Variables

To attempt to derive insights relating to business intelligence from the survey data, multiple machine learning techniques were experimented with using R programming. Due to the low amount of observations, many of these techniques either

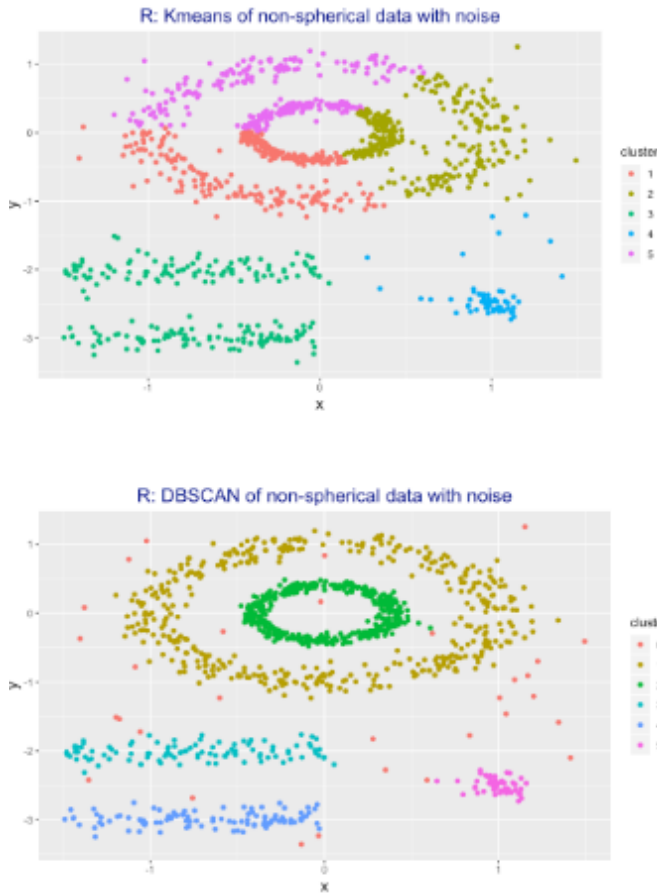


Fig. 2: Example Clustering Result of K-Means and DBSCAN. Source: Adapted from [13]

failed to function properly or reveal definite relationships. Nevertheless, all attempted analyses have been described with the intent to set a framework for further research.

Once the raw data was prepared to adhere to the requirements of the following analysis techniques, factor analysis for mixed data (FAMD) was applied solely on the control variables for dimension reduction purposes. This method performs principal component analysis (PCA) for quantitative data and multiple correspondence analysis (MCA) for qualitative data. Ideally, the data would be reduced to three components. Several clustering methods were then implemented using either the original responses or the reduced data. As FAMD converted the original observations into collections of new quantitative data, both k-means clustering and density-based spatial clustering of applications with noise (DBSCAN) clustering were performed on this data. DBSCAN clustering was necessary due to its ability to better identify and select clusters based on density and detecting outliers, which addresses many of the issues that often arise for k-means. Figure [13, Fig. 2] demonstrates a comparison of the two clustering algorithms. A third clustering technique, known as k-prototype, was also used in the event that the aforementioned dimension reduction proved impractical. Unlike the previous two clustering methods, k-prototype accepts both qualitative and quantitative data, and thus can be used on the original survey data. The intent

Type	Definition
Control	Intent to Extend ~ Gender + Age + Employment Status + Location + Education + Type of Swapfiets
Explanatory	Intent to Extend ~ Psychological Ownership + Environmental Concern + Contract Length + Frequency of Use
Complete	Intent to Extend ~ Psychological Ownership + Environmental Concern + Contract Length + Frequency of Use + Gender + Age + Employment Status + Location + Education + Type of Swapfiets

Fig. 3: Definition of Models

of the preceding data manipulations was to facilitate easier visualization of potential relationships among the measured traits. Dimension reduction allows for data with an excessive amount of components to be easily visualized on a two or three dimensional plane while still maintaining a degree of relations between the observations. Clustering further promotes understanding by classifying connected observations into separate groups. This simplification of complex data into a medium that is easily digestible remains relevant in the modern business realm where companies pursue optimizing the time to achieve useful data insights in an attempt to gain an edge on their rivals [14]. Another motivation for dimension reduction and subsequent clustering is the verification of the choice of control variables. At its core, both processes attempt to extract hidden mathematical relationships among the measured traits. The controls were specified as qualities that may potentially have an unintended effect on the traits focused on in this study. If the previous techniques suggested that specific qualities had a significant influence in the distribution of respondents, than it would indicate that these qualities may need to be accounted for in the study - or addressed in further research - in order for the conclusions to be valid.

### C. Statistical Analysis

Following the exploration of the controls, the validity of the formulated hypotheses were examined through the use of various statistical significance tests for the three logistic regression models specified in Table . To determine the significance of each trait specified in the models and to establish the goodness-of-fit between competing models, a likelihood ratio test for item response theory models was performed. If a level of a categorical trait (its respective dummy variable) was identified as potentially significant, a Wald test was then conducted among all levels to determine whether the overall influence of the trait was truly statistically significant. In addition, the correlations between all measured qualities were found using a package capable of calculating the Pearson, tetrachoric, polychoric, biserial, and polyserial correlations for mixed data types. This was done to observe whether there were any unexpected correlations among the measured traits.

### D. Prediction

The final stage of the data processing involved implementing various common classification methods to rank their performance and examine whether any of these methods

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	68	22.44	2.42	18	21	24	29
cSwapUsed	68	4.68	1.11	2.00	4.00	5.00	7.00
Env	68	4.16	0.78	2.00	3.60	4.80	5.00
cPsych	68	3.57	0.56	2.40	3.20	3.80	5.00

Fig. 4: Quantitative Analysis Results

proved acceptable in predicting a Swapfiets customer's intent to continue their contract. The implemented algorithms included logistic regression, random forests, gradient boosting machines, and naive Bayesian classifiers. The consideration of various classification techniques were limited to those taught in the team's courses. The suitability of each classifier was defined by their ranking in various performance measures (accuracy, precision, specificity, and sensitivity), their receiver operating characteristic (ROC) curves, and their area-under-the-curve (AUC) values. The classifiers were trained using leave-one-out cross validation, in which a single observation was reserved for prediction and all other observations were used for training. This process was repeated until each observation was used for a prediction. This method was chosen over utilizing common holdout ratios (such as 30:70), due to its tendency to produce better results when provided with a small sample size. Because there are a significant number of coefficients in the complete model, leave-one-out cross validation helps address the severity of overfitting by allowing for the maximum amount of observations to be utilized in fitting the model.

## V. ANALYSIS AND RESULTS

### A. Sample

Figure 4 reflects various univariate measures relating to the survey numerical traits, and reveal unexpected distributions. The average respondent age was higher than expected. As a typical first-year student is 17 or 18 and graduates by 22, the survey appeared to have mostly reached undergraduates in the later stages of their degree and many graduate students. Respondents also, on average, scored high for environmental concern, which is somewhat in contrast with Hwang Griffiths' [3], but low for psychological ownership. Participants were found to use their bicycles for the majority of the week, averaging approximately 5 days.

Appendix II gives the counts and frequencies of each categorical variable's levels. There was no significant majority in the distribution of the respondents' genders. The amount of unemployed and part-time employed respondents were similar, with few respondents working full time. The vast majority of respondents lived in urban locations, with very few coming from suburban areas or towns. No respondents lived in rural locations. There was an even number of respondents having completed either high school or a Bachelor's program, with a small proportion having completed a Master's program.

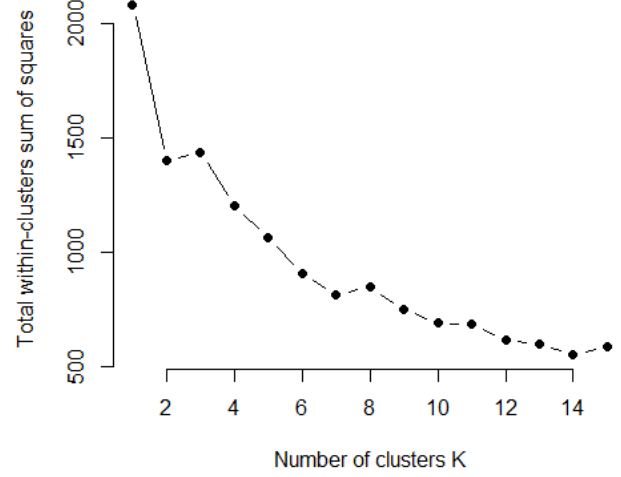


Fig. 5: Elbow Plot for K-Prototype Clustering of Control Variables

No respondents failed to complete high school or achieved a post doctorate degree. Most of the respondents possessed a standard Swapfiets bike, while a select few invested in a deluxe bicycle and no respondents had an e-bike. Those that had a Swapfiets contract for the periods of 3-6 months, 6-12 months, and 1-2 years all shared a similar number of counts. Those with a contract 0-3 months old was twice that count and no respondents had a contract for over 2 years. The majority of respondents were willing to extend their Swapfiets contract, and about a quarter intended to end their contract.

### B. Manipulating the Control Variables

As elaborated upon in the preceding section, the intent of FAMD was to reduce the control features to three dimensions for visualization purposes. However, the resulting components were found to only account for approximately 0.50 of the cumulative variance. This proportion was significantly low enough that the dimension reduction was deemed inappropriate as it indicated that the integrity of the data would potentially be compromised if accepted and used further.

Both k-means and DBSCAN clustering require quantitative data to function properly. The control variables included categorical features, restricting both methods to depend on valid results from FAMD (which accepts mixed data types and returns an output of purely numerical data) to function properly. As this was not the case, clusters determined by these methods are inappropriate to use as evidence for relevant conclusions. This was further supported by the absence of a clear inflection point in the curve of the k-means method's elbow plot.

The previous shortcomings were bypassed through the use of k-prototype clustering, which circumvented the need of FAMD's numerical output. The optimal number of clusters to specify for the method was determined by minimizing the

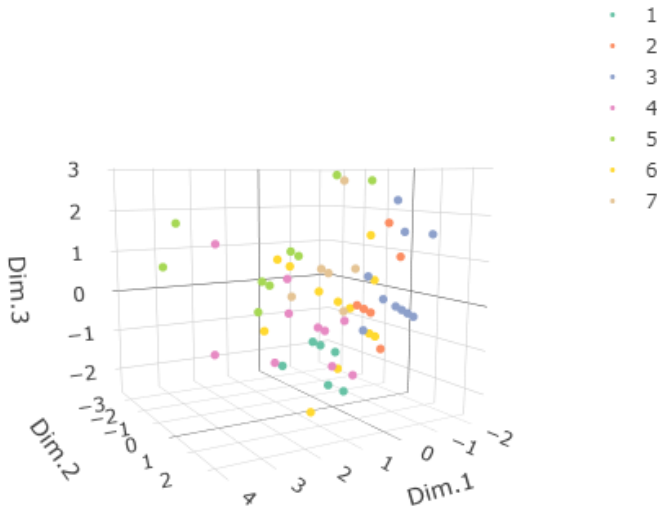


Fig. 6: K-Prototype Clustering for Control Variables Demonstrated on FAMD Coordinates

within-cluster sum of squares. Figure 5 reveals that seven clusters is a sufficient choice. If the intention of clustering was to fit the models separately for each cluster, this selection would be inappropriate. On average, each cluster would contain about 10 observations if using the original survey results. Because the number of samples would not be sufficiently larger than the number of features, the models would be expected to experience severe overfitting and prove impractical for predictions.

Figure 6 illustrates the classification of the respondents using k-prototype clustering. To demonstrate the potential of visualization, the respondents were transformed using their respective three-dimensional coordinates resulting from FAMD. No distinct groupings can be discerned from this spread. The inability to generate reasonable clusters further supports the assumption that the specified control variables lack a significant confounding effect on the explanatory variables.

### C. Statistical Analysis

A conventional classification of correlation magnitudes specifies the range of 0.1 to 0.3 as small, 0.3 to 0.5 as medium, and 0.5 to 1.0 as large [15]. Figure 7 reveals the correlation coefficients between the various features included in the trait. This serves as a preliminary step to examining whether the decisions of control and explanatory variables are justified, and to aid in identifying any unexpected relationships among the features. The following discussion will only focus on correlations relevant to the goals and assumptions of the study.

As the proposed causal diagram and its respective research hypotheses only consider the explanatory and dependent variables, the correlations between these variables and the control variables should be weak. However, this was not the case. Respondents' employment status was found to have a medium negative correlation with psychological ownership ( $r = -0.36$ ), and a large negative correlation with both environmental concern ( $r = -0.56$ ) and the intent to extend their contract ( $r$

	Gender	Age	Work	Home	Education	cType	cSwapUsed	cContractLength	Env	cPsych	cExtend
Gender	1	-0.01	0.02	-0.04	0.04	-0.06	-0.18	0.04	-0.1	0.04	-0.27
Age		1	0.12	-0.16	-0.26	-0.25	-0.11	0.02	-0.03	0.07	0.12
Work			1	-0.32	-0.01	0.35	-0.2	0.18	-0.56	-0.36	-0.51
Home				1	0.09	0.38	-0.01	-0.04	0	0.04	0.26
Education					1	-0.02	0.04	0.33	-0.05	0.01	-0.26
cType						1	0.15	0.54	-0.27	-0.57	0.05
cSwapUsed							1	0.13	0.29	0.24	0.58
cContractLength								1	-0.15	-0.29	-0.26
Env									1	0.39	0.6
cPsych										1	0.32
cExtend											1

Fig. 7: Correlation Matrix for All Variables

$= -0.51$ ). A possible explanation for the negative link between employment and the intent to extend could be that those that work more have more disposable income or need to commute further distances, allowing them to purchase a permanent bicycle or negate the usefulness of a bicycle respectively. Education was also determined to have a medium positive correlation with contract length ( $r = 0.33$ ), which may be due to those at later stages in their academic career having been in Rotterdam and needing a method of transportation for a longer period of time than those in earlier stages of their program. Surprisingly, the type of Swapfiets bicycle appeared to possess a large positive correlation ( $r = 0.54$ ) with contract length, perhaps signaling that the type of bike leased influenced the decision to maintain the contract for certain periods of time. A large negative correlation also existed between the type of bicycle and psychological ownership ( $r = -0.57$ ). The existence of correlations between the control variables and those that are the focus of the proposed causal model suggests that the incorporation of the discussed control variables into the model may prove necessary.

Furthermore, there are numerous instances of medium to large correlations present between the variables included in the original causal scheme. The most significant correlation with the intent to extend among the explanatory variables can be attributed to environmental consciousness ( $r = 0.6$ ), closely followed by frequency of use ( $r = 0.58$ ). There also exists a medium correlation ( $r = 0.32$ ) between psychological ownership and the intent to extend one's contact.. Contrary to the expectations of the study, contract length has only a small correlation with the dependent variable ( $r = -0.26$ ).

To determine whether these correlations provide valuable



	chi2	df	P
cHome	2.042	2	0.360
cLength	1.576	3	0.665
Education	0.792	2	0.673

Fig. 8: Wald Test for Significant Leveled Variables

	Resid. Df	Resid. Dev	Df	Deviance	Pr(> Chi)
Control Model	58	45.042			
Complete Model	52	28.370	6	16.672	0.011

Fig. 9: Likelihood Ratio Test for the Control versus the Complete Model

insight to the actual presence of relationships between the respondents' traits, the influence of the features' coefficients for the logistic regression model were further analyzed. Appendix III provides the study's relevant findings. Several of the categorical variables exhibited significance for at least one of their levels. Figure 8 displays the results of the Wald test for each of these occurrences. No variable was found to have a p-value significant at a significance level of 0.1 or lower. In addition, none of the other control variables were statistically significant at any acceptable significance level. This is in line with the study's expectation of little influence due to control variables, but contradicts the correlations among the control variables previously discussed. In regard to the explanatory variables, several proved to be statistically significant. For the explanatory model, both the frequency that respondents used Swapfiets and their environmental concern were significant at the 0.01 significance level ( $p = 0.007$  and  $p = 0.006$  respectively). Thus, the null hypothesis that these variables have no influence on the intent to extend can be confidently rejected. Similarly, the same variables are also significant in the complete model ( $p = 0.031$  and  $p = 0.042$  respectively), but at the 0.05 significance level. The remaining explanatory variables, psychological ownership and contract length, were unable to achieve a p-value to reject the same null hypothesis for either model. Figure 9 gives the results of a likelihood ratio test used to compare the goodness-of-fit of the control and explanatory model versus the complete model. The null hypothesis is that there is no difference in the fit between the models being compared. The complete model outperforms the control model at the 0.05 significance level (note that it almost achieves this at the 0.01 significance level). This serves as further confirmation that the control variables collectively do not have a significant influence on the dependent variable, and that the addition of the explanatory variables produces a more fitting model.

#### D. Prediction

The performance measures of each of the designated classifiers is stated in Figure 10. In general, there is no clear

	Accuracy	Sensitivity	Specificity	Precision
Log	0.750	0.849	0.400	0.833
Forest	0.750	0.887	0.267	0.810
GBM	0.779	0.849	0.533	0.865
Naive Bayes	0.779	0.887	0.400	0.839

Fig. 10: Performance Measures for Classifiers

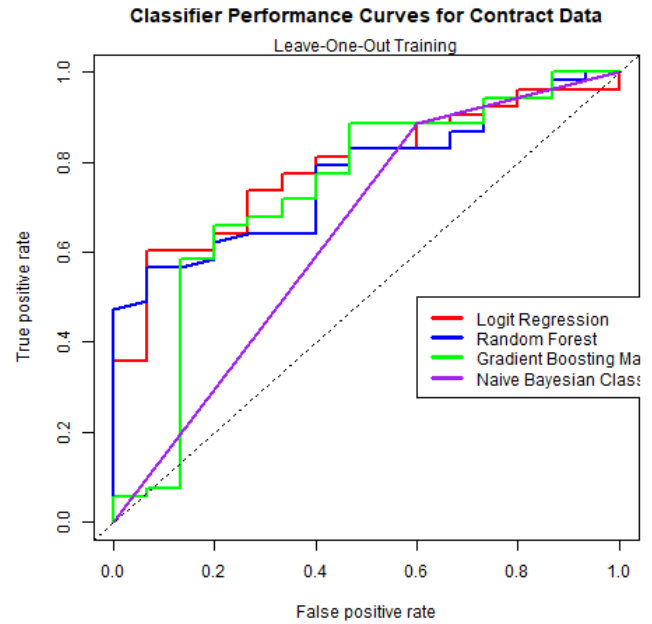


Fig. 11: ROC Plot for Classifiers

top performer as some algorithms excel in one measure while falling short in another. All classifiers fail to perform adequately when considering specificity, or its ability to detect true negatives. The random forest classifier especially suffers here. To achieve the best balance of the performance measures, the gradient boosting machine classifier would be selected. It performs the best in its accuracy, specificity, and precision while maintaining a comparable sensitivity. Aside from the lackluster specificity, any of these classifiers would be appropriate to pursue further for prediction purposes.

In the spirit of demonstrating the benefits of data visualization, the classifiers' performance in terms of its true positive and negative detection rates are displayed in ROC curve plot in Figure 11. It can immediately be seen that the naive Bayesian classifier and the gradient boosting machine do not perform as well as their alternatives. This contradicts the selection of the gradient boosting machine based on the previous four measures, and serves as a reminder that various performance measures should be considered when deciding on methods. It is difficult to distinguish the logistic regression and random forest curves, so Figure 12 is provided to specify the AUC values. The AUC values reflect that the logistic regression

	Log	Forest	GBM	Naive Bayes
AUC	0.786	0.779	0.723	0.643

Fig. 12: AUC Values for Classifiers

classifier performs the best.

## VI. DISCUSSION

### A. Main Conclusions

In terms of the research hypotheses and the questions (a) - (e) stated at the beginning of the report, the presented results and analysis led the team to conclude (with the varying confidence specified in the preceding section) that:

- 1) H1: Contract length had no significant effect on the intent to extend respondents' intent to extend their Swapfiets contacts.
- 2) H2a: Psychological ownership had no significant effect on the intent to extend.
- 3) H2b: Psychological ownership did not have a greater influence on the intent to extend than environmental considerations. Instead, environmental considerations had a greater effect.
- 4) H3: Environmental considerations did have a significant influence over the intent to extend.
- 5) H4: Frequency of use did have a significant effect on the intent to extend.

In consideration of question (f), the team concluded that the logistic regression classifier was the most appropriate selection to be used for the prediction of the dependent variable, the intent to extend one's contract. The fact that psychological ownership nor contract length appeared to have significant influences on whether respondents intended to extend their Swapfiets contracts contradicted the team's expectations.

### B. Additional Comments and Relevance

Since high frequency of use seems to increase the usefulness of Swapfiets' service for its customers (as opposed to leading to churn due to individual bike purchase) it could be inferred that financial considerations play a minor role in the decision of whether to rent or own a bike. It may well be related to the service's high customer service and support, which has long represented its trademark and assured it its initial success.

The fact that environmentally conscious people tend to be more likely to engage in a long term relationship with Swapfiets indicates that they perceive this action as sustainable, supporting their attitude and subsequently influencing their behaviour [5]. Whether Swapfiets or just bike sharing (as opposed to ownership) is perceived as sustainable can not clearly be stated based on the results and could be further investigated by future research. However, the strong influence of environmental considerations on customers continued engagement in Swapfiets services suggests that sustainability is an important factor in value creation for their customers.

Consequently, it would be advisable for Swapfiets to market sustainability more; not just for customer retention but possibly also customer acquisition. So far Swapfiets has mostly emphasized practicality and convenience of their services, while putting little to no focus on sustainability. According to these findings, this could indeed represent a shortcoming in their marketing strategy, since environmentalism seems to be a determining factor in customers considerations to engage in a long term relationship with Swapfiets. This finding in particular could also be useful to other transportation lease companies, especially those serving young consumers.

As neither psychological ownership nor contract length currently play a part in customers' decision to extend their contracts, the findings suggest that contract extension is not based on a loyalty to Swapfiets as a service. The fact that Swapfiets is a relative newcomer to the market suggests this may represent an important area in which to invest. Indeed, though perhaps applied too boundedly in the present study, psychological ownership has the potential to offer additional understanding into phenomena such as "customer satisfaction", "word-of-mouth" and "willingness to pay" [8] and has further been associated with loyalty towards services and platforms as well [9].

While psychological ownership and contract length did not prove to be factors that directly contribute to Swapfiets customers intentions to extend their contracts, they may still be relevant as moderating factors in connection to the findings about environmental awareness. As opposed to short term bike sharing services such as Mobike or Donkey Republic, which function on the basis of a rapid transition of their mobility products between several customers, subscription based 'pseudo'-sharing services like Swapfiets capitalize on equipping their customers long-term. This fosters an increased feeling of psychological ownership of the device on the customers' side, which may consequently lead to a more responsible handling of the object. Although Swapfiets emphasizes the fact that the bikes are repaired (as it is their key service), the company does not communicate this clearly as a sustainable business practice to its customers. By doing so, Swapfiets would create additional value for its customers since apparently, sustainability is an important factor for them and considering their increased feeling of psychological ownership would allow for an increased "perceived consumer effectiveness" as well [3].

### C. Limitations

A major concern for the results involves the small sample size. The size just barely met the requirements to be accepted as statistically significant at a 0.90 confidence level with a 0.10 margin of error. However, such a small sample size is not ideal for the machine learning techniques experimented with throughout the study. This led to problematic results for dimension reduction, impractical clustering, and imprecise predictions.

With more resources and time, the study would be better conducted using random sampling (as opposed to convenience sampling) spanning a much larger sample. This would ideally



allow for successful dimension reduction and proper clustering. If an alternative scenario consisting of a much larger sample size and a successful reduction to three appropriate dimensions generated obvious clusters, researchers could acknowledge that segmentation is possible in their population and develop strategies to address groupings tailored to their clusters' traits.

Furthermore, additional time would allow for sufficient time to explore techniques better suited to the nature of the data. For example, many of the analysis methods used assume a specific distribution of the data, such as normality. It is common to neglect these prerequisites for the sake of producing results. However, this deteriorates the quality of the study's conclusions.

## VII. REFERENCES

- 1) Belk, R. (2014b). Sharing versus pseudo-sharing in Web 2.0, *The Anthropologist* 4(2).
- 2) Head, L. (2013), "Will millennials drive the shift from a consumption-based to a values-based economy?", available at: [www.sustainablebrands.com/news\\_and\\_views/behavior\\_change/will-millennials-drive-shift-consumption-based-values-based-economy](http://www.sustainablebrands.com/news_and_views/behavior_change/will-millennials-drive-shift-consumption-based-values-based-economy).
- 3) Hwang, J., Griffiths, M. A. (2017). Share more, drive less: Millennials value perception and behavioral intent in using collaborative consumption services. *Journal of Consumer Marketing*, 34 (2), 132-146.
- 4) Belk, R. (2014a). You are what you can access: sharing and collaborative consumption online, *Journal of Business Research* 67(8): 1595–1600.
- 5) Ajzen, I. (1991). The theory of planned behavior, *Organizational Behavior and Human Decision Processes* 50(2): 179–211.
- 6) Paundra, J., Rook, L., van Dalen, J. and Ketter, W. (2017). Preferences for car sharing services: Effects of instrumental attributes and psychological ownership, *Journal of Environmental Psychology* 53: 121–130.
- 7) Pierce, J. L., Kostova, T. and Dirks, K. T. (2003). The state of psychological ownership: Integrating and extending a century of research, *Review of General Psychology* 7(1): 84.
- 8) Jussila, I., Tarkiainen, A., Sarstedt, M., Hair, J. (2015). Individual Psychological Ownership: Concepts, Evidence, and Implications for Research in Marketing. *The Journal of Marketing Theory and Practice*, 23(2), 121-139.
- 9) Lee, Yang, Koo, (2018). Exploring the effect of Airbnb hosts' attachment and psychological ownership in the sharing economy, *Tourism Management* 70: 284-294.
- 10) Registered students. (n.d.). Retrieved from <https://www.eur.nl/en/about-eur/facts-and-figures/registered-students>.
- 11) About us. (2016, August 30). Retrieved from <https://www.rotterdamuas.com/about/about-us/>.
- 12) Schonlau, M., Fricker, R., Elliot, M. (2002). Conduction Research Surveys via E-mail and the Web, *RAND Corporation*, 33-34.
- 13) Berhane, F. (n.d.). Data distributions where Kmeans clustering fails. Can DBSCAN be a solution? Retrieved from [https://datascience-enthusiast.com/Python/DBSCAN\\_Kmeans.html](https://datascience-enthusiast.com/Python/DBSCAN_Kmeans.html).
- 14) Stodder, D. (2013). Data Visualization and Discovery for Better Business Decisions, *TDWI Research*, 10-11.
- 15) Pearson Product-Moment Correlation. (n.d.). Retrieved from <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>.