Happiness and its Socio-demographic Determinants Analyzed with Datamining: the Case of a Community at the North of the Border of Mexico

Karla Erika Donjuan Callejo[1], Mario Ricardo Sotomayor[2], Alberto Ochoa-Zezzatti[1]

[1] Universidad Autónoma de Ciudad Juárez , Mexico
[2] Universidad TecMilenio, Mexico
erika@corporacionaem.com, ricardo@corporacionaem.com

Abstract. The topic of happiness and subjective welfare has taken a significant relevance inside the theories of several disciplines such as psychology, economics and politics; the importance of the individual subjective welfare. Happiness is a matter seen throughout many angles and it has an impact in the quality of life, therefore, we can definitely consider happiness multidimensional, as the variables affecting it are complex and diverse. This article analyzes which of the socio-demographic variables such as age, gender, occupation, scholarships, migration intention, and available family income are correlated to the level of Declared Happiness of the people who live in the community at the north of the border of Mexico: Juarez. Making use of the tools from datamining: the classification tree of the type CRT in order to segment the subjects and to correlate the main variable (Declared happiness) with the dependent variables, thus, generating a model able to predict the level of happiness.

Keywords: happiness, subjective welfare, quality of life, classification trees, datamining.

1 Introduction

The analysis and discussion of the quality of life during recent years has increased the number of theorists and investigations focused on the subjective welfare and overall happiness, this represents a boom in such topics around the world. No doubt, this matter remounts to Edward Diener (born in 1946), distinguished for his investigations on happiness and subjective welfare theories, which gave him the title "father of the study of happiness".

Diener is the author of three scales that aid researchers in the evaluation of welfare: The satisfaction with life scale (SWLS) that measures cognitive global judgements of satisfaction with life (Diener, E.; Emmons, R. A.; Larsen, R. J.; Griffin, S, 1985) it has been widely used by diverse researchers, scientists and academics. The scale of positive and negative experiences (SPANE) which evaluates the frequency of experimenting a variety of positive and negative emotions. The scale of flowering that measures the subjective perception of success in important areas of life such as relationships, self-steem and optimism (Diener E; Ryan K., 2009). More and more

researchers talk about "welfare" instead of "happiness", for the word happiness is only related with joy, while "welfare" involves a wide range of complex subjects. Seligman says is more useful to talk about welfare than happiness (Seligman, 2011). Other authors, such as Ed Diener, "father of the study of happiness", use both words interchangeably.

For some, the study of the happiness means a revolution, understood from a methodological, theoretical and political stand (Bruno, Fay, 2010); Fay established the relation between happiness, utility, unemployment and inflation, as well as marriage and self-employment increase, he also proposed a relation between happiness and institutional policies.

The Advisory Forum Scientific and Technological A.C. published a document and points out that it is necessary to define welfare on the basis of what is relevant for people: it presents a view on what the subjective welfare is and how to measure it, based on six principles: (1) it is based on the welfare reported by the individual; the information is recovered from a survey of one or more questions related to happiness or life satisfaction; (2) as the question is made to an individual, the answer is given by concrete human beings; (3) recognizes that the welfare is essentially subjective for it is a self-experience of each individual; (4) the focus implies that each individual has the responsibility of playing the decisive role on happiness; (5) once it is accepted that there is relevant information on the welfare report, the focus follows a quantitative methodology to identify the important factors of the welfare of humans; (6) it is required an transdisciplinary approach or at least an interdisciplinary effort to understand the happiness reported by human beings (Foro Consultivo Cientifico y Tecnologico A.C., October 2012).

In the specific case of this paper, the following topic is developed: happiness focused on the case of a border city at the north of Mexico: Juarez City, municipality in the state of Chihuahua; where a non-lucrative and apolitical organization named *Plan Estrategico de Juarez* has the objective of constructing a citizen force invested and participative on public topics, proposing and demanding a better city (Plan Estratégico de Juárez A.C., 2018); one of its projects has the target of generating a system of city information to measure the quality of life of the citizens of Juarez: the System of Indicators *Asi Estamos Juarez (AEJ) or "This is how we are Juarez"*.

According to the organization, *AEJ* has the purpose of "measuring the quality of life in Juarez, creating a system of data that both general society and the government can use to have a view of what the current situation is, then, define a path to follow and what we ought to do in order to get to an ideal target". This system of data has information from official sources as well as a survey of perception made yearly since 2011 with an statistic representative sample and with a sampling technique that allows to generalize the results to the total population.

Each year, the survey of perception by *AEJ* has a sample of 1500 up to 1600 observations, through an structured survey, which are validated an applied face to face by an interviewer who is previously trained and supervised. The survey contains various topics impacting the quality of life and the subjective welfare of the juarenses, including education, health, corruption, security, citizen participation, among others such as variables of subjective welfare and happiness.

This article focuses in one of the variables: the Declared Happiness as a dependent variable and the correlation between the levels of happiness with socio-demographic

determinants that are provided by the survey such as age, gender, occupation, scholarship, migration possibilities and familiar income. The objective is to find which of those demographic characteristics have an impact on the happiness of the citizens. Definitely the topic of happiness is complex and must be studied from different points of view, therefore, the limits of this particular study are accepted, however, the study pretends to find, firstly, the correlations of the happiness of the juarenses starting with some of their socio-demographic aspects with the possibility of incorporate more variables to the analysis in future studies.

For the analysis of the socio-demographic determinants that may define happiness, a tool of datamining called decision trees is used, this tool creates a model of classification based on flow diagrams whose objective is to classify the cases in groups or prognosticated values of a dependent variable (criteria) from independent variables (predictors). The decision trees are a statistical technique for the segmentation, stratification, prediction, reduction of data and filtering of variables, identification of interactions, categories fusion and discretization of continuous variables (Vaneza Berlanga Silvente; María José Rubio Hurtado; Ruth Vila Baños, 2013). Decision Trees are among the most used algorithms for solving supervised classification problems (Franco-Arcega, Carrasco-Ochoa, & Martínez-Trinidad, 2013). For example, a body shape predictor where they show how an ensemble of regression trees can be used to estimate face landmark positions directly from a sparse subset of pixel intensities, achieving real-time performance with high quality predictions. Similarly, they have trained a shape model by estimating body landmark locations to evaluate (Trejo & Angulo, 2016).

Since the objective is predicting the levels of happiness of a juarense, given the demographic characteristics, the classification and regression trees (CRT) were used, this trees consist of algorithms made from complete binary trees that make partitions of the data and generates accurate and homogeneous subsets; CRT divides the data in segments so it is as homogenous as possible with respect to the dependent variable.

2   Related Works

At present, there are many efforts and works to measure happiness and subjective well-being in the world. Examples of this are:

a) The World Happiness Report is a landmark survey about the state of global happiness. The World Happiness Report 2018, which ranks 156 countries by their happiness levels, and 117 countries by the happiness of their immigrants. was released on March 14th during a launching event at the Pontifical Academy of Sciences in the Vatican. A launch event was also held on March 20th, celebrating International Day of Happiness at the United Nations (Helliwell, Layard, & Sachs, 2018).

The average ladder score (the average answer to the Cantril ladder question, asking people to evaluate the quality of their current lives on a scale of 0 to 10) for each country, based on the average happiness of the targeted countries over the span of the following years 2015-2017.

Karla Erika Donjuan Callejo, Mario Ricardo Sotomayor, Alberto Ochoa-Zezzatti

Furthermore, the Happiness Report correlates and explains the average happiness ranking by country through 6 variables: GDP per capita, social support, healthy life expectancy, social freedom, generosity, and absence of corruption. In their findings regarding world happiness, Mexico can be found in the ranks of the first 25 countries, being positioned place 24th out of 156 countries.
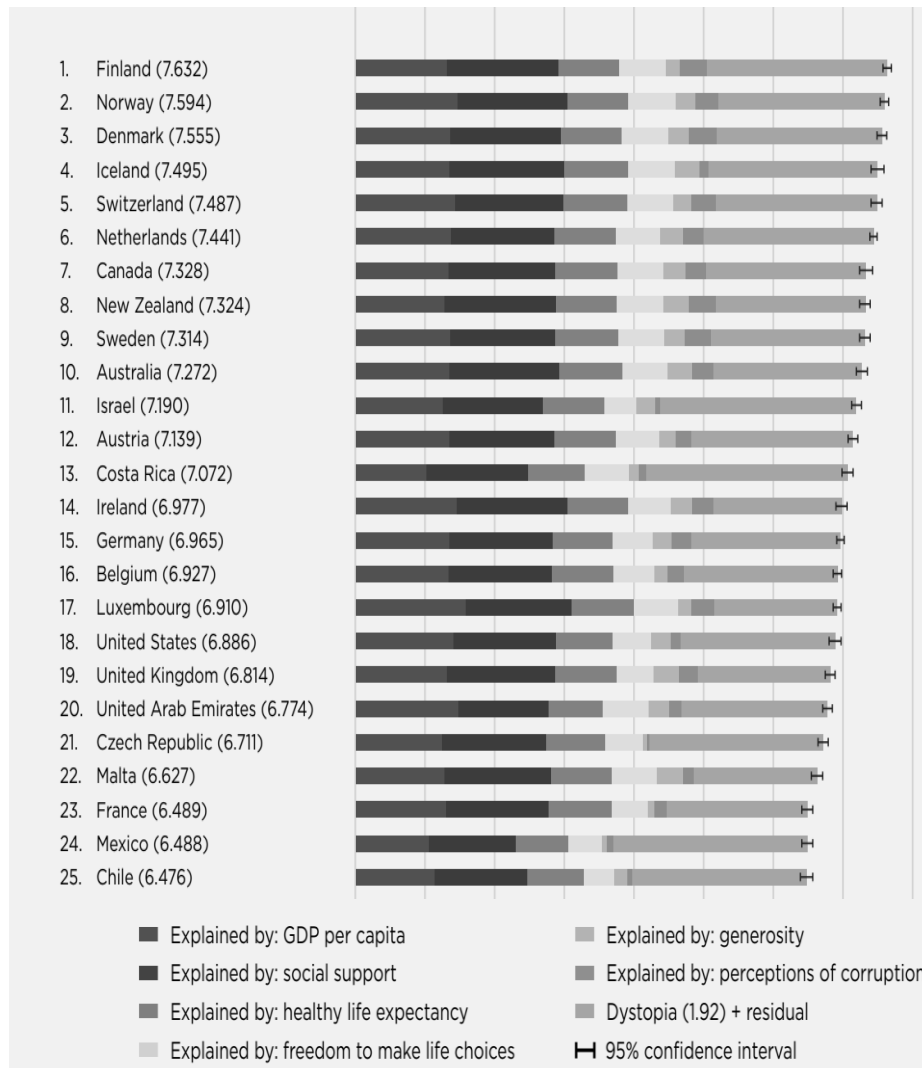

Fig. 1. Ranking of Happiness (Top 25). Source: The World Happiness Report 2015-2017.

b) The Happy Planet Index measures what matters: sustainable wellbeing for all. It tells us how well nations are doing at achieving long, happy, sustainable lives. Wealthy Western countries, often seen as the standard of success, do not rank highly on the Happy Planet Index. Instead, several countries in Latin America and the Asia

Pacific region lead the way by achieving high life expectancy and wellbeing with much smaller Ecological Footprints. The Happy Planet Index provides a compass to guide nations, and shows that it is possible to live good lives without costing the Earth (New Economics Foundation (NEF), 2016). The aforementioned index takes into consideration four determinants: (1)Wellbeing: How satisfied the residents of each country feel with life overall, on a scale from zero to ten. (2) Life expectancy: The average number of years a person is expected to live. (3) Inequality of outcomes: The inequalities between people within a country in terms of how long they live, and how happy they feel. (4) Ecological Footprint: The average impact that each resident of a country places on the environment. In its results about world happiness, Mexico can be found second place out of 140 countries of which it is compared.

Rank	Country	HPI	🙂	❤️	⚖️	👣
1	Costa Rica	44.7	7.3	79.1	15%	2.8
2	Mexico	40.7	7.3	76.4	19%	2.9
3	Colombia	40.7	6.4	73.7	24%	1.9
4	Vanuatu	40.6	6.5	71.3	22%	1.9
5	Vietnam	40.3	5.5	75.5	19%	1.7
6	Panama	39.5	6.9	77.2	19%	2.8
7	Nicaragua	38.7	5.4	74.3	25%	1.4
8	Bangladesh	38.4	4.7	70.8	27%	0.7
9	Thailand	37.3	6.3	74.1	15%	2.7
10	Ecuador	37.0	6.0	75.4	22%	2.2
11	Jamaica	36.9	5.6	75.3	21%	1.9
12	Norway	36.8	7.7	81.3	7%	5.0
13	Albania	36.8	5.5	77.3	17%	2.2
14	Uruguay	36.1	6.4	76.9	18%	2.9
15	Spain	36.0	6.3	82.2	10%	3.7
16	Indonesia	35.7	5.4	68.5	21%	1.6
17	El Salvador	35.6	5.9	72.5	22%	2.1
18	Netherlands	35.3	7.5	81.2	4%	5.3
19	Argentina	35.2	6.5	75.9	16%	3.1
20	Philippines	35.0	5.0	67.9	26%	1.1
21	Peru	34.6	5.8	74.1	21%	2.3
22	Palestine	34.5	4.6	72.6	24%	1.2
23	Brazil	34.3	6.9	73.9	22%	3.1
24	Switzerland	34.3	7.8	82.6	6%	5.8
25	Tajikistan	34.2	4.5	69.0	26%	0.9

Fig. 2. Happy Planet Index. Source: New Economics Foundation. The Happy Planet Index 2016. A Globlan index of sustainable wellbeing.

c) Better Life Index. This Index allows you to compare wellbeing across countries, based on 11 topics the OECD has identified as essential, in the areas of material living

conditions and quality of life: Housing, Income, Jobs, Community, Education, Enviroment, Civic Engagement, Healt, Life Satisfaction, Safety, Work-Life Balance. in its results, Mexico obtained a grade bellow 5, on a scale from 0 to 10.
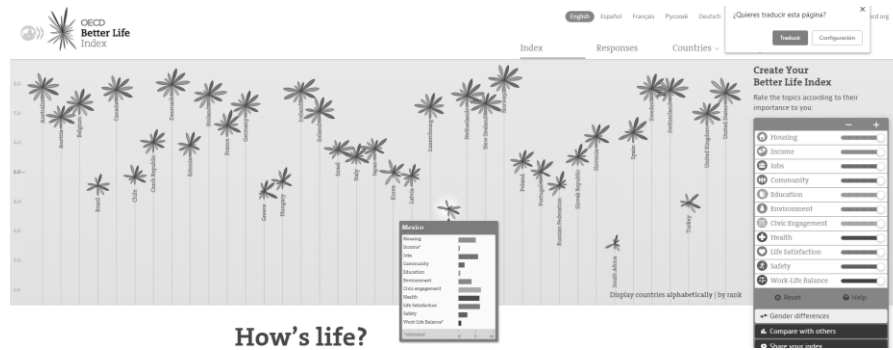

Fig. 3. Better Life Index. Source: OCDE, Better Life Index.

d) In 2007 Moyano and Ramos wrote about Subjective wellbeing: measuring life satisfaction, happiness and health in the Chilean population of the Maule Region. The purpose is to evaluate the subjective wellbeing through its cognitive components, measured as general satisfaction and its domains, and its affective components, measured as happiness, and analyze the relationship with sociodemographic variables. The sample was formed by 927 people, workers and students, between 17 and 77 years old, who answered three instruments.

The results indicate that the persons are satisfied with their life, being their family the principal source of happiness. Married people are happier and more satisfied than singles, and younger persons have lower level of happiness and satisfaction than older persons. Regarding to self-perceived health, men and women did not differ in their perception on general or physical health, although women reported a more negative perception about their mental health than men. Neither men nor women would stop working out of home even if they could do it. A direct and significant relation was found between happiness, self-reported health and vital satisfaction, which is supported by other studies in the field. Finally, there has been a positive correlation between income and welfare (Moyano Díaz & Ramos Alvarado, 2007).

3 Method

Happiness is a subjective concept that depends on diverse factors; for effects of this study, happiness is correlated with socio-demographic characteristics. The objective is to create a predictive model that allows to find out which of the 6 socio-demographic variables contained in the *AEJ* survey are correlated to the Declared Happiness by the juarenses. The model used is a decision tree that allows to classify the cases in groups to prognosticate the levels of happiness (criteria variable) based on the values of the independent variables (predictors). Thus, the classification and regression tree is accurate for the study.

The data comes from an statistical sampling of 1526 individuals, men and women living in Juarez municipality and with legal age, this means older than 18 years. The population (N) and the sample (n) are detailed in the following table:

Table 1. Sample.

Population (1)	Sample (2)	Confidence level and margin of sampling error (2)
1,391,180	1526	- 2.5% statistical error - 95% level of confidence

Source: own elaboration with information of (1) The National Institute of Geography, 2015 intercensal survey and information of *Plan Estrategico de Juarez A.C.* of its system of indicators *Asi Estamos Juarez*, Citizen Perception Survey 2016.

The population proportions statistical formula is used for the calculation of samples.

$$n = \frac{Z^2 * (p * q)}{e^2 + \frac{Z^2 * (p * q)}{N}}$$

(1)

in which:
n=Sample,
Z2=Sampling frame,
p=Success probability (Proportion of the population with the desired characteristics),
q=Failure probability (Proportion of the population without the desired characteristics),
e=error range (Level of error whiling to take),
N= Universe or population size.

The sampling technique used was probabilistic and multi-stage, surveys were applied face to face at the interviewee home address, distributed across multiple areas of the city.

4   Results: The Happiness of Juarenses and its Correlation with their Socio-demographic Determinants

The questions and scales used for each one of the variables are listed as follows:

1. *Dependent Variable: the Declared Happiness*

We've called this variable "Declares Happiness", since it implies the person's capacity to declare how much happy they are in a determined moment, that is to say, it affirms people's capacity of manifesting their happiness and adjust it to a determined scale.

The measuring of subjective welfare of the juarenses is made considering diverse factors, nonetheless, there is an specific question useful for the effects of the study: how happy are you? This question is answered using an scale of 1 to 10, where the extreme values 1 and 10, correspond to "not happy at all" and "very happy". This numeric scale allows the individuals to give a value to a subjective concept such as happiness. Other benefit of this scale is how comprehensible it is to the interviewed individuals thanks to the use of a similar scale in schools at Mexico where a 10 means the maximum scholar grade and 5 implies failing a class or assignment. The third advantage of the scale is it allows to calculate descriptive statistics: mean, mode and median of the Declared Happiness. Finally the problem of the measuring Semantic

Karla Erika Donjuan Callejo, Mario Ricardo Sotomayor, Alberto Ochoa-Zezzatti

Textual Similarity (STS), between words/ terms, sentences, paragraph and document plays an important role in computer science and computational linguistic is avoided (Majumder, Pakray, Gelbukh, & Pinto, 2016).

Table 2. Numeric scale (Declared Happiness).

1 2 Not happy at all	3 4 A little happy	5 6 Somewhat happy	7 8 happy	9 10 Very happy

The next tables show the frequency of answers, the result is an average of 8.27 in the Declared happiness of the people of Juarez.

Table 3. Declared Happiness.

Declared Happiness					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	17	1.1	1.1	1.1
	2	4	.3	.3	1.4
	3	14	.9	.9	2.3
	4	32	2.1	2.1	4.4
	5	75	4.9	5.0	9.4
	6	59	3.9	3.9	13.3
	7	133	8.7	8.8	22.0
	8	404	26.5	26.7	48.7
	9	329	21.6	21.7	70.4
	10	448	29.4	29.6	100.0
	Total	1515	99.3	100.0	
Missing	System	11	.7		
Total		1526	100.0		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 4. Descriptive Statistics Declared Happiness.

Declared Happiness		
N	Valid	1515
	Missing	11
Mean		8.27
Median		9.00
Mode		10

Source of Table 4: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Growing Method: CRT

Dependent Variable: RevealedHappiness

Fig. 1. Mean and Percentile (Dependent variable: Declared Happinness).

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.
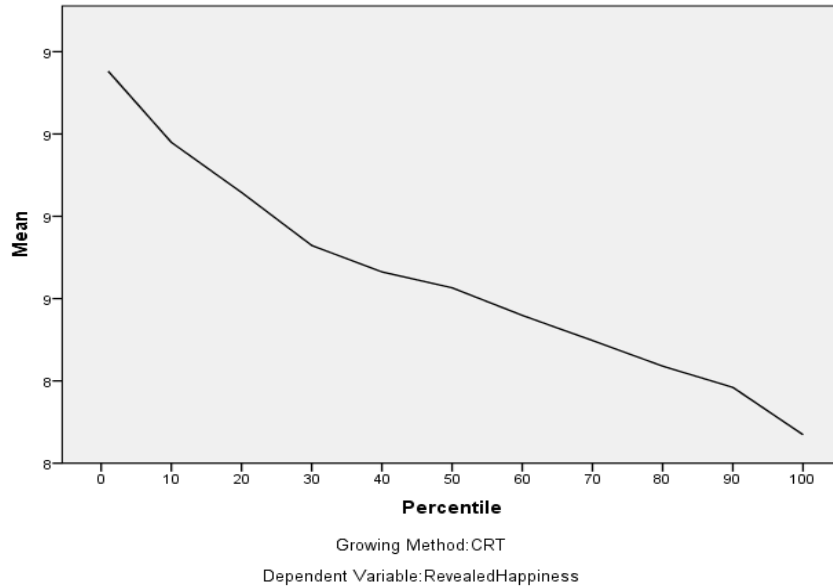
2 Independent Variables

As mentioned before, the study pretends to identify which variables have a correlation with the dependent variable. The socio-demographic data from the survey that belongs to the juarenses and will contribute to finding the Declared happiness is shown in the next table:

Table 5. Independent variables.

#	Independent variable	Observations
1	Age	Numerical variable. Considering the survey is answered by legal citizens, this means people with over 18 years of age.
2	Gender	Codified variables: 1. Woman 2. Man
3	Principal occupation	Codified variables: 1. Housekeeper 2. Employee of private company 3. Employee of the government 4. Unemployed 5. Student 6. Student who works 7. Employer 8. Retired 9. Self-employment

143

#	Independent variable	Observations
4	Scholarship	Codified variables: 1. Illiterate 2. Complete elementary school 3. Incomplete elementary school 4. Complete middle school 5. Incomplete middle school 6. Complete high school 7. Incomplete high school 8. Complete degree (college) 9. Incomplete degree (college) 10. Postgraduate 11. Complete technical career 12. Incomplete technical career 13. Knows how to read and to write but did not went to school
5	Migration intention	This variable is measured with a scale of four points that aims to size how much the population have thought of moving out of the city. The specific question is: In the last year, how often did you thought of moving out of the city? And the codified variables are: – A Lot of times – Sometimes – Few times – Never
6	Familiar income	This variable is measured with a scale of four points as well and is about the familiar income and how much this is or is not enough for the consumption and savings of the families in the city. The answers are listed as follows: With the total of the familiar income, you would say that… 1. It is more than enough and we can save money 2. It is just enough with no difficulties 3. It is not enough and have some difficulties 4. It is not enough and have a lot of difficulties 0. Does not know / Did not answer

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

The following tables contain the frequencies in the responses of the Juarenses, for each of the independent variables.

Table 6. Independent variable: Gender.

Gender					
Code		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	869	56.9	60.1	60.1
	2	577	37.8	39.9	100.0
	Total	1446	94.8	100.0	
Missing	System	80	5.2		
Total		1526	100.0		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 7. Independent variable: Principal Occupation.

Occupation					
Code		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	456	29.9	30.5	30.5
	2	378	24.8	25.3	55.8
	3	159	10.4	10.6	66.5
	4	55	3.6	3.7	70.1
	5	60	3.9	4.0	74.2
	6	54	3.5	3.6	77.8
	7	52	3.4	3.5	81.3
	8	138	9.0	9.2	90.5
	9	142	9.3	9.5	100.0
	Total	1494	97.9	100.0	
Missing	System	32	2.1		
Total		1526	100.0		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 8. Independent variable: Scholarship.

Scholarship					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	26	1.7	1.7	1.7
	2	248	16.3	16.4	18.1
	3	102	6.7	6.7	24.9
	4	372	24.4	24.6	49.5
	5	79	5.2	5.2	54.7
	6	184	12.1	12.2	66.9
	7	62	4.1	4.1	71.0
	8	158	10.4	10.4	81.4
	9	139	9.1	9.2	90.6
	10	25	1.6	1.7	92.3
	11	91	6.0	6.0	98.3
	12	9	.6	.6	98.9
	13	17	1.1	1.1	100.0
	Total	1512	99.1	100.0	
Missing	System	14	.9		
Total		1526	100.0		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Results of the Model: Classification Tree

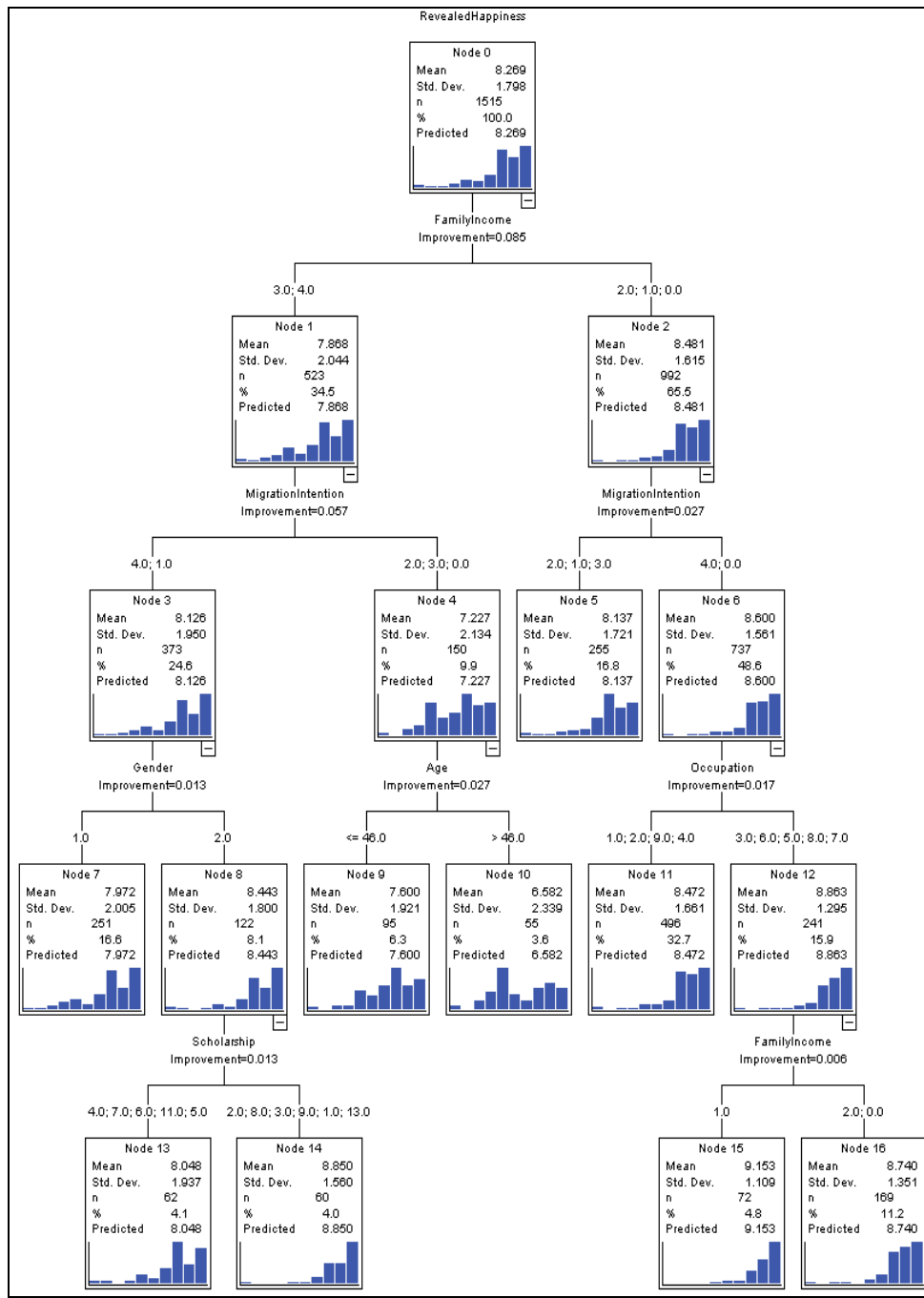Fig. 4 shows the resulting classification tree.

Fig. 4. Classification Tree.
Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 9. Independent variable: Migration intention.

		Frequency	Percent	Valid Percent	Cumulative Percent
		Migration Intention			
Valid	0	24	1.6	1.6	1.6
	1	94	6.2	6.2	7.7
	2	192	12.6	12.6	20.3
	3	161	10.6	10.6	30.9
	4	1055	69.1	69.1	100.0
	Total	1526	100.0	100.0	

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 10. Independent variable: Family Income.

		Frequency	Percent	Valid Percent	Cumulative Percent
		FamilyIncome			
Code					
Valid	0	18	1.2	1.2	1.2
	1	189	12.4	12.4	13.6
	2	793	52.0	52.0	65.5
	3	465	30.5	30.5	96.0
	4	61	4.0	4.0	100.0
	Total	1526	100.0	100.0	

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Interpretation of the classification tree

When the model is tested, the variables of gender, occupation and scholarship are discarded as the main predictive variables of happiness. The two most important predictors of Declared Happiness are: family income and the intention to migrate.

1. The node 0 represents the independent variable which is the happiness of the individual, from 1515 observed cases, the average of happiness was 8.269.
2. Principal predictive variable: Therefore is observed that the dependent variable (happiness) branches into two nodes. The node 1 and node 2 belong to the variable of familiar income, so the classification says this is the principal predictive variable. If the results are analyzed, it is observed that the group who answered their income is "more than enough and we can save" is happier as well as the group of those who said their income is "just enough with no difficulties", both with an average of happiness of 8.48 (this group is the node 2). As for the node 1, the groups whose income is "not enough and have some difficulties" and "not enough and have a lot of difficulties" have an average of happiness of 7.86 out of 10 points.
3. Second predictive variable: The following structure of the tree branches into 4 more nodes, all of them from the second predictive variable: has thought of moving out of the city. From the node 1 are given the nodes 3 and 4, even though in both cases it is observed a lower average of happiness than the

nodes 5 and 6, coincide that the juarenses who have never thought of leaving the city are, in fact, the ones with a higher mean of happiness. The nodes 5 and 6 contain the juarenses with better familiar income and when they are reclassified, they are also the people who have never thought of leaving the city.

4. Third predictive variable. node 3 divides into the third predictive variable which is the gender of the juarense, and, node 4 divides into the predictive variable which is the age. From the node 3 are divided the nodes 7 and 8, the results of predictive variable of the gender say men are happier than women . While from the node 4 are given the nodes 9 and 10, the results of the predictive variable of the age say the happier people have less than 46 years. From the node 5 there are no more predictive variables, and, from de node 6 is divide the predictive variable is de occupation.

5. The last predictive variable from the node 8 is Scholarship and from node 12 the family income.

6. The independent variables importance with respect to the dependent (Declared happiness) sum together 28.7%, that is, happiness is determined by them in that percentage. That is the impact of sociodemographic characteristics on the happiness of the people of Juarez.

Table 11. Independent Variable Importance.

Independent Variable Importance		
Independent Variable	Importance	Normalized Importance
FamilyIncome	.104	100.0%
MigrationIntention	.085	82.2%
Age	.038	36.5%
Occupation	.027	26.5%
Scholarship	.020	19.6%
Gender	.013	12.8%
Growing Method: CRT Dependent Variable: DeclaredHappiness		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

Table 12. Risk.

Risk		
Method	Estimate	Std. Error
Resubstitution	2.991	.173
Cross-Validation	3.202	.181
Growing Method: CRT Dependent Variable: DeclaredHappiness		

Source: own elaboration with information of Plan Estrategico de Juarez A.C. of its system of indicators Asi Estamos Juarez, Citizen Perception Survey 2016.

5   Discussion and Conclusions

With the model of classification regression tree is established that, in the case of the juarense community and the Declared happiness of the year 2016, it is found that gender, occupation and scholarship are not the main variables that influence happiness. Thus, it is established nodes that define the socio-demographic profile for happiness (variables that influence in more happiness) are the variables of familiar income as principal predictive variable as far as the income is more than enough that people may even save money; that the possibility of leaving the city is not in the mind of the individuals and the in terms of age, the younger they are, the happier. That is to say, it is possible to predict that an individual with enough income to satisfy their necessities and to save, that has never thought of leaving the city and has an age less than 46 years, is more likely to have higher levels of happiness than an individual with not enough income and difficulties that also has thought of leaving the city.

However, it is important to note, although family income was the main predictable variable of happiness, there are multiple authors who claim the aforementioned statement is only true when talking about poor and third world countries, since they have found a close to zero correlation between income and happiness in rich countries, they also have found that the relation between income and wellbeing have relevant effects in situations of extreme poverty and that individuals adapt to their corresponding economical level, hence, a lost of economic status can lead to unhappiness (Veenhoven, 1994; Dinner, 1994; Dinner E, Suh, E., , Lucas, R., & Smith, H., 1999).
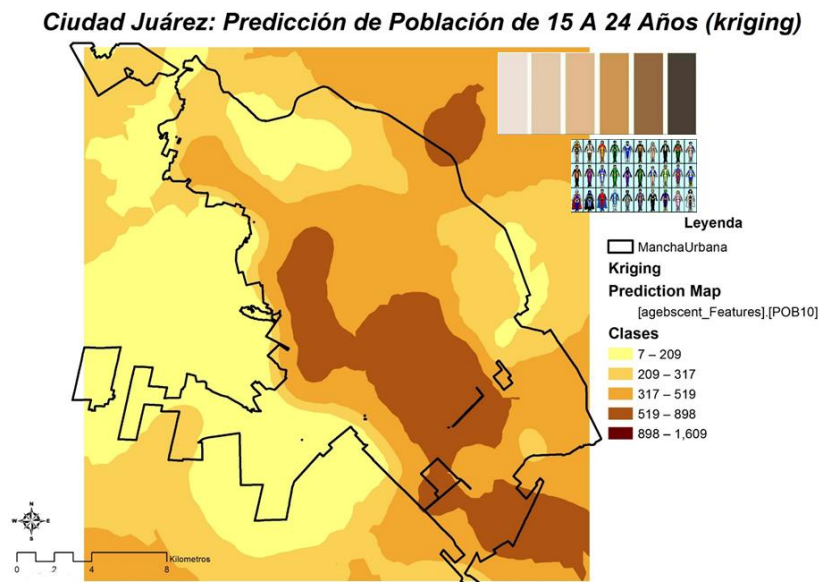

Fig. 4. Prediction model to 2030 including distribution of young people and a future opportunities model. (Source: Adapted of a map using Kriging Model generated in Colech by Luis Cervera.)

Definitely the subject of happiness is complex and is determined by a lot of variables as well as the disciplines that study it. So this study is limited to trying to find out a socio-demographic profile with 6 variables in order to aid in the improvement of happiness research. There is still a lot of work to do to understand everything we need about this topic in order to change the subjective welfare of people and to improve the quality of life.

With the model of classification regression tree is established that, in the case of the juarense community and the Declared happiness in the year 2017, was not found that gender, occupation and scholarships are the main variables that influence happiness. Thus, it is established nodes that define the socio-demographic profile for happiness (variables that influence in more happiness) are the variables of familiar income as principal predictive variable as far as the income is more than enough that people may even save money; that the possibility of leaving the city is not in the mind of the individuals and the age, the younger they are, the happier. That is to say, it is possible to predict that an individual with enough income to satisfy their necessities and to save, that has never thought of leaving the city and has an age less than 46 years, is more likely to have higher levels of happiness than an individual with not enough income and difficulties that also has thought of leaving the city.

Definitely the theme of happiness is complex and is determined by a lot of variables as well as the disciplines that study it. So this study is limited to try finding out a socio-demographic profile with 6 variables in order to help improving the research on happiness. There is still a lot of work to do and to understand about this topic to change the subjective welfare of people and to improve the quality of life.

In our future work, we will seek to generate a shared index values by regions in order to adequately model emerging public policies for each segment of the population, considering their geospatial location as can be seen in Figure 4.

References

1. Diener, E; Ryan, K.: Subjective well-being: A general overview. Journal of Psychology, 391-406 (2009)
2. Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S.: The satisfaction with life scale (1985)
3. Dinner, E, Suh, E., Lucas, R., Smith, H.: Subjetive well-being: three decades of progress. Psychological Bulletin, 276-302 (1999)
4. Dinner, E.: El bienestar subjetivo. Intervención psicosocial. Revista sobre igualdad y calidad de vida, 67-113 (1994)
5. Foro Consultivo Científico y Tecnológico A.C.: Medición, Investigación, e Incorporación a la Política Pública del Bienestar Subjetivo: América Latina. Comisión para el Estudio y la Promoción del Bienestar en América Latina, México (Octubre 2012)
6. Franco-Arcega, A., Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F.: Decision Tree based Classifiers for Large Datasets. Computación y Sistemas 17(1), 95-102 (2013)
7. Helliwell, J., Layard, R., Sachs, J.: World Happiness Report. Reporte, The World Happiness Report was written by a group of independent experts acting.

Recuperado el Junio de 2018, de https://s3.amazonaws.com/happiness-report/2018/WHR_web.pdf (2018)

8. Majumder, G., Pakray, P., Gelbukh, A., Pinto, D.: Semantic Textual Similarity Methods, Tools, and Applications: A Survey. Computación y Sistemas, 20(4), 647-665 (2016)

9. Moyano-Díaz, E., Ramos-Alvarado, N.: Bienestar subjetivo: midiendo satisfacción vital, felicidad y salud en población chilena de la Región Maule. (Talca, Ed.) Universum, 22(2). Recuperado el Junio de 2018, de https://scielo.conicyt.cl/scielo.php?pid=S0718-23762007000200012& script=sci_arttext&tlng=pt (2007)

10. New Economics Foundation (NEF). The Happy Planet Index. NEF. Recuperado el Junio de 2018, de https://static1.squarespace.com/static/ 5735c421e321402778ee0ce9/t/57e0052d440243730fdf03f3/1474299185121/Bri efing+paper+-+HPI+2016.pdf (2016)

11. Plan Estratégico de Juárez A.C. www.planjuarez.org. Obtenido de www.planjuarez.org: https://www.planjuarez.org/index.php/quienes (Junio de 2018)

12. Seligman, M.: Flourish: A Visionary New Understanding of Happiness and Well-being. Free Press (2011)

13. Trejo, K., Angulo, C.: Single-Camera Automatic Landmarking for People Recognition. Computación y Sistemas, 20(1), 19-28 (2016)

14. Berlanga-Silvente, V., Rubio-Hurtado, M.J., Vila-Baños, R.: Cómo aplicar árboles de decisión en SPSS. REIRE, Revista d´Innovació i Recerca en Educació, 65-79 (2013)

15. Veenhoven, R.: El estudio de la satisfacción con la vida. Intervención Psicosocial, 3, 87-116 (1994)