

# Customer Experience in Hotels

23th February 2020

## Objective

**"Maximize the impact in overall satisfaction based on survey results"**

## Survey

The survey contains the following field:

### Fields:

1. Overall Satisfaction (from -1 to 7)
2. Room Comfort Satisfaction (from 1 to 5)
3. Room Cleanness Satisfaction (from 1 to 5)
4. Reception Satisfaction (from 1 to 5)
5. Breakfast Satisfaction (from 1 to 5)
6. Entertainment Satisfaction (from 1 to 5)
7. Nationality (ES, FR, US, DE, UK)
8. Regime (all inclusive, NaN)
9. Segment (family, couple, friends, business)

**Number of Surveys:** 1000

### Percentage of fields not completed:

1. Overall Satisfaction: 0%
2. Room Comfort Satisfaction: 1%
3. Room Cleanness Satisfaction: 5.1%
4. Reception Satisfaction: 0.8%
5. Breakfast Satisfaction: 1.3%
6. Entertainment Satisfaction: 28.9%
7. Nationality: 0%

- 8. Regime: 19.7%
- 9. Segment: 30.4%

Of you can see in the following list. There are a lost of blanks in some fields. Especially, in segment, entertainment satisfaction and regime. The categorical variables was solved replacing null values to mode. In regime, I have assumed that “all inclusive” is the unique option (null vs all inclusive). However, I recommend contrast this information with the business manager.

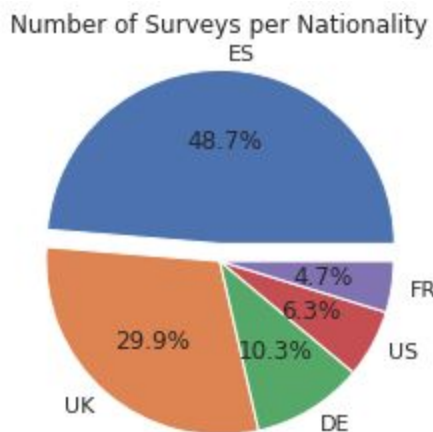
Regarding null “segments” values . I tried different models to replace it with predictions but I didn’t feel comfortable with their accuracy and I decided drop null rows.

### Recommendations

To avoid nulls is recommendable convert all fields into mandatory fields. Maybe UX department can redesign the survey to improve the user experience of these form. If hotel use paper survey, this project is the best excuse to migrate to online version.

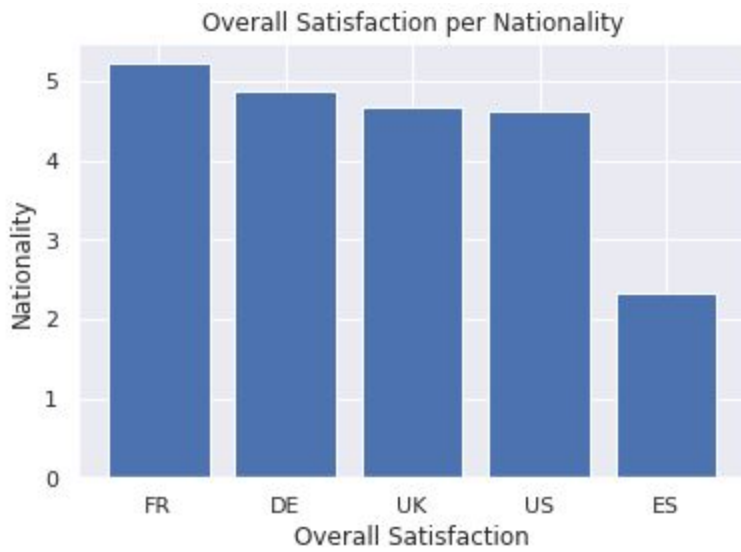
## Results

### Number of surveys per nationality



The 48,7% of surveys have been filled out by spanish clients, followed by english (29,9%) and german (10,3%).

### Overall Satisfaction per nationality

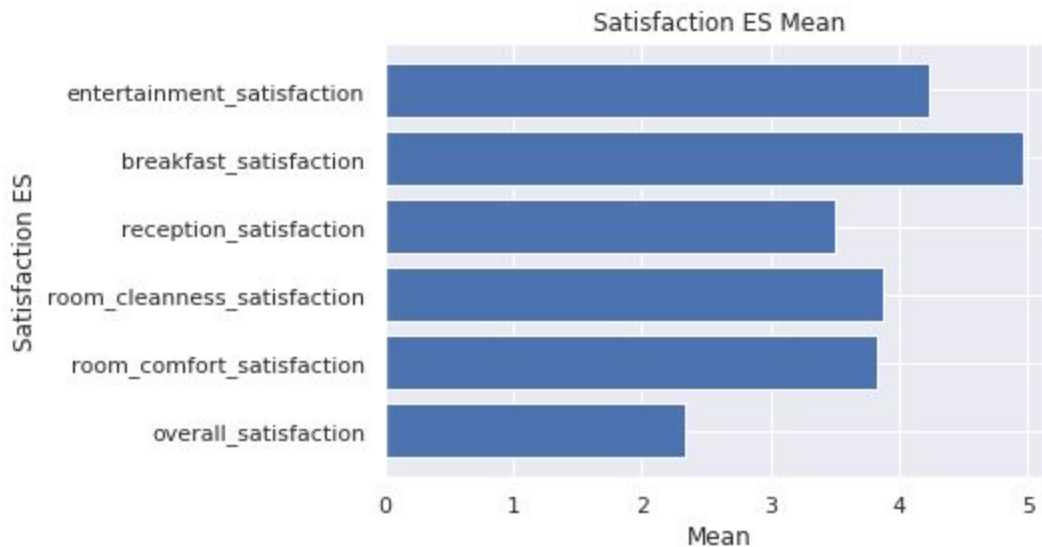


Spanish clients are the most critical. Their overall satisfaction mean are 2,3. While the rest of nationalities are above 4.

nationality	overall_satisfaction
FR	5.219848
DE	4.860083
UK	4.677385
US	4.614977
ES	2.336183

## Spanish Satisfaction

If we deep in spanish surveys we can analyze every satisfaction:

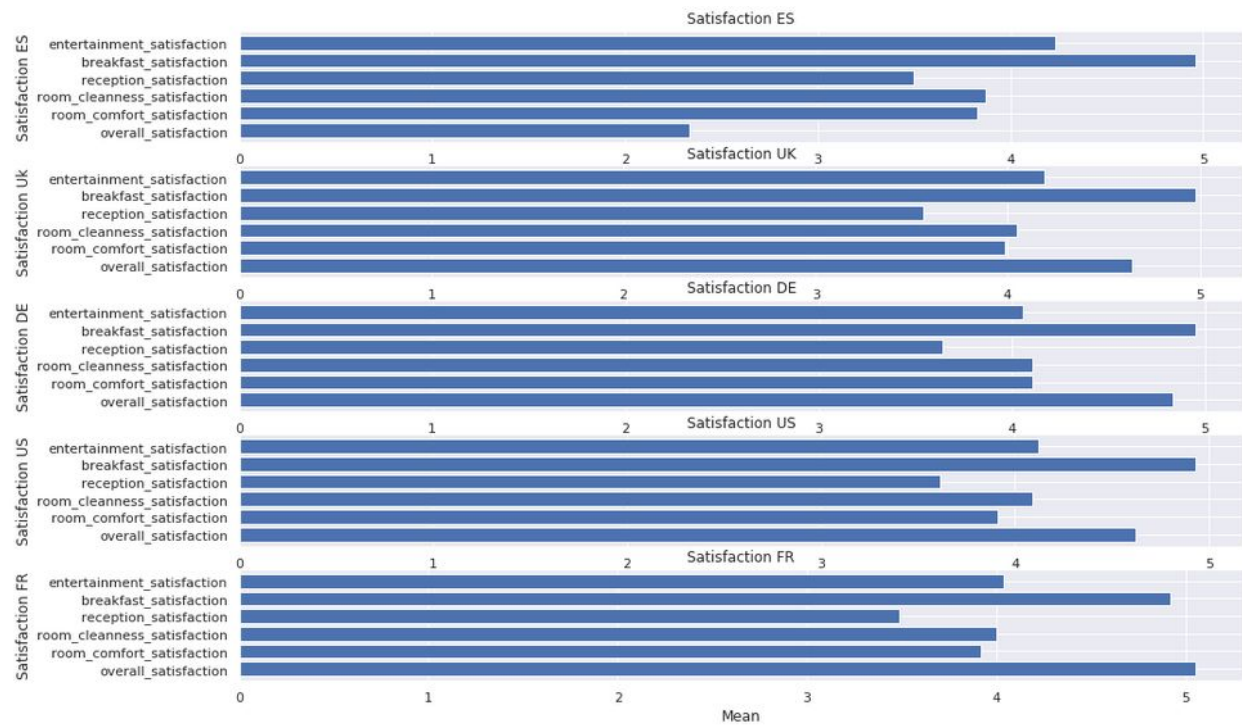


How you can see the worst satisfactions are reception (3,4) and room confort (3,8). However, this number doesn't justify the low score in overall satisfaction.

satisfaction_es	mean
overall_satisfaction	2.336930
room_comfort_satisfaction	3.829569
room_cleanness_satisfaction	3.870637
reception_satisfaction	3.498973
breakfast_satisfaction	4.958932
entertainment_satisfaction	4.232033

For this reason, I recommend introduce new questions to include another kinds of satisfaction. Hotel can use their call center service to speak with spanish customers and analyze the real cause or causes of this dissatisfaction.

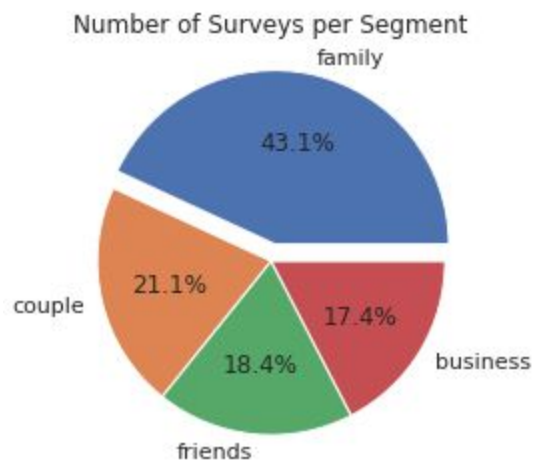
## Summary satisfaction per nationality



All nationalities are satisfied with breakfast and entertainment . However, are dissatisfied with reception and comfort.

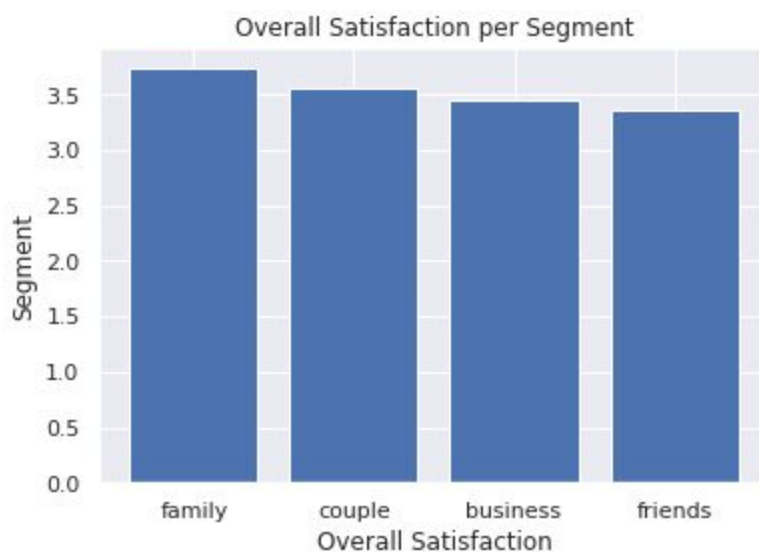
### Number of surveys per segment

The 43,1% of surveys have been filled out by family clients, followed by couples (21,1%), friends (18,4%) and business (17,4%).



### Overall Satisfaction per segment

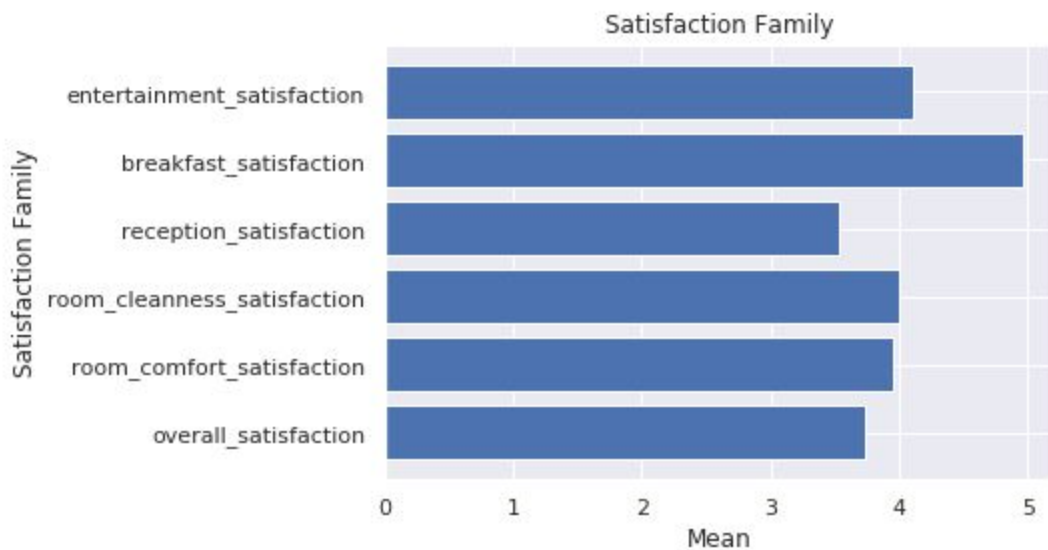
Family (3,7) clients are the most satisfied segment, followed by couples (3,5) and business (3,4). Friends (3,3) are the most dissatisfied group.



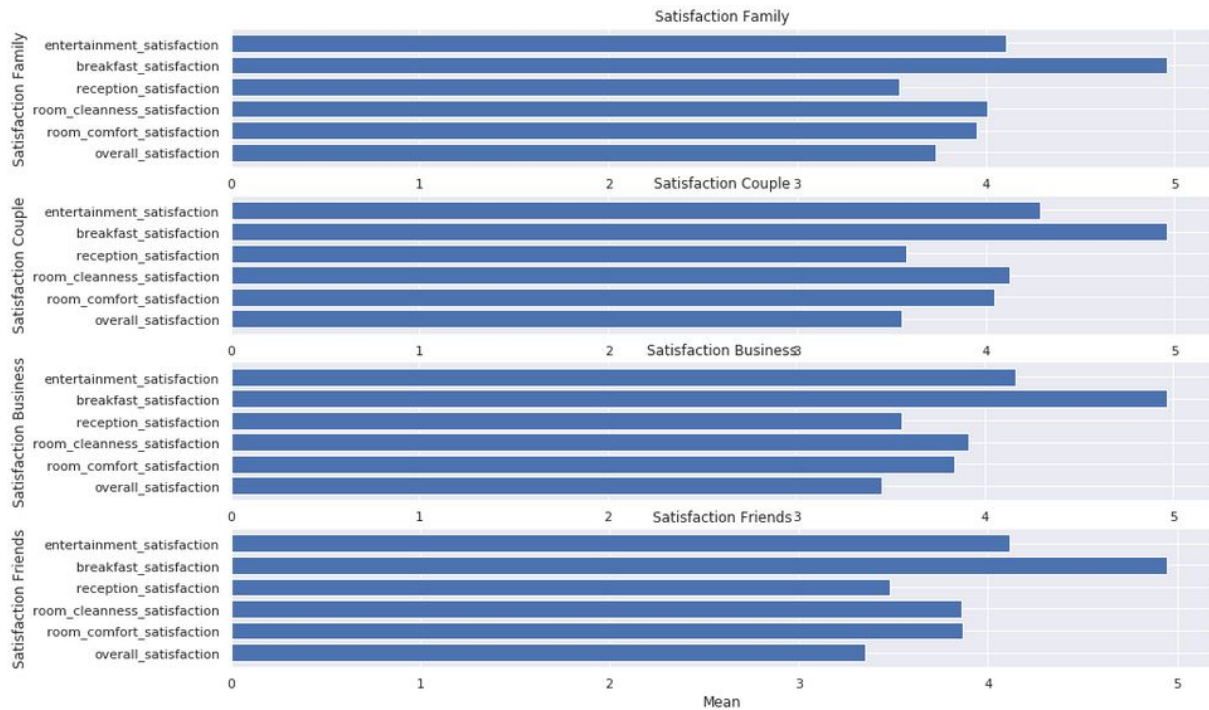
segment	overall_satisfaction
family	3.734290
couple	3.556823
business	3.451736
friends	3.353922

### Family Satisfaction

The same pattern are repeated in families segment, reception and room confort are the main cause of dissatisfaction:



## Summary satisfaction per segment



All segments reflect the situation that we show in the before analysis (reception and comfort are the worst satisfactions).

## Relation between variables

The variables more related to overall satisfaction are room comfort satisfaction, room cleanliness, and reception satisfaction. If these grow overall satisfaction too. Regarding nationalities, being spanish can have a negative effect and being english positive. The rest of variables are not very correlated.



	overall_satisfaction
overall_satisfaction	1.00
room_comfort_satisfaction	0.57
room_cleanness_satisfaction	0.55
reception_satisfaction	0.48
breakfast_satisfaction	0.04
entertainment_satisfaction	0.02
id	0.03
nationality_DE	0.28
nationality_ES	-0.79
nationality_FR	0.24
nationality_UK	0.47
nationality_US	0.18
segment_business	-0.04
segment_couple	-0.01
segment_family	0.09
segment_friends	-0.07

## Model (Linear Regression):

### Model 1: All Variables

In the first model I use all variable:

OLS Regression Results						
Dep. Variable:	overall_satisfaction	R-squared:	0.886			
Model:	OLS	Adj. R-squared:	0.884			
Method:	Least Squares	F-statistic:	442.6			
Date:	Sat, 22 Feb 2020	Prob (F-statistic):	1.64e-312			
Time:	19:29:11	Log-Likelihood:	-531.88			
No. Observations:	696	AIC:	1090.			
Df Residuals:	683	BIC:	1149.			
Df Model:	12					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.6671	0.349	1.913	0.056	-0.017	1.352
room_comfort_satisfaction	0.5397	0.054	10.008	0.000	0.434	0.646
room_cleaness_satisfaction	0.0380	0.049	0.783	0.434	-0.057	0.133
reception_satisfaction	0.2002	0.042	4.751	0.000	0.117	0.283
breakfast_satisfaction	0.0555	0.100	0.554	0.579	-0.141	0.252
entertainment_satisfaction	0.0111	0.017	0.646	0.518	-0.023	0.045
nationality_DE	0.5221	0.090	5.770	0.000	0.344	0.700
nationality_ES	-1.7597	0.078	-22.599	0.000	-1.913	-1.607
nationality_FR	0.9334	0.102	9.152	0.000	0.733	1.134
nationality_UK	0.4687	0.082	5.683	0.000	0.307	0.631
nationality_US	0.5027	0.097	5.156	0.000	0.311	0.694
segment_business	0.0988	0.097	1.018	0.309	-0.092	0.289
segment_couple	0.1788	0.095	1.874	0.061	-0.009	0.366
segment_family	0.2573	0.092	2.799	0.005	0.077	0.438
segment_friends	0.1323	0.094	1.403	0.161	-0.053	0.317
Omnibus:	22.288	Durbin-Watson:	1.905			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	49.271			
Skew:	-0.099	Prob(JB):	2.00e-11			
Kurtosis:	4.288	Cond. No.	6.16e+16			

#### Warnings:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.  
 [2] The smallest eigenvalue is 1.64e-29. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

Model before have multicollinearity because there are a lot of variables that have a big correlation between them. For this reason we must chose the best ones (p values are lower 0,05)

Despite having an optimal R-squared there are a lost of variables with p-values higher than 0,05 and strong multicollinearity problems. For this reason, we must eliminate variables.

**Model 2:**

OLS Regression Results						
Dep. Variable:	overall_satisfaction	R-squared:	0.886			
Model:	OLS	Adj. R-squared:	0.884			
Method:	Least Squares	F-statistic:	760.3			
Date:	Sat, 22 Feb 2020	Prob (F-statistic):	7.61e-319			
Time:	19:53:41	Log-Likelihood:	-533.46			
No. Observations:	696	AIC:	1083.			
Df Residuals:	688	BIC:	1119.			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	1.0584	0.081	13.015	0.000	0.899	1.218
room_comfort_satisfaction	0.5790	0.029	20.116	0.000	0.522	0.635
reception_satisfaction	0.1967	0.042	4.686	0.000	0.114	0.279
nationality_DE	0.5979	0.059	10.194	0.000	0.483	0.713
nationality_ES	-1.6808	0.036	-46.859	0.000	-1.751	-1.610
nationality_FR	1.0083	0.080	12.614	0.000	0.851	1.165
nationality_UK	0.5491	0.041	13.461	0.000	0.469	0.629
nationality_US	0.5839	0.071	8.243	0.000	0.445	0.723
segment_family	0.1172	0.040	2.915	0.004	0.038	0.196
Omnibus:	22.963	Durbin-Watson:	1.905			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	51.495			
Skew:	-0.102	Prob(JB):	6.58e-12			
Kurtosis:	4.317	Cond. No.	2.44e+16			

**Warnings:**

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.  
 [2] The smallest eigenvalue is 3.66e-29. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

In this case, all p-values are lower than 0,05. However, we follow with multicollinearity problems. We must eliminate more variables. Nationality are very correlated for these reason I eliminated someones.

**Model 3:**

OLS Regression Results

Dep. Variable:	overall_satisfaction	R-squared:	0.883
Model:	OLS	Adj. R-squared:	0.882
Method:	Least Squares	F-statistic:	1040.
Date:	Sat, 22 Feb 2020	Prob (F-statistic):	2.11e-318
Time:	19:56:03	Log-Likelihood:	-541.52
No. Observations:	696	AIC:	1095.
Df Residuals:	690	BIC:	1122.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1.7478	0.105	16.643	0.000	1.542	1.954
room_comfort_satisfaction	0.5844	0.029	20.173	0.000	0.527	0.641
reception_satisfaction	0.1885	0.042	4.455	0.000	0.105	0.272
nationality_ES	-2.3650	0.052	-45.286	0.000	-2.468	-2.262
nationality_UK	-0.1359	0.057	-2.391	0.017	-0.248	-0.024
segment_family	0.1236	0.041	3.046	0.002	0.044	0.203

Omnibus:	21.641	Durbin-Watson:	1.902
Prob(Omnibus):	0.000	Jarque-Bera (JB):	47.459
Skew:	-0.090	Prob(JB):	4.95e-11
Kurtosis:	4.267	Cond. No.	31.6

**Warnings:**

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Final model was made up of 5 variables (room\_comfort\_satisfaction, reception\_satisfaction, nationality\_ES, nationality\_UK and segment\_family). It R-squared are 0.88 and it don't have multicollinearity problems.

**Model Formula:**

$y_i = \text{intercept\_} + (\text{room\_comfort\_satisfaction} * \text{room\_comfort\_satisfaction\_}) +$   
 $(\text{reception\_satisfaction} + \text{reception\_satisfaction\_})$   
 $+ (\text{nationality\_ES} * \text{nationality\_ES\_}) + (\text{nationality\_UK} * \text{nationality\_UK\_}) + (\text{segment\_family} * \text{segment\_family\_})$

\_: coefficients

## Conclusions

1. The variables more correlated with overall satisfaction are room comfort satisfaction, room cleanliness satisfaction and reception satisfaction.
2. Spanish clients are the most critical and dissatisfied.
3. French are the most satisfied.
4. Family are the most satisfied segment.
5. Friend clients are the most dissatisfied segment.
6. Hotel should improve comfort rooms and reception service.
7. If they would improve quickly it is important start with comfort rooms.
8. Hotel should analyze the real cause of spanish dissatisfaction.
9. For future analysis is recommendable improve the quality of the polls, using web form and working with the user experience of these module.