

Predicts Individual Life Satisfaction by Demographic Characteristics

Heran Zhou, Xinyu Zhong, Yaqing Feng, Yuhan Gu

Octobor 19, 2020

github: <https://github.com/Kevinzhou0717/STA304-Problem-Set-2>

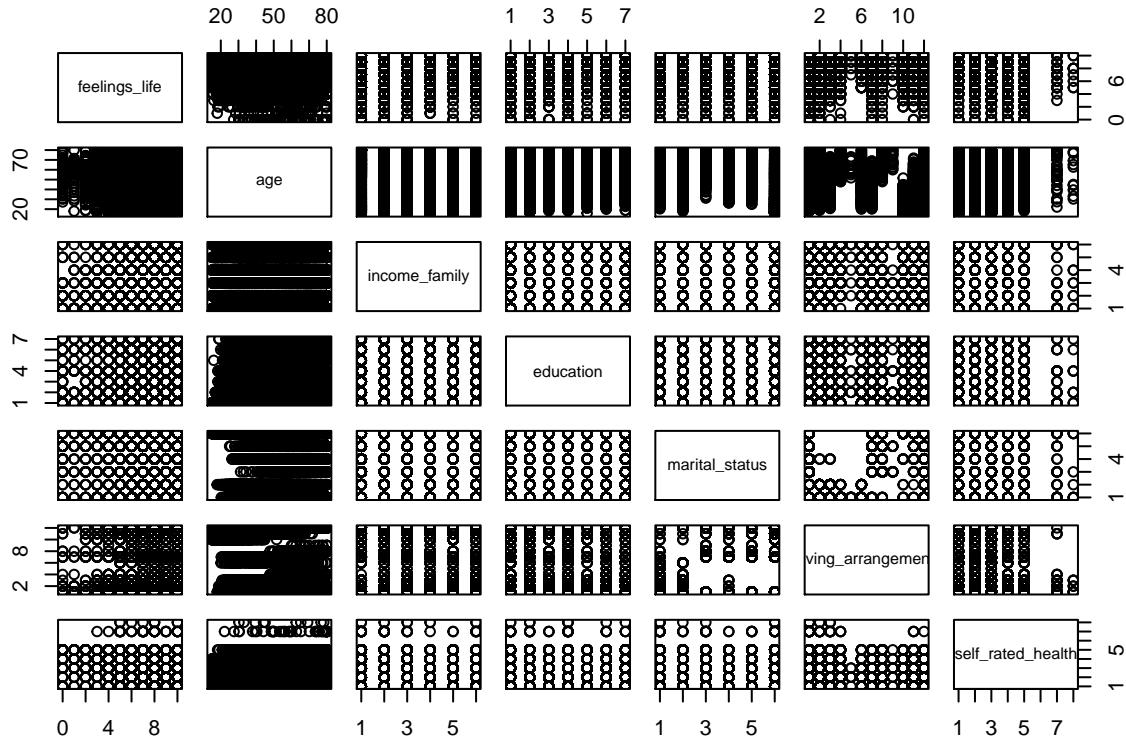


Figure 1: Plot of raw data

Refer to Figure 1.

1 Abstract

2 Introduction

3 Data

4 Model

Hence, our model is

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_3 + \beta_5 x_4 + \beta_6 x_5 + \beta_7 x_6 + \epsilon$$

where p represent the probability to have life satisfaction below average. $x_i (i=1, \dots, 6)$ represent gender, education level, living area, family income, marital status and self_rated_health correspondingly, $\beta_i (i=2, \dots, 7)$ are estimate coefficient, β_1 is intercept, ϵ is random error.

5 Results

As shown in Figure 2, the predictive of final model is 0.69 which is considered acceptable.

Table 1: Summary of Model

	test	train	SD	P value
(Intercept)	0.691	0.58	1.107	8.031e-08
sexMale	1.1	1.068	1.041	1.023e-01
educationLess than high school diploma or its equivalent	0.724	0.674	1.069	3.322e-09
educationBachelor or above	1.237	1.186	1.059	2.687e-03
educationcollege / below bachelor level	1.044	1.081	1.055	1.441e-01
pop_centerRural and non (CMA/CA) or Prince Edward Island	0.853	0.849	1.048	4.891e-04
marital_statusLiving common-law	0.729	0.87	1.099	1.401e-01
marital_statusMarried	0.507	0.614	1.081	4.126e-10
marital_statusSeparated	1.462	1.683	1.156	3.245e-04
marital_statusSingle, never married	1.177	1.432	1.086	1.294e-05
marital_statusWidowed	0.751	0.772	1.1	6.432e-03
self_rated_healthFair	6.145	6.399	1.087	3.862e-108
self_rated_healthGood	2.726	2.837	1.05	5.487e-100
self_rated_healthPoor	10.251	9.65	1.148	8.468e-60
income_family\$25,000 to \$49,999	1.211	1.283	1.067	1.159e-04
income_familyLess than \$25,000	1.396	1.384	1.082	3.990e-05
income_family\$50,000 to \$124,999	1.188	1.089	1.052	9.127e-02

Table 1.

6 Discussion

6.1 Weaknesses

6.2 Next Steps

7 References

8 Appendix

\begin{landscape}\begin{table}

Table 2: Gapminder Data

	Atlantic region (N=4564)	British Columbia (N=2522)	Ontario (N=5621)	Prairie region (N=4073)	Quebec (N=3822)	Total (N=20602)	p value
age							< 0.001
Mean (SD)	53.526 (17.244)	53.052 (17.868)	51.913 (17.822)	50.746 (18.059)	51.974 (17.681)	52.190 (17.747)	
Range	15.200 - 80.000	15.000 - 80.000	15.000 - 80.000	15.000 - 80.000	15.000 -	15.000 -	
						80.000 80.000	
sex							0.132
Female	2537 (55.6%)	1352 (53.6%)	3082 (54.8%)	2157 (53.0%)	2075 (54.3%)	11203 (54.4%)	

	Atlantic region (N=4564)	British Columbia (N=2522)	Ontario (N=5621)	Prairie region (N=4073)	Quebec (N=3822)	Total (N=20602)	p value
Male	2027 (44.4%)	1170 (46.4%)	2539 (45.2%)	1916 (47.0%)	1747 (45.7%)	9399 (45.6%)	
feelings_life							0.007
N-Miss	67	30	85	53	36	271	
Mean (SD)	8.161 (1.635)	8.023 (1.684)	8.064 (1.696)	8.101 (1.678)	8.101 (1.514)	8.094 (1.645)	
Range	0.000 - 10.000	0.000 - 10.000	0.000 - 10.000	0.000 - 10.000	0.000 - 10.000	0.000 - 10.000	
place_birth_canada							< 0.001
N-Miss	18	13	29	24	13	97	
Born in Canada	4237 (93.2%)	1757 (70.0%)	3806 (68.1%)	3300 (81.5%)	3255 (85.5%)	16355 (79.8%)	
Born outside Canada	277 (6.1%)	752 (30.0%)	1780 (31.8%)	738 (18.2%)	550 (14.4%)	4097 (20.0%)	
Don't know	32 (0.7%)	0 (0.0%)	6 (0.1%)	11 (0.3%)	4 (0.1%)	53 (0.3%)	
vis_minority							< 0.001
N-Miss	25	23	43	33	16	140	
Don't know	53 (1.2%)	32 (1.3%)	123 (2.2%)	67 (1.7%)	39 (1.0%)	314 (1.5%)	
Not a visible minority	4376 (96.4%)	1994 (79.8%)	4265 (76.5%)	3504 (86.7%)	3435 (90.3%)	17574 (85.9%)	
Visible minority	110 (2.4%)	473 (18.9%)	1190 (21.3%)	469 (11.6%)	332 (8.7%)	2574 (12.6%)	
citizenship_status							< 0.001
N-Miss	114	202	356	294	177	1143	
By birth	4252 (95.6%)	1779 (76.7%)	3826 (72.7%)	3318 (87.8%)	3264 (89.5%)	16439 (84.5%)	
By naturalization	166 (3.7%)	539 (23.2%)	1429 (27.1%)	453 (12.0%)	377 (10.3%)	2964 (15.2%)	
Don't know	32 (0.7%)	2 (0.1%)	10 (0.2%)	8 (0.2%)	4 (0.1%)	56 (0.3%)	
own_rent							< 0.001
N-Miss	27	18	33	29	13	120	
Don't know	27 (0.6%)	16 (0.6%)	15 (0.3%)	11 (0.3%)	2 (0.1%)	71 (0.3%)	
Owned by you or a member of this household, even if it i...	3467 (76.4%)	1817 (72.6%)	4168 (74.6%)	3162 (78.2%)	2480 (65.1%)	15094 (73.7%)	
Rented, even if no cash rent is paid	1043 (23.0%)	671 (26.8%)	1405 (25.1%)	871 (21.5%)	1327 (34.8%)	5317 (26.0%)	
income_family							< 0.001
\$100,000 to \$ 124,999	434 (9.5%)	267 (10.6%)	634 (11.3%)	464 (11.4%)	359 (9.4%)	2158 (10.5%)	
\$125,000 and more	898 (19.7%)	608 (24.1%)	1510 (26.9%)	1036 (25.4%)	655 (17.1%)	4707 (22.8%)	

	Atlantic region (N=4564)	British Columbia (N=2522)	Ontario (N=5621)	Prairie region (N=4073)	Quebec (N=3822)	Total (N=20602)	p value
\$25,000 to \$49,999	1050 (23.0%)	515 (20.4%)	1024 (18.2%)	821 (20.2%)	935 (24.5%)	4345 (21.1%)	
\$50,000 to \$74,999	844 (18.5%)	419 (16.6%)	993 (17.7%)	715 (17.6%)	725 (19.0%)	3696 (17.9%)	
\$75,000 to \$99,999	679 (14.9%)	366 (14.5%)	778 (13.8%)	559 (13.7%)	539 (14.1%)	2921 (14.2%)	
Less than \$25,000	659 (14.4%)	347 (13.8%)	682 (12.1%)	478 (11.7%)	609 (15.9%)	2775 (13.5%)	
education							< 0.001
N-Miss	71	42	90	80	58	341	
Bachelor's degree (e.g. B.A., B.Sc., LL.B.)	717 (16.0%)	520 (21.0%)	1179 (21.3%)	687 (17.2%)	650 (17.3%)	3753 (18.5%)	
College, CEGEP or other non-university certificate or di...	1068 (23.8%)	519 (20.9%)	1388 (25.1%)	836 (20.9%)	755 (20.1%)	4566 (22.5%)	
High school diploma or a high school equivalency certificate	1065 (23.7%)	620 (25.0%)	1302 (23.5%)	1082 (27.1%)	779 (20.7%)	4848 (23.9%)	
Less than high school diploma or its equivalent	804 (17.9%)	269 (10.8%)	660 (11.9%)	582 (14.6%)	721 (19.2%)	3036 (15.0%)	
Trade certificate or diploma	354 (7.9%)	173 (7.0%)	184 (3.3%)	356 (8.9%)	416 (11.1%)	1483 (7.3%)	
University certificate or diploma below the bachelor's level	135 (3.0%)	122 (4.9%)	158 (2.9%)	178 (4.5%)	139 (3.7%)	732 (3.6%)	
University certificate, diploma or degree above the bach...	350 (7.8%)	257 (10.4%)	660 (11.9%)	272 (6.8%)	304 (8.1%)	1843 (9.1%)	
pop_center							< 0.001
Larger urban population centres (CMA/CA)	2592 (56.8%)	2246 (89.1%)	5067 (90.1%)	2898 (71.2%)	3124 (81.7%)	15927 (77.3%)	
Prince Edward Island	708 (15.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	708 (3.4%)	
Rural areas and small population centres (non CMA/CA)	1264 (27.7%)	276 (10.9%)	554 (9.9%)	1175 (28.8%)	698 (18.3%)	3967 (19.3%)	
marital_status							< 0.001
N-Miss	1	0	1	4	1	7	
Divorced	350 (7.7%)	228 (9.0%)	478 (8.5%)	333 (8.2%)	378 (9.9%)	1767 (8.6%)	
Living common-law	443 (9.7%)	195 (7.7%)	381 (6.8%)	258 (6.3%)	798 (20.9%)	2075 (10.1%)	
Married	2185 (47.9%)	1243 (49.3%)	2708 (48.2%)	2125 (52.2%)	1240 (32.5%)	9501 (46.1%)	
Separated	183 (4.0%)	79 (3.1%)	206 (3.7%)	95 (2.3%)	80 (2.1%)	643 (3.1%)	
Single, never married	903 (19.8%)	562 (22.3%)	1328 (23.6%)	888 (21.8%)	1029 (26.9%)	4710 (22.9%)	

	Atlantic region (N=4564)	British Columbia (N=2522)	Ontario (N=5621)	Prairie region (N=4073)	Quebec (N=3822)	Total (N=20602)	p value
Widowed	499 (10.9%)	215 (8.5%)	519 (9.2%)	370 (9.1%)	296 (7.7%)	1899 (9.2%)	
living_arrangement							< 0.001
Alone	1247 (27.3%)	747 (29.6%)	1544 (27.5%)	1097 (26.9%)	1180 (30.9%)	5815 (28.2%)	
Living with one parent	109 (2.4%)	49 (1.9%)	160 (2.8%)	81 (2.0%)	112 (2.9%)	511 (2.5%)	
Living with two parents	167 (3.7%)	121 (4.8%)	337 (6.0%)	188 (4.6%)	159 (4.2%)	972 (4.7%)	
No spouse and non-single child(ren)	1 (0.0%)	2 (0.1%)	9 (0.2%)	5 (0.1%)	3 (0.1%)	20 (0.1%)	
No spouse and single child 25 years of age or older	41 (0.9%)	22 (0.9%)	78 (1.4%)	41 (1.0%)	31 (0.8%)	213 (1.0%)	
No spouse and single child under 25 years of age	190 (4.2%)	70 (2.8%)	238 (4.2%)	130 (3.2%)	186 (4.9%)	814 (4.0%)	
Other living arrangement	254 (5.6%)	153 (6.1%)	292 (5.2%)	269 (6.6%)	166 (4.3%)	1134 (5.5%)	
Spouse and non-single child(ren)	1 (0.0%)	1 (0.0%)	7 (0.1%)	2 (0.0%)	0 (0.0%)	11 (0.1%)	
Spouse and other	46 (1.0%)	21 (0.8%)	69 (1.2%)	44 (1.1%)	25 (0.7%)	205 (1.0%)	
Spouse and single child 25 years of age or older	91 (2.0%)	40 (1.6%)	137 (2.4%)	61 (1.5%)	44 (1.2%)	373 (1.8%)	
Spouse and single child under 25 years of age	799 (17.5%)	482 (19.1%)	1191 (21.2%)	890 (21.9%)	774 (20.3%)	4136 (20.1%)	
Spouse only	1618 (35.5%)	814 (32.3%)	1559 (27.7%)	1265 (31.1%)	1142 (29.9%)	6398 (31.1%)	
self_rated_health							< 0.001
N-Miss	19	16	33	19	12	99	
Don't know	18 (0.4%)	6 (0.2%)	18 (0.3%)	10 (0.2%)	5 (0.1%)	57 (0.3%)	
Excellent	833 (18.3%)	566 (22.6%)	1206 (21.6%)	839 (20.7%)	932 (24.5%)	4376 (21.3%)	
Fair	542 (11.9%)	248 (9.9%)	517 (9.3%)	426 (10.5%)	345 (9.1%)	2078 (10.1%)	
Good	1303 (28.7%)	771 (30.8%)	1676 (30.0%)	1250 (30.8%)	1162 (30.5%)	6162 (30.1%)	
Poor	216 (4.8%)	116 (4.6%)	234 (4.2%)	155 (3.8%)	95 (2.5%)	816 (4.0%)	
Very good	1633 (35.9%)	799 (31.9%)	1937 (34.7%)	1374 (33.9%)	1271 (33.4%)	7014 (34.2%)	
self_rated_mental_health							< 0.001
N-Miss	19	19	34	22	12	106	
Don't know	16 (0.4%)	4 (0.2%)	16 (0.3%)	9 (0.2%)	12 (0.3%)	57 (0.3%)	
Excellent	1156 (25.4%)	714 (28.5%)	1726 (30.9%)	1134 (28.0%)	1350 (35.4%)	6080 (29.7%)	

	Atlantic region (N=4564)	British Columbia (N=2522)	Ontario (N=5621)	Prairie region (N=4073)	Quebec (N=3822)	Total (N=20602)	p value
Fair	340 (7.5%)	174 (7.0%)	361 (6.5%)	261 (6.4%)	160 (4.2%)	1296 (6.3%)	
Good	1385 (30.5%)	730 (29.2%)	1504 (26.9%)	1196 (29.5%)	998 (26.2%)	5813 (28.4%)	
Poor	79 (1.7%)	54 (2.2%)	99 (1.8%)	70 (1.7%)	24 (0.6%)	326 (1.6%)	
Very good	1569 (34.5%)	827 (33.0%)	1881 (33.7%)	1381 (34.1%)	1266 (33.2%)	6924 (33.8%)	
religion_has_affiliation							< 0.001
N-Miss	48	52	84	71	27	282	
Don't know	39 (0.9%)	35 (1.4%)	32 (0.6%)	42 (1.0%)	11 (0.3%)	159 (0.8%)	
Has religious affiliation	3840 (85.0%)	1494 (60.5%)	4386 (79.2%)	3049 (76.2%)	3386 (89.2%)	16155 (79.5%)	
No religious affiliation	637 (14.1%)	941 (38.1%)	1119 (20.2%)	911 (22.8%)	398 (10.5%)	4006 (19.7%)	
children_in_household							< 0.001
No child	3678 (80.6%)	2015 (79.9%)	4323 (76.9%)	3039 (74.6%)	2942 (77.0%)	15997 (77.6%)	
One child	433 (9.5%)	236 (9.4%)	577 (10.3%)	421 (10.3%)	378 (9.9%)	2045 (9.9%)	
Three or more children	92 (2.0%)	70 (2.8%)	184 (3.3%)	212 (5.2%)	131 (3.4%)	689 (3.3%)	
Two children	361 (7.9%)	201 (8.0%)	537 (9.6%)	401 (9.8%)	371 (9.7%)	1871 (9.1%)	
number_marriages							< 0.001
Mean (SD)	0.845 (0.607)	0.860 (0.651)	0.820 (0.611)	0.842 (0.606)	0.626 (0.610)	0.799 (0.620)	
Range	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	0.000 - 4.000	

\end{landscape}page}

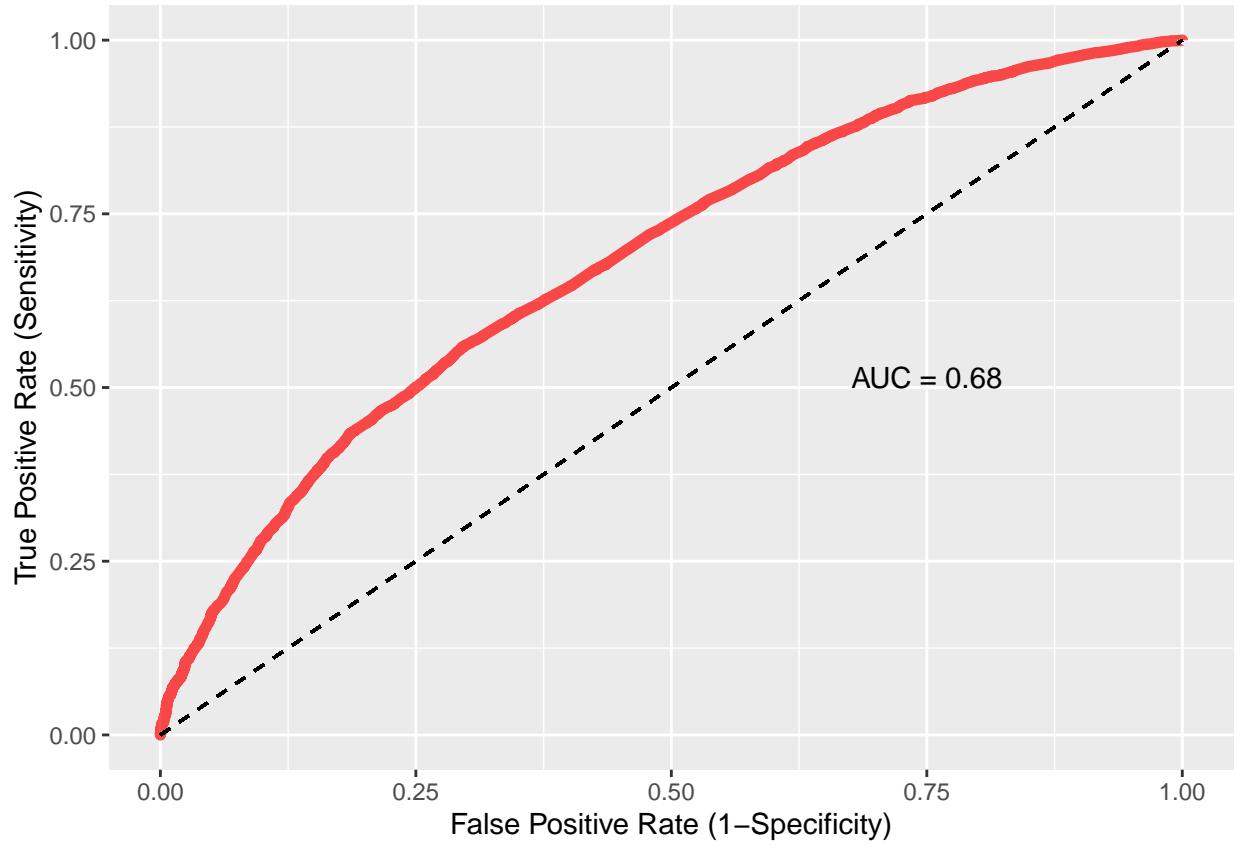


Figure 2: ROC curve