## **City of Austin Community Perception and Satisfaction Analysis**

Author: Eric Hart, MSBA student at Oklahoma State University Mentors: Dr. Goutam Chakraborty, Oklahoma State University Dr. Miriam McGaugh, Oklahoma State University

## ABSTRACT

Austin, Texas is often quoted as one of the best cities to live in within the United States. It is important that the city focuses on key issues to continue to be progressive. The City of Austin performs an annual community survey to assess the residents' satisfaction level with the city. The objective of this paper is to identify which areas are causing the most dissatisfaction for the City of Austin as a place to live for residents by using SAS® Enterprise Miner. The dataset is comprised of over 6,300 responses and 55 variables from 2015 to 2017. The residents of Austin that were selected to take this survey were chosen randomly. The survey covers a broad spectrum of categories such as perception of the community, personal safety, infrastructure, and environmental services. By determining the services that are causing the most dissatisfaction for the city's residents, the City of Austin can be proactive in taking measures to rectify and improve the identified areas of discontent. The city can use these results to focus budgets, improve services for residents, and improve the lives of those who live within the Austin city limits.

## Introduction

Austin, the capital of Texas, is the fourth largest city in the state of Texas. According to the world population review, the metropolitan area contains over 2 million people. The number is rapidly growing. Other city sectors like construction, commercial businesses, and residential areas, have simultaneously increased growth to support the rapidly increasing population. Each year, the City of Austin conducts a survey to assess the overall level of satisfaction or dissatisfaction of its residents. The survey measures satisfaction levels over a wide array of topics. These topics range from public services offered by the city like public transportation and the library system, to the individual's personal feelings towards the city such as how he or she views Austin as a place to raise children, work, retire, and his or her overall outlook as feeling a part of the community. By analyzing the results of the survey, the City of Austin can discover insights into the population's opinion on which areas of Austin need improvement.

## Data Background

This dataset is freely available on the city of Austin website. The survey is randomly distributed to residents of Austin. The survey helps the city of Austin to improve their services and to meet needs related to health, safety, mobility, economic opportunity, culture, learning, and government efficiency. The data is sourced from the 2015-2017 and has over 6,300 observations and 55 variables.

### **Data Preparation**

I downloaded the dataset from the city of Austin's website into an excel file. The responses were available in text format ranging from "Very Dissatisfied" to "Very Satisfied" and were ranked from 1 - 5. I dummy coded male and female to 1 and 0. I transformed the text results into responses from 1 - 5 and used the number 9 for a response of "Don't Know". There were 156 variables initially downloaded from the website, but I was able to reduce the variables to 55. I removed any variables that contained a large amount of missing results or variables that were similar in nature to other remaining questions. I used the vlookup function with the key shown in Figure 1.1 to transform the text from results to numeric responses for data analysis. My dataset contained 55 variables and 6,374 observations when it was ready to be imported into SAS<sup>®</sup> Enterprise Miner. I used the file import node to import the excel file into SAS<sup>®</sup> Enterprise Miner.

	Transformations	_		
		-	Age	
Strongly Agree	Very Satisfied	5	18-34 years	1
Agree	Satisfied	4	35-44 years	2
Neutral	Neutral	3	45-54 years	3
Disagree	Dissatisfied	2	55-64 years	4
Strongly Disagree	Very Dissatisfied	1	65+ years	5
Don't Know	Don't Know	9	Not provided	6
Race			Income	
Hispanic	1		\$40K-59,999	1
African American Only	2		\$60K-79,999	2
Caucasian/White Only	3		\$80K-149,999	3
Other	4		less than \$20K	4
Asian/Pacific Islander Only	5		\$20K-39,999	5
American Indian Only	6		\$150K or more	6
			Not provided	7
Gender				
male	1			
Female	0			

Figure 1.1

#### **Data Exploration**

Descriptive statistics were initially conducted to form a basis of understanding of the data.

Interval Variable Summary Statistics

							Standard		
Variable	Label	Missing	N	Hinimum	Maximum	Mean	Deviation	Skewness	Kurtosis
Animal services	Animal services	0	6374	1	9	4,64	2,200	1,1092	0.11
Availability of affordable housi	Availability of affordable housing for low/moderate income families	0	6374	1	9	3.81	2,985	0.9326	-0.72
Do you own or rent your home	Do vou own or rent vour home?		6374	Û	1	0.72	0.448	-0.9913	-1.02
Employees of the City ofAustin a	Employees of the City ofAustin are ethical in the way they conduct City business.	0	6374	1	9	4.57	2.332	0,9652	-0.13
Energy Conservation program	Energy Conservation program	0	6374	1	9	4.32	2.178	1.2233	0.59
Enforcement of local traffic law	Enforcement of local traffic laws	0	6374	1	9	3.89	1,935	1.4471	2.06
Flood control efforts	Flood control efforts	0	6374	1	9	4.15	2,261	1.2407	0.56
Food Safety Inspection program	Food Safety Inspection program	0	6374	1	9	5,90	2,869	0.0405	-1.77
How many dependents including v	How many dependents (including yourself) did your household claim on its most recentfederal taxes?	36	6338	0	5	1.97	1.247	0.5837	-0.09
I feel safe in my neighborhood a	I feel safe in my neighborhood at night	0	6374	1	9	3.89	1,165	0.0944	2,96
I feel safe in my neighborhood d	I feel safe in my neighborhood during the day	0	6374	1	9	4.37	0.905	0.2343	7.09
Library hours		0	6374	1	9	5,06	2,471	0.7188	-0.94
Medical assistance provided by E	Medical assistance provided by EMS	0	6374	1	9	5.57	2.282	0.6538	-1.15
Weighborhood planning zoning eff	Neighborhood nlanning/zoning efforts	0	6374	1	9	4, 39	2.870	0.7520	-0.98
Number of City of Austin parks	Number of City of Austin narks	0	6374	ī	9	4.22	1.586	1.5217	3, 38
Overall effectiveness of communi	Overall effectiveness of communication	ů.	6374	1	9	3.59	1.801	1,6910	3, 34
Overall management of stormwater	Overall management of stormwater rumoff	ñ	6374	1	9	3,83	2.010	1,4815	1.91
Overall quality of Austin Energy	Overall multiputto of Austin Energy	ň	6374	1	9	3.94	1.653	1.2725	3.03
Overall quality of city librarie	Overall quality of city libraries	ň	6374	1	9	4.81	2,123	1.0918	0.00
Overall quality of city narks an	Overall quality of city marks and recreation	ů	6374	1	9	4 14	1 416	1 5512	4 72
Overall quality of customer serv	Overall quality of customer service	ů	6374	1	9	4 33	1 991	1 3690	1 26
Overall quality of drinking wate	Overall quality of drinking water	ů	6374	1	9	4.05	1 225	0.9734	5.02
Overall quality of fire services	Overall quality of fire services	ů	6374	i	9	5.36	2 162	0.8874	-0.75
Overall guality of health and hu	Overall quality of health and human services	ů	6374	î		4 79	2 592	0.7773	-0.89
Overall quality of municipal cou	Overall quality of municipal court services	ñ	6374	1	9	4.86	2.645	0.6979	-1.03
Overall quality of police servic	Overall quality of nolice services	ň	6374	1	9	4.20	1.677	1,4873	2.95
Overall quality of public safety	Overall quality of public sefety services	ň	6374	1	9	4.16	1.491	1.5349	4.16
Overall quality of services	Overall quality of services	ň	6374	1	9	3 52	1,333	1.3749	5 42
Overall quality of the Airport	Overall quality of the Wirnort	ů	6374	1	9	4 27	1 299	1 6806	5 61
Overall quality of wastewater se	Overall muality of wastewater services	0	6374	1	9	4.05	1.526	1.5057	3, 97
Overall satisfaction with City o	Overall satisfaction with City of Austin swimming nools	0	6374	1	9	5,10	2,662	0.5589	-1.22
Overall value tax dollars and f	Overall value tax dollars and fees	0	6374	ī	9	3.21	1.624	1.5722	4.12
Quality of residential garbage c	Quality of residential marhage collection	ů.	6374	1	9	4.34	1,386	1.4565	4.67
Quality of youth athletic progra	Quality of youth athletic programs offered	ñ	6374	1	9	6.22	2.849	-0.1502	-1.77
Race Ethnicity Recode	Race/Ethnicity Recode	ŏ	6374	1	6	2,35	1,162	0.2824	-0.54
Services provided by the City s	Services provided by the City's 3-1-1 assistance telephone number	0	6374	1	9	5.06	2,276	0.8269	-0.56
Sneed of nolice response	Sneed of nolice resnonse	0	6374	1	9	5.08	2,455	0.6862	-0.90
The water quality of lakes and s	The water quality of lakes and streams	0	6374	1	9	4,19	2,039	1.3989	1.29
Timeliness of EMS response to em	Timeliness of EMS response to emergency location	0	6374	1	9	5.66	2.322	0.5720	-1.29
Timeliness of fire department re	Timeliness of fire department response to emergency location	0	6374	1	9	5.70	2.323	0.5526	-1.34
Traffic flow on major city stree	Traffic flow on major city streets	0	6374	1	9	2.27	1.450	2,2968	7.90
Traffic flow on major highways	Traffic flow on major highways	0	6374	1	9	4.12	3.575	0.5685	-1.57
What is your gender	What is your gender?	0	6374	0	1	0.48	0.500	0.0672	-2.00
Which of the following best desl	Which of the following best describes your annual household income?	0	6374	1	7	3.67	1.902	0.2082	-1.16
Which of the following best desc	Which of the following best describes your AGE?	0	6374	1	6	2.98	1.408	0.0609	-1.21
Year	Year	0	6374	2015	2017	2016.02	0.819	-0.0448	-2.82
Zip Code	Zip Code	0	6374	78613	99999	78739.39	267.079	79.1650	6302.57
place to raise children	place to raise children	0	6374	1	9	4.74	2.129	1.0312	0.17
place to retire	place to retire	0	6374	1	9	4.04	2.361	1.0243	0.23
planning growth	planning growth	0	6374	1	9	2.57	1.781	2.0278	4.86
place_to_live	place to live	0	6374	1	9	4.07	1.112	0.3082	4.55
place_to_work	place to work	0	6374	1	9	4.17	1.344	1.3473	4.80
-									

#### Figure 1.2

I used the data mining database node shown in Figure 1.2 to see the number of missing results, the mean, the skewness, and the kurtosis. As you can see in Figure 1.2, the results for "Traffic flow on major city streets" and how well the city is "planning growth" scored very poor. Austin needs to focus on planning growth, which in turn, will positively affect traffic flow on major streets. Shown below in Figure 1.3 is a histogram of the results for the target variable, Austin, as a place to live.





The graph shown in Figure 1.4 illustrates the correlation of the independent variables to the dependent variable, which is a place to live.



Figure 1.4



Shown in Figure 1.5 is the variable worth of each independent variable to the dependent variable.



Both "traffic flow on major city streets" and "planning growth" as depicted in Figure 1.6 and Figure 1.7 have low results and are areas in which the City of Austin needs to focus its budget and attention.



Figure 1.6



In the output from the variable selection node shown in Figure 1.8, you can observe which variables were not selected by the R-square value due to insufficiency.

R-Squares for Target Variable: place\_to\_live

Effect		DF	R-Square
A0V16:	_place_to_raise_children	5	0.315725
A0V16:	place_to_work	5	0.273464
A0V16: A0V16:	_place_to_retire Overall value tax dollars and f	5 5	0.214152
Var:	place_to_work	1	0.198181
AOV16:	Overall_quality_of_services	5	0.165153
A0V16:	planning_growth	5	0.152089
Var:	Overall_quality_of_services	l	0.142400
Var:	_place_to_retire	1	0.125244
Var:	_planning_growth	1	0.107211
Var:	_place_to_raise_children	1	0.105862
A0V16:	I_feel_safe_in_my_neighborhood_d	5	0.090093
A0V16:	Overall_quality_of_city_parks_an	5	0.087054
Var:	I_feel_safe_in_my_neighborhood_d	1	0.086108
A0V16:	Overall_quality_of_drinking_wate	5	0.081840
A0V16:	Overall quality of public safety	5	0.077528
A0V16:	Traffic_flow_on_major_city_stree	5	0.072696
Var:	Overall_quality_of_drinking_wate	1	0.071941
A0V16:	Overall_quality_of_the_Airport	5	0.070485
A0V16:	Overall_quality_of_municipal_cou	5	0.063966
A0V16:	I_feel_safe_in_my_neighborhood_a	5	0.063556
Var:	Overall_quality_of_Austin_Energy	1	0.062533
Var:	Traffic_flow_on_major_city_stree	1	0.059148
Var:	I_feel_safe_in_my_neighborhood_a	1	0.057781
A0V16: A0V16:	Energy Conservation program	5 5	0.055096
Var:	Overall_effectiveness_of_communi	1	0.053013
A0V16:	Overall_management_of_stormwater	5	0.052131
AUVI6: Var:	Overall quality of public safety	5	0.049778
Var:	Overall_quality_of_city_parks_an	ĩ	0.046975
A0V16:	${\tt Quality\_of\_residential\_garbage\_c}$	5	0.046615
A0V16:	Animal_services Enforcement of local traffic law	5	0.045769
Var:	Overall_quality_of_wastewater_se	1	0.044825
A0V16:	Overall_quality_of_city_librarie	5	0.043794
A0V16:	Traffic_flow_on_major_highways	5	0.041551
A0V16:	The water quality of lakes and s	5	0.038885
A0V16:	Flood_control_efforts	5	0.036750
A0V16:	Services_provided_by_the_City_s	5	0.036345
A0V16:	Overall satisfaction with City o	5	0.033335
A0V16:	Employees_of_the_City_ofAustin_a	5	0.032690
A0V16:	Food_Safety_Inspection_program	5	0.030375
Var:	Overall quality of the Airport	1	0.027008
Var:	Overall_quality_of_police_servic	1	0.026936
Var:	Overall_management_of_stormwater	1	0.021626
AOV16:	Quality of youth athletic progra	5	0.021203
Var:	Quality_of_residential_garbage_c	1	0.020539
A0V16:	Library_hours	5	0.019203
Var: Var:	Overall guality of health and hu	1	0.018825
A0V16:	Medical_assistance_provided_by_E	5	0.018257
Var:	Availability_of_affordable_housi	1	0.017168
AUV16: AUV16:	Timeliness_of_fire_department_re Overall quality of fire services	5	0.017150
A0V16:	Timeliness_of_EMS_response_to_em	5	0.016669
Var:	Overall_quality_of_municipal_cou	1	0.016523
Var: Var:	Animal_services Neighborhood planning zoning eff	1	0.014632
Var:	Services_provided_by_the_City_s	1	0.012831
Var:	Overall_quality_of_city_librarie	l	0.012690
AOV16: Ver:	Which_of_the_following_best_desl	6	0.012539
Var:	Speed_of_police_response	1	0.009634
Var:	Flood_control_efforts	1	0.009182
Var:	Traffic_flow_on_major_highways	1	0.008563
Var:	Number of City of Austin parks	1	0.005476
Var:	Food_Safety_Inspection_program	1	0.005105
Var:	Quality_of_youth_athletic_progra	1	0.005018
Var:	Timeliness of fire department re	1	0.003994
Var:	The_water_quality_of_lakes_and_s	1	0.003980
Var:	Timeliness_of_EMS_response_to_em	1	0.003758
Var:	Overall_quality_of fire services	1	0.003137
Var:	Which_of_the_following_best_desc	1	0.002721
Var:	Overall_satisfaction_with_City_o	1	0.002688
A0V16:	How_many_dependents including v	5 5	0.002284
Var:	Do_you_own_or_rent_your_home_	1	0.002010
A0V16:	Do_you_own_or_rent_your_home_	1	0.002010
AOV16:	Year	1	0.001957
Var:	Year	1	0.001798
Var:	Race_Ethnicity_Recode	1	0.000696
Var: Var:	which_of_the_following_best_desl How many dependents including y	1	0.000625
Var:	What_is_your_gender_	1	0.000158
A0V16:	What_is_your_gender_	1	0.000158

Figure 1.8

# **Model Building**

To predict whether living in Austin will be ranked as "very dissatisfied" to "very satisfying" as a place to live, different models were built using SAS<sup>®</sup> Enterprise Miner.





I used the metadata node to reject the zip code variable as it had too many levels of observations. I used the replacement node to replace any missing values in the dataset. I partitioned the dataset to contribute 50% to training and 50% to validation. The variables shown in Figure 2.1 are those that flowed through the data partitioned node and were used in the models.

Name	Role 🛆	Level
dataobs	ID	Interval
REP Overall quality of drinking	Input	Interval
REP Overall quality of customer	Input	Interval
REP Overall quality of fire serv	Input	Interval
REP Overall quality of city libr	Input	Interval
REP Overall quality of Austin En	Input	Interval
REP Overall quality of city park	Input	Interval
REP Overall quality of public sa	Input	Interval
REP Overall quality of police se	Input	Interval
REP Overall quality of services	Input	Interval
REP Overall quality of health an	Input	Interval
REP Overall quality of municipal	Input	Interval
REP I feel safe in my neighborho	Input	Interval
REP How many dependents includi	Input	Interval
REP Library hours	Input	Interval
REP Flood control efforts	Input	Interval
REP Food Safety Inspection progr	Input	Interval
REP Overall effectiveness of com	Input	Interval
REP Number of City of Austin par	Input	Interval
REP Overall management of stormw	Input	Interval
REP Medical assistance provided	Input	Interval
REP Neighborhood planning zoning	Input	Interval
REP What is your gender	Input	Interval
REP Traffic flow on major highwa	Input	Interval
REP Which of the following best	Input	Interval
REP Timeliness of fire departmen	Input	Interval
REP Timeliness of EMS response t	Input	Interval
REP Traffic flow on major city s	Input	Interval
REP planning growth	Input	Interval
REP place to retire	Input	Interval
REP place to work	Input	Interval
REP Year	Input	Interval
REP place to raise children	Input	Interval
REP Overall value tax dollars a	Input	Interval
REP Overall satisfaction with Ci	Input	Interval
REP Quality of residential garba	Input	Interval
REP Overall quality of the Airpo	Input	Interval
REP Overall quality of wastewate	Input	Interval
REP Speed of police response	Input	Interval
REP Services provided by the Cit	Input	Interval
REP The water quality of lakes a	Input	Interval
REP Quality of youth athletic pr	Input	Interval
REP Race Ethnicity Recode	Input	Interval
REP Do you own or rent your home	Input	Interval
REP Dup2	Input	Interval
REP Animal services	Input	Interval
REP Availability of affordable h	Input	Interval
REP Energy Conservation program	Input	Interval
REP Enforcement of local traffic	Input	Interval
REP Dup4	Input	Interval
REP Employees of the City of Aust	Input	Interval

## Figure 2.1

The models built were a decision tree, gradient boosting, neural network, and a logistic regression. As you can see in Figure 2.2, the decision tree shows the two most important variables being a "place to raise children" followed by a "place to work".

Variable Importance

		Number of		
		Splitting		Validation
Variable Name	Label	Rules	Importance	Importance
REPplace_to_raise_children	Replacement: place to raise children	10	1.0000	1.0000
REP_place_to_work	Replacement: place to work	4	0.6336	0.7102
REPplace_to_retire	Replacement: place to retire	4	0.6268	0.6306
REP_Overall_quality_of_services	Replacement: Overall quality of services	1	0.5219	0.3360
REP_Overall_valuetax_dollars_a	Replacement: Overall value tax dollars and fees	5	0.2217	0.1971
REP_Dup2	Replacement: I feel safe in my neighborhood during the day	1	0.1823	0.0513
REP_Overall_quality_of_police_se	Replacement: Overall quality of police services	1	0.0789	0.0497
REP_Overall_effectiveness_of_com	Replacement: Overall effectiveness of communication	1	0.0655	0.0658

The gradient boosting model came to the same conclusion as the decision tree, showing that the most important variables are a "place to raise children" followed by a "place to work", to predict the score of a "place to live" for the city of Austin. The logistic regression chose the variables in Figure 2.3 in the stepwise selection.

#### Summary of Stepwise Selection

	Effect		Number		
Step	Entered	DF	In	F Value	Pr > F
1	REP_place_to_work	1	1	873.93	<.0001
2	REP_Overall_valuetax_dollars_a	1	2	414.20	<.0001
3	REPplace_to_retire	1	3	157.31	<.0001
4	REPplanning_growth	1	4	97.97	<.0001
5	REP_Dup2	1	5	79.88	<.0001
6	REP_Overall_quality_of_drinking_	1	6	45.25	<.0001
7	REPplace_to_raise_children	1	7	43.59	<.0001
8	REP_The_water_quality_of_lakes_a	1	8	24.18	<.0001
9	REP_Overall_quality_of_services	1	9	24.81	<.0001
10	REP_Do_you_own_or_rent_your_home	1	10	22.07	<.0001
11	REP_Traffic_flow_on_major_city_s	1	11	18.50	<.0001
12	REP_Speed_of_police_response	1	12	15.78	<.0001
13	REP_Overall_quality_of_Austin_En	1	13	5.99	0.0145
14	REP_Overall_satisfaction_with_Ci	1	14	7.00	0.0082
15	REP_Availability_of_affordable_h	1	15	8.37	0.0038
16	REP_Food_Safety_Inspection_progr	1	16	5.62	0.0178
17	REP_Dup4	1	17	5.08	0.0242

#### Figure 2.3

All the models were conjoined to create an ensemble model, and each model, as well as the ensemble model, were flowed into a model comparison node to determine the model that best accurately predicted the ranking of a "place to live", shown in Figure 2.4.

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error
Y	Ensmbl	Ensmbl	Ensemble	place to live	place to live	0.624971
	Tree	Tree	Decision Tree	place to live	place to live	0.641678
	Boost	Boost	Gradient Boosting	place to live	place to live	0.706362
	Neural	Neural	Neural Network	place to live	place to live	0.706678
	Req	Req	Regression	place to live	place to live	0.746088

#### Figure 2.4

The ensemble model, using the fit statistic of average square value, was chosen because it produced the lowest average squared error.

## Conclusion

The City of Austin is growing at a very fast pace and needs to be managed in a proper and organized manner. The overall ratings of the City of Austin show that it is doing very well as a place to live. The results show that the city should focus on improving traffic flow on city streets as well as improve strategy for city growth and planning. These two variables are correlated as follows: if traffic flow improves, there will be a perception that the City is planning growth more effectively. A good predictive model can help city council focus on what is important to the citizens of the city, and what is already sufficient in the cities organizational principles for managing the city. To continue to reach these extremely valuable opinions from residents of Austin, this survey should continue to be distributed annually. By being consistent in the random distribution of the survey to Austin residents, all voices in the city can be heard, and growth of certain variables can be tracked closely. All of these factors will lead to sufficient and steady progress for the city of Austin.

# Further Scope

Run models on certain demographic sections of the observations that can be used to determine what the key variables are in different segments of the dataset. This would allow the city council to be able to understand the problems that exist for subsections of the city.

## References

- Data.austintexas.gov. Community Survey, 2015 2017 | City of Austin | Data Portal.: <u>https://data.austintexas.gov/City-Government/Community-Survey-2015-2016-2017/76qk-igxn</u>
- <u>https://www.forbes.com/pictures/edgl45fkm/no-1-austin-texas/#61262a241a5c</u> by Forbes
  <u>http://worldpopulationreview.com/us-cities/austin-population/</u>

# Acknowledgment

I would like to appreciate Dr. Miriam McGaugh and Dr. Goutam Chakraborty at Oklahoma State University for their support and direction with this research.

# **Contact Information**

Eric Hart, Oklahoma State University Email: <u>Eric.hart@okstate.edu</u> Phone no.: 832-273-2932