Alex Funches
Stat 440
Professor Unger
3/27/2020

Stat 440 Midterm Project Report

Introduction

The objective of this midterm project is to accurately investigate and analyze the connection of socio-economic factors in the city of Boston in 1978 through statistical analysis. The Boston Damaged Dataset consists of 506 observations recorded in 1978 and 21 variables that detail the housing units, residential proportions, and the town's societal conditions. The only character variable is town which is used as a factor to reference the town names with the remaining variables describing their environmental surrounding. At 20 an overwhelming majority of these variables are numeric and characterize the property values, housing unit positions and the town's accessibility to nearby employment centers and highways. More specifically the dataset contains four variables for housing the number of units built prior to 1904 (age), the average number of rooms (rooms), the median values of owner-occupied housing in USD (MedV), and the corrected median values (CMedV).

As a Non-Sas and delimited raw data file the Boston dataset was read in using an infile statement that identified the physical name and location of the datafile with a delimiter to account for the tabs separating the data. Transferring in the information a list input specification was used in addition to a length statement as not all the town names are a single word and include spaces with the longest town name being a length of 24.

Methods

Examining the data, the guidelines used were to see if any observation had a significant impact on the variables and skewed the results, this included missing values, outliers, and illogical values. To detect and remove the potential observations that caused a biased analysis proc mean statement were run for numeric variables like longitude, median values of owner-occupied housing (MedV), and the per capita crime rate (Crim). This would allow me to verify if each variable had a total of 506 observations and conclude if they were any missing observations if this was not the case. At 505 recorded observations the industry variable had one missing value at observation number 136 and at 504 the tax variable had two missing values at observation numbers 213 and 315. Because these missing values were represented by blank spaces the dataset a dsd statement was added to the infile statement so that missing values were instead represented as periods. This adjustment from blanks to periods enabled me to use if statements in the data step. As constants for each town I was able to use this if statement to change the values of Industry and Tax to their corresponding values.

Transitioning to the investigation of outliers I compared the numeric results of the proc means statement for each variable. I could discern if there was an abnormally small minimum or large maximum if it was exceptionally different from the variable's mean and resulted in a very large standard deviation. The Nox variable corresponding to the nitric oxide concentration per town had three recorded observations of 9,999 in observation numbers 82, 116, and 214 while all other 503 observations where between the values of zero and one. Correcting this if statements

were also added for Nox for each observation that had a recorded value of 9,999 that would change it to the values corresponding to the town. While looking through the data I also noticed that not all the longitude and latitude values were in decimal degrees as specified in the midterm assignment. I formatted both variables to the fifth decimal place in the datastep by specifying a comma format that used a width of 9 and 6 digits to the right of the decimal point.

Results

Running a proc means statement with all the numeric variables enabled me to analyze if each variable was recorded for all 506 observations and that their corresponding maximum and means were not considerably different from one another. Checking to see if the standard deviations were reasonable, I could inspect the distribution of the variables with alarming values and ranges. Below is the table with the recorded observations and descriptive values for each variable.

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
ObservationNumber		506	253.5000000	146.2138844	1.0000000	506.0000000
TownNumber		506	47.5474308	27.5934905	0	91.0000000
TractID		506	2700.36	1380.04	1.0000000	5082.00
Longitude		506	-70.2141231	10.9191849	-71.2895000	71.0377000
Latitude		506	42.6116972	6.2753296	42.0300000	142.1374000
MedV	Vector of Median Values of Owner-Occupied	506	22.5328063	9.1971041	5.0000000	50.0000000
CMedV	Housing in USD 1000	506	22.5288538	9.1821759	5.0000000	50.0000000
Crim	Vector of Corrected Median Values of Owner-	506	3.6135236	8.6015451	0.0063200	88.9762000
ZN	Occupied Housing in USD 1000	506	11.3636364	23.3224530	0	100.0000000
Industry	Vector of Per Capita Crime	505	11.1154851	6.8503972	0.4600000	27.7400000
Chas	Vector of Proportions of Residential Land Zoned for	506	0.0691700	0.2539940	0	1.0000000
Nox	Lots Over 25,000 Sq Ft per Town	506	59.8344144	768.3447623	0.3850000	9999.00
Rooms	Vector of Proportions of Non-Retail Business Acres	506	6.2846344	0.7026171	3.5610000	8.7800000
Age	per Town	506	68.5749012	28.1488614	2.9000000	100.0000000
Distance	Factor Indicating if Tract Borders Charles River	506	3.7950427	2.1057101	1.1296000	12.1265000
Rad	Vector of Nitric Oxide Concentration (Parts per 10	506	9.5494071	8.7072594	1.0000000	24.0000000
Tax	Million) per town	504	408.7043651	168.7057796	187.0000000	711.0000000
PTRatio	Average Number of Rooms per Dwelling	506	18.4713439	2.2065225	12.6000000	28.2000000
BProportion	Vector of Proportions of Owner-Occupied Units	506	356.6740316	91.2948644	0.3200000	396.9000000
LStat	Built Before 1940	506	12.6530632	7.1410615	1.7300000	37.9700000
	Vector of Weighted Distances to Five Boston Employment Centers					
	Vector of Index of Accessibility to Radial Highways per Town					
	Vector of Full Value Property-Tax Rate per USD 10,000 per Town					
	Vector of Pupil-Teacher Ratios per Town					
	Proportion of Blacks per Town					
	Vector of Percentage Values of Lower Status Population					

Inspecting these results, I initially found four of the variables to be concerning: Industry, Tax, Nox, and BProportion. Industry is missing one recorded observation at observation number 136 and Tax is missing two recorded observation at observation numbers 213 and 315 and has a large standard deviation at 168.70. The variable Nox is representative of the nitric oxide

concentration in parts per 10 million per town and with a mean of 59.83 the maximum of 9,999 is excessive and causes a very large standard deviation of 768.34. BProportion a numeric vector on the proportion of blacks has a moderate deviation at 91.29. The remaining variables are all recorded for each observation and have reasonable descriptive values.

First looking at the BProportion variable I created a frequency table to see how often the maximum of 396.9 appeared. 121 of 506 of the observations have the maximum value of 396.9 for BProportion and with some observations having values as low as 0.32 as seen by the minimum this could help explain the moderate deviation of 91.29.

Proportion of Blacks per Town				
BProportion	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
396.9	121	100.00	121	100.00

Correcting the missing values, I utilized a dsd statement in the datastep's infile statement which would instead assign missing values to a period. As Industry, Tax, and Nox are constant within the same town I first scrutinized the dataset to figure out the constants for the values represented by periods in Industry and Tax and the values of 9,999 for Nox. Afterwards in the datastep's formatting statement I included if and clauses that would match the values of period for Tax and Industry and the values of 9,999 for Nox to their corresponding Towns and then assign the proper values. For example, in observation 136 when Industry has a value of period and the town is Somerville Industry is then reassigned the value of 21.89 which is the constant for all other observations of Industry in Somerville. Here is a table produce by the proc means procedure for the variables of Industry, Tax, and Nox before and after the changes.

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
Nox	Vector of Nitric Oxide Concentration (Parts per 10 Million) per town	506	59.8344144	768.3447623	0.3850000	9999.00
Industry	Vector of Proportions of Non-Retail Business Acres per Town	505	11.1154851	6.8503972	0.4600000	27.7400000
Tax	Vector of Full Value Property-Tax Rate per USD 10,000 per Town	504	408.7043651	168.7057796	187.0000000	711.0000000

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
Nox	Vector of Nitric Oxide Concentration (Parts per 10 Million) per town	506	0.5546951	0.1158777	0.3850000	0.8710000
Industry	Vector of Proportions of Non-Retail Business Acres per Town	506	11.1367787	6.8603529	0.4600000	27.7400000
Tax	Vector of Full Value Property-Tax Rate per USD 10,000 per Town	506	408.2371542	168.5371161	187.0000000	711.0000000

The mean of Nox drastically dropped from 768.34 to 0.55, the max of 9,999 is now 0.871 and the standard deviation is now more appropriate at 0.115 instead of 768.34. For both Industry and Tax there are now 506 recorded observations, however it had little impact on the descriptive values. The minimum and maximum remained the same in both Industry and Tax.

The mean of Industry slightly rose from 11.115 to 11.136 and the standard deviation increased by .009. The mean of Tax slightly fell from 408.70 to 408.23 and standard deviation decreased by 0.168.

Exercise A

To examine the town that the most recorded number of observations a means procedure was used with the variable of interest being TractID and Town being the class variable. As TractID is a numeric vector of the tract ID numbers it serves as a means of identification and the calculations like the mean and standard deviation are insignificant. The town of Cambridge is represented by the most tracts at 30 observations and attached is the table.

Analysis Variable : TractID						
TOWN	**N Obs**	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
Arlington	7	7	3564.00	2.1602469	3561.00	3567.00
Ashland	2	2	3851.50	0.7071068	3851.00	3852.00
Bedford	2	2	3591.50	0.7071068	3591.00	3592.00
Belmont	8	8	3574.50	2.4494897	3571.00	3578.00
Beverly	6	6	2173.50	1.8708287	2171.00	2176.00
Boston Allston-Brighton	8	8	4.5000000	2.4494897	1.0000000	8.0000000
Boston Back Bay	6	6	104.5000000	2.7386128	101.0000000	108.0000000
Boston Beacon Hill	3	3	202.0000000	1.0000000	201.0000000	203.0000000
Boston Charlestown	6	6	403.5000000	1.8708287	401.0000000	406.0000000
Boston Dorchester	11	11	1006.00	3.3166248	1001.00	1011.00
Boston Downtown	8	8	706.2500000	2.8157719	702.0000000	710.0000000
Boston East Boston	12	12	498.3333333	28.9837990	407.0000000	512.0000000
Boston Forest Hills	7	7	1204.00	2.1602469	1201.00	1207.00
Boston Hyde Park	4	4	1402.50	1.2909944	1401.00	1404.00
Boston Mattapan	6	6	1103.50	1.8708287	1101.00	1106.00
Boston North End	2	2	301.5000000	0.7071068	301.0000000	302.0000000
Boston Roxbury	19	19	811.6315789	6.1843038	801.0000000	821.0000000
Boston Savin Hill	23	23	912.6521739	7.1896940	901.0000000	924.0000000
Boston South Boston	13	13	607.2307692	4.2259880	601.0000000	614.0000000
Boston West Roxbury	4	4	1302.50	1.2909944	1301.00	1304.00
Braintree	8	8	4194.50	2.4494897	4191.00	4198.00
Brookline	12	12	4006.50	3.6055513	4001.00	4012.00
Burlington	4	4	3322.50	1.2909944	3321.00	3324.00
Cambridge	30	30	3535.50	8.8034084	3521.00	3550.00

Analysis Variable : TractID						
TOWN	N Obs	N	Mean	Std Dev	Minimum	Maximum
Canton	3	3	4152.00	1.0000000	4151.00	4153.00
Chelsea	5	5	1603.60	2.0736441	1601.00	1606.00
Cohasset	1	1	4231.00	.	4231.00	4231.00
Concord	3	3	3612.00	1.0000000	3611.00	3613.00
Danvers	4	4	2112.50	1.2909944	2111.00	2114.00
Dedham	5	5	4023.00	1.5811388	4021.00	4025.00
Dover	1	1	4051.00	.	4051.00	4051.00
Duxbury	1	1	5071.00	.	5071.00	5071.00
Everett	7	7	3424.00	2.1602469	3421.00	3427.00
Framingham	10	10	3835.50	3.0276504	3831.00	3840.00
Hamilton	1	1	2151.00	.	2151.00	2151.00
Hanover	1	1	5031.00	.	5031.00	5031.00
Hingham	2	2	5011.50	0.7071068	5011.00	5012.00
Holbrook	2	2	4211.50	0.7071068	4211.00	4212.00
Hull	1	1	5001.00	.	5001.00	5001.00
Lexington	6	6	3584.33	2.1602469	3581.00	3587.00
Lincoln	1	1	3602.00	.	3602.00	3602.00
Lynn	22	22	2061.50	6.4935866	2051.00	2072.00
Lynnfield	2	2	2091.50	0.7071068	2091.00	2092.00
Malden	9	9	3415.00	2.7386128	3411.00	3419.00
Manchester	1	1	2181.00	.	2181.00	2181.00
Marblehead	3	3	2032.00	1.0000000	2031.00	2033.00
Marshfield	2	2	5061.50	0.7071068	5061.00	5062.00
Medfield	1	1	4061.00	.	4061.00	4061.00
Medford	11	11	3396.00	3.3166248	3391.00	3401.00
Melrose	4	4	3362.50	1.2909944	3361.00	3364.00
Middleton	1	1	2121.00	.	2121.00	2121.00
Millis	1	1	4071.00	.	4071.00	4071.00
Milton	4	4	4162.50	1.2909944	4161.00	4164.00
Nahant	1	1	2011.00	.	2011.00	2011.00
Natick	6	6	3823.50	1.8708287	3821.00	3826.00
Needham	5	5	4033.00	1.5811388	4031.00	4035.00
Newton	18	18	3739.50	5.3385391	3731.00	3748.00

Analysis Variable : TractID						
TOWN	N Obs	N	Mean	Std Dev	Minimum	Maximum
Norfolk	1	1	4091.00	.	4091.00	4091.00
North Reading	2	2	3301.50	0.7071068	3301.00	3302.00
Norwell	1	1	5041.00	.	5041.00	5041.00
Norwood	5	5	4133.00	1.5811388	4131.00	4135.00
Peabody	9	9	2105.00	2.7386128	2101.00	2109.00
Pembroke	2	2	5081.50	0.7071068	5081.00	5082.00
Quincy	12	12	4176.50	3.6055513	4171.00	4182.00
Randolph	3	3	4202.00	1.0000000	4201.00	4203.00
Reading	4	4	3342.50	1.2909944	3341.00	3344.00
Revere	8	8	1704.50	2.4494897	1701.00	1708.00
Rockland	2	2	5021.50	0.7071068	5021.00	5022.00
Salem	7	7	2044.00	2.1602469	2041.00	2047.00
Sargus	4	4	2082.50	1.2909944	2081.00	2084.00
Scituate	2	2	5051.50	0.7071068	5051.00	5052.00
Sharon	3	3	4142.00	1.0000000	4141.00	4143.00
Sherborn	1	1	3861.00	.	3861.00	3861.00
Somervile	1	1	3515.00	.	3515.00	3515.00
Somerville	14	14	3507.50	4.1833001	3501.00	3514.00
Stoneham	3	3	3372.00	1.0000000	3371.00	3373.00
Sudbury	2	2	3651.50	0.7071068	3651.00	3652.00
Swampscott	2	2	2021.50	0.7071068	2021.00	2022.00
Topsfield	1	1	2141.00	.	2141.00	2141.00
Wakefield	4	4	3352.50	1.2909944	3351.00	3354.00
Walpole	3	3	4112.00	1.0000000	4111.00	4113.00
Waltham	11	11	3686.00	3.3166248	3681.00	3691.00
Watertown	4	4	3702.50	1.2909944	3701.00	3704.00
Wayland	2	2	3661.50	0.7071068	3661.00	3662.00
Welesley	1	1	4042.00	.	4042.00	4042.00
Wellesley	3	3	4042.67	1.5275252	4041.00	4044.00
Wenham	1	1	2161.00	.	2161.00	2161.00
Weston	2	2	3671.50	0.7071068	3671.00	3672.00
Westwood	3	3	4122.00	1.0000000	4121.00	4123.00
Weymouth	8	8	4224.50	2.4494897	4221.00	4228.00

Analysis Variable : TractID						
TOWN	**N Obs**	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
Wilmington	3	3	3312.00	1.0000000	3311.00	3313.00
Winchester	5	5	3383.00	1.5811388	3381.00	3385.00
Winthrop	5	5	1803.00	1.5811388	1801.00	1805.00
Woburn	6	6	3333.50	1.8708287	3331.00	3336.00

Exercise B

As there are multiple observations that feature the same town in addition to using a proc means procedure with the variable being crime rate (Crim) and the class being Town a mean statement is used so that the mean is the average crime rate across towns. Then I sorted and printed the first 5 observations for both the ascending and descending order by the average crime rate. Below is the table that features the five towns with the highest average per capita crime rate across its tracts. These towns are Boston Charlestown, Boston South Boston, Boston Downtown, Boston Roxbury, and Boston North End.

Obs	TOWN	Average_Crim
1	Boston Charlestown	29.2019
2	Boston South Boston	21.2049
3	Boston Downtown	20.8953
4	Boston Roxbury	17.8646
5	Boston North End	14.8032

This next table features the five towns with the lowest average per capita crime rate across its tracts. These towns are Nahant, Medfield, Millis, Cohasset, and Topsfield.

Obs	TOWN	Average_Crim
1	Nahant	0.00632
2	Medfield	0.00906
3	Millis	0.01096
4	Cohasset	0.01301
5	Topsfield	0.01311

Exercise C

Investigating the distribution of MedV I first ran a means procedure and have attached the table below.

Analysis Variable : MedV Vector of Median Values of Owner-Occupied Housing in USD 1000							
N	Mean	Std Dev	Minimum	Maximum	Lower 95% CL for Mean	Upper 95% CL for Mean	Mode
506	22.5328063	9.1971041	5.0000000	50.0000000	21.7295280	23.3360846	50.0000000

The MedV variable is recorded for all 506 observations and has a mean of 22.53 which is within the 95% confidence level. The maximum of 50 is also the mode. To better see the spread of the distribution I then ran a frequency procedure with a chisquared option. The maximum and mode of 50 is recorded in 16 of the 506 observations while all other values are recorded having less than 8 observations. 313 observations more than sixty percent of the observations only have a frequency of 1 to 3 observations. At a 5% level of significance and a p value of 0.0029 I rejected the chi-square test of equal proportions. Both tables are attached below.

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
5	2	0.40	2	0.40
5.6	1	0.20	3	0.59
6.3	1	0.20	4	0.79
7	2	0.40	6	1.19
7.2	3	0.59	9	1.78
7.4	1	0.20	10	1.98
7.5	1	0.20	11	2.17
8.1	1	0.20	12	2.37
8.3	2	0.40	14	2.77
8.4	2	0.40	16	3.16
8.5	2	0.40	18	3.56
8.7	1	0.20	19	3.75
8.8	2	0.40	21	4.15
9.5	1	0.20	22	4.35
9.6	1	0.20	23	4.55
9.7	1	0.20	24	4.74
10.2	3	0.59	27	5.34
10.4	2	0.40	29	5.73
10.5	2	0.40	31	6.13
10.8	1	0.20	32	6.32
10.9	2	0.40	34	6.72
11	1	0.20	35	6.92
11.3	1	0.20	36	7.11
11.5	1	0.20	37	7.31
11.7	2	0.40	39	7.71
11.8	2	0.40	41	8.10
11.9	2	0.40	43	8.50
12	1	0.20	44	8.70
12.1	1	0.20	45	8.89
12.3	1	0.20	46	9.09

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
12.5	1	0.20	47	9.29
12.6	1	0.20	48	9.49
12.7	3	0.59	51	10.08
12.8	1	0.20	52	10.28
13	1	0.20	53	10.47
13.1	4	0.79	57	11.26
13.2	1	0.20	58	11.46
13.3	3	0.59	61	12.06
13.4	4	0.79	65	12.85
13.5	2	0.40	67	13.24
13.6	2	0.40	69	13.64
13.8	5	0.99	74	14.62
13.9	2	0.40	76	15.02
14	1	0.20	77	15.22
14.1	3	0.59	80	15.81
14.2	1	0.20	81	16.01
14.3	2	0.40	83	16.40
14.4	2	0.40	85	16.80
14.5	3	0.59	88	17.39
14.6	2	0.40	90	17.79
14.8	1	0.20	91	17.98
14.9	3	0.59	94	18.58
15	3	0.59	97	19.17
15.1	1	0.20	98	19.37
15.2	3	0.59	101	19.96
15.3	1	0.20	102	20.16
15.4	2	0.40	104	20.55
15.6	5	0.99	109	21.54
15.7	1	0.20	110	21.74
16	1	0.20	111	21.94

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
16.1	3	0.59	114	22.53
16.2	2	0.40	116	22.92
16.3	1	0.20	117	23.12
16.4	1	0.20	118	23.32
16.5	2	0.40	120	23.72
16.6	2	0.40	122	24.11
16.7	2	0.40	124	24.51
16.8	2	0.40	126	24.90
17	1	0.20	127	25.10
17.1	3	0.59	130	25.69
17.2	3	0.59	133	26.28
17.3	1	0.20	134	26.48
17.4	3	0.59	137	27.08
17.5	3	0.59	140	27.67
17.6	1	0.20	141	27.87
17.7	1	0.20	142	28.06
17.8	5	0.99	147	29.05
17.9	1	0.20	148	29.25
18	1	0.20	149	29.45
18.1	1	0.20	150	29.64
18.2	3	0.59	153	30.24
18.3	2	0.40	155	30.63
18.4	3	0.59	158	31.23
18.5	4	0.79	162	32.02
18.6	2	0.40	164	32.41
18.7	3	0.59	167	33.00
18.8	2	0.40	169	33.40
18.9	4	0.79	173	34.19
19	2	0.40	175	34.58
19.1	4	0.79	179	35.38

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
19.2	2	0.40	181	35.77
19.3	5	0.99	186	36.76
19.4	6	1.19	192	37.94
19.5	4	0.79	196	38.74
19.6	5	0.99	201	39.72
19.7	2	0.40	203	40.12
19.8	3	0.59	206	40.71
19.9	4	0.79	210	41.50
20	5	0.99	215	42.49
20.1	5	0.99	220	43.48
20.2	2	0.40	222	43.87
20.3	4	0.79	226	44.66
20.4	4	0.79	230	45.45
20.5	3	0.59	233	46.05
20.6	6	1.19	239	47.23
20.7	2	0.40	241	47.63
20.8	3	0.59	244	48.22
20.9	2	0.40	246	48.62
21	3	0.59	249	49.21
21.1	2	0.40	251	49.60
21.2	5	0.99	256	50.59
21.4	5	0.99	261	51.58
21.5	2	0.40	263	51.98
21.6	2	0.40	265	52.37
21.7	7	1.38	272	53.75
21.8	2	0.40	274	54.15
21.9	3	0.59	277	54.74
22	7	1.38	284	56.13
22.1	1	0.20	285	56.32
22.2	5	0.99	290	57.31

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
22.3	2	0.40	292	57.71
22.4	2	0.40	294	58.10
22.5	3	0.59	297	58.70
22.6	5	0.99	302	59.68
22.7	2	0.40	304	60.08
22.8	4	0.79	308	60.87
22.9	4	0.79	312	61.66
23	4	0.79	316	62.45
23.1	7	1.38	323	63.83
23.2	4	0.79	327	64.62
23.3	4	0.79	331	65.42
23.4	2	0.40	333	65.81
23.5	1	0.20	334	66.01
23.6	2	0.40	336	66.40
23.7	4	0.79	340	67.19
23.8	4	0.79	344	67.98
23.9	5	0.99	349	68.97
24	2	0.40	351	69.37
24.1	3	0.59	354	69.96
24.2	1	0.20	355	70.16
24.3	3	0.59	358	70.75
24.4	4	0.79	362	71.54
24.5	3	0.59	365	72.13
24.6	2	0.40	367	72.53
24.7	3	0.59	370	73.12
24.8	4	0.79	374	73.91
25	8	1.58	382	75.49
25.1	1	0.20	383	75.69
25.2	1	0.20	384	75.89
25.3	1	0.20	385	76.09

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
26.2	1	0.20	386	76.28
26.4	2	0.40	388	76.68
26.5	1	0.20	389	76.88
26.6	3	0.59	392	77.47
26.7	1	0.20	393	77.67
27	1	0.20	394	77.87
27.1	2	0.40	396	78.26
27.5	4	0.79	400	79.05
27.9	2	0.40	402	79.45
28	1	0.20	403	79.64
28.1	1	0.20	404	79.84
28.2	1	0.20	405	80.04
28.4	2	0.40	407	80.43
28.5	1	0.20	408	80.63
28.6	1	0.20	409	80.83
28.7	3	0.59	412	81.42
29	2	0.40	414	81.82
29.1	2	0.40	416	82.21
29.4	1	0.20	417	82.41
29.6	2	0.40	419	82.81
29.8	2	0.40	421	83.20
29.9	1	0.20	422	83.40
30.1	3	0.59	425	83.99
30.3	1	0.20	426	84.19
30.5	1	0.20	427	84.39
30.7	1	0.20	428	84.58
30.8	1	0.20	429	84.78
31	1	0.20	430	84.98
31.1	1	0.20	431	85.18
31.2	1	0.20	432	85.38

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
31.5	2	0.40	434	85.77
31.6	2	0.40	436	86.17
31.7	1	0.20	437	86.36
32	2	0.40	439	86.76
32.2	1	0.20	440	86.96
32.4	1	0.20	441	87.15
32.5	1	0.20	442	87.35
32.7	1	0.20	443	87.55
32.9	1	0.20	444	87.75
33	1	0.20	445	87.94
33.1	2	0.40	447	88.34
33.2	2	0.40	449	88.74
33.3	1	0.20	450	88.93
33.4	2	0.40	452	89.33
33.8	1	0.20	453	89.53
34.6	1	0.20	454	89.72
34.7	1	0.20	455	89.92
34.9	3	0.59	458	90.51
35.1	1	0.20	459	90.71
35.2	1	0.20	460	90.91
35.4	2	0.40	462	91.30
36	1	0.20	463	91.50
36.1	1	0.20	464	91.70
36.2	2	0.40	466	92.09
36.4	1	0.20	467	92.29
36.5	1	0.20	468	92.49
37	1	0.20	469	92.69
37.2	1	0.20	470	92.89
37.3	1	0.20	471	93.08
37.6	1	0.20	472	93.28

Vector of Median Values of Owner-Occupied Housing in USD 1000				
MedV	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
37.9	1	0.20	473	93.48
38.7	1	0.20	474	93.68
39.8	1	0.20	475	93.87
41.3	1	0.20	476	94.07
41.7	1	0.20	477	94.27
42.3	1	0.20	478	94.47
42.8	1	0.20	479	94.66
43.1	1	0.20	480	94.86
43.5	1	0.20	481	95.06
43.8	1	0.20	482	95.26
44	1	0.20	483	95.45
44.8	1	0.20	484	95.65
45.4	1	0.20	485	95.85
46	1	0.20	486	96.05
46.7	1	0.20	487	96.25
48.3	1	0.20	488	96.44
48.5	1	0.20	489	96.64
48.8	1	0.20	490	96.84
50	16	3.16	506	100.00

Chi-Square Test for Equal Proportions	
Chi-Square	291.4269
DF	228
Pr > ChiSq	0.0029
WARNING: The table cells have expected counts less than 5. Chi-Square may not be a valid test.	