Matthew Martinez

Part I

1. **Cumulative Distribution Function** – the probability of that x is less than or equal to a number x; $F(x) = P(X <= x)$
2. **Nonparametric Test** – a test for different statistics that does not depend on the distribution of the data being tested
3. **Null Hypothesis (H_0)** – this is the research hypothesis or what you seek to reject or fail to reject
4. **Normality** – refers to a mound shaped distribution with a mean of zero and a standard deviation of one
5. **μ and ybar** – mu denotes a population parameter and ybar denotes a sample mean derived from a theoretical population
6. **Critical Value** – this is the value which you set as a researcher which you use to say a test is significant; usually the critical value is set at $p<0.05$
7. **Exploratory Data Analysis** – this is primary data analysis which is done before any tests are run on your data and can include things like histograms of distributions which can give researchers an idea of the variables that are in the dataset and allow them to begin to develop more sophisticated and structured analysis
8. **Relative Frequency Histogram** – a histogram which uses the proportions of each interval to graph the histogram; the sum of the relative frequencies will add up to 1
9. **t-test** – a statistical test which tests for differences in means of two samples to see if they are statistically significant; null hypothesis of no difference
10. **Ecological fallacy** – making assumptions about individuals from aggregate data of the groups in which the individuals belong

Matthew Martinez

Part II

Alamo Area Council of Governments
8700 Tesoro Dr., Suite 700
San Antonio, TX 78217-6228

In this report you will find pooled and individual demographic profiles for members of the

Alamo Area of Council of Governments (AACOG). AACOG consists of 12 county members who include:

Atascosa, Bandera, Bexar, Comal, Frio, Gillespie, Guadalupe, Karnes, Kendall, Kerr, Medina, and Wilson

Counties. The data presented in this report are from the 2006-2010 American Communities Survey 5

year summary file and are computed at the census tract level. Census tracts are small statistical

subdivisions of a county created by the Census Bureau and the 12 counties in AACOG include 475 tracts

with useable data. Four tracts in Bexar County are missing appropriate data and are thus excluded from

this report. There are five demographic categories specified for the purposes of this report which

include sex, age, monetary demographics, race/ethnicity, and education.

Sex is shown by the proportion of males and females in the area while age statistics are

displayed as percentages for ages 0-85+ in five year intervals. Ages were given in such detail so that a

characterization of the area can be generalized with areas being classified as younger or older

depending on the clustering of ages in the intervals. The monetary demographic variables include a

trichotomous household income variable, the poverty rate and a place of work variable. Household

income was divided into three categories with household income being either less than $50,000,

between $50,000 and $99,999, or greater than or equal to $100,000. This will allow the council to see

where there are areas of wealth and need. To further aid the council in seeing which areas are in need,

the poverty rate for the region is included which includes the percent of those whose income in the past

12 months was below the poverty level. The place of work variable is a proportion of workers ages 16

years and older who work outside of their county of residence. Included in this proportion are workers

who work outside of the state of Texas which is included with the assumption that if a worker is working

outside of the state they will be working outside of the Texas county they reside in.

The demographic category for race/ethnicity is divided into five separate racial/ethnic groups

including Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic Asian, and non-Hispanic Native

American. The race/ethnicity variable is given in percentages. It is important to note that these

percentages may not add up to 100% due to rounding error and the exclusion of some minor

racial/ethnic categories like non-Hispanic other. Education is given as a percentage of the population 25 and older which has either some college to an Associates, Bachelor's degree, or has a Master's degree or greater. Once again, percentages in the area will not add up to 100% due to the exclusion of those with less than some college education. By categorizing the education variable in this way, one can see which areas have higher levels of education beyond a high school level degree.

Demographic Profile – AACOG Total Area

Table 1 found at the end of this report includes means, medians, ranges, and standard deviations for all the variables used for the pooled AACOG area. The average census tract in the AACOG area has 4,515 residents with a range of 10,179. The area seems balanced between males and females with males making up approximately 49% of the area population and females 51%. The area is comprised of a younger population with 43.84% of residents being under the age of 30. The largest percentages are in the age categories under 30 with an exception being found in the 45-49 year old age category which comprises 7.18% of the area population. In terms of age, the largest variability is found in the 15-19 and 20-24 year old age categories as evidenced by a range of 37.89 and 39.10 respectively. The 20-24 year old age category also has the largest standard deviation from the mean of 4.42 out of all the age categories. The range and large standard deviation may be explained by the inclusion of a census tract which has zero percent of its residents within this age category. With that said, there are other age categories with similar attributes, so this large variability is not totally explained by this characteristic.

On average, the AACOG area has 16.31% of its residents living in households with incomes below the poverty rate. This poverty rate appears to be skewed with a range of 69.11 meaning that the difference between the maximum poverty rate in the census tracts and the minimum value is 69.11%. In this case, the median of 13.37% may be a better indicator of central tendency or what is more commonly found in the area. A good majority (51.87%) of households in the area have an income of less than $50,000 a year. Only 17.63% of households in the AACOG area have incomes of $100,000 or over. These income measures have similar amounts of variability as the poverty rate, so using the medium as a measure of what is common in the area may be more appropriate. The median value for income over $100,000 in the area is 13.44% which is a value that is not affected by extreme values.

Residents in the area tend to work within their county of residence with 85% of workers 16 years and older working in the county they reside in. Examining this variable further, it appears that this

average is being pulled up by census tracts with over half of their residents working outside of the county they reside in. It should be assumed that these are border tracts and these large proportions are found in census tracts which are on the county border. In order for this assumption to be tested, a map of these areas should be analyzed to see if the border census tracts in fact have a disproportionate amount of their residents working outside of the county.

With 52.64%, Hispanics comprise the greatest racial/ethnic percentage in the AACOG area. Non-Hispanic whites make up the second largest racial/ethnic category with 38.18% of residents in the area being from this category. Less than one percent of residents in the area are non-Hispanic Native American and less than two percent of residents are non-Hispanic Asian. With this demographic profile, services and outreach to Hispanics may be vital in implementing any service or policy in the AACOG area. Lastly, a little over half of the population in the AACOG area has some college education or more. With a quarter of the population having a Bachelor's degree or greater, the area is relatively well educated. It would be helpful to see who those with some college education, but less than a Bachelor's degree are. If these individuals are presently on their way to achieving a Bachelor's degree but have not yet, then there could be services to help them achieve this next step in educational attainment. Considering educational attainment is measured for those 25 years or over, they are not likely to be traditional students. The other alternative is that these individuals achieved their educational ambitions with an Associate's or technical degree and would not benefit from such services. Either way, this group warrants more research if AACOG services are to be directed in this way.

Demographic Profile – AACOG Individual Members

In the last section, the demographic profile for AACOG members as a whole was discussed. In the current section, the averages for each of the 12 member counties will be analyzed. Table 2 found at the end of this report includes the means for all variables discussed early for each of the 12 member counties in the AACOG service area. The county with the highest average census tract population is Medina which has an average of 5,634 people per tract within the county. This may not mean much at face value which is why the number of census tracts is given for each county. To get an approximate number of residents in each county, the census tract average should be multiplied by the number of census tracts. Using this formula, Bexar County would be the most populous county of the AACOG members with approximately nearly 1,650,000 residents. In terms of sex, Karnes County has the

greatest proportion of males with 0.61 and Frio County has the second highest proportion of males with 0.57.

Looking at the age profile for each county, Atascosa County has the greatest percentage of individuals under age five and Karnes County has the smallest percentage of persons under the age of five. On the opposite end of the age spectrum, Kerr County has the largest percentage of persons 85 years and over while Frio County has the smallest percentage of its population 85 years and over. More specific information in terms of dependency ratios and youth ratios are needed and can be calculated from the data presented if the council wants to direct policies or labor resources to those areas with higher dependency ratios. I would also like to highlight Karnes County's age composition as nearly 19% of the population is between ages 20 and 30. This is by far the highest percentage of individuals within this age range for any member of AACOG.

Frio County has the highest poverty rate at 21.65 percent of households earning incomes below the poverty level. Kendall County has the lowest poverty rate at 7.15 and it also has the highest percentage of households making $100,000 or more. Wilson County has the most households in the middle income bracket of $50,000 to $99,999 and has a relatively low poverty rate of about 8.90 percent. Wilson County also has the highest proportion of workers 16 and older working outside of their county of residence and Bexar County has the lowest proportion. The low level of workers 16 and older working outside of Bexar County may be due to the size of the county and the labor market of the county.

In terms of race/ethnicity, Atascosa County has the highest percentage of Hispanics at 65.71% and Bandera County has the lowest percentage of Hispanics at 15.92%. Bandera County and Gillespie County have the highest percentage of non-Hispanic Whites at 81.31% and 80.08% respectively. In terms of higher education, it appears that Atascosa County is the least educated as only 3.22% of its residents have a Master's or greater and 27.61% of its residents have some college to a Bachelor's. With 30.83% of its residents having some college or greater, Atascosa County is well below AACOG's average of 50.83%. Attention should be paid to this area in terms of focusing educational resources to help its residence either have better access to education or to recruit more industries which require the skills that higher education brings.

The demographic profile for AACOG members was presented in this report. Hopefully the council can use the data presented here to inform decisions on which areas may benefit most from the varying resources offered by AACOG. This report is meant to be a brief synopsis about the demographic

Matthew Martinez

characteristics of the area and further avenues for research and investigation have been given to better help AACOG in the decision making process. If you have any questions please feel free to contact me and my associates of the Hard Knock Life Foundation.

Matthew Martinez

Table 1 - Demographic Profile for ANALOG

Characteristics	Pooled Total For All Members				
	Mean	Median	Range	Std. Dev.	N
Total Population -Tract Average	4,515	4,449	10,179	1,817	479
Sex					
Proportion of Male	0.49	0.49	0.42	0.04	475
Proportion of Female	0.51	0.51	0.42	0.04	475
Age - Percentages of					
Persons Under 5 years	7.38	7.05	24.29	3.09	475
Persons 5 to 9 years	7.26	7.17	18.08	2.84	475
Persons 10 to 14 years	7.36	7.23	16.88	2.59	475
Persons 15 to 19 years	7.49	7.09	37.89	3.55	475
Persons 20 to 24 years	7.22	6.63	39.10	4.42	475
Persons 25 to 29 years	7.13	6.76	21.66	3.55	475
Persons 30 to 34 years	6.50	6.24	16.43	2.75	475
Persons 35 to 39 years	6.69	6.33	17.26	2.46	475
Persons 40 to 44 years	6.79	6.66	12.66	2.32	475
Persons 45 to 49 years	7.18	6.99	15.41	2.40	475
Persons 50 to 54 years	6.65	6.54	18.28	2.46	475
Persons 55 to 59 years	5.88	5.56	13.86	2.39	475
Persons 60 and 64 years	4.81	4.62	16.13	2.32	475
Persons 65 and 69 years	3.59	3.30	14.97	2.04	475
Persons 70 to 74 years	2.72	2.54	13.30	1.71	475
Persons 75 to 79 years	2.27	1.94	11.70	1.66	475
Persons 80 to 84 years	1.61	1.30	9.75	1.41	475
Persons 85 years and over	1.46	0.97	14.51	1.73	475
Monetary - Percent (Unless noted)					
Household Income - Less Than 50k	51.87	52.91	93.18	21.30	475
Household Income - 50-99,999k	30.49	29.80	61.72	11.11	475
Household Income - 100k or Greater	17.63	13.37	87.62	16.05	475
Poverty Rate	16.31	13.44	69.11	12.32	475
Proportion of Workers 16 or Older Who Commute Outside of County	0.15	0.05	0.91	0.22	475
Race/Ethnicity - Percent					
Hispanic	52.64	48.72	97.57	25.68	475
Non-Hispanic White	38.18	35.10	95.05	25.27	475
Non-Hispanic Black	5.82	3.27	64.63	8.42	475
Non-Hispanic Asian	1.80	0.81	28.28	2.71	475
Non-Hispanic Native American	0.10	0.00	6.85	0.51	475
Education - Percent					
Some College to Associates	26.38	26.34	57.68	8.01	475
Bachelors	15.60	14.20	46.90	10.61	475
Masters or Greater	8.85	6.59	45.06	8.34	475

Data from 2006-2010 ACS Summary File

Matthew Martinez

Table 2 - Demographic Profile for Individual ACCOG Members (Means)

Characteristics	Atascosa	Bandera	Bexar	Comal	Frio	Gillespie	Guadalupe	Karnes	Kendall	Kerr	Medina	Wilson
Total Population - Tract Average	5,506	4,053	4,508	4,285	5,652	4,839	4,232	3,720	5,253	4,883	5,634	3,751
Number of Census Tracts	8	5	366	24	3	5	29	4	6	10	8	11
Sex												
Male	0.50	0.49	0.49	0.49	0.57	0.49	0.49	0.61	0.49	0.49	0.50	0.49
Female	0.50	0.51	0.51	0.51	0.43	0.51	0.51	0.39	0.51	0.51	0.50	0.51
Age - Percentages of												
Persons Under 5 years	7.96	4.29	7.76	5.81	6.36	4.91	7.04	3.91	5.50	5.48	6.62	6.13
Persons 5 to 9 years	7.47	5.69	7.53	5.97	6.38	4.89	7.05	3.85	5.98	5.43	7.71	6.91
Persons 10 to 14 years	8.09	6.88	7.33	7.51	7.88	6.14	8.26	3.21	8.33	5.40	7.52	8.50
Persons 15 to 19 years	7.43	4.26	7.69	6.65	8.70	5.22	7.60	3.31	6.80	6.31	7.61	7.72
Persons 20 to 24 years	5.66	3.67	7.78	4.81	8.13	4.52	5.82	8.02	4.28	5.60	5.88	5.03
Persons 25 to 29 years	5.86	4.06	7.68	5.06	9.96	3.53	5.76	10.58	4.29	4.19	5.40	4.71
Persons 30 to 34 years	5.82	3.62	6.85	5.15	7.01	3.76	6.17	8.88	4.43	3.96	4.93	5.55
Persons 35 to 39 years	5.53	5.35	6.85	5.77	6.09	4.92	7.07	8.32	5.81	5.74	5.44	6.62
Persons 40 to 44 years	7.20	5.92	6.72	7.46	6.26	5.43	7.51	7.62	7.33	4.52	7.57	7.60
Persons 45 to 49 years	7.66	8.42	6.97	8.17	6.08	7.16	7.78	9.07	8.76	6.95	7.95	7.97
Persons 50 to 54 years	7.45	9.50	6.36	8.42	6.33	7.43	6.99	6.57	7.60	7.19	7.53	7.48
Persons 55 to 59 years	6.99	9.45	5.44	8.33	5.30	8.65	5.97	6.82	7.45	8.21	7.40	6.75
Persons 60 and 64 years	4.94	8.63	4.46	6.21	4.36	7.27	5.31	4.49	7.73	6.13	5.13	6.37
Persons 65 and 69 years	3.72	5.32	3.27	4.86	3.15	7.22	3.96	3.79	4.96	6.30	3.79	4.35
Persons 70 to 74 years	3.37	6.24	2.43	3.69	3.13	5.86	2.73	3.18	3.72	5.87	3.43	2.60
Persons 75 to 79 years	1.73	4.15	2.02	2.96	2.43	5.89	2.32	3.38	2.39	5.02	2.45	3.38
Persons 80 to 84 years	1.67	2.88	1.48	1.49	1.56	3.74	1.56	3.00	1.99	4.30	2.18	1.13
Persons 85 years and over	1.45	1.67	1.39	1.67	0.87	3.47	1.09	1.97	2.64	3.40	1.45	1.19

Matthew Martinez

Table 2 - Demographic Profile for Individual ACCOG Members (Means)

Characteristics	Atascosa	Bandera	Bexar	Comal	Frio	Gillespie	Guadalupe	Karnes	Kendall	Kerr	Medina	Wilson
Monetary - Percent (Unless noted)												
Household Income - Less Than 50k	58.30	57.58	53.86	37.85	64.30	47.47	39.99	62.70	37.70	56.62	50.63	40.26
Household Income - 50-99,999k	29.83	28.56	29.49	34.58	26.34	37.04	37.78	23.43	28.53	28.09	32.21	39.40
Household Income - 100k or Greater	11.87	13.86	16.65	27.57	9.37	15.49	22.23	13.87	33.77	15.29	17.16	20.34
Poverty Rate	19.78	18.68	17.67	9.85	21.65	8.36	9.55	18.48	7.15	13.74	15.86	8.90
Proporiton of Workers 16 or Older Who Communte Outside of County	0.53	0.54	0.05	0.48	0.19	0.18	0.60	0.33	0.46	0.13	0.48	0.63
Race/Ethnicity - Percent												
Hispanic	65.71	15.92	58.51	24.18	76.30	18.16	34.75	43.54	19.58	23.00	48.55	36.75
Non-Hispanic White	32.52	81.31	30.93	72.12	19.94	80.08	56.40	49.40	77.94	73.12	47.99	60.20
Non-Hispanic Black	0.55	0.49	6.80	1.47	3.09	0.04	5.90	5.50	0.98	1.92	1.82	1.76
Non-Hispanic Asian	0.35	0.61	2.15	0.70	0.35	0.21	1.17	0.15	0.76	0.81	0.60	0.23
Non-Hispanic Native American	0.01	0.13	0.11	0.01	0.00	0.11	0.12	0.97	0.00	0.05	0.00	0.00
Education - Percent												
Some College to Associates	21.12	29.05	26.40	26.66	17.62	26.12	27.90	20.98	29.11	26.03	27.12	26.46
Bachelors	6.49	16.99	15.57	21.53	4.73	17.31	15.32	6.55	22.72	17.65	12.45	12.60
Masters or Greater	3.22	7.53	9.05	11.00	2.92	9.85	8.63	2.54	12.83	9.30	6.57	5.26

Data from 2006-2010 ACS Summary File

Matthew Martinez

Part III

Table 1 below shows the definitions for the variables measuring risk for natural disasters for coastal regions in Texas which will be covered in this section. Along with these definitions, are a list of the counties included in each of the areas of the Southern, Northern, and Middle Coasts. These are all counties in Texas along the Texas Gulf Coasts and level of aggregation will be census tracts. The risk index is calculated based on the formula below:

$$risk\ index = -.02 * POver65 + .38 * PPoverty + .52 * PWithout\ High\ School + .25$$
$$* PLinguistically\ Isolated\ + .09 * PWithout\ Phone + .26 * PWithoutVehicle$$
$$+ .66 * PRenting\ ; where\ P = Proportion$$

Table 1 - Variable Definitions

Variables	Definition
Proportion Over 65	Proportion of the population that is age 65 and older
Proportion in Poverty	Proportion of the population in poverty
Proportion Without High School	Proportion of the population 25 years and older without a high school education: Less than 9th Grade
Proportion Linguistically Isolated	Proportion of the population in housholds in which no one 14 and over speaks English only or speaks a language other than English and speaks English 'very well'
Proportion Without Phone	Proporiton of occupied housing units wihtout telephone service
Proportion Without Vehicle	Proportion of occupied housing units without access to a vehicle
Proportion Rented	Proportion of occupied housing units that are rented
Risk Index	An index of risk calculated from the above variables
Areas	
Southern Coast	Census tracts in Cameron, Willacy, Kenedy, Kleberg, and Nueces Counties
Middle Coast	Census tracts in Brazoria, San Patricio, Aransas, Refugio, Calhoun, Jackson, and Matagorda Counties
Northern Coast	Census tracts in Galveston, Harris, Chambers, and Jefferson Counties

Matthew Martinez

Table 2 - Descriptive Statistics of Variables Measuring Risk for Natural Disasters for Coastal Regions in Texas

Variables	Pooled Sample			Southern Coast			Middle Coast			Northern Coast		
	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev	N
Proportion Over 65	0.099	0.056	1196	0.116	0.053	176	0.128	0.066	91	0.093	0.054	929
Proportion in Poverty	0.191	0.139	1192	0.291	0.171	176	0.138	0.074	91	0.177	0.129	925
Proportion Without High School	0.128	0.117	1196	0.189	0.138	176	0.088	0.063	91	0.120	0.112	929
Proportion Linguistically Isolated	0.121	0.127	1192	0.151	0.132	176	0.054	0.051	91	0.122	0.130	925
Proportion Without Phone	0.049	0.049	1192	0.060	0.083	176	0.045	0.036	91	0.047	0.041	925
Proportion Without Vehicle	0.080	0.086	1192	0.103	0.112	176	0.056	0.047	91	0.078	0.082	925
Proportion Rented	0.394	0.241	1192	0.360	0.194	176	0.271	0.124	91	0.413	0.253	925
Risk Index	0.511	0.282	1192	0.582	0.262	176	0.360	0.150	91	0.513	0.291	925

Data from 2006-2010 ACS Summary File

Table 2 above shows the descriptive statistics of the variables used in this analysis for the entire Gulf Coast (Pooled Sample) and each coastal region. As can be seen, the Middle Coast region has the greatest proportion of individuals over the age of 65, but it has the lowest risk index score of 0.360. The Southern Coast region has the highest proportion of individuals in poverty, the highest proportion of individuals without a high school education, highest average proportion of those linguistically isolated, highest proportion of those without a vehicle or phone, and the highest risk index. The Northern Coast region has the highest proportion of occupied housing units that are rented at 0.413. The descriptive statistics can give an idea of which areas are more at risk, but statistical tests are needed in order to determine if these differences are real or non-zero.

Considering we are dealing with three separate groups and variables that are continuous, assumptions for the ANOVA test are first run to see if this test can be used to examine equality between these variable across the three coastal regions. If these assumptions fail, the non-parametric Kruskal-Wallis test will be used to determine equality of each variable across the regions. To begin, I tested for normality of each variable to see if it was drawn from a normal distribution. Shapiro-Wilk tests done on each variable shows that the proportion over 65, proportion in poverty, proportion without a high school education, proportion linguistically isolated, proportion without phone, proportion without a vehicle, proportion renting, and the risk index are not drawn from a normally distributed population which is evidenced by the rejection of the null hypothesis of normality. We can assume that none of the variables in this analysis can be said to be normally distributed. The next section will contain the other assumptions of the ANOVA test for each variable along with a report of the equality test chosen for each variable.

For the proportion over 65 variable, the null hypothesis of the Bartlett test is rejected meaning that the variances are not the same. This variable also fails the last assumption of the ANOVA test as the residuals are not normal or of constant variance as shown by the rejection of the null hypothesis of the Shapiro-Wilk test. Therefore, the non-parametric Kruskal-Wallis test was run to see if there is equality with the proportion over 65 variable in each region. The results are shown in Table 3 and we can assume that the proportion over 65 is not equal among the three coastal regions in Texas.

The proportion in poverty, proportion without high school education, proportion linguistically isolated, proportion without telephone service, proportion without access to a vehicle, and proportion renting variables were also tested to see if they met the requirements for an ANOVA test. It was found that in all cases, the null hypothesis of the Bartlett test was rejected and all variables failed the

assumption of the ANOVA test that the residuals are not normal or of constant variance as shown by the rejection of the null hypothesis of the Shapiro-Wilk test. Therefore, the non-parametric Kruskal-Wallis test was run to see if there is equality with the variables in each region. The results are shown in Table 3 and we can assume that the all variables are not equal among the three coastal regions in Texas. With the critical value set at 0.05, the proportion without a phone variable is still considered to not be equal across coastal regions, but it has a p-value of 0.01 which is larger than the other variables.

Table 3 - Kruskal-Wallis Results for Variables failing ANOVA Test Assumptions

	Proportion Over 65	Proportion in Poverty	Proportion Without High School	Proportion Linguistically Isolated	Proportion Without Phone	Proportion Without Vehicle	Proportion Rented	Risk Index
Chi-Square	62.99	83.43	45.51	43.51	8.9	16.46	27.65	41.38
Degrees of Freedom	2	2	2	2	2	2	2	2
P-Value	<.0001	<.0001	<.0001	<.0001	0.01	0.0003	<.0001	<.0001

Figure 1 below shows a histogram of the risk index for each coastal region along with the pooled histogram of the entire Texas Gulf Coast. In order to test for equality of the risk index across the three areas of the Texas Gulf Coast, a non-parametric test was used due to the failure of the assumptions of the ANOVA test. The results of the Kruskal-Wallis test are shown in Table 3 and it can be assumed that the risk index is different across regions. As stated earlier, the Southern Coast had the highest average risk index at 0.582. Cameron County had the tract with the highest risk index at 1.50 and Harris County had the tract with the lowest risk index at 0.016. There are 37 tracts in four counties with risk indices greater than 1.1. The four counties are Cameron, Galveston, Harris, and Nueces.

Given the information above, Texas should focus on the Southern and Northern Coastal regions as these two regions had the highest values on the risk index. Special attention should be paid to the Southern Coastal region as it has the highest proportions in all variables except proportion over 65 and proportion of occupied housing units that are rented. Overall, Texas should focus primarily on reducing the proportion of those without a high school education in the area as this risk factor is quite large in the area as a whole and especially in the Southern Coastal region. By focusing attention on education in the area, the state could also reduce other proportions associated with risk like the proportion in poverty or the proportion that is linguistically isolated. Also, if technical skills can be learned, this may increase the wealth in the area and give households access to telephone service and vehicles. The state may also consider adding the proportion of households with someone with mobility/disability issues as special assistance may be needed to vacate these individuals when natural disasters such as hurricanes are

Matthew Martinez

imminent. Along with measuring the proportion of those age 65 and over, the state may consider

accounting for the proportion of those under the age of 14 as these individuals may be contribute to the
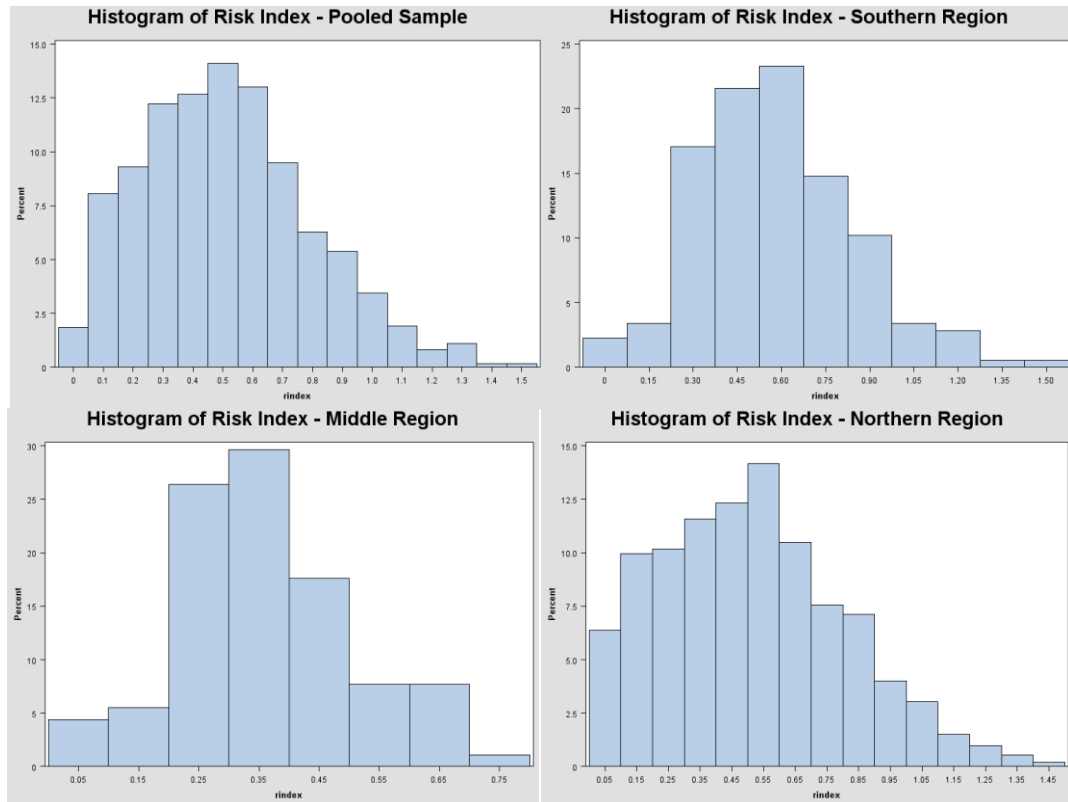
risk factor.

Figure 1 – Histograms of the Risk Index by Region

Matthew Martinez

Appendix

```sas
options helpbrowser=sas;
libname home 'C:\Users\Matt\Documents\Graduate School\Data Set'; run;

****************************PART 2 – AACOG ****************************;

data home.working; set home.acssf_5yr_tx48;

where sumlev='140' AND  (county='013' OR county='019' OR county='029'
OR county='091' OR county='163' OR county='171' OR county='187' OR
county='255'
OR county='259' OR county='265' OR county='325' OR county='493') ;

run;

data now; set home.working;

Male = B01001002;
Female = B01001026;
pmale= (male/B01001001);
pfemale= (female/B01001001);

*Creating five year age intervals;
aunder5= (B01001003 + B01001027);
a5to9= (B01001004 + B01001028);
a10to14= (B01001005 + B01001029);
a15to19= (B01001006 + B01001030+B01001007 + B01001031);
a20to24= (B01001008 + B01001032+B01001009 + B01001033+B01001010 + B01001034);
a25to29= (B01001011 + B01001035);
a30to34= (B01001012 + B01001036);
a35to39= (B01001013 + B01001037);
a40to44= (B01001014 + B01001038);
a45to49= (B01001015 + B01001039);
a50to54= (B01001016 + B01001040);
a55to59= (B01001017 + B01001041);
a60to64= (B01001018 + B01001042+B01001019 + B01001043);
a65to69= (B01001020 + B01001044+B01001021 + B01001045);
a70to74= (B01001022 + B01001046);
a75to79= (B01001023 + B01001047);
a80to84= (B01001024 + B01001048);
a85plus= (B01001025 + B01001049);



*Percentages of Age;
paunder5= (aunder5/B01001001)*100;
pa5to9= (a5to9/B01001001)*100;
pa10to14= (a10to14/B01001001)*100;
pa15to19= (a15to19/B01001001)*100;
pa20to24= (a20to24/B01001001)*100;
pa25to29= (a25to29/B01001001)*100;
pa30to34= (a30to34/B01001001)*100;
pa35to39= (a35to39/B01001001)*100;
pa40to44= (a40to44/B01001001)*100;
pa45to49= (a45to49/B01001001)*100;
pa50to54= (a50to54/B01001001)*100;
```

```
pa55to59= (a55to59/B01001001)*100;
pa60to64= (a60to64/B01001001)*100;
pa65to69= (a65to69/B01001001)*100;
pa70to74= (a70to74/B01001001)*100;
pa75to79= (a75to79/B01001001)*100;
pa80to84= (a80to84/B01001001)*100;
pa85plus= (a85plus/B01001001)*100;

*Income;
less50=sum (of B19001002-B19001010);
more50=sum (of B19001011-B19001013);
more100=sum (of B19001014 - B19001017);

pless50= (less50/B19001001)*100;
pmore50= (more50/B19001001)*100;
pmore100= (more100/B19001001)*100;

*Poverty Rate;
povrate=(B17001002/B17001001)*100;

*Race;
Racetotal=B03002001;

Hispanic=B03002012;
pcthispanic=(Hispanic/Racetotal)*100;
phispanic=(Hispanic/Racetotal);

NhWhite=B03002003;
pctNhWhite=(NhWhite/Racetotal)*100;
pNhWhite=(NhWhite/Racetotal);

NhBlack=B03002004;
pctNhBlack=(NhBlack/Racetotal)*100;
pNhBlack=(NhBlack/Racetotal);

NhAsian=B03002006;
pctNhAsian=(NhAsian/Racetotal)*100;
pNhAsian= (NhAsian/Racetotal);

NhNative=B03002007;
pctNhNative= (NhNative/Racetotal)*100;
pNhNative=(NhNative/Racetotal);

*Workers;
Outsidework=B08007004+B08007005;
pwork=(Outsidework/B08007001);

*Education;
mc=(B15002012 + B15002013 + B15002014);
fc=sum (of B15002029-B15002030);
ac=mc+fc;
pac= (ac/B15002001)*100;

BA=B15002015+B15002032;
pba= (BA/B15002001)*100;

mma=sum (of B15002016-B15002018);
```

```
fma=sum (of B15002033 - B15002035);
maplus=mma+fma;
pmaplus= (maplus/B15002001)*100;

*Renaming Counties;
if county='013' then countyname='Atascosa';
if county='019' then countyname='Bandera' ;
if county='029' then countyname='Bexar';
if county='091' then countyname='Comal';
if county='163' then countyname='Frio';
if county='171' then countyname='Gillespie';
if county='187' then countyname='Guadalupe';
if county='255' then countyname='Karnes';
if county='259' then countyname='Kendall';
if county='265' then countyname='Kerr';
if county='325' then countyname='Medina';
if county='493' then countyname='Wilson';
run;

proc freq data=home.working; tables county; run;
proc freq data=now; tables countyname county tract B08007005; run;

proc univariate data=now; var B19001001; run;

proc means mean median range std n max min data=now; var B01001001 pmale male
pfemale paunder5 pa5to9 pa10to14 pa15to19 pa20to24 pa25to29
pa30to34 pa35to39 pa40to44 pa45to49 pa50to54 pa55to59 pa60to64
pa65to69 pa70to74 pa75to79 pa80to84 pa85plus
pless50 pmore50 pmore100 povrate pwork
pctHispanic pctNhWhite pctNhBlack pctNhAsian pctNhNative
pac pba pmaplus; run;

proc sort data=now; by countyname; run;
proc means mean  data=now; var B01001001 pmale pfemale paunder5 pa5to9
pa10to14 pa15to19 pa20to24 pa25to29
pa30to34 pa35to39 pa40to44 pa45to49 pa50to54 pa55to59 pa60to64
pa65to69 pa70to74 pa75to79 pa80to84 pa85plus
pless50 pmore50 pmore100 povrate pwork
pctHispanic pctNhWhite pctNhBlack pctNhAsian pctNhNative
pac pba pmaplus; by countyname; run;

proc univariate data=now plot; var pa20to24 pwork ; histogram; run;
proc univariate data=now; var B01001001; by countyname; run;


*Finding missing total counts - why are these tracts missing B01001001
counts?;
data nowtest; set now;
keep tract countyname B01001001  B01001026 pa20to24;
run;


***************************PART 3 - Gulf Coast ***************************;

data data home.working1; set home.acssf_5yr_tx48;

where sumlev='140' AND
```

```sas
(county='061' OR county='489' OR county='261'
OR county='273' OR county='355' OR county='039' OR county='409' OR
county='007'
OR county='391' OR county='057' OR county='239' OR county='321'
OR county='167' OR county='201' OR county='071' OR county='245') ;

run;


data now; set home.working1;

*Renaming Counties;
if county='061' then countyname='Cameron';
if county='489' then countyname='Willacy' ;
if county='261' then countyname='Kenedy';
if county='273' then countyname='Kleberg';
if county='355' then countyname='Nueces';
if county='039' then countyname='Brazori';
if county='409' then countyname='San Pat';
if county='007' then countyname='Aransas';
if county='391' then countyname='Refugio';
if county='057' then countyname='Calhoun';
if county='239' then countyname='Jackson';
if county='321' then countyname='Matagor';
if county='167' then countyname='Galvest';
if county='201' then countyname='Harris';
if county='071' then countyname='Chamber';
if county='245' then countyname='Jeffers';

Southern=0;
if countyname='Cameron' then Southern=1;
else if countyname='Willacy' then Southern=1;
else if countyname='Kenedy' then Southern=1;
else if countyname='Kleberg' then Southern=1;
else if countyname='Nueces' then Southern=1;

Middle=0;
if countyname='Brazori' then Middle=1;
else if countyname='San Pat' then Middle=1;
else if countyname='Aransas' then Middle=1;
else if countyname='Refugio' then Middle=1;
else if countyname='Calhoun' then Middle=1;
else if countyname='Jackson' then Middle=1;
else if countyname='Matagor' then Middle=1;

Northern=0;
if countyname='Galvest' then Northern=1;
else if countyname='Harris' then Northern=1;
else if countyname='Chamber' then Northern=1;
else if countyname='Jeffers' then Northern=1;

region=.;
if Northern=1 then region=1;
else if Middle=1 then region=2;
else if Southern=1 then region=3;

*Age;
```

Matthew Martinez

```
M65plus=sum (of B01001020-B01001025);
F65plus=sum (of B01001044-B01001049);
T65plus= (M65plus + F65plus);
p65= (T65plus/B01001001);

*Poverty;
ppov=(B17001002/B17001001);

*Education;
mless=sum (of B15002003-B15002006);
fless=sum (of B15002020-B15002023);
lesshigh=mless+fless;
pwohsedu=(lesshigh/B15002001);

*Language;
liniso=(B16002004 + B16002007 + B16002010 + B16002013);
plingiso= (liniso/B16002001);

*Telephone;
Noservice= (B25043007+B25043016);
phwophone=(Noservice/B25043001);

*Vehicle;
Nocar=(B25044003+B25044010);
phwoveh= (Nocar/B25044001);

*Renters;
rent=B25003003;
prented= (rent/B25003001);

*Risk Index;
rindex= ((-.02*p65) + (.38*ppov) + (.52*pwohsedu) + (.25*plingiso)
+ (0.09*phwophone) + (0.26*phwoveh) + (.66*prented));

keep p65 ppov pwohsedu plingiso phwophone phwoveh prented rindex
county countyname sumlev tract southern northern middle region;

run;

*testing to make sure areas are correct;
data nowtest; set now;
keep tract countyname rindex middle;
run;
proc sort data=now; by countyname; run;
proc freq data=now; tables middle; by countyname; run;
proc freq data=now; tables southern; by countyname; run;
proc freq data=now; tables northern; by countyname; run;
proc freq data=now; tables region; run;

proc means mean std n data=now; var p65 ppov pwohsedu plingiso phwophone
phwoveh prented rindex; run;
proc means mean std n data=now; var p65 ppov pwohsedu plingiso phwophone
phwoveh prented rindex; where southern=1; run;
proc means mean std n data=now; var p65 ppov pwohsedu plingiso phwophone
phwoveh prented rindex; where northern=1; run;
proc means mean std n data=now; var p65 ppov pwohsedu plingiso phwophone
phwoveh prented rindex; where middle=1; run;
```

```
*Tests for Equality;

*Risk;

title 'Histogram of Risk Index - Pooled Sample';
proc univariate data=now; var rindex; histogram;  run;
title 'Histogram of Risk Index - Southern Region';
proc univariate data=now; var rindex; histogram; where southern= 1; run;
title 'Histogram of Risk Index - Middle Region';
proc univariate data=now; var rindex; histogram; where Middle= 1; run;
title 'Histogram of Risk Index - Northern Region';
proc univariate data=now; var rindex; histogram; where Northern= 1; run;

proc means min max data=now; var rindex; run;
proc means min max data=now; var rindex; by countyname; run;
proc univariate   data=now; var rindex; where rindex>1.1; run;

data nownow; set now;
where rindex>1.1;
run;

proc freq data=nownow; tables countyname; run;

*Equality;
proc univariate data=now normal; var p65 ppov pwohsedu plingiso phwophone
phwoveh prented rindex; run;

*p65;
proc glm data=now; class region; model p65=region/solution;
means region/ hovtest=bartlett bon;
output out=p65out student=rstudent;
run;

proc univariate data=p65out normal plots; var rstudent; qqplot/normal(mu=est
sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var p65; run;

*ppov;
proc glm data=now; class region; model ppov=region/solution;
means region/ hovtest=bartlett bon;
output out=ppovout student=rstudent;
run;

proc univariate data=ppovout normal plots; var rstudent; qqplot/normal(mu=est
sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var ppov; run;


*pwohsedu;
proc glm data=now; class region; model pwohsedu=region/solution;
means region/ hovtest=bartlett bon;
output out=pwohseduout student=rstudent;
run;
```

```
proc univariate data=pwohseduout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var pwohsedu;
run;

*plingiso;
proc glm data=now; class region; model plingiso=region/solution;
means region/ hovtest=bartlett bon;
output out=plingisoout student=rstudent;
run;

proc univariate data=plingisoout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var plingiso;
run;

*phwophone;
proc glm data=now; class region; model phwophone=region/solution;
means region/ hovtest=bartlett bon;
output out=phwophoneout student=rstudent;
run;

proc univariate data=phwophoneout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var phwophone;
run;

*phwoveh;
proc glm data=now; class region; model phwoveh=region/solution;
means region/ hovtest=bartlett bon;
output out=phwovehout student=rstudent;
run;

proc univariate data=phwovehout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var phwoveh;
run;

*prented;
proc glm data=now; class region; model prented=region/solution;
means region/ hovtest=bartlett bon;
output out=prentedout student=rstudent;
run;

proc univariate data=prentedout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var prented;
run;

*rindex;
proc glm data=now; class region; model rindex=region/solution;
```

Matthew Martinez

```
means region/ hovtest=bartlett bon;
output out=rindexout student=rstudent;
run;

proc univariate data=rindexout normal plots; var rstudent;
qqplot/normal(mu=est sigma=est) square; run;

proc npar1way data=now wilcoxon plots=wilcoxon; class region; var rindex;
run;
```