

**Creating the  
Somali Data Center**

**A CHANCE Capstone Project**

**2011**

**Created by Kate Bernauer, Christian  
Larson, Charles Miles, and Mike Osberg**

## Contents

Executive Summary .....	3
I. Introduction .....	4
The Issue .....	4
Report Structure.....	4
II. The Somali Data Center: Project Summary .....	5
III. Somali Data Center Clients and Users.....	7
The Somali Justice Advocacy Center (SJAC) .....	7
Dahirjibreel, Executive Director .....	7
Local Advocacy Organizations.....	7
The Somali Community .....	7
The Cedar Riverside Neighborhood .....	8
IV. The Need for the Somali Data Center.....	9
V. Creating the Somali Data Center .....	10
Step 1: Developing theProject .....	10
Step 2: Building an Online Dashboard .....	12
VI. Somali Data Center: Details .....	18
Description of the Data Sources .....	18
Description of Indicators .....	20
Description of the Visualizations of Data Presented .....	21
VII. Sustainability of the Somali Data Center .....	25
VIII. Conclusion .....	27
Appendix I: Meeting Minutes .....	28

## Executive Summary

### The Product

The Somali Data Center is an online, interactive database consisting of information about the status of the Somali population in Minnesota. The goal of the Somali Data Center is to:

- Assist advocates in identifying and articulating the status of the Somali community in Minnesota for the purpose of soliciting funding and developing and evaluating programs and initiatives serving Somalis in Minnesota.
- Assist members of the Somali community seeking to gain knowledge about the demographics of the Somali population residing in Minnesota.

The SDC is hosted on the Somali Justice Advocacy Center website to allow easy access to the information from the website of a trusted organization from within the community. The SDC uses data from the American Community Survey (ACS) and the Minnesota Department of Education (MDE) to illustrate the following indicators of socioeconomic status:

- Age of the population,
- Enrollment in public schools of students whose primary home language is Somali,
- Family income levels
- Poverty levels, and
- Unemployment levels

Using data from the ACS and the MDE, a visualization (graph, table, chart, or map) and a brief overview for each indicator is presented to help explain what the data is showing. The visualizations and graphs are interactive and allow the user to explore the data in greater depth. Additionally, the raw data can be downloaded for further analysis.

### The Process

The SDC was created in the spring of 2011, as a part of a capstone project for the Cedar-Humphrey Action for Neighborhood Collaborative Engagement (CHANCE) course. A group of four graduate students at the Humphrey School of Public Affairs partnered with the Somali Justice Advocacy Center (SJAC) for five months to develop and launch the first comprehensive database regarding the Somali population in Minnesota.

### The Future

To ensure the sustained impact of this resource, the CHANCE group created the SDC in a way that makes it easy to update and replicate. To help this process, a user manual was created with pictures and step-by-step directions on how to update the current visualizations and graphs. Also, the data sources used were selected because they do not require technical skill beyond familiarity with Microsoft Excel for the compiling and developing of visualizations and graphs. This will allow the SJAC to continue to add new indicators which will help achieve the goal of the SDC to provide a holistic picture of the Somali community in Minnesota.

## I. Introduction

### The Issue

Minnesota is home to the largest population of Somali refugees in the United States. Over the past two decades more than 25,000 Somalis have moved into the state. In recognition of this large and still-growing refugee population, many Somali advocacy organizations as well as non-profit and private social service providers are trying to adapt to serve the needs of this community through the integration process. In order to advocate for and serve the Somali community, organizations need to understand the status of Somalis in Minnesota and how they are fairing economically, socially and in schools.

In the fall 2010 semester, Mr. Dahir Jibreel of the Somali Justice Advocacy Center (SJAC) spoke about the issues explained above to the Cedar-Humphrey Action for Neighborhood Collaborative Engagement (CHANCE) course offered by the Humphrey School of Public Affairs. Out of this conversation, a project idea was developed to create an accessible data resource offering a holistic portrait of the status of the Somali community in Minnesota. This would serve the needs of advocates in telling the story of this population and would also provide members of the Somali community a better understanding of the trends occurring within the population.

After a semester-long development of such a resource, the Somali Data Center was released to the public. The SDC is an online, interactive database that provides information on a number of different indicators through visualizations and graphs that will help provide a picture of the status of the Somali population within the state of Minnesota

### Report Structure

This report will first introduce the product of our capstone project, the Somali Data Center. After an introduction to the Somali Data Center, the report describes the clients and users for whom the resource was designed. Next, we outline our process for developing the data center, beginning with the origin and purpose of this project. We follow with a more detailed description of our resource. Finally, we will outline our vision for the future of this project and conclude with lessons learned from our capstone project experience.

## II. The Somali Data Center: Project Summary

Hosted on the website of the SJAC, the Somali Data Center is an online, interactive database of information about the Somali population in Minnesota. The goal of the Somali Data Center is to bring together and make accessible information about the Somali population from official data sources. Table I provides an overview of the sources and information that are included in the Somali Data Center.

The data from each source is presented on the SJAC website with a visualization or graph and a brief overview for each indicator. The visualizations are interactive and allow the user to explore the data in greater depth. Additionally, the raw data can be downloaded for anyone to analyze. Accompanying the Somali Data Center is a manual that provides step-by-step instructions to update the database. Each year, new information from both the Minnesota Department of Education (MDE) and American Community Survey (ACS) will be released, and this manual will allow the SJAC to update and expand the database, the website and the graphs.

In section 5 of this report, more information is provided about the details of the Somali Data Center. This section includes information about the limitations to the data and data sources included in the Somali Data Center.

Table I: Description of indicators used in the Somali Data Center

Source	Indicator	Years	Additional Info	Potential Purposes
MDE*, Primary Home Language	<i>Student Enrollment in K-12 Public Schools (district and county)</i>	1996 - 2009	Includes Somali speaking students as well as top 3 non-English speaking student groups.	Shows growth public school enrollment for students from Somali households (for each district) in MN. Can be used to understand the need for special youth services and programs.
ACS**	<i>Population by Age Groups (0-25, 26-45, 46-100)</i>	2001 - 2009	Includes those survey respondents who identified their ancestry as Somali.	Shows age distribution in the Somali population in MN relative to the total population. Can be used to determine targeted services for population due to relatively young age of Somali population.
ACS	<i>Family Income Levels (\$0-14,000, \$15,000-\$24,999, \$25,000-\$49,999, \$50,000-\$99,999, \$100,000+)</i>	2001 - 2009	Includes those survey respondents who identified their ancestry as Somali.	Shows family income levels for the Somali population in MN relative to total population. Can be used to understand need for special services for Somali population due to disproportionately low levels of income.
ACS	<i>Poverty (Percent below poverty line)</i>	2001-2009	Includes those survey respondents who identified their ancestry as Somali.	Shows poverty level within the Somali population in MN relative to total population. Can be used to understand need for special services for Somali population due to relatively high levels of poverty.
ACS	<i>Unemployment (Percent Unemployed)</i>	2001 - 2009	Includes those survey respondents who identified their ancestry as Somali.	Shows unemployment level within the Somali population in MN relative to total population. Can be used to understand need for special services for Somali population due to relatively low levels of employment.

\* MDE: Minnesota Department of Education

\*\*American Community Survey

### III. Somali Data Center Clients and Users

While the Somali Data Center is a public tool accessible to all Internet users, this online resource was created specifically to serve the vision of our community partner, the Somali Justice Advocacy Center. Based upon this vision, the Somali Data Center was developed for two key groups of users:

- Advocates for the Somali community in Minnesota
- Members of the Somali community in the US and abroad

This section introduces our community partner and the primary users for whom the Somali Data Center was created.

#### **The Somali Justice Advocacy Center (SJAC)**

The SJAC is a community organization based in the Cedar Riverside neighborhood of Minneapolis, MN. The SJAC works throughout the state of Minnesota with the mission to ensure fair treatment of Somalis under the law by advocating for, educating and empowering the Somali community to enjoy full civil rights and civil liberties. The organization was founded in 2001 and is currently led by Executive Director, Dahir Jibreel.

#### **Dahir Jibreel, Executive Director**

Prior to his work at the SJAC, Mr. Jibreel spent almost ten years as a high school social studies teacher and an elementary school teacher in Minnesota before returning to Somalia from 2004-2007 to serve as Chief of Staff to the President of Somalia. Today, in addition to being the Executive Director of the SJAC, Mr. Jibreel is an Executive Leadership Fellow for the Center of Integrative Leadership at the Humphrey School of Public Affairs.

#### **Local Advocacy Organizations**

The SJAC maintains close relationships with many Somali non-profit and service organizations in Minnesota. Nearly a dozen local leaders from public, non-profit and private organizations participated in the development of the Somali Data Center to ensure that it is a useful resource. They expressed their support of the Somali Data Center and enthusiasm for its usefulness for their organizations.

#### **The Somali Community**

The need for data pertaining to Somalis in Minnesota is due to the large influx of Somalis in Minnesota that began nearly two decades ago. Members of the Somali community are interested in accessing accurate data about the Somali community in Minnesota because so many Somalis now call Minnesota their home.

In order to understand the Somali Diaspora in Minnesota, it is necessary to know the circumstances that led to the recent migration of this population. In the late 1980's, a civil war erupted in Somalia as a consequence of the revolution to remove sitting President Siad Barre after 22 year in power. In 1992 the combination of civil conflict and drought led to a widespread famine in Somalia. The famine impacted the lives of most of the population and was a powerful force causing many Somalis to relocate. As a result of the civil conflict, an estimated 1.5 million Somalis have fled Somalia over the past twenty years.<sup>i</sup>The majority of displaced Somalis ended up in refugee camps in neighboring countries of Ethiopia, Kenya, Yemen, Djibouti and Tanzania. Many Somalis with the financial resources to do so have left the refugee camps and moved to Western countries such as the United Kingdom, Netherlands, Norway, Sweden, Finland, Denmark, Canada and the United States. Somalis also moved west through resettlement programs, often sponsored through faith-based initiatives.

Today, nearly one-third of all Somali-Americans live in Minnesota<sup>ii</sup>. Somalis have settled in Minnesota for a number of reasons. Minnesota has traditionally had a strong job market, which has made it a destination for immigrant and refugee populations throughout the state's history. Historically, Minnesota has also had strong social service programs (particularly in terms of access to health care and good educational opportunities) and is home to a number of the faith-based organizations that are often involved in refugee resettlement programs. After the initial establishment of the Somali population in the mid and late 1990s, familial networks have sustained the in-migration of Somalis to the state.

### **The Cedar Riverside Neighborhood**

As mentioned in the introduction, the Somali Data Center was developed through the partnership between the SJAC and the CHANCE capstone course. The goal of the CHANCE course is to build and strengthen the relationship between the Cedar Riverside neighborhood and the University of Minnesota. This project achieves this due to the large Somali population within the Cedar Riverside neighborhood.

The Cedar Riverside neighborhood was one of the main resettlement locations for Somalis in the United States. Due to the resettlement and the continued in-migration of Somalis occurring after the original resettlement, the Cedar Riverside neighborhood has one of the densest Somali populations in the world (outside of Somalia). After the Somalis became established in the neighborhood, the Somali community started a number of businesses and non-profit organizations that provide goods and services to the Somali population.

## IV. The Need for the Somali Data Center

In the fall semester of 2010, Dahir Jibreel, Executive Director of the SJAC, was a guest speaker in the CHANCE capstone class where he provided insight into the issues of the East African population within the Cedar Riverside neighborhood. During this discussion, Mr. Jibreel expressed the need for information pertaining to the socioeconomic status of the Somali community in Minnesota in order to support the advocacy of the SJAC on behalf of the Somali community.

In response to the issues raised by Mr. Jibreel, a project was proposed to create an online resource that would help provide a picture of the status of the Somali population in Minnesota. The rationale of the project was that compiled data based on a variety of indicators (such as economic status, labor force participation and educational attainment) would offer a holistic portrait of the Somali community. The online data resource was designed for two primary purposes:

- To assist advocates in identifying and articulating the status of the Somali community in Minnesota for the purpose of soliciting funding and developing and evaluating programs and initiatives serving Somalis in Minnesota.
- To assist members of the Somali community seeking to gain knowledge about the demographics of the Somali population residing in Minnesota.

## V. Creating the Somali Data Center

After establishing our project scope and our partnership with the SJAC, our group started the process of building a tool to provide data on the Somali community. To guide this process, we developed a set of criteria that would provide direction for this project throughout the spring semester. We determined that our final product for the SJAC should:

- enhance the advocacy mission of the SJAC,
- be useful and easily accessible to the Somali community,
- be a sustainable product that could be easily updated and expanded upon

As agreed upon with the SJAC, our deliverable would be an online resource that would provide easy access to information regarding the Somali population residing in Minnesota. In building this resource, we engaged in many conversations that led to new insights and required us to make decisions and adjustments to produce the most valuable and useful tool possible for our community partner. This section of our paper explains how key conversations re-shaped our final product into the Somali Data Center described above. Throughout this section, we have cited meetings that led to decisions about the project. Detailed information about these meetings is included in Appendix I.

### Step I: Developing the Project

#### ***Concerns about Census Data and the Proposition of an Alternative Count***

As we began to envision the database, our group first considered using official Census data as the primary source for collecting information. We brought this idea to the SJAC and they shared with us their concerns regarding the accuracy of Census data (Meeting 12/21/2010). They believed the Somali population of Minnesota was undercounted by at least 25,000 people. They also explained that this undercount was a concern shared by many Somalis across the state and that high-ranking officials within the State Department had acknowledged the possibility of an undercount<sup>iii</sup>. The belief that the Somalis within Minnesota were undercounted led Mr. Jibreel and the SJAC to suggest that we attempt to perform an alternative population count using familial networks within the Somali community.

In response to the SJAC request to perform a count of the Somali population in Minnesota, our group performed a literature review of non-traditional methods for counting hard-to-reach populations. We also spoke with experts in the field of community-based-research methods (Meeting 1/26/2011). From this research we found several alternative methodologies, including: snowballing methods, respondent-driven sampling methods, and chain-referral methods<sup>iv, v</sup>. We also learned that many of these methods are disputed as a result of the high level of uncertainty involved with the estimates that they produce. Additionally, these methods have not been used on the scale for which we would be attempting, the entire Somali population in Minnesota.

Following the extensive research into alternative population counting methods, our group determined that trying to conduct an alternative count using one of these methods would a) not likely produce credible data and b) be unfeasible given the scope of our project.

### ***Rationale for Using Official Government Data Sources***

1. *Official government data sources are commonly accepted and data from these sources will better serve advocates in telling credible stories.*

Using official government sources of data better served the predetermined criteria we defined for the data center because they are commonly used and accepted as trusted data sources. When the SJAC and other service and non-profit organizations within the Somali community are using this data to support their advocacy, they can be confident that policy makers and funders will view the data as legitimate. By using a trusted source of data, the information becomes much more powerful.

2. *Official government data sources offer a wide variety of indicators to examine and therefore have the potential to offer a holistic portrait of the Somali population.*

By using official government sources of data, we are able to provide a more holistic picture of the status of Somalis in Minnesota than if we had pursued an alternative count methodology (which would have likely produced demographic and population data only). Official government data sources offer a wide variety of indicators, including socioeconomic indicators that would be more difficult to attain using an alternative count method.

3. *Official government data sources provide baseline information for comparing indicators across different population and indicating disparities.*

Because official government data sources employ large-scale collections of data, these sources also provide baselines for comparison (for example, examining economic indicators across subpopulations). This comparability is useful for telling a story and illustrating existing disparities.

4. *Official government data sources are accessible and frequently updated which makes them ideal for sustaining a useful database in the long term.*

Official government data sources are accessible and frequently updated, offering the best prospect for the sustainability of our project. Because our online data resource will be most accurate and useful if it is updated annually, it is critical that our work be replicable. By documenting our data collection process in a manual for updating the online resource, our community partner will be able to keep the database current. Had we performed an alternative count, in contrast, our project would have been a resource-intensive single collection of data.

## Step 2: Building an Online Dashboard

After deciding to build the online database using official sources of data, there were a number of decisions that were necessary to make. We needed to:

1. Choose the indicators that would provide a holistic picture of the status of the Somali community in Minnesota.
2. Choose the specific sources of official government data that would allow us access to those indicators.
3. Decide what platform to use as the interface for our dashboard.

### **Choosing Indicators**

We began choosing indicators by looking at other tools and status reports of subpopulations within the United States. This research allowed us to see what types of indicators that other advocacy organizations and researchers were using, how their data was displayed, and what data sources were used in projects with similar goals.

We reviewed thirteen online databases and status reports<sup>vi</sup>. After reviewing these reports we tabulated the frequency of each indicator's inclusion across the thirteen reports. The results are shown in Figure 2. The review of the online databases and status reports also reinforced our decision to use official government data, because the main data sources used for most of the reports were Decennial Census data and American Community Survey data.

We also spoke with experts to receive feedback on which indicators would be most useful for SJAC advocacy. These experts included Interim Dean Greg Lindsey and Assistant Professor Ryan Allen of the Humphrey School of Public Affairs. Dr. Allen has special expertise in immigrant and refugee policy.

From these conversations we determined that we would need to prioritize indicators and include them in our online resource in a way that was feasible for the scope of our project (Meeting 3/5/2011). Given time constraints, we were limited in the number of indicators that could be included on the dashboard. With this in mind, our conversations with Dean Lindsey and Dr. Allen focused on two questions:

- What will give advocates the most leverage in their work?
- How far can the data be disaggregated within the subpopulation?

From the list of indicators compiled in our review of similar tools and products and through our conversations with the experts at the Humphrey School, we compiled a list of categories of indicators that would be important for advocates. Our list included demographic, economic, education and justice system categories. From this list we decided that the most useful indicators for the SJAC would be demographic and economic indicators, with an understanding that the other indicators could be added in the future to help complete the holistic picture that we were attempting to show.

Table 2: Indicators used in review of online dashboards and status reports

<b>Category</b>	<b>Prevalence</b>	<b>Category</b>	<b>Prevalence</b>
<i>Use of Social Safety Net Programs</i>	1	<i>Home Ownership</i>	7
<i>Transportation</i>	2	<i>Marital Status/Family Data</i>	7
<i>Year of Entry/ Residence 1 year Ago/ Foreign Born</i>	4	<i>Age/ Life Expectancy/ Mortality</i>	8
<i>Fertility</i>	5	<i>Educational Attainment/ Educational Enrollment</i>	9
<i>Language Spoken at Home</i>	5	<i>Income/Poverty</i>	9
<i>Employment Status</i>	6	<i>Population</i>	9
<i>Health/Disability/Abuse</i>	6	<i>Race/Ancestry</i>	10
<i>Sex</i>	6		

Before proposing indicators to the SJAC regarding the indicators, we needed to provide specific indicators within our two main categories. In making our list of more specific indicators, we had to make several decisions:

- From our research of other similar products, we discovered that age distributions of migrant populations was often times used as these populations tend to be younger than the general population, so we adopted this as one of our indicators.
- We also thought that major population centers and growth rates within Minnesota would be a valuable tool for the SJAC, so we proposed that growth rates in the Somali community be included in the development of this resource.
- In examining economic indicators, household income levels, poverty levels, and unemployment rates appeared to be most valuable indicators. When comparing the Somali population to the total population, these indicators would show important disparities in the economic well-being of these populations.

After compiling our list of indicators, we consulted with the SJAC who explained that all of these indicators would be useful for his purposes and for the Somali community (Meeting 3/23/2011). He warned us about using household income levels for the Somali population due to the large informal economic sector within the Somali community. He explained that the informal sector can lead to an under estimation of the income levels of Somalis. In the end, we

decided to use household income levels as it is a measure of access to the formal economic sector while noting actual income levels may not be represented in the data.

### ***Choosing Specific Data Sources***

After agreeing with the SJAC on a list of indicators, we then needed to determine the specific sources of data that would be used for building our online dashboard. We again made this determination using the criteria put forth at the beginning of our project. In particular, ACS data provide a number of demographic and socioeconomic indicators, which would allow us to create a holistic portrait of the Somali population in Minnesota. Additionally, we decided that MDE: Primary Home Language data would also provide useful data for advocacy purposes, because it can be used as a proxy for the size of the population, growth rates within the population and can provide a picture into the geographic distribution of the Somali Population. Also, these data are accessible and comes out yearly, so the database can be updated to provide current and accurate information. Finally, ACS data and MDE data provide the type of information we needed for the indicators we selected.

### ***Working with a Small Sample Size***

After we determined what specific data sources to use, we set out to determine how best to present this data. For the MDE: Primary Home Language data, this was fairly straightforward, since the data was an actual count based on student identification of primary home language. This data source did not make easily accessible any other indicators that would be useful for our database. ACS data, on the other hand, allows the data to be broken down by a number of demographic and socioeconomic indicators. This wide variety of available indicators required more deliberation from the group.

First, we chose the most useful ways to disaggregate each indicator. Our initial conversations led us to believe that we should break each indicator down by sex, age and location (metro/out-state), and that we should compare the results to that of the total population in Minnesota (Meeting 3/11/2011). We believed that with this breakdown, we would be able to provide a holistic picture of the status of the Somali population by additionally providing data on subpopulations within the Somali community.

After pulling the ACS data for subsections of the Somali population in Minnesota, we discovered that the sample size of the Somali community was too small to provide useful data; the confidence intervals were wide and therefore the data could not provide conclusive representations of the status of the population. Even when using the entire Somali population in Minnesota (rather than breaking the population into subpopulations based on sex and age, for example), some confidence intervals reached levels of 50 percentage points for some indicators. The discovery of the high confidence intervals forced us to adapt our vision of the database. After discussing our concerns about confidence intervals with experts at the Humphrey School of Public Affairs and the Minnesota Population Center, we concluded that for the single year samples at least, we would not be able to break down the data beyond the total Somali population as we had hoped (Meetings 3/11/2011 & 3/30/2011).

In response to the issue of large confidence intervals, we were led to make a second decision regarding the ACS data. We needed to determine if a pooled sample of ACS data was more

practical for our purposes, or if we should continue to focus on the single year samples. Both the single-year and pooled samples had advantages and disadvantages.

The single-year data allowed for the ability to see changes over time within the Somali community. As mentioned above, the major disadvantage was the small sample size and large confidence intervals, which eliminates the opportunity to break down the Somali population into subpopulations. Using a pooled sample could solve the confidence interval issue, but by pooling the data, the ability to track changes over time is eliminated. We tried using the three year pooled samples of the ACS data available on MPC's IPUMS-USA database, but found that confidence intervals were not greatly reduced and, because of the late release of 2007-2009 pooled sample, using 3-year pooled samples would only allow us two data points for expressing trends, which is not effective. We then began considering the creation of a 5-year pooled sample from 2005 to 2009. Using a 5-year pooled sample posed an issue with data collection, as IPUMS has not yet released the 5-year pooled samples. This meant that we would need to pool the data and create the confidence intervals ourselves.

Originally, the group decided that we would use both the single year and the 5-year samples. We thought this would provide a holistic picture and be very useful for advocacy purposes. Changes could be tracked over time for the indicators that we had selected for the entire Somali population using single year samples, and we could also provide a "snapshot" of the condition of different subpopulations within the Somali community by using the 5-year pooled sample for 2005-2009.

What we discovered in trying to pool the 5-year data was that it is a lengthy process requiring extensive technical skills with statistical software, especially in the creation of confidence intervals for pooled samples<sup>vii</sup>. As a result of the time-intensive nature of this process and uncertainty as to whether or not the team was producing accurate confidence intervals, we decided to drop the 5-year sample and focus on building our database using the single-year data. This decision was ultimately made based on time constraints, although another factor was that the complicated process of creating the 5-year pooled sample would not be as easy to replicate for a non-expert. Making this decision easier was the fact that IPUMS was set to release the 5-year pooled sample in May, 2011. Once IPUMS releases the 5-year pooled sample, the "snapshot" we originally envisioned can be much more easily produced by following the directions in the attached manual (Meeting 4/20/2011).

### ***Deciding how to Present the Data***

As our group was collecting the ACS and MDE data, we were also developing ideas for how to present the data. This process began by reviewing interactive online databases. The dashboards we found most useful were easy to navigate, provided clear visual representations of the data, offered accessible interpretations of the data, and gave users easy access to the original sources of the data.

From our initial research and discussions with the SJAC we developed a "wish list" of things that we wanted the dashboard to be. This list included:

- User-friendly data visualization software

- Interactive interface for users to explore
- Hosted on the SJAC website

At the most basic level, one option was to use Microsoft Excel to put together tables, graphs and charts that the SJAC could upload onto its website. A limitation of this option was that it would significantly limit our ability to create interactive visualizations. We did not believe such a tool would be as useful as an interactive and visually appealing dashboard. Due to these limitations we continued to explore other options. We researched the idea of partnering with other organizations that had the capacity to build such a website or partnering with a web developer to build the tool. From our research and conversation with CURA, for the purposes of our project, we decided that these options came with high costs and would take too much time (Meeting 1/20/2011).

In our continued search for data visualization software, our group came across Tableau Public. This software allows for the building of interactive online dashboards and provided us with all of the tools we needed to build the Somali Data Center. Using Tableau Public proved useful for our purposes because it was user-friendly and allowed us to quickly embed our data in online, interactive dashboards. Because the software is accessible and comes with several video tutorials, a manual documenting our process of uploading data and creating visualizations in Tableau would be sufficient for assisting future updates to the Somali Data Center by the SJAC. Perhaps most importantly, Tableau Public is a free software.

It should be noted that while working with this software, we have discovered a few limitations to Tableau Public:

- It is not compatible with Mac computers, so it can only be accessed on a PC.
- Occasional glitches causing visualizations to stretch slightly on the screen.
- Infrequent but potentially inconvenient Tableau Public updating cycles, causing the visualizations to be inaccessible temporarily while updates are being made to Tableau.

### ***Feedback from the Community***

After we discovered Tableau Public and collected all of our data, we began to design the dashboards. We initially built the online resource on a blog, so that we had easy access to the development of the tool. We built this prototype of the Somali Data Center as a way to gather feedback on the usefulness of the data provided, the readability of the visualizations and the navigatability of the dashboards.

To gain an understanding of how local advocates would use the tool, we held a feedback session with leaders from the Somali Community. This meeting was held in a computer lab in the Humphrey School in order for each person to have access to a computer for navigating the Somali Data Center. The SJAC requested that the feedback session be hosted at the Humphrey School as it was symbolic of the new relationship that was being built between the Somali community and the Humphrey School. The community leaders that attended the feedback session included professionals in the private, non-profit, and public sectors. The SJAC selected

and invited the group of professionals to join us at the Humphrey School of Public Affairs to test the new tool before adding the final product to the SJAC website.

The response to the Somali Data Center at the feedback session was very positive. The group of community leaders expressed enthusiasm for the prospect of having increased access to data about the Somali community. Additionally, the group indicated their belief that the data center will be widely used, and an asset to advocates for the Somali community. Feedback was also given about additional indicators and sources of data that would be useful to include on the data center in the future. Most importantly, our team gained feedback on what information was unclear on the data center, such as the names of some of the variables and the displayed confidence intervals.

The group also voiced the concern that Mr. Jibreel had mentioned earlier in our process – that levels of poverty and unemployment may be overestimated and income levels underestimated due to Somali participation in the informal economic sector; income made in the informal sector may not be reported when filling out a survey for the government. The suggestions for future indicators included: access to healthcare, housing data, educational attainment and outcomes, and justice system statistics. Although these indicators were not included in the initial release of the Somali Data Center, the user manual provides direction for the future development of the tool (Meeting 4/16/2011).

### ***Going live***

Following the feedback session, our group updated the visualizations as requested in the feedback session and documented the suggestions of the group so that they will be a top priority as further development of the tool occurs. Along with the updates to the SDC and the documenting of future priorities, the group was in contact with the SJAC website administrator to ensure that the data center was ready and correctly formatted to provide a quality experience for those who utilize the tool upon its release.

The website was released to the public at a presentation at the Humphrey Center. The SJAC and the CHANCE team hosted a group of individuals from Somali businesses and non-profit organizations, foundations and service providers, academic representatives, and community members who have interest in using the SDC.

## VI. Somali Data Center: Details

As the Somali Data Center was created with different types of data and has been presented with different types of visualizations, this section presents additional details about the data sources and the visualizations. Descriptions of each of the data sources and the indicators that were used for the Somali Data Center are included below. Also included are the limitations of the data.

### Description of the Data Sources

The two sources of data used for the Somali Data Center are the Minnesota Department of Education: Primary Home Language (MDE) and the American Community Survey (ACS) (accessed through the IPUMS-USA Online Data Analysis tool). The MDE data is a direct count of the students who are in public schools in grades K-12 in Minnesota. The ACS, however, is a survey of only a sample of people in Minnesota. As a result, the numbers that are reported from the ACS are only estimates of the true number for the state.

#### Minnesota Department of Education

The Minnesota Department of Education hosts a data center for information about the students and teachers in the state. For the Somali Data Center, we used data on the primary home language spoken by students who are enrolled in public schools in Minnesota. Each year in October, information is gathered about what the home language is of the students in public school.

#### *Student Enrollment*

For the Somali Data Center, we collected information from the years 1996 - 2010 on the number of students whose home primary language was Somali. The information that resulted is a count of all students in Minnesota whose home primary language was Somali, as well as the school district and county that they were enrolled in.

#### *Advantages:*

- **Geographic Distribution Changes:** Provides information over time of where Somali populations are located around the state. This includes locations outside of the metro area where small populations of Somalis are residing.
- **Provides Enrollment Count:** The MDE data provides a count of all students in the state, not just a sample. Because of this, the data provided here is more reliable than sample data, as errors will only result from mistakes in data collection, not from statistical sampling standard errors.

#### *Disadvantages:*

- **Lacking Education Outcomes:** As the information that MDE provides about Somali speakers is only a count, nothing can be said from this data about the educational attainment and outcomes of the Somali population, such as graduation or dropout rates.
- **No Higher Education Information:** The MDE data does not include any information about students who went on to higher education in the state.

### American Community Survey

The American Community Survey provides information about the American population that previously was only available through the U.S. Census long form survey. Since 2000, the ACS has been conducted as a yearly survey of population centers greater than 65,000 in the U.S. Additionally, starting in 2006, the ACS began surveying smaller geographic areas, to enable more detailed statistical reporting. The surveys of smaller geographic areas are released as 3-year and 5-year pooled data sets, with the first 5-year pooled data set released in 2010<sup>viii</sup>. We have included only single year data sets in the Somali Data Center.

Since the ACS is a survey of a sample population in the U.S., there are important considerations when using and interpreting the data. For each indicator the numbers that are reported are estimates of the true value of each indicator. Each estimate is reported with a confidence interval of 95%, which demonstrates the variability and statistical uncertainty of the estimates of each indicator. A confidence interval provides a range of values from which it can be assumed that 95% of the time the true value for which you are looking lies within that range.

Because of this uncertainty about the estimates, it is suggested that the ACS data is used to measure trends over time about the various characteristics of the Somali population, rather than focus on the specific counts.<sup>ix</sup> See the section below for a more in-depth discussion of confidence intervals, with examples from the Somali Data Center.

The Somali Data Center includes information from the ACS about the Somali population in the state such as: the amount of Somalis who are unemployed and below the poverty line, the family income level, and the population distribution among different ages. As with all data sources the ACS has advantages and disadvantages including the following:

#### *Advantages:*

- **Multiple Indicators:** Allows us to gather information about a number of characteristics about specific populations. Provides over 300 variables to create a holistic picture.
- **Updated Yearly:** As a data source that is updated yearly, the ACS provides the most up-to-date information and changes in a population over time.
- **Ethnicity Information:** Included in each survey are questions about the ethnicity of the respondent. This is more detailed information than is provided on the Census and provides data for over 600 specific ethnic populations in the country.

#### *Disadvantages:*

- **Statistical Uncertainty:** Because the ACS data is taken from a sample of the population, the count estimates have a degree of uncertainty. This makes it difficult to draw inferences about small populations.

## Description of Indicators

### *Population by Age Groups<sup>x</sup>*

In order to get a better understanding of the population distribution of Somalis in Minnesota, the Somali Data Center reports 3 different age groups: 0-25, 26-45, 46-100. Ages in the ACS are collected both by asking the respondent their age at the time of the survey as well as their date of birth.

### *Family Income<sup>xi</sup>*

For the Somali Data Center, incomes are reported in 5 income groups: \$0-14,999, \$15,000-\$24,999, \$25,000-\$49,999, \$50,000-\$99,999, \$100,000+. The information from family income is gathered by asking a series of questions about the potential types of income earned by the respondent. The total of all reported income is then included as the total family income. The questions asked include information regarding:

- Wages, salary, commissions, bonuses, or tips from all jobs
- Self-employment income from own non-farm businesses or farm businesses
- Including proprietorships and partnerships
- Interest, dividends, net rental income, royalty income, or income from estates and trusts
- Social Security or Railroad Retirement,
- Supplemental Security Income (SSI)
- Any public assistance or welfare payments from the state or local welfare office
- Retirement, survivor, or disability pensions. *Does NOT include Social Security*
- Any such other sources of income received regularly such as Veterans' (VA) payments, unemployment compensation, child support or alimony

(Note: Based on information gathered from discussions with community members and our client, it is suggested that these reported numbers are lower than what is the reality in the Somali population.)

### *Poverty<sup>xii</sup>*

Reported in the Somali Data Center is the percentage of Somalis that are currently above and below the poverty threshold. The ACS reports the income of each respondent as a percentage of the poverty threshold from the year prior to the administering of the survey. Thus, a number of 100 is equal to the poverty threshold, where anything below 100 was below poverty and anything above 100 was above poverty. The Somali Data Center reports the percentage of the population who were 'below the poverty threshold'.

### *Unemployment<sup>xiii</sup>*

The Somali Data Center displays information about the percentage of the Somali population that is reported as unemployed. The ACS asks a series of questions that are meant to categorize respondents as either employed, unemployed, not in the labor force or N/A. For the Somali Data Center, unemployment was calculated as the number of respondents who reported as unemployed divided by the total number of respondents who reported as employed or unemployed.

(Note: Some people who are not in the labor force may have stopped seeking employment. For this reason, some experts believe that reported unemployment rates may be lower than reported unemployment rates.)

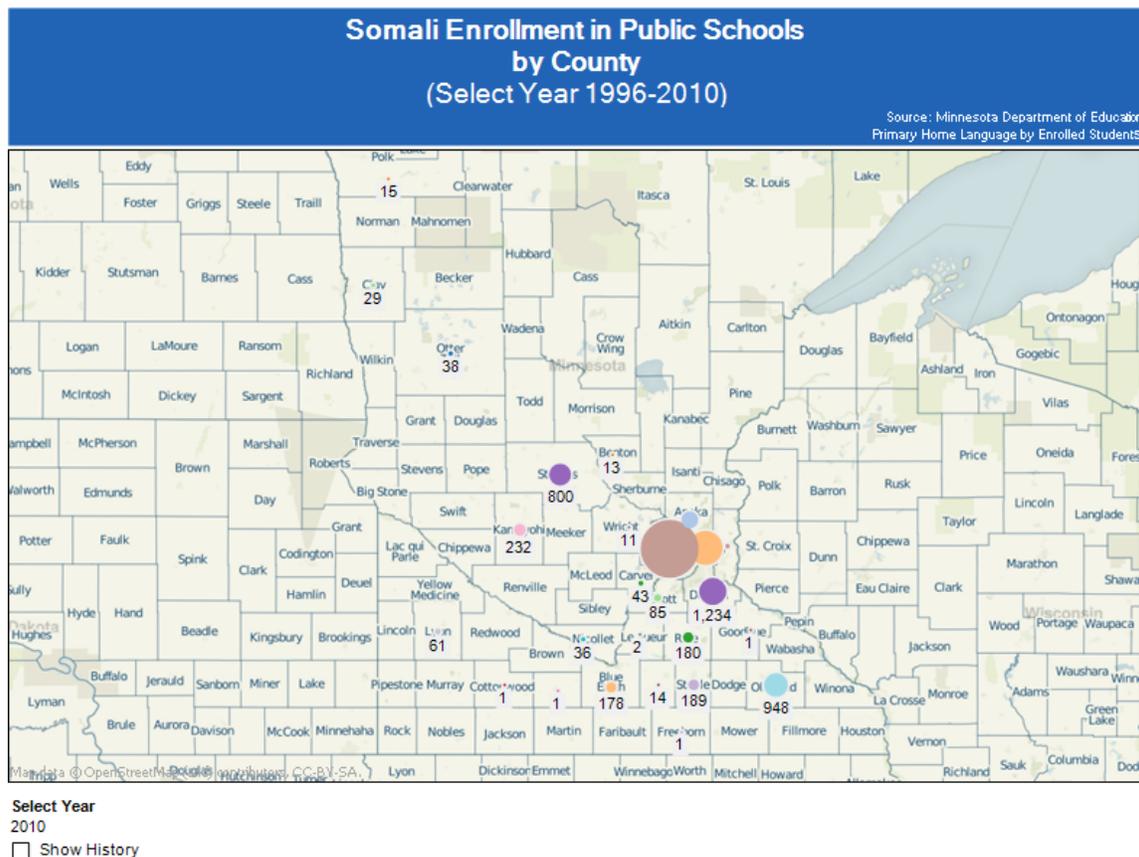
### Description of the Visualizations of Data Presented

Within the Somali Data Center, each of the previously mentioned indicators is presented using various visualization types. Included below are descriptions of each visualization type, and information on reading each visualization.

#### Maps

In order to demonstrate the spatial distribution of the Somali population, a map of the location of student enrollment in Minnesota is included. The user can move from year to year on the map, and the display of the location and size of Somali populations (measured by enrollment) will change on the map. Each bubble on the map grows according to the number of students in the respective county that are enrolled in K-12 and speak Somali as their primary home language.

Figure 3: Sample data visualization: Enrollment in Public Schools of Students whose Primary Home Language is Somali, 2010, map



#### Time Series

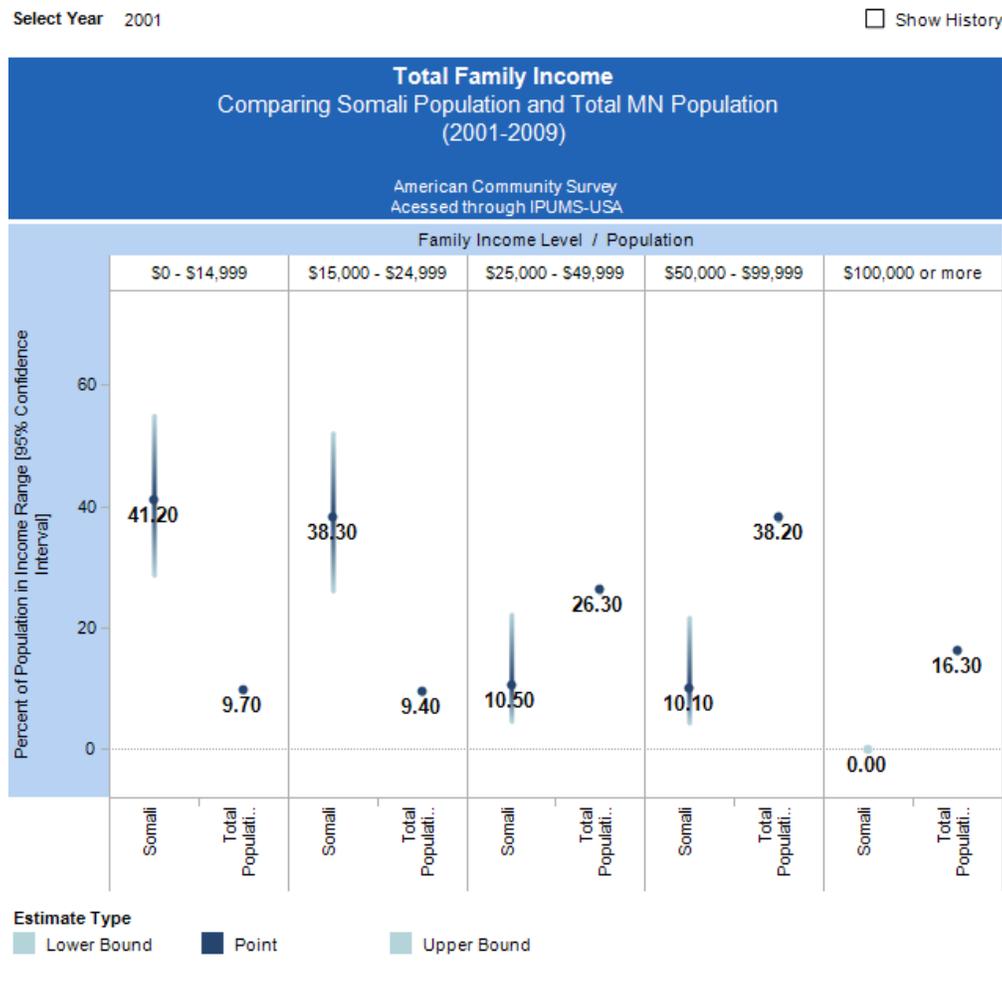


There are multiple visualizations that display 1 year of an estimate that can be navigated from year to year to show changes in the indicator over time. The graphs for poverty, unemployment and total family income are displayed in this way. These graphs show the point estimate for each category, with a vertical bar that displays above and below the point estimate. This vertical bar demonstrates the confidence interval (see Figure 4).

*Confidence Interval*

For the visualizations of the point estimates, a 95% confidence interval is displayed. The confidence interval is a measure of the uncertainty of the statistical estimate. For example in the visualization of family income shown below in Figure 3, in 2001, the Census Bureau estimated that 41.2% of Somalis were between \$0-\$14,999 and that with 95% confidence this estimate was between 28.7% (the low point on the vertical bar) and 55.0% (the high point on the vertical bar). One way to interpret this uncertainty is that 95 times out of one hundred, an estimate of families in the Somali population with income \$0-\$14,999 in 2001 would fall within the bounds of 28.7% and 55.0%.

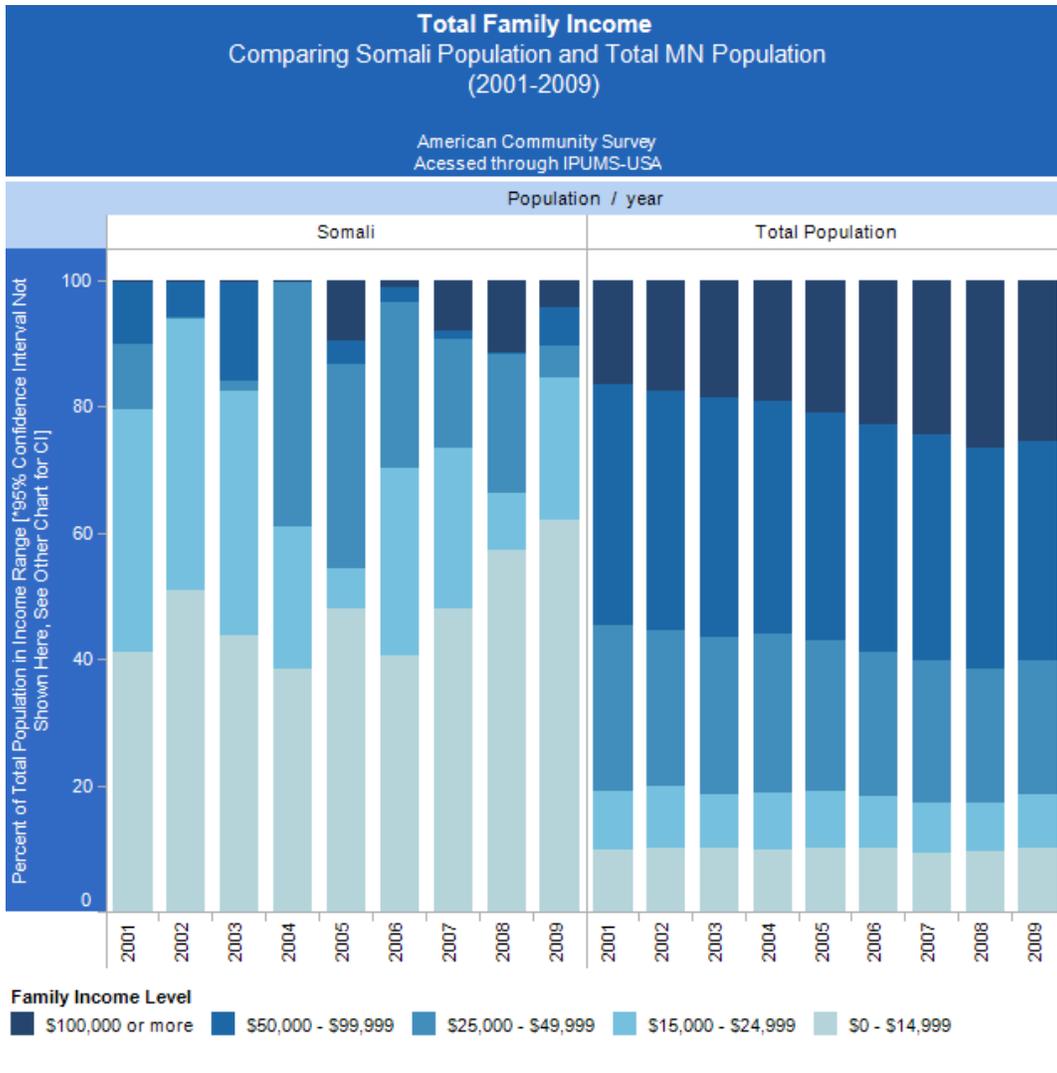
Figure 4: Sample data visualization: Total Famil Income, 2001, showing confidenceintervals



**Bar Charts**

The bar charts are used to display the distribution of the population in given categories. This is shown for both the family income and age group distributions. Each color of the bar represents a different age group, and the size of the colored bars represents how large each age group’s population is (see Figure 4).

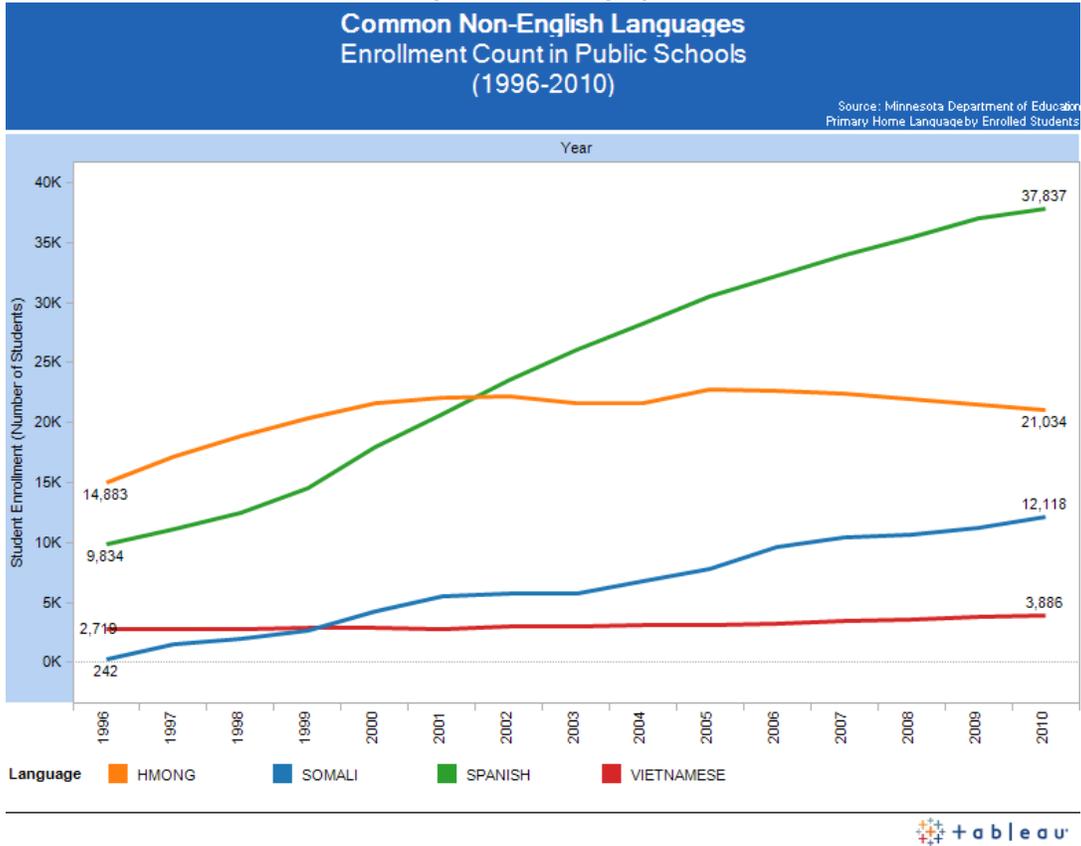
Figure 5: Sample data visualization: Total Family Income, Years 2001 through 2009, bar chart



**Line Graphs**

Line graphs are utilized in the SDC to show changes over time in data that does not require a confidence interval. For the current purposes of the SDC, this includes visualizations of the MDE, Primary Home Language data.

Figure 6: Sample data visualization: Common Non-English Languages in Minnesota Public Schools, Years 1996 through 2010, line graph



## VII. Sustainability of the Somali Data Center

From the beginning of the project, our capstone team aimed to create a product that would be a sustainable resource that will have a continued impact for years after the project was implemented. While it was and still is a goal of our project to have the data center be a tool that is updated and used year-to-year, we recognize the difficulty of creating a sustainable capstone project of this technical nature. For that reason, we have created a manual for updating the data center that provides step-by-step instructions for how to collect the data and update the dashboards to reflect new survey years.

It should be noted that in addition to following the updating instructions, the SJAC staff member, volunteer, intern or research assistant (RA) performing the updates will gain experience in data collection and with Tableau data visualization software. From this experience the person performing the updates could also create their own visualizations for the Somali population in Minnesota, making the project an endeavor that has the potential to build individual as well as organizational capacity. We outline our next steps for the continuation of the Somali Data Center below.

First, the partnership that has been established between SJAC and CHANCE faculty will be key to ensuring the longevity of the project. Not only will the continued support of and collaboration with CHANCE make the project sustainable, but SJAC's willingness to include CHANCE in the continued maintenance of the Somali Data Center will contribute equally to ensuring the project's sustained impact. We suggest that SJAC meet with CHANCE before the beginning of each fall and spring semester to garner support and feedback from CHANCE for the Somali Data Center, as well as to find out which students are interested in working on the continuation of this project. Similarly, CHANCE faculty should make students aware of the opportunity of being involved in the project, specifically mentioning the following key points in the experience the student will gain while working on the Somali Data Center:

- **Updating the site with the easy-to-use manual will provide professional experience in understanding database queries that target small subpopulations in a large population survey.** This experience is important for those interested in policy analysis, population studies, or any other policy field that is interested in quantitative methods complimenting their public policy focus.
- **Data entry in the Tableau Excel spreadsheets we've created to update the Somali Data Center website.** The data entry will be a part of the transfer and load process. The experience will provide a query that will be downloaded onto the computer, sorted by the user, and uploaded onto the Internet. Understanding this product cycle is an essential part of knowing how the construction of online dashboards using large data sets is accomplished.
- **Design and creation of new innovative dashboards that will provide important new information to the community served by SJAC.** These dashboards, whether housing, health care, criminal justice, or school attainment, will be

important indicators for the Somali community and will allow the researcher to carry out a self-designed project that could be personally used for his or her career portfolio.

Lastly, as was mentioned above, the 5 year pooled sample is being uploaded by IPUMS in May of 2011. Comparisons on income with cross-cutting variables like age and sex will be possible with the release of this data, giving new detailed information on the status of Somalis in Minnesota. With the release of this data, CHANCE and SJAC could partner with CURA to update the site with these new variables, since it will create a considerable amount of new detailed dashboards that will be useful to advocates for the Somali community. The ability of CHANCE and SJAC to approach new data releases from the IPUMS-USA website and successfully update the site with an intern or research assistant is the most important factor to ensuring the sustainability of the Somali Data Center.

## VIII. Conclusion

The development of the Somali Data Center proved to be a challenging and informative experience. Our client's strong vision for the project was an asset because the generous contributions of the SJAC informed and legitimized our data center and ensured its usefulness. The SJAC's vision also presented a challenge because we had high ambitions for the data resource but limited knowledge about how to design it, often leading to project scope creep (as illustrated by our consideration of conducting an alternative count of the Somali population in the site rather than using readily available data resources).

The regular communication we had with SJAC and capstone advisors, and with one another as a team, made certain that scope creep did not take a significant toll on the development of our project. But this challenge provided all of us with useful experience in understanding how to best generate solutions that are feasible in size and scope while communicating with stakeholders our goals during every step of the process.

Other key challenges in working on this project were technical in nature. The data on the Somali population in the IPUMS-USA data center are a relatively small sample (n=some number to n=some number; sample size ranges) year-to-year in the state of Minnesota. The result was a narrowing of what dependent variables we could consider given the small sample and, consequently, large confidence intervals with which we were working. Additionally, due to the time-intensive nature of compiling a multi-year comparative analysis, we were not able to include a 5-year pooled sample of Somalis in Minnesota surveyed by the ACS from 2005-2009.

Additionally, project was challenging because we sought to create a product that would be easily accessible and understandable by members of the community with less technical backgrounds. For users less familiar with statistical concepts such as weights and confidence intervals, fully understanding and engaging with the data center may still be a challenge. Also, users with limited experience working in programs such as Microsoft Excel may find updating the data center difficult.

Despite these challenges, from a great relationship with our client, faculty support, and a school-year-long effort to provide a meaningful product for SJAC and the Cedar-Riverside community, we have completed a project that will provide the community with an interactive online database of easily accessible and downloadable information on the demographic and economic status of Somalis in Minnesota. Our hope is that the project will continue in the future since our client and the community leaders we have met with in the Cedar Riverside neighborhood have all expressed their excitement and commitment to the Somali Data Center.

## Appendix I: Meeting Minutes

While our group did engage in all of the activities outlined in our original project scope, conversations with core consultants (listed below) throughout the development and implementation of this project often led to new insights, causing the team to adjust project activities in order to produce the most valuable and useful deliverable possible for our community partner. This appendix will outline the key conversations that re-shaped our final product.

### Core Consultants:

- Dahir Jibreel, Executive Director of the Somali Justice Advocacy Center
- Greg Lindsey, Interim Dean of Humphrey School of Public Affairs,
- Merrie Bennisutti, Coordinator of Community Partnerships, Humphrey School of Public Affairs
- Ryan Allen, Assistant Professor, Humphrey School of Public Affairs
- Brandon Trampe, Graduate Assistant at the Minnesota Population Center

### 12.21.2010

*Meeting Consultant:* Dahir Jibreel

*Meeting Purpose:* To solidify our team's understanding of Mr. Jibreel's vision for this project and to discuss first steps for moving forward.

*Meeting Outcomes:* Our team learned that Mr. Jibreel, in addition to his interest in the creation of an online database, was interested in identifying and employing alternative methods for collecting data on the Somali population. In particular Mr. Jibreel suggested that using a method of surveying Somali familial networks within Minnesota would provide a more accurate and comprehensive dataset than the one produced by the compilation of official government data.

In response to Mr. Jibreel's interest, our team researched alternative data collection methods such as the chain referral method and the snowball method. The team also looked for potential funding opportunities for these types of research.

### 1.20.2011

*Meeting Consultant:* Jeff Corn, Community Programs Assistant at CURA

*Meeting Purpose:* To explore a possible partnership with CURA. Having little web design expertise, the team sought a partner for creating the dashboard and for updating and maintaining the dashboard.

*Meeting Outcomes:* Mr. Corn said that CURA would not likely maintain a database, and recommended that we partner with Humphrey School faculty for our research and connect with the Humphrey School IT Department to potentially find free web design assistance.

### **1.26.2011**

*Meeting Consultants:* Ryan Allen, Dean Lindsey, Merrie Benasutti

*Meeting Purpose:* To gain feedback on alternative data collection methodologies and the feasibility of these methodologies as part of our project.

*Meeting Outcomes:* In this meeting Ryan Allen indicated that our team should be wary of using these methods because they would not likely produce data more credible than official government data. Additionally, it became clear that developing and carrying out a survey using an alternative methodology was unfeasible and beyond the scope of our project. Instead, the idea of creating an accessible online dashboard of statistics (using official government data) was deemed both feasible and likely to be highly useful.

### **2.2.2011**

*Meeting Consultant:* Dahir Jibreel

*Meeting Purpose:* To discuss the unfeasible nature of employing alternative data collection methodologies as part of our project and to share with Mr. Jibreel different online dashboards as examples of what our team aimed to create.

*Meeting Outcomes:* Mr. Jibreel was understanding of our time and resource constraints and remained enthusiastic about the online dashboard as a project deliverable. He was hopeful that an alternative data collection methodology could be employed, perhaps by another set of C.H.A.N.C.E. capstone students, in the future. Greg Lindsey, Merrie Benasutti and Ryan Allen all met with Mr. Jibreel independently to further discuss the limitations of alternative methodologies.

Also at this meeting, Mr. Jibreel offered for team members to attend local community meetings about health concerns, so two team members attended two meetings to meet community members and hear their concerns.

### **2.16.2011**

*Internal Meeting:* Our team decided to divide data collection in order to collect as much raw data as possible from three sources: Minnesota Department of Education (MDE), Department of Homeland Security (DHS), and the American Community Survey (ACS).

**2.23.2011**

*Internal Meeting:* One team member introduced the rest of the team to Tableau – a free online data visualization software for creating interactive dashboards. This software was critical to project success because it allowed our team to design dashboards without a web designer. As a result of this finding, each team member cleaned the data in Excel spreadsheets so that it could be uploaded into Tableau.

**3.5.2011**

*Meeting Consultants:* Ryan Allen, Dean Lindsey, Merrie Benasutti

*Meeting Purpose:* To gain feedback on the data visualizations the team had recently created in Tableau as well as the indicators that might be most valuable for inclusion on the dashboard.

*Meeting Outcomes:* Ryan Allen and Dean Lindsey recommended that our team look closely at confidence intervals while collecting data and think about ways to present confidence intervals clearly using Tableau software. This became an integral part of our data collection process. Additionally, the meeting highlighted the importance of protocol for documenting our data collection process and the limitations of the data. Finally, indicators discussed as top priorities included: population by age, household income, poverty rate, employment rate, and educational attainment.

**3.11.2011**

*Internal Meeting:* After a meeting with Ryan Allen about the potential use of the IPUMS online analysis tool for collecting ACS data, the team decided to use IPUMS as the key source for our data collection. This was due to the large number of indicators available through IPUMS and the ability within IPUMS to display confidence intervals for all data queries. This choice ultimately also led to the decision to exclude Department of Homeland Security data from our final database. We did keep Minnesota Department of Education data, however, as a source not reliant on sample data. Our team also decided to set up a meeting with a representative of Minnesota Population Center to learn more about the IPUMS online analysis tool.

**3.23.2011**

*Meeting Consultant:* Dahir Jibreel

*Meeting Purpose:* To present a tentative list of indicators for the Somali population in Minnesota and gain feedback on the usefulness of those indicators for local advocates of the Somali community in Minnesota.

*Meeting Outcomes:* Mr. Jibreel explained his concern that that household income and employment data may be inaccurate due to Somali participation in the informal economy – a note that we eventually added to our data visualizations. Mr. Jibreel confirmed that all of the indicators on our list would be extremely useful for advocates working to improve the lives of the Somalis in Minnesota.

At this meeting our team also set a meeting date in April to present preliminary dashboards to community members for their feedback.

### **3.30.2011**

*Meeting Consultant:* Brandon Trampe

*Meeting Purpose:* To ask questions and gain technical information for using the IPUMS online analysis tool.

*Meeting Outcomes:* Our team learned that the problem of large confidence intervals would likely be inevitable unless we used 3-year pooled samples or manually pooled 5-year samples in IPUMS. This meeting also confirmed Ryan Allen’s understanding that there is no standard rule of thumb for judging whether a confidence interval is too large or too small to make the data usable or not. As a result of this meeting, our team decided to manually pool a 5-year sample for 2004-2009 and also to collect broad yearly data (for years 2001 through 2009) for 4 key indicators in IPUMS.

### **4.13.2011**

*Internal Meeting:* In preparation for the community feedback session scheduled for April 16, our team decided to spend the week creating online dashboards for all of our indicators. We also decided to upload all dashboards onto a blog, which acted as a preliminary host site for our dashboards. Finally, we drafted descriptions for each visualization to outline the data that each visualization presented and how to interpret the confidence intervals displayed and ran them by Dean Lindsey.

### **4.14.2011**

*Meeting Consultant:* Dahir Jibreel

*Meeting Purpose:* To gain feedback on the preliminary website of dashboards and to confirm plans for the upcoming feedback session.

*Meeting Outcomes:* Mr. Jibreel expressed enthusiasm for the dashboards and the tentative schedule we had created for the feedback session. Additionally, he offered to bring 15-20 community members that would have use for the online resource to the April meeting. Mr. Jibreel recommended that the session take place at the Humphrey School as a means of conveying the

partnership between the Cedar Riverside community and the school. Mr. Jibreel also offered to bring refreshments.

#### **4.16.2011**

*Meeting Consultants:* Dahir Jibreel, Dean Lindsey, Merrie Benasutti, and 12 leaders from the local Somali community

*Meeting Purpose:* To gain feedback from community members on the usefulness of our online dashboards as a means of raising awareness and ultimately improving our website to ensure that it is used.

*Meeting Outcomes:* As a result of the event, our team gained a sense of the comfort of local leaders from the Somali community with the data visualizations and explanations presented. The feedback offered was very positive and constructive, and the team learned that indicators related to housing, healthcare and crime were also of interest to community leaders. Additionally, the presence of Dean Lindsey and Merrie Benasutti helped to demonstrate the commitment of the school as a partner to the community.

#### **4.20.2011**

*Internal meeting:* After consulting Ryan Allen and Brandon Trampe, the team discovered that collecting a 5-year pooled sample would not necessarily produce reliable data and therefore may not prove to be feasible for this project. Additionally, the team learned that within one month of our project's completion, a pre-pooled 5-year sample would be available on IPUMS. As a result, the team decided not to create the pooled sample and instead to collect data and create Tableau dashboards using yearly ACS data for the additional indicators that community members expressed interest in at the feedback session.

<sup>i</sup>Joakim Gundel. *The migration-development nexus: Somalia case study*. International Migration, 40(5), 255-279 (2002).

<sup>ii</sup>Michelle Knoll, *Minnesota Home to Nation's Largest Somali Population*, KSTP Eye Witness News, <http://kstp.com/news/stories/S1882964.shtml> (December, 2010).

<sup>iii</sup>Toni Randolph, *State dept official: Census count of Somalis in MN is too low*, MPR News. <http://minnesota.publicradio.org/display/web/2010/12/16/somali-population-estimates/> (December, 2010).

<sup>iv</sup>Janice Penrod, Deborah Bray Preson Richard E. Cain and Michael T. Starks, *A Discussion of Chain Referral as a Method of Sampling Hard-to-Reach Populations*, Journal of Transcultural Nursing, 14, 100-107 (2003).

<sup>v</sup>Mathew J. Salganik & Douglas D. Heckathorn, *Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling*. Sociological Methodology, 34, 193-239, (2004).

<sup>vi</sup>The following websites were included in our literature review of online dashboards:

- *American Indian Community Profile and Data Center*. American Indian Policy Center. <http://airpi.org/projects/american-indian-community-profile.pdf>
- *Arizona Indicators*. Morrison Institute for Public Policy. <http://arizonaindicators.org/demographics/immigration>
- California Immigrant Policy Center. <http://www.caimmigrant.org/repository.html?topic=3>.
- Chicano Latino Affairs Council of Minnesota. <http://www.clac.state.mn.us/minnesota/origin.html>.
- *Demographic Profile of Hispanics in Minnesota 2008*. Pew Hispanic Center. <http://pewhispanic.org/states/?stateid=MN>
- *Latinos in Massachusetts: A Mid-Decade Status Report*. The Mauricio Gaston Institute for Latino Community Development and Public Policy. <http://www.gaston.umb.edu/articles/LatinMAMidDecRerp.pdf>
- *Looking Forward: Immigrant Contributions to the Golden State, 2010*. California Immigrant Policy Center. <http://www.caimmigrant.org/document.html?id=292>.
- Migration Policy Institute. <http://www.gaston.umb.edu/articles/LatinMAMidDecRerp.pdf>
- *Minnesota Factsheet*. U.S. Census: American Factfinder. [http://factfinder.census.gov/servlet/ACSSAFFacts?\\_event=Search&\\_state=04000US27](http://factfinder.census.gov/servlet/ACSSAFFacts?_event=Search&_state=04000US27)
- *Socioeconomic Status of Nebraska Hispanics or Latinos*. Nebraska Department of Health. <http://www.hhs.state.ne.us/healthdisparities/docs/SESofHispanics.pdf>
- *State of Latinos in the District of Columbia*. The Urban Institute: Metropolitan Housing and Communities Policy Center. <http://www.urban.org/uploadedpdf/412176-state-of-latinos.pdf>
- *Status of Women and Girls in Minnesota*. University of MN Humphrey Institute's Center on Women & Public Policy. [www.wfmm.org/research/2010Reports/2010SWGGM\\_FullReport.pdf](http://www.wfmm.org/research/2010Reports/2010SWGGM_FullReport.pdf)
- *Worldwide Trends in the Human Development Index 1970-2010*. United Nations Development Programming. <http://hdr.undp.org/en/data/trends/>

<sup>vii</sup>U.S. Census Bureau, *2005-2009 ACS 5-year PUMS Accuracy of the Data*, [http://www.census.gov/acs/www/Downloads/data\\_documentation/pums/Accuracy/2005\\_2009AccuracyPUMS.pdf](http://www.census.gov/acs/www/Downloads/data_documentation/pums/Accuracy/2005_2009AccuracyPUMS.pdf)

<sup>viii</sup>For more information about the differences in ACS Pool Sizes, see "When to use 1-year, 3-year, or 5-year estimates". [http://www.census.gov/acs/www/guidance\\_for\\_data\\_users/estimates/](http://www.census.gov/acs/www/guidance_for_data_users/estimates/)

<sup>ix</sup>"What General Data Users Need to Know", U.S. Census Bureau, A Compass for Understanding and Using American Community Survey Data. Page 4. <http://www.census.gov/acs/www/Downloads/handbooks/ACSGeneralHandbook.pdf>

<sup>x</sup>See IPUMS-USA for more information about *AGE*, <http://usa.ipums.org/usa-action/variables/AGE>

<sup>xi</sup>See IPUMS-USA for more information about *FAMILY INCOME*, <http://usa.ipums.org/usa-action/variables/FTOTINC>

<sup>xii</sup>See IPUMS-USA for more information about *POVERTY*, <http://usa.ipums.org/usa-action/variables/POVERTY>

<sup>xiii</sup>See IPUMS-USA for more information about *UNEMPLOYMENT*, <http://usa.ipums.org/usa-action/variables/EMPSTAT>