Data Analysis Scenarios

Demographic Data: Dunklin County

The first section of most health-related reports, including community health assessments and grants, should describe the basic characteristics, or demographics, of a community. Demographic data include age, race, ethnicity, gender, socioeconomic standing, and education level, among others. These characteristics are important because they can impact health. Here demographic data will be used to analyze the population of Dunklin County.

Population MICA is a good source for basic demographic data. Using this resource, the following table was created. The following table shows the comparison 2010 population totals for Dunklin County and the state of Missouri to those in 2015.

An analyst wants to learn how these populations have changed during the five year time span and determine if the trend in Dunklin County is different from the trend for the state as a whole. The analyst chooses to include some of this information in the text but decides against creating a chart or graph. Instead, they choose to calculate the percent change to determine directional difference for the geographies over the given time period, which will then be explained in the opening paragraph of the report.

Percent Change

Dunklin County: (30,895 - 31,953) / (31,953) = .0331 x 100 = - 3.31%

Missouri: (6,083,672 - 5,988,927) / (5,988,927) = - .0158 x 100 = + 1.58%

The opening report paragraph will include a relational sentence identifying the base value, comparison value, and in which direction the base value changed in relation to the comparison value. In the case of Dunklin County, the 2015 population was 3.31 % lower than the 2010 population. Statewide, the 2015 population is 1.58 % lower than the 2010 population.

After calculating the state and county percent change, the analyst decides to compare the age composition of Dunklin County to that of the state of Missouri. Age is a risk factor for many diseases and conditions, so this age structure could be an important determinant of the overall health status of Dunklin County. As seen on the previous table, the populations of Dunklin County and the state differ by over five million people. Therefore, it is impossible to make meaningful comparisons using only the population counts. In order to create a better comparison between the two geographies the analyst chooses to add percentages to the table. To do so, the analyst makes the query selections shown on the next page in the **Choose Your Data** portion of the screen.

In the **Build Your Results** section the analyst changes **Main Row** to Age, **Main Column** to Geography, and **Statistics** to Counts and Percents of Column Total before submitting the previously shown query.

Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row: Age ▼ Row Totals: ☑   **Main Column:** Geography ▼ Column Totals: ☑

Statistics: Counts and Percents of Column Tota ▼

Submit Query

Table Results

Save Table As ▾   **Send Table to Side by Side**

Title:	Missouri Resident Estimated Population			
Data selected in addition to rows and columns below:	Single Year(s): 2015;			
County:	Dunklin	Dunklin	Missouri	Missouri
Statistics:	Count	Percent of Column Total	Count	Percent of Column Total
Age				
Under 15	6,556	21.22	1,151,685	18.93
15 - 24	3,865	12.51	827,190	13.60
25 - 44	6,984	22.61	1,539,442	25.30
45 - 64	8,070	26.12	1,610,433	26.47
65 and Over	5,420	17.54	954,922	15.70
Total for selection	30,895	100.00	6,083,672	100.00
Source:	DHSS - MOPHIMS - Population MICA			
Generated On:	10/11/2017 8:52:38 AM			

Population MICA allows users to download the table into Excel with the **Save Table As** drop down, so the analyst can place the customized table in to my document.

The age distributions are very similar, so the analyst elects to only point out the largest difference between the geographies (circled in red above) and attempts to explain a possible reason for that difference. Since the analyst now knows how the age groups in Dunklin County compare to those in the state overall, they want to determine if those age groups are changing over time. To see if there have been any major changes, the analyst uses **Population MICA** to create a table that provides six years of data. In the **Choose Your Data** section of the screen, the analyst chooses years 2010 through 2015 from the drop down menu. Because the analyst is primarily interested in the changes in Dunklin County, they deselect the "Show State Totals" box. The analyst then navigates to the **Build Your Results** section and changes the **Main Row** variable from Age to Years and **Main Column** from Geography to Age, producing the table shown in the following.

Choose Your Data

Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row: Year ▼ **Row Totals:** ☑   **Main Column:** Age ▼ **Column Totals:** ☑

Statistics: Counts and Percents of Column Tota ▼

[Submit Query]

Table Results

[Save Table As ▾]   [Send Table to Side by Side]

| **Title:** | **Missouri Resident Estimated Population** |
| **Data selected in addition to rows and columns below:** | County: Dunklin; |

Age:	Under 15	Under 15	15 - 24	15 - 24	25 - 44	25 - 44	45 - 64	45 - 64	65 and Over	65 and Over	Total for selection	Total for selection
Statistics:	Count	Percent of Column Total	Count	Percent of Column Total	Count	Percent of Column Total	Count	Percent of Column Total	Count	Percent of Column Total	Count	Percent of Column Total
Year												
2010	6,718	16.55	3,994	16.81	7,526	17.29	8,447	16.96	5,268	16.41	31,953	16.84
2011	6,824	16.81	3,989	16.79	7,435	17.08	8,514	17.10	5,269	16.42	32,031	16.88
2012	6,868	16.92	3,997	16.82	7,354	16.89	8,334	16.74	5,319	16.57	31,872	16.79
2013	6,880	16.94	3,972	16.72	7,221	16.59	8,263	16.59	5,356	16.69	31,692	16.70
2014	6,757	16.64	3,940	16.58	7,019	16.12	8,167	16.40	5,461	17.02	31,344	16.52
2015	6,556	16.15	3,865	16.27	6,984	16.04	8,070	16.21	5,420	16.89	30,895	16.28
Total for selection	40,603	100.00	23,757	100.00	43,539	100.00	49,795	100.00	32,093	100.00	189,787	100.00

| **Source:** | **DHSS - MOPHIMS - Population MICA** |
| **Generated On:** | **10/25/2017 4:09:18 PM** |

When analyzing this table, the analyst discovers that the percentages shown are not the percentages expected. The goal was to see how each age group's percentage of the total population has changed from year to year. Therefore, the age groups in each year should sum to 100%. However, on this table, the total percentage for 2010 is only 16.84%. Closer examination reveals that each age group is summing to 100%, which does not make sense for this analysis. The analyst returns to **Build Your Results** and changes **Statistics** to Counts and Percents of *Row* Total. After submitting the query the analyst can now see (in the table shown on the next page) the percentages based on annual totals and that the age groups in each year total 100%.

MICA
MISSOURI INFORMATION
FOR COMMUNITY ASSESSMENT
DATA MICAS

⌃⌄ Choose Your Data

⌃⌄ Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row: Year ▼ **Row Totals:** ☑ **Main Column:** Age ▼ **Column Totals:** ☑

Statistics: Counts and Percents of Row Total ▼

Submit Query

⌃⌄ Table Results

Save Table As ▾   Send Table to Side by Side

Title:	Missouri Resident Estimated Population
Data selected in addition to rows and columns below:	County: Dunklin;

Age:	Under 15	Under 15	15 - 24	15 - 24	25 - 44	25 - 44	45 - 64	45 - 64	65 and Over	65 and Over	Total for selection	Total for selection
Statistics:	Count	Percent of Row Total	Count	Percent of Row Total	Count	Percent of Row Total	Count	Percent of Row Total	Count	Percent of Row Total	Count	Percent of Row Total
Year												
2010	6,718	21.02	3,994	12.50	7,526	23.55	8,447	26.44	5,268	16.49	31,953	100.00
2011	6,824	21.30	3,989	12.45	7,435	23.21	8,514	26.58	5,269	16.45	32,031	100.00
2012	6,868	21.55	3,997	12.54	7,354	23.07	8,334	26.15	5,319	16.69	31,872	100.00
2013	6,880	21.71	3,972	12.53	7,221	22.78	8,263	26.07	5,356	16.90	31,692	100.00
2014	6,757	21.56	3,940	12.57	7,019	22.39	8,167	26.06	5,461	17.42	31,344	100.00
2015	6,556	21.22	3,865	12.51	6,984	22.61	8,070	26.12	5,420	17.54	30,895	100.00
Total for selection	40,603	21.39	23,757	12.52	43,539	22.94	49,795	26.24	32,093	16.91	189,787	100.00

Source:	DHSS - MOPHIMS - Population MICA
Generated On:	10/25/2017 4:10:41 PM

Although the analyst will need all of these data for the final analysis, there are so many numbers included on the table that it is hard to comprehend. Instead of reproducing the table in the report, the analyst decides to visualize these patterns in a line graph and include this graph in the final report so that readers can more easily see trends. When graphing only a few years of data, a bar chart could be used as an alternative to the line chart. However, if many years of data are to be graphed, line charts are usually the best option.

In **Population MICA** charting using percentages is not an available functionality so the analyst uses the **Save Table As** feature to export the data to Excel and create a line graph based on the percentages.

When developing the line chart in Excel, the analyst knows they need to develop the graph based on the percentages, not the counts, because the issue at hand is whether the age *distribution* has changed over time, not whether the population numbers have changed. Percentages can provide more insight into meaningful variations over time rather than counts. Therefore, once the analyst has downloaded the data from **Population MICA** into Excel they can delete the **Count** columns, graphing only the percentages. This allows readers to see the percentage changes and more clearly conveys the intended message. Also, percentages can be interpreted more easily than potentially large frequency counts. Furthermore, using percentages rather than frequencies will

allow for a fairer comparison if a reader wishes to compare Dunklin County's age distribution to that of another area.

The analyst must include appropriate contextual information in order to complete the graph, including an overall title and axis labels. The vertical axis label specifies that the numbers on that axis are percentages. The analyst also add a source note beneath the graph to inform readers that it was created using data from **Population MICA**.

Injury Data: Boone County

This section of the sample community health assessment will analyze data related to injuries in Boone County. Areas that may be of concern to readers include the types of injuries occurring, different demographic groups involved, whether the number of injuries is increasing or decreasing, and many other related issues. In the following examples, an analyst will use confidence intervals to determine if there are meaningful differences between the injury rates compared.

Community health assessments will usually require that a county address health disparities among different population groups. **One way to determine if a disparity exists is to compare the confidence intervals for different groups.** For example, the two largest racial groups in Boone County are Whites and African-Americans. The analyst would like to determine if injuries are affecting one of these groups more than the other. To find this information, **Injury MICA** is used. The analyst decides to look at the most recent **Year** of data available, which happens to be the default (in this case 2014) and chooses Boone County from the Geography dropdown. Under **Build Your Results** Race could be displayed along the **Main Row** or **Main Column** so both racial categories can be displayed. The analyst leaves the default variable **Year** as the **Main Column** of interest and to determine statistically significant disparities among the two racial groups, select 95% confidence intervals to be displayed.

∧∨ Choose Your Data

∧∨ Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row:	Race ▼	**Row Totals:** ☑
Statistics:	Counts and Rates ▼	
Confidence Intervals:	No Confidence Intervals ▼	

Main Column:	Geography ▼	**Column Totals:** ☑
Age Adjustment Options:	2000 Standard Population ▼	

Submit Query

∧∨ Table Results

Save Table As ▾   **Send Table to Side by Side**

Title:	Missouri Resident Injuries		
Data selected in addition to rows and columns below:	Single Year(s): 2014;		

County:	Boone	Boone	Missouri	Missouri
Statistics:	Count	Rate	Count	Rate
Race				
White	8,653	6,118.59	395,774	7,861.03
Black or African-American	2,379	12,551.48	80,271	10,459.69
All Races	11,649	6,861.45	497,847	8,362.12

Rate:	Injury rates are annualized per 100,000 residents and are age adjusted to the U.S. 2000 standard population.
Source:	DHSS - MOPHIMS - Injury MICA
Generated On:	10/25/2017 4:15:24 PM

The confidence intervals for White individuals and Black/African-American individuals can be compared to determine statistical significance. There is no overlap between the two groups and Black-African/Americans clearly have higher injury rates in Boone County. Therefore, it is determined there is a statistically significantly higher rate of injury for Black residents in Boone County than for White residents. It would also be correct to say the rate of injury for Whites is statistically significantly lower than that for Black/African-American residents.

White	5,989.66 —— 6,247.51
Black/African-American	12,047.11 —— 13,055.86

To compare injury occurrence trends over time additional data years can be added. By returning to **Choose Your Data** and selecting nine additional years preceding the 2014 results, an analyst can then submit a query which allows comparison of the confidence intervals for the last ten years of injury data.

MICA
MISSOURI INFORMATION
FOR COMMUNITY ASSESSMENT
DATA MICAS

▲▼ Choose Your Data

▲▼ Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row: Year ▼   Row Totals: ☑   **Main Column:** Geography ▼   Column Totals: ☑

Statistics: Counts and Rates ▼   Age Adjustment Options: 2000 Standard Population ▼

Confidence Intervals: No Confidence Intervals ▼

[Submit Query]

▲▼ Table Results

[Save Table As ▾]   [Send Table to Side by Side]

Title:	Missouri Resident Injuries			
Data selected in addition to rows and columns below:	None			

County:	Boone	Boone	Missouri	Missouri
Statistics:	Count	Rate	Count	Rate
Year				
2005	13,186	8,953.03	591,255	10,334.75
2006	13,107	8,661.00	583,069	10,118.33
2007	13,561	8,808.90	598,081	10,309.72
2008	12,687	8,161.11	575,981	9,880.31
2009	12,707	8,075.78	564,388	9,637.42
2010	12,369	7,760.60	557,563	9,475.91
2011	11,608	7,113.91	560,048	9,503.92
2012	11,430	6,864.05	558,849	9,476.61
2013	11,460	6,826.45	532,850	8,995.35
2014	11,649	6,861.45	497,847	8,362.12
Total for selection	123,764	7,773.32	5,619,931	9,609.42

Rate:	Injury rates are annualized per 100,000 residents and are age adjusted to the U.S. 2000 standard population.
Source:	DHSS - MOPHIMS - Injury MICA
Generated On:	10/25/2017 4:16:24 PM

The confidence intervals for 2005, 2006, and 2007 overlap, so there were no statistically significant changes in injury occurrence for Boone County during those years. However, the confidence intervals for the years after 2008 do *not* overlap the intervals from the earlier years. Thus, there was a statistically significant decrease between 2007 and 2008. There was another significant decrease between 2010 and 2011. The analyst should note these findings in their report and determine if this significance warrants a visual representation.

When writing a community health assessment or grant application, the needs of the community should be clearly described. A thorough explanation of the community's needs is important because it will allow readers to understand the work that needs to be done and consider the types and amounts of resources that could be utilized to address those needs. However, it is very easy to focus only on problem areas in a community and neglect to describe improvements that have been made. Highlighting positive trends (such as Boone County's improvement in injury rates) in assessments and grant applications is just as important as describing problem areas. A report that is completely negative will only discourage the community. Including positive trends shows

that the community has the potential to make improvements and recognizes the community's prior achievements.

As demonstrated, confidence intervals can be a valuable tool for analyzing data. However, overall context must be kept in mind when using confidence intervals.

1.   Compare injury occurrence in Boone County to that in the State of Missouri and generate the following table using **Injury MICA**.

Injury MICA

∧∨Choose Your Data

∧∨Build Your Results

| Build a Table | Make a Map | Create a Chart | Documentation / Metadata |

Main Row:	Geography ▼	Row Totals: ☑		Main Column:	Statistics ▼	Column Totals: ☑
Statistics:	Counts and Rates ▼			Age Adjustment Options:	2000 Standard Population ▼	
Confidence Intervals:	95% Confidence Intervals ▼					

Submit Query

∧∨Table Results

Save Table As ▾   Send Table to Side by Side

| | Title: Missouri Resident Injuries |
| Data selected in addition to rows and columns below: | Single Year(s): 2014, 2013, 2012, 2011, 2010; |

Statistics:	Count	Rate	Lower 95% Conf Limit	Upper 95% Conf Limit
County				
Boone	58,516	7,078.18	7,020.83	7,135.53
Missouri	2,707,157	9,163.45	9,152.53	9,174.36

Rate:	Injury rates are annualized per 100,000 residents and are age adjusted to the U.S. 2000 standard population.
Source:	DHSS - MOPHIMS - Injury MICA
Generated On:	10/25/2017 4:18:37 PM
Confidence Intervals:	95% confidence intervals are displayed.

In this example, Boone County's rate of injury occurrence is significantly lower than the state rate of injury occurrence. Does this mean that Boone County definitely does not have a problem with injuries?

Suppose the analyst researches this topic further and finds that Missouri's rate of injury occurrence is statistically significantly higher than the US rate. Therefore, even though Boone County's rate is significantly lower than the Missouri rate, it could still be significantly higher than the rate for the rest of the nation!

Boone County	7,020.83------7,135.53	
State of Missouri		9,152.53---9,174.36
United States*	6,632.91—6,700.62	
*Fictional confidence interval		

2. Whether the desired rate is statistically significantly higher or statistically significantly lower depends on the indicator involved. For example, having a statistically significantly lower rate of undesirable conditions, such as injury-related deaths, is considered to be good. However, if Boone County had a statistically significantly lower rate of lead testing than the state, that would indicate a problem, because Boone County would not be testing as many children as the state overall.

3. Be careful when defining an issue and selecting data for analysis. Suppose the user would like to compare injury hospitalizations to hospitalizations for some other cause, such as heart and circulation problems. **Inpatient Hospitalization MICA** is selected to compare these two diagnoses. The table on the following page shows the number of injury and heart/circulation hospitalizations for Boone County residents from 2010-2014.

Inpatient Hospitalization MICA

∧∨Choose Your Data

∧∨Build Your Results

Build a Table	Make a Map	Create a Chart	Documentation / Metadata

Main Row:	Diagnosis ▼	**Row Totals:** ☑	**Main Column:**	Statistics ▼	**Column Totals:** ☑
Statistics:	Counts and Rates ▼		Age Adjustment Options:	2000 Standard Population ▼	
Confidence Intervals:	95% Confidence Intervals ▼				

Submit Query

∧∨Table Results

Save Table As ▾   Send Table to Side by Side

Title:	Missouri Resident Inpatient Hospitalizations
Data selected in addition to rows and columns below:	Type of Data: Hospital Discharges; County: Boone; Single Year(s): 2014, 2013, 2012, 2011, 2010;

Statistics:	Count	Rate	Lower 95% Conf Limit	Upper 95% Conf Limit
Diagnosis				
Heart and circulation	9,412	131.72	129.06	134.38
Injury and poisoning	6,796	89.10	86.98	91.22
Total for selection	16,208	220.83	217.43	224.23

Rate:	Inpatient hospitalization rates are annualized per 10,000 residents and are age adjusted to the U.S. 2000 standard population.
Source:	DHSS - MOPHIMS - Inpatient Hospitalization MICA
Generated On:	10/25/2017 4:21:13 PM
Confidence Intervals:	95% confidence intervals are displayed.

The 95% confidence interval for heart/circulation hospitalizations is statistically significantly higher than the 95% confidence interval for injury-related hospitalizations. According to these data, heart/circulation problems may be more of an issue in Boone County than injuries.

However, when analyzing some other type of data, a different picture may emerge. For instance, use **Emergency Room MICA** to find the 95% confidence intervals for injury and heart/circulation ER visits from 2010-2014.

Thus, although injury-related *hospitalizations* are statistically significantly lower than heart and circulation *hospitalizations*, injury *ER visits* are statistically significantly higher than heart and circulation *ER visits*.

One interpretation of these data is that heart and circulation problems are usually more severe and more likely to require hospitalization, but injuries affect a much larger number of people. In most cases, multiple types of data should be considered when setting priorities.

NOTE: Always check to see if rates (and confidence intervals) are based on the same constant when making comparisons. In these examples, rates were per 100,000 on **Injury MICA**, 10,000 on **Inpatient Hospitalization MICA**, and 1,000 on **Emergency Room MICA**. Comparisons between rates were originally made on the same table. To make comparisons between tables from two different MICAs, the user would need to convert one set of rates to the constant used on the other table. For example, to compare an ER Visit rate of 525.3 per 1,000 to a rate from **Death MICA**, the rate must be converted to the **Death MICA** default constant of 100,000. Move the decimal to the right two more spaces for a rate of 52,530.00.

$100{,}000 \div 1{,}000 = 100 * 525.3 = 52{,}530$ per 100,000

A final comment on resource allocation as it relates to data: This handbook focuses on how to most effectively analyze and present data, but data should not be the sole determinant in the prioritization/allocation of time and funding. The data found in MICA or in other sources are only one piece in a complex process used by communities to allocate limited resources. Factors such as amenability to change and community support of particular programs, among other things, may at times trump data-based findings; however, a solid understanding of the numbers will ensure that more informed decisions can be made.