

## Effective Use of Bubble Charts

- |                 |   |
|-----------------|---|
| <b>Purpose</b>  | This tool provides guidelines and tips on how to effectively use bubble charts to communicate research findings.  |
| <b>Format</b>   | This tool provides guidance on bubble charts and their purposes, shows examples of preferred practices and practical tips for bubble charts, and provides cautions and examples of misuse and poor use of bubble charts and how to make corrections.                        |
| <b>Audience</b> | This tool is designed primarily for researchers from the Model Systems that are funded by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). The tool can be adapted by other NIDILRR-funded grantees and the general public. |

The contents of this tool were developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DP0012-01-00). The contents of this fact sheet do not necessarily represent the policy of Department of Health and Human Services, and you should not assume endorsement by the Federal Government.

# Bubble Charts

- ▶ Bubble Charts resemble XY Scatter Charts - but can convey information regarding a third data element per observation, using the size of each traditional XY plotted point (expressed as a “bubble” instead of as a “dot”) to express the magnitude of the third variable.
- ▶ Bubble charts thus plot triplets of linked data elements per observation.
- ▶ Example: A typical XY Scatter chart might be used to display the relationship between income and life expectancy at birth, with county, state, or country as the unit of analysis (generally, higher income per capita or per household is associated with longer life expectancy).
- ▶ If you wished to also express the population size of each county etc on the same chart, you could proportionally enlarge or shrink each plotted data point in the form of a bubble (counties with larger populations would have larger bubbles on the XY Scatter plot) – so as to simultaneously highlight the relationship between income and life expectancy, but also to highlight where the largest populations live within this relationship.
- ▶ Thus, the third value in the data triplet per observation determines the size of the bubble.

# Bubble Charts

- ▶ Bubble Charts are available in Excel 2010 and later.
- ▶ Bubble Charts are also available in most dedicated data visualization software packages.

# Standard XY Scatter Chart

Median Household Income and Life Expectancy Across NC Counties

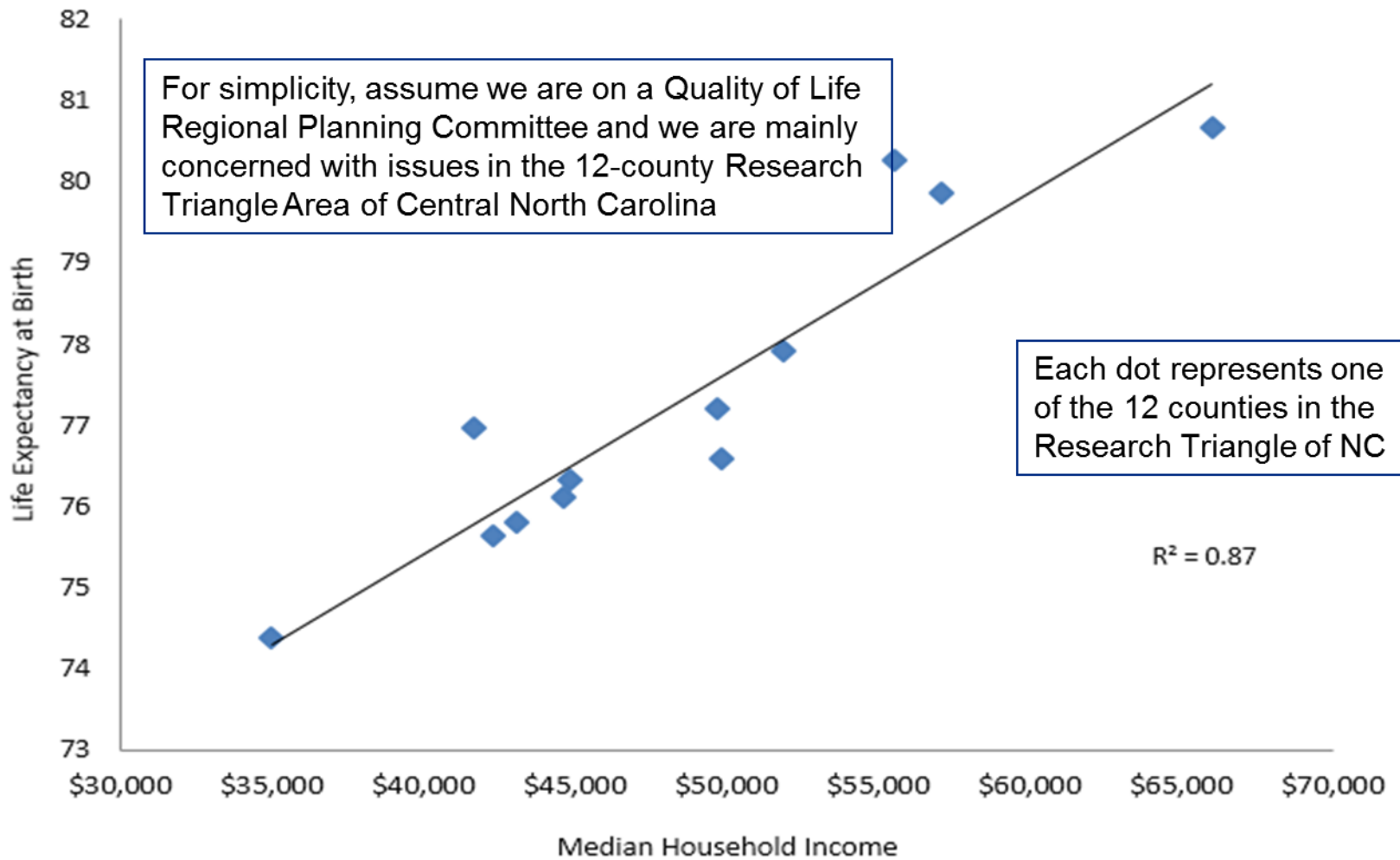


Income: Census ACS five-year aggregation ending in 2013.

Life Expectancy: Institute for Health Metrics and Evaluation IHME 2010 (male and female combined)

# Standard XY Scatter Chart

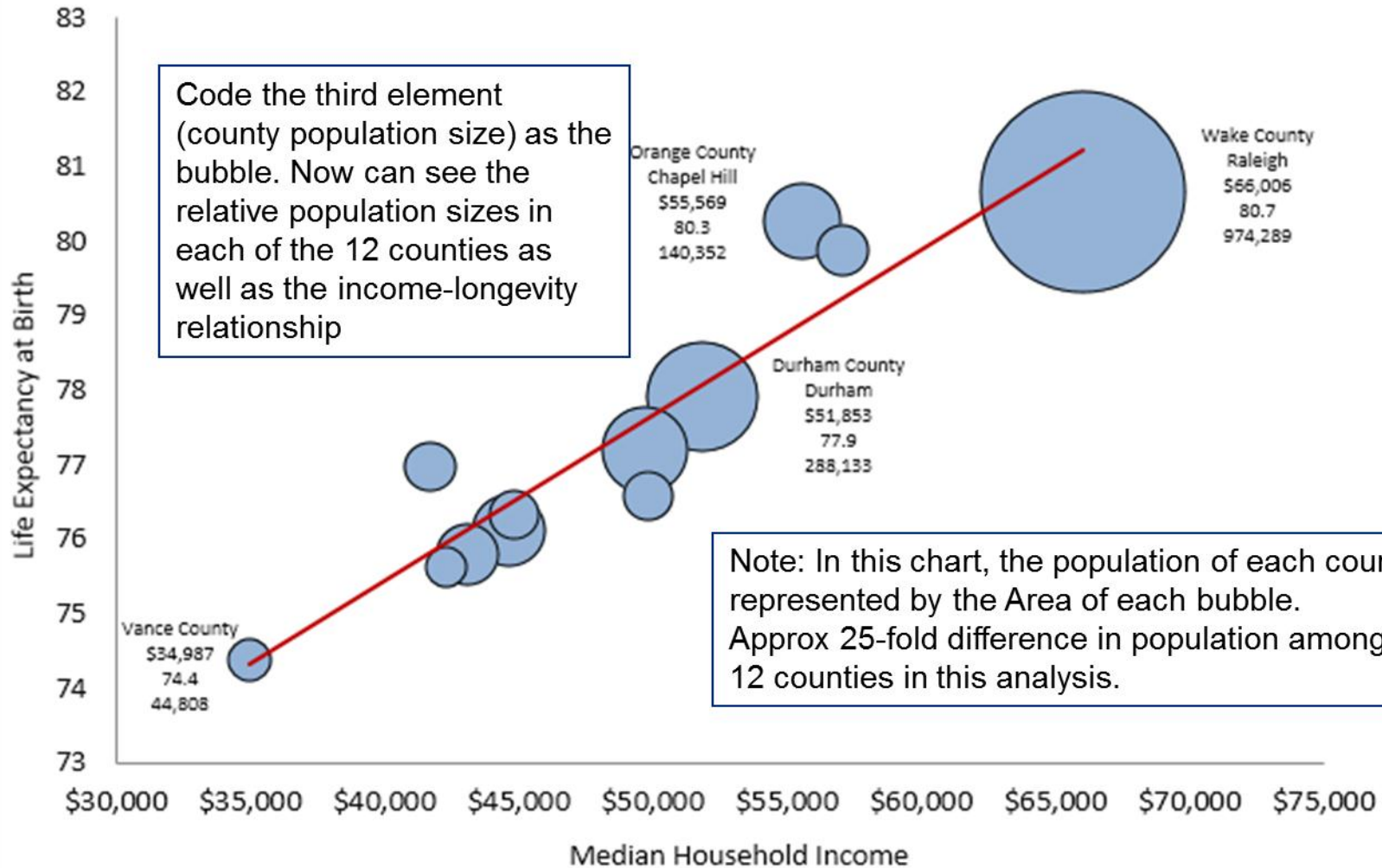
Median Household Income and Life Expectancy - Research Triangle Area NC



Income: Census ACS five-year aggregation ending in 2013.  
Life Expectancy: IHME 2010 (male and female combined)

# Bubble Chart

Median Household Income - Life Expectancy - And Total Population  
By County in the Research Triangle Area North Carolina

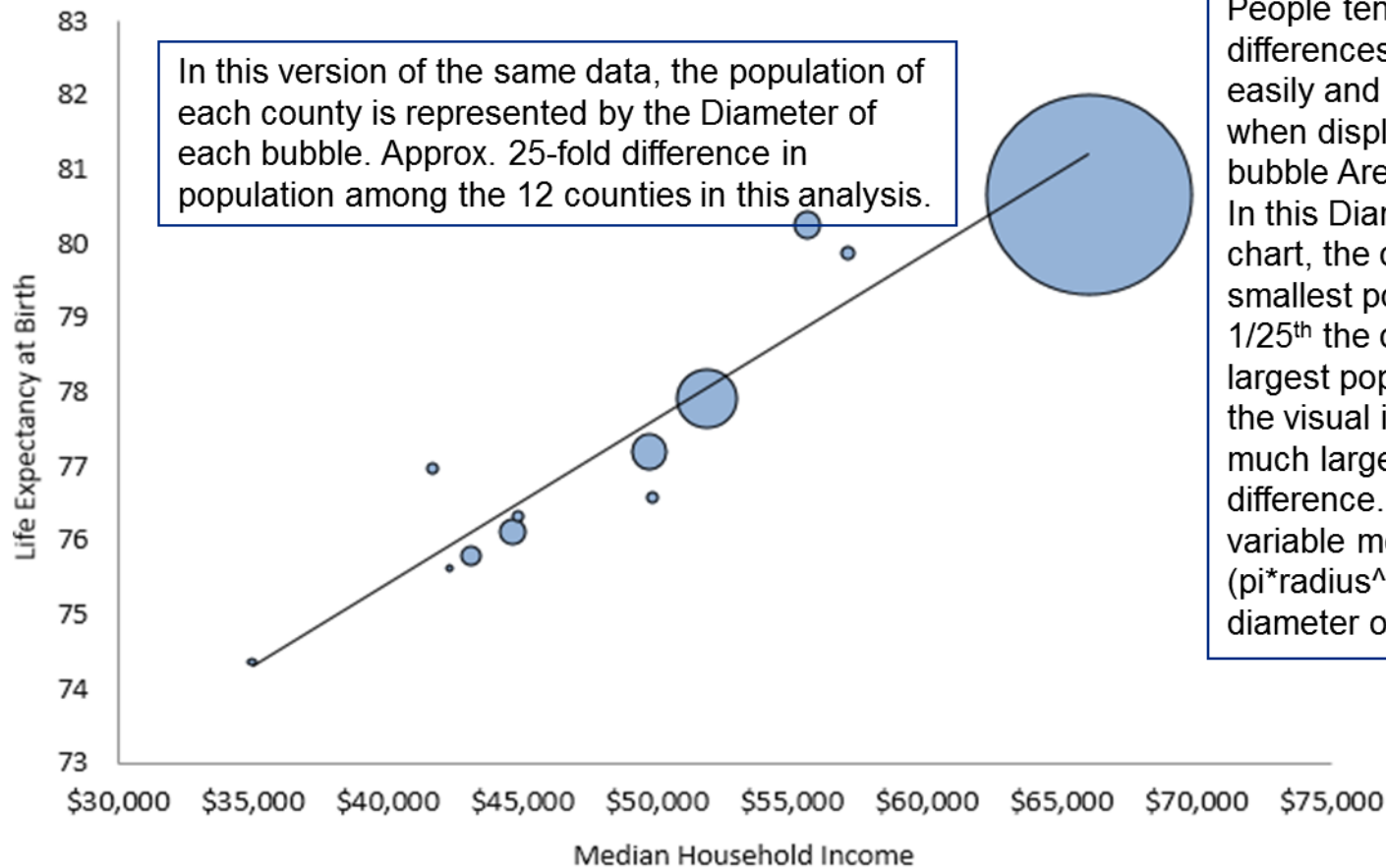


Population: Census 2013 Est.  
Population Coded as Bubble Area.

# Bubble Chart

## Cautions, Misuse, and Poor Use

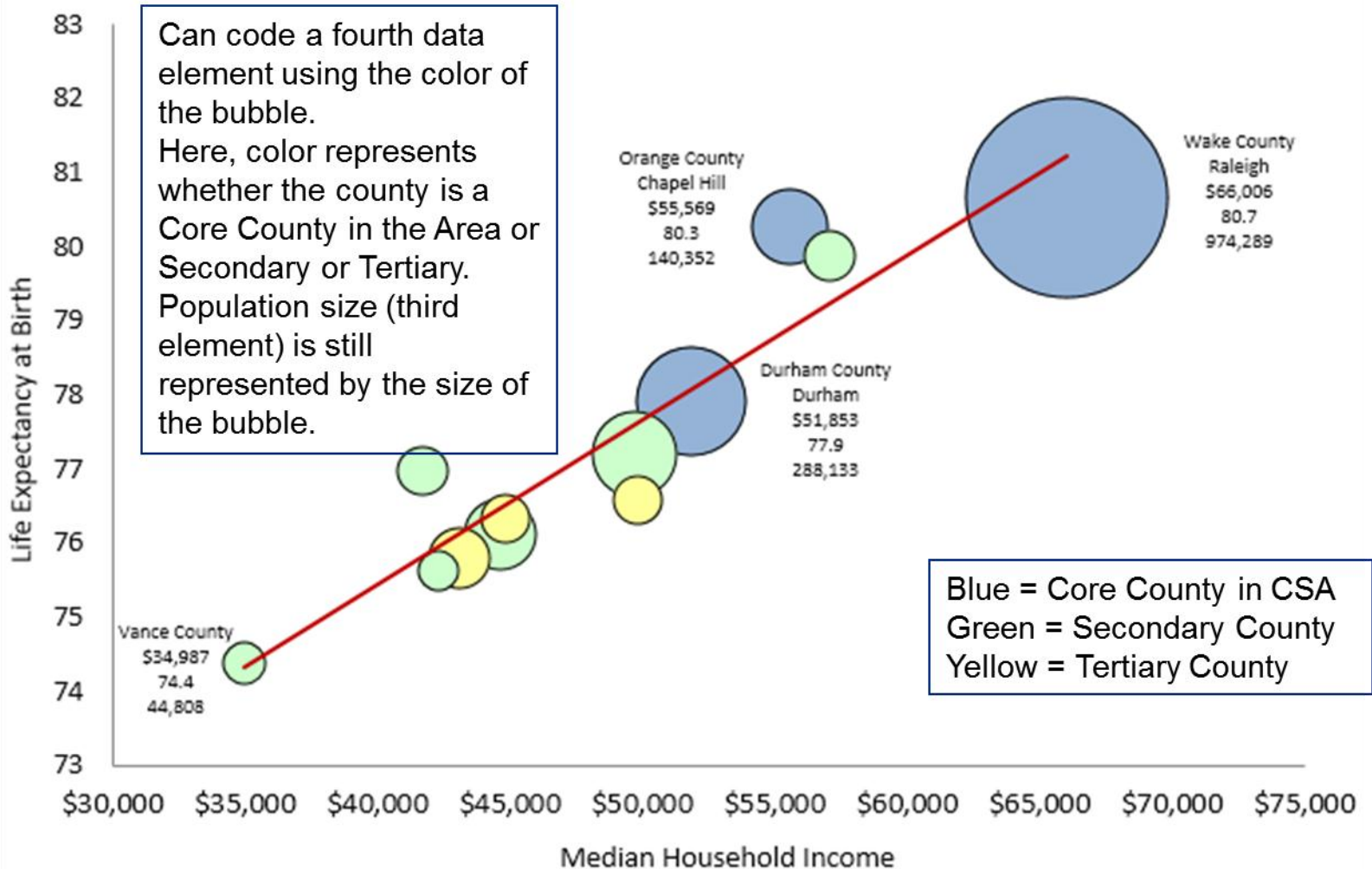
Median Household Income - Life Expectancy - And Total Population  
By County in the Research Triangle Area North Carolina



People tend to associate differences in magnitude more easily and more accurately when displayed in terms of bubble Area (previous version). In this Diameter-based bubble chart, the diameter of the smallest population county is 1/25<sup>th</sup> the diameter of the largest population county, but the visual impression is of much larger population difference. So plot the third variable metric as Area ( $\pi \cdot \text{radius}^2$ ) rather than as diameter or radius.

# Bubble Chart – Fourth Data Element

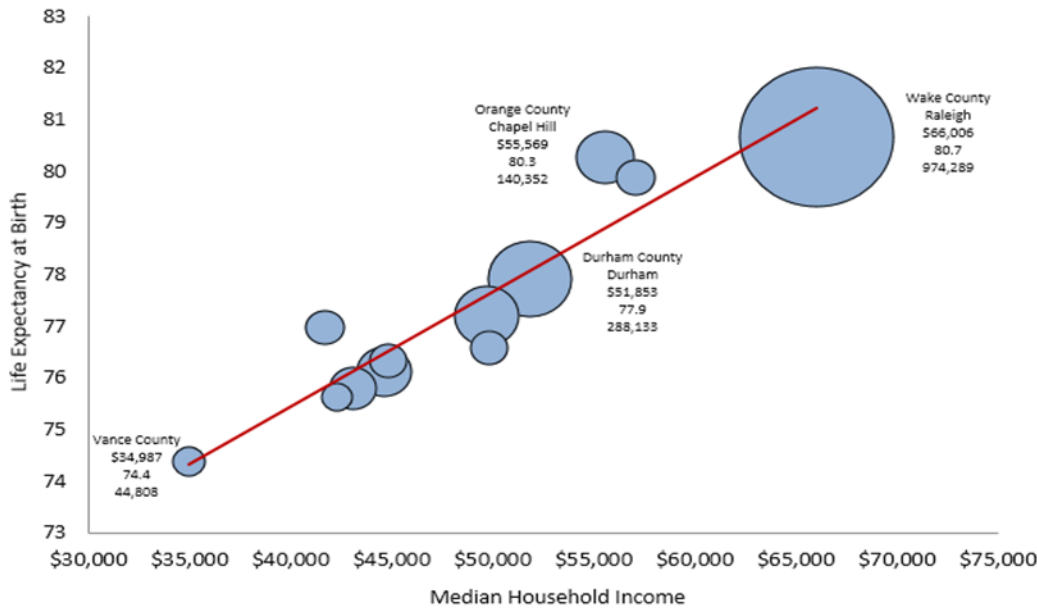
Median Household Income - Life Expectancy - And Total Population  
By County in the Research Triangle Area North Carolina





# Bubble Chart

Median Household Income - Life Expectancy - And Total Population  
By County in the Research Triangle Area North Carolina



- ▶ Bubble charts are mainly useful for coding the general impression of the third element in the triplet.
- ▶ Footnote whether the third data element is coded as bubble area (preferred) or bubble diameter.

Note that bubble charts can become more difficult to interpret when the magnitude differences in the bubble data elements are large.

Note that bubble charts with a large number of triplet data points will tend to have bubbles that overlap and obscure other bubbles, thus limiting the usefulness, completeness, and accuracy of interpretation of the bubble chart.

If your third variable contains negative values or zeros, probably best to avoid bubble charts. Can finesse negative and zero values with special symbols etc, but such slows interpretation considerably.