

## SASEM – Creating Histograms

### With specified Bins

(Fall 2016)

**Sources** (adapted with permission) -

Ron Freeze Course and Classroom Notes

Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville

Applied Analytics Using SAS® Enterprise Miner Course Notes & Workshop, 2015

Data source is the Association for computing Machinery's (ACM) 1998 KDD-Cup competition. The data set and other details of the competition are publicly available at the UCI KDD Archive at [HTTP://kdd.ics.uci.edu/](http://kdd.ics.uci.edu/).

*For educational uses only - adapted from sources with permission. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission from the author/presenter.*



## Charity Direct Mail

*A national veterans' organization seeks to better target its solicitations for donation. By soliciting only the most likely donors, less money is spent on solicitation efforts and more money is available for charitable concerns. Solicitations involve sending a small gift to an individual and include a request for a donation. Gifts to donors include mailing labels and greeting cards. The organization has more than 3.5 million individuals in its mailing database. These individuals are classified by their response behaviors to previous solicitation efforts. Of particular interest is the class of individuals identified as **lapsing donors**. These individuals made their most recent donations between 12 and 24 months ago. The organization seeks to rank its lapsing donors based on their responses to a greeting card mailing sent in June of 1997. (The charity calls this the 97NK Campaign.) With this ranking, a decision can be made to either solicit or ignore a lapsing individual in the June 1998 campaign.*

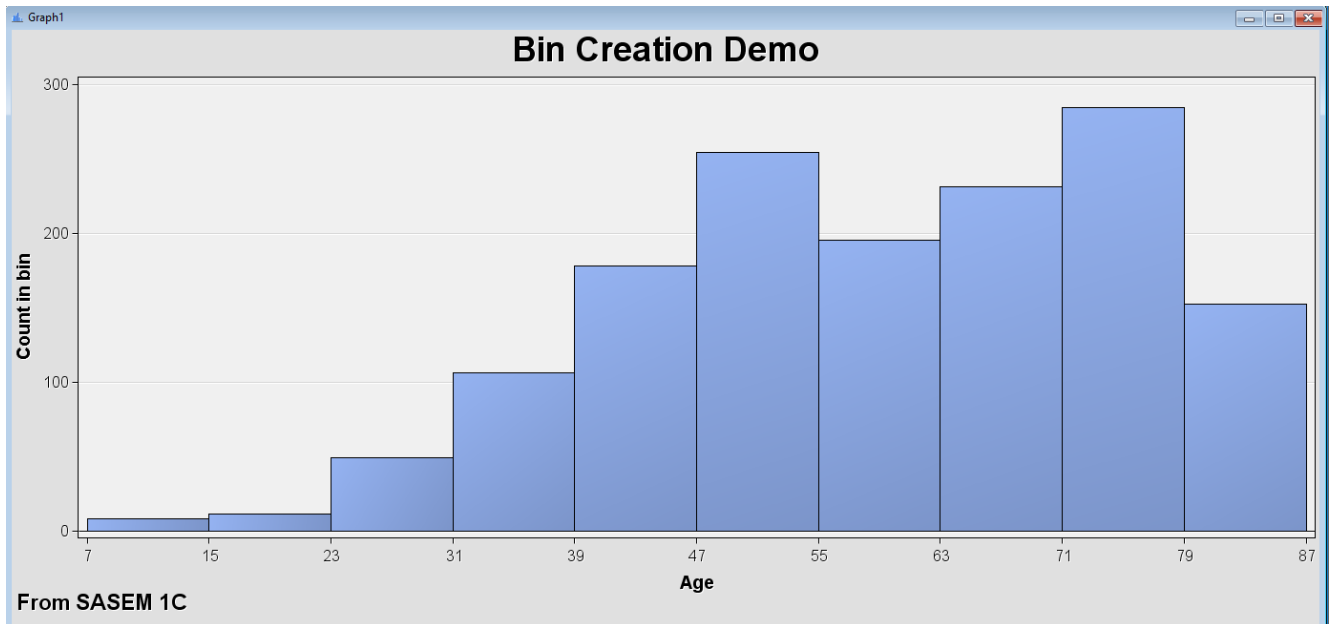
## Connecting a Data Source

1. Connect to the PVA97NK data source referred to in the Enterprise miner section of the documentation to login and connect to the data source (SASEG 0).

## Explore the Data Source – Creating a Histogram

2. Right click on the data source and select Explore (2000 rows are fetched by default)
3. Select Actions -> Plot
4. Select Histogram and then Next
5. Click the box under Role in the DemAge row
6. Select X in the DemAge under Role
7. Select Next
8. We will not be applying a WHERE Clause so Select Next
9. Place a Title, Footnote, X Axis Label and Y Axis Label
  - a. Title – Bin Creation Demo
  - b. Footnote – From SASEM 1C
  - c. X Axis Label – Age
  - d. Y Axis Label – Count in Bin
10. Select Next

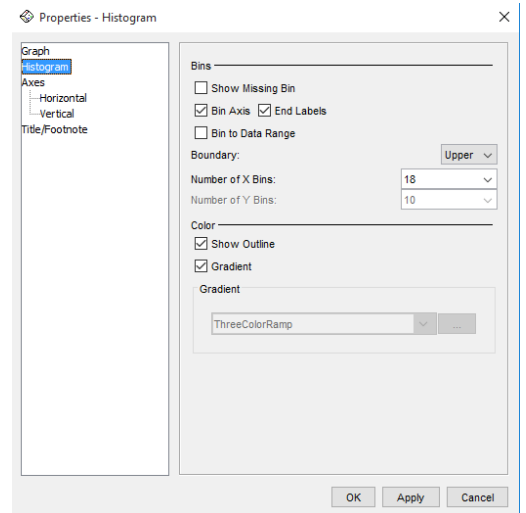
11. Select Finish and the Histogram should look similar to the screen shot



## Setting Bin Size

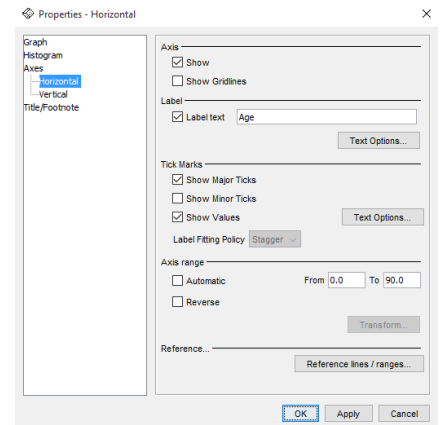
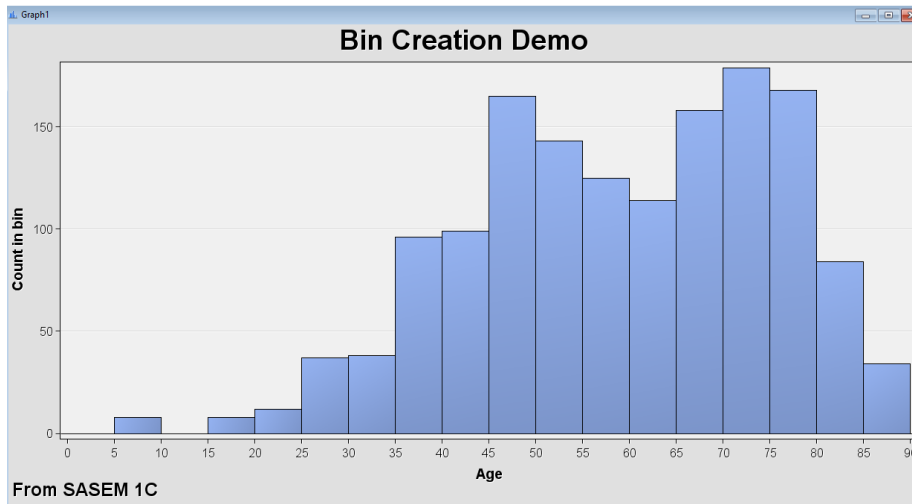
The histogram however is not separated correctly for the number of bins we wish to display. We would like to display the histogram in 5-year increments starting at 0 and going through age 90. This means we will have 18 bins.

12. Right click in the chart area and select graph properties.
13. Select Histogram
14. Change Number of X Bins: to 18.
15. Select Axes-Horizontal
16. Uncheck the Axis range box and set the range as follows
  - a. From: 0
  - b. To: 90



17. Select Apply and OK

18. Your chart should now have 18 bins with each column representing groups of 5 and look like the following screen shot



## Saving your Histogram

The histogram can be saved as a bitmap from the Explore window.

19. Select File -> Save As

20. Locate your student folder

21. Name the Histogram as <lastname>BinCreationDemo