

A Picture Is Worth A Thousand Data Points - Increase Understanding by Mapping Data Midwest SAS® Users Group (MWSUG)

Paul Ciarlariello, Sinclair Community College, Dayton, OH

ABSTRACT

If you find yourself faced with difficult to interpret data it may be helpful to “paint” the data onto a map. This paper presents information for transforming numeric data into visual data using SAS 9.1 and Microsoft Excel and MapPoint. There are also step by step instructions on how to automatically update your map for presentation on the SAS Information Delivery Portal.

INTRODUCTION

The following paper deals with visualizing student enrollment differences between one year and the next. Data was overlaid onto a map by ZIP Codes. And although the example presented here specifically relates to the education market, this process can easily be used whenever the data is geographic in nature. Examples of other potential applications are provided at the end of this paper.

This paper is broken down into three sections...

Section 1 - shows how to quickly create a mapped ZIP code HTML file using SAS 9.1, Microsoft Excel, and Microsoft MapPoint.

Section 2 - shows how to automate updating the map and making it accessible on a website.

Note - To accomplish the automated part of this process you will need to have a working knowledge of your network infrastructure. This is necessary if you want to publish your results to a website.

You will also need...

- to place the resultant html file in a publically accessible location on the web server
- to have administrative rights on the SAS Information Delivery Portal
- Microsoft MapPoint software to generate the HTML map file
- to schedule the update task to run in the operating system
- to produce some basic VB (Visual Basic) Scripting for the update
- to optionally generate some HTML code to customize the map file if necessary.

Section 3 - shows an option for animating your map using Adobe Photoshop and ImageReady.

SCENARIO SET UP

For the purposes of this exercise a scenario is being used to help illustrate the process. The enrollment example is used throughout as a simple way to relate to the steps.

Background – When data is reported in summary fashion, it gives an overall impression and makes it easier to understand at a high level. If I were to say for example, enrollment for fall term is up 3% from the previous year, it is fairly easy to understand. If however you needed to look at details regarding that 3% you could be faced with an incomprehensible list of numbers. So let’s set up a scenario to illustrate how we might deal with detailed data.

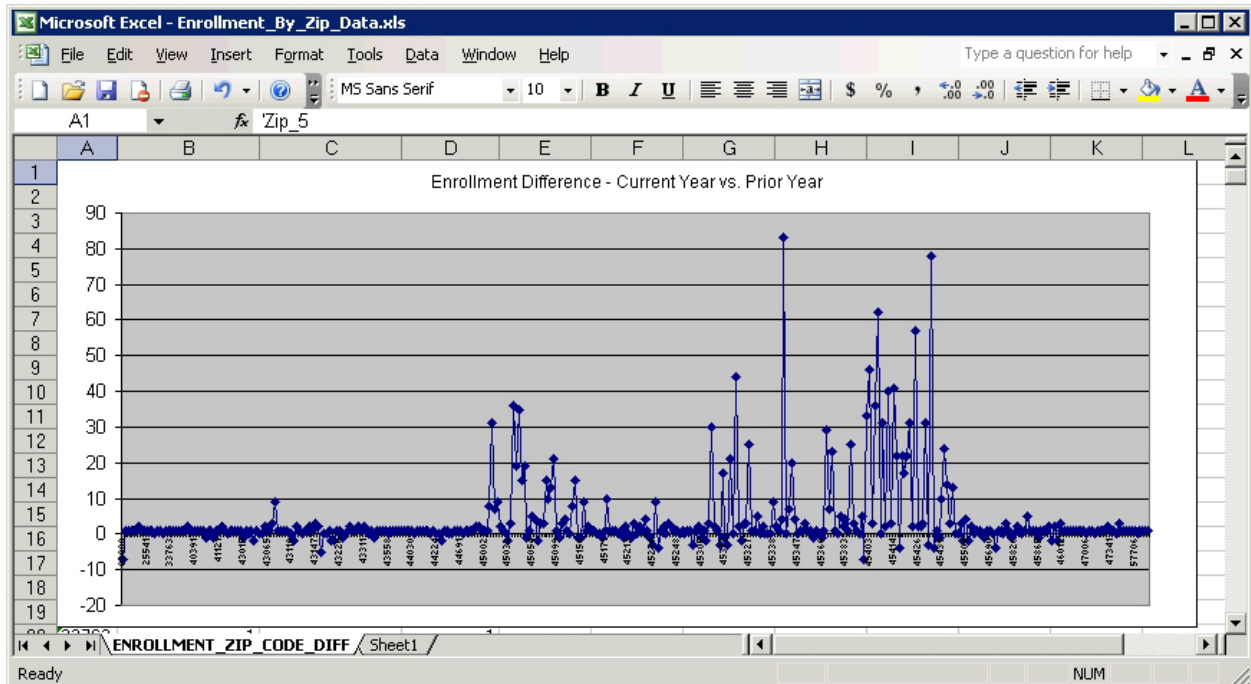
Scenario - You are the admissions office manager. You have been given a mandate to increase enrollment by 5% over the previous year. You’ve just been given a report which says the college managed a 3% increase last year. You want to formulate a marketing campaign, but you would like to see if the report data can yield any useful information. You decide that you would like to look at the geographic breakdown of that 3% enrollment report data. You request a breakdown and here is what you get back...

NOTE – DATA VALUES REFERED TO IN THIS PAPER ARE FICTICIOUS; THEY ARE FOR ILLUSTRATIVE PURPOSES ONLY.

RAW DATA AND GRAPH FROM 3% ENROLLMENT INCREASE REPORT

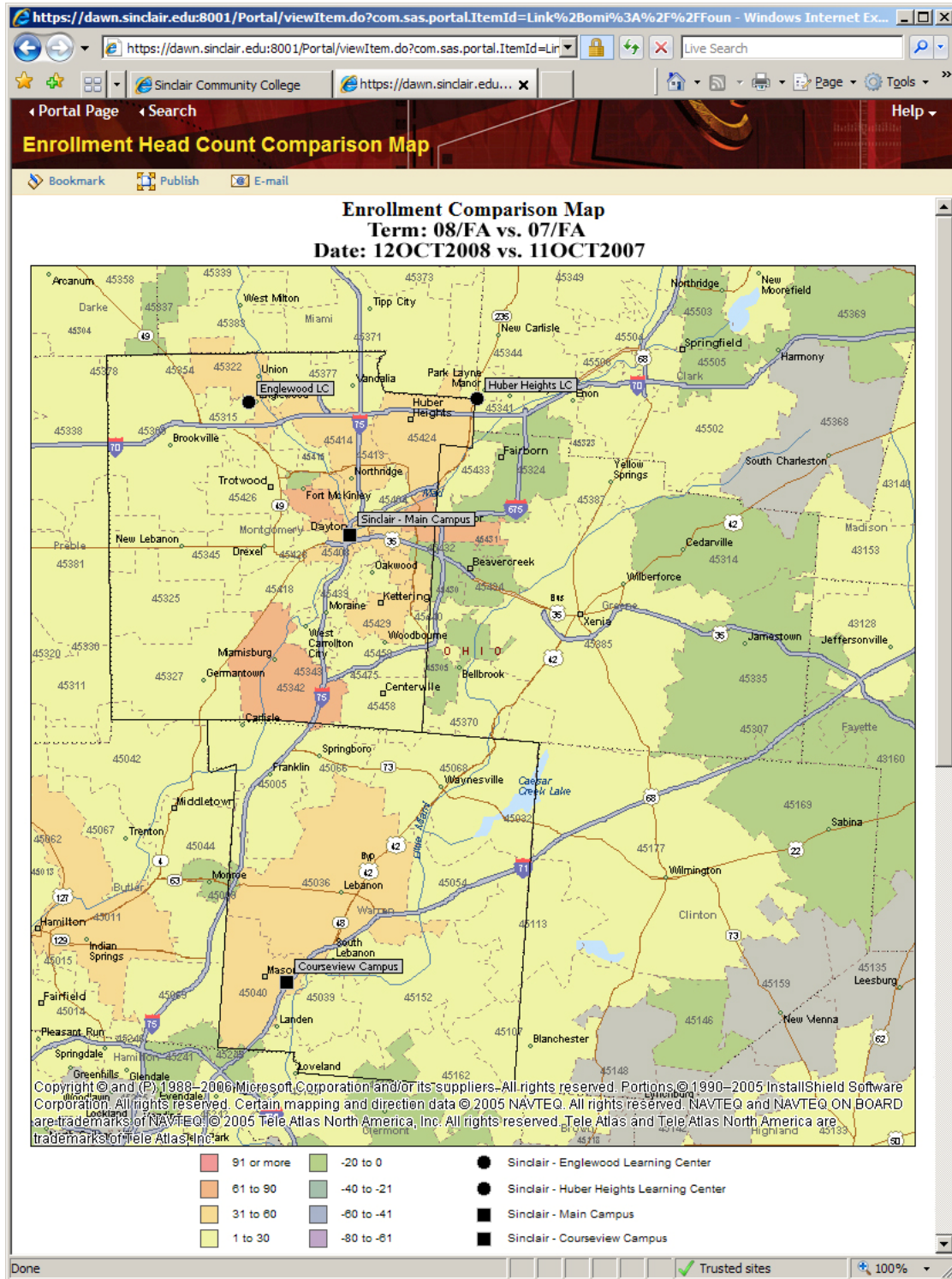
| | A | B | C | D | E | F |
|----|-------|--------------|--------------|------------|---|---|
| | Zip_5 | CY_Headcount | PY_Headcount | Difference | | |
| 1 | | | | | | |
| 2 | 00000 | 19 | 26 | -7 | | |
| 3 | 06040 | 1 | | 1 | | |
| 4 | 11106 | 1 | | 1 | | |
| 5 | 16248 | 1 | | 1 | | |
| 6 | 17870 | 1 | | 1 | | |
| 7 | 22066 | 1 | | 1 | | |
| 8 | 25501 | 2 | | 2 | | |
| 9 | 25510 | 1 | | 1 | | |
| 10 | 25524 | 1 | | 1 | | |
| 11 | 25541 | 1 | | 1 | | |
| 12 | 25559 | 1 | | 1 | | |
| 13 | 25564 | 1 | | 1 | | |
| 14 | 25704 | 1 | 1 | 0 | | |
| 15 | 28540 | 1 | | 1 | | |
| 16 | 29403 | 1 | | 1 | | |
| 17 | 30240 | 1 | 1 | 0 | | |
| 18 | 31522 | 1 | | 1 | | |
| 19 | 32607 | 1 | | 1 | | |
| 20 | 33763 | 1 | | 1 | | |
| 21 | 38583 | 1 | | 1 | | |
| 22 | 40010 | 1 | | 1 | | |

The detailed data consists of Postal ZIP codes, current year head count (CY) and previous year head count (PY) and the difference between the two. Looking at the columns of numbers yields little useful information. Given the limited value presented here you decide to graph the results which look like this...



Again the results make it rather difficult to determine how this data can help you formulate a strategy for the enrollment issue. At this point you decide that the data may be more understandable if it were overlaid onto a map. You submit a request to your friendly Research Analytics and Reporting Department and here and they provide you with the following...

MAPPED DATA GENERATED FROM A SAS DATA SET AND USING MICROSOFT MAPPOINT SOFTWARE



Mapping the data by ZIP code clearly shows the breakdown of activity. At a glance you can see each campus location and which areas have positive, neutral, or negative activity. With this information you can now target your efforts. You decide to formulate three marketing plans, one each to deal with each activity type shown on the map. So, how was this mapping accomplished? Let's take a look...

STEP-BY-STEP INSTRUCTIONS – MAPPING GEOGRAPHIC DATA FOR PRESENTATION ON THE SAS PORTAL

Section 1- The Manual Process: Steps for quickly creating an HTML Map file showing data broken down by ZIP code and making the results available on the SAS Information Delivery Portal

You will need SAS 9.1 to generate the data and Microsoft MapPoint to create the HTML map file.

To gauge whether or not mapping the data will be useful, start by running the process manually; if the results are useful check out the steps for fully automating the creation, update, and publishing process.

Step 1. Create ZIP Code data and save as an MS Excel spreadsheet or comma separated value (.csv) file. Sample SAS code is provided in the Code Samples Section of this paper to generate a csv file.

Step 2. Open Microsoft MapPoint and follow the Import Data Wizard instructions under Data on the menu bar. Make sure and change the file type to .csv if that is the file type you created. Browse to the data file location and finish the import process. You will need to zoom into the map area where your results are to view them. You can tweak certain aspects of the MapPoint file like – the legend type and color, map size, and others. Make changes as necessary until you get satisfactory results.

Step 3. Complete the MapPoint process by clicking on *Save as Web Page* under *File*. Save the file to a location on a web server that is accessible to the outside world. Test to make sure you can get to the file by entering the URL for the location into a browser window. To add the map to managed content, log onto the SAS Information Delivery Portal (you will need administrative rights) create a link to the HTML map file URL on the web server. Portal users should be able to access the file.

Section 2 – Automating the Process: If your data doesn't change often you can run the process manually to update the HTML map file. If however you have data that changes often, say on a daily basis, you may want to consider automating the process. This section deals with doing just that. It illustrates the steps necessary to automate map updates and making it accessible on the SAS Information Delivery Portal.

As mentioned before you will need to have a working knowledge of your network infrastructure and have administrative rights to certain areas of your network, or get someone who has these capabilities to perform the tasks. Automating the data mapping process can be accomplished by writing some SAS Code, a little VB Scripting, and using the operating system Task Scheduler. Note – make sure to keep track of where the various files are kept.

Step 1. Start by writing SAS code to create your data set, this is assuming of course that the data exists in a SAS data library. Add the code segments in the Code Sample Section of this paper to export the data and update the HTML Map file. Save the file.

Step 2. Create the VB Script file to refresh the data. Copy and paste the VB Script Code found in the sample code section of this paper into a text editor and save the file to the same location as the other files. Be sure to change the filename designations to the one you are using and make the file extension .vbs.

Step 3. Run the SAS Code manually. This is an intermediate step and can be omitted once the process is up and running. Make sure that the code runs without error.


Step 4. Create link on the SAS Information Delivery Portal to the HTML map file location. Make sure that the file is updated after the SAS code has run.

Step 3. Open the Task Scheduler on the computer where the files are stored and schedule a time to run the SAS code file you created. Once this process has been established the map file will automatically update whenever the scheduled task runs.

Section 3 - Animate your map: A series of HTML Map files can be combined to form an animated graphic file, which can add another dimension to your data. By extracting the graphic image from the HTML map files and using Adobe Photoshop ImageReady you can animate the maps to see the effects of time.

Step 1 – Save a series of the graphic images (the map part of the HTML file) to a common location. You can right click on the graphic map while the file is open in a browser and do a Save As to accomplish this.

Step 2 – open Adobe Photoshop and copy each of the graphic files to the project. They will be on separate layers.

Step 3. Switch to edit with ImageReady making sure that the Animation and Layers palettes visible. Most of the activity for the animation process will take place in the Animation window. Click on the Duplicate Animation Frame button  to start adding the graphic images and adjust the duration time setting for each frame as needed. A play back button is provided to preview the animation as you work.

Step 4. When finished use the Save Optimized As option and choose Images Only (*.gif), note where the file is saved. Open the .gif file with a web browser to replay the animation.

Step 5. Save the file on the web server in the same location as your HTML map file and create a link on the SAS Information Delivery Portal to view the file.

CONCLUSION

Viewing geographic data on a map can increase understanding. Once you have established the process it is relatively easy to apply many applications.

APPENDICIES

APPENDIX A - CODE SAMPLE SECTION

```
/* SAS Code sample for generating ZIP code .csv data file used with the manual process in the mapping exercise */
```

```
data ZIPdata;

infile datalines dlm=',';
input ZIPCode $ CY_Headcount
      PY_Headcount Difference;
datalines;
45401,10,17,-7
45402,189,158,31
45403,184,141,43
45404,77,76,1
45405,297,257,40
45406,345,282,63
45407,2,3,-1
45408,123,101,22
45409,102,98,4
45410,228,192,36
;
run;

ods csv file="ZIPdata.csv";

proc print data=ZIPdata noobs;
run;

ods csv close;
```

EXPORT DATA

```
/* add this code to the SAS program used by the automated process to create your data set. It will export the data file for use with the MapPoint software */
```

```
Proc Export Data = WORK.filename outfile="E:\maps\filename.xls" replace;
RUN;
```

DATA REFRESH

```
/* add this code to the SAS program used by the automated process to create your data set. SAS code calls VB Script file which runs MapPoint to refresh and export the map */
```

```
filename cmd pipe "call E:\maps\filename.vbs";
data _null_;
  infile cmd;
  input;
  put _infile_;
run;
```

Stand Alone VB Script file Code - save this code with a .vbs extension

' VB Script code to refresh the map

Option Explicit

MAIN

Sub MAIN()

Dim objApp

Dim objMap

Dim objDS

Dim Map

Dim MapCount

Dim IntLoop

Set objApp = CreateObject("MapPoint.Application")

'Open the MapPoint project file

Set objMap = objApp.OpenMap("E:\maps*filename*.ptm")

'Refresh the Data (you will need to identify the correct "Item" for it to display properly)

Set objDS = objApp.ActiveMap.DataSets.Item(5)

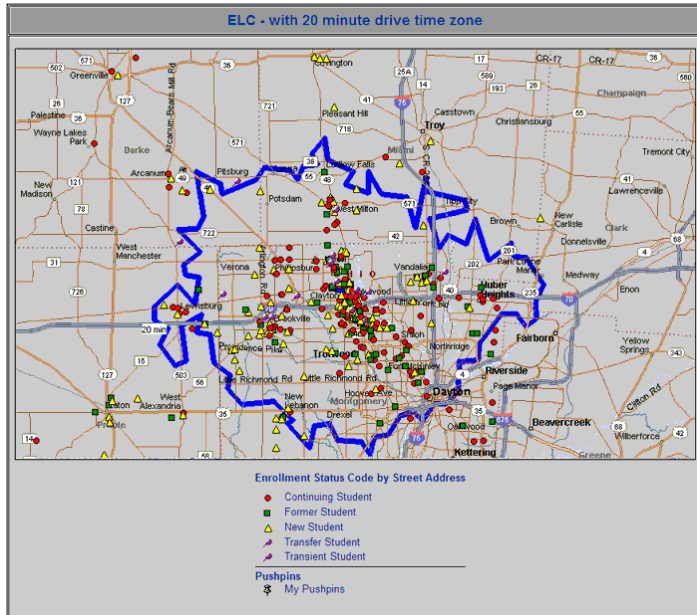
objDS.UpdateLink

objMap.Save

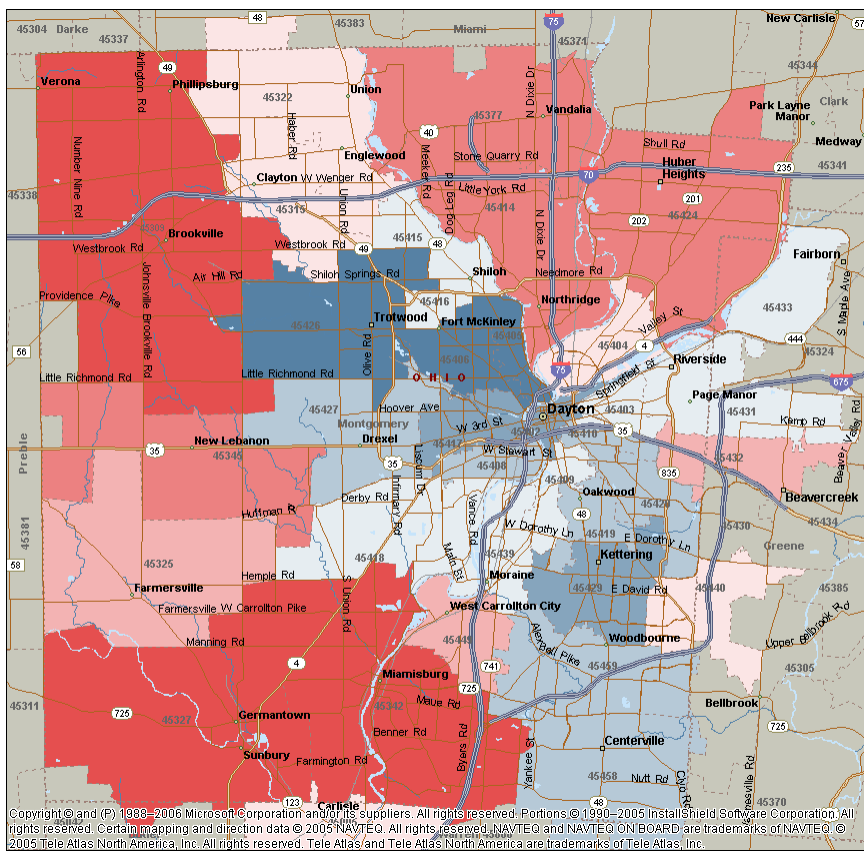
Set objApp = Nothing

End Sub

APPENDIX B – ADDITIONAL MAPPED DATA EXAMPLES



Map of student enrollment showing 20 minute drive time



Election results showing relative strength of voting activity

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Paul Ciarlariello
Sinclair Community College
444 West Third Street
Dayton, OH 45402
Work Phone: 937 512-2213
Fax: 937 512-2049
E-mail: paul.ciarlariello@sinclair.edu
Web: www.sinclair.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.