ABSTRACT
Have you ever been intrigued by a table, a chart, or a map and wanted to explore deeper into the data behind it? Have you ever produced a table, a chart, or a map, and then been asked for another just like it, but with changes – looking at a different population, subpopulation, or geographic area? If you're working in a hard-copy environment, you can either modify the program to handle each new request for additional information, or you can proactively produce a report for every possible level of the data. If there are numerous requests, the former course of action wastes your time; if there's little interest in delving deeper into the data for a particular subcategory, then taking the latter course wastes a great deal of resources. As programming becomes increasingly web-oriented, a third option opens up: adding exploratory capabilities to programs and output. This course allows you to anticipate the end user's needs for further analysis of a particular item of data and to meet that need with only the click of a mouse.

INTRODUCTION
This paper will look at two web-based techniques for allowing users to delve into your output: pop-ups (boxes containing additional information that pop up whenever a mouse is dragged over the appropriate area), and drill-downs (output areas that are linked to other web pages or that execute a program to provide additional details). To provide a concrete foundation for our illustrations, we will be looking at an application that uses data from higher education institutions in the United States, as well as population data from the 2000 census. From this data, we will be able to generate reports, tables, and maps, all with the capability of exploring deeper into the details of our data. Only the initial web page, where the user enters the most basic information about the desired output, is static; the remaining web pages are created dynamically by our program.

THE DATA
MAPPING DATA
Like most applications that deal with map graphics, our program needs two types of data – location and substantive. SAS® Institute supplies the first type through two datasets that ship as part of SAS Base: MAPS.COUNTIES and SASHELP.ZIPCODE. The first gives us the unprojected coordinates of the boundaries for all 3,000+ counties in the United States, while the second pinpoints the centroid of each ZIP code area. We will use these centroids to determine the location various higher education institutions, with each institution placed at the centroid for its ZIP code. While it might be possible to locate each institution more precisely, the benefit for our program would not be outweighed by the effort.

SUBSTANTIVE DATA
Our application allows you to map and report three different types of substantive data – total population, total fall headcount, or degrees conferred.

- Total Population comes from the U.S. Census Bureau’s 2000 headcount at the census tract level. Tracts are intended to represent between 1,000 and 8,000 persons, and each tract is contained within a county’s borders. The tract boundaries are not used in this application, but the tract centroid is; thus, the entire population of a tract is assumed to be located at its centroid.

- Total Fall Headcount is supplied by the U.S. Department of Education's Integrated Postsecondary Education Data System (IPEDS) fall enrollment data. All educational institutions whose students receive federal financial aid are required to supply data through this system. For our purposes, we have excluded all proprietary schools and all institutions which do not offer at least an associate’s degree. Some institutions, generally religious colleges, choose to refuse federal funds, and so exclude themselves from the reporting requirements.

- Degrees Conferred information also comes from the IPEDS collection, and covers an entire academic year, from July 1st through June 30th. The exclusions applicable to the fall headcount are also in effect for this data. Degrees are further divided by CIP codes (Classification of Instructional Programs), a six-digit number commonly used in education circles to identify a particular program or major field. Fine Arts in general would be designated by a CIP of 50.xxxx, while Art History has a CIP code of 50.0703.
STEPPING THROUGH THE APPLICATION
THE STATIC WEBPAGE

The static webpage contains some basic required fields and optional fields that help shape the remainder of our application. The initial required field of the center of the first area will help determine the next set of questions our program will ask to aid in the customization of the output, while the initials will be used to aid in distinguishing one set of hardcopy output from another.

Since this is a mapping tool, one of the most important pieces of information is where to begin the mapping; you have the choice of selecting by named places within a state, or if you already know the specific location, by ZIP code. A second optional area allows you to do a comparison of data between two points, a particularly useful function for economic recruitment. Not surprisingly, North Carolina is the default destination for the first area; we’ll specify the ZIP code for Chapel Hill as the center for this example. Since we would like to do a comparison with Portsmouth, we’ll begin by selecting “Virginia” from the pull-down menu of states for the optional second area.

MORE CHOICES
When we submit our choices from the initial menu, we’re presented with detailed options, beginning with the distance from our selected locations. Circles can be drawn around the center points at intervals from 25 miles to 1,500 miles. Since we specified a ZIP code for our first location, the only option we have for that location is “Chapel Hill”. We select our second location from a pull-down menu of locations in our previously-specified state of Virginia; the default choice is the alphabetically first town, city, or educational institution, except in North Carolina, where the default is “Research Triangle Park”. The final mapping choice is whether to map the areas separately or together, allowing you to determine how detailed the maps are; if no second location is specified, this choice is not presented. Since we’re looking at locations in adjoining states, both of them will fit nicely on a single map; if we were comparing Chapel Hill to San Diego, we’d probably prefer a separate map for each area.
Having specified our mapping locations and options, we turn our attention to the data to be reported. Given the title of this paper, naturally we’ll select degrees conferred in the fields of Art and Art Studies (CIP code 50.0701) and Art History, Criticism and Conservation (50.0703) as well as the related ones of Arts Management (50.0704) and Fine Arts Studies (50.0799). Since the number of degrees awarded annually in these fields is likely to be relatively small, we’ll boost our output by selecting all the available years. We’ll also look at total fall headcount enrollment for the most recent year; since IPEDS doesn’t collect enrollment information for all CIP codes, can’t examine the number of students in the educational pipeline to become art historians. While there’s no option for it shown, we’ll also take a look at population patterns from the 2000 census within the specified circles.

OUTPUT #1 – DRILLABLE MAPS
The maps at the heart of our output are generated by relatively simple PROC GMAP code shown below. The ‘FLAG’ variables are used to control whether each of our three substantive variables produces one map or separate maps for each of our two areas.

```
PROC GMAP map=work.map(where=(flag=&flg))
data=dum1(where=(0<=degrees and flag&flg=1)) all;
id state county;
format degrees degfmt.;
choro degrees / discrete coutline=ligr cempty=white legend=legend1 html=href
annotate=work.anno(where=(flag&flg=1));
```
This code produced the map shown below from the options we selected earlier. While Portsmouth doesn’t have any institutions of higher education, its neighbor Norfolk does. From the legend below the map, we know that Norfolk produced 123 or more graduates in our selected fields (the breakpoints for the formats being determined by the data). While we could generate the same map in printed form, the interactive nature of this application allows us to explore it further simply by positioning the cursor.

The pop-up box is a product of the ‘HTML=HREF’ parameter in our PROC GMAP code; this parameter also allows us to drill down into the data underlying a particular county on our map. The variable ‘HREF’ is simply a character variable made of up a series of HTML tags passed as text strings; the tags are associated with a particular portion of the map, so that when the cursor is positioned over that area, the HTML tags are activated. To create the pop-up box, the crucial tag is called ‘ALT’. For our maps, we will first build a character variable called ‘ALT’, and then attach it to the other HTML tags when we build ‘HREF’.

Since we have three sources of data – degrees conferred, fall enrollment, and population – and (potentially) multiple degree observations for each county, we’ll use the RETAIN statement along with FIRST.variable and LAST.variable processing to reduce our data to one observation per county within our range. After we merge this dataset onto one containing all the counties that we intend to map, we’ll build a truncated ‘ALT’ variable for the counties outside the range, and then build ‘HREF’ for all the counties on our map. If you position the cursor on a county within the range, you’ll see a display similar to the one above; if you select a county outside the 100-mile circle, you’ll see only the county and state name.
data dum10 dum11 dum12;
set dum0;
by flag state county;
length alt $1500;
retain alt ' ' tot89 tot99 0;

/* Initialize variables for each county */
if first.county then do;
  alt=trim(countynm) || ' ' || trim(fipstate(state));
  tot89=0; tot99=0;
end;

/* 89.xxxx is a fake cipcode for totaling all the specified CIPs */
/* 99.0000 is the cipcode for all degrees granted at an institution */
/* '0d'x forces the output to a new line; '09'x is a tab character */
if cipcode='89.xxxx' then tot89=sum(of tot89,degrees);
else if cipcode='99.0000' then tot99=sum(of tot99,degrees);

/* Degree data will sort so that our totaling CIPs come last after */
/* all the specific cipcodes we have selected */
if first.county and 0<degrees then
  alt=trim(alt) || '0d'x || 'Degrees Conferred';
if cipcode ne '99.0000' and cipcode ne '89.xxxx' then
  alt=trim(alt) || '0d'x || 'Degrees in ' || trim(cipcode) || '09'x ||
  put(degrees,comma15.);
else if cipcode='89.xxxx' then
  alt=trim(alt) || '0d'x || 'Total Specified Degrees' || '09'x ||
  put(degrees,comma15.);
else if cipcode='99.0000' then do;
  if degrees then
    alt=trim(alt) || '0d'x || 'Total Degrees Conferred' || '09'x ||
    put(degrees,comma15.);
  else
    alt=trim(alt) || '0d'x || 'Total Specified Degrees' || '09'x ||
    put(degrees,comma15.);
end;

/* Enrollment data is already summarized to one observation per county */
/* The different degree levels are represented by different variables */
if enroll then
  alt=trim(alt) || '0d'x || 'Undergraduates' || '09'x ||
  put(ugrad,comma15.) ||
  '0d'x || 'Graduates' || '09'x ||
  put(grad,comma15.) ||
  '0d'x || 'First Prof.' || '09'x ||
  put(fprof,comma15.) ||
  '0d'x || 'Total Fall Headcount' || '09'x ||
  put(enroll,comma15.)
;
/* Population is also already summarized to one observation per county */
if pop then
  alt=trim(alt) || '0d'x || 'Total Population in Range' || '09'x ||
  put(pop,comma15.);
end;

/* Keep the final observation for each county */
/* If 99.0000 was selected, we use that for the number of degrees. */
/* If not, we go with the total of the selected cipcodes. */
if last.county;
  degrees=max(of tot89,tot99);
  if flag=0 then output dum10;
  else if flag=1 then output dum11;
  else if flag=2 then output dum12;
run;
/* Attach the ALT variable to the mapping data */
data dum10;
merge dummy(in=in1)
dum10(in=in2);
by state county;
if flag1=0 then flag1=.;
if flag2=0 then flag2=.;
flag0=1;
array one(*) degrees enroll pop ugrad grad fprof;
do i=1 to dim(one);
   if one(i)=. then one(i)=0;
end;
drop i;
length href $ 1500;
countynm=put(compress(put(state,z2.) ||
 put(county,z3.)),$cnty.);
/* Build an ALT & HREF for all the counties outside the range. */
/* This HREF is not drillable. The ALT information (county */
/* and state name) appears when you mouseover the county. */
if in2=0 then do;
   alt=trim(countynm) || 
       trim(fipstate(state));
   href='alt="' ||
     trim(alt) || 
     "';
end;
/* Use the ALT to build HREF for all the counties inside the range. */
/* This HREF passes parameters to the program PUB.MAP1.SAS for */
/* drilldown. When you mouseover a county, the ALT information */
/* appears. */
else if in2 then
href='href="' ||
   "&action=" ||
   '?_service=default&_program=pub.map1.sas' ||
   '&_debug=' || trim("&_debug") ||
   '&_inits=' || trim("&_inits") ||
   '&loop=' ||
   '&_state=' || put(state,z2.) ||
   '&_county=' || put(county,z3.) ||
   '&cips=' || trim("&cips") ||
   '"' ||
   'alt="' ||
   trim(alt) || 
   '"';
run;

OUTPUT #2 – DRILLABLE REPORTS
While the 'ALT' part of the 'HREF' variable makes the pop-up box possible for all the counties on our map, the HREF tag and its associated parameters play an equally vital role in making our map drillable. If the number of possible combinations of data were relatively limited, we could generate static web pages and link those to our 'HREF'; since we'd like accommodate each county in the United States and each CIP code, building a dynamic web page is preferable.

When you click on one of the counties within our 100-mile circles, the program MAP1.SAS executes with the parameters 'INITS', 'LOOP', 'STATE', 'COUNTY', and 'CIPS'. While moving the cursor over Norfolk produces the pop-up box with its summary information, clicking on that location will allow us to delve deeper into the county's higher education statistics, as shown below.
<table>
<thead>
<tr>
<th>INSTITUTION NAME</th>
<th>Fall Enrollment</th>
<th>Degrees Awarded in CIP Codes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Virginia Medical School</td>
<td>699</td>
<td>645</td>
</tr>
<tr>
<td>Johnson &amp; Wales University-Nor</td>
<td>619</td>
<td>635</td>
</tr>
<tr>
<td>Norfolk State University</td>
<td>6,668</td>
<td>6,731</td>
</tr>
<tr>
<td>Old Dominion University</td>
<td>18,969</td>
<td>19,627</td>
</tr>
<tr>
<td>Old Dominion University</td>
<td>20,184</td>
<td>21,379</td>
</tr>
<tr>
<td>Virginia Wesleyan College</td>
<td>1,421</td>
<td>1,412</td>
</tr>
<tr>
<td>DEGREE LEVEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associates</td>
<td>20,303</td>
<td>21,914</td>
</tr>
<tr>
<td>Baccalaureate</td>
<td>20,131</td>
<td>20,418</td>
</tr>
<tr>
<td>Master's/Doctoral</td>
<td>7,225</td>
<td>7,500</td>
</tr>
<tr>
<td>First Professional</td>
<td>411</td>
<td>417</td>
</tr>
<tr>
<td>INSTITUTION TYPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Sector</td>
<td>25,637</td>
<td>26,648</td>
</tr>
<tr>
<td>Private Senior</td>
<td>2,120</td>
<td>2,057</td>
</tr>
<tr>
<td>Public Junior</td>
<td>20,184</td>
<td>21,379</td>
</tr>
<tr>
<td>Private Junior</td>
<td>619</td>
<td>625</td>
</tr>
<tr>
<td>CARNEGIE CLASS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 - Doctoral / Research Extensive</td>
<td>18,969</td>
<td>19,627</td>
</tr>
<tr>
<td>21 - Masters I</td>
<td>6,668</td>
<td>6,721</td>
</tr>
<tr>
<td>31 - Baccalaureate -- Liberal Arts</td>
<td>1,421</td>
<td>1,412</td>
</tr>
<tr>
<td>40 - Associates</td>
<td>20,303</td>
<td>21,914</td>
</tr>
<tr>
<td>52 - Medical Schools</td>
<td>619</td>
<td>625</td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28,717</td>
<td>29,885</td>
</tr>
<tr>
<td>Female</td>
<td>19,843</td>
<td>20,434</td>
</tr>
<tr>
<td>RACE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Resident Alien</td>
<td>1,200</td>
<td>1,317</td>
</tr>
<tr>
<td>African-American</td>
<td>15,684</td>
<td>16,324</td>
</tr>
<tr>
<td>Native American</td>
<td>302</td>
<td>276</td>
</tr>
<tr>
<td>Asian-American</td>
<td>2,251</td>
<td>2,393</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1,201</td>
<td>1,391</td>
</tr>
<tr>
<td>White</td>
<td>27,827</td>
<td>27,622</td>
</tr>
<tr>
<td>Unknown</td>
<td>95</td>
<td>1,086</td>
</tr>
<tr>
<td>CIP - DISCIPLINE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50.0701 - Art/Art Studies, General</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>50.0703 - Art History, Criticism and Conservation</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>99.0000 - Award Level Total</td>
<td>48,560</td>
<td>50,319</td>
</tr>
</tbody>
</table>

**Educational Profile for Norfolk City County VA**

**Census 2000 Population: 234,403**

---

**UNC-OP SDA Maps: GG038 G25JUN05**

Printable (Adobe PDF) version of reports
As you might guess by now, the table above uses an ‘HREF’ variable as well, to provide for additional drilling. Clicking on Norfolk State’s link produces the table below, profiling that institution, detailing the degree levels available and breakdowns by gender, race, and discipline. If, instead we selected CIP 50.0703 – Art Criticism, History, and Conservation, the program would produce a similar table detailing the degree levels, gender, and race of students in that discipline, as well as information about the types of institutions offering that major, their Carnegie classification, and the state in which they are located.

### Profile of Norfolk State University

*A Public Senior Institution in Norfolk, Norfolk City County, Virginia*

*Carnegie Class: 21 - Masters I*

<table>
<thead>
<tr>
<th>Degree Level</th>
<th>Fall Enrollment</th>
<th>Degrees Awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACADEMIC YEAR</td>
<td>ACADEMIC YEAR</td>
</tr>
<tr>
<td>Certificate</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Associate</td>
<td></td>
<td>150</td>
</tr>
<tr>
<td>Baccalaureate</td>
<td>5,880</td>
<td>1,538</td>
</tr>
<tr>
<td>Master’s/Doctoral</td>
<td>778</td>
<td>689</td>
</tr>
<tr>
<td><strong>GENDER</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4,345</td>
<td>1,460</td>
</tr>
<tr>
<td>Female</td>
<td>2,533</td>
<td>1,201</td>
</tr>
<tr>
<td><strong>RACE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Resident Alien</td>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>African-American</td>
<td>2,953</td>
<td>1,592</td>
</tr>
<tr>
<td>Asian American</td>
<td>49</td>
<td>24</td>
</tr>
<tr>
<td>Hispanic</td>
<td>35</td>
<td>24</td>
</tr>
<tr>
<td>Native American</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>White</td>
<td>535</td>
<td>370</td>
</tr>
<tr>
<td>Unknown</td>
<td>62</td>
<td>194</td>
</tr>
</tbody>
</table>

**CIP 01.0301 - Economics, General**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Certificate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Associate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**CIP 01.0301 - Political Science and Government, General**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Certificate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Associate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
OUTPUT #3 – NON-DRILLABLE MAPS AND REPORTS

While the SAS Output Delivery System (ODS) allows us to build interactive applications with a minimal knowledge of the actual HTML behind them, it also provides options for producing hardcopy versions of our maps and report. The maps and the final reports can be replicated in RTF and PDF formats (among others) by the same procedures that generate our browser content. Since the drillable reports make more extensive use of HTML-specific variables, their alternative versions require invoking the REPORT or TABULATE procedure a second time, after the HTML destination has been closed.

CONCLUSION

What’s next? In the current incarnation, our final report is dynamically produced, but static – we can’t drill any further. This does not have to be the case; since many of the categories are the same as those in the previous report, we could just as well add another round of ‘HREF’ variables to our data, producing a perpetually self-referencing loop. For those of us who can spend an enjoyable hour bouncing from one dictionary entry to another (forgetting in the process the word we were looking up originally), the idea of discovering that Obscure State University offers Art History, then drilling into Obscure State’s data to find that it offers a basket-weaving major has great appeal.

As with many, if not most, SAS programming tasks, the bulk of the work lies is preparing the data. The tools for analysis and presentation, such as ODS, SAS/IntrNet, GMAP, REPORT, and TABULATE, are already in place and waiting to be used. I hope that you will use some of the techniques discussed here as a jumping-off point to build applications to explore your own data.

REFERENCES

SUGI 29: Paper 090-29
Pop-Ups, Drill-Downs, and Animation
Mike Zdeb

SUGI 30: Paper 137-30
Stretching the Bounds of SAS/GRAPH Software
Mike Zdeb and Robert Allison

CONTACT INFORMATION

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

Keith J. Brown
UNC-Office of the President
P.O. Box 2688
Chapel Hill, NC 27515-2688
Work Phone: 919.962.4547
Fax: 919.962.4316
Email: kjb@northcarolina.edu
Web: http://www.northcarolina.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.