ABSTRACT
UNC General Administration's Division of Institutional Research (UNCGA IR) collects student-level data on enrollment, graduation, courses, grades, financial aid, facilities, and personnel from the 16 public universities. The university system is transitioning from data collection and processing of flat files submitted via SAS/IntrNet® to an Oracle® Data Mart that retrieves the data from the campuses' source systems. After collecting the data, either via flat files or data mart, SAS® data sets are created and used for analysis and reporting for public and private universities, policy makers, elected officials, and the general public. After a complete turnover in programming staff, UNCGA IR decided to make some major upgrades and improvements to a server and system that were quickly becoming unsupported and outdated. This presentation covers the following topics: a background of our office functions, including our SAS environment; the decision-making process to upgrade to a SAS® Business Intelligence and SAS® Visual Analytics environment, using SAS® 9.3 and eventually SAS® 9.4; the factors in determining which processes should be moved to a PHP interface, to SAS® Stored Processes, and to SAS Visual Analytics (including whether the user was internal or external, whether the user had access to secure data, and how much time was involved in making the changes); the process and challenges of implementing the changes; and the lessons learned. This presentation has a universal audience but is specifically intended for anyone who has gone through or plans to go through a SAS upgrade, has inherited spaghetti code and used it to do their job, or has used SAS/IntrNet®, SAS® Visual Analytics, SAS® Stored Processes, or a PHP interface to initiate SAS® programs.

INTRODUCTION
The Institutional Research (IR) Division in the University of North Carolina General Administration collects data from its sixteen universities and one high school, provides reports for the general public and policy makers, and handles federally mandated reporting. Many of IR’s data collections and methods of sharing information with the public were through the use of SAS/IntrNet. SAS programs were all written using an outdated text editor and were run on the command line. Data existed on two separate servers with two different versions of SAS. Documentation was non-existent. With complete programming staff turnover and many changes needing to be made to the current system, IR requested the help of SAS consultants to help define our needs and upgrade our system with up-to-date technology. This paper discusses the transition and decision making process to move from SAS 9.1.3 and 9.3 environments heavily dependent on SAS/IntrNet to a SAS 9.4 environment using SAS® Enterprise Guide®, SAS Visual Analytics, a PHP interface, and SAS Stored Processes.

BEGINNING PROJECT STATUS AND OFFICE FUNCTIONS
The IR office is comprised of a data manager, three analysts, and five SAS programmers to maintain our data collection system and handle reporting. All new staff was responsible for running spaghetti code on outdated systems. The former team operated solely in a UNIX environment, initiating SAS from the command line and coding in a user-unfriendly text editor that was developed in 1974, and all of this was saved on a server that was no longer supported. Data, programs and format catalogs were strewn all over two servers with many of the essential data sets being stored on individuals’ home drives—and these individuals were no longer employees. The new IR programmers had experience programming in the SAS® Enhanced Editor or in a modern text editor with full use of keyboard shortcuts. The system administrator gave IR access to Windows® based SAS 9.3 but format catalogs could not be used because of the cross-environment errors.
Campuses accessed the server through the internet, by confirming their campus account credentials. They could SFTP flat data files to their folders on the server then access a SAS/IntrNet page that allowed them to submit their flat data file, check the file for errors, build a series of reports for data validation, and produce a SAS data set. This process of submitting data was about to become obsolete for the UNC system campuses, since IT was developing an Oracle Data Mart that would retrieve data directly from the campus source data management system. However, private universities and community colleges also reported data to IR using SAS/IntrNet and people outside the university system used SAS/IntrNet and a series of dropdown boxes to run reports.

**HOW TO DEFINE NEEDS**

To define our needs for a new operating environment, we considered the users that will need access, the reports or data they will need to access, and our IT-controlled secure server constraints. If you have a SAS administrator, then they can help define the system requirements. Since UNCGA does not have a SAS administrator, we had SAS consultants assist with our system development. After meeting multiple times with the consultants and our system administrators to discuss our user and required business processes, we defined the following requirements for our new system:

1. A secure environment for data collection,
2. Publicly accessible environment for reports,
3. Connection to our Oracle® Data Mart, and
4. Seamless changes for the campuses.

**USER REQUIREMENTS**

The first step in identifying the structure of our new system was to define our user requirements and business practices. IR has both internal and external users. Internal users are defined as people within IR who need access to all data and reports. As programmers with full access to data, we have access to data on every student who has ever applied to a UNC University. By moving IR’s data collection of faculty, student, course, grade, and financial aid to an Oracle Data Mart that retrieves the data from the campus source system, an additional requirement for IR’s new system was the ability to access this data from SAS to use for analysis and reporting.

All of this data needs to be stored on a secure environment, but we also need to be able to meet the expectations of the audience for our reports including the federal government, state legislature, Board of Governors, and the general public in a timely manner. Ad hoc reports must be completed quickly, while other information can be in an environment where it can be easily shared with the public. External users are defined as people outside of IR who need access to data and reports. External users can further be broken down into users with secure access to individually identifying data and then the general public who would not need access directly to data, but would need the ability to create aggregate reports. In this example, external users with secure access include the campus users who need to submit data, retrieve data, and view and download reports. IR did not want to recreate the wheel for these users since their data collections were already in the process of changing from flat file submissions to the Oracle based retrieval system. A system that mimicked the SAS/IntrNet data collection system from SAS 9.1.3 was acceptable to be used until these processes could be moved to a Stored Process interaction with the Oracle Data Mart. For the external users who only need access to reports, like the Board of Governors or the general public, IR needed something to replace the SAS/IntrNet beginning page of dropdown boxes leading to static data tables with more visually appealing dynamic reports.

**PROPOSED SOLUTIONS**

Taking into consideration the SAS consultants’ recommendations, we decided to move from SAS 9.1.3 to SAS 9.3 before moving to SAS 9.4 because of the age and complexity of most of our code. The Business Intelligence platform met most of our needs. Enterprise Guide would allow us to access our data sets and format catalogs without errors. SAS/IntrNet could be replaced with the Stored Process Server, and our public reporting could be moved to the latest version of SAS hosted SAS Visual Analytics. We knew our options included:
• Continuing with SAS/IntrNet,
• Moving some business processes to the Stored Process Server,
• Converting public reports to interactive SAS Visual Analytics, and
• Driving around the state collecting data on paper or via flash drive and delivering physical reports to external users.

We know that driving around the state collecting data and distributing reports is not a viable solution, although, at times during this process, it did seem like a simpler solution.

PREPARING FOR THE CHANGE TO SAS 9.3

In order to prepare for the upcoming changes, we needed to complete several steps:

1. Programmers needed to document our SAS/IntrNet Processes.
2. IT needed to install a test environment where programmers could update code without it affecting our work in production.
3. Programmers needed a plan for testing in the new environment.
4. IT and programmers needed to work together to develop a plan for moving all of the data and programs to the new environment.
5. Programmers attended training classes on the new SAS products they would be using.

DOCUMENTING SAS/INTRNET

Documenting the current SAS/IntrNet processes proved to be quite challenging. Some of the programmers had exposure to HTML which helped, but the spaghetti code was difficult to decipher. Campuses used SFTP to move flat files to our servers and then SAS/IntrNet would initiate the SAS programs to process those flat files for error checks, validation reports, and create a SAS data set for analysis purposes. For this process, campuses would start on this html page, select items from the drop down boxes that would be passed to SAS and then run a series of programs based on the selection of the radio buttons. The page edit.html can be seen in Display 1.

Display 1 HTML Page for Campus Data Submission

Campuses enter their FICE code, IPEDS Unit ID, email address, then select the year and term from the drop-down boxes. The radio button on the left indicates which program SAS should run, and the submit button starts the process.

Figure 1 shows the map of a data submission process. Campuses begin on index.html, and after logging
in with a campus ID and password allowing them access to the secure server, they can access edit.html. This is the page from *Error! Reference source not found.*, where the campus enters their FICE code, IPEDS ID, email address and then selects a flat file to submit and process. The submit button triggers scfedit1.sas to run which includes three programs head.sas, edsecure.sas, and mailto.sas. Head.sas assigns some macro variables to be applied to LIBNAME statements and ODS OUTPUT statements; edsecure.sas checks the FICE and IPEDS Unit ID for a match; and mailto.sas displays errors if any exist. Scfedit1.sas also creates a new SAS program called campus.scf.sas which will be called to run in a later step and creates a temporary webpage where the campus can select the errors they would like to check and the submit button triggers scfedit2.sas to run. Scfedit2.sas includes campus.scf.sas, which in turn includes nine programs; five of which output a variety of error reports, data validations and the SAS data set for analysis and four that build a webout page, send an email to the campus with the links to the logs and the output, and add a record to a data file that tracks when data is submitted. Clearly, this needed to be simplified.

![Figure 1 Map of Campus Data Submission Process](image)

For one data submission, a campus will automatically run at least fifteen programs by clicking the “Submit” button, this number could increase depending on the number of validation reports that are included for the data type.

**TESTING, TESTING, 1, 2, 1, 2**

We were lacking a test environment. With as complicated as our processes were, we needed a test environment to try to update our code to SAS 9.3. The main issue with the SAS code was producing the number of temporary webpages throughout each process. Strange things were happening and online documentation and tech support were no help. There was an issue with the way the forms were being created, but code followed white papers and tech support’s documentation. The temporary webpages
were being produced, but it appeared as though the submit button was not operational. One of our consultants contacted a developer who had the solution and directed us to this documentation indicating that the code and ODS were outputting information at the same time to the same place (http://support.sas.com/kb/23/970.html). The corrected code is shown below:

```sas
ods _all_ close;
   data _null_;  
   file _webout;  
   *** Some PUT statements here;  
run;
ods html body=_webout (no_bottom_matter) path=&_tmpcat (url=&_replay) rs=none;
ods html close;

***DATA Step does not write to _WEBOUT until ODS is finished writing to _WEBOUT;
   data _null_;  
   file _webout;  
   *** Other PUT statements here;  
run;
ods html body=_webout (no_top_matter) path=&_tmpcat (url=&_replay) rs=none;
   proc print data=anydata;  
   run;
   quit;
ods html close;
```

We fixed the forms but decided to move only items necessary for the data collections to continue to the new environment. We run many of our reports on an annual basis and decided to move those pieces to the new environment as the work came along.

As we adapted to testing on the new SAS BI environment, we realized we were going to face several challenges to moving our business processes including:

- User capability and knowledge,
- IT cooperation and knowledge,
- User-friendliness and escaping information overload,
- Server security and accessibility to internal and external users, and
- Other random challenges like Google crawler.

**PLANNING OUR MOVE**

By defining which current processes would be remaining in SAS/IntrNet and which would be moving to SAS Visual Analytics, we built a GANTT chart of all of our business processes defining what needed to move, who would move it, and when we would move it.
SAS CLASSES


REALITY OF THE SAS 9.3 UPGRADE

No matter how much you test and how much you prepare for change there are always unexpected hiccups to go along with the benefits. We also knew this change to SAS 9.3 was only a stepping stone and to get the full functionality we wanted, we would need to move to SAS 9.4 in a few months.

Not having a SAS administrator on staff has proved to be quite challenging. System administrators do not necessarily know how to get the Stored Process Server to work with our security needs for access via Shibboleth SSO, or how to get the SAS server to talk to the Oracle Server and vice versa. The SAS consultants were needed for much of the work with these types of issues.

SAS Enterprise Guide solved our format catalog issues, but we do not always use the metadata server.

Then we came to the realization that it was not in our best interest to try to convert all of our SAS/IntrNet code to Stored Processes. Many of these data submission processes would be obsolete in the next year or two with the implementation of the Oracle Data Mart.

Only new collections, outside of data mart would use the Stored Process server. We had a low-pressure, 8 cell data collection to test the Stored Process Server. The SAS consultants worked with us, and including working out all of the kinks with the Stored Process Server, metadata connections, security issues with Shibboleth SSO, it took about 20 hours to produce a simple Stored Process that did not meet all of our needs. The idea was that we would eventually have multiple processes so the campus would begin on a webpage similar to the ones that had used before with a radio button to select their process, and a submit button. After selecting their FICE code, IPEDS ID, and Academic Year from the drop down boxes, a campus user would enter their email address and then select the radio button for Veteran’s Benefits, and then click Submit. Then Display 2 would appear for them to enter their eight cells of data.

Display 2 Veteran’s Benefits Web Collection Form

After they clicked submit, they would get a report of their data, click another button to officially submit their data, and receive an email with a link to the output reports. We wanted anyone on the campus to be able to access this form, but because of our IT rules and server security where the stored process server sits, only campus IR users who were granted access could enter data. We also had to set up a proxy username and password to gain access to the metadata server. We also wanted the page we had designed to be the beginning page through a direct link, but users were always brought back to the
Stored Process server log in page and then had to click multiple links to get to the correct page. We quickly realized this would never be a solution for the general public to run reports. Our campus users had difficulty with screen shot instructions for how to get to the correct page.

Our Oracle Data Mart was also supposed to be able to connect with the SAS Stored Process Server. We needed data currently collected in the data mart and data from our metadata server for parallel testing reports. Because of the Shibboleth SSO issues with the Stored Process Server, we needed to upgrade to SAS 9.4 before this would work. Ironically, we ended up building this process in SAS/IntrNet.

PREPARING FOR THE MOVE TO SAS 9.4

We went through some of the same steps for the move to SAS 9.4 as we did for the move to SAS 9.3, except that we already had taken the SAS classes and documented our SAS internet processes. However we still needed:

1. IT to install a test environment,
2. IR to develop a plan for testing the new environment, and
3. IR and IT to develop a plan for moving to the new environment.

MOVING TO SAS 9.4

With the move to SAS 9.4 we expected full functionality of our Stored Process Server. Outside users should be able to have access to a usable webpage without logging in with any credentials. Users in Oracle should be able to initiate a Stored Process. Unfortunately, neither of these processes was possible. Since the Stored Process Server was located on our secure drives, the outside world would never be able to access it.

The intention was for a user, who had already logged into data mart using Shibboleth SSO, to select options from the drop down boxes for type of report, institution, term, and other selections as shown in Display 3.

Display 3 Data Mart Interface

PHP would pass these selections as cookies to the SAS Stored Process Server and run a program based on selections, then SAS would pass the output back to PHP. But SAS was unable to accept the parameters from PHP and users would get the SAS Stored Process login page.

We needed a different solution. We needed to think outside the box. What if PHP could go to the command line and run a SAS program? We set up an intermediary server with access to both Oracle and SAS. We were able to pass the parameters through to a PHP middle-man that was able to initiate SAS from the command line and pass through the parameters. The SAS program checks the user that is passed through via PHP on a table of registered users from Oracle Data Mart and then includes other programs based on the selected parameters from the PHP page in Display 3. Access to the metadata server is possible by use of the proxy account with an encrypted password. Output is in either html, pdf, or Excel based on selections in the PHP and placed on the intermediary server where the PHP application retrieves it and makes it available for the user. This met our needs and although the SAS programmers did not code the PHP applications, I am happy to share the code with anyone who would like it. Our final output to PHP is shown in Display 4.
Because we moved to SAS 9.4 and a new Stored Process Server, we needed to recreate our Veteran’s Benefits Collection Stored Process. We were still unable to give anyone on campus access because of server security and the multiple clicks needed to get to the page.

Based on the above experiences, we came to the conclusion to not use the Stored Process Server since it was unable to meet our needs.

The SAS product that exceeded our needs was SAS Visual Analytics. This replaced the SAS/IntrNet processes that created reports for anyone including the outside world. Because of our struggles with our internal IT and not having a SAS administrator on staff, we have SAS host our SAS Visual Analytics site for us. Our old SAS/IntrNet page for viewing freshmen measures is shown in Display 5 ("Freshmen Measures," 2016). Users could select a high school or system summary, a report year, and a type of report.

After selecting from the dropdown boxes, the output would look something like the static output in Display 6.
This has been replaced by SAS Visual Analytics ("Freshmen Admissions and Performance," 2016). Tabs across the top of the page allow the user to select a variety of dynamic graphs and tables by campus and then select demographic options in the checkboxes on the left. Display 7 shows with line charts the number of applied, accepted, and enrolled students over time as well as acceptance and yield rates.

### Display 6 Former SAS/IntrNet Output

The interactive page in Display 8, shows through a bar chart the number of applied, accepted and enrolled students in our university system which is dramatic improvement over Display 6. Users are able to select the type of high school with the check boxes on the left.

### Display 7 New Freshmen Measures Dashboard Overview Page

### Display 8 Improved Dashboard Page on UNC Applicant Data
IR receives many questions about graduation and retention. Display 9 shows retention and graduation rates through a stacked bar chart so you can see the number of students who return, graduate, or are no longer enrolled from year to year.

Display 9 Graduation Rate Charts

Display 10 shows through bar charts the graduation and retention rates of a selected university, in this case, North Carolina State University, versus its peers.

Display 10 Peer Comparison Charts on Graduation and Retention
Another SAS/IntrNet process that we moved to SAS Visual Analytics was reports on transfer students coming from community colleges. Our SAS/IntrNet beginning page is shown in Display 11 ("Transfer Student Performance," 2016). Users select report type, year, community college, and a UNC institution.

Display 11 Former Transfer Students SAS/IntrNet Page
Former retention, graduation, and persistence reports are shown in Display 12.

Display 12 Former Transfer Students Report

Display 13 shows the new transfer student dashboard overview page with tabs at the top allowing the user to select the enrollment trends, transfer trends, GPA information, and graduation rates (“Transfer Students,” 2016.). This page also displays if transfer students are coming from and community colleges, private institution, other UNC universities, or from out of state, and gives more detail on the community colleges feeding our university system.
Display 14 shows transfer student enrollment trends over time.

MOVING FORWARD

The SAS Visual Analytic interactive dashboards have reduced the number of ad hoc requests to IR for data and reports because so much information is readily available. Thus far, we have developed dashboards on enrollment, freshmen admissions and performance, transfer students, research and development, degrees, educator quality, facilities usage, and we have a financial aid dashboard in development.

IR is currently still using SAS/IntrNet for a data collection with the community colleges and private universities; although the project has begun to transition this process to something simpler and more user-friendly. The collection requires these campuses to enter data in temporary webpages, some of which contain over one hundred cells, without an option to save data and return to the form. In addition to this challenge, Google crawler has stopped the SAS/IntrNet service multiple times over the past few months.

SAS Visual Analytics is our solution for external reporting. Through training, programmers are now adept at creating dashboards, however there have been some challenges uploading data to the test and production sites since we are dependent on the SAS hosting group for assistance. External users on PCs or MACs are able to access the webpage directly and select what they would like to display. The only disadvantage of this process is that many members of our Board of Governors would like this to be easily accessible via iPad or tablet, and currently you have to download an application in order to view the dashboards.

PHP is our solution for letting campus users interact through the data mart with SAS. This meets our needs for the time being.
We are currently not using the Stored Process Server for any of our business processes due to the lack of accessibility with our current server design.

We have not yet needed to drive across the state to retrieve data and distribute reports.

Table 1 below details our internal, external with secure access, and external public users; the different types of SAS software we considered for our business processes; and how they measured on our checklist of challenges.

<table>
<thead>
<tr>
<th>User Type</th>
<th>Business Solution</th>
<th>User Capability and Knowledge</th>
<th>IT Cooperation and Knowledge</th>
<th>User-friendliness/ Escaping Information Overload</th>
<th>Server Security and Accessibility</th>
<th>Other Random Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>SAS/IntrNet</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>SAS Visual Analytics</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>PHP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Stored Process Server</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Manual Collection and Distribution</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>External User with Secure Access (Campus IR Offices)</td>
<td>SAS/IntrNet</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>SAS Visual Analytics</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>PHP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Stored Process Server</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Manual Collection and Distribution</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>External Users – General Public</td>
<td>SAS/IntrNet</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>SAS Visual Analytics</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>PHP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Stored Process Server</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Manual Collection and Distribution</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1 User Types, Business Solutions, and Challenges

CONCLUSION

The Institutional Research (IR) Division in the University of North Carolina General Administration needed to update SAS/IntrNet for multiple data collections and methods of sharing information with the public.
We have shared our decision-making process to upgrade and the criteria for moving specific business processes to SAS Visual Analytics, a PHP interface, and SAS Stored Processes. This included whether the user was internal or external, whether the user had access to secure data, and how much time was involved in making the changes, the process and challenges of implementing the changes, and the lessons learned. Our solution for displaying data to external users is SAS Visual Analytics, while our solution for internal users and external users with access to secure data is PHP. This process has taken us nearly three years and we should have all of our solutions in place by summer 2016. If not, we will drive round the state collecting data on forms or flash drives and distribute paper reports in enormous three-ring binders to the general public.

REFERENCES


ACKNOWLEDGMENTS
I’d like to thank my colleagues at UNCGA IR who have assisted with testing and development through multiple system upgrades and the SAS consultants who helped us throughout this process, in particular Zach Brunk and David Franklin.

RECOMMENDED READING
- Building Business Intelligence Using SAS®
- Web Development with SAS®
- The 50 Keys to Building SAS Stored Processes

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

Laura Gelo Simpson
University of North Carolina General Administration
lgsim@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.