Paper 1731-2014

Using SAS® to Analyze the Impact of the Affordable Care Act

John J. Cohen, Advanced Data Concepts; LLC, Meenal (Mona) Sinha, Independence Blue Cross

ABSTRACT

The Affordable Care Act that is being implemented now is expected to fundamentally reshape the health care industry. All current participants--providers, subscribers, and payers--will operate differently under a new set of key performance indicators (KPIs). This paper uses public data and SAS® software to illustrate an approach to creating a baseline for the health care industry today so that structural changes can be measured in the future to establish the impact of the new laws.

INTRODUCTION

The Affordable Care Act, of considerable interest to politicians and the general public, is also going to receive great attention from political scientists, health economists, and the medical provider community. The effort to analyze the impact of these new laws will be aided by a wealth of data and the array of data manipulation, analysis, and data visualization tools available to users of the SAS System. Considerable attention must be paid to taking the data from disparate sources and creating useful metrics. A second consideration is that there are considerable lags in the preparation and reporting of many of these data sources. As such, the preparation of a baseline becomes the most immediate goal in analysis. Finally, there is considerable variation between states in their implementation strategies, whether the election to create a state-based marketplace, reliance on the Federal offerings, or a hybrid version; the decision to extend or not State Medicaid benefits; along with the underlying variation in State populations with respect current uninsured population, distribution of respective populations by age, health care requirements, access to health care, and the like. Thus our emphasis will be to suggest a set of State-level metrics.

SELECTING THE INDUSTRY BASELINE METRICS

The Healthcare Industry baseline will include the number, size and reach of various providers as well as current and new KPIs related to operating efficiency, bundled payments, etc. In addition, a set of Affordable Care Act-specific KPI's will be established so that the success of the Affordable Care Act can also be estimated in terms its stated goals.

More specifically, we will want to identify a set of metrics which will allow us to assess certain goals of the Affordable Care Act, the expansion of healthcare coverage to people currently uninsured altogether or underinsured, to improve the quality of coverage, and to identify bases for optimizing the healthcare treatment provided.

Insurance Coverage - To do so we will need data providing measures of overall insurance coverage, including patients enrolled, patients with no health insurance coverage, and finally a measure of those patients under-insured.

Healthcare Utilization – We will want data on the number of patients and families in our population (using the Insurance Industry terminology, the number of covered lives). We will want measures of Dollar expenditures in total, per patient, and per procedure. Finally we will need data on the number of treatments, broken out by provider categories (physician office, out-patient clinic, hospital, pharmacy, etc.).

Disease Incidence – Here measures of morbidity and mortality, co-morbidities, and the like are of interest. These can be broad disease states such as diabetes, asthma, or metabolic syndrome or instead specific issues such as gangrene (a challenge with diabetes patients), asthma attacks necessitating ER visits, knee replacement surgery, blood pressure prescriptions, and the like.

Healthcare Outcomes – With greater coverage, we would expect increased utilization. Confirming that people are availing themselves of treatment options previously unavailable to them, or at greater levels than before, would in the short run likely lead to higher expenditures. The critical question asks whether these "better" expenditures, more consistent with ideal treatment regimens rather than late attention to issues at harder-to-treat stages? Does better access to preventive medicine improve health outcomes? Is Emergency Room utilization better? Are people healthier, leading longer lives with better quality of life?

Control Factors - Variation is expected in any reasonable set of measures regardless of the impact of the Affordable Care Act. To tease out the independent effect of this "treatment" we will want to consider controlling for factors such as age, gender, income, prior health, prior coverage, and ethnicity. Further, local variations in government (state, county, local, non-profit) coverage, in addition to regional variations in disease patterns (e.g., varying seasonality for asthma and hay fever due to climate variation) and treatment modes are other confounding factors.

Model for Investigation – Where the goal is to measure the value of preventive medicine, changes in patterns of utilization (e.g., use of emergency rooms), are newly-insured patients being treated at earlier stages of disease states, changes in provider utilization, better attention to "unmet need. What changes do we expect to see and how can we bring data to bear to measure whether or not the changes occur?

COLLECTING DATA

Numerous public sources can be tapped to collect relevant data for the Baseline Exercise. Data requirements will be developed to align with Baseline metrics established above. Selected representative data sources include:

- Federal and State Government Websites
- National Health and Nutrition Examination Survey (NHANES)
- National Health Interview Survey (NHIS)
- Behavior Risk Factor Surveillance System (BRFSS)
- Medical Expenditure Panel Survey (MEPS
- Medicare Claim Public Use Files

Investigation will be required to understand the breadth and depth of each source, the update schedule – elements may be several years old – the strengths and gaps in these data, and the comparability across sources. (Comparability is critical if datasets are to be combined.) Displayed in Figure 1 is a more detailed treatment of some of these same sources. (For additional discussion see Gorrell 2012, and Scerbo, et. al. 2001)

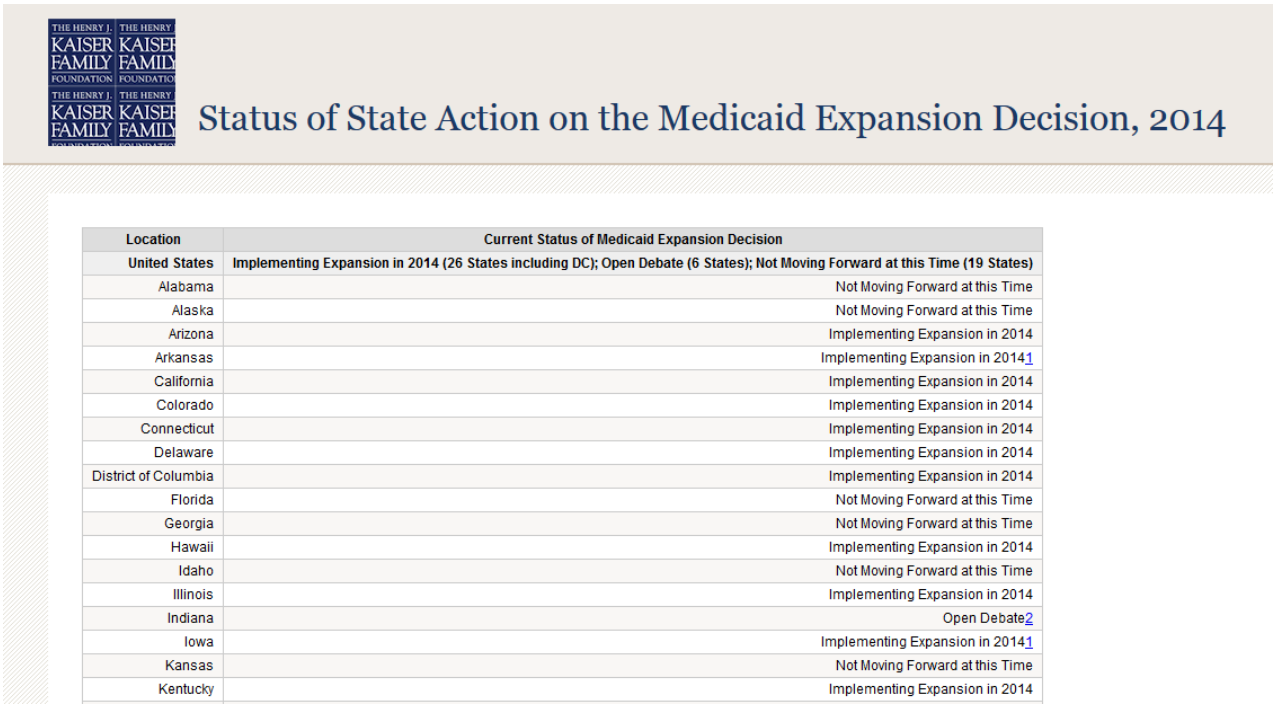
Figure 1 – Sample of Publicly Available Data Sources

Data Source	Description	Strengths	Weaknesses	Available Data Points	Source
Household Component	Health insurance coverage of the U.S. population by person and family-level characteristics is collected through the MEPS Household Component.	MEPSnet Query Tool	(1) Most recent data is for 2012; (2) State Level only through 2010	Age/Sex/Employment/ Race/ Marital_Status/ Census Regions // Out of Pocket/ Private Insurance/ Medicaid/ Medicare // Total Health/ Dental/ Prescription/ Office Based	http://meps.ahrq.gov/mepsweb/data_stats/quick_tables_search.jsp?component=1&subcomponent=0
Insurance Component	Job-based health insurance coverage offered by employers is collected through the MEPS Insurance Component.	(1) MEPSnet Query Tool; (2) Very comprehensive data	(1) Most recent data is for 2012; (2) State Level only through 2010	Industry Group/ Ownership/ Age of Firm/ Number of Locations/ Percent of Fulltime Employees/ Union Presence/ Percentage of Low Wage Earners	http://meps.ahrq.gov/mepsweb/data_stats/quick_tables_search.jsp?component=2&subcomponent=1
Quality of Healthcare	Several quality of health care items are asked in the self-administered questionnaire (SAQ) of the MEPS-HC. The quality of care measures were taken from an AHRQ-sponsored instrument, the Consumer Assessment of Health Plans (CAHPS®).		2011 Data	Various Quality Indicators Tables. Topics include experiences with getting an appointment, experiences at the appointment, and smoking and blood pressure related topics.	http://meps.ahrq.gov/mepsweb/data_stats/MEPS_topics.jsp?topicid=16Z-1
State and Metro Area Estimates	MEPS provides state-level health insurance data for employer-sponsored plans through the MEPS-IC.		The MEPS-HC public use data files that are available for download do not contain state-level identifiers. This information is available through the AHRQ Data Center for approved research projects	Percentage of businesses offering health insurance/ Percent of employees enrolled in health insurance/ etc	http://meps.ahrq.gov/mepsweb/data_stats/MEPS_topics.jsp?topicid=19Z-1

NORMALIZING AND RECODING DATA

Raw data collected from various sources will need to be scrubbed and matched to relevant time-periods. Some amount of recoding will need to be done in order to derive the Baseline Metrics. All these activities require a set of assumptions that need to be called out so that appropriate diligence can be exercised while reviewing the results of the exercise. By way of illustration, we will gather one dataset – the status of each State's decision regarding expansion of State Medicaid under the Affordable Care Act – and demonstrate "scrubbing", descriptive statistics, and data visualization.

Figure 2 – Sample Data Source (Partial)



Status of State Action on the Medicaid Expansion Decision, 2014

Location	Current Status of Medicaid Expansion Decision
United States	**Implementing Expansion in 2014 (26 States including DC); Open Debate (6 States); Not Moving Forward at this Time (19 States)**
Alabama	Not Moving Forward at this Time
Alaska	Not Moving Forward at this Time
Arizona	Implementing Expansion in 2014
Arkansas	Implementing Expansion in 20141
California	Implementing Expansion in 2014
Colorado	Implementing Expansion in 2014
Connecticut	Implementing Expansion in 2014
Delaware	Implementing Expansion in 2014
District of Columbia	Implementing Expansion in 2014
Florida	Not Moving Forward at this Time
Georgia	Not Moving Forward at this Time
Hawaii	Implementing Expansion in 2014
Idaho	Not Moving Forward at this Time
Illinois	Implementing Expansion in 2014
Indiana	Open Debate2
Iowa	Implementing Expansion in 20141
Kansas	Not Moving Forward at this Time
Kentucky	Implementing Expansion in 2014

The first step is to get the data into SAS. We download the data from the Kaiser Family Foundation website (see link below) into spreadsheet format using the convenient "Download Raw Data" button (see Figure 3).

Figure 3 – Downloading Data from Website

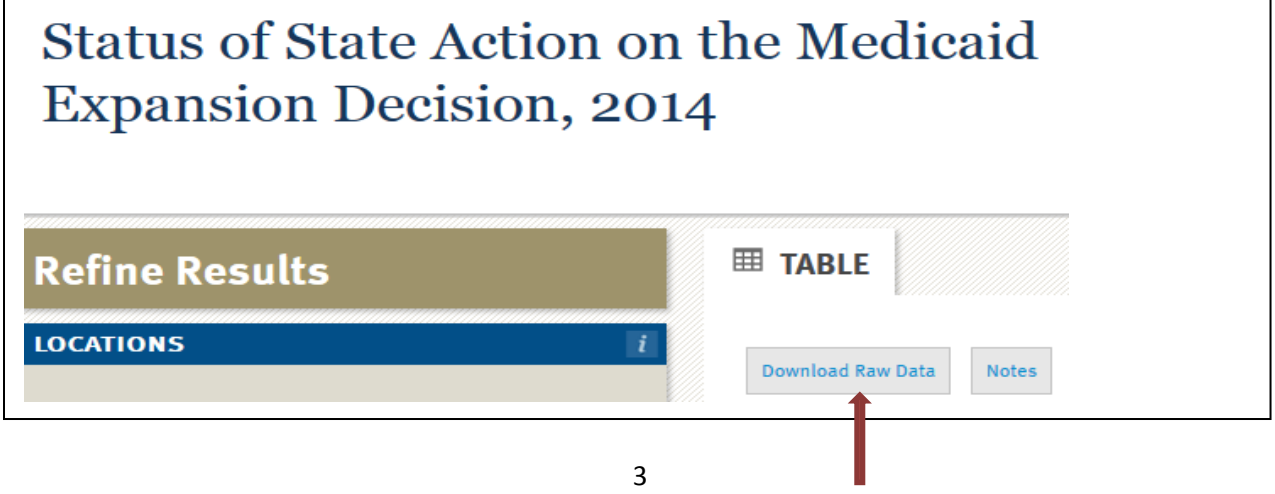


Figure 4 – SAS Import Wizard

```
NOTE: The infile 'C:\SGF\2014\Data\Status of State Action on the Medicaid Expansion Decision, as
      of December 11, 2013 (KFF).csv' is:

      Filename=C:\SGF\2014\Data\Status of State Action on the Medicaid Expansion Decision, as of
      December 11, 2013 (KFF).csv,
      RECFM=V,LRECL=32767,File Size (bytes)=2847,
      Last Modified=06Jan2014:17:51:25,
      Create Time=06Jan2014:17:47:53

NOTE: 58 records were read from the infile 'C:\SGF\2014\Data\Status of State Action on the
      Medicaid Expansion Decision, as of December 11, 2013 (KFF).csv'.
      The minimum record length was 1.
      The maximum record length was 175.
NOTE: The data set ACA.MEDICAID_DECISION has 58 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.38 seconds
      cpu time            0.04 seconds


58 rows created in ACA.Medicaid_Decision from C:\SGF\2014\Data\Status of State Action on the
Medicaid Expansion Decision, as of December 11, 2013 (KFF).csv.



NOTE: ACA.MEDICAID_DECISION data set was successfully created.
NOTE: The data set ACA.MEDICAID_DECISION has 58 observations and 2 variables.
```

A quick look at the data in Figure 5 indicates one set of challenges – we have imported into our SAS dataset footnotes, totals, and other elements of the spreadsheet which are not appropriate once we start our data analysis.

Figure 5 – A Quick Look at Our Data (Partial)

	Location	Current_Status_of_Medicaid_Expan
1	United States	Implementing Expansion in 2014 (26 States including DC); Seeking to Move Forward with Expansion Post-2014 (2 States); Not Moving Forward at this Time (23 States)
2	Alabama	Not Moving Forward at this Time
3	Alaska	Not Moving Forward at this Time
4	Arizona	Implementing Expansion in 2014
5	Arkansas1	Implementing Expansion in 2014
6	California	Implementing Expansion in 2014
7	Colorado	Implementing Expansion in 2014
8	Connecticut	Implementing Expansion in 2014
9	Delaware	Implementing Expansion in 2014
10	District of Columbia	Implementing Expansion in 2014
11	Florida	Not Moving Forward at this Time
12	Georgia	Not Moving Forward at this Time
13	Hawaii	Implementing Expansion in 2014
14	Idaho	Not Moving Forward at this Time
15	Illinois	Implementing Expansion in 2014
16	Indiana2	Seeking to Move Forward with Expansion Post-2014

A gentle round of data scrubbing is demonstrated in Figure 6, where we remove rows meaningful in a spreadsheet but in a SAS dataset not so much. Next we remove the footnote numbers from our Location variable (values of 1, 2, 3) and then from our renamed Status variable (value of 4) using contrasting methods for illustration.

Figure 6 – Data Scrubbing

```
*************************************************************;
*** SGF 2014 ************************************************;
*** Health Care Outcomes and the ACA ************************;
*************************************************************;

Libname ACA 'C:\SGF\2014\Data';

Data Medicaid_Decision;
   set ACA.Medicaid_Decision;                  *** "normalize" the data ***;
   if Location in: ('United States',' ','Footnotes','1','2','3','4')
      then delete;                             *** remove "non-data" rows ***;
     else Location = compress(Location,'123'); *** remove footnote numbers ***;
   if Current_Status_of_Medicaid_Expan = 'Not Moving Forward at this Time4'
      then Current_Status_of_Medicaid_Expan = 'Not Moving Forward at this Time';
   label Current_Status_of_Medicaid_Expan = 'Status of State Medicaid Expansion';
   rename Current_Status_of_Medicaid_Expan = Status
          Location = statename;
run;
```

PRESENTING BASELINE METRICS

The Baseline Metrics will be calculated and presented in a graphical display as well as a set of detail tables. Figure 7 shows an example of a quick frequency distribution using PROC FREQ, with the distribution of the three values at the time we pulled these data.

Figure 7 – Descriptive Statistics

```
Proc Freq data=Medicaid_Decision order=freq;
   table status;
run;
```

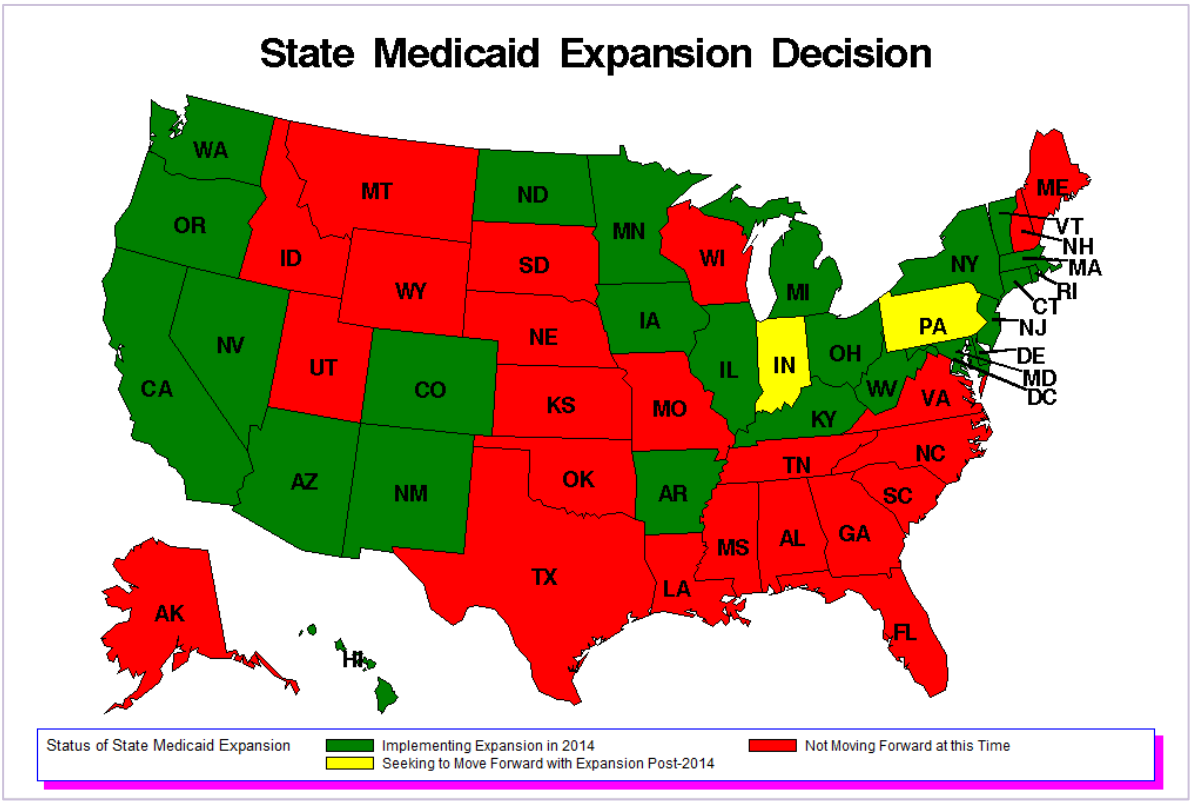
The FREQ Procedure
Status of State Medicaid Expansion (as of December 11, 2013)

Status	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Implementing Expansion in 2014	26	50.98	26	50.98
Not Moving Forward at this Time	23	45.10	49	96.08
Seeking to Move Forward with Expansion Post-2014	2	3.92	51	100.00

These data are fairly straightforward. More complex data may warrant additional study at a descriptive level, especially when multiple datasets are to be combined. And additional data scrubbing may then be required. (A key variable in one dataset may be character format while the equivalent variable in the next dataset is numeric – even though the nominal values are identical. Or they may be stored in different lengths. Or perhaps some of the values do not correspond and must be recoded, say, from "1" and "2" to "Y" and "N", or to be all uppercase, or the like.)

A productive step in research can be data visualization – information presented graphically may be suggestive in ways that simple data tables are not. In Figure 8 we show an example of a map of the United States showing each States' status with respect to Medicaid expansion visually. The SAS programming to create this map (along with a simpler version) are presented in Appendix A. In this instance, as the data are at a state level, the geographic representation of these same data offers a productive complement.

Figure 8 – Data Visualization



INFERENTIAL STATISTICS

A desired step, once data are identified, scrubbed, and well-understood, is to employ some of the powerful statistical tools available to us to measure the differences between baseline and the Healthcare profile of the United States post- Affordable Care Act.

Inferential Statistics infer from the sample to the population based on the probability that sample characteristics represent the population as a whole. They can be used to determine the strength of relationship between independent (causal) variables and dependent (effect) variables. One of the key success criteria in this analysis is the correctness and adequacy of the sample itself. Some common techniques that are relatively easy to interpret include one-sample hypothesis test, t-Test or ANOVA, Pearson Correlation, Bivariate Regression and Multiple Regression.

Figure 9 shows a very simple t-Test example that compares 'before' and 'after' data using 'paired after*before' statement. The output from t-Test shows that the Mean has increased by 1.7 with a p-value of 0.022 – and since 0.022 < 0.05, we can infer that the increase in Mean value observed in the sample can be extended for the population as a whole with a 95% confidence level.

The t-test can be extended to multiple pairs of data. For example, 'Paired A*B' compares A-B; 'Paired A*B C*D' compares A-B and C-D, and 'Paired (A B)* (C D)' compares A-C, A-D, B-C and B-D.

Figure 9 – Sample output from PROC tTest

```
/*  The Before and After data can be analyzed using PROC t-
test    */

data demo;
     input Mbr Before After;
     cards;
     1 22 25
     2 33 33
     3 40 45
     4 30 30
     5 28 30
     6 32 36
     7 35 35
     8 25 24
     9 27 29
     10 22 24

;

title "Paired T -test";
proc ttest data= demo;
paired after * before;
run;
```

Paired T -test

The TTEST Procedure

Difference: After - Before

N	Mean	Std Dev	Std Err	Minimum	Maximum
10	1.7000	1.9465	0.6155	-1.0000	5.0000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1.7000	0.3076	3.0924	1.9465	1.3389	3.5536

DF	t Value	Pr > |t|
9	2.76	0.0220

CONCLUSION

The Affordable Care Act is a sweeping change to the fundamental construct of the Healthcare Industry in the United States. Various elements of this law are being implemented at varying degrees of sophistication and on very different timetables across states and across business entities. This makes the measurement of effectiveness of the law a moving target. In addition, the data generated by various players in the industry is disparate, incomplete and dated. So, any attempt to measure the effectiveness of the law has to necessarily start with establishing a baseline.

The recommended approach for establishing a baseline involves selection of suitable industry metrics, collection of relevant data, normalization and recoding of data, presentation of baseline metrics and optionally, inferential statistics.

This data analysis exercise for establishing a baseline is expected to run into the following issues. There will be long lags in data being reported by the industry resulting in delayed analysis. Data formats and level of completeness are expected to be choppy initially. The data analysis has to be fine-tuned to be able to tell the difference between noise/ small differences vs. large differences. In trying to establish a causal relationship, it will be difficult to isolate the effects of the law from other factors not specific to ACA. It would be appropriate to do a separate study design for independent effects of ACA.

Some aspects of the ACA programs will exhibit better outcomes than others. And some will lend themselves to measurement better than others. One hope, as data analysts, is that there is sufficient variation in the data to allow us to come to reasonably confident conclusions. Random 'noise' and indeterminate results would be disappointing from both a data analytics and a policy perspective. "

SOURCES

Greg Allen, "In Florida, Insurer And Nonprofits Work On Enrollment", NPR - October 1, 2013 http://www.npr.org/blogs/health/2013/10/01/228196533/in-florida-insurer-and-nonprofits-work-on-enrollment?sc=17&f=1128

Sharon Frazee, "Impact of Wealth on Rx Utilization", Healthcare Insights From The Express Scripts Lab, 6.14.13

Jenny Gold and Sarah Varney, "Obamacare Day One: A Tale Of Two States", KHN - October 2, 2013, http://www.npr.org/blogs/health/2013/10/02/228249901/obamacare-day-one-a-tale-of-two-states?sc=17&f=1001

Paul Gorrell, "Generating Estimates for U.S. Healthcare Costs and Use", NESUG 2012

Louise S. Hadden, Mike Zdeb, and Robert Allison, "Wow! You Did That Map with SAS/GRAPH®?", Paper 215-2009, SAS Global Forum, 2009

Kaiser Family Foundation, "Status of State Action on the Medicaid Expansion Decision, 2014", http://kff.org/health-reform/state-indicator/state-activity-around-expanding-medicaid-under-the-affordable-care-act/

Steve Miller, "MAP: Bad Rx Decisions Cost Poor US States Most", Healthcare Insights From The Express Scripts Lab, 4.23.13

Stephanie O'Neill, "Small Businesses May Find Insurance Relief In Exchanges", SCPR - October 3, 2013, http://www.npr.org/blogs/health/2013/10/03/228522598/small-businesses-may-find-insurance-relief-in-exchanges?sc=17&f=1001

SAS Institute Inc 2005, "Sample *25558:* Label the states on a U.S. map using PROC GMAP", http://support.sas.com/kb/25/558.html

SAS Institute Inc 2012. **SAS/GRAPH® 9.3: Reference**, Third Edition. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2012. **SAS® 9.3 Language Reference: Concepts**, Second Edition Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2013. **SAS/STAT® 12.3 User's Guide**. Cary, NC: SAS Institute Inc.

Scerbo, Marge,Dickstein, Craig,Wilson, Alan, **Health Care Data and SAS**, SAS, Cary, NC, 2001

Adeline J. Wilcox. "Computation of Medicare Part A Skilled Nursing Facility Reimbursement Rates", NESUG 16

Adeline J. Wilcox, "Length of Intensive Care Unit Stay Computed from the VHA Corporate Data Warehouse", NESUG 2012

John Ydstie, "Part-Time Workers Search New Exchanges For Health Insurance", NPR - October 4, 2013, http://www.npr.org/blogs/health/2013/10/04/229103233/part-time-workers-search-new-exchanges-for-health-insurance?sc=17&f=1128

http://247wallst.com/special-report/2013/09/27/10-states-with-the-worst-health-coverage/2/

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. You may contact the authors at:

John Cohen
Advanced Data Concepts LLC
Newark, DE
(302) 559-2060
jcohen1265@aol.com

Meenal (Mona) Sinha
Independence Blue Cross
Downington, PA
(484) 467-7503
meenalsinha99@gmail.com

Appendix A - Program to Create Map Displayed in Figure 8

Task: Create the map displayed in Figure 8 above in two steps:

1. Create a simple map using SAS/GRAPH®
2. Create an ANNOTATE dataset and generate a more complex graph

Create a simple map

- Use Sashelp.zipcode SAS-supplied dataset to allow us to easily associate FIPS code for each state with our downloaded data. All three programming components are displayed in Figure 10.
- Merge Sashelp.zipcode with our Medicaid Decision dataset on STATENAME variable. The resulting dataset – our Medicaid Decision dataset with FIPS code appended – is displayed in Figure 11.
- Set colors for responses and generate simple map using maps.us, a SAS-supplied dataset defining state borders/boundaries, as the map dataset, and our Medicaid Decision dataset as the response dataset. (Color values on the map will be determined be each states value on the STATUS variable.) Variable STATE (FIPS code) is the common/ID variable in both map and response datasets. The map is shown in Figure 12.

Figure 10 – Create a Simple Map

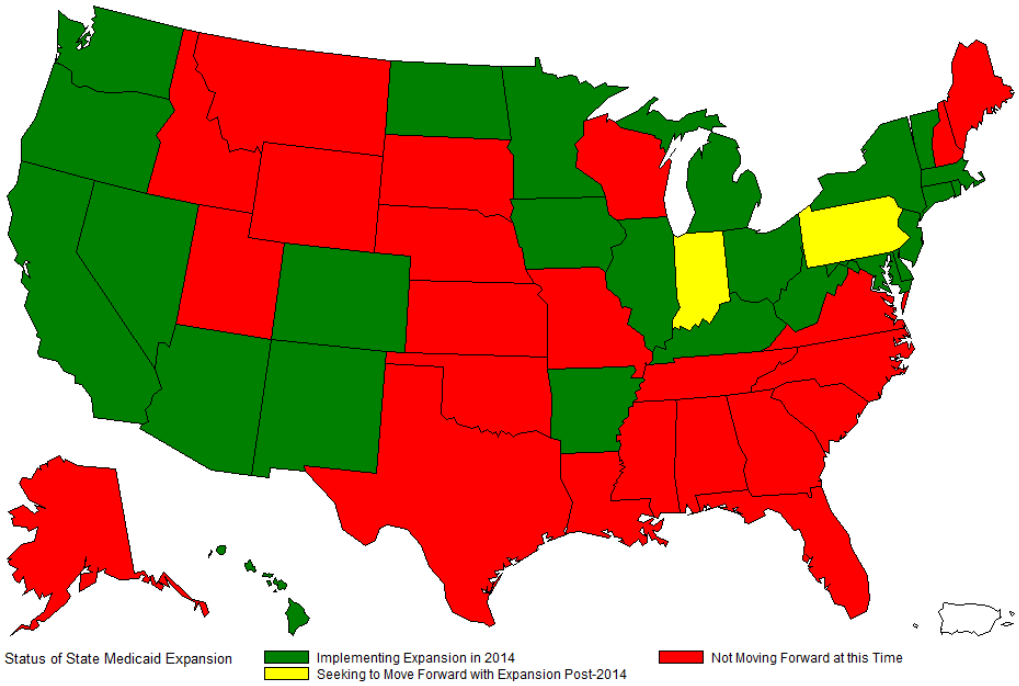
```
            /*** bring in Sashelp.zipcode for FIPS code (variable state) ***/
Proc Sort data=Sashelp.zipcode(keep=state statename) nodupkey
          out=statename;
   by statename;
run;
          /*** merge Sashelp.zipcode with our Medicaid Decision dataset ***/
Data Medicaid_Decision_fips;
   merge Medicaid_Decision(in=on_list)
         statename(in=have_fips_code);
   by statename;
   if on_list;
   if not have_fips_code then put 'ERROR:' _all_;
run;
                                  /*** setting colors for SAS/GRAPH (GMAP) ***/
pattern1 v=ms c=green;
pattern2 v=ms c=red;
pattern3 v=ms c=yellow;
                                                          /*** generate map ***/
Proc Gmap map=maps.us
         data=Medicaid_Decision_fips;
   id state;
   choro status / coutline=black;
run;
quit;
```

Figure 11 – Medicaid Decision with FIPS Code appended (partial)

VIEWTABLE: Work.Medicaid_decision_fips

	Full name of state/territory	Status of State Medicaid Expansion	Two-digit number (FIPS code) for state/territory
1	Alabama	Not Moving Forward at this Time	1
2	Alaska	Not Moving Forward at this Time	2
3	Arizona	Implementing Expansion in 2014	4
4	Arkansas	Implementing Expansion in 2014	5
5	California	Implementing Expansion in 2014	6
6	Colorado	Implementing Expansion in 2014	8
7	Connecticut	Implementing Expansion in 2014	9
8	Delaware	Implementing Expansion in 2014	10
9	District of Columbia	Implementing Expansion in 2014	11

Figure 12 – Simple Map



Create an ANNOTATE dataset

- Bring in the SAS-supplied maps.uscenter dataset. This contains the mapping coordinates for the center of each state (see Figure 13). We use the FIPSTATE function to create a variable STATE1 containing the 2-character state abbreviation. We will be using the coordinates expressed in Radians.
- As shown in Figure 14, the maps.uscenter dataset is the basis for creating a new, specialized ANNOTATE dataset which we will call MAPLABEL. Several special variables are created (with special key names) which are processed as plotting/mapping data by SAS/Graph. We will take each states' center coordinates and position the state abbreviation at that location – unless that location is flagged in the dataset as being external (state too small to fit) – ocean = yes – in which case we will position the state abbreviation external to the state borders and draw a line back to the state's center.
- Create a slightly customized legend called **legend1**.
- Generate the more complex map (see Figure 8), with three additions to the simple map program:
  - Exclude Puerto Rico: `where=(state^=72)`
  - Call the customized legend: `legend=legend1`
  - Invoke the annotate dataset: `annotate=maplabel`

Figure 13 – maps.uscenter dataset

	Unprojected Latitude in Degrees	Unprojected Longitude in Degrees	Ocean: Y/N	State FIPS Code	Projected Longitude from Radians	Projected Latitude from Radians
1	32.799987793	86.599975586	N	1	0.1352076079	-0.066475149
2	65	152	N	2	-0.321886144	-0.121835931
3	34.5	111.5	N	4	-0.222906551	-0.024993309
4	35	92.5	N	5	0.047765589	-0.033613471
5	37	120	N	6	-0.33099946	0.0425571086
6	39	105.5	N	8	-0.129846581	0.0422959775
7	40.599975586	72	Y	9	0.3116508108	0.1028083968
8	41.599975586	72.75	N	9	0.2978662858	0.117414451

Figure 14 – Create an Annotate Dataset

```
       /*** SAS-supplied SAS dataset containing State centers mapping coordinates ***/
data states;
   set maps.uscenter(where=(fipstate(state) not in ( 'GU','PR')) drop=long lat);
   state1= fipstate(state);
run;
                                          /*** create specialized ANNOTATE dataset ***/
data maplabel;
   length function $ 8;
   retain flag 0 xsys ysys '2' hsys '3' when 'a' style 'swissb';
   set states;
   by state;
   if state_app=. then state_app=0;
   function='label';
   text= state1; *<== this is the place where we place the annotated text;
   size = 2.7; position='5';
   /* The FIPSTATE function creates the label */
   /* text by converting the FIPS codes from */
   /* MAPS.USCENTER to two-letter postal codes. */
   if ocean='Y' then do;
      position='6'; output;
      function='move';
      flag=1;
   end;
   /* If the labeling coordinates are outside the state (OCEAN='Y'), Annotate */
   /* adds the label and prepares to draw the leader line. Note: OCEAN is a */
   /* character variable and is therefore case sensitive. OCEAN='Y' must specify */
   /* an uppercase Y. */
   /* When external labeling is in effect, Annotate */
   /* draws the leader line and resets the flag. */
    else if flag=1 then do;
       function='draw'; size=.5;
       flag=0;
    end;
   output;
run;
                    /*** set certain additional formatting details for the legend ***/
legend1 cborder=blue cshadow=magenta;
                                      /*** generate more complex map (see Figure 8) ***/
Proc Gmap map=maps.us(where=(state^=72))
          data=Medicaid_Decision_fips;
   id state;
   choro status / legend=legend1 annotate=maplabel coutline=black;
title1 f=swissb height=25 pt "State Medicaid Expansion Decision";
run;
quit;
```