

5-15-2016

De-Scribing Administrative Law Case Data: From Sparklines to Dashboards to Analytics

Steven Placek

Follow this and additional works at: <http://digitalcommons.pepperdine.edu/naalj>



Part of the [Administrative Law Commons](#)

Recommended Citation

Steven Placek, *De-Scribing Administrative Law Case Data: From Sparklines to Dashboards to Analytics*, 36 J. Nat'l Ass'n Admin. L. Judiciary 98 (2016)

Available at: <http://digitalcommons.pepperdine.edu/naalj/vol36/iss1/3>

This Article is brought to you for free and open access by the School of Law at Pepperdine Digital Commons. It has been accepted for inclusion in Journal of the National Association of Administrative Law Judiciary by an authorized administrator of Pepperdine Digital Commons. For more information, please contact paul.stenis@pepperdine.edu.

De-Scribing Administrative Law Case Data: From Sparklines to Dashboards to Analytics.

By Steven Placek¹

I. INTRODUCTION	99
II. CURRENT STATE OF DATA DISPLAYS.....	101
III. SPARKLINES AND DASHBOARDS FOR LEGAL ADMINISTRATION STATISTICS.....	107
IV. ANALYTICS FOR ADMINISTRATIVE LAW AGENCIES.....	118
A. <i>The Transition to Analytics</i>	118
B. <i>The Analytics Framework</i>	120
V. ANALYTICS CASE STUDY AND SOME LESSONS-LEARNED AT THE NATIONAL APPEALS DIVISION	125
A. <i>The Current System: Goals and Expectations for Analytics</i>	125
B. <i>Description of NAD Analytics</i>	130
C. <i>Datasets</i>	132
D. <i>Dashboards and Beyond</i>	137
VI. CONCLUSION.....	140

¹Steven Placek is the Special Assistant for Legal Administration at the National Appeals Division (NAD). He is a former Associate Professor of English at the United States Military Academy at West Point. The views and opinions in this article are those of the author and do not reflect the policy or position of the United States Department of Agriculture.

I. INTRODUCTION

In our age of big data, advancements in robust hardware and scalable software have made it easier and more affordable to collect, store, and analyze large amounts of information. These developments have created expectations for administrative law agencies to display more visual metrics about their mission and operational activities. Moreover, most electronic case management systems (CMS) and platforms now provide an opportunity to engage in the kinds of analytics formerly the province of large marketing entities or internet business giants, like Google or Amazon, that analyze customer transactional business data.

If you think about it, administrative law agencies also engage in many transactions with their customers: First, they gather various demographic, personal, and legal data about appellants and their cases; second, they “transact business” with appellants during the life-cycle of a case, including activities such as issuing notices, making rulings, or scheduling hearings. Finally, they adjudicate and dispose of cases, generating a bulk of data and reportable metrics about timeliness, the outcomes of adjudications, and other measures that contextualize, compare, or contrast a given case with other cases. These activities provide a rich reserve of collectable data and metadata for input into a CMS, which is subject to analysis and display.²

For many years, administrative law agencies commonly employed, what I call, the inscribed method to display data about their activities. The inscribed method usually takes two forms. The first form is the classical narrative that provides analysis with numbers embedded in sentences, usually with a noun or noun / modifier accompanied by a number, such as *Cases (2,304)*, or *FY ‘15 Cases Filed (2,304)*. However, agencies now provide more visual statistical displays on websites, in agency reports, or in white papers. This trend evolves amidst a data-rich culture that generates an

² In fact, with the progress of electronic filing now permeating the industry, many appellants are actually inputting this kind of data themselves into case tracking systems.

astonishing amount of data tables, micro-charts, pictures, and maps that we consume on our smart devices, cable news, sports and business networks, or by reading the finance and weather sections of newspapers.

Unfortunately, as agencies keep pace with this trend, the second form emerges, which requires readers to confront many bulky and stock-colored Excel-type charts. Although both inscribed forms of data display may have utility at times, they often display limited variables and lack context and density of data that could otherwise be presented more efficiently with a small data table or through a brief paragraph to transmit the same information. In relying solely on the inscribed method, many agencies miss opportunities to provide more meaningful information and insightful analysis about their activities.

In this article, I present some examples of opportunities for agencies to depart from the inscribed method. These examples of display “de-cribe” administrative law data, making the data more visually active, multi-variate—with additional context and greater density. They create word-like picture effects for the reader and provide more meaningful insights for agencies to convey their performance metrics. De-described data displays complement the age of big data. A commitment to these displays encourages agencies to collect not only more data, but also different kinds of data, that can increase customer engagement and become embedded in operational decisions that affect agency operations.

In Part II of this article, I review examples of the current state of data display for administrative law agencies and show how the inscribed method limits complex displays of data. These data displays ultimately stymie the agency collection and display of data and the choice of performance metrics. Concurrently, these agency displays limit persuasive arguments about their business cases. In Part III, I introduce the concept of the sparkline, which is the fundamental unit of visual data display in the big data era that appears in organizational dashboards and analytics deployments. I discuss how sparklines can apply to the display of legal administration agency data and, when applied, encourage agencies to establish performance metrics and display data that cut across a greater number of dimensions for their activities. This application reflects an agency transition to the beginning of an analytics mindset. In Part IV, I address how an agency decision to implement formal analytics software and platform tools for its data fits within the

broader context of business intelligence and processes. An effective implementation of analytics heralds an agency transformation of how it views and presents its operational activities that goes well beyond the technical installation and use of fashionable software. In Part V, I discuss how the principles of complex data display give rise to dashboards and analytics modules that track, display, and operationalize case data. In this regard, I will share some lessons learned through the experience of establishing an analytics capability at the National Appeals Division (NAD) at the United States Department of Agriculture (USDA).

II. CURRENT STATE OF DATA DISPLAYS

Federal and State administrative law agencies present data about their activities to comply with statute or regulation, secure funding, respond to congressional inquiries, and inform the public, the legal community, or potential new appellants. Data displays also demonstrate an agency value of transparency.³ As transparency has evolved, agencies have begun to display statistics and post large datasets online about their activities to meet the obligations of accountability, but often this release of information is discretionary.⁴ Ironically, the proliferation of data release, through its sheer volume, may breed a new type of opacity.⁵ This transparency evolution reflects another facet of big data challenges for administrative law agencies, and for government generally, that underpins agency efforts to employ more technology to collect, parse, and display an increasing volume of data. The triad of more data, technology, and transparency creates an inflationary cycle that moves agencies to reexamine their metrics and data stories about their activities. Agencies that demonstrate more facility with complex data collection, and support their performance metrics with meaningful data displays, make stronger cases about their accountability to the

³ For a comprehensive discussion about transparency, which includes the advantages and disadvantages of statistical displays, see Elizabeth Figueroa, *Transparency in Administrative Courts: From the Outside Looking In*, 35 J. NAT'L ASS'N ADMIN. L. JUDICIARY 1 (2015), <http://digitalcommons.pepperdine.edu/naalj/vol35/iss1/1>.

⁴ *Id.* at 14.

⁵ *Id.* at 14-15.

public and customers. For these reasons, agencies have an interest in making insightful and meaningful arguments about their performance.

Generally, the traditional inscribed method of providing statistics for legal activities is similar to the paragraph below, found in the 2014 Annual Report of the United States Courts, which explains the negative impact government sequestration had upon its Defender Services staffing levels:

The Defender Services program continued its recovery from the budget cuts that occurred during sequestration, and the resulting loss of staff and employee furloughs. Staff levels in federal public defender organizations reached a high point of 2,779 in October 2012, and began to decrease rapidly during sequestration, bottoming out at 2,358 in March 2014, a decrease of 421 (15 percent). During this time, it was a challenge for existing staff to handle caseloads. As of October 2014, the current staffing level is 2,514 – still significantly below the October 2012 on-board level. It is a continuing priority of Defender Services to ensure that the effects of the budget cuts do not damage the right to counsel guarantees of the Criminal Justice Act and the Sixth Amendment.⁶

The passage above makes a clear argument for a trend over time, with statistical information integrated into the sentence structure. Once again, the reader sees the noun / modifier combined with numbers (e.g. *Staffing Level is 2514* or *Decrease of 421 (15 percent)*). This form of explanation, and thus its persuasiveness, is limited, however, because the paragraph displays only two variables (time and number of staff). Moreover, the impact (“damage”) upon the Defender Services program is vague or lacks context. As inscribed, the data is static: answers to questions, such as what precisely the impact was, or how the program has recovered, are not apparent in the data display. Yet, even if the writer had included information in this narrative form about how the Defender Services measures its ability to guarantee the right to counsel, how that right

⁶ U.S. CTS., ANNUAL REPORT (2014), <http://www.uscourts.gov/statistics-reports/defender-services-annual-report-2014>.

diminished over a corresponding period, or how sequestration unreasonably increased staff workload, then the paragraph might have been too demanding for a reader to absorb through linear, sequential reading.

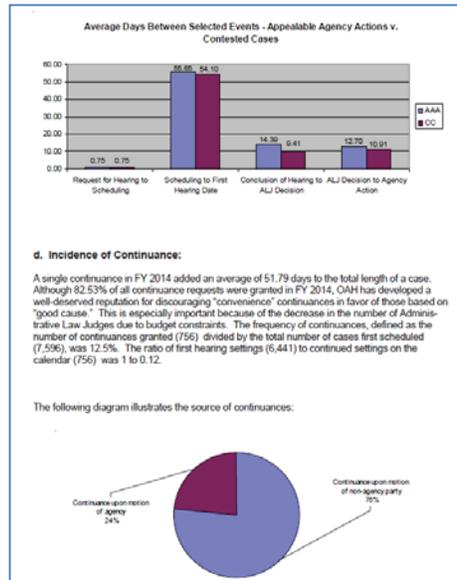
In attempts to convey statistics with more impact, many agencies now provide more visual displays. A simple count of the number of tables and charts in agency reports through the years verifies this trend. For example, as with many state and federal appeal agencies, the Arizona Office of Administrative Hearings (AZOAH) and the Board of Veterans Appeals (BVA) post annual reports on the web, dating back over fifteen years, that include routine administrative appeal statistics about cases filed, decisions issued, timeliness, and other metrics.⁷ In the earliest posted reports, both agencies relied more upon the inscribed narrative form to cite trends or workloads. As time progressed, these agencies included more visual displays. For example, the BVA 1991 report contained a single data table; by 2014, the report contained six data tables and four column graphs. Similarly, the 1996 AZOAH report contained one data table, with six column charts and one pie chart. At the height of its affection for visual data display, however, the AZOAH 2014 report contained seven tables, seven column charts, and four pie charts.⁸

One section of the 2014 AZOAH report (below) shows how most agencies introduce, display, and explain performance metrics.⁹ In this section, visual displays, interrupted by an expository paragraph, take up most of the reading space. Although the column and pie charts spatially demand most of the reader's attention, they contribute little more insight than would a sentence or two in the inscribed form; or even simpler, a small table with two or three rows and columns that would be more effectively convey the data. For example, on the column chart at the top of the page, the agency presents data about the timeliness of events in various phases of an appeal. This data includes events, such as the average time from when an appellant

⁷ I wish to acknowledge the Veteran's Board of Appeals and the Arizona Office of Hearings and Appeals for exemplifying the highest standards of transparency throughout the years in posting these reports. The specific references or examples taken from their annual reports are intended to make broader points about visual data and are not intended to make judgments about agency operations.

⁸ The 2015 AZOH report reduces the number of charts.

⁹ 19 AZOAH ANN. REP. 4 (2014).



requests a hearing to the time a hearing is scheduled, or the average time from the conclusion of a hearing until an appellant receives an Administrative Law Judge (ALJ) decision, and so forth. This chart follows a paragraph on the preceding page that references the state codes that require a "rigorous timeline" for Appealable Agency Actions (AAA) and Contested Cases (CC)—the abbreviations in the legend of the column chart.¹⁰ Although it takes up over one

third of the page, the column chart displays only a single data point for each of the annual averages for those phases. Unfortunately, there is nothing in the report or the chart that shows how agency appeals compare to those mandated timelines, which is information that would have made the chart more meaningful. You assume the agency would want to make this comparison, because it appears the averages in these categories fall below the required timelines—something positive to emphasize for its constituencies. Moreover, there are other contextual data, which, if displayed, could inform this area of analysis: Have these averages improved over time? Or what is the range or standard deviation of days for cases at each phase? It may be that the agency chooses not to provide this information or additional complexity. However, the large amount of space, ink, and color on the chart devoted to the relatively sparse data suggest that it will be more meaningful.

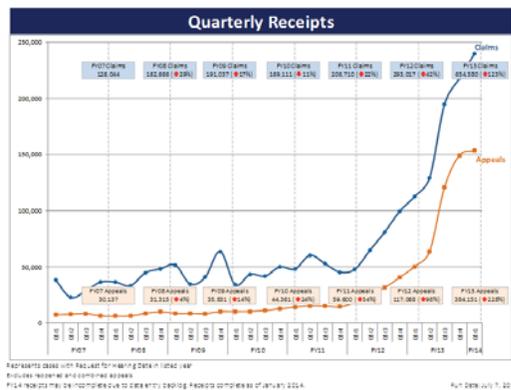
Below the column chart, the section about continuances (Section d) addresses delays that extend the number of days in the hearing phase.¹¹ The section about continuances on the same page contains a paragraph and a large pie chart, which, upon scrutiny, reveals no

¹⁰ 19 AZOAH ANN. REP. 3 (2014).

¹¹ Interestingly, data about the average number of days for continuances figures into the number of days it takes to hold a hearing, which the agency chose *not* to display in the column chart above.

cause and effect relationship between the chart and the preceding paragraph. Further, the pie chart only presents two variables about continuances (percentage and the party requesting the continuance).¹² Unlike the pie chart, the paragraph about continuances is potentially the most variable and analytical data description on the page: in five sentences, it discusses averages, the percentage of granted continuances, frequency rate, ratios—and a statement about the “well-deserved” reputation of ALJs to ensure continuances are granted with good cause. Although these variables might make for some rich and insightful visual displays of data, the Agency chose to present this material in the classical inscribed narrative form, thereby limiting its ability to make discerning arguments and a favorable case about how it conducts this phase of the hearing process.

Charts like the one above are not aberrations for administrative law agencies. Everywhere you look, you can find very large data-



Excel-type charts that fall short of telling an effective agency story. For example, the Office of Medicare and Appeals (OMHA) prominently displays a full screen trend line chart (left) that plots quarterly receipts for claims and appeals submitted from 2007-

2014.¹³ This chart falls under the section about Adjudication Timeframes. Although more complex than most agency trend line charts, there is no adjudication time variable plotted in the chart.¹⁴ Instead, the agency chose to provide some adjudication averages in a

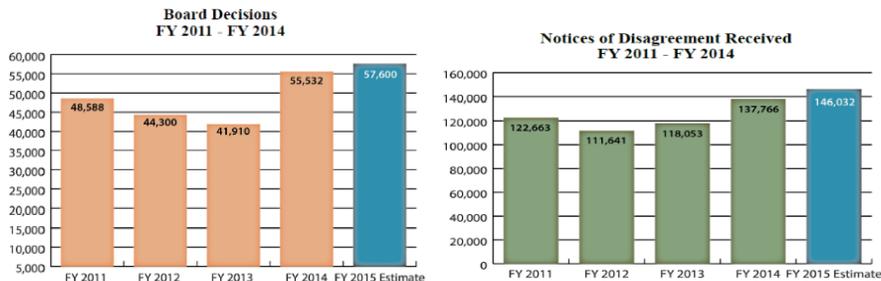
¹² Rarely does a pie chart improve data display over a simple data table.

¹³ *Office of Medicare Hearings and Appeals (OMHA)*, U.S. DEPT. OF HEALTH & HUMAN SERVS. (Apr. 29, 2015), http://www.hhs.gov/omha/important_notice_regarding_adjudication_timeframes.html#adjudication.

¹⁴ The chart does have the virtue of showing the OMHA has received over 600,000 claims and appeals since 2010, but the quarterly totals—and how they compose a percentage of the annual totals—are difficult to interpret.

few sentences preceding the chart and in a brief table in other parts of the website.¹⁵ The purpose of the chart seems to send applicants one vague general message: Don't expect an outcome soon, because this agency is overwhelmed, and it's getting worse. On this section of the webpage, the Agency seems like it wants to show a cause and effect relationship between quarterly receipts and the average time of adjudication, but as presented, these variables are not displayed together and the story is fragmented.¹⁶

A fragmented story is also on display in the 2014 BVA Annual report that appends two Excel charts (below) at the end of its report to present some statistical information about the number of Notices of Disagreement Received and the number of Board Decisions the Agency issued from FY 2011-2014.¹⁷



Each chart takes almost a full page worth of space. The location of these column charts at the back of the report suggests they are obligatory, however, and any substantial analysis about the BVA challenge in meeting its significant appeals mission can be found in other parts of the report in the traditional inscribed narrative form.¹⁸

One might argue that charts like these referenced above have at least some visual impact and draw a reader's attention to the data,

¹⁵ *Chapter I-7 Adjudication Timeframes*, U.S. DEPT. OF HEALTH & HUMAN SERVS. (Dec. 2, 2015), http://www.hhs.gov/omha/OMHA_Case_Processing_Manual/ocpm-d1c7.pdf.

¹⁶ *Id.*

¹⁷ *Annual Report Fiscal Year 2014*, BD. OF VETERANS' APPEALS 25, 28 (2015), http://www.bva.va.gov/docs/Chairmans_Annual_Rpts/BVA2014AR.pdf. [hereinafter BD. OF VETERANS' APPEALS ANNUAL REPORT].

¹⁸ *Id.*

because they contain color-shaded pictures of blocks and pie slices that take up the percentage of space in proportion to the dataset shown.¹⁹ But they really are providing no more information than would a sentence written with a noun / modifier accompanied by a number. Instead of these column charts containing more data and context, they triple-down on displaying the value of each of their relatively few data points: The first indication of the data value is when the reader engages the size of the area of space taken up by the color-shaded column; second, the label at the top of the column repeats the data value; and finally, the height of the column crosses a line that is labeled on the Y-axis, providing yet a third confirmation of the data point.²⁰ It is almost as if these charts draw repeated attention to their limits by repeating inscriptions of the paucity of the data they contain.

Understandably, the increased user-friendliness of standard templates with software tools like Excel encourage these data displays.²¹ Agencies have better case tracking systems and can download data to spreadsheets, taking advantage of software-provided charts that transfer data into boilerplate displays. However, there are other examples of data display in our routine observances that provide examples for agencies to model, empowering them to describe their data with greater insight and meaning.

III. SPARKLINES AND DASHBOARDS FOR LEGAL ADMINISTRATION STATISTICS

It only takes a quick glance of the financial, weather, and sports pages of newspapers to see that a reader can absorb an astounding amount of meaningful information in a relatively small space. Sections that group this information are dashboards, and one of the fundamental units of composition for the dashboard is the sparkline.²²

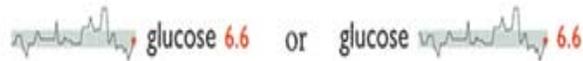
¹⁹ *Id.*

²⁰ *Id.*

²¹ *Sparklines in Excel – Mini Graphs that Make Data Analysis Fun*, VERTEX, 42 (May 21, 2013), <http://www.vertex42.com/blog/help/excel-help/sparklines-in-excel.html>.

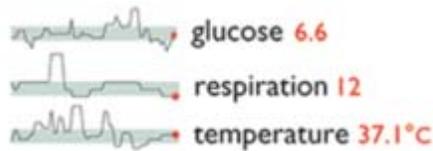
²² *The Darlings of Dashboards*, DUNDAS (Nov. 28, 2013), <http://www.dundas.com/blog-post/the-darlings-of-dashboards/>.

As Edward Tufte explains, “[s]parklines are *datawords*: data-intense, design simple, word-sized graphics.”²³ The main advantage of a sparkline is that it can describe data trends that include many data points and move within multi-variate spaces, like a table, a graph, or in combination with other sparklines.²⁴ Tufte’s classic example of a meaningful sparkline is the glucose level of a medical patient’s clinical record:²⁵



The sparkline above tracks the last eighty readings of the patient, and a gray band, which shows the range for a normal glucose reading, gives the data display context.²⁶ Words, pictures, context, colors, and fonts are critical for sparklines. In the sparkline above, for example, the most recent glucose reading is 6.6, highlighted in red font and connected to the red dot on the line.

Moreover, a data display of grouped sparklines can efficiently display the overall health of a patient.²⁷



The sparklines above for the health of a patient can apply to the health of an administrative legal agency. In the agency case, the ‘readings’ display would comprise agency metrics, such as cases docketed, decisions issued, or averages for timeliness. For example,

²³ For a thorough discussion of the history, theory, and practice of Sparklines, see EDWARD ROLF TUFTE, BEAUTIFUL EVIDENCE 47-63 (Cheshire, Connecticut: Graphics Press LLC, 2006). His website also contains a sparkline discussion thread that provides fascinating and enlightening commentary on the use of sparklines in business and industry. See *Sparkline Theory and Practice Edward Tufte*, EDWARD TUFTE, http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR.

²⁴ *Id.*

²⁵ *Id.*

²⁶ *Id.* at 47.

²⁷ *Id.*

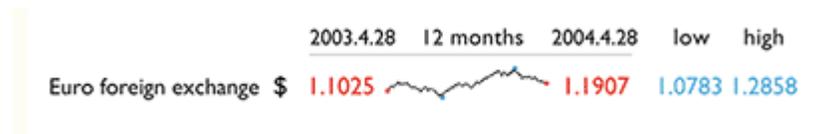
a simple, small, and efficient column sparkline for BVA data can effectively display the last five years of performance for the metrics previously referenced above by the large Excel charts at the end of its annual report.



Or a trend sparkline can display the metrics for claims and appeals displayed on the very large OMHA trend chart:



As efficient as the group of sparklines above may appear, they still lack context that might provide more meaning. For example, over what period were these ‘readings’ taken? Time context is critical for endeavors such as financial management and investment, and financial sparklines often show trends in context over time.²⁸



Precisely conveyed context shows the period of the data trend and the range of data. Once again, the color of the font indicates the **beginning** and **end** values, as well as the **high** and **low** values. As a reader can see, sparklines can be arranged, standardized, and scaled in numerous ways. Seemingly, the more data points, context, and trends displayed, the more sparklines take life, as the ones taken from the January, 3, 2016 Market Watch website below.²⁹

²⁸ *Id.* at 50.

²⁹ MARKET WATCH (Jan. 3 2016), <http://www.marketwatch.com/>.



With additional temporal context, the sparklines for BVA and OMHA can also gain meaning.

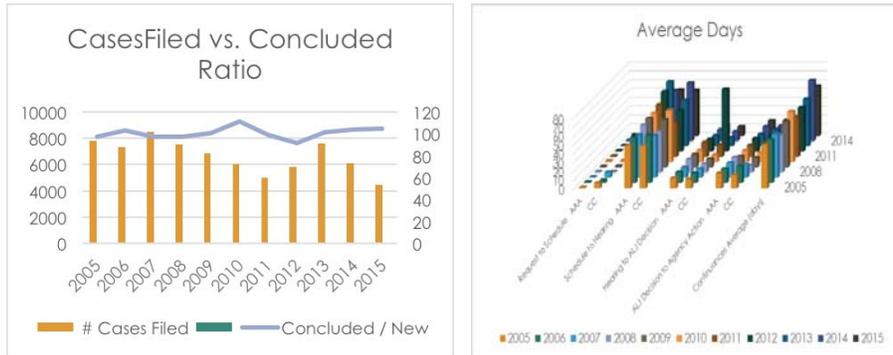
BVA Metrics	2011	Five Years	2015	low	high
Notices of Disagreement	122,663		146,032	111,641	146,032
Board Decisions	48,588		57,600	41,910	57,600

One may wonder at first whether administrative law agencies have the sophistication of data that rises to the complexity of the financial markets; but federal agencies, such as OMHA, BVA, and the U.S. Citizens and Immigration Services adjudicate many thousands of annual appeals.³⁰ Even for modest sized agencies, data density, context, insight, and meaning can increase when analyzing trends over time. For example, I previously noted that the AZOAH has posted annual reports going back to 1996.³¹ Each report displays single year data about the number of cases filed, the yearly conclusion ratio (number of cases incoming / number of cases disposed), and several related metrics about timelines while a case is in the hearing phase.³² If we extract ten-year data from those reports, then the display of these metrics might create some very clunky traditional Excel charts.

³⁰The Office of Social Security Administration posts Health Data sets for its Office of Hearings and Appeals. A user can download over a hundred thousand of data points into a .csv file that can make for some rich sparklines. *Appeals to the AC as a Percentage of Appealable Hearing Level Dispositions*, SOC. SEC. ADMIN., https://ssa.gov/appeals/DataSets/AC01_RR_Appealable_HO_Dispositions.html (last visited Apr. 7, 2016).

³¹ See AZOH *supra* notes 8-10 and accompanying text. <http://www.azoah.com/stats.html>

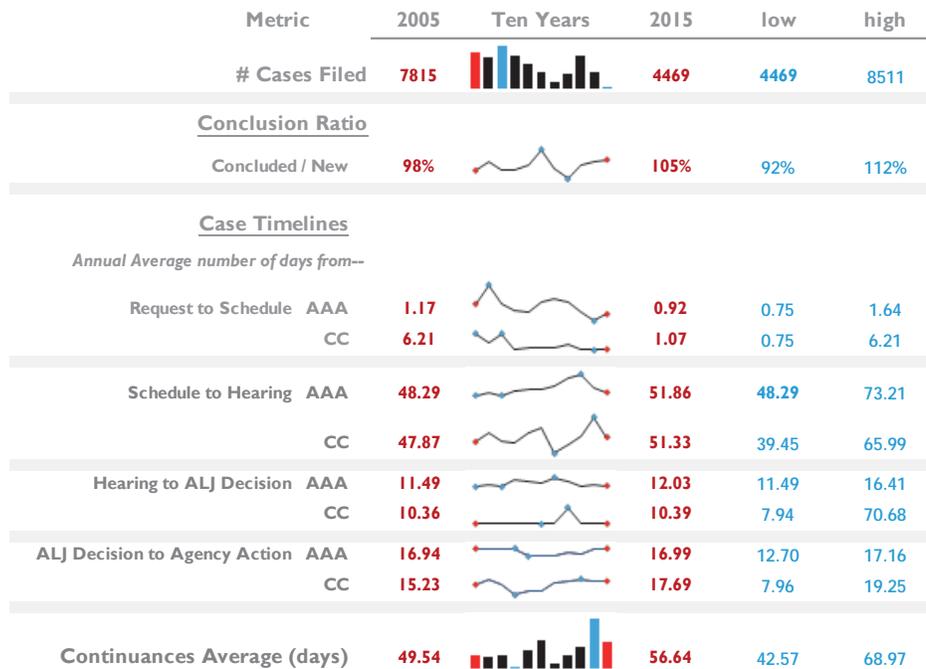
³² *Id.*



The charts above may exaggerate the point that standard spreadsheet software does not facilitate effective displays of dense data. As you find these cluttered charts in many reports, however, it is not overstated very much. At least, the combination chart on the left shows a relationship between the number of cases that are filed each year compared to the conclusion ratio of cases; combination charts such as this one are rare in agency reports, which seldom display cause and effect relationships. Of course, the chart on the right that displays the average number of days a case is in a particular phase of the appeal is undecipherable, even when the chart is magnified and the font adjusted. Further, both charts have to be enlarged on a page for a reader to pick out the details of the values, legend, and axis information.

A group of sparklines below, however, can effectively transmit the ten-year performance for AZOAH for the same data contained in the above excel charts.

Arizona Office of Hearings and Appeals Dashboard³³



It is helpful to remember that the display above presents AZOAH data for ten years that the agency showed for only one year in the section of the report I discussed in Part II. The sparkline dashboard above has many advantages over traditional agency displays that one might see for metrics displayed in a single year or for multiple years using spreadsheet software graphs and column charts. The ten-year history of the metrics above reduces recency bias, which may cause a reader to overweight recent case data in making conclusions about agency operations.³⁴ A reader can scan the table, make comparisons, and look for patterns or deviations in patterns of the metrics through

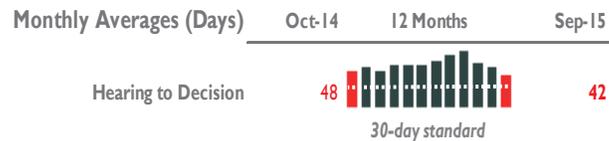
³³To show that these displays are possible with accessible tools, I have created all sparklines for case data using MicroSoft Excel, which began offering the sparkline tool with Office 2010. Office 2013 versions include some financial templates with sparklines. All the sparklines in this section use Gil Sans font and use the Excel sparkline function for highlighting first and last points, and high and low points. I completed the rest of the formatting with Excel cell drawing tools.

³⁴ TUFTE, *supra* note 22, at 50.

the years.³⁵ Moreover, this analysis engages the reader in the performance story of the agency and provokes questions about data patterns, outliers, or how one set of metrics might provide a context for other sets of data. Searching for answers to these questions might provide some more nuanced context and insights for future data collection and display.³⁶

Individual sparklines can also be embedded in sentences to make inscribed displays more descriptive. For example, in a sentence analyzing case trends, the AZOAH might state that the number of cases filed in 2015  4,496 reflect a 10-year low. As Tufte explains, the sparkline “depicts like a graph but means like a word,” as if to say, “Here are the last ten years of cases filed, with the most recent number of cases filed indicated by the red column.”³⁷

Agencies can provide additional context to sparklines that display metrics about case intake, average processing times, or the number of decisions issued. For example, many agencies have timelines mandated by regulatory authority. In the example below, an agency shows that its average monthly time from the close of a hearing to the time a judge issues a decision does not meet the 30-day regulatory standard.



Averages may often mask the full data range of a variable and depicting ranges within data can provide additional context.³⁸ For example, assume an agency has calculated that in a given year, it can

³⁵ There appears to be a cause and effect relationship between the conclusion rate and the number of cases filed each year. Further, it appears that the continuance average does not significantly affect significantly the conclusion ratio.

³⁶ For example, what is the average number of days a case is in the hearing phase? And how does the continuances average affect that average?

³⁷ TUFTE, *supra* note 22, at 49.

³⁸ TUFTE, *supra* note 22, at 60-61.

dispose of 300-450 cases per month to meet its obligations.³⁹ In making a point about the need for more resources to reduce its backlog, the agency can display the actual number of cases received each month compared to its output capability. Bands that indicate ranges can be raw numbers, standard deviations, or high and low averages. Here, the grey band indicates the agency's ability to dispose of cases each month.



Using the sparkline above, an agency may argue that for most months its intake exceeds its output capacity, which causes a backlog. An additional sparkline could show the backlog over the same period. Agencies are only limited by their creativity and analytical perspectives when employing sparklines to display trend data. The best models come from newspapers and websites that display financial data.⁴⁰

Appeal agencies often track data about rates of favorable or unfavorable rulings. This provides another sparkline opportunity to display win / loss data.⁴¹ For example, at NAD, a participant in an agricultural program may seek an appealability ruling from NAD that challenges an agency view that an adverse decision is not appealable. NAD may rule that the agency decision is appealable (win for the appellant) or not appealable (loss for the appellant). If NAD wants to encourage potential appellants to challenge such an agency ruling, it could describe that NAD ruled favorably for challenges to the agency decision in 78% of cases  42-12 in FY 2014.⁴²

With thoughtful color shading, win whiskers on sparklines can also indicate **partially favorable** rulings.⁴³

³⁹ The Board of Veterans Appeals calculates and presents such a metric in every annual report: see BD. OF VETERANS' APPEALS ANNUAL REPORT, *supra* note 14, at 26.

⁴⁰ TUFTE, *supra* note 22, at 51.

⁴¹ *See Id.*

⁴² *Id.*



Like NAD, many appeal agencies have a two-step appeals process. A simple underline of a win/loss whisker can show that the final agency appeal outcome was the result of a second-level appeal.⁴⁴



Win/loss sparklines gain power and context, when combined with other trend sparklines.⁴⁵ For appeals agencies, this may be helpful when comparing the wins and losses for various programs or appeal categories. For example, to display whether an appellant will be successful at the first- and second-level of the NAD appeals process, the agency can display a five-year trend rate, accompanied by the individual rulings it issued for a specific assistance farm program under its jurisdiction. Below, a combination of sparklines about appellant success rates can provide a more nuanced picture of agency appeal services.⁴⁶



The display above provides information about individual program rulings in 2015, as well as opportunities for an appellant to receive relief in other forms. A provision in the NAD statute provides that, even though an appellant may eventually lose a ruling in the appeals

⁴³ The Office of Medicare and Appeals presents this data on its website at http://www.hhs.gov/omha/important_notice_regarding_adjudication_timeframes.html#adjudication.

⁴⁴ See TUFTE *supra* note 22, at 54. This is similar to indicating that a baseball team played an away game.

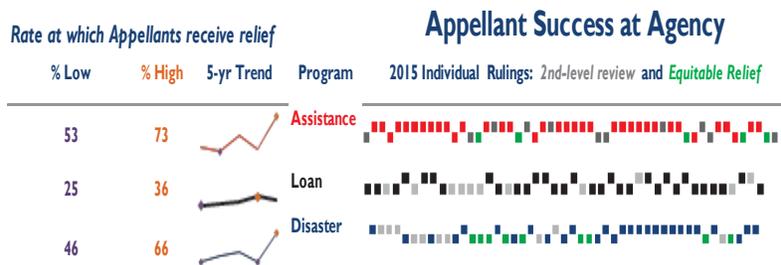
⁴⁵ *Id.* “A useful strategy for data displays is to *multiply a good design.*” (emphasis in original).

⁴⁶ *Id.*

process, NAD may still grant the appellant equitable relief at the second-level review. Equitable relief grants the appellant a waiver of program deadlines or grants money if appellant showed good faith effort to achieve a farm program requirement. The agency wants to encourage farmers and ranchers to appeal, even if they do not believe they have complied with all technical aspects of a program, because appellants may still receive some form of relief. Although appellants may lose NAD’s program ruling, the trend sparklines to the left portray statistics that also include NAD’s authority to grant equitable relief. Thus, either a positive program ruling or the granting of equitable relief both qualify as forms of relief.

The trend line above shows a recent rise in the percentage of cases where appellants receive relief for the assistance program. The individual win/loss whisker sparklines for cases in 2015 to the right confirm the reason for the recent upward rise in relief for the assistance program. As demonstrated by the loss whiskers to the right, it is true that appellants may lose a ruling in the appeal process, but the green loss whiskers show cases in which the agency ruled against the appellant in 2015 but still provided some form of equitable relief. Moreover, the reader can analyze all losses at the second-level review, because the gray-shaded whiskers mark wins or losses at the second-level review. As stated previously, the regulations provide that equitable relief can only be granted at the second level of the appeals process; so it can be assumed that green whiskers are also second-level review outcomes. Finally, an appellant who receives a favorable ruling does not require equitable relief, and that is why there are only green loss whiskers—no equitable relief-designated wins.

The nuances of winning or losing an appeal, but still receiving relief, presents a data display challenge for NAD to communicate to potential appellants who are making risk-based decisions about a full range of insurance, disaster prevention, and other activities for their farms. The sparkline below includes the assistance program displayed above and integrates specific 2015 win/loss data and five-year equitable relief information trends into one data display for three groups of programs.



The three sparklines above tell a story about appellants' chance of success (relief) when appealing to the agency for assistance, loan, and disaster programs. As the win/loss sparkline for disaster programs shows, equitable relief was granted nine out of fifteen times the appellant lost in a NAD ruling for a disaster program appeal that reached second-level review in 2015. This high relief rate contributed to the overall rise in the success rate of an appellant in a disaster program, which is confirmed by the upward rising sparkline trend to the left. Judging by the display, an appellant in a disaster or an assistance program has a fair chance of success or relief in some form through an appeal. Finally, unlike the assistance and disaster programs, where green whiskers are displayed, the loan program has no regulatory provisions that permit NAD to grant equitable relief, confirmed by the absence of green **loss** whiskers for 2015 rulings. Sparkline displays, like the one above, challenge readers to engage multi-variate data, contextualize the data in time or for other variables, such as a specific program or an individual appeals win or loss, and draw conclusions.⁴⁷ These conclusions are not arbitrary, however, because NAD has designed the display to persuade the reader to make an analysis about favorable outcomes of NAD appeals and how the nuances of equitable relief increase those chances.⁴⁸

In the discussion above, sparklines demonstrate that they can improve or replace many of the standard statistical charts and graphs provided by spreadsheet software packages, because they are more descriptive, efficient, and multivariate. They compel a reader to make analytical contrasts and comparisons at very granular levels about the data presented.⁴⁹ When engaging a trend line with many data points, often with a range provided, the data seems active, urging the reader to wonder, if not predict, the next data points.⁵⁰ Agencies committed to presenting this new interactive data become rigorous about the data they collect and display. They begin to ask questions about how they might contextualize their normal performance metrics, or how can they track more data, drill down to the smallest data element, and visually display it. Most important,

⁴⁷ *See Id.* at 50-51.

⁴⁸ *See Id.*

⁴⁹ *See Id.* at 54.

⁵⁰ “*The idea is to be approximately right rather than exactly wrong.*” *Id.* at 50 (emphasis in original).

they ask how they can leverage these descriptions to make persuasive arguments to their constituencies about their organizational values, mission, and activities. These questions create a mindset that give rise to the use of analytics for a CMS.

IV. ANALYTICS FOR ADMINISTRATIVE LAW AGENCIES

A. The Transition to Analytics

Once an agency commits to information systems that collect more frequent, voluminous, and variable data, and it begins to parse and integrate this data into complex displays, it dons an analytics mindset. This mindset generates sophisticated and nuanced descriptions that support an agency case about transparency and accountability. As a result, agency performance metrics also become more refined. What were previously simple agency metrics now become more intricate. Where previously an agency may have been content to tabulate and report the number of cases it receives annually, in the new analytics mindset, it wants to know the number of cases it receives yearly, quarterly, daily, and even hourly. It wants to display those receipts in sparklines, as one might see the stock market being tracked on the Bloomberg cable channel. Further, it wants to know how the pattern of these receipts by appeal type, program, state, or county compare with other data. It may even want to analyze these patterns against the background of other data categories: Do appeals for different categories of veterans' benefits, for example, have a cause and effect relationship with the age groups of veterans from different periods of war? Do appeals for farming programs match the growing seasons or insurance filing deadlines for different crops? Each agency operates within its own context.

Analytics is not solely devoted to the visual display of data; but its broader goal of clustering, analyzing, scoring, and transforming data into insight complements the principles of effective visual display of data. Analytics discovers and communicates meaningful patterns in data, and many case management systems are business intelligence applications that contain analytics modules. Legal Administration agencies do not require integrated case management systems, however, to 'do analytics' or present meaningful visual

displays of data.⁵¹ For many agencies that do not have sophisticated case management systems, the discussion about the principles of sparklines and insightful data display can drive data collection and analysis at the granular level. Agencies can collect and present different kinds of data from a variety of databases, spreadsheets, or other data sources to which agencies may have access. Good command of mathematical and statistical knowledge, combined with a sophisticated use of spreadsheet tools, can foster very powerful analytics for those agencies. However, when an agency adopts the analytics mindset, a CMS with an analytics module will follow soon.

The most commonly recognized components of analytics modules are visual. Analytics modules afford users the ability to create dashboards that group together various micro-graphs, sparklines, charts, and tables. Moreover, the features of analytic modules take user data interactivity to the next level. Rather than engaging readers to scan and compare data points, trend lines, and context at detailed levels with sparklines, an analytics user can point and click on charts and drill down, or segment data and see the data values and the format of the display change right in front of their eyes.⁵² An agency without analytics software may need to rely on several system queries to create a meaningful sparkline, which can be time-consuming. Further, although I have described sparklines as more active than other data displays, in the end, they are fixed in time and require reiteration to keep the display current. After constructing the data relationships for an analytics module, however, the user can see the data and the display evolve in real time—with each point and click on the fly—a powerful demonstration of an agency ability to govern its data collection and visual displays. It is no mystery why agencies desire this level of proficiency.

Before reaping the benefits of analytic software features, however, a smooth contractor demonstration of an analytics module may cause agency planners to underestimate the extensive data planning and analysis that is required to implement this initiative. In addition to purchasing and installing the software, analytics can place difficult demands on many administrative law agencies that have

⁵¹ John W. Foreman devotes an entire book on using spreadsheets to achieve the same goals of analytics: JOHN W. FOREMAN, *DATA SMART* (2014).

⁵² Anyone who has attended a contractor presentation about analytics can testify that this feature is the core deliverable of its sales pitch.

modest budgets, sparse technical resources, and an incomplete ability to marry agency data metrics with business processes required for analytics data development.

B. The Analytics Framework

The ensuing discussion about an analytics culture, and our experience with analytics at NAD, targets a typical Legal Administration agency that is installing an enterprise-wide CMS with analytics capability, or an agency that operates an existing CMS and foresees purchasing an analytics module for its current agency data structure. Consciously or unconsciously, an agency in one of these categories is adopting an analytics mindset that will eventually transform its operations and metrics, reflected by the display of its data. The size of agencies in this category may vary, depending on the number of system users, workload, the multitude of its programs, and other agency IT structures. Some very large federal agencies have several IT systems that collect data and transmit information to a centralized database solely for the purpose of data management. Although many of the principles in this discussion can apply to all levels of data structures, NAD's experience is most relevant to agencies with an existing or envisioned enterprise CMS that collects and reports all case management data under one system.

Legal Administration agencies that progress to an analytics mindset conform to an analytics evolution found in other business organizations. This evolution provides a broader context to help agency analytics initiatives, because the number of Legal Administration agencies evolving into analytics is sparse. For example, even though NAD is on the cusp of evolving analytics, its effort has comprised many trials and errors, and it now only begins to comprehend the longer term implications of embedding analytics into its operational activities. This can be a humbling realization.

Bill Franks, in the *Analytical Revolution*, categorizes three generations of analytics, a framework created by the International Institute for Analytics.⁵³ The first generation, Analytics 1.0, is primarily descriptive, composed of internal sources of small,

⁵³ BILL FRANKS, *THE ANALYTICS REVOLUTION: HOW TO IMPROVE YOUR BUSINESS BY MAKING ANALYTICS OPERATIONAL IN THE BIG DATA ERA*, 3-31 (2014).

structured data that often requires management analysts or other expert employees to aid decisions makers.⁵⁴ The data is mostly transactional, and the reporting of data requires metadata, or a reformatting of the data, to make sense of reports; therefore, it takes a while to make the data available for analysis.⁵⁵ In this generation, developing reports are time consuming and expensive, because creating a report requires some analyst from a headquarters to gather the requirement, configure the report, and enable it for users.⁵⁶ Different users may need variations of the same report.⁵⁷ Analytics 1.0 characterizes organization activities that have existed for many years, and most businesses have moved beyond this stage.⁵⁸ Businesses that require data for decision support to react quickly to customer needs for services cannot survive in the Analytics 1.0 generation.

Yet, most Legal Administration agencies that have an existing CMS, without an analytics module, are in the Analytics 1.0 generation (If an agency plans to install a new CMS with analytics, then it will probably design a data and reporting structure that bypasses this generation). Historically, a typical agency CMS in this generation probably migrated legacy system information into a new case management data structure that reflects its business process. There are multiple screens, user roles, system administrators, and search tools for users to enter data and retrieve information, based upon the agency mission, activities, and user needs. These agency systems have standard reports, developed when the systems were installed, which probably contain tables and charts or standard search tools. These reports may include depictions of data points for a limited number of variables. They are very similar to the types of reports found in spreadsheet software, which I have discussed as being static. A CMS in the Analytics 1.0 generation will also have *ad hoc* reporting capabilities that may require a user, although it will probably be some analyst, to learn Structured Query Language (SQL) or Oracle Query skills to query the CMS and download results into

⁵⁴ *Id.* at 10.

⁵⁵ *Id.*

⁵⁶ *Id.* at 10-11.

⁵⁷ *Id.* at 12

⁵⁸ *See Id.* at 10.

spreadsheets for further analysis and data display. As these systems evolve over time, they generally gather more data but not different kinds of data. For example, they may add or modify data fields for programs, or they may add a data entry screen or two, but, for the most part, new data is confined to the existing data structure.⁵⁹ Similarly, there may be additional, or varying reports, but the basic metrics for agencies, such as the number of cases docketed, assignment information, statistics about timeliness, etc. remain the same.

Analytics 2.0 generation systems emerged with big data challenges.⁶⁰ These systems access more complex, larger, and unstructured data sources.⁶¹ They contain modules that provide new analytical and computational capabilities, and they usually create more online offerings for products or services.⁶² Organizations with these systems become more involved in data governance processes and hire employees, or change the duties of existing employees, to focus on the quality and relevance of data.⁶³ They invest resources into analyzing how the data integrates into strategic planning and customer services. Data in this generation of systems may include social media data, documents, photos, images, or sensor data.

Most agency case management systems with analytic features probably fall into the Analytics 2.0 generation.⁶⁴ These systems collect more information on appellants, activities about the case process, and store indexed documents or audio files.⁶⁵ Systems in this category may have document management modules that store and manage documents through a separate module and process within the CMS. Agencies may have partial or complete electronic case file management. Moreover, in creating new documents within the agency, such as notices, letters, or draft judgments, agencies

⁵⁹ : U.S. Census Bureau, CSPro Date Entry User Guide, INTERNATIONAL PROGRAMS CENTER FOR TECHNICAL ASSISTANCE POPULATION DIVISION, <http://www.census.gov/population/international/files/cspro/csent50.pdf>.

⁶⁰ FRANKS, *supra* note 55 at 25.

⁶¹ *Id.*

⁶² *Id.*

⁶³ *Id.*

⁶⁴ *See* FRANKS, *supra* note 55 at 3-33.

⁶⁵ *Id.*

establish a document business process and users have access or editing privileges, depending upon their roles and authorities. Further, there may be a more sophisticated reporting module that provides an increased user-friendly interface for a broader array of users to query the system for information. The reporting module may include more computational data criteria, so that a report can display mean, median, mode, sums, min, and max calculations that can be visually displayed. Finally, an analytics 2.0 agency probably has some capability that permits customers to file cases electronically (eFile) and integrates the eFile data into its database. Other electronic services in an agency 2.0 generation would include email notifications, with attached documents, or other media, that are triggered by data field entries or other conditions.⁶⁶

Agencies that fall into the 2.0 generation are collecting different kinds of data and making the information more accessible to users.⁶⁷ They can calculate how long a notice or a judgment remains in the drafting phase. They can compare the demographics and timing of individuals using eFile against those that file cases through other means. They embed content from database information in email notices and attach other forms of media, such as documents.⁶⁸ Thus, they can measure and analyze the frequency and content of their customer contacts.⁶⁹ The potential data display for these activities are richer and more meaningful, and they enhance an agency case about transparency and accountability.⁷⁰ These behaviors reflect the beginning of a deeper analytics mindset.

In addition to the agency CMS capabilities described above for a 2.0 generation, I would also include an agency CMS that has a specific analytics software module. Agencies with specific analytics software have advanced more on the Analytics Framework—perhaps it is an agency in the 2.5 generation. The analytic software provides an agency the ability to manipulate visual data and displays that go one-step further than static reports and outputs, like those found from query tools or reporting modules in a typical CMS. Analytics features

⁶⁶ *Id.* at 25.

⁶⁷ *Id.*

⁶⁸ *Id.*

⁶⁹ FRANKS, *supra* note 55 at 25.

⁷⁰ Omer Tene & Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 NW. J. TECH. & INTELL. PROP. 239, 243 (2013).

provide alternative data views, both data points and computational, which permit users the ability to adjust these views on the fly. The most common use of these tools is when users drill down on the visual presentation of data to filter data directly from the charts or visual displays. For example, it is typical that an agency would permit a user to point and click on a line graph depicting the number of cases received.⁷¹ Additional clicks would allow the user to drill down to the subset of cases received in a region, by program, or date. More than the tool itself, the ability to structure data for visual display, making it more accessible to the entire agency or customers, marks an agency in the 2.5 generation.

Analytics 3.0 evolves the previous two generations (three generations, if you count 2.5) but, additionally, seamlessly blends the data from all data sources and uses the results to become embedded processes or business rules to make operational decisions. The analytics in this generation are a strategic asset in themselves, available to decision-makers or customers at every level of the organization or at the point of service delivery.⁷² Because analytics in this generation are more prescriptive⁷³, rather than descriptive, it is difficult to envision how this generation of analytics would apply to case management systems or legal administrative agency operations. In my review of systems throughout federal and state agencies, I have not identified any Analytic 3.0 models.⁷⁴ In fact, very few agencies have implemented analytic modules that rise to the level of 2.5.⁷⁵ A thought experiment on this issue, however, might generate the list below, which proposes five hypothetical behaviors that signify an Analytics 3.0-level at an administrative law agency:

- The CMS assigns a Judge a case based on metrics related to timeliness, workload, or area of expertise. Or the person responsible for assigning cases can point and click at various assignment options and visually assess how each option will affect other agency performance metrics.

⁷¹ *Id.*

⁷² *Id.*

⁷³ FRANKS, *supra* note 55 at 17.

⁷⁴ This review is by no means exhaustive. It is based upon my attendances at conferences or through various outreach activities with other administrative law agencies.

⁷⁵ Tene, *supra* note 72, at 243.

- Adjudicators access web-based analytic interfaces to explore outcomes of past cases based upon the program, legal issues, semantic searches, and prior outcomes of similar cases;

- Accessing the same data above, potential agency customers or appellants can explore agency data with analytic tools to perform a risk assessment for their appeal activities.

- When customers file an appeal electronically and enter appropriate data, they receive prior agency decisions that fit their case profiles. Further, customers can adjust their profiles to receive alternative decisions.

- Potential customers have access to analytics tools about hearing outcomes and the length of hearings to make a decision about selecting the type of hearing (e.g. in-person, video, telephonic, review of the record, etc.) they choose for an appeal.

The few examples of Analytics 3.0 above may foster agency goals of transparency and high quality service to customers and constituencies, but many agency leaders, adjudicators, and employees may think these examples cross the line of objective adjudication.⁷⁶ As part of an adjudicative organizational culture, agencies value fair and impartial adjudications, based upon consideration of the law and the factual pattern of individual cases one by one, even though, in aggregate, outcomes and other insights can be derived from data patterns.⁷⁷ Although it is uncertain how these issues in the future will be resolved, there can be little doubt that the growth of analytics will transform how agencies make operational decisions and interact with the public.⁷⁸

V. ANALYTICS CASE STUDY AND SOME LESSONS-LEARNED AT THE NATIONAL APPEALS DIVISION

A. *The Current System: Goals and Expectations for Analytics*

The National Appeals Division (NAD) is an emerging 2.5 analytics generation agency. As a modest-sized appeals agency, by

⁷⁶ *Id.*

⁷⁷ Tene, *supra* note 72 at 29.

⁷⁸ FRANKS, *supra* note 55 at 25.

many standards, the NAD CMS is robust and flexible. Through data and metadata entry, it tracks hundreds of case activities during the life cycle of a single case.⁷⁹ The system has tabs, screens, parent-child relationships, and business rules that generally follow the phases of a case, as an appellant goes through the appeals process.⁸⁰ The system also organizes data tables and hierarchy accordingly.⁸¹ For example, an appellant goes through the phases of a pre-hearing, first-level appeal, second-level appeal (review), and post appeal adjudications.⁸² Thus, there are separate tables for data collected at each phase—usually indexed by an automatically-generated case number. The system also stores and indexes documents and audio files. Through the CMS, NAD also provides full eFile services.⁸³ NAD also engages customers through other electronic means, as data input into the system triggers email notifications, with appropriate attachments, to customers, agency employees, and members of the public.⁸⁴ These emails contain information about assignments, case information, and provide other information about all aspects of case activities.⁸⁵ Further, NAD's CMS has a reporting module, supported by extensive search capabilities that provide real time outputs for every data field in the system.⁸⁶ Finally, the CMS is linked to the NAD website for public search of prior NAD decisions, filtered by the type of decision, decision date, and with text searching capability.

⁷⁹ I wish to acknowledge our private partners at MicroPact Engineering for electronic case management at NAD. The NAD CMS replaced legacy systems and has grown in size and complexity since 2005. MicroPact continues to provide technical assistance to NAD for the development of analytics in a collaborative partnership. Its website can be found at <http://www.micropact.com/>

⁸⁰ National Appeals Division, *NAD Rules of Procedure and Regulations*, U.S. DEP.T OF AGRICULTURE, http://www.nad.usda.gov/lr_rules_procedures.html. [hereinafter Nat'l Appeals Div.].

⁸¹ *Id.*

⁸² *Id.*

⁸³ *Id.*

⁸⁴ *Id.*

⁸⁵ *Id.*

⁸⁶The volume and variability of data is comprehensive, but these reports look similar to the charts and graphs one might find in spreadsheet packages. Typical report modules in a CMS suffer from the same paucity of density, context, and activity as do stock spreadsheet templates.

Although NAD managers were satisfied with system performance, the promise of an analytics module impressed the NAD project team. In concept, NAD employees could access the module and engage some striking, preset visual displays of data in categories that were already previously established in existing CMS reports. In fact, the NAD team's first impression from an analytics demonstration was that this feature was simply a more user-friendly reporting module for all users. Using analytic features, employees could manipulate the data visually to gain further insight about a specific area of interest.⁸⁷ Moreover, users could point and click their way across several sets of data, focusing on results, rather than worrying about how to run various reports or queries.

One existing limitation of NAD's reporting structure, as is the case with most current CMS systems, was that NAD users had to run multiple reports or do a sophisticated query and download data into spreadsheets—and then manipulate the data again with spreadsheet tools—to retrieve multiple sets of data.⁸⁸ Often NAD managers and other administrative employees had to seek help from management analysts to set up customized searches or to run queries.⁸⁹ There are over 120 data tables with dozens of fields per table.⁹⁰ In reality, most employees only needed to access a small sample, but even posted queries would provide results with column names, such as “C.Date.Decision.due,” which many found inscrutable.⁹¹ Of course, these queries and fields could be reformatted to become more accessible, but NAD did not have unlimited resources for extensive data massage across the spectrum of reporting.

Reports and queries were also increasing.⁹² As NAD had introduced more data variations into the system to support its increased services, this proliferation emerged to track and report activities.⁹³ When the CMS was first implemented, there were approximately twelve queries and reports, which did not include

⁸⁷ Tene, *supra* note 72 at 29.

⁸⁸ Nat'l Appeals Div. *supra* note 80, at 32.

⁸⁹ *Id.*

⁹⁰ *Id.*

⁹¹ *Id.*

⁹² *Id.*

⁹³ Nat'l Appeals Div. *supra* note 80, at 32.

internal CMS search functions. Several years later, the number increased to over fifty. Presently, when all forms of reports, text search capabilities, and queries are included, the number of different kinds of searches and established reports and queries reach into the hundreds.

The demands for training were also increasing, especially in the areas of providing employees and managers a more thorough understanding about how the data related to the business process in NAD. For example, what does an input / output ratio mean for its cases? How can there be more hearings held in a year than there were cases filed? Can we run and present 'Report X' to an outreach organization to portray whether its members meet or exceed the success rate of other organizations? These were the challenges that had developed because of NAD's internal big data dilemma, and they are common to many agencies that begin to collect more case and customer information.⁹⁴ The data was getting more voluminous, variable, and frequent. The goal for analytics was to reduce the reliance on reports and decrease training needs.⁹⁵ Analytics would provide interactive metrics information to all employees and become a part of the metrics dialogue of the organization.

The first phase of NAD's analytics deployment has been operating in the development environment for several months. NAD will launch in production in spring 2016. Even before analytics launches, the NAD project team has Phase II and III modifications planned. In the NAD experience, future changes with analytics outpace changes that occurred with other CMS activities for traditional agency data collection and reporting. For the most part, legal administration agencies have more process stability over time, when compared to other businesses. Like most agencies, NAD's business process and case management is inexorably tied to its statutory authority and procedural rules, which are slow to change. In the analytics 2.5 generation, however, the data about an aspect of the business process is less important than user interaction with the data—and that interaction privileges visual display. The difference may seem subtle, but the ramifications are consequential.

⁹⁴ Nat'l Appeals Div. *supra* note 80, at 32.

⁹⁵ *Id.* at 29.

A quick vignette can clarify the growing importance of visual data interaction in analytics. Before the implementation of analytics software, a NAD Regional Director (RD) might want to know the number of cases assigned to a region that had a disposition before a full hearing. In other words, the RD wants to know how many cases dropped out of the appeals pipeline before an administrative judge conducted a hearing. Traditionally, a report could show the number of cases received in the region and the number of hearings conducted for those cases. Probably, with minimal management analyst help, there would be a stacked column chart created with both variables, and there might even be columns that span several years to reflect a trend, but the analysis would probably stop there.

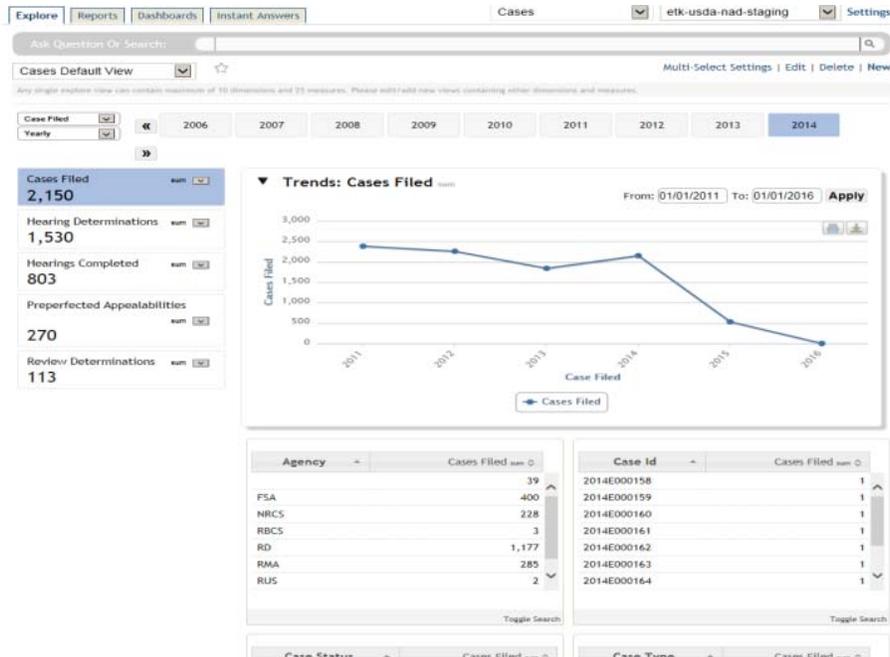
In the new analytics environment, however, the question differs slightly, and it provokes an internal conversation, because instead of the RD running a stock report or requesting a variation of a report from a management analyst, the RD is now clicking several times on a relevant analytics view to drill down and see various visual displays of the information. The internal dialogue might sound something like this:

I want to see the trend line of the number of cases filed for the past five years in my region. Click. Ok, show me a line for the number of hearings held on the same chart. Click. Oh, that's interesting. I wonder if that result is a constant rate for all the programs in my region. Click. Hmmmm. I wonder if these cases go to a pre-hearing. Click. You know, it would be nice to see this one line displayed as an area chart, and then I could see a trend line plotted against that area chart. Click. Oh, I guess I can't do that.

In the internal dialogue above, the RD's visual interaction with the data becomes the fundamental element of analytics design and delivery. Future phases of enhanced analytics implementation will privilege visual interaction. In order to accommodate the RD's desire to see the data display a certain way, few changes, if any, to the data design, data entry, or computation of data values will be necessary. Instead, the analytics module will have to be modified to show existing data the way the RD wants to see it. These are interactive reports, one level removed from traditional data displays, and they mark agencies moving through the analytics framework.

B. Description of NAD Analytics

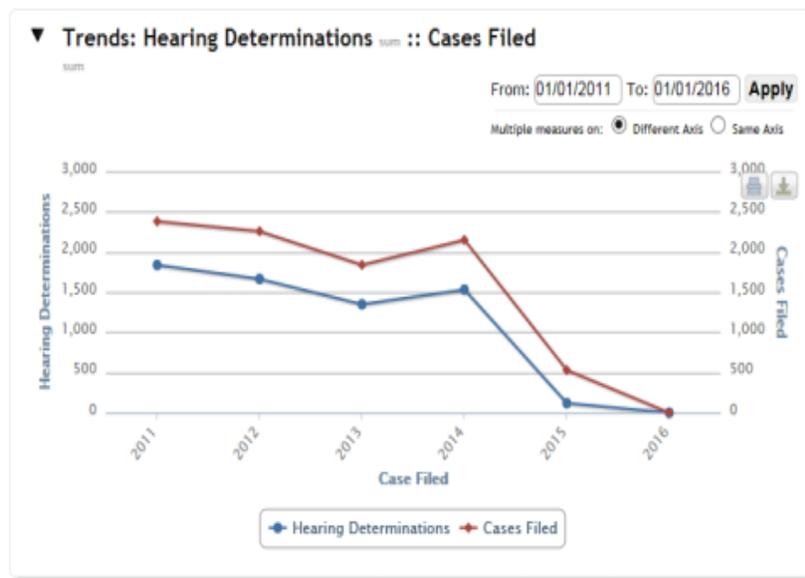
For the NAD CMS, the basic analytics interface is the Explorer Screen, which, as the name suggests, permits users to explore data within a dataset:



In the initial implementation, NAD has created datasets that include general case information, prehearings, hearings (first-level appeals), reviews (second-level appeals), activities, and electronic filing. Information at the top of the Explorer tab relates to time dimensions, while information on the left side are measures (metrics) of quantitative data (sum, count, average, etc.) computed within a dataset for the time dimension selected. For example, the Explorer tab above shows measures of counts of cases filed, hearing determinations issued, hearings completed, and second-level review determinations issued in the cases dataset for 2014. The bottom of the Explorer tab, partially depicted, shows Dimensions, which are business categories of potential analysis within a dataset. For example, the ‘Cases Filed’ dataset, depicted above, includes

dimensions for agency, Case ID, program, outcome, region, state, and other information. When a user selects one or more dimensions, the measures and the trend chart adjust to reflect the data values associated with the dimensions selected. This is analogous to the sort and filter functions found in spreadsheet tools, although more visual and interactive. When a user selects one of these measures, a trend line depicts the five-year history of the measure.

Users have the option to depict trends of several measures simultaneously.



For example, the trend chart above depicts the number of cases filed with the number of hearing determinations issued over a five-year period. Finally, users are able to drag a cursor over points on the trend line and see a numerical presentation of the data value.

Although one goal of analytics is to reduce long-term training, as with any new application, employees will require initial training on the interface. Concepts and terminology about *datasets*, *dimensions*,

and *measures* may create a new paradigm for CMS users.⁹⁶ Further, employees will need a sound understanding of how the data is organized in the datasets, which may differ from the normal data hierarchy and entity relationships used for traditional CMS reports or queries. But the NAD experience is that, after some initial training and practical exercises, users are quick to learn the various points and clicks necessary to reveal some interesting and useful information.

C. Datasets

As an analytics view leverages a domain of data from the CMS database, datasets become the lynchpin of analytics. For agencies planning a new installation with analytics functions, as much as possible, the desired analytics view should drive the design of the database.⁹⁷ In other words, the design should be data-driven, not process-driven. For example, if an agency wants to view a trend line of the number of ‘cases received’ and plot it against an area chart of case backlogs, then project managers need to ensure that data tables are designed or indexed with the appropriate fields to plot both variables together.⁹⁸ Additional consideration in data design would ensure that the drill down and filter options (e.g. by date, program, geography, sub-agency region, etc.) for these variables are also possible. Further, project managers should pay special attention to data relationships that cut across traditional case processing boundaries. Often, CMS planning silos the data fields for activities in these case phases that may limit complex views of this data in analytics.⁹⁹ For example, it may be helpful for an appeals agency to

⁹⁶ These terms are used in some software packages, such as Tableau or Crystal Reports, but most agencies use typical spreadsheet tools to manipulate data that do not employ these concepts.

⁹⁷ An agency planning a new CMS will most likely be migrating from a legacy system and creating a new data design anyway.

⁹⁸ Agencies have special challenges displaying visual data that describes a backlog as a variable. This is because the data design needs to capture the date a case became docketed AND null values that show a final disposition has not been rendered or some other signifier that shows a case is still active. It is also helpful to distinguish a backlog from an inventory, because not all cases docketed are late.

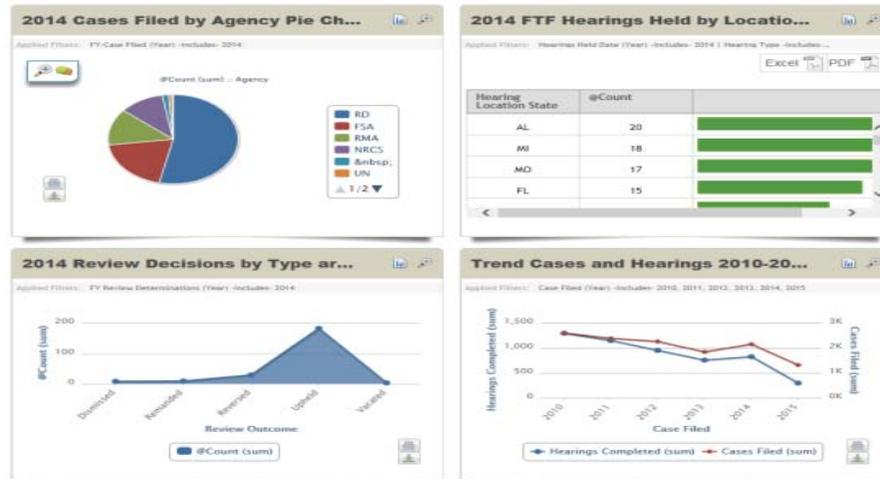
⁹⁹ The CMS challenge here is to maintain one-to-one and one-to-many relationships of the data. For example, because cases may be remanded back to a

view the number of ‘cases filed’ on the same chart as the number of first-level and second-level appeals. This analysis may cut across several proposed data tables and data entry screens that segregate these activities. Finally, the design of tables should take into account computational data too. For example, a CMS may track the beginning and ending date of a hearing, but does not have a separate data field for the computation of the number of days in the hearing phase. Data designers need to pay special attention to these critical agency metrics and account for them by incorporating these computations in analytics measure functions.

Many agencies, however, may already operate a CMS but are just beginning to integrate analytics into the system. The analytics will still require new datasets, but current CMS table design and data relationships may limit more complex visual presentation and manipulation of data that would augment the power of analytics. On the one hand, an existing data design can be an advantage, because an agency in this category probably does not have to figure out how to integrate data design with its business process; the current CMS already reflects those decisions. However, as a disadvantage, the current data design probably creates a silo that will limit a complex analytics view, because the datasets will probably be confined to the existing tables.

There are two workarounds in analytics for agencies to address this limitation of a data silo that limits complex visual displays of data. The first workaround employs the superficial principle that analytic views, just like sparklines, can sit side-by-side on a dashboard. This arrangement alone provides context, because a thoughtful dashboard displays a variety of metrics that reflect the performance of an agency. Just as Tufte displayed a group of sparklines that reflect the health of a patient, through analytic dashboards, an agency can display its overall health through its performance metrics. The dashboard below assembles four displays about ‘cases filed,’ ‘face-to-face hearings by state,’ ‘second-level review outcomes,’ and a five-year trend of ‘hearings and cases filed.’

lower appeal tribunal, there could be many hearings associated with a single case; or several second-level reviews for first-level appeal.



The dashboard above draws from four separate datasets and thus visually overcomes the inability of the agency to display these metrics through one chart. The side-by-side view, if properly designed, provides the reader coherence and context for these metrics. Further, analytics modules usually have an attribute that permits designers to “link” views, similar to linking reports in software packages, such as Tableau and Crystal Reports, which pass parameters, such as a date, from one dataset analytics view to another. When viewing a chart in analytics, the user can select the linked report as a drill-down option, just like drilling down on a dimension, to see the various views appear in sequence, providing a user more context and multivariate data.

The second workaround for data silos is more technical: The agency may want to invest in the design of some new intermediate tables that collect data from several other tables for the purposes of creating robust datasets only for analytics. For example, if the data relationships of an agency do not permit a dataset that unifies all the activities associated with both the first- and second- level appeals processes in a case, then a third table can link the tables and group selected first- and second- level appeal information into one dataset. This technique may also facilitate computational data.

One simple example illustrates how creating intermediate data tables produces a more robust analytics design. In the NAD initial CMS data design, it is not possible to show both the number of cases received and the number of hearings held in the same period on one

chart in a cohesive analytics view.¹⁰⁰ There is a table and dataset for general case information, which contains the date a case was filed; additionally, there is a table and dataset for all activities relevant to hearing information, including the date a hearing was held.¹⁰¹ Moreover, these analytics datasets are organized around the existing table designs before analytics was envisioned.¹⁰² The tables are linked in the CMS, but in analytics, a trend line showing the number of cases filed and the number of hearings held in FY 2014 would show different results for each respective dataset.¹⁰³ In the case information dataset, the number of hearings held would be for those cases filed in FY 2014, which would not include data on hearings held for cases filed before FY 2014.¹⁰⁴ In the hearings dataset, the data for the number of cases filed would only show cases that were filed that had a hearing held in FY 2014, which would exclude cases filed that did not go to a hearing.¹⁰⁵ Therefore, both analytic views are incomplete.



One way to resolve the dilemma, as described before, is simply to include both views from the two datasets as part of one dashboard (left). This result may not be sufficient, however, because managers may want to drill down and view unified results on some common measures and dimensions for both

‘cases filed’ and ‘hearings held,’ such as region, state, or program. Thus, in this instance, it is necessary to create a new table in the CMS

¹⁰⁰ *Id.*

¹⁰¹ *Id.*

¹⁰² *Id.*

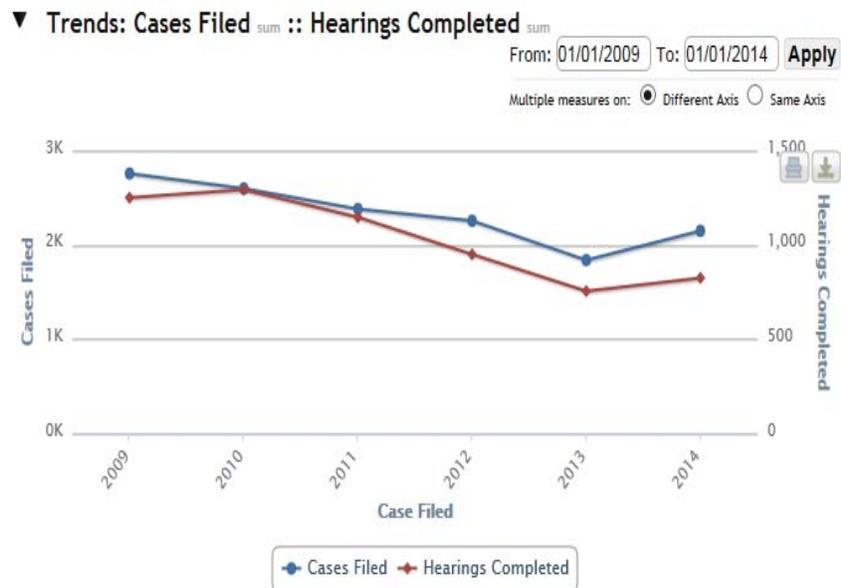
¹⁰³ One tip: An agency should ensure that the analytics module can show trends in both calendar and fiscal years.

¹⁰⁴ *Id.*

¹⁰⁵ *Id.*

that contains information from both the case information table and the hearings table.

The third table is for data purposes only and will be the domain for a new dataset that can show both measures on one chart in a more unified analytics view:

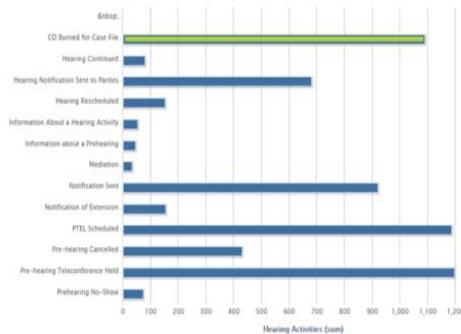


Additionally, a user will be able to drill down on dimensions that relate to both measures.

In implementing analytics, NAD initially underestimated the effort required to check data results that would emerge from analytics. Analytics creates datasets primarily from existing data tables and relationships. Based on this principle, there should be little variation in data counts or computations in the analytics view. But since the data domain is selective and analytics provides the ability to create separate measures with increased computations, data checking protocols are critical. The first method for agencies to validate analytics data is to ensure that the test data encompasses a domain of data for a given period. For example, in NAD, all data quality checking pertains to 2013 data. The second method is to compare analytics data with the standard reports in the existing CMS. In the testing protocol, the analyst team should run the most commonly used reports for the period and query the CMS with all available tools

to verify that the analytics provide consistent and accurate results that other reporting tools confirm in the CMS.

A new agency analytics initiative has the advantage of creating a data mining opportunity for its activities. For example, at NAD,



agency administrative staff employees routinely enter data, previously unanalyzed, in the CMS about a myriad of administrative activities in the first-level appeals process. It was envisioned that the new agency analytics would contain a new ‘Hearing Activities’ dataset.

After deploying this dataset, it was enlightening to engage one analytics view of over 7,000 of these administrative activities performed in 2014 (above). These activities included creating and mailing notices to appellants about the hearing schedule, postponements, extensions, and a multitude of other interactions that NAD has with a customer in the hearing phase of a case. Even more striking was how analytics reinforced that over 1,000 of these activities related solely to the burning of audio CDs of the hearing for the case file or to respond to customer requests.¹⁰⁶ This is the kind of discovery an agency can look forward to that may cause it to question staffing priorities or procedures. As an agency implements analytics for an existing CMS, it may be surprised to see what is in its data.

D. Dashboards and Beyond

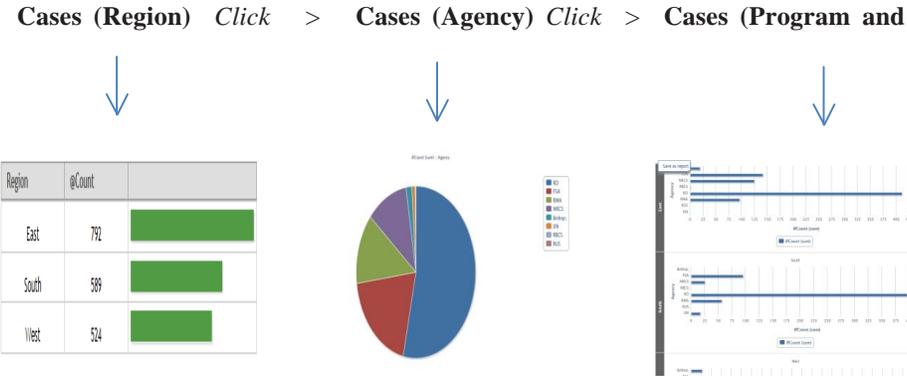
Most analytics permit CMS managers to create views that can be modified on the fly by pointing and clicking.¹⁰⁷ These interactive modifications are drill-downs of measures and dimensions. Further,

¹⁰⁶ NAD appellants have right to request a copy of the audio. Since many are rural participants, they need a copy they can play on a CD player.

¹⁰⁷ I have seen a couple of analytics modules call these views “Reports,” which can be confusing, because they are not the same as the traditional reporting modules in the CMS; nor are the datasets related.

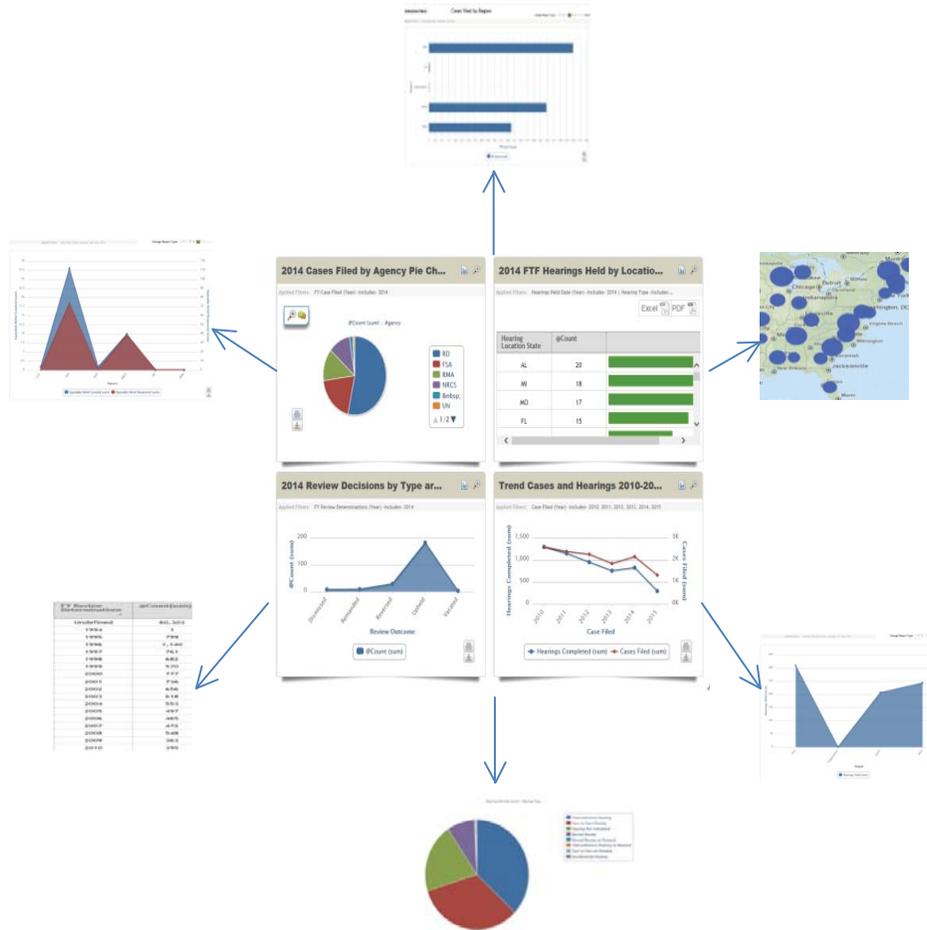
the display (chart, table, map, etc.) can also be changed with a point and click.

For instance, suppose a user wants to know the number of cases docketed in an agency in a year. With one or two clicks, the user can drill down in the data and change the dimension and the format of the display:



As the analytics trail depicts above, a user can engage each view separately and drill down for more dimensional data or further computations. Grouped together, these views become dashboards at the center of an agency data display. A dashboard should be designed as the starting point for a user to take a journey through the CMS, each stop only a click away from the next. Dashboards should be designed to facilitate that journey. The collection of analytic views ought to provide the overall ‘readings’ of agency health, while still providing a springboard for other searches and analysis too. For example, the NAD dashboard below (shown previously) displays metrics in four different areas of interest in the NAD process.¹⁰⁸ Each area, however, has potential for further investigation.

¹⁰⁸ See Nat’l Appeals Div. *supra* note 80,.



The arrows in the analytics trail above depict the first steps in user journeys leading to other dimensions, which are only one click away from one of the quadrants in the basic dashboard. The bottom left quadrant of the center dashboard above, for example, shows an area chart of the number and outcomes of second-level reviews issued by the agency in 2014. This leads to an analysis of the number of instances the agency provided equitable relief in 2014 for second-level reviews — and then from 2003 to 2014. The display above shows that all quadrants have similar opportunities for additional visual displays that show alternative dimensions from the dashboard. By the time an agency is crafting—and it really is an art—an analytics dashboard, it invokes all the principles of effective data display: density, context, and multi-variability. Through analytics

software, it is able to parse and activate the data, making it more accessible to employees in the organization.

VI. CONCLUSION

Analytics creates a “seeing is believing it” organizational mindset. While analytics have not been a large component of Legal Administration agency portfolios in the past, the availability and affordability of systems with analytics software will change expectations for agencies to collect, analyze, and display data about its activities. As agencies take on this mindset, however, it is best not to jump into analytics without having some foundation in place.¹⁰⁹ Agencies should start with some simple approaches to case management data and metrics and then make them more sophisticated over time. An analytics implementation does not have to reproduce or invigorate all past agency reports at first. Agencies should choose a couple of dimensions that cut across the agency or provide context for trend lines in many areas of activity. For example, an agency with analytics that parses data about its backlog can be a good starting point for much analysis, because other metrics, such as average days to issue a determination, number of adjudicators assigned to the agency, and input / output ratios all have a cause and effect relationship to a backlog. A dataset with backlog information can have many dimensions to it, providing for rich analytics exploration and data display.

Second, agencies should be prepared to devote more time and resources to data governance to ensure data quality, reduce errors, and monitor that what users see is what they want to see and that users believe what they see. Remember, analytics will evolve an organization into valuing the analytics data as an asset in itself, and future system modification will more likely address user interactions with the analytics module, rather than the availability of the data alone. Management analysts who previously ran reports, queried databases, and developed spreadsheets will transition to a more data immediate environment as they support these organizational needs. These employees will probably become the core of future data governance in the agency.

¹⁰⁹ FRANKS, *supra* note 55 at 263-65.

Finally, agency leaders with a CMS should look for opportunities to operationalize analytics, permitting analytics to play a bigger role in decision-making, whether it is assigning a case, investing money in outreach activities, or conducting training. It may be as simple as including a sparkline or an analytics display trail in an agency report occasionally. This is the mark of an agency aspiring to evolve into a 3.0 agency. Some of these applications may be difficult to envision initially, especially for Legal Administration agencies, but keeping open to that possibility is critical to the full implementation of an analytics initiative. For when agencies embrace analytics, which they should, they are embracing the inseparability between performing their activities and the visual display of information that portrays their performance. And as they become empowered in this analytics endeavor, they become the designers of visual displays and new metrics that communicate the quality of their services to the public and the complexity of their missions.