

**THE
STATISTICAL
CONSULTANT**



Section on Statistical Consulting
Kim Love, Editor

American Statistical Association
Spring 2016; Vol 32, No. 1

IN THIS ISSUE...

Chair's Column – Update #2 on Section Initiatives	3
Reaching Out to Data Scientists	3
Headlines and Announcements	5
Have an idea for a 2017 JSM Session?	5
Leadership positions in CNSL Executive Committee	5

Executive Committee of the ASA Section on Statistical Consulting

Chair	Chris Holloman cholloman@icct.com	
Chair-Elect	Charles Kincaid chuck.kincaid@experis.com	
Past Chair	Eric Vance ervance@vt.edu	
JSM Program Chair	Isabella Ghement isabella@ghement.ca	
JSM Program Chair-Elect	H. Dean Johnson dean_johnson@wsu.edu	
Secretary / Treasurer	Ralph Turner mac.turner@hotmail.com	
Publications Officer	Vaneeta Kaur Grover vkgrover@yahoo.com	
Webmaster	Ravi Makhija rm1535@georgetown.edu	
Newsletter Editor	Kim Love kim@krloveqcc.com	
Assistant Newsletter Editor	Christopher Swearingen chris@samumed.com	
At Large Members		
<u>2014–2016</u> Scott Berry scott@berryconsultants.com	<u>2015–2017</u> Edward Jones edward.jones@tamstatservices.com	<u>2016–2018</u> George Luta gl77@georgetown.edu
Council of Sections Representatives		
<u>2015–2017</u> MaryJo Smith drmaryjosmith@gmail.com	<u>2016–2018</u> Sarah Kalicin sarah.kalicin@intel.com	

Additional Officer Information can be found at www.amstat.org/sections/cnsl/officers.cfm

Chair's Column – Update #2 on Section Initiatives



Chris Holloman
Information Control Company

cholloman@icct.com

First, I want to say thank you to those who provided some feedback about the three initiatives that I developed for the section this year. Several members of the section contacted me to say that they liked these ideas and felt they should be pursued. As a reminder, the three initiatives I suggested for the section were as follows:

1. Develop a collection of online resources useful for private consultants.
2. Broaden the base of participants in discussions of statistical topics.
3. Improve the perception of statistical consulting's usefulness in the broader business community.

Unfortunately, despite the positive response about the value of the initiatives, no one provided ideas for implementing them, and there weren't any volunteers to lead the charge on any of them. Because I still believe they are important, I've pivoted to working directly with the Executive Committee Members-At-Large to make some progress. In addition, I believe that initiative 3 is too ambitious, so we will focus on one or both of the first two in the list. I'll keep the section updated through the eGroup as we make progress.

One new idea that has been implemented is the "official" discussion topic for the eGroup. I led the first one in March/April, and Chuck Kincaid (Chair-elect) kicked off a new discussion for July. If you have topics you'd like to suggest for discussion, please send them to me and let me know if you'd like to moderate the discussion or whether you'd prefer that someone from the Executive Committee conduct it.

Finally, I'm looking forward to seeing many of you at JSM next month. We have a great program that Isabella Ghement worked hard to pull together, and I think anyone who is able to go will get a lot out of it.

Reaching Out to Data Scientists



Kim Love
Newsletter Editor

kim@krloveqcc.com

As Chris mentioned in his Chair's Column, our Chair-elect Chuck Kincaid recently began a discussion within our eGroup regarding data science. To paraphrase his initial email, it remains unclear what a data scientist is, whether statisticians are data scientists or vice versa or neither, whether there is a place for traditional "statisticians" in the future of data analysis, etc. He also mentioned the ASA's plans to "facilitate further collaboration between statisticians and other data scientists", and asked members of the eGroup to comment with some of their strategies for facilitating collaboration between statisticians and data scientists.

Through the thoughtful and varied answers provided by the community, it became very clear that the definition of "data scientist" is absolutely not agreed upon. Some participants in the discussion clearly felt that statisticians are data scientists and the two terms are either interchangeable, or one is a subgroup of the other. For those who did feel there is a difference, some common implications and understandings were that data scientists tend to work with bigger data sets than general statisticians (think of the also somewhat ill-defined term "Big Data"); that data scientists have more conceptual knowledge of data storage and database structures; and/or that data scientists are focused on prediction rather than explanation. Data scientists may come from computer science and IT areas as well as the field of statistics. There is also a potential negative connotation that data scientists often do not understand probability or the deeper underpinnings of statistical methodology.

While there have been some negative responses to the concept of "data science", many who joined in the conversation did feel it was important to interact with those who consider themselves data scientists but not necessarily statisticians, and those who are not currently members of the ASA. There does seem to be some consensus that everyone who works with data applications in the modern world would benefit from a mix of skills, including an understanding of statistical principles, the ability to store and efficiently manipulate large and/or complicated data sets, and expertise with varied and advanced approaches to analysis, including computationally intensive techniques that many of us do not learn in our programs of study in statistics. Rather than expecting each individual to retain these skills, collaborative efforts appear to be a high priority.

It also was clear from the discussion that even among those who currently work with "data scientists" (by title or self-identification), the distinction is far from clear. This made it pretty difficult to answer the question that Chuck originally posed, which was regarding strategies for facilitating collaboration between statisticians and (other) data scientists. To learn more about the ASA's perspective on data science and the relationship of data science to statistics, you may be interested in the [ASA Statement on The Role of Statistics in Data Science](#), which Chuck included in the original post. One direct quote from this statement that perhaps lends some further context to the ASA's current efforts is reproduced here:

While there is not yet a consensus on what precisely constitutes data science, three professional communities, all within computer science and/or statistics, are emerging as foundational to data science: (i) Database Management enables transformation, conglomeration, and organization of data resources; (ii) Statistics and Machine Learning convert data into knowledge; and (iii) Distributed and Parallel Systems provide the computational infrastructure to carry out data analysis.

The ASA encourages statisticians to engage and collaborate with "researchers with expertise in data organization and in the flow and distribution of computation." The statement also says that "[s]tatistical education and training must continue to evolve—the next generation of statistical professionals ... must be more able to engage with database and distributed systems experts."

Given the ASA's view of the role of statisticians in data science, have you had the opportunity to collaborate with other data scientists who may not consider themselves statisticians? Please feel free to add to the conversation: <http://community.amstat.org/cnsl/home> (must be logged in to view and comment).

Headlines and Announcements

Have an idea for a 2017 JSM Session?

The Statistical Consulting Section is pleased to announce that it is offering a travel award to the top submitted topic-contributed proposal as determined by a judging committee. The recipient receives a plaque and \$750 prize. If you feel you have a topic that would have some appeal to the statistical consulting community, please email your proposal to dean_johnson@wsu.edu by January 11, 2017.

Note that online submission of invited sessions for JSM 2017 is already open! [Click here](#) to view the online submission website; you are also encouraged to email dean_johnson@wsu.edu prior to submitting your idea online to obtain feedback.

Leadership positions in CNSL Executive Committee

The Statistical Consulting Section is searching for individuals to fill several non-elected leadership positions on its executive committee. If you are interested in becoming more involved in the section, this may be a great opportunity for you! Please email cholloman@icct.com for more information about these positions.

The Statistical Consultant Contact Information

Kim Love

Editor

K. R. Love Quantitative Consulting

and Collaboration

337 S. Milledge Ave, Suite 208

Athens, GA 30605

☎ 706-850-2210

✉ kim@krloveqcc.com

Christopher Swearingen

Assistant Editor

Samumed, LLC

Sr. Director, Biostatistics

9381 Judicial Drive, Suite 160

San Diego, CA 92121

☎ 858-926-2900

✉ chris@samumed.com

Section on Statistical Consulting Microsite

<http://community.amstat.org/CNSL/Home>

You'll find all information relating to the Section on Statistical Consulting, including our charter, officer list, section activities, past issues of *The Statistical Consultant*, minutes of past meetings, and more.