PPAS - BEST PRACTICES

Presenter: Jean-Marc Fix, FSA, MAAA

September 23, 2020
SOA Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

• Do not discuss prices for services or products or anything else that might affect prices
• Do not discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
• Do not speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
• Do leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
• Do alert SOA staff and/or legal counsel to any concerning discussions
• Do consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone’s responsibility; however, please seek legal counsel if you have any questions or concerns.
Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.
Agenda

• CRISP-DM methodology
  • Business understanding
  • Data understanding
  • Data quality
• ASOPs: 23-Data & 56-Modeling
• Bias
• Reproducibility aka Real Documentation
CRoss InduStry Process for Data Mining

• Business understanding
• Data understanding
• Data preparation
• Modeling
• Evaluation
• Deployment

Source: www.sv-Europe.com/crisp-dm-methodology/
Business Understanding

• What is the business problem you are trying to solve?
  • Business objective
  • Business success criteria

• Where are you now?
  • Assess current situation
  • Costs and benefits to move to the new structure

• Only then assess the technical goals (e.g. desired accuracy, speed...)

Source: www.sv-europe/crisp-dm-methodology/
Data Understanding

• The data is key so overdocument
  • Data collection
  • Data description
  • Data exploration
  • Data quality

Source: www.sv-europe/crisp-dm-methodology/
Data Quality

- Emphasis on measures and acceptability threshold
- **Completeness**: % complete, relative to what
- **Uniqueness**: # real world/# in dataset
- **Timeliness**: time between occurrence and recording
- **Validity**: how much of the data is in the correct form/format
- **Accuracy**: data represents what it is supposed to be
- **Consistency**: with other data sets
- Implied: data is safe, legal and appropriately protected

Source: The Six Primary Dimensions for Data Quality Assessment DAMA UK Working Group
ASOP 23: Data

• Reliance
• Review of data for *reasonableness* and consistency
  • Definitions (of the data elements)
  • Identify questionable data values
  • Review of prior data (maybe less relevant in our context)
• Use of Data (as is, enhance, judgment, audit, inadequate)
• Professional judgment!

Source: ASOP 23 Data Quality Doc 185
ASOP 56: Modeling

• Started in 2010, 4 exposure drafts later, ASOP released 12/19
• Wording adjusted to encompass predictive models (as distinct from model office or pricing models)
• Appropriate for intended purpose
• Understand weaknesses and assumptions
  • Of data
  • Of technique
• Governance and control

Source: ASOP 56 Modeling Doc 195
Bias

• Watch out for bias leading to unfair discrimination
• Bias can be due to not understanding correlations
• Harder to detect in machine learning approaches
• One way to guard against bias is to verify that explicitly adding the variable does not change your model results
• Generally due to problems in training set
Reproducibility

Code First
Think Later
Reproducibility

Code First
Think Later
Reproducibility

Data + Environment + Code = Results
Data

• How did you get your data?
  - Where (what URL/system)?
  - How?
  - When?
• Keep one clean copy of your starting data as downloaded
• Beyond data cleanup, add fields rather than replace
• Keep an eye on size
• Where is it stored?
• How will you get it next time?
Environment

• R and packages change -> sessionInfo()
• Technological platform
  - Communication
  - Storage
• Outside data source
  - Links
  - Moved
  - Changed
Code

• Do NOT clean by hand
• Version
  - Use GITHUB
• Code clearly
  - # use comments and indentations
• Use R-Markdown to code and produce your output
Questions?

RESULTS RESULTS RESULTS
RESULTS RESULTS RESULTS
RESULTS RESULTS RESULTS