

The SAS Analytics Shootout Annual Student Competition

Sponsored by SAS and
The Institute for Health & Business Insight at
Central Michigan University

Overview

2



- What is the SAS Analytics Shootout Competition?
- Tasks and Requirements
- Integration to the classroom/Professional Development
- Rules and Timeline



IHBI at Central Michigan University

3

- The Institute for Health and Business Insight at Central Michigan University, founded in 2001 as a center of excellence specializing in data mining and predictive analytics.
- Purpose
 - ▣ Promote the application of predictive analytics to solve social, economic and business challenges
 - ▣ Build bridges between the public and private sector
 - ▣ Train the next generation of data scientists
- Team
 - ▣ PhD and MS in physics, statistics, economics, computer science, geography, etc.

What is the Analytics Shootout?

4

- For the past 9 years, The Institute for Health & Business Insight has partnered with SAS to create and judge the SAS Analytics Shootout student competition
- Annual competition that challenges students to tackle a complex problem using SAS software and advanced analytics

What is the Analytics Shootout?

5

- Student teams are given a collection of datasets and a “Problem Statement” built and put together by IHBI
- Use of any variety of analytical methods and approaches they see fit to address the problem
- Produce and submit a comprehensive report and SAS flows/code to be considered for grading

What is the Analytics Shootout?

6

- A team of judges review and grade all submitted reports and SAS code
 - ▣ Judging team is collection of data scientists, analysts, and professors
 - ▣ A blind judging process; all identifying information from the reports and code are removed before the judges begin their review process
- After several rounds of review, the judges decide on a final ranking the Top 3 teams and any number of honorable mentions
- Winning teams are given complementary admission to the annual SAS Analytics Conference where they are awarded for their work and give a presentation on their project

What is the Analytics Shootout?

7

- Recognition for winning and honorable mention teams on both SAS and IHBI press releases and social media
- Financial donations to the schools of the Top 3 teams



What is the Analytics Shootout?

8

- 2016 will be the 10th annual competition
- Open to teams from universities across the US and Canada
 - ▣ 2015 Shootout competition had 96 teams registered from 36 different universities

Past Problems

9

Complex problems dealing with a wide range of topics:

2008

Developing contingency plans for forecasting flight delays and cancellations to increase customer loyalty

2009

Propose processing facility locations based on predicted 'energy grass' (new biofuel source) crop acreage and yields

2010

Diabetes early detection: costs and benefits

2011

Impact of weather and storms on hospitalizations due to infectious diseases

2012

Developing improved flu vaccine plans and predicting the cost and the benefit in terms of lives saved, reduced hospitalizations and reduced infections.

2013

Predicting costs and benefits of disease prevention programs

2014

Understanding and reducing crime in urban areas

2015

Predicting city-wide wind and solar energy production and consumption

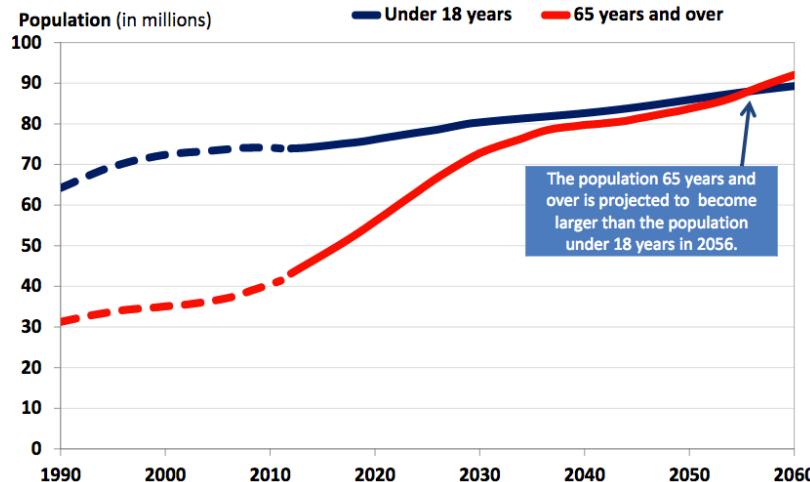
2016 Shootout Problem

“Impacts of an Aging American Population”



Impacts of an Aging American Population

- As the Baby Boomer generation ages and the median age in the US continues to rise, the effects will be widespread
- By 2056, the number of individuals in the US 65 years and older is projected to exceed those under 18



Impacts of an Aging American Population



The future availability of health care and advancements in technology may contribute to an overall reduction of disease incidences and lead to individuals living longer

Impacts of an Aging American Population

With both an increase in the overall population and individuals living longer, the labor force landscape may experience significant changes



Impacts of an Aging American Population

Changes in the labor force landscape and an increasing number of individuals retiring will present challenges to the Social Security program and the US tax system



Impacts of an Aging American Population

- Teams were challenged with understanding and modeling future impacts on the areas of:
 - Disease Diagnoses
 - Labor Force landscape
 - Impacts on Taxes and Social Security.

- Teams examined scenarios based on predicting future trends for:
 - Reducing costs in health care technology
 - The possibility of a Social Security Crisis



Integration to the Classroom/Professional Development

18

- The Shootout problem is a great teaching tool
 - ▣ A realistic and relevant advanced analytics problem
 - ▣ Use of real data and it's complexities
 - ▣ Multiple approaches to addressing the problem with no clear-cut, “correct” answers
 - ▣ A start-to-finish analytics problem

Integration to the Classroom/Professional Development

19

- A relevant advanced analytics problem
- Diverse topics
 - Topics are wide-ranging and address problems in a variety of industries
- Real-world and current issues
 - Problems are chosen to be realistic and currently relevant

Integration to the Classroom/Professional Development

20

- Use of real data and its complexities
 - ▣ Use primarily real data (with some simulated data, occasionally) that come from a variety of sources
 - Previous datasets used in the Shootout have been both collected from public sources, supplied by businesses/institutions, or built by IHBI
 - ▣ Students have to examine, clean, and harmonize their data carefully in order to gain useful insights
 - The strategy for cleaning the data and performing thorough EDA is a large part of the Shootout problem

Integration to the Classroom/Professional Development

21

- Multiple approaches to addressing the problem with no clear-cut, “correct” answer
- ▣ Several types of modeling approaches may and can be used to come up with a good solution to the problem
 - Techniques ranging from regression to machine learning to the effective use of data visualization can be incorporated throughout the process
- ▣ Successful teams justify their analytical approaches with thoughtful, well-formed explanations
 - Strong emphasis on explaining the analytical choices used to address the problem

Integration to the Classroom/Professional Development

22

- A start-to-finish problem
- Students work through every step of an analytics problem solving process
 - Data cleaning and harmonization, EDA, modeling choices, training and scoring, writing a report detailing the entire process
- Real-world problems that can be expanded upon further in continuing research or in a professional setting

Basic Rules

- Open to students in the US or Canada currently enrolled as a full-time student as an accredited academic institution
- Teams may only use the data provided in the Shootout package
 - ▣ Additional research about the topic is encouraged, but additional datasets can not be used
- Must use SAS software during the process

Timeline



- Set to run concurrently with a typical university semester schedule
- Early January – Registration open and data package available for download
- Early June – Submission deadline

So, Why Participate in the Shootout?

25

- Gives students the opportunity to solve **real-world** and relevant advanced analytic problems
- Great teaching tool for professors, great professional development experience for students
- No risk to register and attempt the problem...FREE



If you would like more information, or if you would like to provide data for use in a Shootout problem, please contact:

Joseph Pomerville
Research Analyst / Advanced Analytics
The Institute for Health & Business Insight
at Central Michigan University
pomer1jb@cmich.edu

Or contact Jenna Green, SAS Student Programs Manager : jenna.green@sas.com