

# AN OVERVIEW OF THE ANNUAL SAS 'ANALYTICS SHOOTOUT' STUDENT COMPETITION

# Overview



- What is the SAS Analytics Shootout Competition?
- Types of work involved
- Integration to the classroom/Professional Development
- Tools and Requirements
- Timeline

# What is the Analytics Shootout?

- Student teams are given a collection of datasets and a business problem to solve using SAS software and advanced analytical methods
- Teams produce a comprehensive report and SAS flows/code to be considered for grading

# What is the Analytics Shootout?

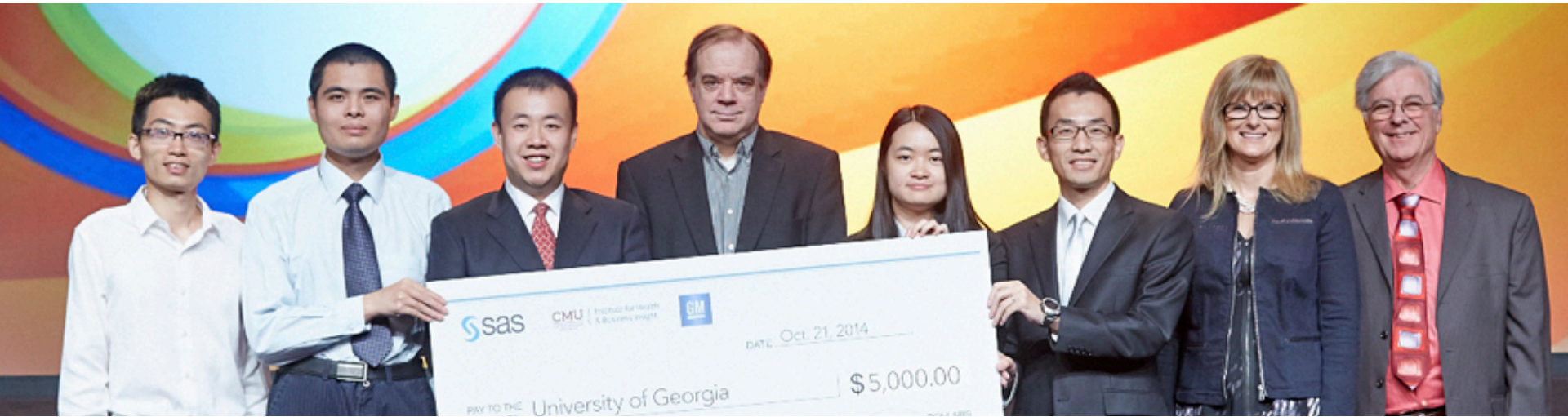
- 2016 will be the 10<sup>th</sup> annual competition
- Teams from universities across the US and Canada
  - 2015 Shootout competition had 96 teams registered from 36 different universities

# What is the Analytics Shootout?

- Winning teams are given complementary admission to the annual SAS Analytics Conference where they are recognized and give a presentation on their project
- Recognition for honorable mentions on both SAS and IHBI press releases and social media

# What is the Analytics Shootout?

- Financial donations to the schools of the Top 3 teams



# What is the Analytics Shootout?

- The Shootout gives students the opportunity to solve real-world advanced analytic problems
- A great way for students to put their modeling skills to the test, gain recognition for their work, and add valuable experience to their resumes

# Types of work involved

## 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



# Types of work involved

- Datasets are provided but must be cleaned and harmonized to use
- Modeling approaches are wide-ranging and multiple techniques can be incorporated
- An academically written, professional report to present and summarize the approach and solutions to the problem

# Example Set of Data

## Shootout 2015 Data Dictionary

### “Power City” – A study in Energy Consumption and Renewable Energy Production

This document lists the datasets provided to accompany the 2015 SAS Shootout Competition. Each dataset is described for content, the dataset name is provided, and the data elements are described.

#### Provided Data:

1. Sector\_Use\_Matrix.xlsx
2. Calendar\_Days\_Consumption.sas7bdat
3. Calendar\_Days\_Scenario.sas7bdat
4. Car\_Charging.sas7bdat
5. Powercity\_Consumption.sas7bdat
6. PowerCity\_Population.sas7bdat
7. PowerCity\_SolarAngle\_Consumption.sas7bdat
8. PowerCity\_Weather\_Consumption.sas7bdat
9. PowerCity\_Weather\_Scenario.sas7bdat
10. SolarArray\_Production.sas7bdat
11. SolarArray\_Weather.sas7bdat
12. SolarArray\_SolarAngle.sas7bdat
13. WindFarm\_Production.sas7bdat
14. WindFarm\_WindSpeed.sas7bdat

# Example Set of Data

7. **PowerCity\_SolarAngle\_Consumption.sas7bdat**: Solar Elevation during each hour in Power City during the consumption year. Contains 6 Columns and 8,783 Rows.

- **City**: Name of the city
- **Year**: Year of consumption data labeled as "Consumption"
- **Month**: Numeric month (1 = January, 2 = February, etc.)
- **Day**: Numeric day (1 = 1<sup>st</sup> of the month, 2 = 2<sup>nd</sup> of the month, etc.)
- **Hour**: Numeric hour (1-24)
- **Solar\_Elevation**: The **solar elevation (angle)** is the altitude of the sun, measured as the angle between the horizon and the center of the sun's disc.

8. **PowerCity\_Weather\_Consumption.sas7bdat**: Hourly weather variables over the course of a year. Contains 12 Columns and 8,760 Rows.

- **City**: Location of weather variables
- **Year**: Year of weather variables
- **Month**: Numeric month (1 = January, 2 = February, etc.)
- **Day**: Numeric day (1 = 1<sup>st</sup> of the month, 2 = 2<sup>nd</sup> of the month, etc.)
- **Hour**: Numeric hour (1-24)
- **Cloud\_Cover\_Fraction**: Amount of cloud cover (decimal from 0 being no clouds to 1 being fully cloudy)
- **Dew\_Point**: Temperature of the dew point in degrees Celsius
- **Humidity\_Fraction**: Fraction of humidity in the air (0 to 1)
- **Precipitable\_Water**: The total precipitable water contained in a column of unit cross section extending from the earth's surface to the top of the atmosphere in millimeters
- **Pressure**: Pressure reading in millibars

# Integration to the classroom/Professional Development

- The Shootout problem is a great teaching tool
  - ▣ A real-world and relevant advanced analytics problem
  - ▣ Use of real data and its complexities
  - ▣ Multiple approaches to addressing the problem with no clear-cut, “correct” answers
  - ▣ A start-to-finish analytics problem
  - ▣ Publication potential by extending the concepts learned into future studies

# Integration to the classroom/Professional Development

- A real-world and relevant advanced analytics problem
  - ▣ Diverse topics
  - ▣ Real-world issues
  - ▣ Sophisticated modeling approaches and explanations to justify choices

# Integration to the classroom/Professional Development

- Use of real data and its complexities
  - Use primarily real data (with some simulated data) that come from a variety of sources
  - Data formats vary and must be harmonized
  - Students have to carefully examine and clean their data in order to achieve the full potential of the data provided

# Integration to the classroom/Professional Development

- 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 solution
  - Successful teams justify their analytical approaches with well-thought explanations

# Integration to the classroom/Professional Development

- A start-to-finish problem
  - Students work through every step of an analytics problem solving process
  - Extensive research is used to fully understand the problem and suggest potential improvements for future studies



# Integration to the classroom/Professional Development

- Potential to extend the concepts learned to future studies
  - Topics are common business problems
  - Solutions and approaches can be expanded to include additional data and even more sophisticated modeling approaches

# Tools and requirements

- 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
- 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

# Main Takeaways

- Realistic and relevant challenges for students
- 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 donate data for use in a Shootout problem, please contact:

## Joseph Pomerville

Research Analyst / Advanced Analytics

The Institute for Health and Business Insight at  
Central Michigan University

[pomer1jb@cmich.edu](mailto:pomer1jb@cmich.edu)