STUDENT DATA HACKATHON 2017

Inspiring Tomorrow's Data Heroes

Student Data Heroes 2017 Hackathon Information

Participant Groups and Competition Levels
Hack-a-Thon participants can choose to participate at whatever level of difficulty makes sense to them, as an individual or in a group of 4 or less. All participants should begin at the level of descriptive analysis and then progress to the highest level of complexity that they can complete

Judging and Scoring Information
Final presentations will be judged by multiple judges (subject matter experts and industry data science experts) to determine if their mathematical approach is sound and to determine if their results are relevant and make sense given the topic area. Students' final scores will reflect an average of the two scores. Each Hackathon judge will be grading using the same rubric with 20 possible points.

The highest scoring submissions will be given the opportunity to receive an internship or a cash prize. Internships are reserved as prizes for individual competitors. The internships will be awarded by Biocom's industry partners and will be negotiated between the student and the organization providing the internship. Both paid and unpaid internships are included as prizes and best attempts at successful matches will be made. Winning high school contestants receive a cash prize regardless of whether competing as an individual or part of a team.

Data Release Dates and Submission Deadlines
All students will be given access to the data during the kickoff event around 11:00am on Saturday, September 30th and will be due by 11:59pm on Saturday, October 7th. The winners will be notified on October 11th and announced on October 12th. Winning presentations will be required to present their findings at the Hackathon Awards Ceremony on Oct. 12th from 6 - 8pm.

Note, students unable to attend the September 30th kickoff can expect to receive access to the data and competition rules by 1pm that day.

Final Submission Guidelines
All individual competitors/student teams competing will be required to write up their findings in a short report highlighting the most important elements of their analysis. Reports should be max 5 pages plus a title page.

Reports should include:

- Title page (see sample for specifications)
- Technical methodology section where tools used to complete analysis are identified and, if applicable, to merge data sets
- Analytical techniques used
- Any assumptions made in the analysis
- Identify any additional data that students merged into the data set and any transformations made to the data. If additional data was merged into the provided data set, please upload the final data set used to complete the analysis to your Dropbox folder.

Each individual competitor/student team will also be required to produce a high-level PowerPoint type presentation that highlights the findings that might be most important to decision makers in the educational arena. Focus on the relevance of the findings for the school district administrators, your audience, and include any possible recommendations based on the findings. It is not necessary to go into high level of technical detail for the PowerPoint.

To be eligible to compete, students must upload the following to their competition Dropbox folder by 11:59pm on Saturday, October 7th:

- Report
- PowerPoint
- Final (merged) data set used (if different than what was supplied by the completion)
- Rubric

For each file uploaded the following file name configuration should be used:
Individual competitors: IndHack2017_Q*#*_lastname_firstname_ *type of file*
Group competitors: GroupHack2017_Q*#_teamname_type of file*

 # = enter the number of the highest level question answered (1, 2 or 3)

For example, Ann Smith competes as an individual and answers questions 1 and 2. Ann's rubric file name should read: IndHack2017_Q2_smith_ann_rubric

Teams should only upload these solution components to one folder of the selected team member. All other team members should upload a copy of the title page only to their Dropbox folder.

All winners will be sharing a few key slides from their PowerPoint and should be prepared to talk about what were the most fun/rewarding and the most challenging parts of the competition at the awards ceremony on Oct 12th. **Individual winners and at least one member of each winning team must be present at the awards ceremony.**

The grading rubric is included in this packet.

Sample Title Page

Competition Category: *(enter your category choice, from below, in which you wish to compete)*

Choices

- **High School Competitor** *(in the high school category, individuals and teams are judged together)*
- **Individual College Student competing for cash prize** *(this individual category will compete amongst college teams)*
- **Individual College Student competing for internship**
- **Team of College Students**

Your Name: _____

School Name: _____

Current Level in School: _____

School Program/Department, if applicable:_____

Major, if applicable:_____

Gender (optional): _____

For Teams Only:

Team Name: _____

Name of Team Member whose Dropbox folder contains the team's submission:_____

Student Data Heroes 2017 Hackathon Competition Questions

I am a Superintendent. I am in charge of the success of our high schools. I am trying to find out how to improve the AP scores at the high schools in our district. What factors might impact the number of passing AP scores we have at our school?

Descriptive Questions (answer both)

- What are the pass/fail rates for each high school in San Diego County?
- Are there any trends by sub-group of student/type of school/school location?

Correlational Questions (answer both)

- Is there a relationship between elementary ELA test scores or Math Test Scores and AP pass/fail rates for those schools that feed into each high school?
- Are there any other relationships between student/school characteristics and AP pass/fail rates?

Predictive Question

- Predict the AP pass/fail rate for each high school in my district based on elementary school characteristics.
  ** Please note, this question will require participants to merge the ELA and Math Scale Scores into the data set. This data will be in Dropbox in a separate file. The easiest way will be to join the files using CAISC code as the variable they are joined on

Hack-A-Thon Rubric Scoring Sheet

Team Name: (to be completed by student) **Names of Individuals in the Team:** **(to be completed by student)**	**Total Score Out of 20 Possible Points:**

Mathematical Modeling –14 Total Possible Points	**Scores and Comments**
Mathematical Accuracy • Mathematically accurate (2pt.s) • Uses an appropriate mathematical technique to answer the question (4 pt.) • Findings report appropriate statistics and coefficients for the method being used (2 pt.) • Findings are represented appropriately for the method being used, i.e. a correlational question is not presented as being able to account for causation (2pt) • Listed transformations they did on the data (2pts) • Listed assumptions they made during their analysis (2pts)	

Solution Relevance – Total of 6 points	**Scores and Comments**
Grammatical Accuracy and Clarity - up to 2 points • Presentation is free from spelling errors, typos and is grammatically correct (1 pt.) • Conveys a clear message (1 pt.)	
Understanding of Data Applications – up to 2 points • Findings demonstrate an accurate understanding of how the data will be applied in the real world (1 pt.) • Presentation's findings are logical (1 pt.)	
Communicates Data in an Interesting Way – 2 points • Presentation uses data to tell a story about the findings that's meaningful • Presentation presents information that can help shed light on the social issues being examined	

For Questions:

Andrea Yoder Clark, PhD
Dir. of Research & Evaluation
Inst. Entrepreneurship in Education
University of San Diego
ayoderclark@sandiego.edu

Liisa Bozinovic
Executive Director
Biocom Institute
lbozinovic@biocom.org

STUDENT DATA HACKATHON 2017

Inspiring Tomorrow's Data Heroes

Partners

We wish to thank

Rady School of Management
UNIVERSITY OF CALIFORNIA SAN DIEGO

for its support!