



Two Sigma Connect – Data Analysis and Beaker Contest FAQs

About Us

We are Two Sigma. We find patterns, meaning, and relationships in all the world's data. And we use these insights to make intelligent decisions in the global financial markets. All of which helps us achieve our ultimate goal: to deliver value for our clients.

Visit us at www.twosigma.com/careers to learn more and apply.

Competition Schedule

Thursday, April 16th: The competition kicks off at 6pm. Activities for the night include networking sessions, dinner, and an overview of tournament rules before prepping laptops and returning to the hotel at 9pm.

Friday, April 17th: The competition starts at 9am and will run for 24 hours. Food will be provided throughout the day and Two Sigma mentors will be on hand to give advice and guidance.

Saturday, April 18th: After closing the competition period at 9am there will be a Project Expo for teams to demo their submissions. Lunch, networking, and the award ceremony will follow, wrapping up at 2pm.

Competition Structure

Will this be an individual or a team competition?

This is a team competition; you will be randomly assigned to a 3-person group by an impartial algorithm.

Will I be able to work solo?

Unfortunately, you will not be able to participate solo. At Two Sigma, we constantly collaborate across multiple departments to create products and we hope the team environment will provide you with a glimpse of the nature of our work.

What type of data will I be using? What kind of data will I have access to?

New York City's transportation network provides vast amounts of data to work with, and we will be providing data sets related to it on Friday morning. You'll have access to data such as taxi trip data, subway entrances and exits and bike share rides. Please feel free to use any publicly available datasets that are relevant to produce your insights

You will be provided with a few datasets from NYC Open Data on Friday morning when the competition begins. Please feel free to use relevant data sets as you see fit to produce your insights.



Technical Requirements

Will I need my own laptop?

Yes, you will need to bring your own laptop with any necessary programs already installed. If you would like to use Beaker, please [click here](#) to install it in advance. The Beaker Development Team will be on site during the tournament to help out should you face any issues.

What languages would I need to compete? What kind of libraries am I allowed to reference?

You are free to use any programming language of your choice for analysis and visualization. Same goes for libraries. If you would like to use Beaker, we support a variety of commonly used data science languages from Python to R to JavaScript and many others. For a full list of supported languages, please visit <http://beakernotebook.com/>

Do I need to use Beaker for the competition?

No, using Beaker is not required (although recommended!). It is our open source notebook-style development environment, and we are excited to see how you might use it to conduct sophisticated data analysis and create innovative visualizations. Keep in mind as well that one of the three award categories is for the Best Beaker Application.

Travel and Accommodations

Is Two Sigma organizing travel and accommodations?

Yes.

In appreciation of your participation, Two Sigma will cover the cost of your trip if you choose to go through our travel agency and stay in our reserved hotel (Thu 4/16 arrival & Sat 4/18 departure).

Please note: if you elect to make your own arrangements, you will be responsible for them individually. Any costs incurred for independent travel and/or accommodations will not be included or reimbursed.

Where will I be staying?

We will have rooms reserved at the [SoHo Grand Hotel](#). All participants will have a room (to sleep, relax, recharge), regardless of if you had to travel to get here or live locally.

Will my meals be covered?

We will provide meals and beverages throughout the competition.

Subject: Invitation to Two Sigma Connect!
Attachments: Two Sigma Connect_Data and Beaker Contest FAQs.pdf

Hello !

You're invited to participate in Two Sigma's data analysis competition – Two Sigma Connect – based on your strong quantitative and programming background.

The event will take place Thursday, April 16, 2015 to Saturday, April 18, 2015 at our headquarters in SoHo, NYC. We hope that you'll be able to join us for what is sure to be a fun and intellectually challenging event. See below for more information.

If you're interested in competing and working with some of the top students in your field, please reply to this email by noon on Monday, March 23, 2015. We hope to see you soon!

Sincerely,
Two Sigma Campus Recruiting Team

Two Sigma Connect

Demonstrate your creative data analysis and visualization skills at Two Sigma Connect! Collaborate with fellow students to explore data's possibilities and find meaningful patterns and powerful predictions. *Show us your skills.*

Throughout the 2-day competition, our modelers and engineers will be present to mentor and discuss the power of data. You'll also have the opportunity to meet with some of our most innovative leaders in the field of big data and data visualization.

Competition Details

Date & Time: Thursday, April 16, 2015 at 6:00pm – Saturday, April 18, 2015 at 2:00pm

Location: 101 Avenue of the Americas, New York, NY 10032

Prizes include: Best Data Interpretation, Best Visualization, and Best Beaker Application (Beaker is our polyglot web-based IDE allowing you to work in many languages simultaneously. It's targeted at research, data analysis, and visualization, but it's open source and you can use it to do almost anything.)

**Please see the attached FAQ for further information.*

Getting Started

Tips to help you stay on track!

Team #: _____

Milestone	Description	Estimated Time	Complete?
Ideation and idea selection	<p>Get a plan and coherent sense of direction that is exciting for everyone on the team. Look briefly at the schema for the datasets while going through this process to get an idea of what's available.</p> <p>Team Check-ins: be ready for a check-in at 10:30am when each team will share a 1 min summary with every one of the idea they have chosen.</p>	9-10:30am**	[]
Ingestion and relevant dataset profiling	Get your data into a workable format, identify outliers and decide on outlier handling,	10:30-noon	[]
Hypothesis testing / idea verification	Prove to yourselves that your original idea is interesting. If it's not, try to pivot to another idea that can also utilize the datasets you've already ingested and profiled.	noon-1pm	[]
MVP/Presentation skeleton/"ghost deck"	<p>Figure out the structure of what you want to present--it could be a notebook you walk through, a demo of some web app, or a report/presentation. Sketch out the structure of what you want and what questions you need to be able to answer to construct your final presentation/product.</p> <p>1:30pm Mentor Check-in: you should have a sketch of all the work needed to complete your project, along with what you think will take the most time.</p>	1pm-1:30pm**	[]
Completed Presentation (all the data filled in)	<p>Answer all of the questions identified above, and fill them in in whatever skeleton notebook/app/presentation you were using. This is the fun part, so it should be the bulk of the work, and how you work on this will be very dependent on your idea and imagined finished product.</p> <p>5:30pm Mentor Check-in: make sure you are steadily checking off questions you need to answer during this whole time. By 5:30pm you should have half of the answers (data/charts/processes) you need filled in.</p>	1:30-9pm (5:30pm**)	[]
Polished Presentation	Keep exploring, iterating, improving, and polishing if you have more time!	9pm-midnight	[]

**denotes a check in. Check-ins are:

- idea check in: 10:30-10:45am
- idea verification and "ghost deck" check in: 1:30-1:45pm
- progress check-in: 5:30-5:45pm

Check-ins can take less than 15 minutes, but should not take more.

