CITADEL

Summer Invitational Datathon 2024

Competition Schedule and Guidelines

Welcome to Summer Invitational Datathon

Today's Webinar

Introduction

Competition Schedule and Instructions

- Review of competition rules and schedule

Problem Statement

- Introduction of the problem statement and datasets

Tips for Success

- Best practices for working virtually and building a successful report

CITADEL

Competition Schedule
Summer Invitational Datathon

Timeline	Activity
Monday, July 29th 6pm ET	**Introductory Webinar** ● Introduce Problem Statement and Competition Schedule ● Participants will be able to ask questions at the end
Tuesday, July 30th 10am ET	**Competition Begins** ● Participants will receive the datasets for the competition
Sunday, August 4th 5pm ET	**Submissions Due** ● Participants will submit their work by this time
Monday, August 5th 6pm ET	**Final Day** ● Citadel Info Session, Q&A, Awards Ceremony

CITADEL

Visit our careers sites @
citadel.com/careers
citadelsecurities.com/careers

Apply for Opportunities

→ **https://www.citadel.com/careers/open-positions/positions-for-students/**
→ **https://www.citadelsecurities.com/careers/open-positions/**
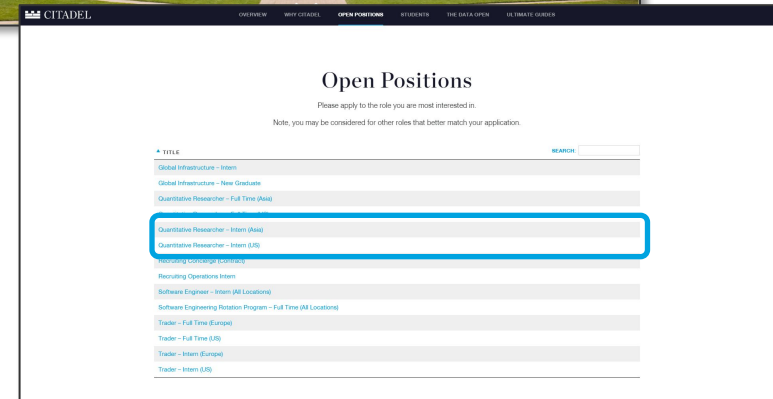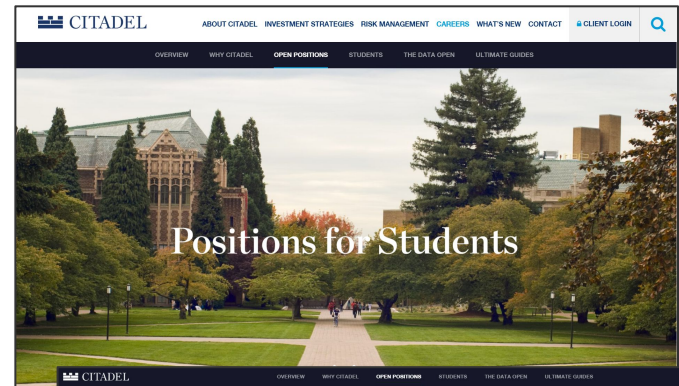
1 ➡ Careers

2 ➡ Open Positions

3 ➡ Quantitative Research

Event Follow-Up

- Post event, you'll be contacted by a Campus Recruiter
- If you're ready to be considered for positions at that time, they will help you get started
- If you're not ready, they will act as a resource
- Keep in touch!

Competition Schedule
Last Day of the Event

The week-long Datathon will culminate on Monday, August 5th with networking events followed by the Awards Ceremony (all in ET)

Time	Agenda	Notes
6:00pm - 6:15pm	Citadel Careers Info Session	Online info session and Q&A on career opportunities
6:15 - 7:00pm	Q&A with Citadel	Q&A session to learn more about specific teams and roles at Citadel and Citadel Securities
7:00 - 7:30pm	Awards Ceremony Webinar	Zoom meeting with gallery view of all participants, and awards presented by C1 Team

CITADEL

Problem Statement
Summary Invitational Datathon

There are approximately 750,000 restaurants in the United States representing myriad countries and cultures. Despite this, American cuisine remains closely linked to indulgent foods like oversized hamburgers and sugary desserts. This connection is deeply embedded in the economy, supporting many jobs from food service to agriculture. However, the health impact is concerning, as evidence continues to mount against sugar-laden, processed products as being bad for our overall health. In 2022, 22 states had an adult obesity rate at or above 35%. While healthy eating slowly makes inroads, processed foods don't appear to be going away anytime soon, nor the economy that supports it. With all of the data available, is there a way to better understand the true costs associated with this lifestyle?

Your goal is to analyze the provided datasets, potentially in combination with supplementary datasets, in order to better understand the footprint of processed food in the United States.

You will be provided several pre-cleaned datasets, including:

- Datasets on nutrition physical activity and obesity data.

- Datasets on processed food stocks and commodities (HLOC).

- Datasets on meat statistics and the American Community Survey (ACS) about various social and economic characteristics.

Three sample questions can be found in your packet, but you are welcome to design your own. You are also welcome to utilize relevant external datasets (< 2GB) with proper citation.

CITADEL

Problem Statement
Summary Invitational Datathon

There are approximately 750,000 restaurants in the United States representing myriad countries and cultures. Despite this, American cuisine remains closely linked to indulgent foods like oversized hamburgers and sugary desserts. This connection is deeply embedded in the economy, supporting many jobs from food service to agriculture. However, the health impact is concerning, as evidence continues to mount against sugar-laden, processed products as being bad for our overall health. In 2022, 22 states had an adult obesity rate at or above 35%. While healthy eating slowly makes inroads, processed foods don't appear to be going away anytime soon, nor the economy that supports it. With all of the data available, is there a way to better understand the true costs associated with this lifestyle?
Your goal is to analyze the provided datasets, potentially in combination with supplementary datasets, in order to better understand the footprint of processed food in the United States.

You will be provided several pre-cleaned datasets, including:

- Datasets on nutrition physical activity and obesity data.

- Datasets on processed food stocks and commodities (HLOC).

- Datasets on meat statistics and the American Community Survey (ACS) about various social and economic characteristics.

Three sample questions can be found in your packet, but you are welcome to design your own. You are also welcome to utilize relevant external datasets (< 2GB) with proper citation.

CITADEL

Problem Statement
Summer Invitational Datathon

There are approximately 750,000 restaurants in the United States representing myriad countries and cultures. Despite this, American cuisine remains closely linked to indulgent foods like oversized hamburgers and sugary desserts. This connection is deeply embedded in the economy, supporting many jobs from food service to agriculture. However, the health impact is concerning, as evidence continues to mount against sugar-laden, processed products as being bad for our overall health. In 2022, 22 states had an adult obesity rate at or above 35%. While healthy eating slowly makes inroads, processed foods don't appear to be going away anytime soon, nor the economy that supports it. With all of the data available, is there a way to better understand the true costs associated with this lifestyle?

Your goal is to analyze the provided datasets, potentially in combination with supplementary datasets, in order to better understand the footprint of processed food in the United States.

You will be provided several pre-cleaned datasets, including:

- Datasets on nutrition physical activity and obesity data.

- Datasets on processed food stocks and commodities (HLOC).

- Datasets on meat statistics and the American Community Survey (ACS) about various social and economic characteristics.

Three sample questions can be found in your packet, but you are welcome to design your own. You are also welcome to utilize relevant external datasets (< 2GB) with proper citation.

CITADEL

Submission Content
Summer Invitational Datathon

Submissions have two components:

1. **Written report**, with two sections

    a. Non-technical executive summary

    b. Technical exposition

2. **Code** - all relevant code that generated your results

CITADEL

Submission Instructions
Summary Invitational Datathon

As a reminder, submissions must be submitted by 5PM ET on Sunday, August 4th. Make sure to send your submissions in an email to datathonsubmissions@correlation-one.com.

When you are ready to submit, please ensure that your submissions conform to the guidelines below:

- Your report should be named "Team_[n]_report", where [n] is your team number, making sure it is in a universally accessible format (.pdf, .docx, .ppt, .HTML).
- Place all code into a folder and zip the file, named in the format "Team_[n]", where [n] is your team number. Make sure the zip format is universal.
- The email subject line should be in the following format: "Team [n] Submission - Summer Invitational Datathon 2024" where [n] is your team number.
- **Only <u>one of your team members</u> should send us the email** (feel free to CC your teammates). You can resend the email before 5PM on Sunday if you would like to make any changes - we will look at your latest submission before the deadline.
- **Please ensure to give yourselves extra time to send your submissions, especially if the size of your files is large.**

CITADEL

Competition Guidelines
Tips and Suggestions for Participants

Here are the guidelines to ensure a smooth completion and delivery of your submission:

1. ***Schedule your work time and plan ahead*** - you might want to spend the first couple of days exploring the data, reading up on the topic and testing hypotheses about the data

2. ***Commit to an approach*** – after the exploratory stage is completed, you might want to pick a narrower topic that can set you apart

3. ***Start working on your report early*** – the report you submit will need to stand on its own, so we strongly recommend starting on it early and working on it as you develop your analysis

CITADEL