CDP Climate Hackathon

July 20-27, 2020

CDP
DISCLOSURE INSIGHT ACTION

Agenda

- CDP Overview
- Purpose
- Challenges Overview
- Hackathon Operational Overview
  - Resources available
  - Timeline and key dates
  - Submission requirements
  - Judging
- Appendix: Challenges Detail

ABOUT CDP

CDP is a global environmental impact non-profit working to secure a thriving economy that works for people and planet.

We focus investors, companies and cities on taking urgent action to build a truly sustainable economy by measuring and understanding their environmental impact.

Name | @Twitter

How CDP Works

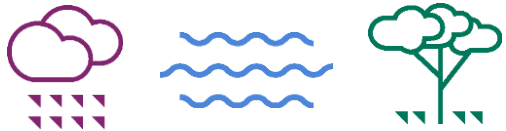
Investors, purchasing companies, cities, states/regions use CDP to make informed decisions, reward companies with superior performance and drive action.

CDP, PARTNERS, INVESTORS & COMPANIES

Request environmental information through CDP

COMPANIES, CITIES, STATES/REGIONS & SUPPLIERS

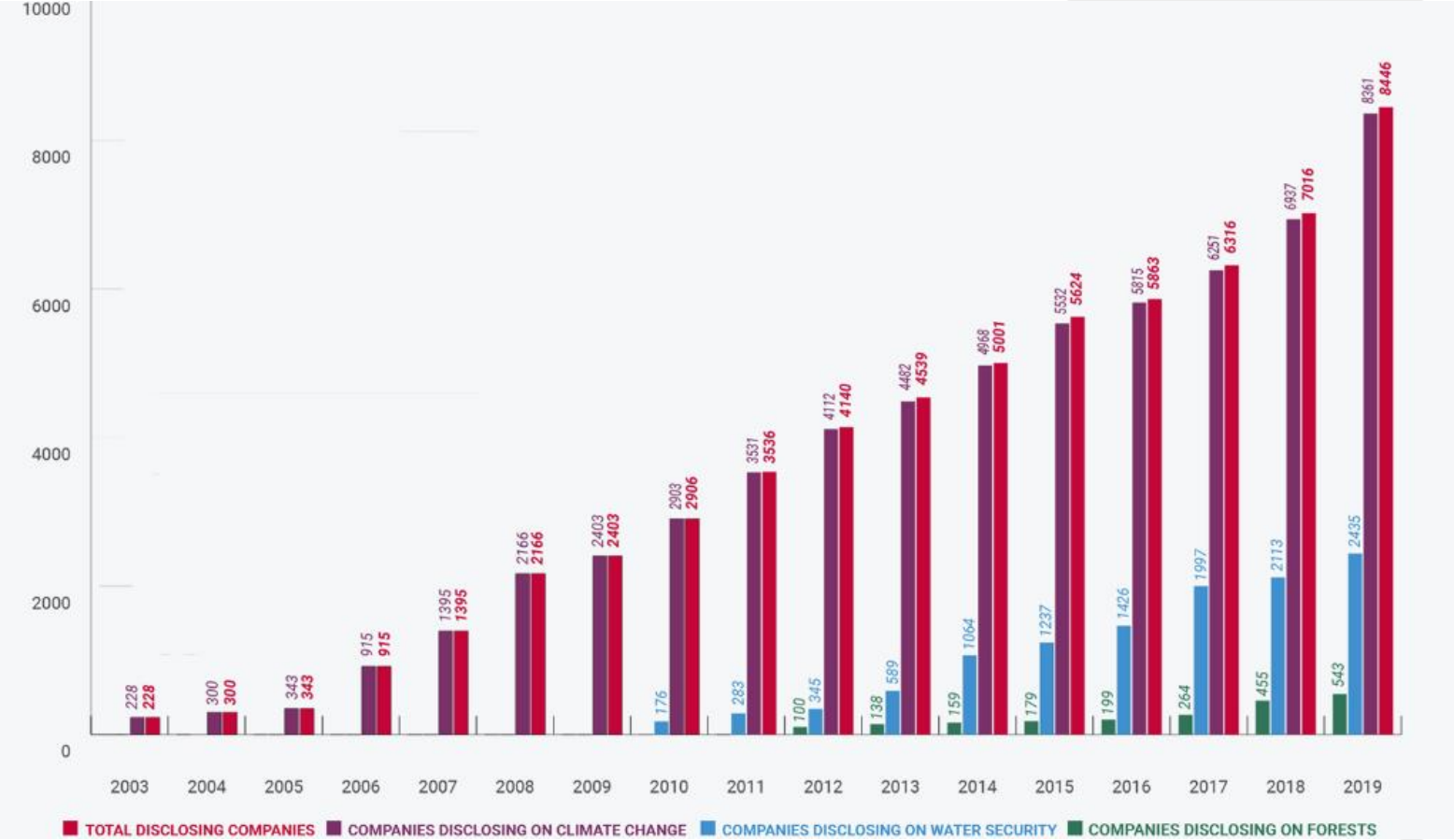
Measure their impact to improve performance



Companies, cities, states/regions take action to tackle climate change, safeguard water resources and prevent the destruction of forests.

Companies, suppliers, cities, states/regions provide data on environmental impacts, risks, opportunities, investments and strategies.

Growth in Corporate Disclosure



Cities Disclosure Drives Action

Disclosure

950+

Cities and States/Regions disclosed to CDP in 2019

Insight

90%

More cities have set emissions reduction targets since 2015

Action

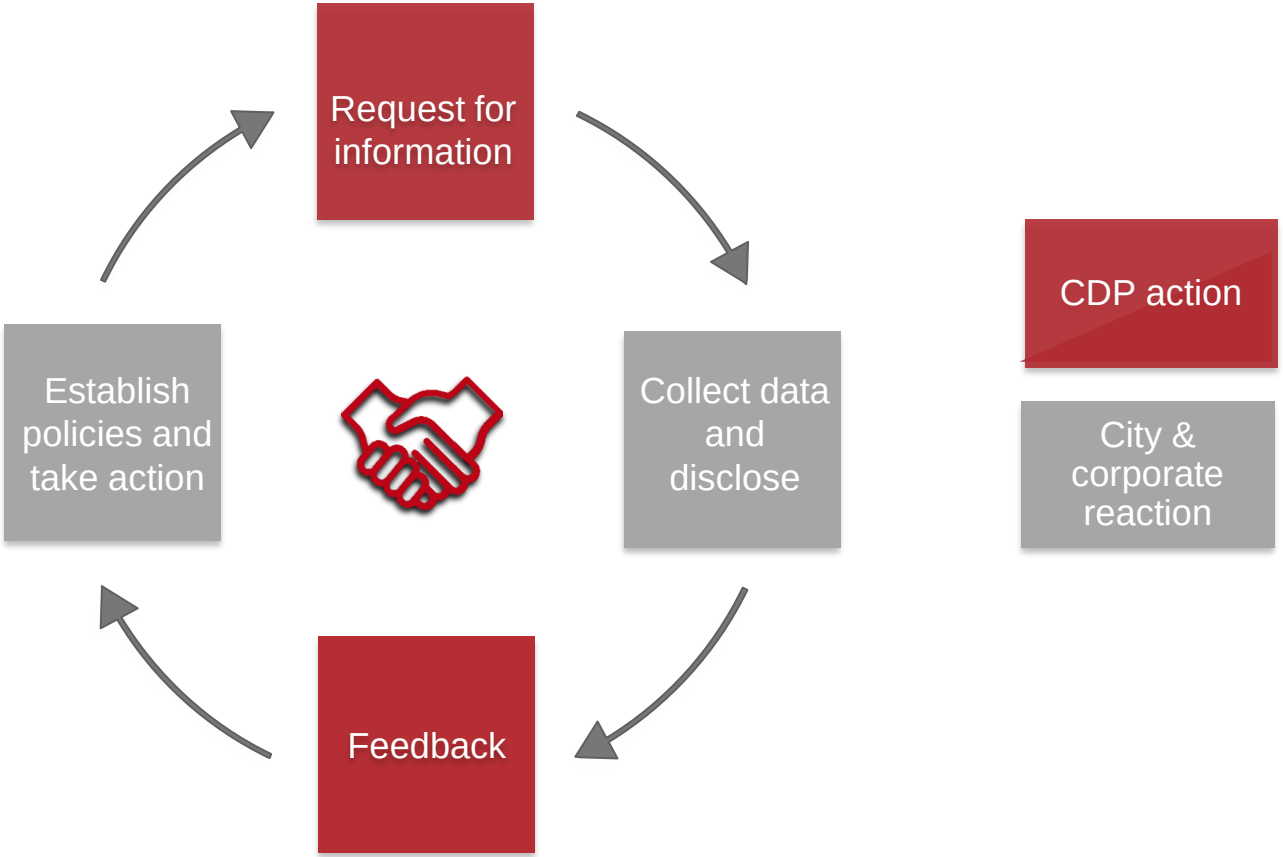
2x

Twice as many cities have created climate action plans since 2015

< 2° World

A climate safe, water secure, deforestation free world

Disclosure Drives Action



How cities can tackle climate change

- **Cities house more than half the world's population, and by 2050, nearly 70% of the Earth's inhabitants will live in cities.** Cities have the potential to play a crucial part in making this transition happen, and many of them are working toward ambitious climate goals.
- **Cities can meet these goals much faster, cheaper and more effectively if they enlist the help of the private sector.** For example, when retrofitting its building stock, it's important for a city to also address the commercial and private sector; or, when transitioning its vehicle fleet to electric vehicles, cities will only have limited success without corporate fleets also making the shift. Additionally, environmentally ambitious companies will have learnings to share with cities – including their headquarter city, cities in which they operate today, and the cities they are considering expanding to in the future.
- In 2019 **more than 60% of cities disclosing to CDP reported that they are already engaged in city-business collaboration** for the low carbon transition, and we are seeing continued demand for partnerships across stakeholders. This is especially important for climate adaptation projects that will help "future proof" cities for both residents and businesses, as well as unlocking new funding and technological opportunities that can only be harnessed when cities and companies collaborate.

PURPOSE

Goals of this hackathon

- The overall goal is to ***identify opportunities for collaboration between various stakeholders***. Since each stakeholder accesses, calculates and submits their CDP response independently, the data can remain siloed. By applying data analysis tools, we hope to identify and unlock opportunities for collaboration, comparison and learnings across companies, cities, states, regions, investors and other interested parties.

- Secondary goals of the data competition include validating/verifying data being reported, predictive models for future scenarios, strengthening of guidance for how to report data, identification of additional datapoints to gather and request, etc.

CHALLENGES OVERVIEW

The importance of the CDP data challenges

- CDP's unique data is used globally by capital markets, corporations, policymakers, and other institutions to inform purchasing decisions, investment portfolio selections and environmental performance benchmarks. **However, in many applications, CDP's data remains underutilized due to the growing volume and complexity of data available.**
- This hackathon is intended to allow expert users to **get hands-on with CDP data, identify opportunities within the data and across other datasets, derive insights and build new tools and visualizations to unlock new potential from this information**.
- Additionally, we hope this competition helps raise awareness of climate change, water security, deforestation risk and city resilience, and the role corporate and city data can play in finding solutions. We hope to **identify areas of impact that CDP data can have on climate change mitigation strategies and solving real-world climate challenges**.

Challenge 1: City-Business Collaboration

- To identify opportunities for companies and cities to collaborate on sustainability solutions, we need to first understand how the data they each report is aligned or divergent by using data science and text analytics techniques.
- **Task**: Utilize data science and text analytics techniques to improve readability of CDP cities and corporate data.

  Visualize shared sentiment between cities and companies on specific topics (such as clean energy, sustainable buildings, clean transport, waste and circular economy).

  Create a KPI model that measures the propensity of cities to collaborate with companies. Provide the justification of your weights and indicators.

Challenge 2: Water Security & Climate Risk

- To help companies and cities fully understand their current water security and future climate risks, we must first inspect the valuable free text data collected via CDP disclosures. This data in its qualitative state is difficult to compare at scale but needs to be compared and assessed with companies' and cities' peers, suppliers, and neighbors to address water and climate risk challenges.
- **Task:** Utilize a subset of freeform response features to inspect from the water security and/or climate risk response.

  Apply your choice of NLP application (Sentiment Analysis, Classification, Question Answering, etc.) to the subset of data you choose – explain the metrics you used to benchmark the success of your algorithm.

  Visualize your data set by combining it with third party geospatial data and applying graph analytics to establish risk patterns (for inspiration, review the CDP 2018 Global Water Report).

Challenge 3: Renewable Energy Demand Projects

- Currently only just over 10% of U.S. electricity comes from new renewable sources and this needs to ramp up quickly to meet growing renewable energy targets and demand. By focusing on matching renewable energy supply to meet demand signals in a given metropolitan area or electric utility's service area, we can help green the electric grid more quickly.
- **Task:** Using CDP and other external datasets, identify where the current city and corporate demands/commitments are for renewable energy in the US and where they will be by 2023.

  Select a city or region and build a predictive model that shows how that city will make progress towards meeting renewable energy supply and demand by 2023.

  Explain the assumptions you've made to build your model and comment on the additional information needed to scale your model beyond 2023.

HACKATHON OPERATIONS

Hackathon Resources

- **Devpost**: the central location for hackathon updates and announcements https://cdp-climate-hackathon.devpost.com/

- **Github:** we will be using a private Github repository for accessing the data and submitting your solutions. **In order to receive access to this repository, you must provide your Github handle and agree to the terms and conditions of the participant agreement here: https://cdphackathon.splashthat.com/** In addition to the data access, here you will find data catalogs and additional recommendations on external datasets relevant to the challenge topics.

- **Slack community:** for community conversations between other participants and submitting questions to the hackathon support team, we will use Slack. You can join the Slack community here.

Hackathon Timeline

July 20th
Hackathon begins

July 27th
Submissions due

July 28th
Judging begins

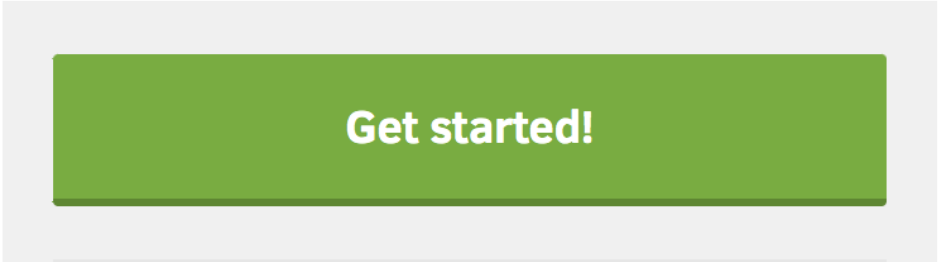
August 4th
Winners announced

Submission Requirements

- The following files will be required for submissions to be eligible for review:
  - **Video** – 2 minutes max; should explain your approach, solution and any interesting findings
  - **PowerPoint deck** – 5 slides max demonstrating your findings
  - **Solution** – should include all required data files, analysis and code (to be submitted on Github)

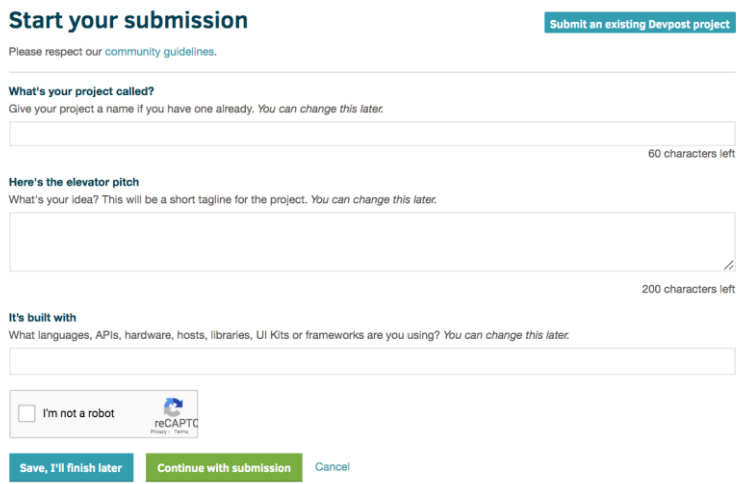
Submission Deadline

- The submission deadline is *Monday July 27th at 12PM EST*,
- Start your submission on Devpost, by selecting “Get Started”

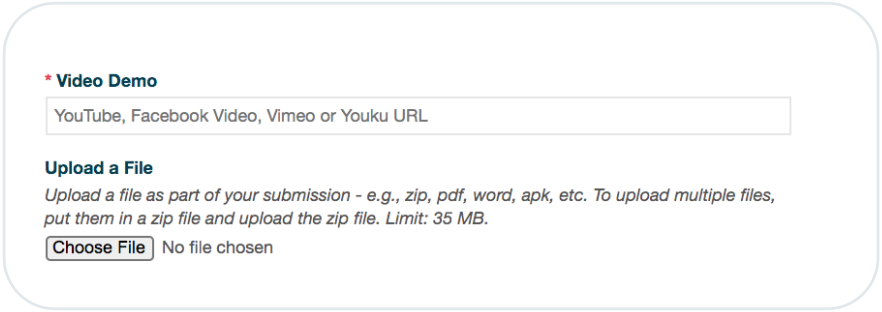


Submission Instructions

Step 1. Begin Submission on Devpost



Step 2. Include video (link) and presentation file



Step 3. Create a branch in the CDPHackathon Github Repository, upload your solution to that branch, and include a link to the branch in Devpost. Detailed instructions on next slide

Github Upload Instructions

1. Create a new branch for your team in the CDP Hackathon repository
Create a branch of the master repository.
Need help? Follow these instructions

2. **Name your branch: [challenge #]-[team name]**
Here's an example: Challenge 1 – CDP

3. **Push your changes**
Need help? Follow these instructions

Judging

- Submissions for each challenge will be reviewed by a panel of CDP staff, including data scientists, analysts, and subject matter experts
- Judges will be evaluating submissions based on the criteria shown on the next slide
- One winner per challenge (3 total winners)
  - Prize for each winning team: $1,500

Judging Criteria

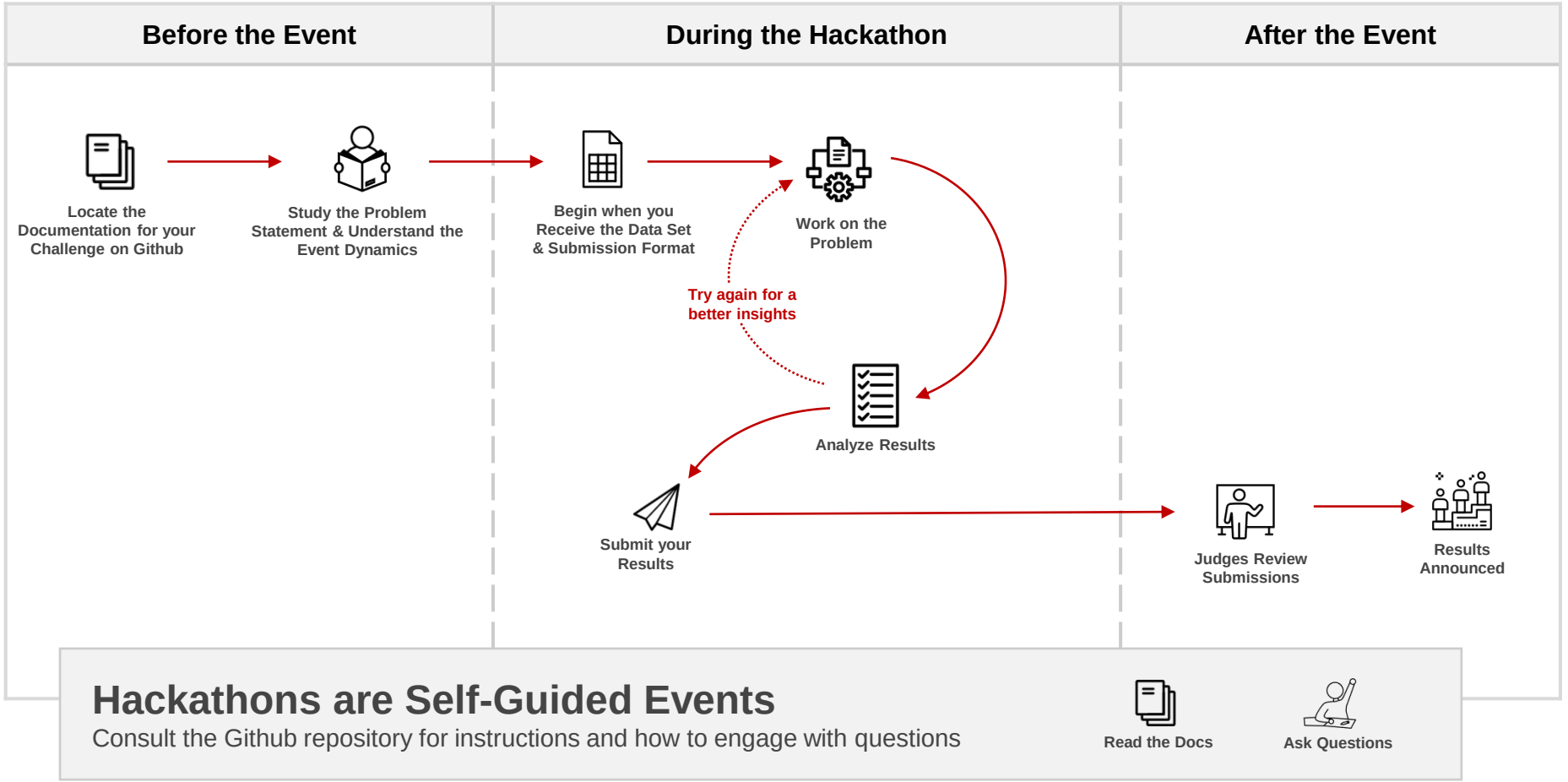
Area	How to score	Score
Use of CDP data and external data: Does the team utilize the CDP data provided in combination with external data sources?	5 = Used CDP data AND external data to develop the solution 3 = Used either CDP data OR external data to develop the solution 1 = Not clear whether CDP data or external data was used	/5
Outcome: Does the team have a clear understanding of the challenge? Do they address the challenge prompt?	5 = Team has a clear understanding of the challenge and addresses the challenge prompt 3 = Team has a clear understanding of the challenge, but does not articulate how they address the challenge prompt (or vice versa) 1 = Team is unclear or demonstrates a limited understanding of the challenge	/5
Feasibility of idea: Could their solution viably be implemented and sustained in the real world?	5 = Feasible and could be applied today with minor cosmetic improvements. 3 = Requires fundamental changes to become feasible. 1 = Submission is not feasible.	/5
Innovation: Are they thinking outside of the box? Does the solution implement trends and best practice from other innovative areas?	5 = Breakthrough innovation that potentially changes how impact is achieved 3 = Moderately innovative; different but not game changing. 1 = Interesting idea, but not fundamentally or distinctively innovative	/5
Technical explanation: Each challenge asks teams to explain their technical decisions (i.e. Justify weights, explain assumptions, etc). Are clear justifications or explanations provided?	5 = Provides a clear explanation/justification of technical decisions 3 = Attempts to provide an explanation/justification of technical decisions 1 = Does not provide an explanation/justification of technical decisions	/5
Total		/25

CDP
DISCLOSURE INSIGHT ACTION

RECAP

Recap on CDP data and this hackathon

- CDP is an international nonprofit organization that runs the global disclosure system that enables companies, cities, states and regions to measure and manage their environmental impacts.
- You will be examining public North American CDP corporate and city data alongside other external datasets to solve one of the three challenges presented.
- The overall goal is to ***identify opportunities for collaboration between various stakeholders***.
- We hope you learn about how companies and cities are tackling environmental issues, utilize data analytics techniques to identify insights from the data, and bring new thinking to solving real-world climate challenges.
- We look forward to receiving and judging your submissions to identify the winners!


Before the Event
During the Hackathon
After the Event
Locate the Documentation for your Challenge on Github
Study the Problem Statement & Understand the Event Dynamics
Begin when you Receive the Data Set & Submission Format
Work on the Problem
Try again for a better insights
Analyze Results
Submit your Results
Judges Review Submissions
Results Announced
Hackathons are Self-Guided Events
Consult the Github repository for instructions and how to engage with questions
Read the Docs
Ask Questions

APPENDIX

CHALLENGE 1: CITY-BUSINESS COLLABORATION

- Over 8,000 companies and almost 1,000 cities, states and regions disclose valuable sustainability data through CDP each year, however each prepares and submits their report independently. Many companies and cities are reporting data identifying very similar risks and opportunities, however they are not able to connect to address them together effectively.

- Public-private partnerships can enable collaboration on things like clean energy, sustainable buildings, clean transport, waste and circular economy that are too big for one entity to tackle alone. For example, CDP has been a partner in the City Business Climate Alliance, which identifies actions businesses within a city can take to help the city achieve its climate plan.

- In order to connect companies and cities together, we need to first understand how the data they each report is aligned or divergent by using data science and text analytics techniques. Then, we can begin to identify where there might be opportunities for collaboration and co-investment for mutual sustainability goals that will ultimately benefit all citizens of the city.

CHALLENGE 1: CITY-BUSINESS COLLABORATION

1. Utilize data science and text analytics techniques to improve readability of data.

2. Visualize shared sentiment between cities and companies on specific topics (such as clean energy, sustainable buildings, clean transport, waste and circular economy).

3. Create a KPI model that measures the propensity of cities to collaborate with companies. Provide the justification of your weights and indicators.

CHALLENGE 1: CITY-BUSINESS COLLABORATION

CDP Data Guidance:

CDP will provide a subset of corporate and city disclosures from entities based in North America. This subset will include qualitative and quantitative self-reported data on emissions reduction targets and other climate- and energy-related targets, reported emissions, and identified climate-related risks and opportunities from the corporate datasets. It will also include data on city population and emissions reduction targets from the cities datasets.

Recommended External Data Sources

Enrich your information with publicly available external data sets such as city-level information on electric power users, employers, planned economic investment, business registries, corporate city taxpayers, members of local chambers of commerce, City Business Climate Alliance which is a consortium of CDP, WBCSD and C40, location data from Google's Geocoding API

CHALLENGE 2: WATER SECURITY & CLIMATE RISK

- Water security is one of the most basic human rights, yet there are a surprising number of examples of U.S. cities – from Flint to Newark to Dos Palos, CA – where it is not a guarantee. Water security issues today can be considered a leading-edge indicator of deeper climate risks in the near future – droughts, floods, hurricanes, coastal flooding – which will strengthen in intensity and frequency.

- CDP collects great quantities of information from cities and companies on water security and expected climate risks, and a lot of that information is in the form of free form text. This makes it hard to analyze and compare, as companies and cities might report the same risks but with different magnitudes, likelihoods or potential impacts. Furthermore, CDP's recent report highlights that future and long-term risk is critically underreported. In fact, the biggest risk is when a key stakeholder does not acknowledge a risk at all!

- In order to help companies and cities fully understand their current water security and future climate risks we must first inspect this free text data in comparison with their peers, suppliers and neighbors. Cities will benefit from an analysis of reported long-term risk with projected climate risk from specific water-stressed river basins.

CHALLENGE 2: WATER SECURITY & CLIMATE RISK

CDP
DISCLOSURE INSIGHT ACTION

1. Utilize a subset of freeform response features to inspect from the water security and/or climate risk responses

2. Apply your choice of NLP application (Sentiment Analysis, Classification, Question Answering, etc.) to the subset of data you chose. Explain the metrics you used to benchmark the success of your algorithm.

3. Visualize your data set by combining it with third party geospatial data and applying graph analytics to establish risk patterns.

   For inspiration, review the CDP 2018 Global Water Report.

CHALLENGE 2: WATER SECURITY & CLIMATE RISK

CDP
DISCLOSURE INSIGHT ACTION

CDP Data Guidance:

CDP will provide a subset of corporate and city disclosures from entities based in North America. This subset will include qualitative and quantitative self-reported data on facility-level water withdrawals and consumption, water usage reduction targets, and water- and climate-related risks and opportunities from the corporate datasets (hint: there may be relevant water-related risks such as floods and droughts reported in the climate change questionnaire disclosures). It will also include data related to cities' water supplies, risks to those water supplies, and climate adaptation measures from the cities datasets.

Recommended External Data Sources:

WRI's Aqueduct tool API, population, geographic factors, climate risk models, Notre Dame Global Adaptation Initiative (ND-GAIN), C40, 100 Resilient Cities, location data from Google's Geocoding API

CHALLENGE 3: RENEWABLE ENERGY DEMAND PROJECTS

- Renewable energy is the fastest-growing energy source in the United States, doubling from 2000 to 2018. Much of this growth is driven by shifting demand signals from cities and companies who are committing to source 100% renewable energy such as through RE100.

- However, most of the new solar, wind, biomass, ocean power, etc. has been going to fulfill new electricity demands, not replacing old, dirtier baseload power. And far too few metropolitan areas or electric utilities are leading the effort, nor are there enough supportive state-level policies (such Renewable Portfolio Standards) guiding the path forward.

- Currently only just over 10% of U.S. electricity comes from new renewable sources and this needs to ramp up quickly. In Canada, large-scale hydro power contributes base load but is not the most sustainable option in the future. CDP's company and city data provide a window into the demand side of renewable energy, both today and in the future. By focusing on matching renewable energy supply to meet demand signals in a given metropolitan area or electric utility's service area, we can help green the electric grid more quickly.

CHALLENGE 3: RENEWABLE ENERGY DEMAND PROJECTS

1. Using CDP and other external datasets, identify where the current city and corporate demands/commitments are for renewable energy in the US and where they will be by 2023.

2. Select a city or region and build a predictive model that shows how that city will make progress towards meeting renewable energy supply and demand by 2023.

3. Explain the assumptions you've made to build your model and comment on the additional information needed to scale your model beyond 2023.

CHALLENGE 3: RENEWABLE ENERGY DEMAND PROJECTS

CDP Data Guidance:

CDP will provide a subset of corporate and city disclosures from entities based in North America. This subset will include qualitative and quantitative self-reported data on emissions reduction targets and other climate- and energy-related targets, emissions management and energy usage (including energy consumption in MWh), and climate-related opportunities from the corporate datasets (hint: there may be information on a company's renewable energy credits and investments in renewable energy in the qualitative responses to questions on low-carbon electricity usage [C8.2f] and climate related opportunities [C2.4a]). It will also include data related to cities' renewable energy and electricity targets, renewable energy installations in MW, and other energy efficiency/renewable energy projects from the cities datasets.

Recommended External Data Sources:

Utility service areas, RPS standards and state-level renewable energy installations, REBA map of where renewable energy installations are happening, Sierra Club's Ready for 100 and Beyond Coal campaigns, map of power plants, WRI/RMI City Renewable Accelerator, ACEEE, location data from Google's Geocoding API

In partnership and funded by

CDP
DISCLOSURE INSIGHT ACTION