ASA ENVR Section 2021 Student Data Challenge

1. ASA ENVR Section 2021 Student Data Challenge

0 Recommend

Christopher Wikle

Posted 4 days ago

Reply Reply Privately

The ASA Section on Statistics and the Environment (ENVR) is excited to announce the 2021 ENVR Data Challenge. The competition is open to student groups of up to 4 participants (and, may include one additional non-student senior collaborator). This contest challenges participants to analyze the environmental data sets described below using statistical, machine learning, and visualization tools and methods. The submission due date is July 1, 2021, and each team will provide one submission (see below). This contest considers the same core data set as the 2020 challenge given some teams had to suspend their participation due to the COVID-19 pandemic, and because the data set is very rich with possibilities.

The data sets for the ENVR Data Challenge 2021 are generously provided by Jupiter Intelligence (https://jupiterintel.com/) and are focused around multiple model-based outputs of maximum temperature, minimum temperature, and precipitation. A description of the data sets is available in the attached document. You may access the data and data description at www.dropbox.com/sh/f667nr0pe2nqubv/... (Note, be careful – there is on the order of 230GB of data in these files!). Contestants must use some portion of these data but are encouraged to consider other data sources for their analyses. Some possible research questions and possible extra data sets are included in the data document as well. At least one member from winning teams will be invited to present their results at the 2021 ENVR Workshop in Provo, UT (September 30 – October 2, 2021, community.amstat.org/envr/events/envr2020workshop).

Contest Rules and Submission Requirements:

1. Teams of up to 4 students (anyone actively enrolled in K-12, undergraduate, or graduate school who has not yet obtained a PhD; note, a team can have a mixture across different experience levels), and one additional non-student senior collaborator if desired (not mandatory). The team will be placed in the category of the most senior student participant (e.g., if the team is made up of 3 Masters students and 1 PhD student, its submission will be judged in the PhD category, etc.).
2. Teams that participated in the 2020 Challenge ARE allowed to participate in the 2021 Challenge but must work on a completely different application.
3. Submissions must be received by July 1, 2021 at 6pm EDT for full consideration.
4. **Submission Requirements:** A report (pdf format) including:
    1. A cover page including the title and listing the educational institution, level, and contact information for each participant, as well as the contact information and title of a senior collaborator if warranted.
    2. An executive summary (no more than one page); in addition to the one-page executive summary, teams can submit an optional video (duration < 5 min) that describes their project (see Submission instructions).
    3. A written report no longer than 10 pages following a scientific paper format (please use the sample PLoS One Word or LaTeX templates located in the data folder)
    4. Code that can be used to replicate the results and graphics presented in the report (see Submission instructions).

1. **Submission Instructions**
    1. Submissions will be uploaded through DropBox at the link given below, with 4.1-4.3 being included in one .zip file with the name: ENVR2021_datachal_report_#####.zip, where ##### corresponds to a short team name (or last name). In addition, submit 4.4 in a separate .zip file entitled: ENVR2021_datachal_code_#####.zip (do NOT include data files in these .zip files – please just provide a link (e.g., to github)). Project and Code Submission Link:

www.dropbox.com/request/013G0KoG438yPrE0cnH2

2. In addition, one team member should **register your team** by sending an email to Chris Wikle (wiklec@missouri.edu) with the following information:
    1. Subject line: ENVR2021 Data Challenge Registration
    2. Team Contact First and Last name
    3. Team Contact Email
    4. University of School Affiliation
    5. Team members and level (e.g., pre-university level, undergraduate, masters, PhD); include non-student advisor if relevant
    6. Team name (optional)

Criteria for judging: Judging will be based on four team categories: K-12, undergraduate, Masters, and PhD, with mixed teams being judged based on the level of the most senior participant. A panel from the ENVR Section will judge the entries based on:

- *Scientific merit.* Winning submissions will be those which make an important contribution to a scientific field, using existing and/or novel statistical methods. This contribution could be a methodological contribution to Statistics, or a contribution to another scientific field, such as Climate Science, Ecology, Environmental Science, or other fields related to the mission of the ASA Section on Statistics and the Environment.
- *Appropriate and creative use of data.* Winning submissions will be those which appropriately use the climate data provided in novel and interesting ways. Use of additional external datasets is encouraged, but not required.
- *Clarity of presentation.* Winning submissions will be those for which the written report is clear and professional, and for which supporting visualizations are helpful, interesting, and clear. Reports should be formatted as a standard academic journal article (see the sample template in the data folder). Additional supplemental visualizations (i.e., videos or Shiny apps) or other research products that help convey important information are encouraged.
- *Adherence to contest rules.* Winning submissions will be submitted on time and be within the required length. All code used for a project must be made available to the judges as part of the submission.

Information: If you have questions, please contact Chris Wikle at wiklec@missouri.edu

---

Christopher Wikle
Curators Distinguished Professor and Chair
Department of Statistics
University of Missouri
