**GWU Data Science September Datathon**

1. **Dates:**

* The problem statement will be released on Friday, September 24th, 10:00 am
* The deliverables must be submitted by 11:59 PM on September 26th

1. **Data:**

You are encouraged to use the following data resources:

* [NOAA Weather Tracking](https://www.ncdc.noaa.gov/stormevents/ftp.jsp)
  + This dataset lists all the severe weather storms in the US from 1950-2021 and tells us how many people died from these events, how many people were injured, and the economic damage caused by each event.
  + **It contains the target parameter ‘DAMAGE\_PROPERTY’**
* [EPA Climate Change Indicators](https://www.epa.gov/climate-indicators/climate-change-indicators-sea-level)
  + The EPA has documented data on climate change based on many different indicators such as coastal flooding rates, ocean acidity, greenhouse gas emissions, heat/cold-related deaths, and much more.
* You may use additional data sources, but they must be properly cited in your report

1. **Problem Statement:**

* We are interested in identifying, understanding, and predicting the effects of climate change on the economic development of America. Your task is:
  + to identify trends in different storms/events and how climate change may be related
  + build a model that predicts future economic damages (‘DAMAGE\_PROPERTY’) brought on by climate change

1. **Report:**

The report (in the form of Jupyter notebook) should include the following sections:

* **Abstract**: Summarize the key findings
* **Pipeline**: Implement the full pipeline of the project, including:
  + Data preprocessing
  + EDA
  + Hyperparameter tuning and / or model selection
  + Interpretation: e. g., feature’s predictive power over the target
* **Discussion**: Suggest policy (to address climate change) based on the findings, must be supported by report from the literature (preferably from top conference / journals)
* **Citation**: if you use any other data sources, you must provide proper citations

1. **Video:**

Record a 5-minute-long presentation, which should include a focus on the interpretation and a discussion of the report

1. **Deliverables:**

As this datathon is individual-based, each participant should submit a txt file named

firstname\_lastname.txt, to the email and include:

* The github repo link of the report (in the form of Jupyter notebook)
* The link of the video
* You submission will be emailed to [jakelieberfarb@gwmail.gwu.edu](mailto:jakelieberfarb@gwmail.gwu.edu) by 11:59 9/26/2021