**Project Description**

**Overview**

You will do a survey study on climate change perceptions. You’ll collect responses from N people (N being... however many you can guilt into doing it), then poke around (i.e., apply Data Science) looking for links, for example, between their personalities and what they reckon about the climate. Earth's dying - let’s get some charts about it, shall we?

**Deliverables? Of course:**

* A report (maximum 8 pages for main paper, because my attention span is like a goldfish’s)

Submission details are [below](https://researchswinger.org/teaching/data_science_project_assignment_2.html" \l "submission).

**Survey Design & Data Collection (How to Annoy Strangers with Forms)**

For each participant of your N, you must collect data on the following:

* **Climate Survey**(Italian Version by ChatGPT) – What people say they care about vs. what they actually do. Classic.
* [**Big Five Personality Test**](https://researchswinger.org/teaching/data_science_big5.html) ([Italian](https://researchswinger.org/teaching/data_science_big5_it.html) Version by ChatGPT) – Because who wouldn’t want to know if they’re a neurotic introvert saving the planet? This is [how you score the Big5](https://researchswinger.org/teaching/data_science_big5_scoring.html).
* **An [Additional Survey](https://researchswinger.org/teaching/data_science_additional_survey.html)** ([Italian](https://researchswinger.org/teaching/data_science_additional_survey_it.html) Version by ChatGPT) – Made it for you to make your life more exciting.
* **Personal Characteristics** – Your survey participant's gender, country of origin, age, zipcode (to consider their area's media income, to correlate their answers to their city's size) , and any chracteristic you decide (surprise me!). NOT names or emails. Bonus tip: You can give them a random ID so you know who’s who without actually knowing who’s who. GDPR and all that jazz.
* **One question (or very short survey) of your choice**. Have a creative moment. In my "moment", I thought about asking the [Moral Foundation Questionnaire](https://researchswinger.org/teaching/data_science_MFQ20.html) (and this is [how you score MFQ](https://researchswinger.org/teaching/data_science_MFQ20_scoring.html). This is the [Italian MFQ](https://researchswinger.org/teaching/data_science_MFQ20_it.html) provided by your best friend, ChatGPT).
* **Signed Consent**. Each participant needs to sign a [consent form](https://researchswinger.org/teaching/Consent_Form_Climate.docx). If a minor, a parent needs to sign a [parental consent form](https://researchswinger.org/teaching/Consent_Form_Climate_Parental.docx). All signed consent forms need to be attached in the Appendix of your report.

**Data Analysis Tasks (Let’s Pretend We’re Scientists)**

You’re going to do some “analysis”. Sounds fancy, but it’s pretty basic:

* **Correlation Matrix** – Draw a pretty table showing which questions are connected to what (ALL questions).
* **Stratified Analysis** – Split your data by personal stuff like gender or country and answer these:
  + Do men and women answer differently?
  + Which questions caused the biggest “aha!” moments?
  + Any obvious clusters, like a bunch of anxious vegans in one corner?
* **Predictive Analysis** – Basic requirement: here you must use both linear regression and logistic regression, as requested in [Homework 4](https://researchswinger.org/teaching/datascience_homework_4.html). Beyond that: Get creative. Use any classification/predictive algorithm you like. Can certain answers predict if someone’s hopeful or just... crying about the future?