Data Science for Environmental Health (DaSEH) Post - online course survey

Thank you for taking the time to participate in this survey. We are collecting data about user experience with our course to learn more about how to improve the data science education experience. This data may ultimately be used for a research publication and reporting to the NIH.

DATA CONFIDENTIALITY: We will collect your name in order to track your progress through the course. We will also ask optional demographic information. The potential risks to you are small. The potential benefits to the community of data scientists, developers, and professors are very high – we will be able to learn how we can improve data science/environmental health education material.

DATA SHARING: We plan to release the data collected from this survey only in aggregate form using GitHub (https://github.com/). No personally identifiable information (such as your name or response timestamp) will be shared.

PARTICIPATION: You must be 18 or over to participate. You will be asked a series of questions about you, your institution type, your prior experience, and your expectations for the course. You can stop participating in the survey at any point. You do not need to answer optional questions. Only answers that are submitted by clicking the submit button at the end of the survey will be recorded. Depending on how you answer the survey, this will take ~10 minutes to complete. You will have an opportunity before submission to edit your responses. Any answers that are ultimately submitted (and not removed by the participant) will continue to be used for research.

General information about traffic to the case study websites and our main website (https://daseh.org/) is being tracked by Google Analytics (https://analytics.google.com). This provides us with summary information such as the number of visitors from different countries. We will not use this to try to identify who has submitted a survey response.

RISK: Risk of participation is minimal. The questions we ask pose minimal security risk to participants as they focus on your personal experience and expectations. Demographic questions are optional. While we have taken measures to ensure the security of the survey responses, please be aware that no online system is completely secure. By participating in this survey, you understand and accept the inherent risks associated with transmitting information over the internet.

CONSENT: The first questions will ask about your age and if you consent (or agree) that the data from your responses be used for research purposes.

QUESTIONS OR PROBLEMS: If you have questions or problems, email Ava Hoffman at ahoffma2@fredhutch.org or contact us through the anonymous Code of Conduct Reporting Form.

You may also contact the Fred Hutchinson Cancer Center IRB Office if you have questions about your rights as a participant. Contact the IRB if you feel you have not been treated fairly or if you have other concerns.

This study has been deemed exempt by the Fred Hutchinson Cancer Center IRB, specifically Exempt status (Category (2)(ii) Tests, surveys, interviews, or observation (low risk)). Thus, the proposed activity does not qualify as human subjects research as defined by DHHS regulations 45 CFR 46.102, and does not require IRB oversight.

IRB No.: FHIRB0020064
Date Approved: 3/19/2024

The IRB contact information is:
Address: Fred Hutch Cancer Center Institutional Review Office
1100 Fairview Ave. N. Mail Stop J2-100, Seattle, WA 98109
Telephone: 206-667-5900
E-mail: iro@fredhutch.org

* Indicates required question

1. **Are you 18 or over?** *

   You cannot participate in this survey if you are under 18.

   Mark only one oval.

   ◯ Yes

   ◯ No   *Skip to section 7 (**Review Responses**)*

2. **Do you consent to your responses being used for research purposes?** *

Please see the description at the beginning of the survey for more information. Use the back button within this survey if you wish to see it again.

Mark only one oval.

◯ Yes

◯ No   *Skip to section 7 (**Review Responses**)*

General Information

Here we are asking general information about you.

3. **First or Given Name:** *
This will only be used to match any of your survey responses through time.

_____

4. **Last or Family Name:** *
This will only be used to match any of your survey responses through time.

_____

Post-Online Course Knowledge

Here we are asking about your experience following the online course.

5. **Now that you have taken DaSEH's Online Course, how would you rate your current comfort with data science, generally?** *

Mark only one oval.

	1	2	3	4	5	
Less	◯	◯	◯	◯	◯	More comfortable

6. **Now that you have taken DaSEH's Online Course, how confident are you in your ability to use data science tools (e.g., R programming, GitHub) to answer Environmental Health related questions?** *

Mark only one oval.

	1	2	3	4	5	
Less	◯	◯	◯	◯	◯	More confident

7. **Please rate how you feel your skills have improved as a result of the online course.** *

Mark only one oval.

	1	2	3	4	5	
Less	◯	◯	◯	◯	◯	More Improvement

8. **Now that you have taken DaSEH's Online Course, how would you rate your current confidence with these specific tools or topics?** *

Mark only one oval per row.

	1 (Less confident)	2	3	4	5 (More confident)
Rmarkdown / Reproducible documents	◯	◯	◯	◯	◯
RStudio	◯	◯	◯	◯	◯
Basic R Syntax	◯	◯	◯	◯	◯
Importing or loading data	◯	◯	◯	◯	◯
Subsetting datasets	◯	◯	◯	◯	◯
Summarizing datasets	◯	◯	◯	◯	◯
Data classes	◯	◯	◯	◯	◯
Cleaning datasets	◯	◯	◯	◯	◯
Reshaping or combining datasets	◯	◯	◯	◯	◯
Data visualization	◯	◯	◯	◯	◯
Statistics / modeling	◯	◯	◯	◯	◯
Functional programming	◯	◯	◯	◯	◯
Github / version control	◯	◯	◯	◯	◯

9. How did you interact with the online course material? *

Check all that apply.

☐ Synchronously (attended set class times with the instructors)

☐ Asynchronously (watched recordings)

☐ Completed optional homeworks

☐ Other: _____

10. What material did you engage with from the online course? *

Check all that apply.

☐ All of the material

☐ Day 1 - R Studio/Basic R

☐ Day 2 - Basic R/Reproducibility/Data Input

☐ Day 3 - Subsetting Data

☐ Day 4 - Summarizing Data / Data Classes

☐ Day 5 - Cleaning Data

☐ Day 6 - Manipulating Data /Esquisse

☐ Day 7 - Data Visualization / Factors

☐ Day 8 - Statistics/ Data Output

☐ Day 9 - Functions

☐ None of the material

Data Skills

Here we will ask you some questions about data.

11. **Now that you have taken DaSEH's Online Course, what changes, if any, would you make to this dataset in order to do analysis or make a visualization?**

Please provide a brief answer of 1-2 sentences.

country	1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910
1 Afghanistan	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2 Albania	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
3 Algeria	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
4 Andorra	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5 Angola	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6 Antigua and Barbuda	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7 Argentina	2070	2490	2820.0	2860	3.80e+03	3990	6.26e+03	6.26e+03	7.62e+03	5.94e+03	8.91e+03
8 Armenia	131	135	130.0	127	1.42e+02	126	1.44e+02	1.61e+02	1.62e+02	1.72e+02	1.68e+02
9 Australia	10200	11400	11400.0	11200	1.16e+04	12100	1.44e+04	1.55e+04	1.68e+04	1.46e+04	1.75e+04
10 Austria	27700	28400	25700.0	25600	2.69e+04	28100	3.36e+04	4.22e+04	5.90e+04	4.22e+04	5.76e+04
11 Azerbaijan	1270	1300	1260.0	1230	1.37e+03	1220	1.39e+03	1.56e+03	1.57e+03	1.66e+03	1.62e+03
12 Bahamas	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
13 Bahrain	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
14 Bangladesh	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
15 Barbados	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
16 Belarus	1970	2020	1950.0	1910	2.12e+03	1900	2.16e+03	2.42e+03	2.44e+03	2.57e+03	2.52e+03
17 Belgium	49700	46700	48900.0	52700	5.04e+04	50300	5.75e+04	5.84e+04	5.79e+04	5.78e+04	6.13e+04
18 Belize	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
19 Benin	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
20 Bhutan	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
21 Bolivia	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
22 Bosnia and Herzegovina	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
23 Botswana	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

12. **Now that you have taken DaSEH's Online Course, in your own words, describe practices you might use to make your work more transparent and reproducible in R.**

Please provide a brief answer of 1-2 sentences.

13. **Now that you have taken DaSEH's Online Course, in your own words, describe a process you would typically take when investigating a new dataset.**

Please provide a brief answer of 1-2 sentences.

Open Feedback

Please provide your thoughts below.

14. **Please provide any feedback that you think could help us improve the DaSEH Online Course.**

_____

_____

_____

_____

Review Responses

Click "Submit" to finish.