

Extra Credit: Statistics is Everywhere

Introduction

Throughout the term, we have (and will continue to) include a recurring “Statistics is Everywhere” segment, highlighting statistical content you may see in the news, social media, or other places. The goal is to demonstrate the relevance of statistics to the world around us. One of the skills we hope you will develop is an ability to think critically about the quality of statistical information and how it is presented in popular (non-scientific) media. We hope that through this course we can equip you to be more discerning consumers of statistical information, and to share your critical thinking with others.

For this extra credit assignment, we are asking you to provide your own “Statistics is Everywhere” example and demonstrate this kind of critical thinking.

You may submit *one* course extra credit assignment during the semester to earn up to a 1% bump in your overall course grade (either the “Statistics is Everywhere” *or* the “Statistical Humor” assignment).

You are not permitted to use generative AI for this assignment.

Due Date

Submissions are accepted on a rolling basis until **December 5th at 11:59pm.**

Guidelines

Find a recently published (within the last year) article in the popular press (more traditional print or online) which discusses some statistical information and which also has cited source material from a scientific peer reviewed publication that is publicly available. For example, this could be a newspaper article or blog post covering a scientific study for which the paper is published.

For the purposes of this assignment, the source material should be a peer reviewed publication in a scientific journal rather than a government or organizational report. The original study material should include a type of statistic/method that we have covered in class.

The statistic/method should not just be a p-value or confidence interval, but could be a method that results in p-values or confidence intervals. For example, a study where a two sample t-test was used would be appropriate, but a study that presents ROC curves, or p-values from a survival analysis would not be in the scope of this course. The objective here is to use the knowledge you have gained in this course to interpret a published study and critique the way the information was presented in a lay publication.

Provide a copy of the source material for both the original material and the popular press article. Please provide a full copy (no credit will be given if only a link is provided).

Apply our PPDAC framework to your critique.

1. In 3-4 sentences, briefly summarize the source study: what is the problem/question being addressed? What type of a problem is this (causal/predictive/descriptive)? What kind of a study design was used? How were the data collected? Do you have any concerns about how the population was selected, or how the sampling was conducted? Does the target population fit the question being asked?
2. In 2-3 sentences, give your assessment of the overall quality of the data collection. Is there any reason to suspect that there is a bias in the way that the data were collected or that variables were measured? Are there other concerns about the data quality?
3. In 2-3 sentences, identify what relevant statistical method/concept that we have discussed in the course was used in the original study, and your assessment of whether this was appropriate. Was the type of analysis appropriate? Are any assumptions needed for that method that were not met? For example, if a linear model was used, do the data fit the necessary assumptions?
4. In 3-4 sentences, describe how the original data/analyses were presented in the popular press (tone/detail etc). For example, did the report make extreme claims not supported by the data? Is enough information presented about the limitations? Is there anything about how the data were presented in writing or graphically that was misleading?

Requirements

You must cite at least one newspaper article, magazine article, or blog post.

You must cite at least one original data source that is from a scientific publication. Provide a citation for the peer reviewed paper in the format requested by the American Journal of Epidemiology.

You may not use an example we have already discussed in class.

You may not work with anyone else on this assignment.

- Note: We expect articles, data sets, and functions to be chosen independently. If there is evidence that students are submitting the same things we will ask you to change your project. To avoid this, decide your idea independently so that your independent thought shines through.

Submission

To receive full credit you must submit your write-up as well as a full copy of both the original study and the popular press article on Gradescope. Please submit a single PDF by combining the original study and popular press article pdfs with the pdf of your answers to the questions.