

Modelling Social Media Poll Data Finds Correlation Between Age and Liberalization of Sexual Opinions*

Analysis of internet influencer Aella’s Kink Taboo Ratings (v3) Dataset

Russell Luchin

April 25, 2024

Data from a survey conducted by internet influencer Aella on taboo sexual attitudes was analyzed to examine the influence of age on perceptions of various sexual acts. The analysis revealed a significant trend indicating that younger individuals tend to rate sexual acts as more taboo compared to older respondents, suggesting a younger generations hold more conservative views. However, the study highlighted substantial issues with data collection methods, including a lack of controls and potential response biases, which may skew results. Recommendations made that future research on similar topics implement stronger data collection frameworks, particularly adapted for online environments.

Table of contents

| | | |
|----------|---------------------------|----------|
| 1 | Introduction | 2 |
| 2 | Data | 3 |
| 2.1 | Raw Data | 3 |
| 2.2 | Cleaned Data | 3 |
| 2.2.1 | Analysis Data | 4 |
| 2.2.2 | Average Data | 4 |
| 2.3 | Data Weaknesses | 4 |
| 2.3.1 | Collection | 4 |
| 2.3.2 | Control | 6 |

*Code and data are available at: <https://github.com/r1uchin/taboosextopics>. Thank you to the R core team for making the analysis of this data possible and to Rohan Alexander for his guidance.

| | | |
|----------|---|-----------|
| 2.4 | Model | 6 |
| 2.4.1 | Explanation | 6 |
| 2.4.2 | Justification | 7 |
| 3 | Results | 7 |
| 3.1 | Age Distribution | 7 |
| 3.2 | Ratings by Category | 8 |
| 3.3 | Average Rating by Age | 9 |
| 3.4 | Average Rating by Age Group | 10 |
| 4 | Discussion | 11 |
| 4.1 | Significance of Findings | 11 |
| 4.2 | Importance of Good Collection Practices | 12 |
| 4.3 | Conclusion | 13 |
| 4.4 | Acknowledgement | 13 |
| | References | 14 |

1 Introduction

The proliferation of online data collection tools has vastly expanded our capacity to gather information on sensitive topics. This paper critically analyzes the Kink Taboo Rating (v3) dataset collected by internet influencer Aella through a Google Forms survey, primarily disseminated via the social media platform X (formerly known as Twitter). Aella’s dataset, hosting responses to various sexual acts deemed taboo, provides a unique lens into the sexual mores of a specific segment of the internet-using public. However, anarchic nature of her data collection practices raises significant questions about the reliability of the insights derived from it and their application to general populations.

The primary focus of this paper is to analyze the impact of age on perceptions of sexual taboos. Taboos were rated from zero to five, with zero being the lowest rating (not taboo) and five being the highest rating (extremely taboo) Initial inspection of the raw data, comprising over 48,000 responses, reveals a significant volume of unusable data due to missing age identifiers, necessitating extensive cleaning. The analysis was conducted on a refined dataset of approximately 3,800 responses, segmented by respondents’ self-reported age. A secondary dataset was forked from the primary cleaned set by aggregating the ages of respondents and the mean rating per aggregate age. This secondary dataset was then used to create a model to explore how perceptions of taboo vary across age groups. The estimand this paper tracks is the correlation of age with average taboo score.

Analysis identified a clear, statistically significant trend: as age increases, the average taboo rating assigned to various sexual acts decreases. This relationship is quantified through a linear regression model, which captures notable, significant variance in responses based on age alone.

The implication of this finding suggest a potential shift in sexual norms over generations, or alternatively a reflection of the more conservative attitudes held by younger internet users.

Moreover, the structure of this paper is designed to tell a story through our comprehensive methodology, from data collection and cleaning to detailed statistical analysis and model building. Each section is crafted to build upon the last, closing with a discussion that contextualizes our findings within a broader methodological landscape associated with data collection.

By dissecting the method of data collection used by Aella and critically assessing its impacts on the data quality and analysis outcomes, this paper not only provides insights into sexual norms but also underscores the crucial importance of rigorous data collection methodologies in social research.

2 Data

2.1 Raw Data

Table 1: First 6x6 Subset of Raw Data

| Run | Program Version | User | Time Started (UTC) | Time Finished (UTC) | Minutes Spent |
|----------|-----------------|------|---------------------|---------------------|---------------|
| 10152615 | 6 | NA | 2023-05-04 19:16:26 | 2023-05-04 19:23:07 | 6.67 |
| 10152616 | 6 | NA | 2023-05-04 19:16:30 | NA | 0.35 |
| 10152617 | 6 | NA | 2023-05-04 19:16:33 | NA | 188.37 |
| 10152619 | 6 | NA | 2023-05-04 19:16:34 | 2023-05-04 19:25:10 | 8.60 |
| 10152620 | 6 | NA | 2023-05-04 19:16:21 | 2023-05-04 19:23:08 | 6.78 |
| 10152621 | 6 | NA | 2023-05-04 19:16:36 | NA | 0.23 |

Raw data was collected by internet influencer Aella, using a Google Forms survey primarily distributed through the social media platform X (formerly known as Twitter). Raw data is hosted on Aella’s personal website “Knowingless”, from which this data was downloaded. Data does not appear to have seen any controls for the respondents, leading to over 48,000 rows of data (most of which becomes irrelevant to us due to an “NA” answer being present in the age column). Table 1 shows a snippet of how much noise is present in the data, which will need to be cleaned prior to any analysis conducted.

2.2 Cleaned Data

Cleaned data resulted in close to 45,000 rows of responses being eliminated from the sheet. Due to the sheer amount of data that was eliminated between the raw data and the analysis data a short breakdown of how this cleaning was conducted will be provided.

2.2.1 Analysis Data

Table 2: First 6x6 Subset of Cleaned Data

| selfage | biomale | gendermale | cis | gendered | firsttime |
|---------|---------|------------|-----|----------|-----------|
| 31 | 1 | 1 | 1 | 1 | 1 |
| 68 | 1 | 1 | 1 | 1 | 1 |
| 35 | 1 | 1 | 1 | 1 | 1 |
| 56 | 1 | 1 | 1 | 1 | 1 |
| 42 | 1 | 1 | 1 | 1 | 1 |
| 43 | NA | NA | NA | NA | NA |

Cleaning for the primary analysis data was mostly done through filtering metadata from the sheet (time of response, program version, empty points columns, etc). However, the majority of the cleaning ended up when empty “selfage” identifiers were filtered out. As we intend to track the effect of age on taboo rating, empty age identifiers left us with junk data which we could not use. The end sheet was approximately 3800~ rows from the 48,000~ rows which were present in the raw data. Additionally, the data was converted from .csv to .parquet - this had no impact on the responses in the data when they were compared directly. Table 2 displays the first 6x6 entries in our cleaned data.

2.2.2 Average Data

A secondary cleaned dataset was created by aggregating responses by the selfage variable and finding the mean taboo rating for each aggregated age. This data set was used to create a model, which we discuss in Section 2.4. However, based on the simple plot we create in Figure 1, preliminary analysis suggests that the data collected for younger ages (<50) is more consistent than the later ages (>50). This is likely due to weaknesses in data collection and control methodology.

2.3 Data Weaknesses

2.3.1 Collection

As mentioned prior, data was collected by internet influencer Aella of X (formerly known as Twitter) fame. Data was collected through a Google Form disseminated through X and the majority of respondents accessed the relevant Google Form through this platform. While there is nothing inherently wrong with collecting data through Google Forms, there lies a problem in bias when said Google Form is distributed through one platform primarily.

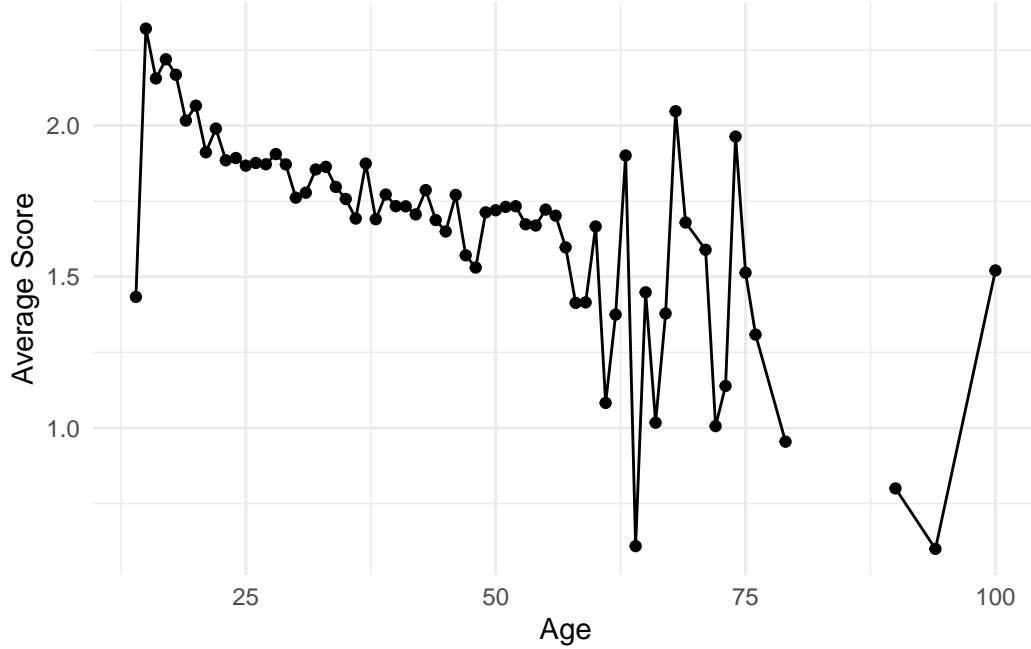


Figure 1: Average taboo rating of sexual acts by age.

Aella is a prominent commentator on sex and sexuality, regularly “pushing the envelope” on controversial sex topics such as age of consent or non-traditional sexual relationships (polygamy, for example). While Aella’s reach is wide due to the size of her X following (204,000~ follows as of the time of writing (Aella, n.d.)), the primary audience she is reaching would be those who already follow her and hold her opinions in a higher regard. Furthermore, the average age of X users is 25-34 years old, with other age groups skewing younger. We can assume that Aella’s followers are of a similar, if not younger, distribution. Age distribution is further discussed in Section 3.1.

This is all to say that Aella’s data is likely to bias towards younger individuals with edgier sexual opinions than someone who has zero knowledge of Aella. This creates a limited set of respondents which cannot be used to gauge any “general” consensus on certain sex topics. While this data is phenomenal for gauging the opinions of Aella’s followers, or even X users, applying this data to any broad population outside the internet is a significant stretch of credibility. Due to the way the data was collected, the applications of this data to the world outside of X is extremely limited if not outright inapplicable. The internet is not real life, therefore this data should not be treated as the opinions of the average Joe you see at Walmart.

2.3.2 Control

Aella's data collection lacked any form of meaningful control, leading to the vast majority of the data her Google Form collected being junk. As mentioned before, the raw excel sheet provided by Aella has around 48,000 rows of responses. Most of these responses lack a selfage answer; which is the only identifier Aella uses to distinguish responses in her data. The lack of identifiers, combined with an inherent lack of control to ensure responses are authentic, inject a "poison pill" into the data where every point needs to be taken with a grain of salt.

To elaborate further, refer to Figure 1. The average data in the pre-50 self age plots neatly and regularly - so much so that a gradual decrease can be found correlated as the age pushes towards the 50 mark. However, passed the 50 mark, the data become irregular. While this can partially be attributed to lower distribution of older ages in the data (discussed in Section 2.4), this can also be due to "fake" ages being inputted in the selfage category. As selfage is the only reliable identifier within the data, any question to the authenticity to this identifier puts the credibility of all the data in jeopardy.

To Aella's credit, this data was never intended for serious academic study. While the lack of control puts the credibility of responses into question, for Aella's intended purpose for the data (interesting discussions on taboo sexual topics) this lack of control won't take away from it. However, this data has the potential to be incredibly insightful - even a small amount of authentication for the selfage variable would've went a long way to making the responses concrete and reliable. That isn't the case, unfortunately, and the extreme irregularities for the >50 selfage respondents puts the authenticity of all responses into question, even if we were to "clean" the irregular respondents from the data.

2.4 Model

2.4.1 Explanation

A model was created to explore the correlation of age with the "taboo rating" Aella uses to gauge opinion on sexual acts. It is a simple linear model, where Y is the average score and X is the age. Below is the basic formula derived from our model:

$$\text{average} = 2.2286 - 0.0121 \times \text{selfage}$$

This model suggests that as age increases, the average score tends to decrease by about 0.0121 for each additional year.

2.4.2 Justification

The linear model developed to track selfage with average score indicates a statistically significant relationship between the two variables. The model above finds that $P < 2e-16$ for the intercept and $P = 1.12e-11$ for the slope. These incredibly small P-scores signify that both the intercept and slope are significantly different from zero, and due to a negative coefficient (-0.0121) in the slope the model finds that the relationship is statistically reliable (as there is a clear negative correlation between the Y and X axis).

Furthermore, about 51.6% of the variability (0.516 in the R-squared coefficient) is explained by age. For a simple linear regression with one variable, this is an incredibly significant amount. It also aligns neatly with what we discuss in Section 2.4, where the younger respondents seem to plot neater than the older respondents.

These scores indicate that our model is statistically robust and therefore practically useful for predicting the average score based on age alone.

3 Results

3.1 Age Distribution

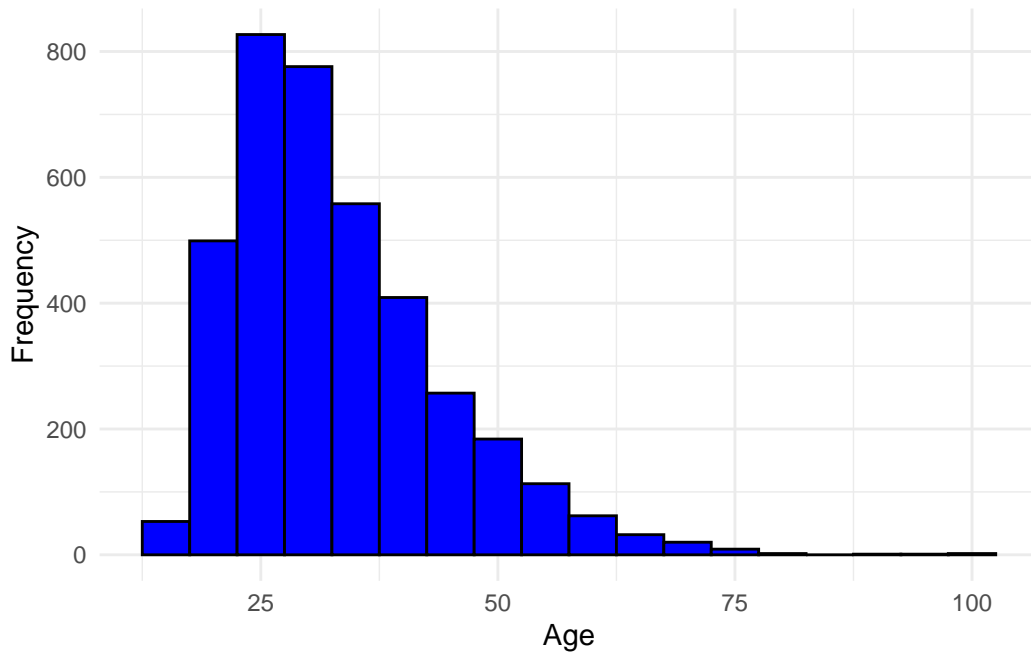


Figure 2: Distribution of respondents by age.

Figure 2 plots the distribution of responses based on age within the cleaned data file. What we find is that Aella’s data seems to be collected from younger respondents (in relation to the overall data), centered around the average X user age range of 25-34 (Duarte 2023).

This has several key implications for the data. The primary one is, as we predicted in Section 2.3.1, there is a substantial bias in the data towards the average user of the X platform. This leads to the data being highly-specific towards X users and not applicable for much else (other than making interesting graphs and starting discussions, as was Aella’s intent). Furthermore, we see a wide range of ages (specifically, age 14 to age 100) despite the vast majority of responses being in the 20-40 age range. This wide range could result in a slight skew in certain analysis, specifically pertaining to model accuracy due the model being built off a simple selfage/average rating correlation without accounting for amount of responses per selfage variable.

Essentially, analyzing the age distribution gives us the clearest indicator of data bias. As it correlates with average X user data (Duarte 2023), the biases we discuss in Section 2.3.1 become validated to a degree. Furthermore, it explains why our model has a 51.6% R-score; the midpoint of this data is coincidentally the point where respondent concentration levels off dramatically, leading to less accurate averages due to a significantly smaller sample size.

3.2 Ratings by Category

Figure 3 displays a heatmap (boxplot) of the concentration of responses to the individual sex acts present in the data. The Y-axis, low to high, displays the concentration of the scores from zero to five. Figure 3 demonstrates the strengths of the data - due to the large amount of responses we can visualize quite neatly what we culturally consider “taboo” and what is considered a “normal” sex act. While the variables are self explanatory, some are esoteric to users of social media or sex enthusiasts like Aella; if an esoteric variable comes is discussed, it will be briefly explained for context.

Interestingly, we find that the majority of sex acts gauged have a middling opinion of “taboo-ness”, where the scores seem to aggregate in the two-to-four score range and very few having a definitive zero (objectively normal) or five (objectively taboo). For the objective categories, there are little surprises - variables “normalsex” (where what ‘normal sex’ isn’t defined), “sensuality”, and “romance” are entirely uncontroversial and variables that are actual crimes, such as “executions” and “pedophilia” are definitively taboo.

Beyond the extreme highs/lows, we see very stable plotting around certain ranges for each sexual topic. This indicates a form of consensus; we don’t have any instances where there is a range of zero to five as a whole, rather the scores all seem to aggregate around two to three taboo scores. For example, in the bottom left cluster, we see three distinct sexual acts/categories - “creepy”, “dirty”, and “futa (sexual media involving depictions of characters who possess all male and female sexual organs)” - all three of these categories have the exact

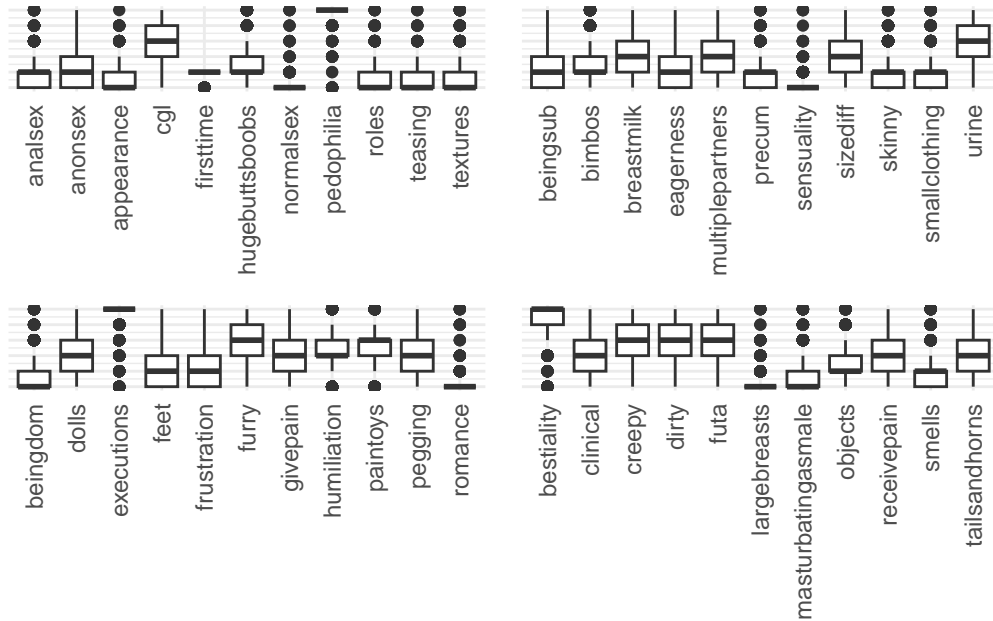


Figure 3: Heat map of average response to each taboo sexual category from 0 (lowest) to 5 (highest), with 0 being the lowest point on the Y-axis and 5 being the highest.

same heatmap aggregating in the middle-high range of taboo score. From a data perspective, this is fascinating, as these are very distinct categories from each other. This plays to the strengths of the data; it finds meaningful insights on opinions of which sex acts are considered taboo due to the sheer amount of data points (3800~ cleaned) it utilizes.

3.3 Average Rating by Age

Figure 4 builds off of Figure 1 by superimposing a line of best fit to the data directly. As this graph is built using the same two variables as our model, the line of best fit is near identical to our model due to it being created through linear regression.

Observing this graph we see significant correlation with what we discussed in Section 2.4.2 - our line of best fit is near perfect in predicting the gradual decrease for ages up to the 50 mark, but then loses accuracy afterwards. While there isn't much unique to discuss, the close association of the line of best fit to our model is further indication that there is a severe bias towards the younger ages while simultaneously having unreliable data found in the higher ages (or, ages with lower proportions of distribution).

Furthermore, the way the data plots in Figure 4 indicates that certain age clusters might be using false/inauthentic selfage values. Observe the "highs" of the ages between 50 and 75 -

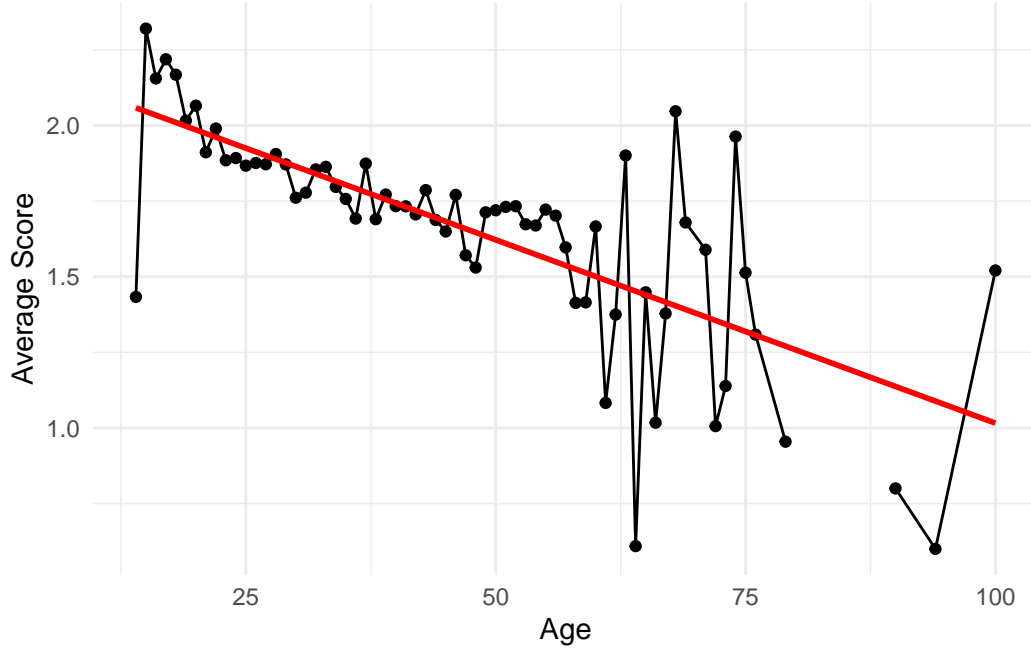


Figure 4: Average rating by age with model/line of best fit superimposed onto graph.

we can draw a nearly perfect line to equal highs in earlier age groups. Due to the extremely low distribution of higher ages seen in Figure 2, alongside the inherent unlikeliness of there being respondents over the age of 60, we can hypothesize that these data points come from younger respondents who decided to falsify their age when asked to disclose it. This further corroborates our discussion in Section 2.3.2, where the lack of controls for the data make it difficult to glean meaningful, definitive conclusions from the data.

3.4 Average Rating by Age Group

Figure 5 separates the average data based on age group. Ages were split directly down the middle, and were based not on the total range of ages but rather the total amount of selfage variables we had in our sample (approximately 60 aggregate selfage data points). We find, interestingly, that despite the model being unable to predict the latter half of ages in Figure 4, the model's accuracy holds up through clustering.

Our model in Section 2.4 finds a strong negative correlation with the selfage and average variables. Splitting our selfage variable into two, between low and high, manages to hold this correlation between the two groups. The higher age group, 44-100, is notably lower (by about .5) than the lower age group of 14-43.

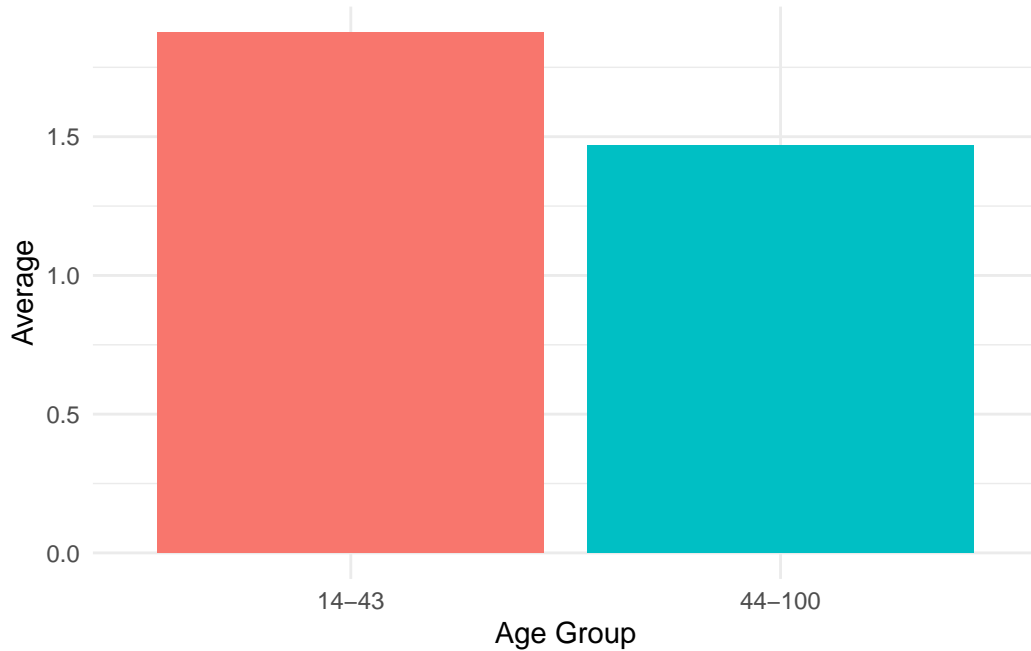


Figure 5: Comparison of average rating between age groups; ages grouped by ‘younger (14-43)’ and ‘older(44-100)’ based on selfage variable from the data.

This speaks to the strength of our model: despite a semi-arbitrary clustering of our selfage variable we find the same correlation between the two groups as our model does between the two variables altogether.

4 Discussion

4.1 Significance of Findings

The significance of this paper’s findings lies less in the trends found in the opinions of the sex acts themselves but rather in meta-analysis of the how the collection practices affected the data. Aella’s dataset does many things well: it is robust, includes identifiers to separate sets of responses, and has an original sample size of over 48,000 individual responses. Based on these factors alone, we’d expect the data to be incredibly insightful with broad application to cultural opinions of various taboo sex acts.

However, the data has a built-in poison pill: a distinct and pervasive lack of any controls. This is most clear in Figure 1 and Figure 4, where our data models very neatly until we reach age groups of low distribution as seen in Figure 2. Even through eye-balling the data, based

on X’s average user group (where the survey was disseminated) it becomes incredibly unlikely that there are as many respondents in the 70-100 age group.

This lack of control stems from Aella’s preferred survey medium alongside how she distributes the survey in the first place. Aella utilizes Google Forms; inherently, there is nothing wrong with this. However, Aella uses no function to ensure the responses are genuine or authenticated; there is no use of email validation, date of birth proof, or even something as simple as name validation through the form. All of these functions are possible through Google Forms, but require manual review from Aella herself. Whether this comes from laziness or simply lacking the need for it for her purposes can be debated ad nauseum, but fact remains that the lack of controls present in the survey allow for bad actors to poison the data as a whole. There’s no way to distinguish which specific data-points are “bad”, despite having overwhelming indication to there being so as Figure 4 shows.

Furthermore, the data is inherently biased due to how Aella distributed the survey. Aella is an X influencer and the majority of her public communication comes from the social media platform. Because of this, it was natural that she found respondents by announcing it through the platform. As Aella’s audience are a subset of X users, who are a subset of the general population, Aella’s data has an inherent bias in who is responding to this survey. This isn’t bad, per say, but results in limited application of the dataset beyond Aella’s X followers who likely share certain “edgy” opinions on sex as she does. For example, the bestiality variable which tracks opinion of the crime of bestiality is not the highest taboo score of five in Figure 3; rather, it falls in a range of four to five on the taboo scale. Aella is known for controversial opinions on the subject of bestiality (Aella 2022), which would explain why bestiality has a substantial proportion of lower taboo ratings than other sexual crimes like pedophilia on the heatmap. This shows how Aella’s sampling created a skew which might not be accurate to the general population.

4.2 Importance of Good Collection Practices

Good collection is the backbone of good data. The US General Social Survey, for example, despite its weaknesses (Beveridge 2007), has extremely strong and consistent data collection practices. Despite the inconvenience, especially in a connected age, the GSS is collected mostly through in person interviews.

The benefits of this are widespread, but primarily lie in the authenticity of the data. In person surveys allow for identifiers to be validated on the spot; this means that data like age and gender can be verified immediately which lends credibility to the responses collected by the survey. Does this prevent specific actors from lying regardless? No, but anecdotally we can assume that individuals who agree to an in-person survey, validate their information, and actively participate in the survey will be answering honestly and in good faith.

This contrasts with Aella’s collection drastically. Once again, Aella collected this dataset through anonymous surveys distributed through X, a platform known for having a veritabily

high amount of bad actors and trolls (Breen 2023). While functional for Aella’s purposes, its not difficult to see how flawed her collection practices are. Instituting any number of simple controls would change this dramatically, but due to there not being any Aella’s practices end up highlighting how much better other social opinion surveys are conducted.

4.3 Conclusion

In conclusion, Aella’s Kink Taboo Rating (v3) dataset provides an extremely robust dataset that is functionally useless. Very pretty graphs can be created with very interesting data points being plotted - this is an objective strength to Aella’s reach and approach to data collection. However, to only praise her data based on how attractive/interesting the data is would be narrow minded.

For every strength this dataset has, a glaring flaw will undermine it. The data is collected with a “quick and dirty” approach, seeming to focus on amount of respondents over verifiable, authentic responses. Here lies the underlying question: is data really meaningful if we can’t apply it generally? If I go to the fictional “Pee-In-Milkshake Convention”, and I sample how many people enjoy pee in their milkshakes, I will likely get a high amount of responses indicating enjoyment of pee milkshakes. However interesting this data might be it is functionally useless - if I try to apply it any population outside of the milkshake pee enthusiasts I would be incredibly dishonest.

Therein lies the flaw in Aella’s data: despite being robust on paper, her sampling and authenticity controls make her data useless. And useless data, no matter how interesting, shouldn’t be used to glean useful insights to our culture. Unfortunately, this isn’t the case with Aella’s data: many other influencers and prominent tech industrialists treat her data as a socially valuable (Aella 2024), where we can glean important insights to how we view sex acts. This is misguided and dangerous, and highlighting these bad practices by analyzing the data independently is how can push against misleading and dangerous data being disseminated.

4.4 Acknowledgement

The analysis conducted in this paper would not be possible without the R programming language (R Core Team 2023) or the accompanying libraries. Thank you to the creators of ggpubr (Kassambara 2023), ggplot2 (Wickham 2016), openxlsx (Schauberger and Walker 2023), here (Müller 2020), patchwork (Pedersen 2024), dplyr (Wickham et al. 2023), corrplot (Wei and Simko 2021), nortest (Gross and Ligges 2015) and knitr [Xie (2023)] for making this paper possible.

References

- Aella. 2022. “Who’s into Bestiality.” <https://aella.substack.com/p/the-great-bestiality-infodump>.
- . 2024. “My Birthday Gangbang.” <https://aella.substack.com/p/my-birthday-gangbang>.
- . n.d. “Twitter Account @Aella_girl.” https://twitter.com/aella_girl.
- Beveridge, Andrew A. 2007. “The General Social Survey and Its Impact in Sociology and Other Social Sciences.” https://www.nsf.gov/pubs/2007/nsf0748/nsf0748_6.pdf.
- Breen, Kerry. 2023. “100 Jewish Leaders Call Out Elon Musk for Antisemitism on x, Formerly Twitter: ”We Have Watched in Horror”.” <https://www.cbsnews.com/news/elon-musk-antisemitism-on-x-twitter-jewish-leaders/>; CBS News.
- Duarte, Fabio. 2023. “X User Stats.” <https://explodingtopics.com/blog/x-user-stats>.
- Gross, Juergen, and Uwe Ligges. 2015. *Nortest: Tests for Normality*. <https://CRAN.R-project.org/package=nortest>.
- Kassambara, Alboukadel. 2023. *Ggpubr: ’Ggplot2’ Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Schauberger, Philipp, and Alexander Walker. 2023. *Openxlsx: Read, Write and Edit Xlsx Files*. <https://CRAN.R-project.org/package=openxlsx>.
- Wei, Taiyun, and Viliam Simko. 2021. *R Package ’Corrplot’: Visualization of a Correlation Matrix*. <https://github.com/taiyun/corrplot>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.