An Exploratory Data Analysis of 80.000hours.org 2-year Review 2021-2022

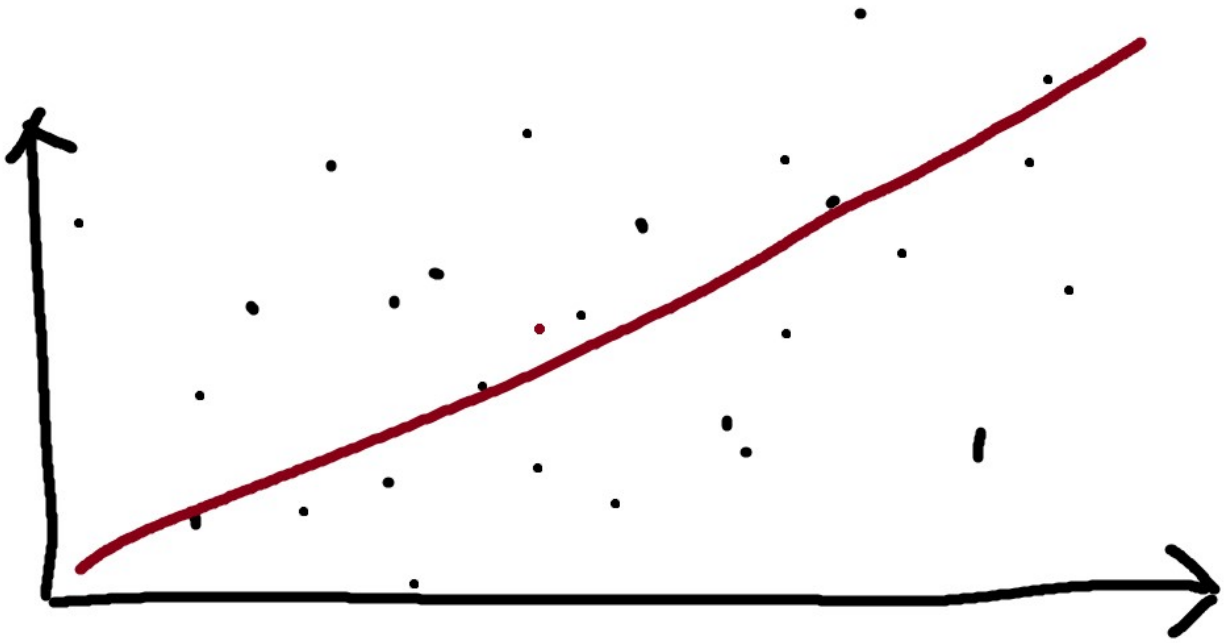
Data Science, Prediction, and Forecasting

David Fjendbo

02/06/2023

201907395

Github: https://github.com/au639194/DataSci



Abstract

This paper serves as a precursor for an API interface to ChatGPT, aimed at providing high impact career guidance in accordance with the principles of 80.000hours.org. The study examined relations and differences in an open-access dataset obtained from a two-year review of the organization. Significant correlations were found between Full-Time Equivalent (FTE) advising and the number of one-on-one calls completed, as well as between financial expenses and vacancy clicks on the website. Furthermore, a t-test revealed a significant difference between the number of advising applications and completed advising sessions. However, the study found no significant relations between FTE advising and the average usefulness score, as well as between FTE web and unique visitors on the website. Other methods and datasets that could have improved the present findings are discussed. The findings suggest that investments of time in various resources aimed at providing high impact career guidance yield moderate payoffs. This implies that dedicating time and effort to this field could potentially result in significant positive social impact, given adequate attention and commitment.

Introduction

This paper serves as a precursor for an API interface made for OpenAI's ChatGPT that aims to provide career guidance. This guidance will be primarily aimed at high school and graduate students or people wanting to transition into new jobs. The framework in which the guidance will be given was created by 80.000hours.org (80k henceforth) (*Start Here*, n.d.) an Effective Altruism (*What Is Effective Altruism?*, n.d.) aligned organisation providing high impact career guidance. Their guidance is primarily provided through the website and 1-on-1 calls with applicants from all over the world. Due to time limitations this paper will be a precursor rather than the actual API interface providing an overview of the data collected by 80k. The dataset that is used in this paper is publicly available through 80k's own two-year review between 2021-2022 (*[PUBLIC] 80,000 Hours two-year review*, 2023). The main purpose of this paper is to provide an exploratory data analysis of this dataset to get an idea of the market for high impact career guidance. Since the dataset is rather small the individual statistical analysis and data manipulation is not very complex. Thus, emphasis will be placed on possible alternatives to these methods, had the dataset been bigger and had more time been available.

Due to GDPR rules and personalised content the field of career counselling does not collect big amounts of data on its users. The available data is thus mainly averages of user input leading to few datapoints (*[PUBLIC] Historical metrics by programme (Static 2022 version)*, n.d.). This results in difficulties making accurate forecasts and predictions of the need for career guidance in the future. Thus, for the limited amount of available data simple correlations and t-tests were chosen to shed light on the possible need for career counselling in the future. The 80k dataset was chosen since the aforementioned API interface for ChatGPT will be aimed at providing high impact career guidance. The content of the 80k website emphasises the impact one can have with career choices (Todd, 2023b) and is thus the best freely available dataset for examining questions related to this application.

Effective Altruism and 80.000hours.org

Effective altruism is an international community of people trying to do the most possible good with the available resources (MacAskill, 2015). The community separates itself from ordinary altruistic movements by trying to do good effectively as indicated by the name. The 80.000 hours organisation is a great example of this aim to do good effectively. Rather than perceiving a career as a means to earn money the approximately 80.000 hours spent working can be utilised to have a positive impact on the world (Todd, 2016). Other than 1-on-1 calls and the website, 80k

provides a job board containing positions where the possession of such a role has a pre-estimated high positive impact on the world. Currently such positions could be as Research Manager for one of the most effective charities Give Directly, working on artificial intelligence (AI) alignment for ARC Evals, Alignment Research Centre, Evaluations Project, etc (*80,000 Hours Job Board*, n.d.). The estimated impact of specific positions is estimated based on their definition of social impact:

> *"Social impact" or "making a difference" is (tentatively) about promoting total expected wellbeing — considered impartially, over the long term (Todd, 2023a).*

In addition, the key principles of effective altruism are used to account for uncertainty, namely importance, neglectedness and tractability. Importance covers over the amount of impact making progress on a specific cause would improve expected wellbeing. Neglectedness covers over effort already made or being made on a specific problem by others, resulting in efforts to solve the problem being more or less impactful. Lastly, tractability accounts for the likelihood of invested resources to actually make a difference towards solving a specific problem (Wiblin, 2019). These three aspects help avoid pitfalls in a landscape of uncertainty.

As an organisation committed to maximising its positive impact on the world 80k utilizes a commonly used work-load quantification, Full-Time Equivalent (FTE) (*[PUBLIC] 80,000 Hours two-year review*, 2023). In the context of workforce management, it serves as a metric to quantify the total work hours contributed by an individual or a group of individuals (CFI Team, 2023). FTE allows for a standardized assessment of workload across different positions and departments, this aligns with the principles underlying the 80k framework and the goal of maximizing impact. In this analysis FTE is expressed as a decimal, indicating the proportion of a full-time workload attributed to an employee or group. A full-time workload is defined by the standard number of hours worked by a full-time employee. For instance, it might be 37 hours per week or an alternative predefined value. To calculate FTE, the total hours worked by an individual or group are divided by the number of hours constituting a full-time workload. This calculation ensures a consistent representation of workload across departments and individuals. For example, if an employee contributes 25 hours per week, the FTE value would amount to 0.675 (25 hours divided by 37 hours, assuming a 37-hour workweek).

ChatGPT

As evident from the burst in number of users in a historically short time for an internet app, OpenAI's ChatGPT has unforeseen potential (Milmo, 2023). This potential and the money it brings have both the opportunity to be utilised for good and for bad. The severity of ensuring that these large language models are used for good have recently been highlighted with Sam Altman, the CEO of OpenAI, participating in a senate hearing in the US (Kang, 2023). Similarly, the Future of Life Institute published an open letter to pause giant AI experiments like those made with ChatGPT and GPT-4 which 31,810 people have signed including e.g., Elon Musk, Stuart Russell, and Steve Wozniak ('Pause Giant AI Experiments', 2023). Some of the leading researchers at Microsoft even claim that models such as GPT-4 might be a precursor for artificial general intelligence (Bubeck et al., 2023). Specifically for the purpose of data science ChatGPT has great potential for increasing productivity especially in programming and translation tasks (Hassani & Silva, 2023). This potential is typically followed by the fear that these large language models will take away learning potential if not utilized responsibly (Mhlanga, 2023). Furthermore, the application of ChatGPT for career counselling can be improved, especially for high impact career guidance, as can be seen in Appendix A.

ChatGPT uses a neural network with transformers, a type of deep learning architecture specifically designed for processing sequential data, such as text. This neural network architecture was first introduced in a paper by Google in 2017 called "Attention Is All You Need" (Vaswani et al., 2017). Explaining the entire neural network architecture is not within the scope of this paper, so a quick overview is given instead. As seen with the rapid growth in use of ChatGPT and DALL-E 2 the transformer model has been successful in various tasks, including language translation, text generation, image generation, and question-answering (Cretu, 2023). The transformer model uses an attention mechanism, which allows the model to selectively focus on different parts of the input sequence. This attention mechanism enables the model to capture dependencies and relationships between tokens in the input. ChatGPT, is trained using a large dataset that consists of input-output pairs labelled by human annotators, also known as supervised learning. The model was then asked to generate multiple responses to user prompts and human annotators were asked to rate the usefulness of these responses. Lastly, the previous model's responses to user queries are evaluated using reinforcement learning, a process that tries to maximise some reward function by updating its output generation process (Cretu, 2023).

Career guidance

The amount of freely available data in career guidance is very limited. This is the result of most career guidance aiming at being personalised to a degree where saving data easily violates e.g., the Danish GDPR rules (Cardoso et al., 2012). Fortunately, for data scientists some organisations, such as 80k, also include metrics that are not personalised and thus provide open access datasets. Despite the lack of datapoints the variables included in the data create a foundation for an interesting exploratory data analysis.

This paper will look at differences and relationship in the dataset. The first thing will be to test whether there is a difference between completed advising calls and advising applications. This could provide insight into whether 80k and future organisation should utilise a more scalable design for providing high impact career counselling. In part, the website does this, but usually people prefer in-person career counselling (Khurumova & Pinto, 2023).

Next will be an examination of a possible relationship between the Full-Time-Equivalent (FTE) spent their website by employees and number of unique visitors to the website. A positive correlation seems likely since time invested in the website by the organisation ought to create an incentive for new users to visit the website or that more new users encourage the organisation to spend more time on developing the website. This relationship is interesting mainly to have a data-driven encouragement to spend more time on developing a well-functioning API interface for the career guidance model.

Third is a possible relation between FTE spent on advising and number of one-on-one calls completed. Fourth is FTE spent on advising but examining the relation to the average usefulness rating from people who participated in one-on-one calls. These two relationships have a similar function to the one above, namely whether investing more working hours into specific functions of the high impact career guidance pays off. This is true no matter the causal directionality of the correlations.

Fifth and last is a possible relationship between financial expenses and vacancy clicks on the website. The vacancy clicks are here used as a proxy for “revenue”, since successful career counselling often ends in either employment, a chosen education, or the like. Here we hope again to see a positive correlation since spending more money should lead to more people using 80k’s services e.g., the job board where vacancies are placed (*80,000 Hours Job Board*, n.d.), or more vacancy clicks should lead to more investments into creating more “revenue”. This relationship is mainly interesting to double check whether money spent in the organisation are also high impact.

Also, financial investments into career guidance can be hard to track and this examines one possible outcome variable invested resources in career guidance, namely vacancy clicks.

Methods

Data

The dataset used for this study was obtained from 80.000hours.org (*[PUBLIC] Historical metrics by programme (Static 2022 version)*, n.d.). As mentioned, their main purpose is to provide high impact career guidance. This aligns with the intent to make an API interface for ChatGPT that aims to provide high impact career guidance. Thus, due to the high impact framework of 80k and their focus on career guidance, this dataset was ideal for an exploratory data analysis. The included time periods in the analysis vary according to the available data. In order to have more datapoints for some tests the quarterly dataset was used as opposed to the yearly one. Both the yearly and quarterly datasets are freely available through 80k's 2 year-review 2021-2022 (*[PUBLIC] 80,000 Hours two-year review*, 2023). Multiple methods were used for the analysis since the datapoints are partially paired in the dataset, e.g., for the difference between number of applications and number of completed advising calls, some of the people representing applications are repeated in the number of completed advising calls. In order to be conservative in the analysis more than one test was performed for each possible relationship in the data.

Software

The statistical tests and plots were made using the R programming language (R version 4.2.2, 2022-10-31 ucrt) (R Core Team, 2020). Plotting functions were provided by *ggplot* and statistical functions such as *cor.test()* and *t.test()* were included in base R as cited above. Other relevant packages and the code are available on the linked GitHub repository. All pre-processing steps and data manipulations are also available in the code.

Analysis

Two main methods for examining relationships and associations in the data were used: t-tests and correlation test. T-tests were used to identify whether there was a significant difference between the number of applications for advising and the number of completed advising calls. Three correlation tests were used to examine other relationship in the data. The first one is total FTE for the website and the number of unique visitors to the website. The second is total FTE for advising and the number of completed one-on-one calls. The third and last is financial expenses and number of vacancy clicks on the website. Furthermore, to visualise possible relations in the data, plots are

provided below. Prior to any statistical test being made assumptions of the involved variables were tested. These included testing normality with a Shapiro-Wilk normality test, and testing for homoscedasticity with both qq-plots and plotting residuals vs predicted values based on a linear regression model. The significance level of the statistical tests was set at alpha = 0.05 coherent with the standard in most non-health sciences. Despite this p-value approach being increasingly problematic, for the current dataset a more rigorous Bayes approach was not possible (Cristea & Ioannidis, 2018; Hubbard & Lindsay, 2008).

Results

Difference between call applications and completed calls.

Initially a paired t-test was performed to examine the difference between the number of advising applications and completed advising calls between 2017 and 2022. Since some applicants are also in the completed calls data the samples are not independent. This test used the yearly dataset with fewer datapoints. The analysis showed a non-significant difference $t(5) = 2.3279$, $p = 0.0674$ between the mean number of advising applications and the mean number of completed advising calls with a mean difference of 1367.2. The 95% confidence interval for the difference in means ranged from -143.1 to 2889.5.

Looking at the plots in figure 1, the same trend seems to be present in the increase of both number of completed calls and number of applications. Yet the gap between indices on the y-axis in figure 1A is rather large. To further examine whether the first test was inaccurate due to a lack of data, another t-test was performed using the quarterly dataset (figure 1B). This dataset includes slightly more datapoints, yet the trends are still similar in both the right and left-hand side of figure 1B. Another paired t-test was performed indicating a significant difference between the number of applications and number of completed calls with $t(21) = 4.607$, $p = 0.00015$. The 95% confidence interval for the difference in means ranged from 199.9 to 529.0 and a mean difference of 364.5. Lastly, a non-linear Wilcox rank sum test was performed which was also significant. All test result outputs produced in R are included in Appendix B.

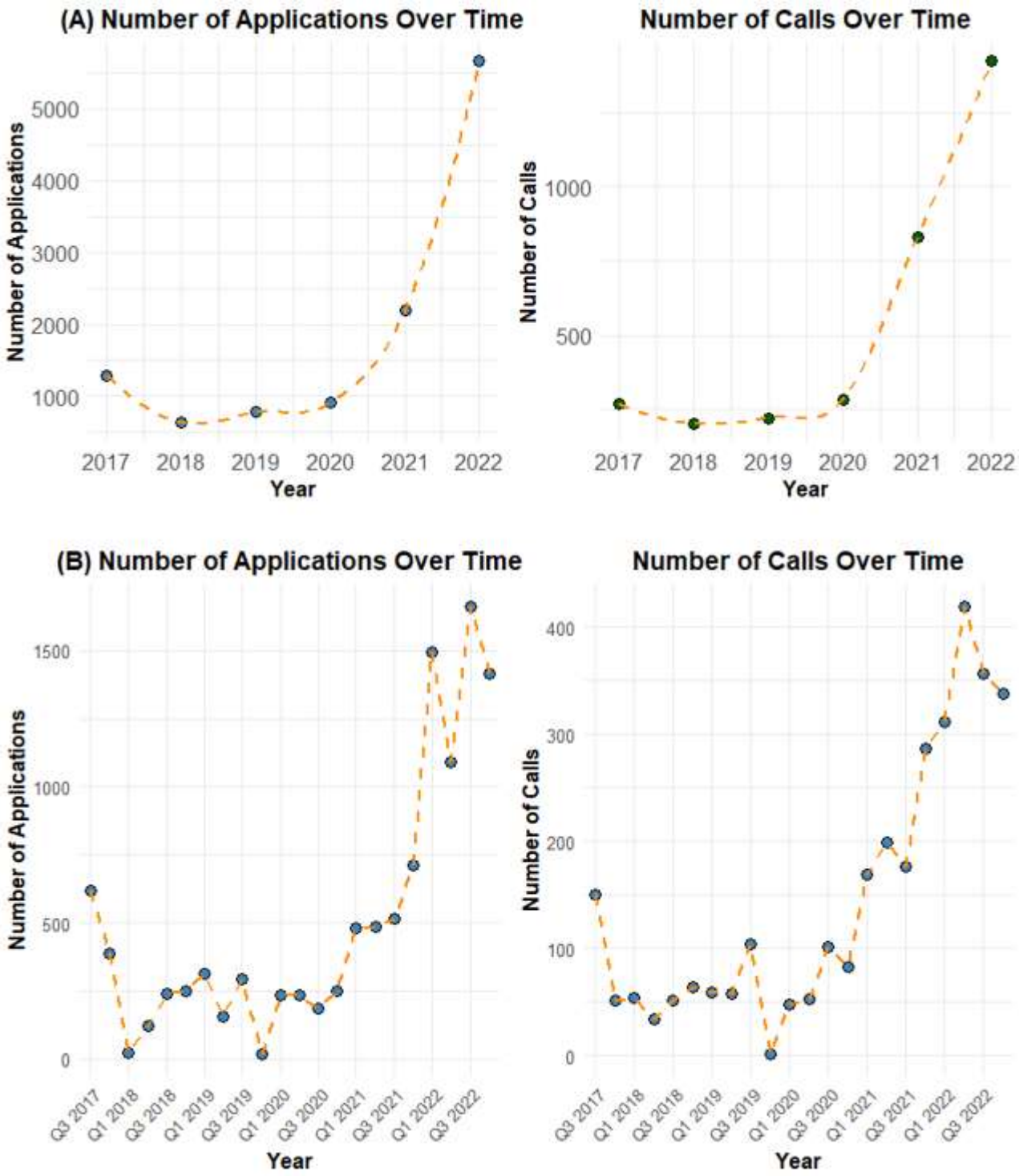


Figure 1: (A) Yearly Number of Calls and Applications, (B) Quarterly Number of Calls and Applications

Correlation between Total FTE on web and Unique visitors.

The relationship between total FTE on the website and number of unique visitors between 2016-2022 was examined using a correlation test based on the initial plotting of the data

shown in figure 2. This analysis also used the quarterly dataset to include more datapoints to improve accuracy of the tests. The Pearson correlation test was significant, revealing a slight positive correlation between total FTE for the website and number of unique visitors r(26) = 0.382, p = 0.044. The 95% confidence interval for the correlation coefficient ranged from 0.011 to 0.661. Yet since the total FTE on the website is not normally distributed, violating one of the assumptions of the Pearson correlation test, a Spearman rank correlation was also performed. This test was insignificant with S = 3538.5, p = 0.87 and a rho value of 0.0316. This test is non-parametric making it more conservative and reliable when the data violates assumptions of a regular Pearson correlation test.

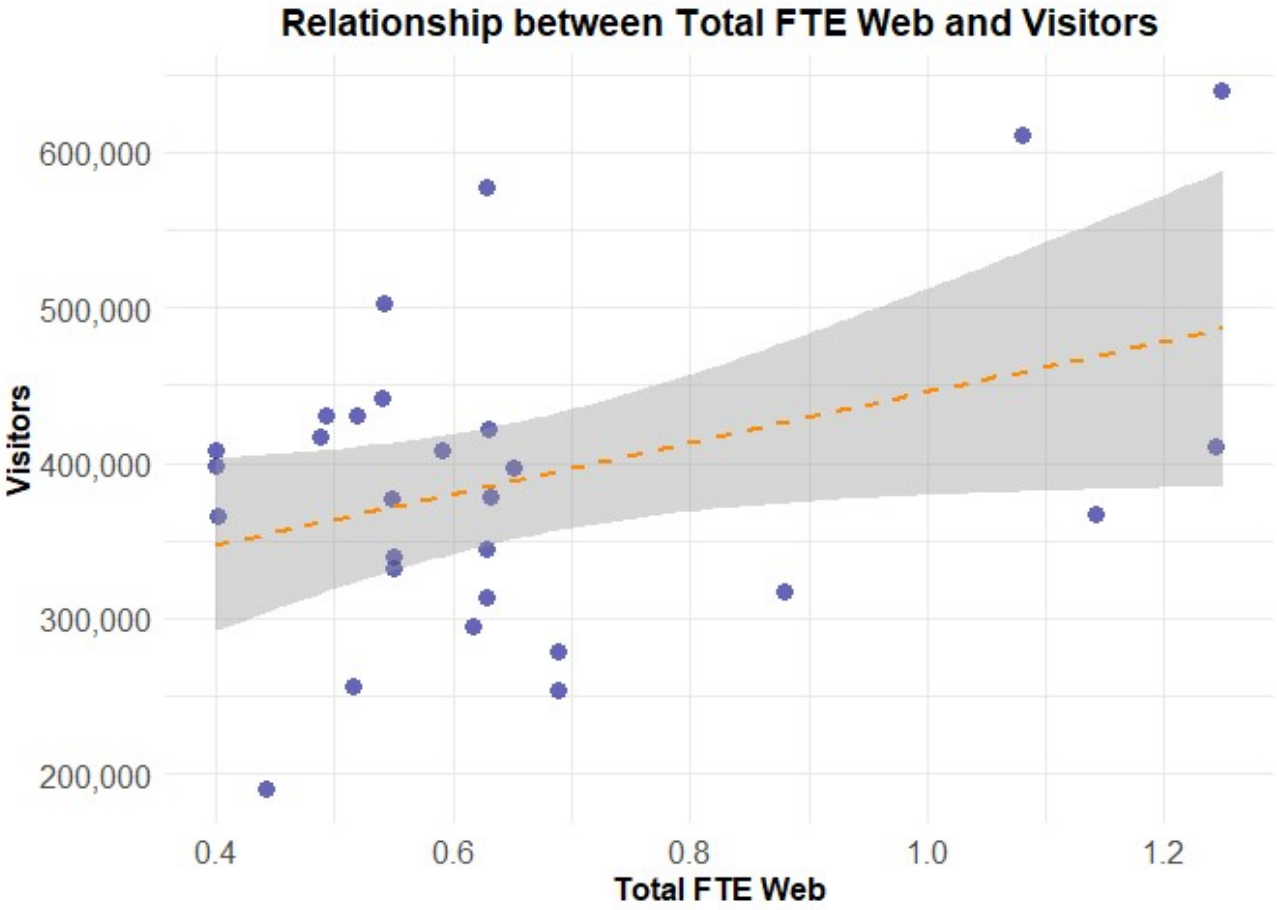


Figure 2: Total FTE Web and Number of Unique Visitors. Quarterly datapoints from 2016-2022.

Correlation between Total FTE advising and Number of one-on-one calls.

To see whether FTE spent on advising affected the number of completed calls, the second relationship, between total FTE for advising and number of one-on-one calls completed was examined using a correlation test. Here the quarterly dataset was also used. The FTE advising

variable was not normally distributed, yet since the data was homoscedastic, and the number of one-on-one calls was normally distributed a Pearson correlation test was conducted. The test revealed a significant positive correlation between the two variables r(18) = 0.822, p = 8.764e-06, with 95% confidence intervals at 0.596 and 0.927.

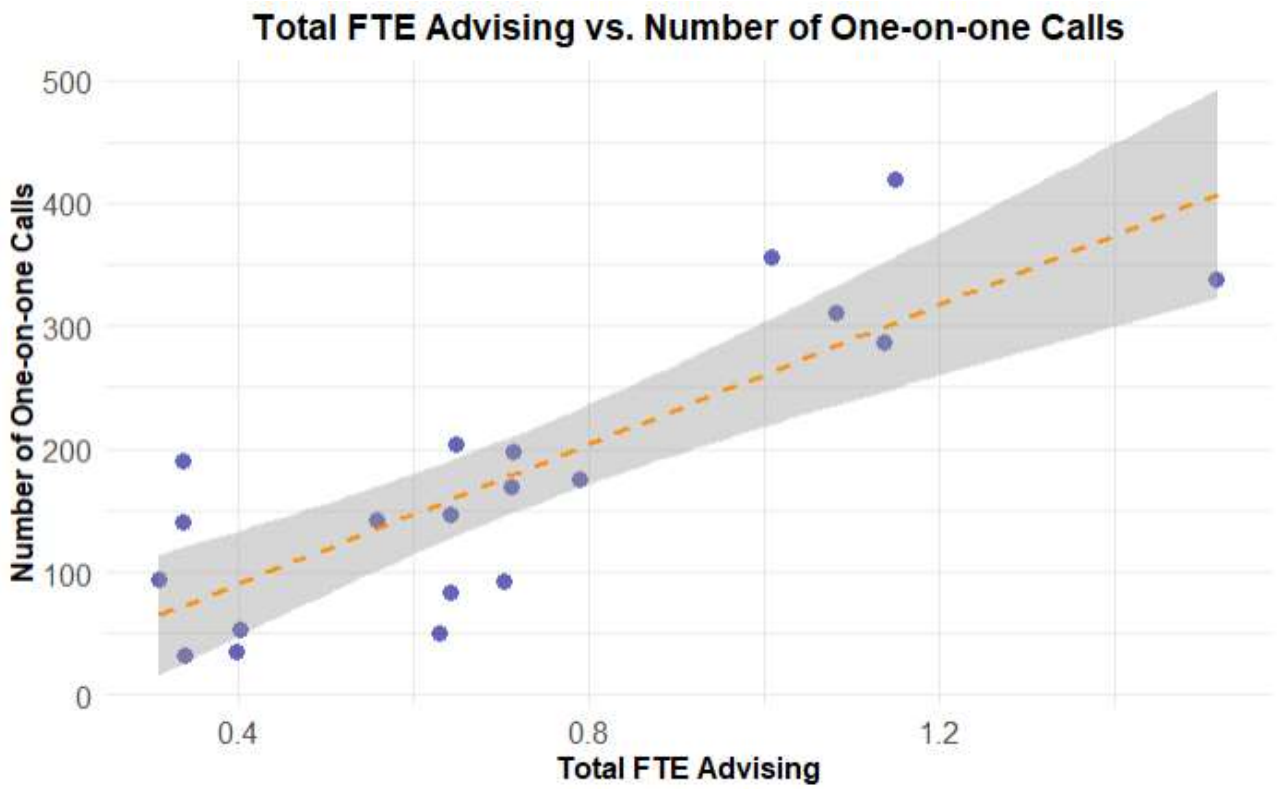


Figure 3: Relationship Between Total FTE Advising and Number of One-on-one Calls. Quarterly timepoints from 2018-2022.

Correlation between Total FTE advising and Average usefulness rating.

The relationship between total FTE and the average usefulness rating (between 1-7) is used as a proxy for examining the quality of time spent on advising calls. The quarterly dataset was used. As can be seen in figure 4 there seems to be a slight negative correlation between FTE and usefulness. A Pearson correlation test was conducted since data for both variables was normally distributed and rather homoscedastic. The results for the correlation test were insignificant with r(6) = -0.409, p = 0.31 and 95% confidence intervals for the correlation coefficient at -0.865 and 0.415. These confidence intervals reflect well how few datapoints influence the reliability of the results.

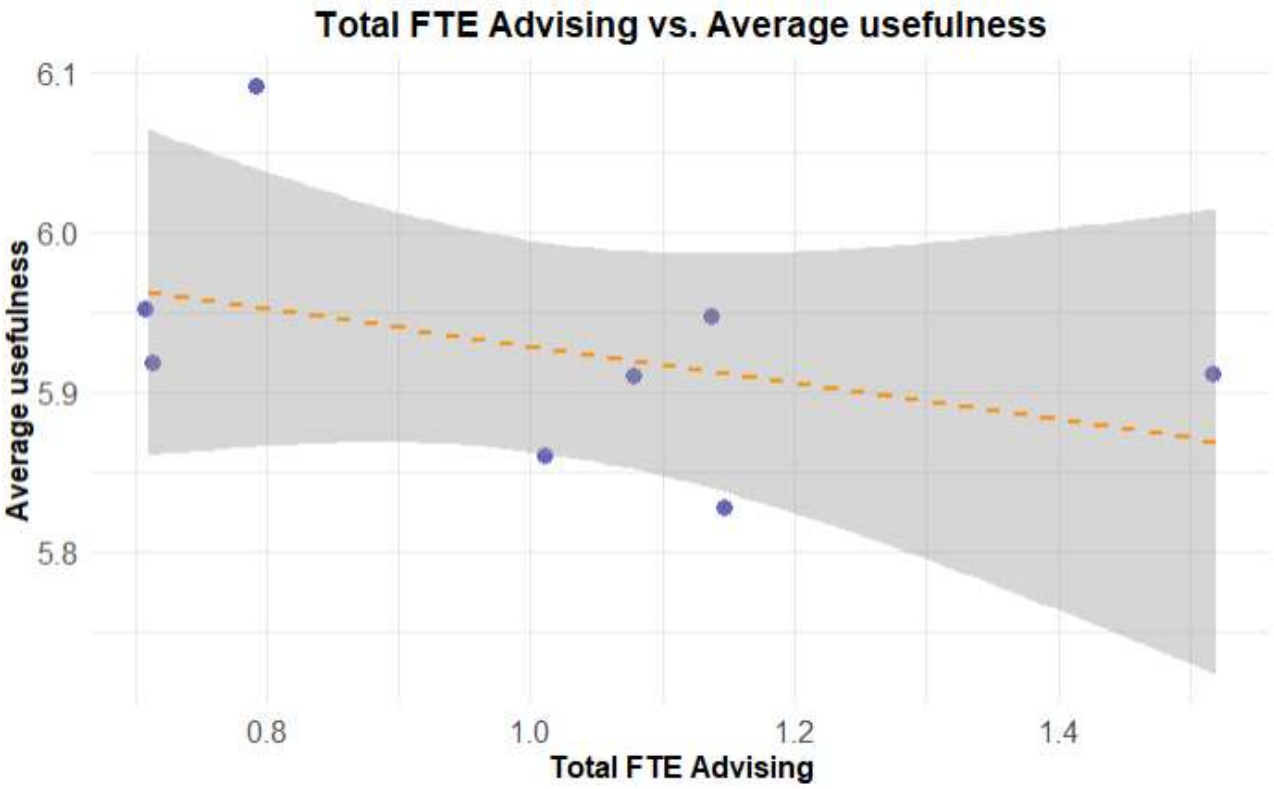


Figure 4: Relationship between Total FTE and Average usefulness score of 1-7. Quarterly timepoints between 2021 and 2022.

Correlation between financial costs USD and Vacancy clicks.

To see whether spending more money inside the organisation leads to more engagement from users, the financial expenses and number of vacancy clicks are used. Here due to a lack of data in the quarterly dataset, the yearly one was used. To test whether there was a significant relationship between financial expenses and vacancy clicks a correlation test was conducted. Since none of the assumption tests were very informative due to a limited number of observations, a Spearman correlation was used. The results revealed a perfect positive correlation with S = 4.44e-17, p = 0.017 and a rho value of 1. Since there are very few datapoints this perfect correlation is somewhat spurious.

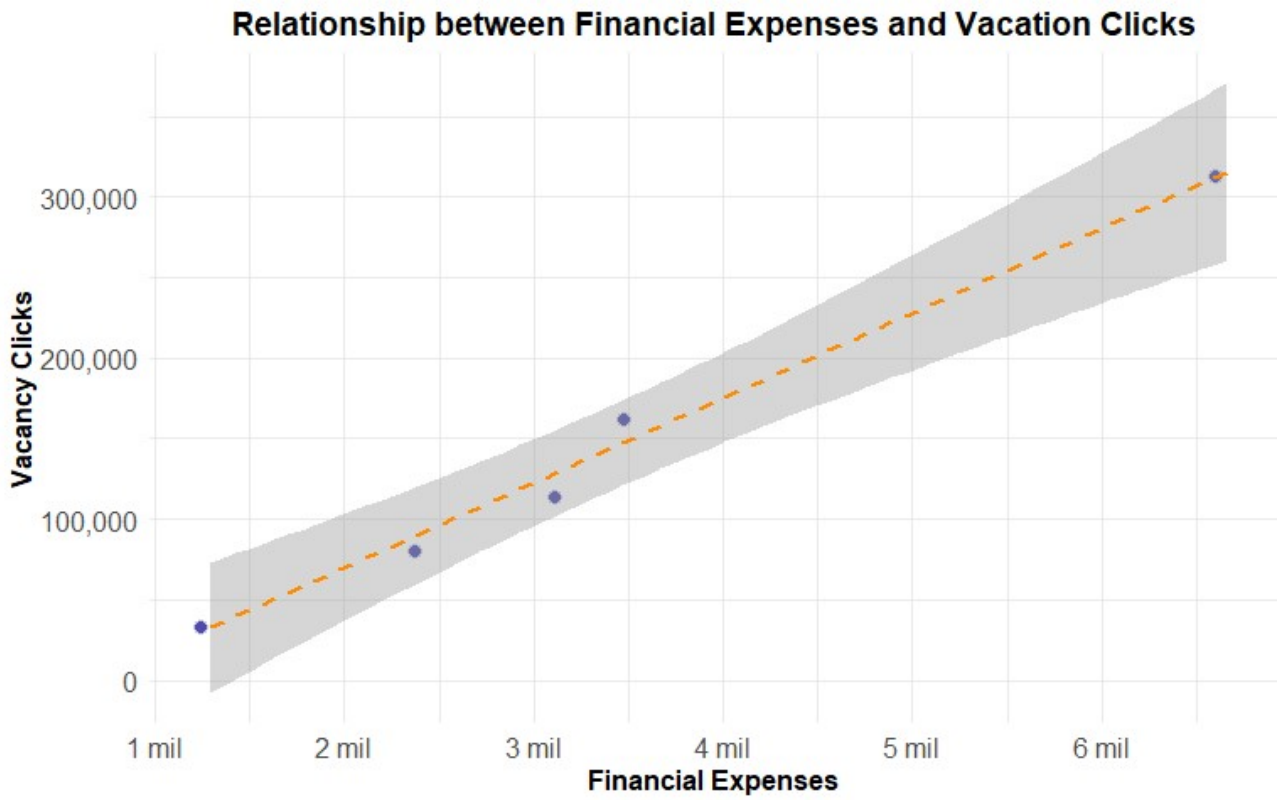


Figure 5: Financial Expenses and Vacation Clicks. Yearly timepoints from 2018-2022.

Discussion

A general overview of the results clarifies the lack of datapoints in the chosen dataset. Specifically, the 95% confidence intervals indicate that both the Pearson and Spearman correlation coefficients and the test statistic from the t-test are very unsure whether the correct relation or difference is detected from the data. For this exploratory data analysis of this open access dataset the certainty of the tests does not pose a significant problem. The purpose of the tests and plotting is to gain preliminary insight into relations and differences between variables in the data. This purpose is fulfilled by the present analysis and plotting. Given more time and resources a more thorough methodology would have been chosen and another non-public dataset could have been selected. Alternate dataset include data pulled from Statistikbanken (*Statistikbanken*, n.d.), a Danish website with a variety of freely available data. Other options could be to ask hiring companies such as Jobindex (*Jobindex*, n.d.) for data, reach out to career counsellors or create an experiment/questionnaire to gather data.

Alternative Methods

As an alternative to the t-test conducted to find whether there was a difference between advising application and completed advising calls an optimal pooled t-test could be performed. This has previously been recommended for small samples sizes and data with missing values (Guo & Yuan, 2017). Similarly, if a specific theory had justified the use of regression rather than a correlation test this would have been preferrable since it allows for incorporation of more uncertainty in the model and allows one to interpret directionality of causal relations. This is also true for the permutation test; by reshuffling the data both more uncertainty is accounted for and the limited amount of datapoints is somewhat circumvented.

Ideally enough data and time had been available to make a Bayesian model. Such a model would allow for the incorporation of uncertainty and prior knowledge directly into the model (Mathys et al., 2011). With a small dataset from 80k relying more on expert knowledge from career counsellors or using universal priors could make the analysis more rigorous. Another way the Bayesian model accounts for uncertainty is by providing a posterior distribution. As distributions naturally incorporate uncertainty in the standard deviation this is a great way to perform exploratory data analysis. Yet where the Bayesian approach has the most potential is to explore complex relations in the data (Gal et al., 2022). Thus, for this initial examination of the 80k dataset t-test and correlation test provide an overview for further research into which aspects of high impact career guidance are most important to optimise.

Ethical Considerations and Social Impact

When providing career guidance ethical considerations and the social impact that such advice can have, is important to reflect on. Specifically, aiming to give high impact career counselling through an AI model such as ChatGPT has some important implications. Without thinking about the impact, one has on the world it can be hard to find the right fit and maintain a healthy work-life-balance. Adding the layer of pushing people in the direction of having an impact may lead them to deprioritise other important aspects of their life. Yet the present exploratory analysis indicates that high impact career counselling is an underprioritizes field of work with high demand. Given enough work and time the results also indicate that even more people would like to get pushed towards a more impactful career path, e.g., seen in relations between time spent on the website and number of visitors and financial spending and vacancy clicks. Thus, improving the ways in which high impact career guidance is available may have a positive social impact by utilising current data science methods such as ChatGPT.

Conclusion

This precursor for an API interface to ChatGPT aimed at providing high impact career guidance in accordance with the principles of 80.000hours.org looked at relations and differences in the open access dataset from a 2-year review of the organisation. Significant correlations were found for 2 of 4 correlation tests, specifically for FTE advising and number of one-on-one calls completed, and financial expenses and vacancy clicks on the website. The t-test was also significant examining the difference between number of applications for advising and completed advising session. The relations between FTE advising and average usefulness score and FTE web and unique visitors on the website were both insignificant. In conclusion it seems like invested time in various resources aimed at providing high impact career guidance have moderate payoff. The most important finding is that the field of high impact counselling has high demand and few resources. This could indicate that time invested in this field would result in a positive social impact if done with the proper time and work.

References

80,000 Hours Job Board. (n.d.). 80,000 Hours. Retrieved 2 June 2023, from https://jobs.80000hours.org

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4* (arXiv:2303.12712). arXiv. http://arxiv.org/abs/2303.12712

Cardoso, P. M., Taveira, M. C., Biscaia, C. S., & Santos, M. G. (2012). Psychologists' dilemmas in career counselling practice. *International Journal for Educational and Vocational Guidance*, *12*(3), 225–241. https://doi.org/10.1007/s10775-012-9232-9

CFI Team. (2023). *Full Time Equivalent (FTE)*. Corporate Finance Institute. https://corporatefinanceinstitute.com/resources/management/full-time-equivalent-fte/

Cretu, C. (2023, May 9). *How Does ChatGPT Actually Work? An ML Engineer Explains*. Scalable Path. https://www.scalablepath.com/data-science/chatgpt-architecture-explained

Cristea, I. A., & Ioannidis, J. P. A. (2018). P values in display items are ubiquitous and almost invariably significant: A survey of top science journals. *PLOS ONE*, *13*(5), e0197440. https://doi.org/10.1371/journal.pone.0197440

Gal, Y., Koumoutsakos, P., Lanusse, F., Louppe, G., & Papadimitriou, C. (2022). Bayesian uncertainty quantification for machine-learned models in physics. *Nature Reviews Physics*, *4*(9), Article 9. https://doi.org/10.1038/s42254-022-00498-4

Guo, B., & Yuan, Y. (2017). A comparative review of methods for comparing means using partially paired data. *Statistical Methods in Medical Research*, *26*(3), 1323–1340. https://doi.org/10.1177/0962280215577111

Hassani, H., & Silva, E. S. (2023). The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data and Cognitive Computing*, *7*(2), Article 2. https://doi.org/10.3390/bdcc7020062

Hubbard, R., & Lindsay, R. M. (2008). Why P values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, *18*, 69–88. https://doi.org/10.1177/0959354307086923

Jobindex. (n.d.). Jobindex. Retrieved 2 June 2023, from https://www.jobindex.dk/

Kang, C. (2023, May 16). OpenAI's Sam Altman Urges A.I. Regulation in Senate Hearing. *The New York Times*. https://www.nytimes.com/2023/05/16/technology/openai-altman-artificial-intelligence-regulation.html

Khurumova, V., & Pinto, J. C. (2023). Online and computer-assisted career guidance: Are students prepared for it? *Frontiers in Psychology*, *14*, 1117289. https://doi.org/10.3389/fpsyg.2023.1117289

MacAskill, W. (2015). *Doing Good Better: How Effective Altruism Can Help You Make a Difference*. Penguin.

Mathys, C., Daunizeau, J., Friston, K., & Stephan, K. (2011). A Bayesian Foundation for Individual Learning Under Uncertainty. *Frontiers in Human Neuroscience*, *5*. https://www.frontiersin.org/articles/10.3389/fnhum.2011.00039

Mhlanga, D. (2023). *Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning* (SSRN Scholarly Paper No. 4354422). https://doi.org/10.2139/ssrn.4354422

Milmo, D. (2023, February 2). ChatGPT reaches 100 million users two months after launch. *The Guardian*. https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app

Pause Giant AI Experiments: An Open Letter. (2023, May 22). *Future of Life Institute*. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[PUBLIC] 80,000 Hours two-year review: 2021-2022. (2023, May 8). Google Docs. https://docs.google.com/document/d/1mZmox0G910jw7I9jCoDC6vowz_PcrHRfFoW04G59XMA/edit?urp=gmail_link&usp=embed_facebook

[PUBLIC] Historical metrics by programme (Static 2022 version). (n.d.). Google Docs. Retrieved 24 May 2023, from https://docs.google.com/spreadsheets/d/1XVeiYoKjG-wQWdR3LGlXCVMCAKqgASx-aFoeiiaKSrM/edit?usp=embed_facebook

R Core Team. (2020). *R: A Language and Environment for Statistical Computing* (R version 4 . 0 . 3). R Foundation for Statistical Computing. https://www.R-project.org/

Start here. (n.d.). 80,000 Hours. Retrieved 2 June 2023, from https://80000hours.org/start-here/

Statistikbanken. (n.d.). Retrieved 2 June 2023, from https://statistikbanken.dk/statbank5a/default.asp?w=1922

Todd, B. (2016). *80,000 Hours: Find a Fulfilling Career that Does Good*. Centre for Effective Altruism.

Todd, B. (2023a, March 13). *What is social impact? A definition*. 80,000 Hours. https://80000hours.org/articles/what-is-social-impact-definition/

Todd, B. (2023b, May 23). *How much good can one person do? Here's what the evidence says.* 80,000 Hours. https://80000hours.org/career-guide/can-one-person-make-a-difference/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (arXiv:1706.03762). arXiv. http://arxiv.org/abs/1706.03762

What is effective altruism? | Effective Altruism. (n.d.). Retrieved 2 June 2023, from https://www.effectivealtruism.org/articles/introduction-to-effective-altruism

Wiblin, R. (2019). *How can you figure out which global problem is most pressing?* 80,000 Hours. https://80000hours.org/articles/problem-framework/

Appendix A

20 Considering I soon have a masters degree in cognitive science can you provide me with career guidance based on the 80.000hours.org framework

Certainly! Based on the 80,000 Hours framework, which aims to help individuals have a significant positive impact through their careers, here are some potential career paths that align with your background in cognitive science:

1. Researcher: Pursue a career in academia or industry research, focusing on areas such as cognitive psychology, neuroscience, human-computer interaction, or artificial intelligence. You could contribute to our understanding of cognition and develop innovative solutions to real-world problems.
2. User Experience (UX) Researcher: Apply your knowledge of cognitive science to enhance user experiences in various domains. You can work with companies to design intuitive interfaces, conduct user studies, and optimize products and services for better user engagement and satisfaction.
3. Data Scientist: With your background in cognitive science, you can analyze and interpret complex data sets to gain insights into human behavior, decision-making processes, and cognitive functioning. This can be valuable in areas like marketing, consumer research, or policy analysis.
4. Human Factors Specialist: Work on improving the interaction between humans and technology by considering cognitive and ergonomic factors. This field involves designing user-friendly interfaces, conducting usability testing, and optimizing human-machine systems to enhance performance, safety, and efficiency.
5. Behavioral Economist: Apply insights from cognitive science to the field of economics. Study how cognitive biases and heuristics influence decision-making and design interventions to nudge individuals towards more rational choices. This field has applications in policy-making, marketing, and organizational behavior.
6. Education Consultant: Use your knowledge of cognitive science to improve educational practices and learning outcomes. You can work with schools, universities, or educational technology companies to develop evidence-based teaching methods, design effective educational programs, or conduct research on learning processes.
7. Science Communication: Share your expertise in cognitive science through various mediums like writing, public speaking, or multimedia platforms. By translating complex concepts into accessible language, you can help bridge the gap between scientific research and the general public, promoting understanding and awareness.

Remember to consider the specific skills, values, and impact potential that resonate with you the most. Conducting informational interviews, internships, or volunteering in relevant fields can help you gain practical experience and further refine your career goals. Good luck on your career journey!

Regenerate response

Appendix B

T-test and Wilcoxon rank: number of applications vs calls

```
        Paired t-test

data:  df_calls_quar$nb_apps and df_calls_quar$nb_calls
t = 4.607, df = 21, p-value = 0.0001523
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 199.9369 528.9722
sample estimates:
mean difference
       364.4545

Warning: cannot compute exact p-value with ties
        Wilcoxon signed rank test with continuity correction

data:  df_calls_quar$nb_apps and df_calls_quar$nb_calls
V = 251, p-value = 5.654e-05
alternative hypothesis: true location shift is not equal to 0
```

Correlation: Total FTE web and number of visitors

```
        Pearson's product-moment correlation

data:  df_uni$total_fte_web and df_uni$visitors
t = 2.1132, df = 26, p-value = 0.04434
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.01140694 0.66145333
sample estimates:
      cor
0.3828546

Warning: Cannot compute exact p-value with ties
        Spearman's rank correlation rho

data:  df_uni$total_fte_web and df_uni$visitors
S = 3538.5, p-value = 0.8731
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
0.03161108
```

Correlation: Total FTE advising and number of one-on-one calls

```
        Pearson's product-moment correlation

data:  df_one$FTE_adv and df_one$calls_one
t = 6.1233, df = 18, p-value = 8.764e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5963845 0.9272260
sample estimates:
      cor
0.8219756

Warning: Cannot compute exact p-value with ties
        Spearman's rank correlation rho

data:  df_one$FTE_adv and df_one$calls_one
S = 315.33, p-value = 9.153e-05
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.7629126
```

Correlation: Total FTE advising and average usefulness rating

```
        Pearson's product-moment correlation

data:  df_us$FTE_adv and df_us$avg_us
t = -1.098, df = 6, p-value = 0.3143
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.8645242  0.4153512
sample estimates:
       cor
-0.4090432

Warning: Cannot compute exact p-value with ties
        Spearman's rank correlation rho

data:  df_us$FTE_adv and df_us$avg_us
S = 127.27, p-value = 0.1914
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.515161
```

Correlation: Financial expenses and vacancy clicks

```
        Pearson's product-moment correlation

data:  df_fin$fin_cost and df_fin$vac_clicks
t = 11.403, df = 3, p-value = 0.001447
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8329108 0.9992869
sample estimates:
      cor
0.9886607


        Spearman's rank correlation rho

data:  df_fin$fin_cost and df_fin$vac_clicks
S = 4.4409e-15, p-value = 0.01667
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
  1
```