Exploring the Causal Relationship Between Income and Job Satisfaction

Dhruv Goyal, 1005282203

19/06/2021

Abstract

This paper examines the unidirectional casual effect of Income on Job Satisfaction amongst working class Americans, diving into the personal, cultural and industrial significance that the their association holds. The 2020 Public Stack Overflow Survey data is used which contains 64,461 responses to 61 questions. The technique of Propensity Score Matching is implemented with the Nearest Neighbor approach which creates treatment and control groups to find the strength of the association between Income and Job Satisfaction by calculating the ATT. The paper proves a weak causal relationship between Income and Job Satisfaction and ends with possible caveats in the approach used. The results suggest that other external factors such as working environment, working cultural and social relationships in the office may be more important than determining the level of satisfaction an employed person associated with their job.

Introduction

Background, Context and Importance

Every so often a new study claims to have quantified the relationship between Income and Job Satisfaction[1]. However, the complexity of their association is deep rooted in their subjective nature, cultural and geographic divergences, nature of work performed at the job and external factors like tangible and intangible benefits attached with income and job satisfaction. Thereby, it is not only difficult to isolate the above two variables but studies that aim to analyze the link between income and job satisfaction often present contrasting views. A survey taken by 221,000 Glassdoor users which collected respondents annual salary and a rating from one to five of their level of satisfaction, reported that 51% of Glassdoor users earning more than $120,000 USD left either a four or a five star rating compared to 40% of those users making less than $30,000 every year[2], suggesting a positive relation between the two factors. Contrastingly, a meta-analysis published in the U.S. *Journal of Vocational Behavior* which reviewed 120 years of research experience to synthesize the findings from 92 quantitative studies suggests a statistically insignificant correlation factor of $r = 0.14$ between income and job satisfaction, signalling an independent relationship[3]. Therefore, to examine the 'effect' of income on job satisfaction, a study must adjust for variables that confound with income and job satisfaction so that far-reaching personal, cultural and industrial consequences can be identified and examined.

As stated, the importance of analyzing the relationship between income and job satisfaction has extensive personal, cultural and industrial significance. Job satisfaction could mean different things to different people - while some might enjoy the work they are involved with, the industrial community they are part of or take pride in the company they are working for, others might be incentivized and satisfied solely by the monetary benefits and aspects of a job. Some may be in content with their responsibilities and roles at a job, while others may have a strong desire to advance and enjoy the fiscal advantages of their hard work and thus strive for higher salaries. In totality, the degree to which an individual interrelates satisfaction and income speaks to their personal, cultural and socioeconomic beliefs. Therefore, it is safe to say subjectivity lingers around the two factors and influences studies that examine their relationship. On a more industrial level, investigating the relationship between income and job satisfaction has influential consequences as well. A common fact that interrelates businesses and success is that employees' satisfaction and morale is one of the key factors determining it. More motivated, interested, incentivized employees' are, more efficient, productive and profitable a business is. Managers usually prefer employees who associate a high level of satisfaction with their jobs, employees that love their jobs so much so to devote their private time to work activities, employees that are creative, committed, feel acknowledged and safe in their work environment and those who are prepared to cross any obstacle to realize their roles. Due to the subjective nature of actualizing job satisfaction outlined above, it becomes significant for businesses to understand the 'effect' of income on job satisfaction so that a safe, inclusive working environment can be established and salaries can be distributed in a cost-effective rational way that takes into account the 'effect' of it on job satisfaction, and, indirectly a businesses success and growth.

To tie the points above, the way an individual measures the 'effect' of income on satisfaction associated with their job gives an insight into their peronsal, cultural and socio-economic beliefs while the ability of a business to understand this 'effect' along with other factors that influence job satisfaction determines its prosperity. This outlines the importance of analyzing the tedious and complex relationship between income and job satisfaction on a broad level and justifies the importance it serves.

Terminology, Research Question and Hypothesis

The overarching topic of this paper is to analyze the association between income and job satisfaction. The definition of job satisfaction we follow is: *"Any combination of psychological, physiological, and environmental circumstances that cause a person to truthfully say that they are satisfied with a job."*[4]. Furthermore, we consider job satisfaction to be a binary variable. The definition of income we follow is : *all income collectively received by individuals. Income includes compensation from a number of sources, including salaries, wages, and bonuses received from employment or self-employment, dividends and distributions received from investments, rental receipts from real estate investments, and profit sharing from businesses*[5]. In addition, the paper blurs the boundary between income and economic class following the rationale that a person with a higher annual income belongs to a higher class in comparison to a person with a lower annual income who belongs to a lower class. More specifically, we consider three economic classes - low, medium and high. Clouding the boundary between income and economic class introduces possible caveats and assumptions, and they have been discussed towards the end of the paper.

The goal of this paper is to *determine the uni-directional causal effect of income on job satisfaction.* To shape our analysis, we use the 2020 Stack Overflow Developer Survey which contains 64,461 responses to 61 questions answered by users of Stack Overflow. Stack Overflow is a question and answer website for professional and enthusiast programmers. Under its open license, this survey data is free to share, adapt and create derivative works from. Although the survey has been completed by global users of Stack Overflow, most responses have been recorded from the United States. **Therefore, the research question of this paper is to examine *if income significantly determines job satisfaction amongst working class Americans*.** For our analysis we use the technique of *Propensity Score matching* based on *nearest neighbors* to create treatment and control groups and then calculate the ATT (average treatment effect among the treated) to determine the causal effect of income on job satisfaction. Propensity score matching (PSM) is a statistical matching technique that attempts to estimate the effect of a treatment, policy, or other intervention by accounting for the covariates that predict receiving the treatment[6]. ATT is the effect of the treatment on the outcome among those who received the treatment[7]. Furthermore, we include two treatment groups - high class or high income, medium class or medium income, and contrast each with a single control group - low class or low income. Propensity scores are used to identify the treatment and control groups in the two scenarios and then ATT is calculated for each.

Preliminary Hypothesis: *The causal effect of income on job satisfaction is strong and substantial* i.e. a person with a higher income is more satisfied than a person with a lower income. This also implies that a person with medium income or belonging to the middle class is more satisfied than a person with a low income or belonging to a lower class.

The rationale of our hypothesis follows from the assumption that a person earning a higher income is free of stressors like making rent, putting food on the family table, paying mortgage, saving enough for themselves or their family etc. The only reason they have a higher income is because of the job that they have, which, in most cases should imply, that they are happy, satisfied and thankful for their job and the economic benefits it brings with it. It is important to note that this a highly subjective assumption and the reader may/may not agree with it depending on their interpretation of the interrelation between income and job satisfaction.

This section is followed by the 'Data' Section which first describes and provides a birds eye view on the data and useful insights that highlight the importance of our analysis. It then outlines the data cleaning and management process. Next, the 'Methods' section introduces the statistical methods used, more specifically, logistic regression, propensity score matching and the rationale behind calculating ATT. After which, we discuss and share the outcomes of our analysis in the 'Results' sections. Lastly, this is followed by 'Conclusion', which discusses the implications of our results, biases in our study, caveats introduced by assumptions and wraps the paper by imparting a holistic significance of our study.

Data

The data used for the purposes of this paper comes from the *Public 2020 Stack Overflow Developer Survey Results*. It contains responses of 64461 respondents to 61 survey questions. Before making it public under its open license, the data has been cleaned to remove any personal identifying information to protect the privacy of the respondents. The survey was fielded from 5 Feburary 2020 to 28 Feburary 2020 and the median time of all of the qualified responses was 16.6 minutes. Respondents were recruited and targeted primarily through channels owned and controlled by Stack Overflow. The top 5 sources of respondents were onsite messaging, blog posts, email lists, Meta posts, banner ads, and social media posts. Due to the way respondents were targeted, those users who extensively used Stack Overflow were more likely to respond to the survey.

Note: A copy of the Public Stack Overflow survey has been attached in the appendix.

Data Cleaning

The main focus of this paper is to examine the causal inference between Income and Job Satisfaction using propensity scoring matching. As explained in the **Methods** section of this paper, this entails creating a creating a logistic regression model. This model is used to calculate the propensity scores and proceed with the matching approach. The propensity scores then dictate and identify the treatment group and the control group. To create a logistic regression model that is suitable for our approach, we need our model to predict a respondents income based on certain covariates/variables. Out of the 61 covariates/variables available to choose from in the Stack Overflow data, we identified *Age, Gender, Employment status , Education level* to explain a respondents Income. Along with Income and Job Satisfaction, the question that these variables correspond to in the survey have been outlined below:

- Income(CompTotal): What is your current total compensation (salary, bonuses, and perks, before taxes and deductions), in $? Please enter a whole number in the box below, without any punctuation. If you are paid hourly, please estimate an equivalent weekly, monthly, or yearly salary. If you prefer not to answer, please leave the box empty.
- Age(age): What is your age (in years)? If you prefer not to answer, you may leave this question blank.
- Gender(Gender): Which of the following describe you, if any? Please check all that apply. If you prefer not to answer, you may leave this question blank.
- Job Satisfaction(Job_Sat): How satisfied are you with your current job? (If you work multiple jobs, answer for the one you spend the most hours on.)

- Education Level(Ed_level): Which of the following best describes the highest level of formal education that you've completed?
- Employment(Employment) : Which of the following best describes your current employment status?

From the above, Income and Age are numerical variables, while Gender, Job Satisfaction, Education Level and Employment are categorical variables. The categories used to generate responses have been outlined below:

- Gender: Man, [Man;Non-binary, genderqueer, or gender non-conforming], [Non-binary, genderqueer, or gender non-conforming], Woman, [Woman;Man], [Woman;Man;Non-binary, genderqueer, or gender non-conforming], [Woman;Non-binary, genderqueer, or gender non-conforming]
- Job Satisfaction: Neither satisfied nor dissatisfied, Slightly dissatisfied, Very dissatisfied, Slightly satisfied, Very satisfied.
- Education Level: Associate degree (A.A., A.S., etc.), Bachelor's degree (B.A., B.S., B.Eng., etc.), I never completed any formal education, Master's degree (M.A., M.S., M.Eng., MBA, etc.), Other doctoral degree (Ph.D., Ed.D., etc.), Primary/elementary school, Professional degree (JD, MD, etc.), Secondary school (e.g. American high school, German Realschule or Gymnasium, etc.), Some college/university study without earning a degree.
- Employment: Employed full-time, Employed part-time, Independent contractor, freelancer, or self-employed.

To explain our rationale behind choosing the above variables to predict Income, consider the following: As 'Age' increases, the time spent on the job, experience and skills also increase which translates to an increase in Income. In the general case, a Ph.d graduate is more likely to score a higher paying job than a fresh graduate with a Bachelor's degree, this implies that as the level of education increases ('Education Level'), Income increases respectively. A person that is employed full-time is more likely to earn more than a person who is only employed part-time. There is evidence of gender inequality in Income[8], with the general stereotype that men are paid more than woman and thus we incorporate 'Gender' to explain the income of a respondent.

Proceeding with the data cleaning, we needed to align our data with the research goal and the research question. To do this, we filtered the Stack Overflow data to only contain respondents from the United States using the *filter()* function in the tidyverse library of *R 4.0*. This required filtering the data for the variate 'Country = United States'. This reduced our data to include a total of 12,469 responses. Next, we selected the above mentioned six variables - *Age, Gender, Job Satisfaction, Education Level, Employment, Income* using the *select()* function in the same *R* library. Lastly, for the purposes of this paper and to perform propensity score matching, we dropped those responses which were not recorded, skipped by a respondent or had been marked as "NA" by the publishers of the Stack Overflow data. This was done using the *drop_na()* function in R. A glimpse of the dataset after filtering and selecting the above mentioned varibales is shown below:

Table 1: Survey data after some cleaning

Age	Gender	Employment	EdLevel	JobSat	CompTotal
36	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Slightly dissatisfied	116000
27	Man	Employed full-time	Associate degree (A.A., A.S., etc.)	Slightly satisfied	66000
25	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Slightly dissatisfied	79000
32	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Very satisfied	105000
24	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Slightly dissatisfied	83400

Next, to use propensity score matching we needed to transform Job Satisfaction into a binary variable. The rationale and the reason behind this is explained in the Methods section of this paper. To do this we combined and labeled *Neither satisfied nor dissatisfied, Slightly dissatisfied, Very dissatisfied* levels of the Job

Satisfaction variable as *Dissatisfied* while *Slightly satisfied, Very satisfied* were labeled as *Satisfied*. The logic behind classifying a respondent who is 'Slightly dissatisfied',' Very dissatisfied' with their job as Dissatisfied is rather obvious, and, with it, is the logic behind classifying a respondent who is 'Slightly satisfied', 'Very satisfied' as Satisfied. We interpreted those respondents who are not sure of whether they are satisfied or not satisfied with their job and thus fall in the 'Neither satisfied nor dissatisfied' category as being Dissatisfied because of their skepticism behind answering the question in the survey. A respondent who is happy/satisfied with their job, would be sure of it, while the very fact that a respondent is skeptic about the level of their satisfaction signals that they are not satisfied. Consequences of caveats resulting from this classification are explained in the 'Conclusion'. To perform the above classifications, we used the *mutate()* function in the *tidyverse* library. The function *casewhen()* was used inside mutate to specify the classification desired. For example, to classify all respondents who where 'Very Satisfied' as 'Satisfied' we would use the query 'mutate(casewhen(Job_Sat == "Very Satisfied" ~ "Satisfied"))

Furthermore, the propensity score matching technique requires us to treat Income as a categorical variable. As touched upon in the introduction, we blur the boundary between Income and Class. We divide income into three categories - High, Medium and Low. Respondents with an income of less than or equal to $39,500 are in the low income/class category, respondents with an income of more than $118,000 are in the high income category and respondents who have an annual income between this bracket belong to the medium income category. These cutoffs have been determined by following *Pew Research Centre* which classifies Americans into socioeconomic classes based off their annual income [9]. *Pew Research Centre* is a 'nonpartisan fact tank' that 'conducts public opinion polling, demographic research, content analysis and other data-driven social science research'. To proceed with the above classification, we first filtered all respondents based on their Income to create 3 new subsets. So, we had a new subset containing all respondents with low income, a new subset containing all respondents with medium income and another subset containing all respondent with high income. This was done using the *filter()* function in *tidyverse*.

Additionally, to perform propensity score matching pertaining to our research goal, we needed to combine our cleaned data to create final two datasets each containing treatment and control groups. To do this, we simply combined and merged the subset containing high income respondents with the subset containing low income respondents, forming our first final cleaned subset, and then we merged the subset containing medium income respondents with the subset containing low income respondents to form our second final cleaned subset. Adding to the data cleaning process, in the two final datasets, high income respondents and medium income respondents were marked with '1' to signify that they belonged to the treated category while low income respondents where marked with '0' to signify that they belonged to the control category. The significance of creating a treated group and a control group is discussed in the 'Methods' section of this paper. Subsets were joined using the *rbind()* function and control and treated group were marked '0' and '1' usind a combination of *mutate()* and *casewhen()* as described before. As an example, consider 'mutate(Income.Class = case_when(CompTotal > 118000 ~ 1))' which transforms all high income respondents to be represented by 1.

A glimpse of the cleaned data is provided below:

Table 2: Final Cleaned Dataset 1 with High Income and Low Income

Age	Gender	Employment	EdLevel	Satisfied	Income.Class
28	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
30	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
36	Man	Independent contractor, freelancer, or self-employed	Associate degree (A.A., A.S., etc.)	Dissatisfied	1
30	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
42	Man	Employed full-time	Some college/university study without earning a degree	Satisfied	1

Table 3: Final Cleaned Dataset 2 With Medium Income and Low Income

Age	Gender	Employment	EdLevel	Satisfied	Income.Class
31	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
39	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Dissatisfied	1
32	Man	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Dissatisfied	1
35	Man	Employed full-time	Some college/university study without earning a degree	Dissatisfied	1
47	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1

Data analysis

To analyze the data, draw insights from it and realize the importance of our research, we first show a summary of the uncleaned data selecting all the necessary variables as outlined in our paper.

	Age	Gender	Employment	EdLevel	CompTotal	Satisfied
	Min. : 1.00	Length:7110	Length:7110	Length:7110	Min. :0.000e+00	Length:7110
	1st Qu.:27.00	Class :character	Class :character	Class :character	1st Qu.:8.000e+04	Class :character
	Median :32.00	Mode :character	Mode :character	Mode :character	Median :1.100e+05	Mode :character
	Mean :34.36	NA	NA	NA	Mean :8.236e+05	NA
	3rd Qu.:39.00	NA	NA	NA	3rd Qu.:1.450e+05	NA
	Max. :99.00	NA	NA	NA	Max. :5.000e+09	NA

After selecting the necessary variables, dropping the 'NA" values, we have the above summary statistics. The mean age of respondents is between 34-35 years. The mean annual Income of a respondent is $823,600.

The table groups the uncleaned data by satisfied and dissatisfied respondents. It shows the count for each grouping and statistics based on Income for each group:

Satisfied	Count	Average.Income	StdDEv.Income
Dissatisfied	2035	2575029.4	110835190.7
Satisfied	5075	121322.3	77101.1

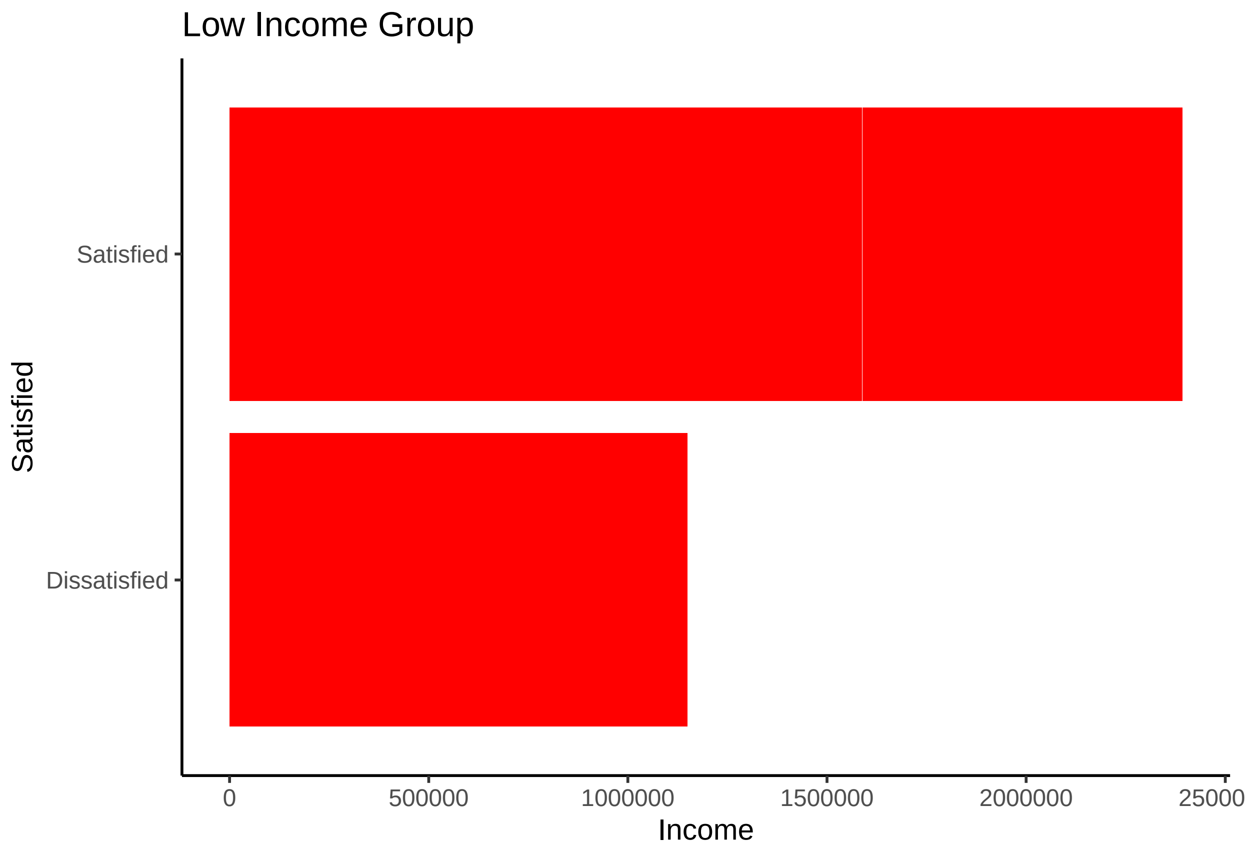
} As shown above, the number of satisfied respondents exceed the number of dissatisfied respondents. This provides an insight into corporate America, signalling that the number of employees happy and satisfied with their work exceed those that are dissatisfied. Furthermore, as is evident from the 'Average.Income' column, satisfied respondents have a lower average income while dissatisfied respondent have a higher average income. This is again a reflection to the working class individuals of America, were those that are working tiresomely, day-in and day-out, exhausting themselves at their higher paying jobs/positions tend to be dissatisfied more often than those who have a lower income, but tend to be more happy/satisfied at their jobs. Note: It is important to realize that we connect solely base our analysis on these statistics because of their nature of being derived from a survey without rather than a study that places control on confounding variables (Refer to the Methods Section for reasoning).

The table below provides the modes or the most frequently occurring values of categorical variables in our analysis:

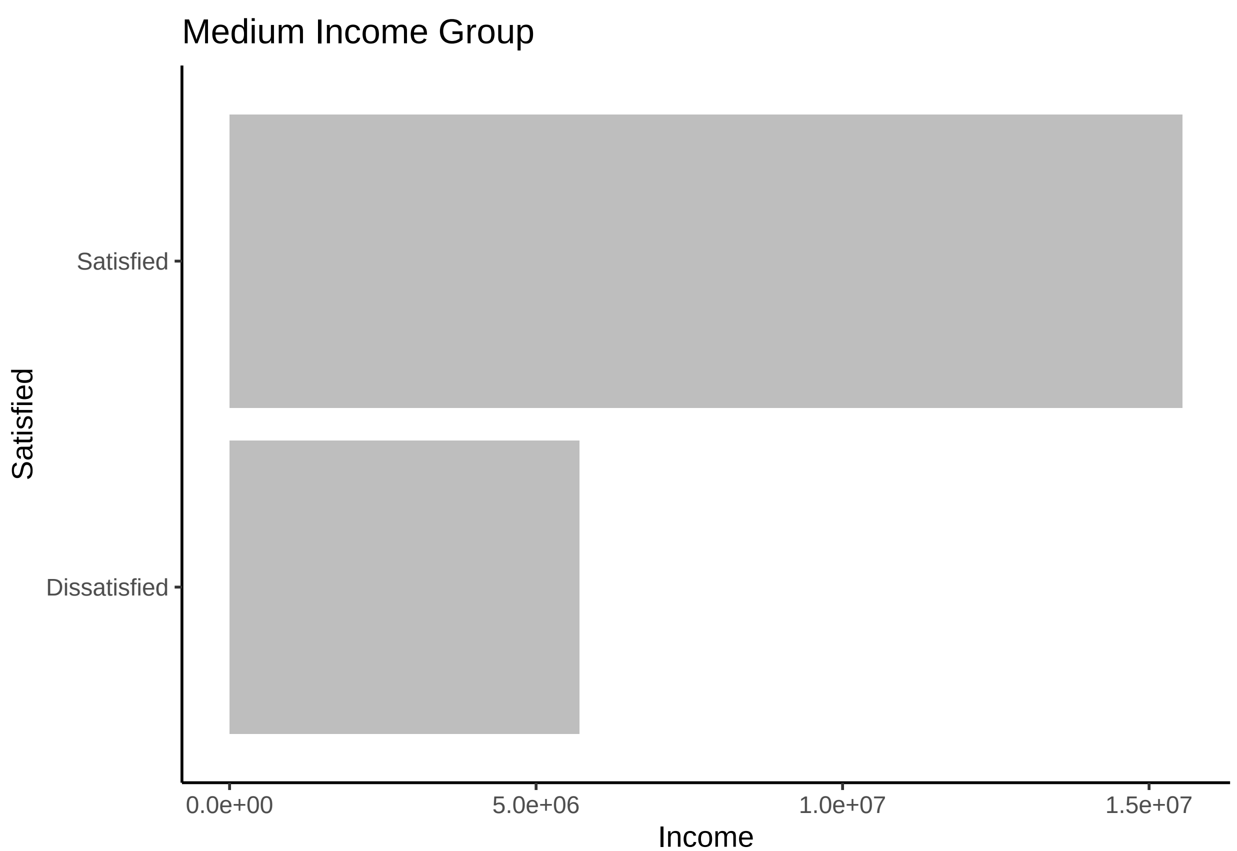
Covariate	Mode
Gender	Man
Employment	Employed full-time
Education Level	Bachelor's degree (B.A., B.S., B.Eng., etc.)

} As is evident from the above table, our survey is answered by those individuals who identify themselves as 'Man'. Most of the respondents are 'Employed full-time' and majority have a 'Bachelors degree (B.A., B.S., B.Eng., etc.)' degree. At first, this suggest that most users of Stack Overflow are full-time employed men with a bachelors degree, secondly, it suggests that our analysis and research is biased towards this category.

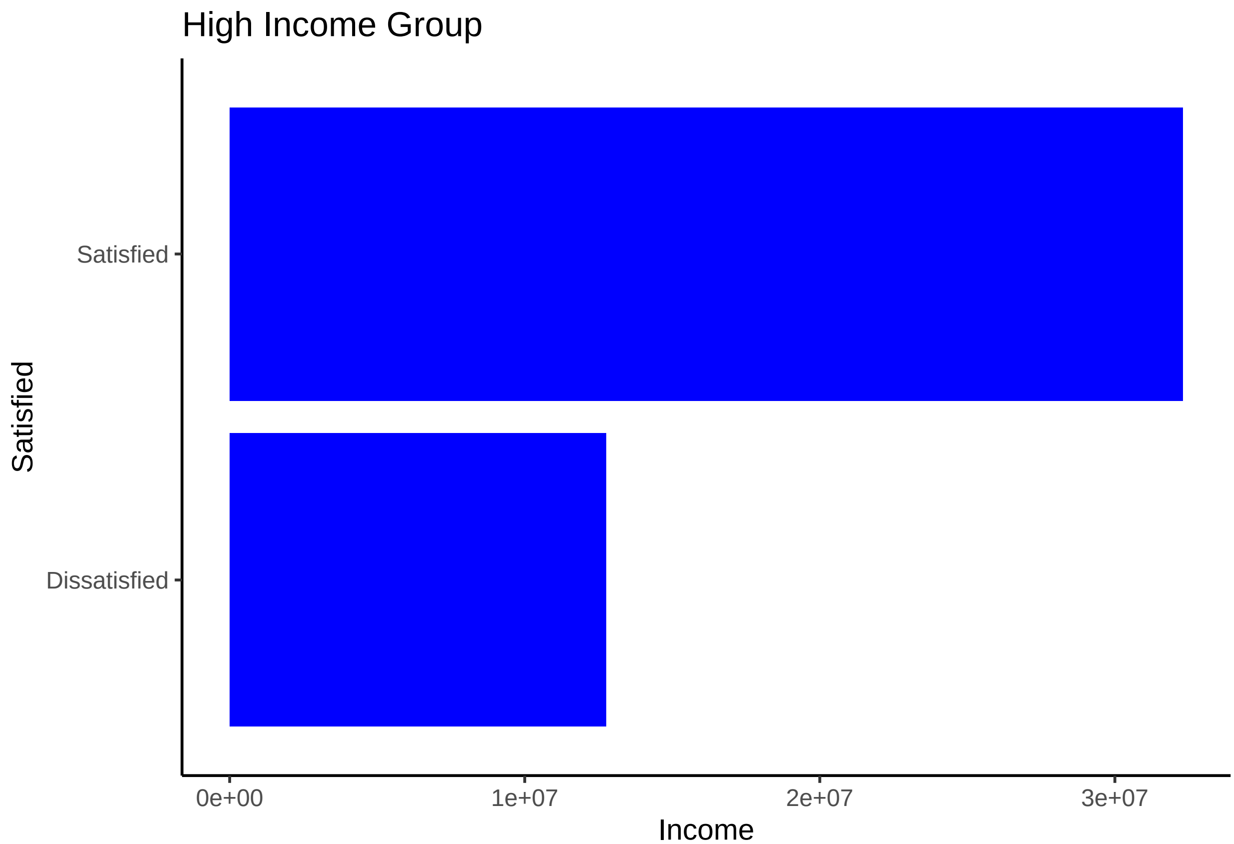
Next, we visualize Job Satisfaction based on low, medium and high income categories.



Low Income Group
Satisfied
Dissatisfied
Satisfied
0
500000
1000000
1500000
2000000
25000
Income



Medium Income Group
Satisfied
Dissatisfied
Satisfied
0.0e+00
5.0e+06
1.0e+07
1.5e+07
Income



The above column graphs shows the differing levels of Satisfaction among Low, Medium and High Income earners. The x-axis plots Income and the y-axis plots the binary Satisfaction variable. In each category, income of those individuals are higher who belong to the 'Satisfied' group.

The above statistics and plots make us ponder if actually Income is positively related with Satisfaction or if as the income of an individual increases, the level of their satisfaction increases as well. The above statistics and plots support this statement however, it would be too naive to conclude our analysis based on the above result. As witnessed the above data is biased to men who are employed full-time and have completed their Bachelors Degree. More importantly, due to the nature of this preliminary analysis being based on survey data, we cannot possibly determine if the difference in Satisfaction between the high income group(treatment group) and the low income group(control group) or the medium income group(treatment group) and the low income group is due to higher incomes (treatment) or the difference between respondents based on other characteristics like Age, Gender, Education, Employment etc. Therefore, to determine if there exist a causal relation between Income and Job Satisfaction we need to control for these 'other' variables/characteristics. This is where propensity score matching comes into play.

Methods

Drawing from our last idea, it may not be possible to determine the causal inference or the 'effect' of Income on Job Satisfaction solely based on survey data due to lack of randomization leading to a confounding bias, where a secondary variable other than Income and Job Satisfaction is associated with the two. Note that there can be more than one variable which leads to this confounding bias. This requires the use of propensity score matching:

Propensity Score Matching

Definitions

Treatment: It refers to any predictor in the survey context about which we wish to estimate a causal effect.

Confounder: A variable that predicts the treatment and the outcome and therefore may weaken our ability to make causal inferences about the effect.

Propensity Score: 'Statistical tool that allows researchers to make accurate causal inferences by balancing non-equivalent groups that may result from using a non-randomized design'[8]. It can be understood as an individuals probability to have received a treatment (e.g. belong to high income group) conditional on a host of potential confounding variables. This score can be used to adjust for confounding in an analysis so that plausible causal inferences can be made.

More formally, propensity scores can be defined as $\hat{\pi}_i = P(T_i = 1|X_i)$. Where T_i denotes an individual i who receives the treatment and X_i denotes the potential confounders.[8]

Propensity score matching

It refers to the pairing of treatment and control units with similar propensity scores, and possibly other covariates, and the discarding of all unmatched units. Propensity score matching involves four steps:

Step 1: Estimate propensity scores Propensity scores are typically estimated using a logistic regression model. This paper includes two logistic regression model. One model predicts the probability of an respondent belonging to a high income category or a low income category based on their Age, Gender, Employment, Education Level. Another model predicts the probability of whether a respondent belongs to the middle income category or a low income category based on the same variables.

- When estimating propensity scores, one should keep in mind that all confounders that are predictive of selection into the treatment groups or the outcome group should be included in the logistic regression model. Additionally, regardless of statistically significant differences between the treatment and the control group, variables that are theoretically related to selection and outcome should be included in the model. Lastly, those instrumental variables strongly related to the treatment but not the outcome, should also be included.

The logistic regression models we use in our analysis to calculate propensity scores are:

$log(\frac{p}{1-p}) = \beta_0 + \beta_{Age}x_1 + \beta_{Gender}x_2 + \beta_{emplt}x_3 + \beta_{Edlevel}x_4$

- p : probability of earning a high income.
- β_{Age} : change in log odds of p for every one unit increase in x_1 or age. -β_{Gender} : change in log odds of p for every one unit change in x_2 or gender.
- β_{emplt} : change in log odds of p for every one unit change in x_3 or employment level.
- $\beta_{Edlevel}$: change in log odds of p for every one unit change in x_2 or education level.

$log(\frac{t}{1-t}) = \beta_0 2 + \beta_{Age2}x_1 + \beta_{Gender2}x_2 + \beta_{emplt2}x_3 + \beta_{Edlevel2}x_4$

- t : probability of earning a medium income.
- β_{Age2} : change in log odds of t for every one unit increase in x_1 or age. -$\beta_{Gender2}$: change in log odds of t for every one unit change in x_2 or gender.
- β_{emplt2} : change in log odds of t for every one unit change in x_3 or employment level.
- $\beta_{Edlevel2}$: change in log odds of t for every one unit change in x_2 or education level.

Step 2

Matching or using propensity scores to adjust for confounding This involves matching one individual from the treatment group to one or more individuals in the control group. Matching can be done with or without replacement, although it is more common to match without replacement. To perform this, we use the nearest neighbor approach. Under nearest neighbor matching, assume that the individuals in the treatment group are randomly sorted. The first individual in the treatment group is then matched with another individual with the closest propensity score in the control group. Similarly, the next individual is taken from the treatment group and is matched with an individual in the control group having the closest propensity score estimate. This process continues until each individual in the treatment group is matched with an individual from the control group. Remaining individuals in the control group are discarded.

- Note that in our analysis, upon creating the treatment groups and the control group, we found that our treatment groups included more respondents, i.e. a higher number of respondents fell into high and medium income categories. More specifically, 3089 respondents belonged to the high income category and 3608 respondents belonged to the medium income category, while only 413 respondents fell into the low income category. As explained by step 2, to match individuals in the treatment group to the control group without replacement, we took a random sample of 250 respondents from the original 3468 and 3273 high and medium income respondents. This meant our datasets for high and medium income respondents was reduced to include 250 respondents, while low income respondents included 413 observations.

Note: Matching using the nearest neighbor approach is done using the 'matching()' function in the arm package.

Step 3

Evaluating the quality of matching. This step requires assessing the balance on the potential confounders across the treatment groups. The goal of this step is to determine whether difference between the treatment group and the control group remain on the confounders even after the data has been adjusted using propensity scores. However, this paper represents a naive analysis based on propensity score matching. This step has been skipped and the caveats resulting from this are discussed in the 'Conclusion'.

Step 4

Assessing the Causal effect of Income on Job Satisfaction. Once balance is accepted (in our case assumed) to be sufficient, the causal effect of Income on Job Satisfaction can be determined using the propensity score adjusted data to calculate ATT or the average treatment effect on the treated. The value of ATT allows us to answer our research question.

Explanation of ATT

Assume there is a set of treatments $T \in 0, 1$. For example in our analysis, $high\ income = 1$, $medium\ income = 1$ and $low\ income = 0$. For each respondent i there are corresponding potential outcomes: $Y_i(0)\ and\ Y_i(1)$, where Y corresponds to the potential outcome(e.g. job satisfaction). For the adjusted data using propensity scores, unit-level causal effect of the treatment is given by: $Y_i(0) - Y_i(1)$ i.e. it is the difference between the outcome being exposed to the treatment and not being exposed.[10]

Let $W_i = 1$ denote the set including all the respondent who received treatment in the adjusted propensity score data derived from Step 2. Then the ATT is given by

$ATT = \frac{1}{N_t} \sum_{i:W_1=1}^{N} (Y_i(0) - Y_i(1))$

Where, N_t denotes the number of respondents exposed to the treatment. N is the number of units in the adjusted propensity score data from step 2.[10]

Assumption of Propensity score matching

- Unconfoundedness Assumption: Selection of being exposed to the treatment or not should be solely based on observable characteristics. It should be assumed there is no selection bias from unobserved characteristics.
- Common Support: Observations with similar characteristics are present both in the treatment group and the control group.

Results

As mentioned in step 2 of the Methods section, to proceed with our matching approach without replacement. We randomly sampled 250 responses from high income and medium income control group. Glimpse of this data is presented here:

Age	Gender	Employment	EdLevel	Satisfied	Income.Class
25	Woman	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Satisfied	1
40	Man	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Dissatisfied	1
30	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
60	Man	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Satisfied	1
34	Woman	Employed full-time	Some college/university study without earning a degree	Satisfied	1

Age	Gender	Employment	EdLevel	Satisfied	Income.Class
26	Woman	Independent contractor, freelancer, or self-employed	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
22	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
38	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1
34	Man	Employed full-time	Associate degree (A.A., A.S., etc.)	Satisfied	1
24	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	1

After creating the logistic regression model, we predicted the propensity scores for both datasets. This follows step 1 of the 'Methods section'. A glimpse is shown here:

Age	Gender	Employment	EdLevel	Satisfied	Income.Class	propensityScore
60	Man	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Satisfied	1	0.8409087
25	Man	Employed full-time	Some college/university study without earning a degree	Dissatisfied	1	0.2881321
40	Woman	Employed full-time	Other doctoral degree (Ph.D., Ed.D., etc.)	Dissatisfied	1	0.7037792
34	Man	Independent contractor, freelancer, or self-employed	Some college/university study without earning a degree	Satisfied	0	0.0403695
27	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	0	0.4797543

Age	Gender	Employment	EdLevel	Satisfied	Income.Class	propensityScore
60	Man	Employed full-time	Master's degree (M.A., M.S., M.Eng., MBA, etc.)	Satisfied	1	0.8409087
25	Man	Employed full-time	Some college/university study without earning a degree	Dissatisfied	1	0.2881321
40	Woman	Employed full-time	Other doctoral degree (Ph.D., Ed.D., etc.)	Dissatisfied	1	0.7037792
34	Man	Independent contractor, freelancer, or self-employed	Some college/university study without earning a degree	Satisfied	0	0.0403695
27	Man	Employed full-time	Bachelor's degree (B.A., B.S., B.Eng., etc.)	Satisfied	0	0.4797543

Then we use the above adjusted data for propensity scores to create matches. For every person who has a high income(treated) we want the low income (untreated) person who was considered as similar to them (based on propensity score) as possible. Similarly for middle income earners and low income earners. The unmatched units/respondents were discarded. This follows Step 2 of the 'Methods' section.

Lastly we predicted the ATT. This follows step 4. A visualization of average treatment effects is provided:

Treatment : High Income, Control : Low Income, Outcome : Job Satisfaction.

Income.Class	Mean_Satisfied	var(Satisfied)
0	0.688	0.2155181
1	0.728	0.1988112

}

Treatment : Middle Income, Control : Low Income, Outcome : Job Satisfaction.

Income.Class	Mean_Satisfied	var(Satisfied)
0	0.716	0.2041606
1	0.668	0.2226667

}

The ATT values for the effect of income on job satisfaction for the high income as the control group and low income as the control group is 0.04 while, when we treat medium income as the treatment group the ATT value is -0.048 . This suggests that income has an 'effect' on job satisfaction when we consider high income earners and low income earners, i.e. as income transitions from low to high, the level of satisfaction increases. However, it is plausible to note that this effect is not strong and substantial as hypothesized, in fact we have a weak effect. This is evident from the low ATT value observed. Interestingly, the ATT value when medium income is the treatment group and low income is the control group is given by -0.048. This implies that there is an inverse effect and as the income level transitions from low income to medium income, the level of satisfaction fails to increase.

Conclusions

To tie up our research, we successfully created treatment and control groups using propensity score matching based on the nearest neighbor approach and evaluated the ATT to determine the causal effect of income on job satisfaction. We can say that the analysis presented in our paper provides weak evidence of the effect of income on job satisfaction. While this completes our research goal, it does not reflect our preliminary hypothesis. Specifically, there does not exist a strong and substantial effect of income on job satisfaction. This is further reflected in the low ATT values calculated where the income seems to effect job satisfaction when we consider a gap of $118,000 - $39,500 = $78,500, or almost 3 times the threshold for the low income category, for which the average treatment effect of the treated is 0.04. While, this does suggest that as income increases, level of satisfaction also increases, the ATT value of -0.048 for the analysis comparing the job satisfaction of medium income earners to low income earners suggest an opposite relationship. Specifically, that as income increases from low income to medium income, level of satisfaction does not increase. This suggests it is typically difficult to isolate the two variables and perform causal inference analysis on the two.

The analysis presented in this paper allow us to get an insight into the effect of income on job satisfaction. The reason for a negative ATT between medium income earners and low income earners, and, a positive ATT between high income and low income earners may be due to a number of reasons. Firstly, the difference between the threshold used to categorize a respondent as a high income earner is substantially more than the threshold used to categorize a respondent as a medium income earner. So, we would expect that as income increases three folds, one may be relatively more confident about their economic situation which allows them to focus on their job than those earning lower by three folds. These individuals would also be the ones for whom the economic tangible and intangible benefits associated with their job matter and is significantly associated with their job satisfaction level. Again, we trace back to the subjectivity factor that lingers and affects the strength of association between income and job satisfaction between individuals. Secondly, the overall weak evidence we found through our analysis may reflect the evolving factors that affect a persons level of satisfaction at their job. Rather than associating their level of satisfaction with their income, it is possible that people are more likely to associate it with their corporate working environment, their working culture, their relationship with their peers, whether or not the work that they do is meaningful and fulfilling to them etc. Therefore, our paper reflects that these external factors that people tend to associate with their job satisfaction, may be more influential than just income.

To conclude, this paper examines the unidirectional causal effect of Income on Job Satisfaction and concludes with evidence of a weak association. Propensity score matching and ATT calculations aid our analysis and allow us stufy this causal effect from a survey conducted by the online website Stack Overflow.

Limitations and Future Analysis

There exist limitations in our analysis due to the assumptions we have made. We present the drawbacks and the limitations here. We also provide suggestions for future analysis to circumvent the limitations of our paper.

- As stated in the data section, the stack overflow survey is only answered by those who are heavy users of the website rather than being distributed on a random platform. This means, we are particularly generalizing the incomes, ages, education levels, employment status, sex and job satisfaction levels of stack overflow website to the entire working American population. This makes our study biased and may be a reason for obtaining a weak effect of income on job satisfaction. A solution for this may be create a multi-level logistic regression model when predicting income with job satisfaction in the stack overflow dataset and then post-stratifying it to an unrepresentative census. After this process we can find the population level ATTs which may be more accurate in determining the effect of our variables in study.
- There are caveats of blurring the boundary between income and socio-economic class. As explained in the introduction, a person with a low income does not need to be strictly part of a lower class. It is possible that the person in question comes from a family of wealth and is not motivated enough to earn a high income and thus belong to a higher income class category in our analysis. A person like this may not typically associate income with their level of job satisfaction, while it would be more likely for them to associate their corporate working enviornment, their relationships with thier peers and their work culture more strongly with their level of satisfaction. This may also be a reason to the lower ATT we get when comparing medium income to low income groups.
- To perform propensity score matching, we had to transform the Job Satisfaction variable with 5 levels into a binary variable. Doing so, we classified all respondents who where neither satisfied nor dissatisfied as being dissatisfied which may be a reason for the weak association between income and job satisfaction we report. To adjust for this limitation there can possibly be two solutions. First, we can choose to not incorporate the respondents who choose to reply with a 'neither satisfied nor dissatisfied' option. However this is only possible when majority of the data for the job satisfaction column is binary (satisfied or not). Second, we can create a model based on all other covariates in our study to predict whether or not a person will be satisfied. This can be used when the number of binary responses are low.
- We present a naive approach for calculating propensity scores. We skip step 3 which requires assessing the balance on the potential confounders across the treatment groups. The goal of this step is to determine whether difference between the treatment group and the control group remain on the confounders even after the data has been adjusted using propensity scores. Since this step is particularly important, it is possible that although we perform matching based on propensity scores to form treatment and control groups, one of the 4 covariates is still a confounder which skews our results and our final calculation of ATT. For example, gender or age may affect the level of satisfaction a person associated with their jobs and this may not be completely controlled for after we adjust for propensity scores. For future analysis, a study can more accurately predict the ATT and thus determine the causal effect on job satisfaction with greater accuracy than this paper presents by incorporating this step into their analysis.

Bibliography

[1]. Greenwood, Beth. “Relationship Between Job Satisfaction & Salary.” Career Trend, 21 Nov. 2017, www.careertrend.com/relationship-between-job-satisfaction-salary-30445.html.

[2]. Nuñez, Mario. “Does Money Buy Happiness? The Link Between Salary and Employee Satisfaction.” Glassdoor Economic Research, 17 Oct. 2019, www.glassdoor.com/research/does-money-buy-happiness-the-link-between-salary-and-employee-satisfaction/#.

[3]. “Does Money Really Affect Motivation? A Review of the Research.” Harvard Business Review, 7 Aug. 2014, www.hbr.org/2013/04/does-money-really-affect-motiv.

[4] Bourne, Joshua PsychD. “What Is Job Satisfaction and Why Is It Important?” PositivePsychology.Com, 3 Dec. 2020, www.positivepsychology.com/job-satisfaction.

[5]. Kagan, Julia. “What Is Personal Income?” Investopedia, 22 Jan. 2021, www.investopedia.com/terms/p/personalincome.a

[6]. Wikipedia contributors. “Propensity Score Matching.” Wikipedia, 22 May 2021, www.wikipedia.org/wiki/Propensity_sco

[7]. Lanza, Stephanie T et al. “Drawing causal inferences using propensity scores: a practical guide for community psychologists.” American journal of community psychology vol. 52,3-4 (2013): 380-92. doi: 10.1007/s10464-013-9604-4.

[8]. Lanza, Stephanie T et al. “Drawing causal inferences using propensity scores: a practical guide for community psychologists.” American journal of community psychology vol. 52,3-4 (2013): 380-92. doi: 10.1007/s10464-013-9604-4

[9].Bennett, Jesse, et al. “Are You in the American Middle Class? Find out with Our Income Calculator.” Pew Research Center, 23 July 2020, www.pewresearch.org/fact-tank/2020/07/23/are-you-in-the-american-middle-class.

[10]. “Average Treatment Effects on the Treated (ATT) • Zelig.” Christine Choirat, Christopher Gandrud, James Honaker, Kosuke Imai, Gary King, Olivia Lau., 29 Oct. 2017, docs.zeligproject.org/articles/att.html.

Appendix

A glimpse of the original Public Stack Overflow Dataset:

```
##   Respondent                                               MainBranch Hobbyist
## 1          1                           I am a developer by profession      Yes
## 2          2                           I am a developer by profession       No
## 3          3                              I code primarily as a hobby      Yes
## 4          4                           I am a developer by profession      Yes
## 5          5 I used to be a developer by profession, but no longer am      Yes
## 6          6                           I am a developer by profession       No
##   Age Age1stCode CompFreq CompTotal ConvertedComp            Country
## 1  NA         13  Monthly        NA            NA            Germany
## 2  NA         19     <NA>        NA            NA     United Kingdom
## 3  NA         15     <NA>        NA            NA Russian Federation
## 4  25         18     <NA>        NA            NA            Albania
## 5  31         16     <NA>        NA            NA      United States
## 6  NA         14     <NA>        NA            NA            Germany
##     CurrencyDesc CurrencySymbol DatabaseDesireNextYear
## 1  European Euro            EUR   Microsoft SQL Server
## 2 Pound sterling            GBP                   <NA>
## 3           <NA>           <NA>                   <NA>
## 4   Albanian lek            ALL                   <NA>
## 5           <NA>           <NA>       MySQL;PostgreSQL
## 6  European Euro            EUR                   <NA>
##                        DatabaseWorkedWith
## 1 Elasticsearch;Microsoft SQL Server;Oracle
## 2                                    <NA>
## 3                                    <NA>
## 4                                    <NA>
## 5            MySQL;PostgreSQL;Redis;SQLite
## 6                                    <NA>
##                                                              DevType
## 1 Developer, desktop or enterprise applications;Developer, full-stack
## 2                              Developer, full-stack;Developer, mobile
## 3                                                                 <NA>
## 4                                                                 <NA>
## 5                                                                 <NA>
## 6                     Designer;Developer, front-end;Developer, mobile
##                                                                             EdLevel
## 1                                   Master's degree (M.A., M.S., M.Eng., MBA, etc.)
## 2                                      Bachelor's degree (B.A., B.S., B.Eng., etc.)
## 3                                                                              <NA>
## 4                                   Master's degree (M.A., M.S., M.Eng., MBA, etc.)
## 5                                      Bachelor's degree (B.A., B.S., B.Eng., etc.)
## 6 Secondary school (e.g. American high school, German Realschule or Gymnasium, etc.)
##                                               Employment
## 1 Independent contractor, freelancer, or self-employed
## 2                                   Employed full-time
## 3                                                 <NA>
## 4                                                 <NA>
## 5                                   Employed full-time
## 6                                   Employed full-time
##                      Ethnicity Gender
```

```
## 1 White or of European descent    Man
## 2                         <NA>   <NA>
## 3                         <NA>   <NA>
## 4 White or of European descent    Man
## 5 White or of European descent    Man
## 6 White or of European descent    Man
##
## 1                Languages, frameworks, and other technologies I'd be working with;Remote work optio
## 2
## 3
## 4                               Flex time or a flexible schedule;Office environment or company cultu
## 5
## 6 Diversity of the company or organization;Languages, frameworks, and other technologies I'd be worki
##                  JobSat
## 1    Slightly satisfied
## 2     Very dissatisfied
## 3                  <NA>
## 4 Slightly dissatisfied
## 5                  <NA>
## 6    Slightly satisfied
##                                                         JobSeek
## 1                I am not interested in new job opportunities
## 2                I am not interested in new job opportunities
## 3                                                          <NA>
## 4 I'm not actively looking, but I am open to new opportunities
## 5                                                          <NA>
## 6                I am not interested in new job opportunities
##     LanguageDesireNextYear       LanguageWorkedWith
## 1   C#;HTML/CSS;JavaScript   C#;HTML/CSS;JavaScript
## 2             Python;Swift         JavaScript;Swift
## 3 Objective-C;Python;Swift Objective-C;Python;Swift
## 4                     <NA>                     <NA>
## 5          Java;Ruby;Scala        HTML/CSS;Ruby;SQL
## 6 HTML/CSS;Java;JavaScript HTML/CSS;Java;JavaScript
##            MiscTechDesireNextYear MiscTechWorkedWith
## 1                .NET Core;Xamarin     .NET;.NET Core
## 2 React Native;TensorFlow;Unity 3D       React Native
## 3                             <NA>               <NA>
## 4                             <NA>               <NA>
## 5                     Ansible;Chef            Ansible
## 6                             <NA>               <NA>
##            NEWCollabToolsDesireNextYear
## 1 Microsoft Teams;Microsoft Azure;Trello
## 2                          Github;Slack
## 3                                  <NA>
## 4                                  <NA>
## 5  Github;Google Suite (Docs, Meet, etc)
## 6                          Github;Slack
##                                      NEWCollabToolsWorkedWith NEWDevOps
## 1               Confluence;Jira;Slack;Microsoft Azure;Trello        No
## 2                        Confluence;Jira;Github;Gitlab;Slack      <NA>
## 3                                                       <NA>      <NA>
## 4                                                       <NA>        No
## 5 Confluence;Jira;Github;Slack;Google Suite (Docs, Meet, etc)      <NA>
```

```
## 6                             Confluence;Github;Slack;Trello  Not sure
##       NEWDevOpsImpt                         NEWEdImpt
## 1 Somewhat important                  Fairly important
## 2              <NA>                   Fairly important
## 3              <NA>                               <NA>
## 4              <NA> Not at all important/not necessary
## 5              <NA>                     Very important
## 6              <NA>                   Fairly important
##                                                             NEWJobHunt
## 1                                                                 <NA>
## 2                                                                 <NA>
## 3                                                                 <NA>
## 4 Curious about other opportunities;Wanting to work with new technologies
## 5                                                                 <NA>
## 6                                                                 <NA>
##   NEWJobHuntResearch      NEWLearn NEWOffTopic NEWOnboardGood NEWOtherComms
## 1               <NA>   Once a year    Not sure           <NA>            No
## 2               <NA>   Once a year    Not sure           <NA>            No
## 3               <NA> Once a decade        <NA>           <NA>            No
## 4               <NA>   Once a year    Not sure            Yes           Yes
## 5               <NA>   Once a year          No           <NA>           Yes
## 6               <NA>   Once a year          No             No            No
##                                               NEWOvertime
## 1                        Often: 1-2 days per week or more
## 2                                                    <NA>
## 3                                                    <NA>
## 4 Occasionally: 1-2 days per quarter but less than monthly
## 5                                                    <NA>
## 6                                                   Never
##
## 1
## 2
## 3
## 4
## 5 Start a free trial;Ask developers I know/work with;Visit developer communities like Stack Overflow
## 6
##      NEWPurpleLink
## 1           Amused
## 2           Amused
## 3             <NA>
## 4             <NA>
## 5 Hello, old friend
## 6           Amused
##
## 1
## 2
## 3
## 4
## 5 Stack Overflow (public Q&A for anyone who codes);Stack Exchange (public Q&A for a variety of topic
## 6
##
## 1                                                               Visit Stack Overflow;Go for a walk o
## 2                                                                                                 Vi
## 3
```

```
## 4
## 5 Call a coworker or friend;Visit Stack Overflow;Watch help / tutorial videos;Do other work and come
## 6                                                                                Play games;Visit Stack Overflow;Wa
##        OpSys                  OrgSize
## 1    Windows         2 to 9 employees
## 2      MacOS 1,000 to 4,999 employees
## 3 Linux-based                    <NA>
## 4 Linux-based       20 to 99 employees
## 5    Windows                     <NA>
## 6    Windows                     <NA>
##                          PlatformDesireNextYear
## 1   Android;iOS;Kubernetes;Microsoft Azure;Windows
## 2                     iOS;Kubernetes;Linux;MacOS
## 3                                           <NA>
## 4                                           <NA>
## 5 Docker;Google Cloud Platform;Heroku;Linux;Windows
## 6                                        Android
##              PlatformWorkedWith                   PurchaseWhat
## 1                       Windows                           <NA>
## 2                           iOS    I have little or no influence
## 3                          <NA>                           <NA>
## 4                          <NA> I have a great deal of influence
## 5 AWS;Docker;Linux;MacOS;Windows                           <NA>
## 6      Android;Docker;WordPress            I have some influence
##               Sexuality SOAccount          SOComm
## 1 Straight / Heterosexual        No  No, not at all
## 2                  <NA>       Yes Yes, definitely
## 3                  <NA>       Yes   Yes, somewhat
## 4 Straight / Heterosexual       Yes Yes, definitely
## 5 Straight / Heterosexual       Yes   Yes, somewhat
## 6 Straight / Heterosexual       Yes   Yes, somewhat
##                          SOPartFreq                     SOVisitFreq
## 1                              <NA>          Multiple times per day
## 2 Less than once per month or monthly          Multiple times per day
## 3     A few times per month or weekly           Daily or almost daily
## 4     A few times per month or weekly          Multiple times per day
## 5 Less than once per month or monthly A few times per month or weekly
## 6     A few times per month or weekly            A few times per week
##                   SurveyEase          SurveyLength Trans
## 1 Neither easy nor difficult Appropriate in length    No
## 2                       <NA>                  <NA>  <NA>
## 3 Neither easy nor difficult Appropriate in length  <NA>
## 4                       <NA>                  <NA>    No
## 5                       Easy             Too short    No
## 6 Neither easy nor difficult Appropriate in length  <NA>
##                                                  UndergradMajor
## 1 Computer science, computer engineering, or software engineering
## 2 Computer science, computer engineering, or software engineering
## 3                                                           <NA>
## 4 Computer science, computer engineering, or software engineering
## 5 Computer science, computer engineering, or software engineering
## 6                                                           <NA>
##   WebframeDesireNextYear   WebframeWorkedWith
## 1           ASP.NET Core ASP.NET;ASP.NET Core
```

```
## 2                     <NA>                    <NA>
## 3                     <NA>                    <NA>
## 4                     <NA>                    <NA>
## 5   Django;Ruby on Rails        Ruby on Rails
## 6               React.js                    <NA>
##                                   WelcomeChange WorkWeekHrs YearsCode YearsCodePro
## 1  Just as welcome now as I felt last year          50        36           27
## 2 Somewhat more welcome now than last year          NA         7            4
## 3 Somewhat more welcome now than last year          NA         4         <NA>
## 4 Somewhat less welcome now than last year          40         7            4
## 5  Just as welcome now as I felt last year          NA        15            8
## 6                                      <NA>          NA         6            4
```

The schema of the Dataset:

Column	QuestionText
Respondent	Randomized respondent ID number (not in order of survey response time)
MainBranch	Which of the following options best describes you today? Here, by "developer" we mean "someone who writes code."
Hobbyist	Do you code as a hobby?
Age	What is your age (in years)? If you prefer not to answer, you may leave this question blank.
Age1stCode	At what age did you write your first line of code or program? (e.g., webpage, Hello World, Scratch project)
CompFreq	Is that compensation weekly, monthly, or yearly?
CompTotal	What is your current total compensation (salary, bonuses, and perks, before taxes and deductions), in ‘CurrencySymbol‘? Please enter a whole number in the box below, without any punctuation. If you are paid hourly, please estimate an equivalent weekly, monthly, or yearly salary. If you prefer not to answer, please leave the box empty.
ConvertedComp	Salary converted to annual USD salaries using the exchange rate on 2020-02-19, assuming 12 working months and 50 working weeks.
Country	Where do you live?
CurrencyDesc	Which currency do you use day-to-day? If your answer is complicated, please pick the one you're most comfortable estimating in.
CurrencySymbol	Which currency do you use day-to-day? If your answer is complicated, please pick the one you're most comfortable estimating in.
DatabaseDesireNextYear	Which database environments have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the database and want to continue to do so, please check both boxes in that row.)
DatabaseWorkedWith	Which database environments have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the database and want to continue to do so, please check both boxes in that row.)
DevType	Which of the following describe you? Please select all that apply.
EdLevel	Which of the following best describes the highest level of formal education that you've completed?
Employment	Which of the following best describes your current employment status?
Ethnicity	Which of the following describe you, if any? Please check all that apply. If you prefer not to answer, you may leave this question blank.
Gender	Which of the following describe you, if any? Please check all that apply. If you prefer not to answer, you may leave this question blank.
JobFactors	Imagine that you are deciding between two job offers with the same compensation, benefits, and location. Of the following factors, which 3 are MOST important to you?
JobSat	How satisfied are you with your current job? (If you work multiple jobs, answer for the one you spend the most hours on.)
JobSeek	Which of the following best describes your current job-seeking status?
LanguageDesireNextYear	Which programming, scripting, and markup languages have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the language and want to continue to do so, please check both boxes in that row.)
LanguageWorkedWith	Which programming, scripting, and markup languages have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the language and want to continue to do so, please check both boxes in that row.)
MiscTechDesireNextYear	Which other frameworks, libraries, and tools have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the framework and want to continue to do so, please check both boxes in that row.)
MiscTechWorkedWith	Which other frameworks, libraries, and tools have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the framework and want to continue to do so, please check both boxes in that row.)
NEWCollabToolsDesireNextYear	Which collaboration tools have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you worked with the tool and want to continue to do so, please check both boxes in that row.)
NEWCollabToolsWorkedWith	Which collaboration tools have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you worked with the tool and want to continue to do so, please check both boxes in that row.)
NEWDevOps	Does your company have a dedicated DevOps person?
NEWDevOpsImpt	How important is the practice of DevOps to scaling software development?
NEWEdImpt	How important is a formal education, such as a university degree in computer science, to your career?
NEWJobHunt	In general, what drives you to look for a new job? Select all that apply.
NEWJobHuntResearch	When job searching, how do you learn more about a company? Select all that apply.
NEWLearn	How frequently do you learn a new language or framework?
NEWOffTopic	Do you think Stack Overflow should relax restrictions on what is considered "off-topic"?
NEWOnboardGood	Do you think your company has a good onboarding process? (By onboarding, we mean the structured process of getting you settled in to your new role at a company)
NEWOtherComms	Are you a member of any other online developer communities?
NEWOvertime	How often do you work overtime or beyond the formal time expectation of your job?
NEWPurchaseResearch	When buying a new tool or software, how do you discover and research available solutions? Select all that apply.
NEWPurpleLink	You search for a coding solution online and the first result link is purple because you already visited it. How do you feel?
NEWSOSites	Which of the following Stack Overflow sites have you visited? Select all that apply.
NEWStuck	What do you do when you get stuck on a problem? Select all that apply.
OpSys	What is the primary operating system in which you work?
OrgSize	Approximately how many people are employed by the company or organization you currently work for?
PlatformDesireNextYear	Which platforms have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the platform and want to continue to do so, please check both boxes in that row.)
PlatformWorkedWith	Which platforms have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the platform and want to continue to do so, please check both boxes in that row.)
PurchaseWhat	What level of influence do you, personally, have over new technology purchases at your organization?
Sexuality	Which of the following describe you, if any? Please check all that apply. If you prefer not to answer, you may leave this question blank.
SOAccount	Do you have a Stack Overflow account?
SOComm	Do you consider yourself a member of the Stack Overflow community?
SOPartFreq	How frequently would you say you participate in Q&A on Stack Overflow? By participate we mean ask, answer, vote for, or comment on questions.
SOVisitFreq	How frequently would you say you visit Stack Overflow?
SurveyEase	How easy or difficult was this survey to complete?
SurveyLength	How do you feel about the length of the survey this year?
Trans	Are you transgender?
UndergradMajor	What was your primary field of study?
WebframeDesireNextYear	Which web frameworks have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the framework and want to continue to do so, please check both boxes in that row.)
WebframeWorkedWith	Which web frameworks have you done extensive development work in over the past year, and which do you want to work in over the next year? (If you both worked with the framework and want to continue to do so, please check both boxes in that row.)
WelcomeChange	Compared to last year, how welcome do you feel on Stack Overflow?
WorkWeekHrs	On average, how many hours per week do you work? Please enter a whole number in the box.
YearsCode	Including any education, how many years have you been coding in total?
YearsCodePro	NOT including education, how many years have you coded professionally (as a part of your work)?