CODEBOOK

IL181 Gender inequality in online marketplaces

Project description:

This project seeks to investigate whether women have to undersell themselves in online job marketplaces to compete against men. We focus on the online job platform Freelancer.com to gather talent listings of individuals offering their services and compare the hourly rate offered by men vs. women, controlling for factors like job function, experience, job completion rate, etc.

Codebook description:

This codebook is meant for the "original" version of the dataset before any matching-specific preprocessing occurred. The data file corresponding to this version of the codebook is "cleaned-gender-annotated-v5.csv".

VARIABLES

Name	Column label	Description	Format	Collection method	Missing values
Search query	search_query	The search query that was used to retrieve the user profile. For example, "software engineer" or "copywriter".	String	User inputted	No
Person name	name	The name of the person. Only the first name is provided. We will use this to determine the gender of the person.	String	Scraped	No
Profile link	profile_link	The Freelancer.com link to the user's profile.	URL	Scraped	No
Profile tagline	tagline	The self-described tagline of the user on their profile. For example, "Experienced CPA".	String	Scraped	No
Profile description	user_description	The self-written description of the user on their profile. They often describe the services offered, their level of experience, and other things here. It is freeform text	Freeform text	Scraped	No
Profile location	location	The location of the user. It is usually in the format "City, Country".	String	Scraped	No
Join date	join_date_from_earliest	The date the user joined the Freelancer.com platform, represented as the number of days after the earliest date in the dataset. For example, a value of "10" indicates the user joined 10 days after the earliest join date in the dataset. The earliest join date has a value of 0.	Number	Scraped	No

Hourly rate	hourly_rate	The hourly rate charged by the user to perform their services. All rates are in USD.	Number	Scraped	No
Pay grade	pay_grade	The amount of money earned by the user doing projects in the skill/category they belong in. It is a number from 0 to 10, and increases the more money the freelancer earns in the category.	Number	Scraped	No
Average rating	avg_rating	The user's average rating received by the people they have worked for. It is a number from 0 to 5. The higher the number, the better the rating. Some users may have a rating of 0, which means they have not been reviewed yet.	Number	Scraped	No
Number of reviews	num_reviews	The number of reviews that the user has received.	Number	Scraped	No
Number of recommendations	num_recommendations	The number of recommendations that have been written on the user's profile by people that have contracted them before.	Number	Scraped	No
Percentage of jobs completed	pct_jobs_completed	The percentage of jobs the user has completed successfully. It is a number from 0 to 100.	Number	Scraped	Yes
Percentage of jobs on budget	pct_on_budget	The percentage of jobs performed by the user that have stayed within the agreed-upon budget. It is a number from 0 to 100.	Number	Scraped	Yes
Percentage on time	pct_on_time	The percentage of jobs completed on time by the user. It is a number between 0 and 100.	Number	Scraped	Yes
Repeat hire rate	repeat_hire_rate	The rate at which this user has been hired again by individuals who have contracted with them. The number is a percentage from 0 to 100.	Number	Scraped	Yes
Preferred freelancer	verification_preferred_freelancer	Whether the user has received the "preferred freelancer" verification from Freelancer.com. This is described by the website as an "elite group" of the best freelancers.	Boolean	Scraped	No
Identity verified	verification_identity_verified	Whether the user has verified their identity on the platform. This consists of a government-issued ID check and a proof of address.	Boolean	Scraped	No
Payment verified	verification_payment_verified	Whether the user has been verified to have a valid payment method	Boolean	Scraped	No
Phone verified	verification_phone_verified	Whether the user has been verified to have a valid phone number	Boolean	Scraped	No
Email verified	verification_email_verified	Whether the user has been verified to have a valid email address	Boolean	Scraped	No
Facebook connected	verification_facebook_connected	Whether the user has been verified to have a Facebook account	Boolean	Scraped	No

Skills	skill_[skill_name]	Indicates whether the user has listed this [skill_name] as a skill on their profile. The value of the column indicates the amount of projects completed on the platform that have been related to this skill. "0" indicates no jobs related to the skill have been completed yet, and N/A indicates the skill is not listed on this user's profile.	Number	Scraped	Yes
Certifications	pct_certification_[certification_name]	Indicates whether the user has listed the score obtained (as a percentage) on [certification_name] on their profile. These certifications range from accounting exams to language proficiency tests. N/A indicates the certification is not listed on their profile.	Number	Scraped	Yes
Badges	badge_[badge_name]	Some users have badges they have earned on the platform listed next to their username on their profile. These indicate whether they have been verified, have some kind of professional membership, amongst others. N/A indicates the user does not have the badge listed on their profile.	Boolean	Scraped	Yes