**Data Set:**

**Description**

Cross-section data about resume, call-back and employer information for 4,870 fictitious resumes.

**Usage**

data("ResumeNames")

**Format**

A data frame containing 4,870 observations on 27 variables.

name

factor indicating applicant's first name.

gender

factor indicating gender.

ethnicity

factor indicating ethnicity (i.e., Caucasian-sounding vs. African-American sounding first name).

quality

factor indicating quality of resume.

call

factor. Was the applicant called back?

city

factor indicating city: Boston or Chicago.

jobs

number of jobs listed on resume.

experience

number of years of work experience on the resume.

honors

factor. Did the resume mention some honors?

volunteer

factor. Did the resume mention some volunteering experience?

military

factor. Does the applicant have military experience?

holes

factor. Does the resume have some employment holes?

school

factor. Does the resume mention some work experience while at school?

email

factor. Was the e-mail address on the applicant's resume?

computer

factor. Does the resume mention some computer skills?

special

factor. Does the resume mention some special skills?

college

factor. Does the applicant have a college degree or more?

minimum

factor indicating minimum experience requirement of the employer.

equal

factor. Is the employer EOE (equal opportunity employment)?

wanted

factor indicating type of position wanted by employer.

requirements

factor. Does the ad mention some requirement for the job?

reqexp

factor. Does the ad mention some experience requirement?

reqcomm

factor. Does the ad mention some communication skills requirement?

reqeduc

factor. Does the ad mention some educational requirement?

reqcomp

factor. Does the ad mention some computer skills requirement?

reqorg

factor. Does the ad mention some organizational skills requirement?

industry

factor indicating type of employer industry.

**Details**

Cross-section data about resume, call-back and employer information for 4,870 fictitious resumes sent in response to employment advertisements in Chicago and Boston in 2001, in a **randomized controlled experiment** conducted by Bertrand and Mullainathan (2004).

The resumes contained information concerning the ethnicity of the applicant. Because ethnicity is not typically included on a resume, resumes were differentiated on the basis of so-called “Caucasian sounding names” (such as Emily Walsh or Gregory Baker) and “African American sounding names” (such as Lakisha Washington or Jamal Jones). **A large collection of fictitious resumes were created and the pre-supposed ethnicity (based on the sound of the name) was randomly assigned to each resume**. These resumes were sent to prospective employers to see which resumes generated a phone call from the prospective employer.

**Source**

Online complements to Stock and Watson (2007).

**References**

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, **94**, 991–1013.

Stock, J.H. and Watson, M.W. (2007). *Introduction to Econometrics*, 2nd ed. Boston: Addison Wesley.

**See Also**

StockWatson2007

**Examples**

data("ResumeNames")

summary(ResumeNames)

prop.table(xtabs(~ ethnicity + call, data = ResumeNames), 1)

One of the most common—and one of the trickiest—challenges in data analysis is deciding how to include multiple predictors in a model, especially when they’re related to each other.

Here’s an example. Let’s say you are interested in studying the relationship between work spillover into personal time as a predictor of job burnout.

You have 5 categorical yes/no variables that indicate whether a particular symptom of work spillover is present

A picture containing text

Description automatically generated

**Fig:** The blue numbers in each column are the second type of parameters, equivalent to [factor loadings](https://www.theanalysisfactor.com/principal-component-analysis-negative-loadings/) in [confirmatory factor analysis](https://www.theanalysisfactor.com/structural-equation-modeling-first-step-confirmatory-factor-analysis-2/). Each is the conditional probability that someone in a particular class would respond ‘yes’ to a certain item. These parameters are used to interpret the classes.

While you could use each individual variable**, you’re not really interested if one in particular is related to the outcome**. Perhaps **it’s not really each symptom that’s important, but the idea that spillover is happening**.

**One possibility is to count up the number of items to which each respondent said yes. This variable will measure the degree to which spillover is happening**. In many studies, this is just what you need.

**But it doesn’t tell you something important**—

**whether there are certain combinations that generally co-occur, and is it these combinations that affect burnout?**

In other words, **what if it’s not just the degree of spillover that’s important, but the type**?

Enter Latent Class Analysis (LCA).

**LCA is a measurement model in which individuals can be classified into mutually exclusive and exhaustive types**, **or latent classes, based on their pattern of answers on a set of categorical indicator variables**. (Factor Analysis is also a measurement model, but with continuous indicator variables).

**-------------------------------------------**

**PROJECT OUTLINE**

**-------------------------------------------**

**Data debrief:**

The data was generated, and names were assigned to resumes with randomized sets of skills and sent to prominent employers. The randomized controlled experiment occurred in 2001. Variables chosen will hopefully provide information on whether or not the relationship exists at all.

**Latent Class Analysis:**

We are interested in ***investigating the degree and types of (highlighted items below) of employer ads***

**Categorical “yes-no” variables:**

***Applicant Related***

Was the applicant called back?

factor indicating gender.

Did the resume mention some honors?

Did the resume mention some volunteering experience?

Does the applicant have military experience?

Does the resume have some employment holes?

Does the resume mention some work experience while at school?

Was the e-mail address on the applicant's resume?

Does the resume mention some computer skills?

Does the resume mention some special skills?

Does the applicant have a college degree or more?

***Items for Classifying Employer ads***

Is the employer EOE (equal opportunity employment)?

Does the ad mention some educational requirement?

Does the ad mention some computer skills requirement?

Does the ad mention some organizational skills requirement?

factor indicating type of employer industry

Additional items

-Callback?

-Callback of particular implied ethnicity

1. Groups made with Employer + additional items in order to create classes of employers

|  |  |  |  |
| --- | --- | --- | --- |
| **Probability of ‘Yes’ response for each Class** | | | |
| **Item** | **Class 1:Unfair**  **(??%)** | **Class 2: Sort of fair**  **(??%)** | **Class 3: Fair**  **(??%)** | **Class 4: Very Fair**  **(??%)** |
| 1. Is the employer EOE (equal opportunity employment)? | NO = 0 | NO | YES = 1 | YES |
| 2. Does the ad mention some educational requirement? | YES | NO | NO | NO |
| 3. Does the ad mention some computer skills requirement? | YES | YES | YES | NO |
| 4. Does the ad mention some organizational skills requirement? | YES | NO | NO | NO |
| 5. Was the applicant called back? | NO | NO | YES | YES |
| 6. factor indicating gender. | FEMALE = 0 | FEMALE | MALE = 1 | MALE |
| 7. factor indicating ethnicity (i.e., Caucasian-sounding vs. African-American sounding first name). | Afam = 0 | Cauc = 1 | Afam | Cauc |

|  |  |  |  |
| --- | --- | --- | --- |
| **Probability of ‘Yes’ response for each Class** | | | |
| **Item** | **Class 1:Unfair**  **(??%)** | **Class 2: Sort of fair**  **(??%)** | **Class 3: Fair**  **(??%)** | **Class 4: Vary Fair**  **(??%)** |
| 1. Is the employer EOE (equal opportunity employment)? | .?? | .?? | .?? | .?? |
| 2. Does the ad mention some educational requirement? | .?? | .?? | .?? | .?? |
| 3. Does the ad mention some computer skills requirement? | .?? | .?? | .?? | .?? |
| 4. Does the ad mention some organizational skills requirement? | .?? | .?? | .?? | .?? |
| 5. factor indicating type of employer industry | .?? | .?? | .?? | .?? |
| 6. Was the applicant called back? | .?? | .?? | .?? | .?? |
| 7. factor indicating ethnicity (i.e., Caucasian-sounding vs. African-American sounding first name). | .?? | .?? | .?? | .?? |

Opportunity to improve on presentation of data

Run the LCA with different number of CLASSES (Assuming that employers are independent due to no employer ID)

**Choosing Classes name**: Subjective under proper representations (e.g. visual or otherwise)