Jobs:

<https://cran.r-project.org/web/packages/mediation/mediation.pdf>

Page 10

| Variable name | Role | Range of values | Units of measurement |
| --- | --- | --- | --- |
| depress1 | Response variable  (numerical) | 1~3 | Measure of depressive symptoms pre-treatment. |
| comply | Response variable  (categorical-binary) | 1 = participation  0 = no participation | Whether participant actually participated in the JOBS II program. |
| econ\_hard | Explanatory variable  (numerical) | 1~5 | Level of economic hardship pre-treatment |
| sex | Explanatory variable  (numerical) | 1=female  0=male | Indicator variable for sex |
| age | Explanatory variable  (numerical) | 17~72 | Age in years |
| occp | Explanatory variable  (categorical) | 7 categories | Factor with seven categories for various occupations |
| marital | Explanatory variable  (categorical) | 5 categories | Factor with five categories for marital status |
| nonwhite | Explanatory variable  (categorical) | 1 = nonwhite  0=white | Indicator variable for race |
| educ | Explanatory variable  (categorical) | 5 categories | Factor with five categories for educational attainment |
| income | Explanatory variable  (categorical) | 5 categories | level of income |
|  |  |  |  |
|  |  |  |  |

Response variable: depress2-depress1, work1

Explanatory variable: age, sex, marital, comply, education, occupation,

Stats 212 Group Project

Cristen McCann, Yilin Song, and Vitaly Ievlev

Stage 1: Proposal

I. Research Question: What socio-economic factors influence depression level and to what extent?

Rationales for each variable:

Such factors as: economic hardship, occupation, marital status, income, gender, education and even race can all influence person’s depression score. We would like to investigate to what extent these factors correlate with depression level. To gain a complete understanding of which factors influence depression we would also look for additional linkages between these variables.

* We expect economic hardship to be positively correlated with depression.
* The role of occupation in depression level is hard to estimate without statistical analysis.
* In terms of marital status - we expect married people to be less depressed.
* Income most likely negatively correlates with depression score but up to a certain limit after which the strength levels.
* Role of gender is hard to correlate with depression without running statistical analysis.
* For education level, highly educated people may be more depressed since their job might be more challenging.
* Role of race is hard to correlate with depression without running statistical analysis.

We want to find the factors that influence depression level. We expect that all the variables above may affect depression level.

Age is a potential confounder for all variables above with the exception of gender and race.

II. References

1) <http://bjp.rcpsych.org/content/190/4/293.full-text.pdf+html>

LORANT V., CROUX C., WEICH S, DELIÈGE D., MACKENBACH J, ANSSEAU M. (2007).

Depression and socio-economic risk factors: 7-year longitudinal population study. The British Journal of Psychiatry, 190(4), 293-298. doi:DOI: 10.1192/bjp.bp.105.020040

Socio-economic factors were assessed with regard to material standard of living, education, employment status and social relationships. A lowering in material standard of living between annual waves was associated with increases in depressive symptoms and caseness of major depression. Life circumstances also influenced depression. Ceasing to cohabit with a partner increased depressive symptoms and caseness. Overall, there was a clear relationship between worsening socio-economic circumstances and depression.

2. <https://link.springer.com/article/10.1007%2Fs00127-015-1166-3>

Patten, S., Williams, J., Lavorato, D., Wang, J., Bulloch, A., & Sajobi, T. (2016). The

association between major depression prevalence and sex becomes weaker with age. Social Psychiatry and Psychiatric Epidemiology, 51(2), 203-10.

10 national studies of Canadian people over age 14 were conducted from 1996-2013. The results of the study indicate that as age increases, the association between major depression and sex (gender) is weaker. In other words, there is less difference in the prevalence of depression between female and male subjects as age increases. The p-value was 0.002, making it a significant finding. From previous studies, it is known that depression scores are not different between sexes before puberty. After puberty, female depression scores are significantly higher than male. This particular study gives new insight, by suggesting that the gap between female and male depression scores closes as the subjects reach middle age. A handful of the studies even indicated that mean depression scores were higher in males for some age groups. This change may be attributed to the onset of menopause in women causing lower depression scores. The study provides insight for our own research, because it is evident that the effect of sex on major depression varies with age, so our variables may interact and influence each other in ways that we may not expect. Age is proven to be a confounding variable that we will have to take into account.

III. Variable Chart

| Name | Variable Role | Type | Values | Units |
| --- | --- | --- | --- | --- |
| Depression status1 | Response | Quantitative | 1-5 | N/A |
| Economic hardship | Explanatory | Quantitative | 1-5 | N/A |
| sex | Explanatory | Categorical | 1=female  0=male | N/A |
| age | Confounder/  Explanatory | Quantitative | >0 | years |
| occupation | Explanatory | Categorical | Seven categories: Professionals; operatives; managerial; sales workers; craftsmen; clerical;  laborers/service workers | N/A |
| marital | Explanatory | Categorical | Five categories:  Married; never married; separated; divorced; widowed; | N/A |
| nonwhite | Explanatory | Categorical | 0= white  1= non-white | N/A |
| education | Explanatory | Categorical | Five categories:  Less than high school; high school; community college; bachelors; graduate; | N/A |
| income | Explanatory | Categorical | Five categories:  50k+; 40 to 49k; 25 to 39k; 15 to 24k; less than 15k; | dollars |
| comply | response | binary | 1-participated  0-not | n/a |

Depending on which two variables we are investigating, other variables can be potential confounders. For example if we are correlating marital status and depression level, age might be a confounding variable.

IV. Data Description

We found the descriptions of datasets in R packages in <https://vincentarelbundock.github.io/Rdatasets/datasets.html>.

By looking at the number of columns and rows, we found this Job2 data. Then, we installed the [mediation] package in R server.

**The dataset is in Yilin Song’s Submit Section B > Project folder.**