BIVARIATE RELATIONSHIPS

Assignment 1 – SOCI 328

Stefanie Serka,

Deryn Gahagan, 16731135

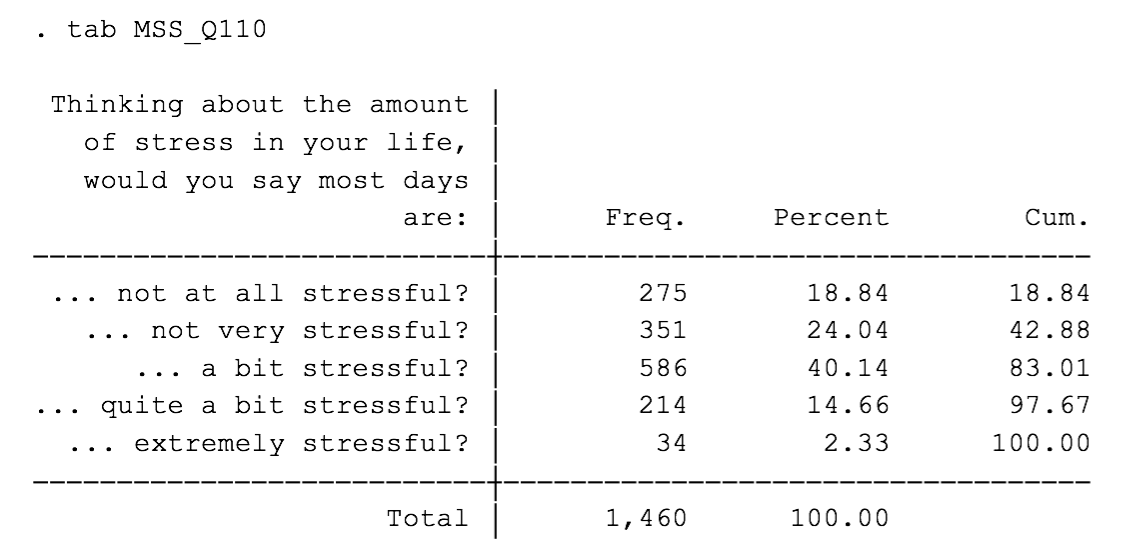
We would begin by assigning the variable NOCS2006\_C10 (occupational classification) as the independent variable and the variable MSS\_Q110 (self-reported stress) as the dependent variable. We would expect occupation to have a direct causal relationship to stress levels. Ie. X (NOCS2006\_C10) → Y (MSS\_Q10). We would expect these variables to be related because we spend an extensive amount of hours per week working, so the type of work that entails and the responsibility that goes along with it are quite likely to have an effect on self-evaluated mental state. Additionally, one’s job is typically their only source of income and any variable that is in relation to one’s standard of living can realistically be seen as a direct cause of stress.

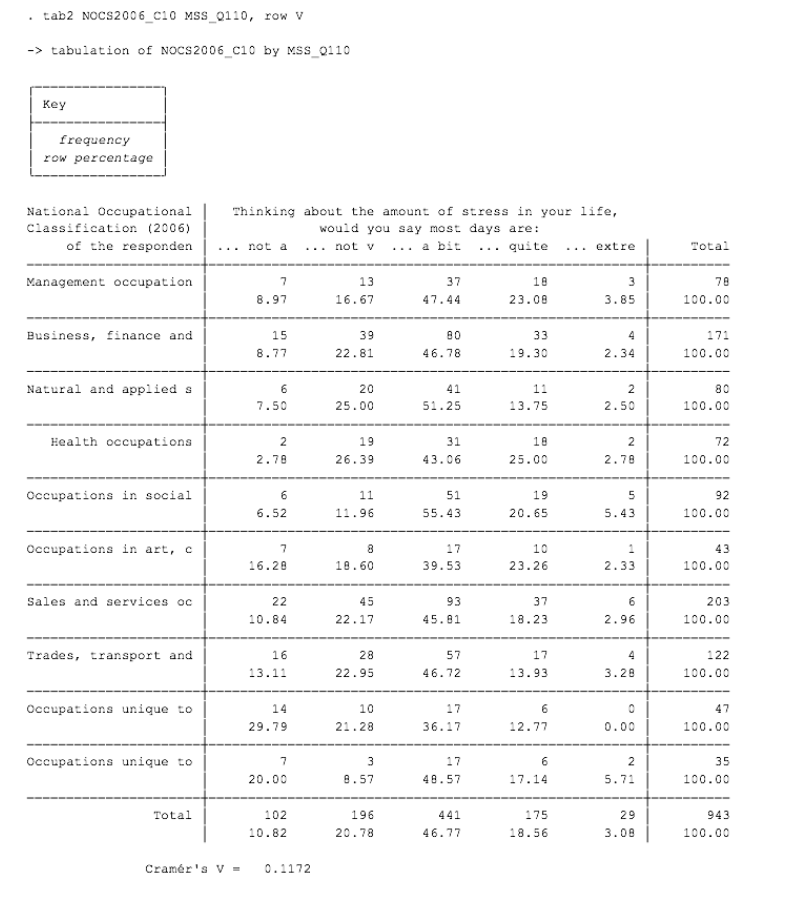
The variables measure and assess the classification of occupation and the amount of stress individuals feel they are experiencing most days in their life. They both are self-reported. The level of measurement for Occupational Classification is nominal as they are categorical. The level of measurement for Self-Reported Stress is ordinal, the categories have an order but undefined distances between. The data set NOCS2006\_C10 has a low frequency of total responses while MSS\_Q110 is still missing responses, but a much higher frequency of total responses is available.  While looking at distribution within each variable we can see that variable NOCS2006\_C10, Occupational Classification, has higher responses to occupations involving sales and services (21.50%) as well as business, finance and administrative jobs (18.02%). Lowest responses being manufacturing (3.79%) and art, culture and recreation (4.53%).

Our initial frequency table for NOCS2006\_C10:

|  |
| --- |
| National Occupational Classification |
| (2006) of the responden | Freq. | Percent | Cum. |
|  |  |  |  |
| Management occupations | 79 | 8.32 | 8.32 |
| Business, finance and administrative oc | 171 | 18.02 | 26.34 |
| Natural and applied sciences and relate | 80 | 8.43 | 34.77 |
| Health occupations | 74 | 7.80 | 42.57 |
| Occupations in social science, educatio | 93 | 9.80 | 52.37 |
| Occupations in art, culture, recreation | 43 | 4.53 | 56.90 |
| Sales and services occupations | 204 | 21.50 | 78.40 |
| Trades, transport and equipment operato | 122 | 12.86 | 91.25 |
| Occupations unique to primary industry | 47 | 4.95 | 96.21 |
| Occupations unique to processing, manuf | 36 | 3.79 | 100.00 |
|  |  |  |  |
| Total | 949 | 100.00 |  |

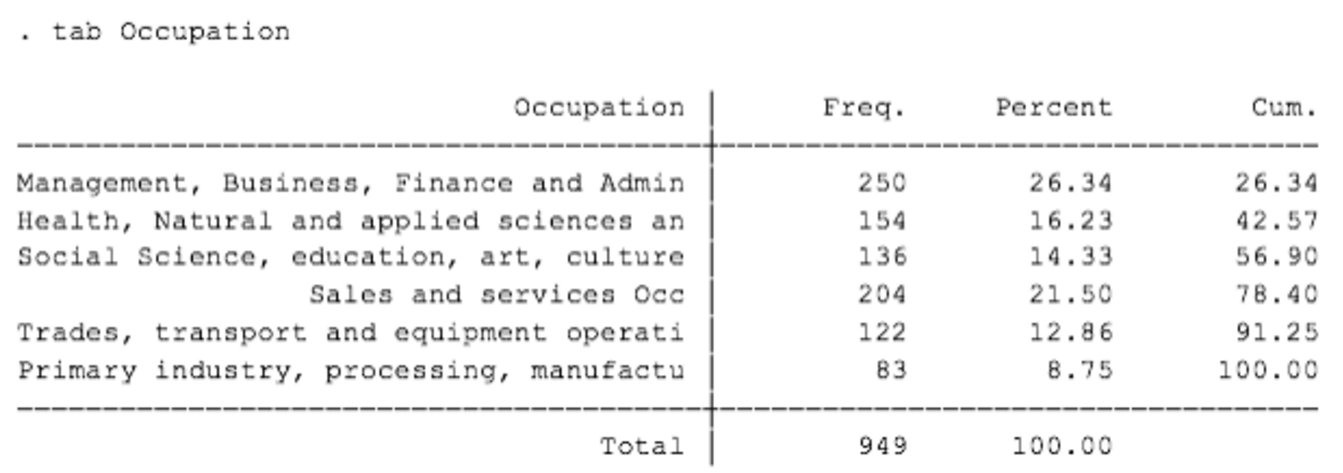
Our Frequency table for MSS\_Q110:



The distribution of the data set representing the variable MSS\_Q110, Self-Reported Stress, has a higher variance among respondents, the highest frequency of which responding ‘a bit stressed’ (40.14%). However, these distributions alone aren’t very insightful so we began by combining the two frequency distributions into one cross tabulation. We produced both column and row percentages to view the data and chose the row percentages as being more conducive to our understanding of the variable relationship. This allows us to see the relationship of the job type across each stress value.   


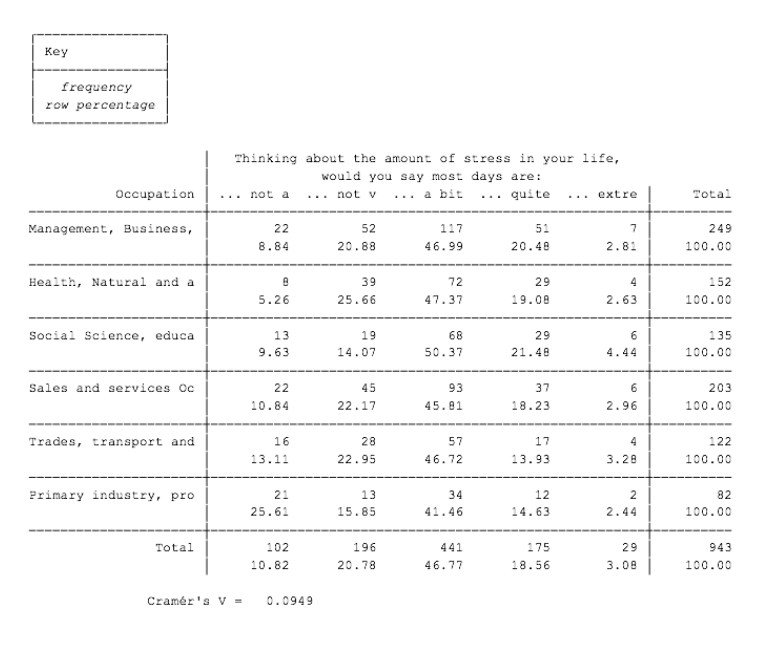
As you can see, the quantity of numbers makes for a difficult read of the material so we recoded Occupational Classification variables in order reduce the number of rows and create a better understanding of differences in stress levels between occupation types. To compare the variables better, we created a new variable “Occupation” from NOCS2006\_C10 then recoded the values of Occupation. We used command ‘recode Occupation (1=2) (3=4) (5=6) (9=10)’ to combine these categories. Then, we label defined the variables by using the command ‘label define Occupationk 1 ”Management, business, finances Occ” 3 “Health, Natural and applied sciences Occ” 5 “Social Science, education, and art Occ” 9 “Primary industry, production Occ”’, followed by ‘label values Occupation Occupationk’. We chose to combine mangagment occupations with business, finance and administrative occupations because for this comparison they would be undergoing a similar level of responsibility, work environment and job hours, etc. The same reasoning goes for combining the others as we did. By recoding the variables, we are then able to understand the relationship between occupation and stress level more clearly.

Our recoded Frequency table for Occupation (formerly NOCS2006\_C10):



The same data is available as on page 2, but it is group in a way that we feel does not distort the results but allows for a clearer read of the information.

Finally, we recombined the data in to a new cross tabulation to view the relationship between our two variables, Occupational classification and Self-Reported Stress. Below is this cross tabulation. Here we can see the distribution of each occupational category row across the columns of stress quantities.

Tabulation of Occupation by MMS\_Q110: Looking at this final representation of the data we conclude that our original theoretical rational was incorrect. The distribution between the categories is quite similar. Everyone appears to find their days ‘a bit stressful’ within all job types. By looking at the percentages we can see a consistency among the higher stress options with a slight variance upwards for social science and educational and sales and services, but it is not overwhelmingly higher. It is when looking at who self-reports their days as being ‘not at all stressful’ that we see a bit more variance. We would conclude that there is not enough consistency within the job categories to show one occupation being a direct cause of higher self-reported stress levels or the opposite. Therefore, Occupational Classification is not directly causal of increased Self-reported Stress.