**Proposal**

**Summary**

The senior population in South Korea is rapidly growing. As a result, seniors continue to participate in the labour market, and often engage in educational programs to further develop their workplace skills. The objective of this study is to identify how demographic, organizational and learning factors are associated with South Korean senior workers’ skill usage and proficiency scores. Identifying what is the association between the aspects of senior workers and their skill use is a critical research issue because it will help Human Resource officers to develop training programs.

**Data**

The data set was obtained from the open source of the Programme for the International Assessment of Adult Competencies. These data were collected from questionnaires that measured two key cognitive skills: literacy and numeracy. Based on the study’s objective, employers aged 50-65 year-old on the private sector in South Korea were included.

**Dependent variables***:*

There are four dependent variables: Skill usage for work (numeracy or literacy) and Proficiency test score (numeracy or literacy). The skill usage for work states the frequency that numeracy/literacy skills are applied. These variables are calculated by taking the average of related sub-items. These sub-items are measured by likert scale: 1,2,3,4,5 (from strongly disagree to strongly agree).

The proficiency test score is the average test score of numeracy/literacy. They are continuous and seem to be Gaussian distributed. Therefore, our analysis would focus on the average test score first.

**Independent variables***:* we have 13 independent variables that need to be test.

|  |  |  |
| --- | --- | --- |
| **Demographic Factor** | **Organizational Contexts** | **Learning/Education** |
| **Continuous:**  Age  Work Experience (in years)  **Categorica**l:  Gender  Education level  Employment type: full-time or part-time | **Continuous**:  Work Flexibility  Learning opportunity | **Continuous**:  Active learning strategies  number of hour of participation  **Categorica**l:  Participation in non-formal education  Participation in job-related education  Participation in job-related adult education  Participation in non-job-related adult education |

**Analysis Outline**

**Explanatory analysis**

1. overall understanding of the dataset
   1. statistical summary
   2. some plots / graphs for visualization
2. identify potentially important variables
   1. check correlation between the response variable and explanatory variables
   2. perform ANOVA and t-test to do group comparison

**Research questions**

1. how demographic, organizational, and learning factors are associated with senior workers’ numeracy skills (measured by numeracy test score)
2. how demographic, organizational, and learning factors are associated with senior workers’ literacy skills (measured by literacy test score)

**Methods**

To address the two research questions above, we will use a stepwise linear model with AIC criterion because AIC helps to choose the best ‘predictive’ model. AIC helps to seek a model that has a good fit to the truth but few parameters. Before running AIC, we will check collinearity between covariates because AIC picks only one of the collinear variables. We need to be aware of the fact that some variables are not picked by AIC due to collinearity. After running stepwise AIC, we run model diagnostic and check goodness of fit of selected model.

Another model selection method we will use is stepwise linear model with BIC criterion. BIC penalizes model complexity more heavily. Using multiple approach to do model selection could provide further evidence supporting the inclusion or exclusion of certain variables.

The final result could be the intersection of variables picked by AIC and lasso united with possible important variables obtained from univariate test. We could also have two separate models for AIC and BIC and then make comparisons.