SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCST1043
Sem: 2 Session: 2024/25

SECTION A: Project Information.

Program Name:	**Masters of Science (Data Science)**
Subject Name:	**Project 1 (MCST1043)**
Student Name:	LI HONGLIN
Metric Number:	MCS241031
Student Email & Phone:	Lihonglin@graduate.utm.my&601126779007
Project Title:	Re-employment Analysis of Older Adults Aged 50–70 Using Kaggle Data
Supervisor 1:	
Supervisor 2 / Industry Advisor(if any):	

SECTION B: Project Proposal

Introduction:

As estimated by the United Nations World Population Outlook, the worldwide fertility rate is expected to fall to 2.2 by 2050, approaching the critical global alert level. [1]., From a social demography perspective, higher levels of education are associated with lower fertility rates, increased investment in children, reduced mortality, improved health outcomes, slower population growth, and greater gender equality.[2] That lead to a shortage of the labor force. Meanwhile,population aging has become an unavoidable challenge for the majority of nations worldwide[3]. Facing these social problems, some countries have introduced policies of delaying retirement or re-employment after retirement[4]. The group aged between 50 and 70 is in the middle of middle age and old age, and it's very crucial. This group still has a certain amount of labor ability before and after retirement, and they also have the intention to re-employ[5] But they encounter many problems, such as the need to update their skills, experiencing age discrimination, having poor health conditions, and the changes in industry demands. We use the existing employment data on Kaggle and conduct an empirical study on the employment situation of this group through data analysis.

Problem Background:

The proportion of middle-aged and elderly people in the world population decreased from 8.0% in 1955 to 6.4% in 1990, and then remained basically unchanged until 2005, and then gradually rose to 8.0% in 2025 and then to 10.9% in 2045. From a new perspective, before 1990, the world population was experiencing a process of

younger age and then aging. From the traditional point of view, the world population is aging, and by 2045, the proportion of the elderly population will be 4.3 percentage points higher than from the new point of view[6]. Current employment research primarily focuses on two groups: new college graduates and the general elderly workforce [7]. However,studies on re-employment for people aged 50-70 remain limited [7]. This pre-elderly group faces a unique dilemma:while aging brings physical and cognitive decline [8], and inflation particularly reduces purchasing power for lower-income seniors [9], many remain capable and willing to work [10] yet face systemic employment barriers [9]. Although most discussions center on pension adequacy, the issue is more complex - even with sufficient pensions, labor shortages among younger workers can drive inflation that erodes retirees' purchasing power [8,9]. This highlights the dual importance of supporting re-employment for this group: protecting individual welfare while maintaining economic stability [8,10]. This study utilizes publicly available datasets from the Kaggle platform to conduct an in-depth quantitative analysis addressing the significant research gap in the field of elderly re-employment. Through systematic examination of the employment characteristics, barriers, and career development potential of the pre-elderly population (aged 50-70), the research aims to establish robust empirical evidence in three key areas: (1)revealing the actual employment status of this demographic in labor markets; (2) identifying critical determinants affecting their re-employment success rates; and (3) Based on the data, give suggestions to the government, enterprises and social organizations to help them formulate reemployment policies and support measures.

Problem Statement:

Solomon Barkin [11]pointed out that the middle-aged and elderly groups have long faced discrimination in many aspects and encountered a lot of difficulties when seeking reemployment. And the reemployment of this group is of great significance to the social economy, but it also brings challenges to enterprises and the government. At present, the existing literature and data analysis on the overall reemployment situation of the group aged 50 to 70 are still insufficient. The Antti J. Tanskanen model predicts [12]that its state variables include employment status, age, accumulated pension, last salary, state duration, etc., but there is a lack of comparative analysis especially in aspects such as industry, income and skills. The core problem that this study aims to solve is: how to use public data and apply Data cleaning and preprocessing, Data visualization analysis methods to have an in-depth understanding of the characteristics of this group in terms of employment quality, occupational distribution and reemployment success rate.

References

[1] Davis, D., Hartley, D., Li, C., 2022. Firm locations and declining US fertility rates. Quant. Econ.
[2]. Lutz, W. & KC, S. Global human capital: integrating education and population. Science 333, 587–592 (2011)
[3] Joan E. Madia,Francesco Moscone,Asieh Hosseini Tabaghdehi,Jong Chol An & Changkeun Lee.(2025).Fertility decline and tax revenues in South Korea.Research in Economics,79(2),101025-101025.
[4] Sun, S. D., & Gao, W. S. (2025). Delayed retirement, fertility support policies, and birth rates. Population & Economics, (2), 16-30.
[5] Bai Yue. "The Impact Mechanism of Work-Life Balance on Re-employment of Middle-aged and Older Adults and Policy Implications." Lanzhou Academic Journal, 1-27.

[6] Chen, W. (2022). Forecasting China's Population Decline and Aging Trends. Social Science Journal, (5), 133-144.
[7] BLUNDELL R,BRITTON J,DIAS M C,et al.The impact of health on labor supply near retirement[J].Journal of Human Resources,2023,58(1):282-334.
[8] Ning, G. (2025). The differential impact of population aging on residents' income distribution and its mechanisms. Academic Forum, 48(1), 99-109. https://doi.org/10.16524/j.45-1002.20250116.001
[9] THERN，MATILLA－SANTANDE R N，BODIN T，et al．Precarious employment at a young age and labor－market marginalization during middle－adulthood: a register－linked cohort study［J］．Scandinavian Journal of Work，Environment ＆ Health，2023，49(3) : 201
－210．
[10] KANTACHOTE K，WI ROONS RI N．Do elderly want to work? modeling elderly’s decision to fight aging Thailand［J］．Quality ＆
Quantity，2023，57(1) : 509－539．
[11]Solomon Barkin.(2012).The Employment of the Older Worker—The Economically Middle Aged and Aged.Journal of the American Statistical Association,27(177A),102-108.
[12] Tanskanen, A. J. (2020). Deep reinforced learning enables solving rich discrete-choice life cycle models to analyze social security reforms. arXiv:2010.13471v3

Aim of the Project:

1.Talk about how people aged 50 to 70 are distributed in the job market, such as which industries they are in, their salaries,skills and educational levels.

2. identifying critical determinants affecting their re-employment success rates.

3. Based on the data, give suggestions to the government, enterprises and social organizations to help them formulate reemployment policies and support measures.

Objectives of the Project:

1.Find employment, age - related datasets on Kaggle. Screen and clean the data.

2.Analyze re - employment of 50 - 70 - year - olds in different industries and income levels with descriptive stats and group comparison.

3.Create an interactive dashboard with Tableau or Power BI to show employment trends and data insights. Summarize results, write a report, and give policy suggestions.

Scopes of the Project:

1.Data source: Only use the employment - related data that can be publicly obtained on Kaggle.

2.Research object: Study the middle - aged and elderly group whose age is between 50 and 70.

3.Analysis method: Mainly use descriptive statistics, group - comparison analysis, and data visualization methods.

Expected Contribution of the Project:

1.A complete analysis report on the re-employment situation of the 50 - to - 70 - year - old group, including industry distribution, salary levels,.

2.An interactive employment data visualization dashboard that can dynamically present the comparison results of data from different aspects.

3.Summarize the key factors and impact indicators to provide targeted re-employment policies and support suggestions for the government, enterprises, and social organizations.

Project Requirements:

Software: -Python（Pandas、Matplotlib、Seaborn）、Tableau or Power BI

Hardware: -Personal computer with sufficient processing power and storage capacity

Technology/Technique/ Methodology/Algorithm: -Data cleaning and preprocessing (using Pandas)
-Descriptive statistical analysis and group comparison (with the help of Python and Excel)
-Data visualization (using Matplotlib, Seaborn, Tableau, Power BI

Type of Project (Focusing on Data Science):

[ ] Data Preparation and Modeling

[✓] Data Analysis and Visualization

[✓] Business Intelligence and Analytics

[ ] Machine Learning and Prediction

[ ] Data Science Application in Business Domain

Status of Project:

[✓] New

[ ] Continued

If continued, what is the previous title?

SECTION C: Declaration

I declare that this project is proposed by:

[✓] Myself

[ ] Supervisor/Industry Advisor ( )

Student Name: LI HONGLIN

Signature 07/April/2025 **Date**

SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

I/We agree to become the supervisor(s) for this student under aforesaid proposed title.

Name of Supervisor 1: .....

Signature **Date**

Name of Supervisor 2 (if any): .....

Signature **Date**

SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

Result:

[ ] FULL APPROVAL [ ] CONDITIONAL APPROVAL (Major)*

[ ] CONDITIONAL APPROVAL (Minor) [ ] FAIL*

***** Student has to submit new proposal form considering the evaluators' comments.

Comments:

Name of Evaluator 1: .....

Signature **Date**

Name of Evaluator 2: .....

Signature **Date**