

## **Admission Test to “The 2<sup>nd</sup> Empirical Research Immersion” at SUFE-Finance**

### **Instruction:**

1. This test consists of two exercises. The first one mainly requires data processing, while the second one requires both data processing and analytical writing. You must hand in your work in English language only.
2. You are free to choose any data processing software or econometrics package. Source dataset for exercise I is provided in .CSV format and that for exercise II is in STATA format.
3. You are free to search for information from reference books or internet. However, you must provide attribution if you build on someone else’s work.
4. Your submitted output will include three parts: (1) your code to generate required dataset for exercise I; (2) An essay that contains both tables and text to address the questions in exercise II; (3) your codes for exercise II. Please document your codes clearly so that another person can read and understand your logic.
5. The difficulty level of each question is marked for your reference. We encourage you to submit your report as long as you are able to finish the “Basic” level. No regressions are necessary for any of the questions; however, you are free to use the regression tools if you think they help.
6. A template for your report is provided in the next page. When you produce your report, put yourself in the position of a potential reader and try to make your report as clear, illustrative, and professional-looking as possible.
7. There is no minimum of maximum page requirement. When your time is about to run out, please wrap up and produce a self-containing report with what you have. It is more important that you provide a logically complete report than providing many unsynthesized analyses.
8. You must finish the work independently, i.e., with no help from or consultation with anyone else. If you have clarification questions, please contact [2015210425@live.sufe.edu.cn](mailto:2015210425@live.sufe.edu.cn) .

## **Report Template**

[Essay Title]

[Author Name]

[Author Department/School affiliation]

[Opening paragraph outlining the research question.]

[Describe how you address question #1]

[Insert Table X or Figure X here.]

[Describe how you address question #2]

[Insert Table X or Figure X here.]

....

[Concluding paragraph.]

## Exercise I: Data Manipulation (Basic)

### Dataset

The dataset scp-1205.csv contains information on Medicare Advantage (MA) enrollment by contractcounty pair. The variables in the dataset are:

**countyname:** name of the county

**state:** state postal code

**healthplanname:** name of the health plan

**typeofplan:** type of health plan

**countyssa:** Social Security Administration county code

**eligibles:** number of individuals in the county that are Medicare eligible

**enrollees:** number of individuals enrolled in the specific health plan

**penetration:** percent of individuals in the county enrolled in the plan, defined as  $100 \times \text{enrollees} / \text{eligibles}$

**monthlypayments:** Medicare's monthly payments to the health plan

A row in the data corresponds to a contract-county pair. For example the row that begins

countyname	state	contract	healthplanname
------------	-------	----------	----------------

AUTAUGA	AL	H0150	HEALTHSPRING OF ALABAMA INC.
---------	----	-------	------------------------------

indicates that there are 6883 eligibles in Autauga, AL and that 313 of them are enrollees in the Healthspring of Alabama plan.

County-plan pairs that have missing values for eligibles, enrollees, and penetration should be considered as having zero values for these variables.

### Objective

The goal of the exercise is to write a short script that produces a county-level dataset that identifies the number of plans and total enrollment in each county (you may exclude territories such as Puerto Rico and Guam). In particular, we would like you to produce a dataset that has following variables:

**countyname:** name of the county

**state:** state postal code

**numberofplans1:** number of health plans with *more than 10 enrollees*

**numberofplans2:** number of health plans with *penetration > 0.5*

**countyssa:** Social Security Administration county code

**eligible:** number of individuals in the county that are Medicare eligible

**totalenrollees:** number of individuals in the county with a MA health plan

**totalpenetration:** percent of individuals in the county enrolled in a MA plan, defined as  $100 \times \text{totalenrollees} / \text{eligibles}$

The dataset should include one observation for each county and should be sorted by state and then by county.

## Exercise II: Labor Market Analysis

**File:** cps wages LFP.zip contains a stata .dta file from the U.S. Current Population Survey since 1976. It provides individual-level data on year, state, month, survey weight, demographic characteristics, and labor market outcomes.

**Goal:** We'd like you to use these data to produce your best answer to the following question:

how have hourly wages (“wage”) and labor force participation (“lfp”) evolved for skilled and unskilled workers since 1976?

- (a) Please summarize the key trends for wages and labor force participation. (**Basic**)
- (b) Among men older than age 25, which groups of people have had the biggest changes in labor force participation? (**Basic**)
- (c) What factors do you think are driving these patterns? (**Intermediate**)
- (d) What evidence might you want to assemble to test these hypotheses if you were to investigate them further? (**Difficult**)

### Output:

- (a) Please provide a short note in .pdf format that concisely answers these questions. Please be sure to include a few graphs and/or tables to support your conclusions.
- (b) Please provide the code that you used to produce these figures and/or tables.

Note: You should feel free to use whatever techniques you want. The goal here is not to show off hi-tech econometrics, but rather to show us how you think about data. Sometimes something as simple as a graph can do more for an argument than all the estimators in the world.