Basic Data Analysis

Stephen Turnbull

Business Administration and Public Policy

Homework 8: Due July 7, 2011

Requirements

Homework assignment about multivariate models, including regression and factor analysis.

Homework 8

Due Thursday, July 7, 2011, 11:45 am. Submit by email to
`data-hw@turnbull.sk.tsukuba.ac.jp`. Your header should look like this:

```
From:  a-student@sk.tsukuba.ac.jp
To:  data-hw@turnbull.sk.tsukuba.ac.jp
Subject:  Basic Data Analysis HW#8
```

The subject should be all half-width Roman letters (ASCII).

Get the data

1. Get the data set `Section1All_csv.csv` from the home page.

2. This data set has several sections with different kinds of data. *After reading and thinking about the rest of the problems*, pick one section; using data across sections is a bad idea.

3. Input the data into your statistical package, and print out the data of the section (only!—no fair printing everything and editing the output) you have picked.

Correlation matrix

4. Generate a correlation matrix for all the variables in your section.

5. Think of some way in which *some* of the variables in your section are related. Refer to scientific theory where possible.

Define and estimate a model

6. Define a regression model for *the variables you picked*. Explain why you picked the dependent variable.

7. Write down your regression model.

8. Estimate the regression model using your statistical package.

Add an unrelated variable

9. Use Excel or your statistical package to generate a series of random numbers, enough to make a new variable for your data set.

10. Add it to the data set, and print out the data set (*i.e.*, your model variables plus the random variable).

11. Add the random variable to your model of problem 3 as an explanatory variable, and estimate the new regression model.

12. Define and execute a hypothesis test that the new variable is in fact statistically unrelated to the model.