

Visualization and Statistical Modeling of Financial Data with R

Masayuki Jimichi^{1*}, Shinnosuke Maeda²

1. Prof. School of Business Administration, Kwansei Gakuin University, Japan

2. M.B.A. candidate, Graduate School of Business Administration, Kwansei Gakuin University, Japan

*Contact author: jimichi@kwansei.ac.jp

Keywords: Data Visualization, Statistical Modeling, Exploratory Data Analysis, Financial Data

We treat visualization (*e.g.*, [6]) and statistical modeling (*e.g.*, [2]) for financial longitudinal data (*e.g.*, sales, employee, assets) of Japanese firms which belong to the first section market of the Tokyo Stock Exchange based on exploratory data analysis [7] with R. They are extracted from a database system of Nikkei NEEDS financial data. (See [5].) As a result of data visualization from temporal and cross-sectional aspects, we know the joint distribution of the data sets at each closing day is a multivariate lognormal (*e.g.*, [3]) by using R packages **ggplot2** [8], and **googleVis** [4]. We build a statistical model based on the result. Under the condition of fixed time, a lognormal linear model (*e.g.* [1]) with dummy variables which denote the middle classification of industries is very useful for explaining sales by employee and assets. Furthermore, this result is valid in terms of time variation. Note that we can fit the model to the dataset by using the basic R function `lm` only.

References

- [1] Bradu, D. and Y. Mundlak (1970). Estimation in lognormal linear models. *Journal of the American Statistical Association* 65(329), 198–211.
- [2] Chambers, J. M. and T. J. Hastie (Eds.) (1991). *Statistical Models in S*. Chapman and Hall/CRC.
- [3] Crow, E. L. and K. Shimizu (Eds.) (1988). *Lognormal Distributions: Theory and Applications*. Marcel Dekker.
- [4] Gesmann, M. and D. de Castillo (2011). Using the Google visualization API with R. *The R Journal* 3(2), 40–44.
- [5] Jimichi, M. (2010). Building of financial database servers. Technical report. <http://kgur.kwansei.ac.jp/dspace/handle/10236/6013>.
- [6] Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut.
- [7] Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Co.
- [8] Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer Verlag.