PROSPECTS OF TEACHING STATISTICS
USING INTERACTIVE APPLICATIONS BY SHINY

Takafumi Kubota
Tama University, Japan
kubota@tama.ac.jp

This study contributes to the teaching of statistics using Shiny's interactive applications. Using Shiny's application named Radiant enables students with insufficient training in mathematics or statistics to concentrate on the content of the analysis. This study also proposed the use of Shiny application in fields such as disaster prevention and marketing. Specifically, the application will be effective in disaster prevention because it places the necessary information on a map at the time of a disaster. This project also created a calendar to visualize the annual sales of the application. Using this content made it possible for the user to obtain insights, which were not available when a data table was used.

INTRODUCTION

The use of data science and artificial intelligence has been increasing in recent years. Students who do not discuss mathematics or statistics in these contexts need to first practice mathematics or statistics using analysis tools, because an intuitive approach can be gained through the repeated use of medium of the analysis.

This study offers the use of the R package radiant, a package based on the R package Shiny, in order to enable students to concentrate on the content of their analysis. This study also proposes the use of Shiny's application in fields, such as disaster prevention and marketing. Specifically, the application will be effective in disaster prevention because it places the necessary information on a map at the time of a disaster. This project also created a calendar to visualize the annual sales of the application. Using this content made it possible to obtain insights that were not available when a data table was used.

SHINY AND APPLICATIONS

We proposed a format to analyze data using several packages created by R and Shiny. Here we used the following packages.

- radiant
- ggplot2
- leaflet

Radiant is an open-source R package application that runs on a platform-independent browser-based interface and is used for business analytics. The application is based on the Shiny application package (also an R package) and can be run locally or on a server. Nijis (2017) states that the packages' key features are as follows; Explore, Cross-platform, Reproducible, Programming and Context.

For plotting on open street maps, we used a leaflet, an R package application to draw maps and over rays for several statistics. We also used the ggplot2 package to visualize high dimensional marketing data.

The motivation for choosing these R packages was as follows:

Radiant is a good tool for students who are not familiar with CUI, while ggplot2 is useful for creating elegant data visualizations using the grammar of graphics, and leaflet is the only application in the Shiny package that can handle maps.

These packages are therefore suitable for students, and particularly for those who are not familiar with CUI and are not well-versed with complex graphics. There is an advantage in using maps with Shiny's applications in these three packages.

CASE STUDIES

In the case studies, we introduced several Shiny's applications, such as decision tree analysis, visualization of refuges, and sales calendar, over a period of two years.

Decision tree

In a university course of Machine Learning (ML), radiant was used as an analysis tool. In this paper, data on diamonds are used as example data to introduce the radiant application. Using radiant, we could easily conduct a correlation analysis to obtain the results in a scatter plot matrix and a correlation matrix, as in Figure 1. Figure 1 shows that diamond price and carat have a strong correlation of 0.93. From these results, the user can easily determine which two variables have high or low correlations in each pairs of variables.
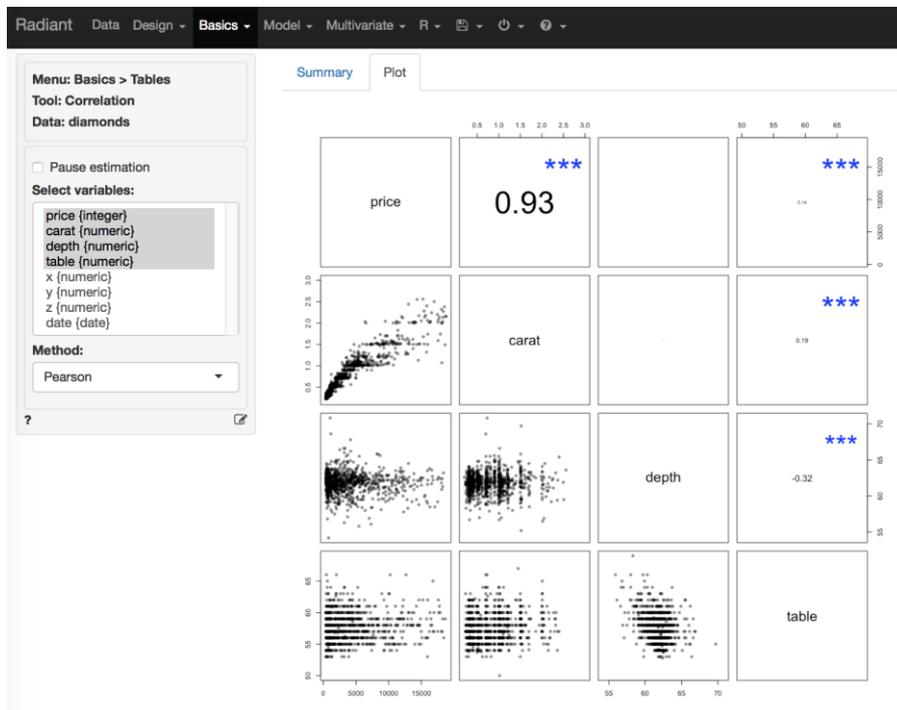

Figure.1 Hard copy of results of correlation analysis using radiant

Using this data, we could then conduct a decision analysis for a regression tree. Figure 2 shows the hard copy of the results of regression tree analysis using radiant.
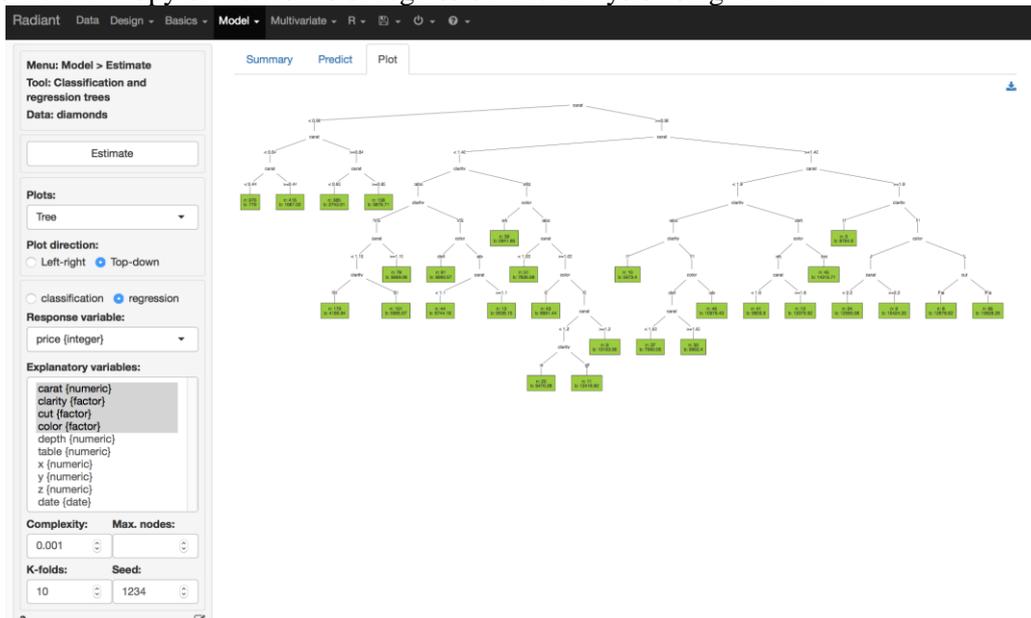

Figure.2 Hard copy of results of regression tree analysis using radiant

In a university ML course, it is helpful to determine the appropriate number of node for a decision tree. Figure 3 thus shows the appropriate number of nodes using K-folds cross-validation. This figure shows the minimum error achieved with 28 nodes using the results of the evaluate tree pruning based on cross-validation.
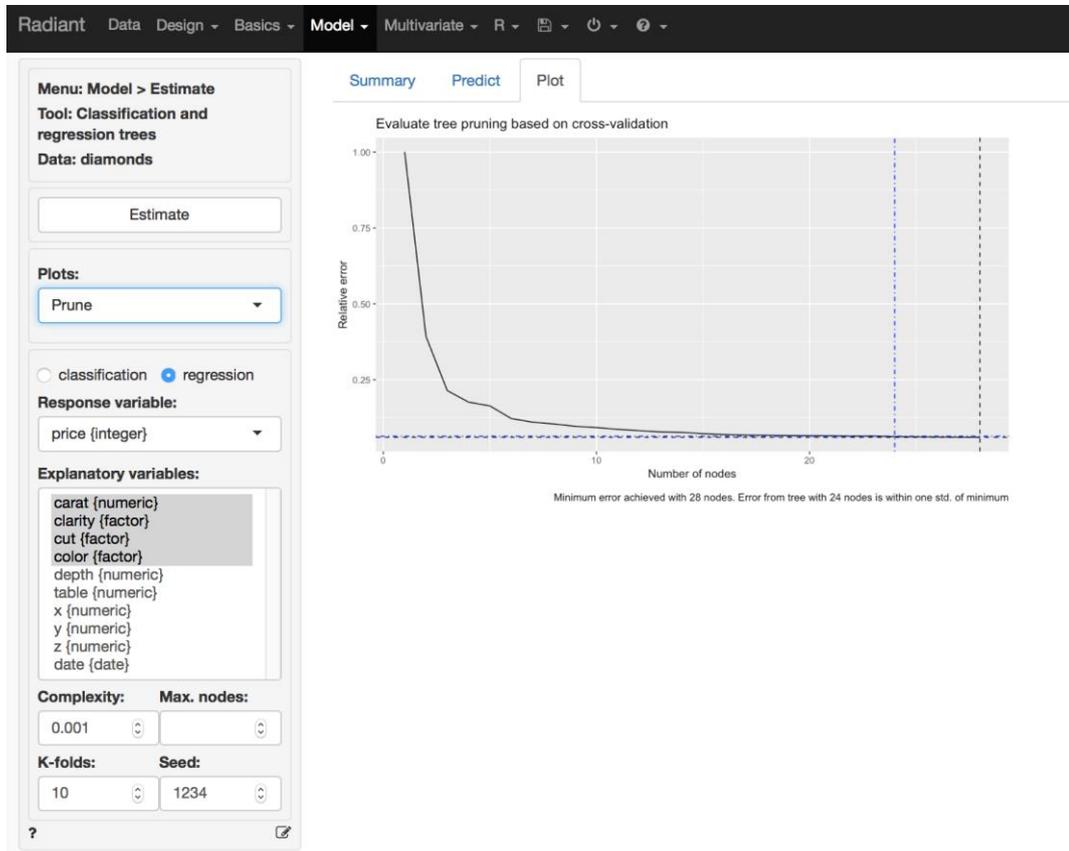

Figure. 3 Hard copy of other of regression tree (pruning) results using radiant

Visualization of refuges

We introduced leaflet and Shiny to the graduate students of my research laboratory. The study used open data published by Ministry of Land, Infrastructure and Transport. The objective of the study offered to use map application for people of the refugee to find out the seats of refuges. We proposed using the leaflet and Shiny web applications for this purpose. Figure 4 shows one web application used to view interactive changes from Hinanjo (a place of disaster refuge) to Koiki-Hinanbasho (a place of disaster refuge in large area) and the menu button we made using R and Shiny.
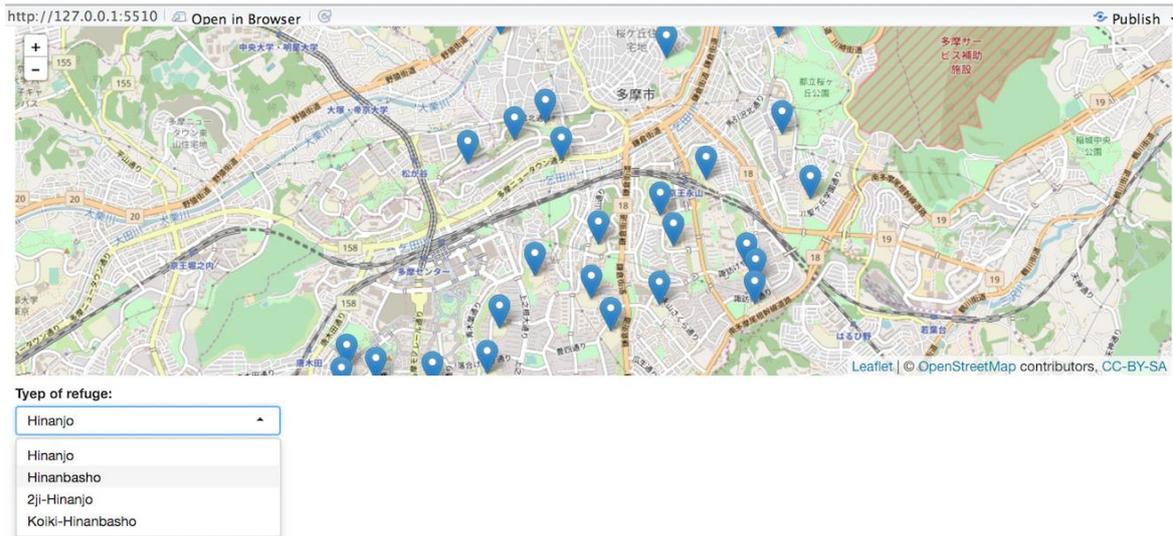
Figure. 4 Hard copy of visualization of refuges and interactive changes for Tama city, Tokyo, Japan.

Marketing

Figure 5 shows the sales results from one shop's calendar from July 1 of 2015 to Jun 30 of 2017. The values of this graph categorized by colors from yellow to red by the sales and are dummy data for this paper.

The figure uses R's ggplot2 to separate facet rows by year, including the date in each cell and facet columns by months, including the weeks and days of the month in each cell from 2015 to 2017.
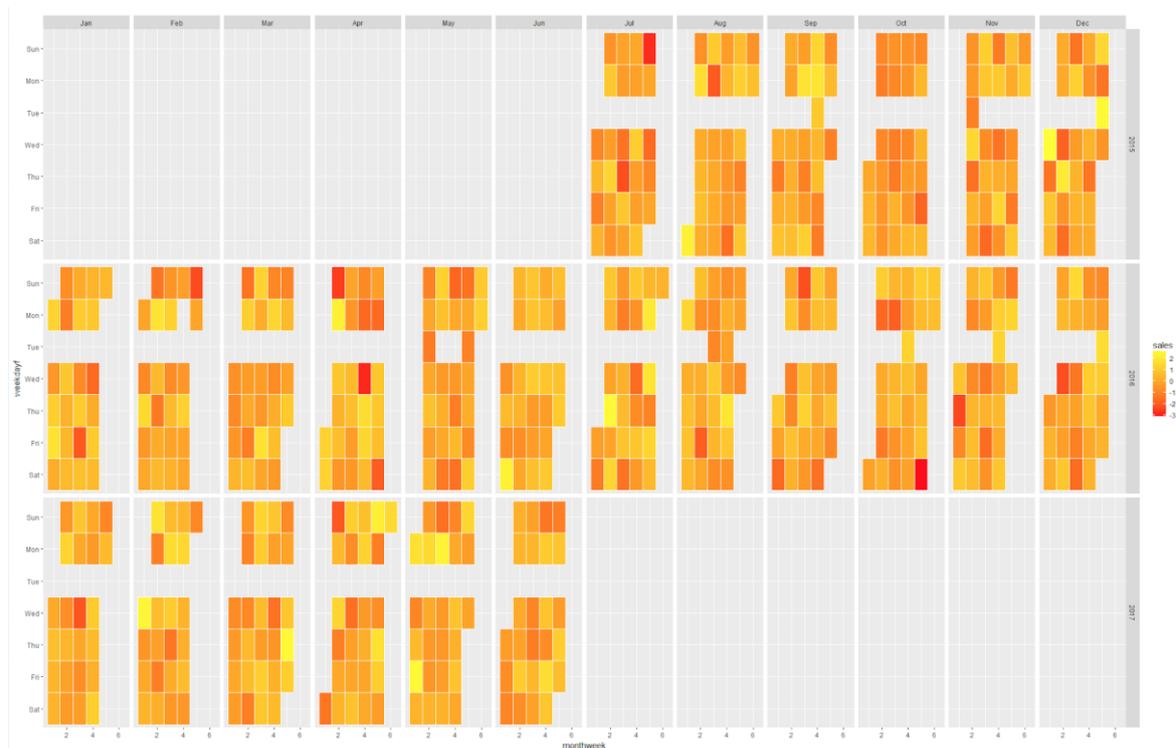

Figure. 5 Examples of sales from July 2015 to June 2017

SUMMARY AND FUTURE STUDIES

We proposed using R applications of radiant, leaflet and ggplot2 for multivariate statistical analysis, visualizing characteristic values on maps, and visualizing sales data. Using this content, it was possible to obtain insights not available by only using a data table.

REFERENCE
Nijis, V. (2017). Radiant - Business analytics using R and Shiny. Retrieved from vnijs.github.io/radiant/