Sales Analysis with Application to Sparkling Beverage Products Sales in Southern Thailand

Wassana Suwanvijit (Corresponding author)

Faculty of Economics and Business Administration, Thaksin University

140 Karnchanavanich Road, Muang, Songkla, 90000, Thailand

E-mail: swassana@gmail.com.

Chamnein Choonpradub

Faculty of Science and Technology, Prince of Songkla University

181 Charernpradit Road, Rusamilar, Pattani, 94000, Thailand

E-mail: cchamnein@bunga.pn.psu.ac.th

Nittaya McNeil

Faculty of Science and Technology, Prince of Songkla University

181 Charernpradit Road, Rusamilar, Pattani, 94000, Thailand

E-mail: nittaya@bunga.pn.psu.ac.th

Abstract

Managers need to know about their business performance and how to gain rapid insight into fast-growth products and profit opportunities. This case study used monthly sales revenue from sparkling beverages of data collected routinely from 14 provinces of Southern Thailand during years 2000-2006. Linear regression models of log-transformed sales revenue per 1000 population were used to analyse per capita revenue and study product preference with respect to flavour and branch location. Population data was obtained from the 2000 Population and Housing Census of Thailand. The model contains quarterly effects and interactions associated with flavour-branch and branch-year. The average annual per capita consumption value rate was 297 baht for the population of 8.087 million residents in Southern Thailand. Branches in Samui and Phuket had higher rates than other branches. In addition, areas with different proportions of Muslims in their population had different beverage preferences.

Keywords: Sales revenue, Linear regression model, Sparkling beverages, Consumption rate, Product preference, Southern Thailand

1. Introduction

Generally, business owners and managers need to know about their sales trends and profitability, not only overall but by departments or products within each department. They also have to consider the potential of their business improvement from different areas, by gathering, classifying, comparing, and studying company sales data. Most managers don't have the time or expertise to analyse their sales data themselves. Nevertheless, there are many studies about sales analysis. In the case study of carbonated soft drink consumption and bone mineral density in adolescence by McGartland et al. (2003), adjusted regression modeling was used to investigate the influence of carbonated soft drinks on bone mineral density. Residual demand analysis (Higgins et al. 2005) could be applied to test whether carbonated soft drinks were a relevant product market under the Merger Guidelines. Goktolga et al. (2006) analyzed socio-demographic factors affecting a decrease in consumption of chicken meat because of Avian Influenza (Bird Flu) in Turkey using a multiple bounded probit model. Descriptive and multiple regression analyses (Probart et al. 2006) had some good features to identify the factors associated with the offering and sale of competitive foods and school lunch participation. Helasoja et al. (2007) used logistic regression to fit socio-demographic patterns of drinking and binge drinking in Estonia, Latvia, Lithuania and Finland, 1994-2002. Ratio analysis, historical trend analysis and linear regression analysis used by Bureau

(2007) were found to be useful for analyzing the factors and examining the opportunities critical to the success of the food and beverage industry in India.

This study illustrates methods for graphing and statistical modeling available to businesses and demonstrates how each can be implemented using freely available software. Linear regression models are used to see how well these methods can be applied to a case study in southern Thailand. From an academic perspective the study demonstrates how easily business analysis can be taught to both students and managers. From a managerial perspective, the study demonstrates that developing accurate analysis capability need not be expensive or overly time-consuming. The purpose of this study is to understand the main factors that influence sparkling beverage sales in Southern Thailand and to find the annual per capita consumption value rate. The study investigates the associations between product preference with respect to flavour, branch location and socio-demographic factors that are useful for planning and decision making. Our study aims are to find a suitable statistical model to describe the variation in quarterly sales data of a major sparkling beverage company in southern Thailand reported from 2000 to 2006.

Thailand is divided into four geographical regions. The southern region occupies about 14% of the total land area. There are 14 provinces in southern Thailand with a total area 71,798 square kilometers. Sparkling beverages are traditionally popular products in the south, although some consumers prefer more healthy beverages. Sparkling beverages are also primarily used as mixers for consumption with alcoholic drinks.

2. Methods

The outcome in this study is sales revenue. Determinants are flavour, branch location and quarterly period. Flavours were identified for five types of products, namely "Cola flavour", "Orange flavour", "Red flavour", "Green flavour" and "Lime flavour". Two types of returnable packages (10 ounce and 1 liter) and three types of non-returnable packages (1.25 liter, 2 liter and 325 ml) were identified. In our study area 20 branches were sampled from 14 provinces. Provinces with more than one branch were Surat Thani, Nakorn Sri Thammarat, Phang-nga, Chumphon and Narathiwat. Branch locations were grouped into 3 areas (Muslim, tourist and normal areas) based on product preferences and consumer behavior in each branch location.

Data for each month were available in computer files with records for sales revenue separated by flavour and branch location. Records from years 2000 to 2006 were stored in a MySQL database. SQL was used to create sales revenue data in baht by month, flavour and branch location. All graphical and statistical analyses were performed using R software (R Foundation for Statistical Computing, Vienna, Austria).

This quantitative research focused on using statistical graphics and statistical models, including linear regression, for analyzing sales revenue and its dependence on flavour, branch location and quarter. Sales revenue data generally have positively skewed distributions so it is conventional to transform them by taking logarithms. Cells with zero counts were adjusted to avoid the problems of taking the logarithm of zero.

Linear regression models of log-transformed sales revenue per 1000 population were used to analyze the sparkling beverage sales per capita revenue and to study the main factors including product preference with respect to flavour and packaging.

The annual per capita consumption rate was computed using sales revenue divided by the number of years and population obtained from the 2000 Population and Housing Census of Thailand.

The Southern Thailand market can be grouped by branches location and consumer preferences into tourist area, Muslim area and normal areas as shown in Figure 3.

3. Statistical model

The simplest linear model takes the additive form

$$Y_{ijtk} = m + b_i + f_j + a_t + q_k, \tag{1}$$

where Y_{ijtk} is the natural logarithm of the quarterly revenue in 1000s of baht per 1000 population for branch i (b_i), flavour j (f_j), and quarter k (q_k) of year t (a_t). In this formula m is the overall mean of the log-transformed quarterly revenue.

Since this additive model does not allow for different flavour preferences in different regions, we also consider a more general model of the form

$$Y_{ijtk} = m + c_{ij} + a_t + q_k \tag{2}$$

In this model c_{ij} is an interaction between branch and flavour (c_{ij}). Generalizing further, we also consider the model

$$Y_{ijtk} = m + c_{ij} + d_{it} + q_k \tag{3}$$

Model (3) thus allows for interactions between branch-flavour (c_{ij}) and branch-year (d_{it}).

These statistical methods are described in further detail by Venables and Ripley (2002, chapter 6). After fitting the models, we plotted confidence intervals for parameters after back-transforming so that the parameter estimates were expressed in

terms of the original data, that is, in 1000 baht per 1000 population. To do this, it was necessary to incorporate an additional scale parameter for each factor to ensure that the mean revenue associated with each factor based on the fitted model matched the overall observed mean revenue.

4. Results

Figure 4 shows the overall distribution before and after transforming the data by taking natural logarithms of sales revenue per 1000 population. The sales revenue data were symmetry with homogeneously of variance. This log-normal distribution can be used to provide an estimate of consumption rate in each quarter.

From Table 1, the results indicate that Model (3) is the best regression model to use in this study because it has the highest r-squared value, which was very close to 1.

From Figure 5, it is clear that the residual plot of Interaction Model (3) was the most linear.

Figure 6 shows plots of sales revenue in baht (left panel) and per 1000 population (right panel) versus fitted value. The model predicts the proportions in the 2800 cells very well.

Table 2 shows an annual per capita consumption rate in each branch location compared with population in the corresponding catchment area. The average annual per capita consumption rate was 297 baht for the total population of 8.087 million residents.

Figure 7 shows sales revenue per 1000 population by branch location and flavour. Samui (Sm) and Phuket (Pk) branches had higher rates than the other branches, while Nakhon Sri Thammarat (Nk) branch had the lowest consumption rate.

Figure 8 shows sales revenue of each flavour grouped by branch location and consumer preferences for Muslim, tourist, and normal areas. Areas with different proportions of their population being Muslim, had different beverage preferences due to consumers preferring colour and lime products.

5. Discussion and conclusion

The log-transformed quarterly sales revenue trends can be modeled using linear regression. From this model, we found that quarterly effects and interactions associated with flavour-branch and branch-year are main factors that influence the sparkling beverage sales in southern Thailand.

We also found that the average annual per capita consumption value rate was 297 baht for the population of 8.087 million residents in southern Thailand. The annual per capita consumption rate can be used to find market opportunities which lead to greater share and growth in each area. Samui and Phuket branches, both tourist areas, had higher consumption rates than other branches. This may be because most tourists prefer these drinks over other types, are familiar with sparkling beverages. Nakhon Sri Thammarat (Nk) branch had the lowest consumption rate, thus this branch has an opportunity to increase their sales growth. This province also has the highest population in southern Thailand. A company would need to do the market development including more activities or promotions to drive volume and sales growth rate.

We studied the associations between product preference with respect to flavour, branch location and socio-demographic factors by grouping branch location into 3 types, namely tourist, Muslim and normal areas. Areas with Muslim population had different preferences compared with other areas because the Muslims prefer cola products and colour flavours (orange, red, green) of sparkling beverage as well. This is very useful fact findings for managers to know where is the right place to push each product in and to understand how to create more sales into each area.

Other studies have also used linear regression analysis (McGartland et al, 2003; Probart et al, 2006., and Bureau, 2007) but their studies did not take the location into account and did not include any analysis on per capita consumption rate. In our study, linear regression models were applied to both sales revenue and consumption rate analysis with application to a case study of sparkling soft drink products in southern Thailand. The statistical model used in this study was very suitable to answer our research questions and the results are very useful for managers to understand their performance, consumers needs in each branch location and opportunities to boot up the sales growth in each areas. Having the best model, managers can provide a useful basis for sales analysis, incorporating the results into their company plans and strategies. However, further study in this data will be expanded to more complex business forecasting.

Acknowledgements

References

Bureau. (2007). Indian Food and Beverages Forecast (2007-2011). [Online] Available: http://www.bharatbook.com/detail.asp?id=50059 (January 12, 2008)

Cox, D.R. & Wermuth, N. (1996). Multivariate Dependencies-Models, Analysis and Interpretation. London: Chapman & Hall.

Goktolga, Ziya Gokalp., Gunduz, Orhan. (2006). The Analysis of Socio-Demographic Factors Affecting Decrease in Consumption of Chicken Meat Because of Avian Influenza (Bird Flu) in Turkey: Case of Multiple Bounded Probit Model. *Journal of Applied Sciences Research*, 2(12), 1350-1354.

Helasoja, Ville et al. (2007). The sociodemographic patterning of drinking and binge drinking in Estonia, Latvia, Lithuania and Finland, 1994-2002. BMC Public Health, 7, 241.

Higgins., Richard, S., Kaplan., David, R.and McDonald., Michel, J. (2005). Residual Demand Analysis of the Carbonated Soft Drink Industry. [Online] Available: http://www.springerlink.com/content/h7n6w33821up1w24 (July 13, 2007)

Kleinbaum, D.G., Kupper, L.L., Muller, K.E., & Nizam, A. (1998). Applied regression analysis and multivariable methods, 3rd ed, Duxbury Press, Pacific Grive, C.A.

McGartland, C., Robson, P.J., Murray, L., Cran, G., Savage, M.J., Watkins, D., Rooney, M., & Roreham, C. (2003). Carbonated Soft Drink Consumption and Bone Mineral Density in Adolescence: The Northern Ireland Young Hearts Project. *Journal of bone and mineral research*, 18, 1563-1569.

National Statistical Office. (2002). The 2000 population and housing census, southern region. Bangkok: Statistical Data Bank and Information Dissemination Division, National Statistical Office.

Nichols, Donald. (2005). The Simple Regression Model. [Online] Available: http://dnichols.wustl.edu/413sp2005/simple%20regression%20sp2005.pdf (June 5, 2007)

Probart C., McDonnell E., Hartman T., Weirich J., & Bailey-Davis L. (2006). Factors associated with the Offering and sale of competitive foods and school lunch participation. *Journal of the American Dietetic Association,* 106, 242-247.

R Development Core Team. (2007). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statisical Computing, Vienna.

Syspro. (2007). Sales Analysis. [Online] Available: http://www.syspro.com/PDF/sales_analysis.pdf (March 24, 2008)

Venables, W.N., & Ripley, B.D. (2002). Modern Applied Statistics with S. Springer. Queensland.

Venables, W.N., & Smith, D.M. (2004). An Introduction to R. [Online] Available: http://cran.r-project.org/doc/manuals/R-intro.pdf (March 30, 2007)

Table 1. Result of fitting linear regression model

Model	R^2	df	RSS
1	0.913	2767	211.9
2	0.959	2691	100.7
3	0.973	2577	65.1

Table 2. Population (2000 census) of sparkling beverage branch locations and annual consumption in Southern Thailand

Branch	Symbol	Population (x 1000)	Annual consumption (Baht/population)
Nakhon Sri Thammarat (Exclude: Thung Song district)	Nk	1,375	92
Hat Yai	Hy	1,258	267
Surat Thani (Exclude: Phunphin and Samui district)	Sr	746	183
Narathiwat (Exclude: Kolok district)	Nv	600	69
Pattani	Pn	599	92
Trang	Tr	597	142
Phatthalung	Pl	501	84
Yala	Yl	418	172
Chumphorn (Exclude: Lang Saun district)	Cp	380	76
Krabi	Kb	338	169
Phuket	Pk	250	697
Satun	St	219	144
Phang-nga (Exclude: Takuapa district)	Pg	192	105
Ranong	Rn	162	208
Thung Song	Ts	148	325
Phunphin	Pp	91	428
Lang Saun	Ls	69	291
Kolok	Kl	65	386
Takuapa	Tk	44	583
Samui	Sm	35	1,424
Total		8,087	
Average annual consumption (Baht per population)			297

Determinants

- ▪ Branch location
- ▪ Flavour
- ▪ Time
  - ○ Year
  - ○ Quarter
    - ○ Month
  - ▪ Environmental factors

Outcomes
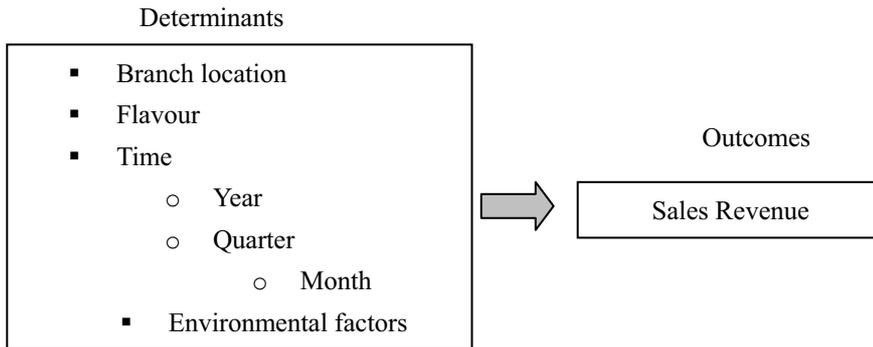
Sales Revenue

Figure 1. Data path diagram

Collected Raw Data

Transforming and coding the data -> MySQL Database, Excel

MySQL

Separate data into each package and order by month, flavour and branch

Excel

Get rid of the outliers

R

Graphical and statistical analysis
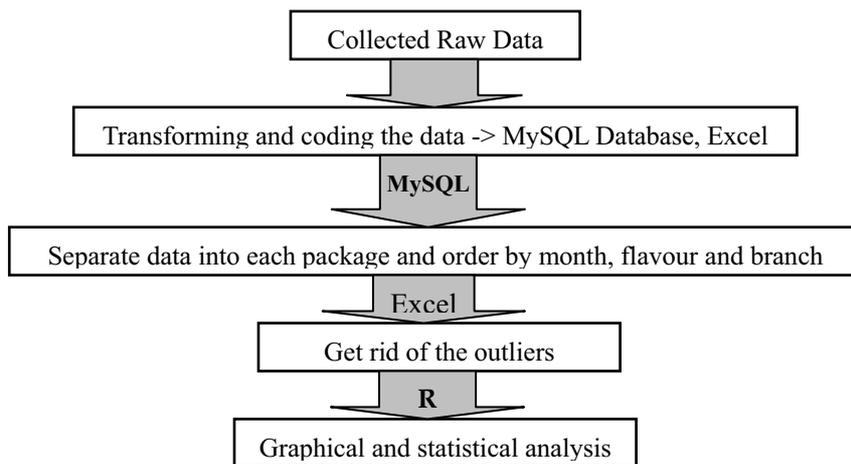
Figure 2. Data collection and management
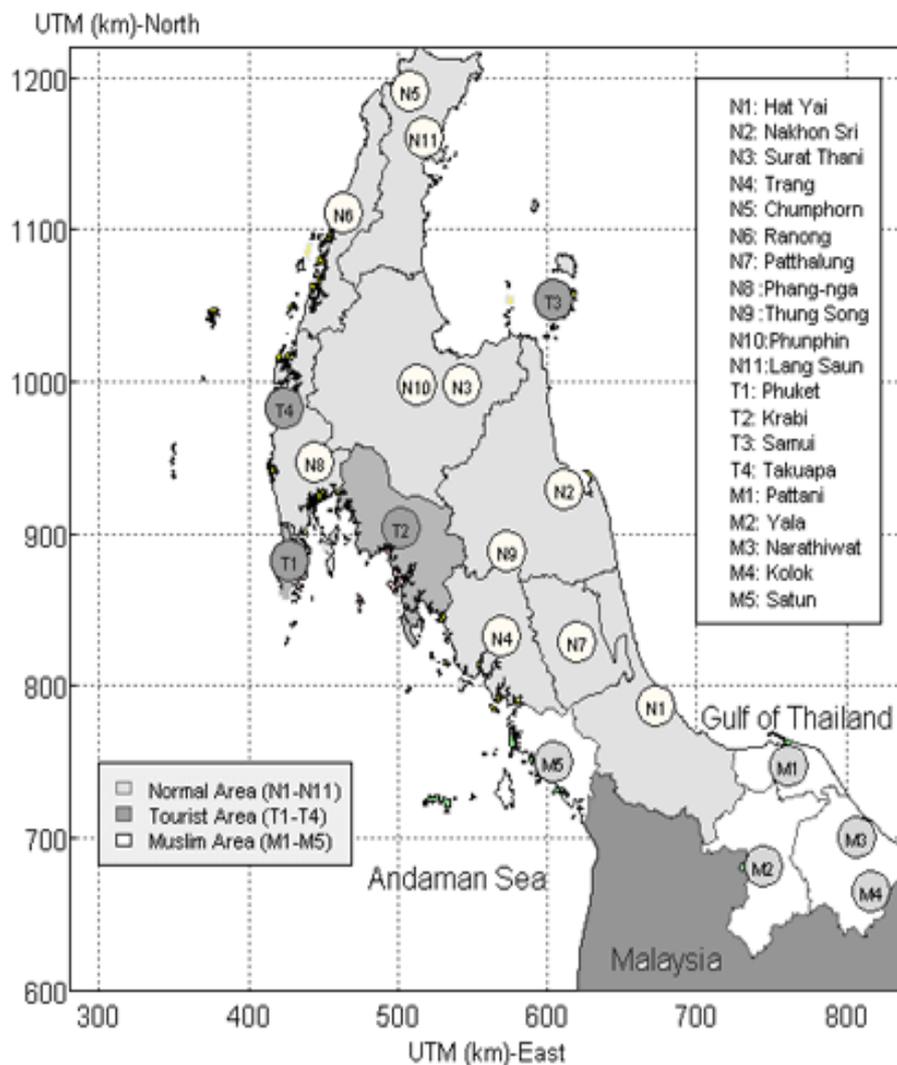
Figure 3. The locations of the branches and their definitions as

tourist, Muslim or normal areas
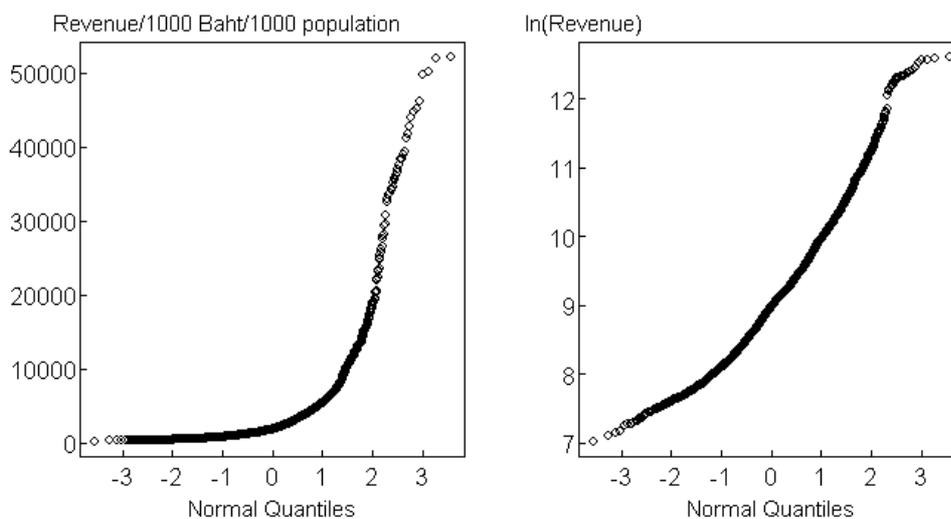
Figure 4. Sales revenue distribution before and after transforming to

ln (Baht/1000 population)
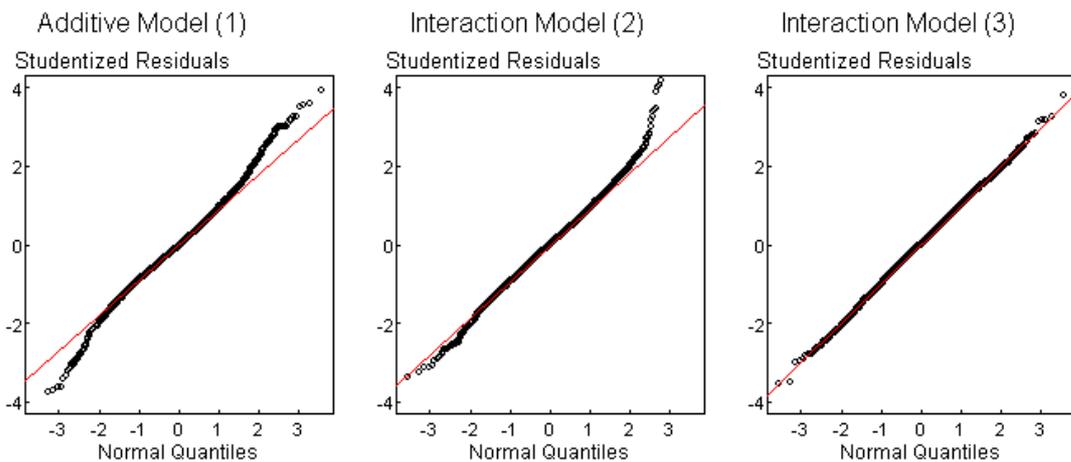
Figure 5. Comparison between residual plots for three models
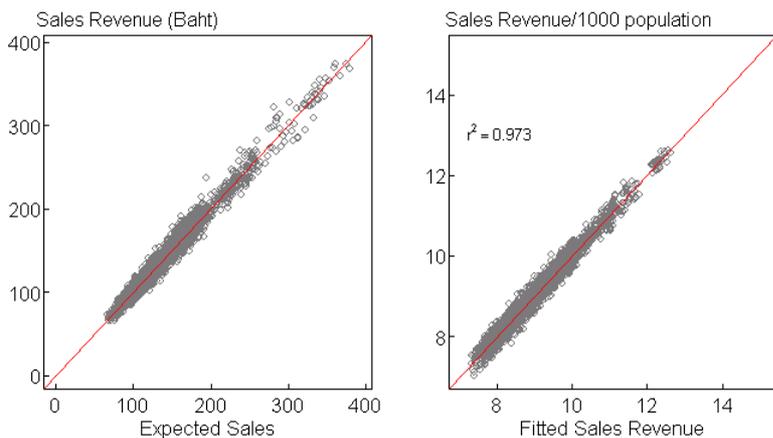
Figure 6. Plots of sales revenue and fitted values for Model (3)
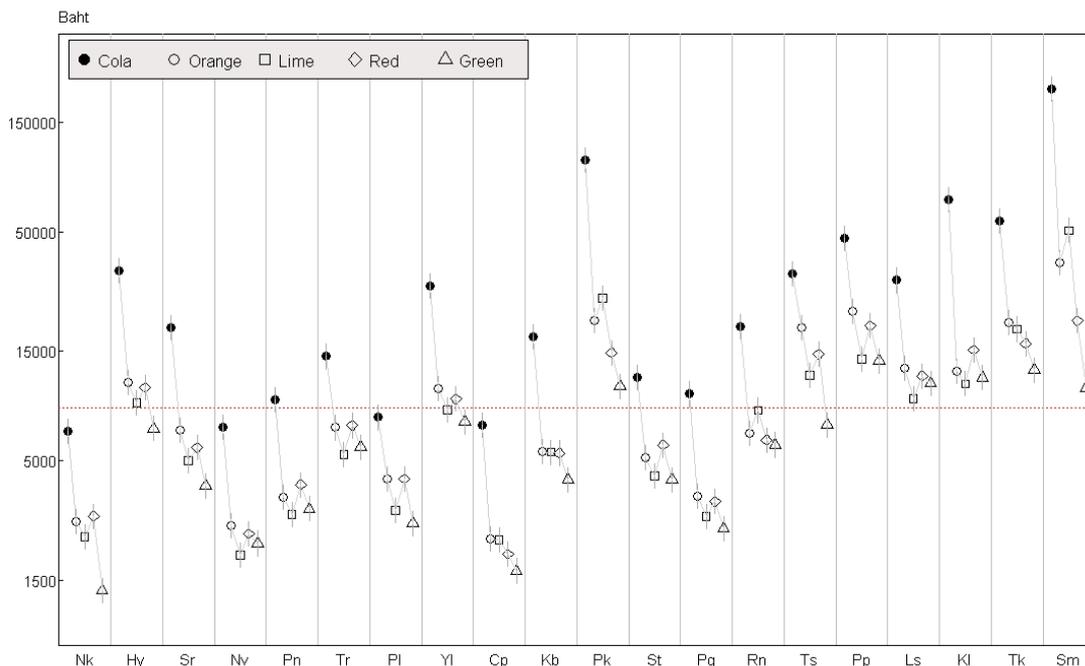
Figure 7. Sales revenue per 1000 population by branch location and flavour
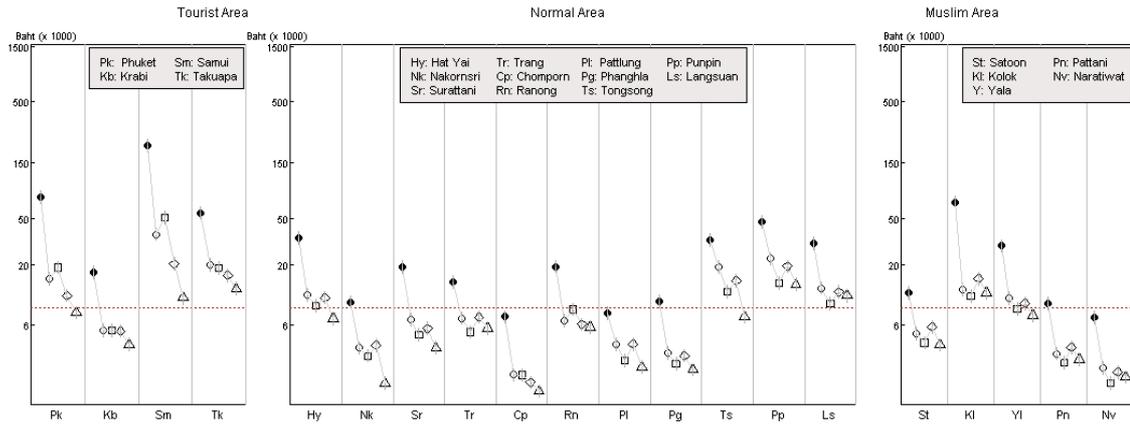
Figure 8. Sales revenue grouped by area per 1000 population