Rossmann Store Analysis

Introduction

Data Set Descriptors

A. Data set identity
        Title: Database of Rossmann stores

B. Data set description

Abstract:
The purpose of this data set was to compile daily sales and number of customers information for Rossmann's 1115 stores located across the Europe and see what patterns arise when variables such as competition and promotion are considered against the sales or the number of customers each store had for a that particular day. We were interest in seeing what variables have the positive of negative effect in changing sales or number of customers, and measuring correlation coefficient/r^2 values to see how big of an impact each variable has on the sales and the number of customers.

**B. Variable definitions**

**-Store**

| Variable name | Variable definition | Units | Storage type | Range numeric values (-999 not incl.) | Missing value codes |
|---|---|---|---|---|---|
| store_num | 1115 different stores spread around the European area. | NA | Integer | 1 to 1115 | NA |
| Store_type | There are four types of stores and we assumed that these fall under Remedies, Beauty, Foods, and House. These stores focus more on providing these products. | NA | Character | NA | NA |
| competition_distance | Distance between a closest competition store. | Meters | Integer | 20 to 75860 | NA |
| competition_op | Month in a single year | Month | Integer | 1 to 12 | NA |

| Variable name | Variable definition | Units | Storage type | Range numeric values (-999 not incl.) | Missing value codes |
|---|---|---|---|---|---|
| en_since_mont h | when the competition store opened. | | | | |
| competition_op en_since_year | Year that the competition store opened | Year | Integer | 1900 to 2015 | NA |
| comp_my | Year number/Month number of since competition opened | NA | Character | NA | NA/NA |
| promo2 | Whether the store holds promotion type 2 or not. | NA | Integer | 0 or 1 | NA |

-Train

| Variable name | Variable definition | Units | Storage type | Range numeric values (-999 not incl.) | Missing value codes |
|---|---|---|---|---|---|
| store_num | 1115 different stores spread around the European area. | NA | Integer | 1 to 1115 | NA |
| day_of_week | Day of the week | Week | Character | Mon, Tues, Weds, Thurs, Fri, Sat, Sun | NA |
| my | Year and month of sales | NA | Character | NA | NA |
| date | Date of sales | NA | Character | NA | NA |
| sales | Sales of the day | € | Integer | 0-41551 | NA |
| customers | Number of customers of the day | NA | Integer | 0-7388 | NA |
| promo1 | Status of the promotion | NA | Integer | 0 or 1 | NA |