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	1115 different stores spread around the European area.	NA	Integer	1 to 1115	NA
StoreType	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
Assortment	There are four types of assortments within a single store. There was not a representation of	NA	Character	NA	NA

	these variables available.				
CompetitionOpenSinceMonth	Month in a single year when the competition store opened.	Month	Integer	1 to 12	NA
CompetitionOpenSinceYear	Year that the competition store opened	Year	Integer	1900 to 2015	NA
Promo2	Whether the store holds promotion type 2 or not.	NA	Integer	0 or 1	NA
Promo2SinceWeek	Calendar Week when the store started participating in promo2	Week	Integer	NA	NA
Promo2SinceYear	Year when the store started participating in promo2	Year	Integer	2009-2015	NA
PromoInterval	Consecutive intervals promo2 is started, naming the months the promotion is started anew.	Month	Character	NA	NA

-Train

Variable name	Variable definition	Units	Storage type	Range numeric values (-999 not incl.)	Missing value codes
Store	1115 different stores spread around the European area.	NA	Integer	1 to 1115	NA
Date	Date (Year-Month-Day)	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
Open	Status of the stores	NA	Integer	0 or 1	NA
Promo1	Status of the	NA	Integer	0 or 1	NA

	promotion				
StateHoliday	Status of the state holiday	NA	Integer	0 or 1	NA
SchoolHoliday	Status of the school holiday	NA	Integer	0 or 1	NA