

Newspaper Kiosk Revenue Increase Strategies

Phase 1: Business & Data Understanding

true

true

26/01/2021

Abstract

This document details the process of business and data understanding. It points out data cleaning issues, showcases descriptive and exploratory data analysis, variable transformations and formulates insights gained from them.

Data Cleaning

1. We converted the store_name variable which was stored as numeric value to strings. We did this in case of all 3 dataset (turnover, cost, stores).
2. We converted nominal variables which were stored as numeric to factor variables to effectively use them later during potential regression modeling.
3. We converted some variables which were stored as character but are indeed numeric.
4. We replaced “#MISSING” values with NAs.
5. We flipped the sign of costs
6. We divided the Gross Margin and Costs variables with 10,000 to make interpretation easier.

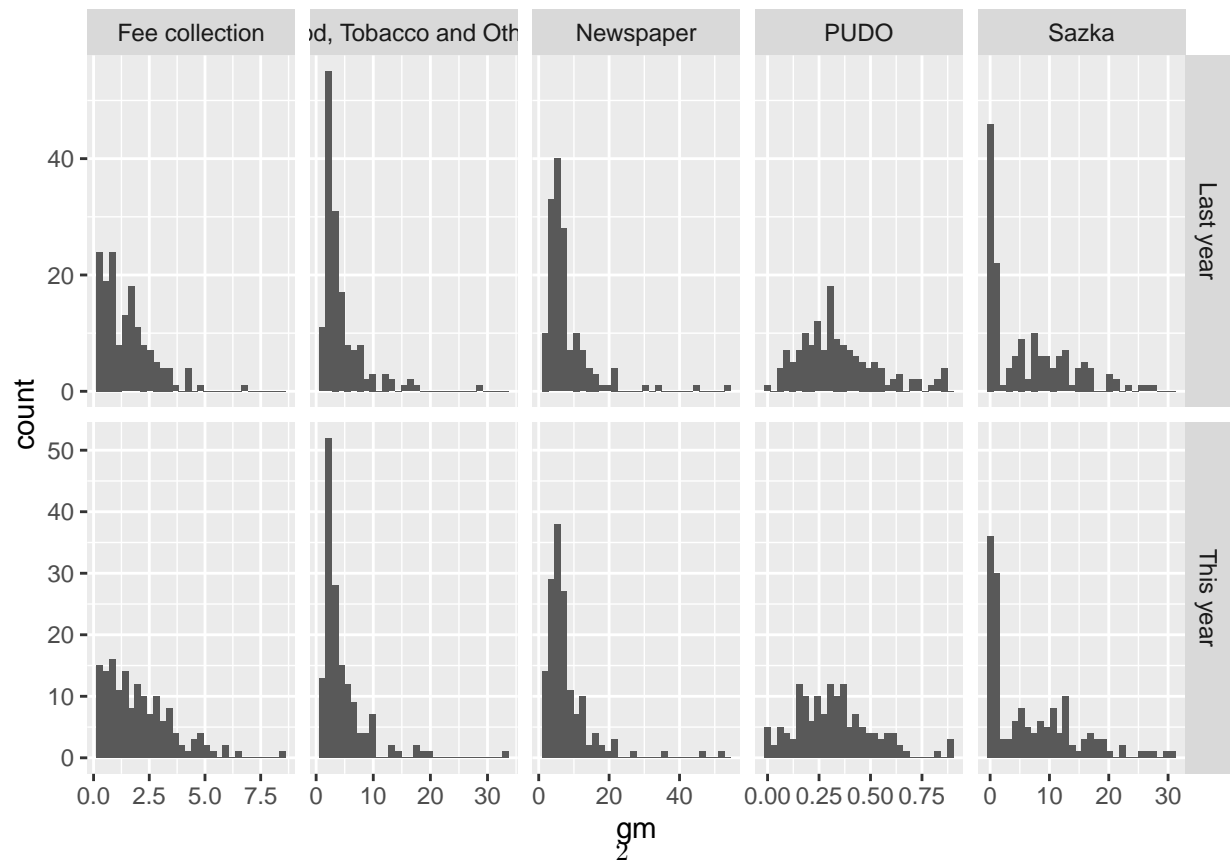
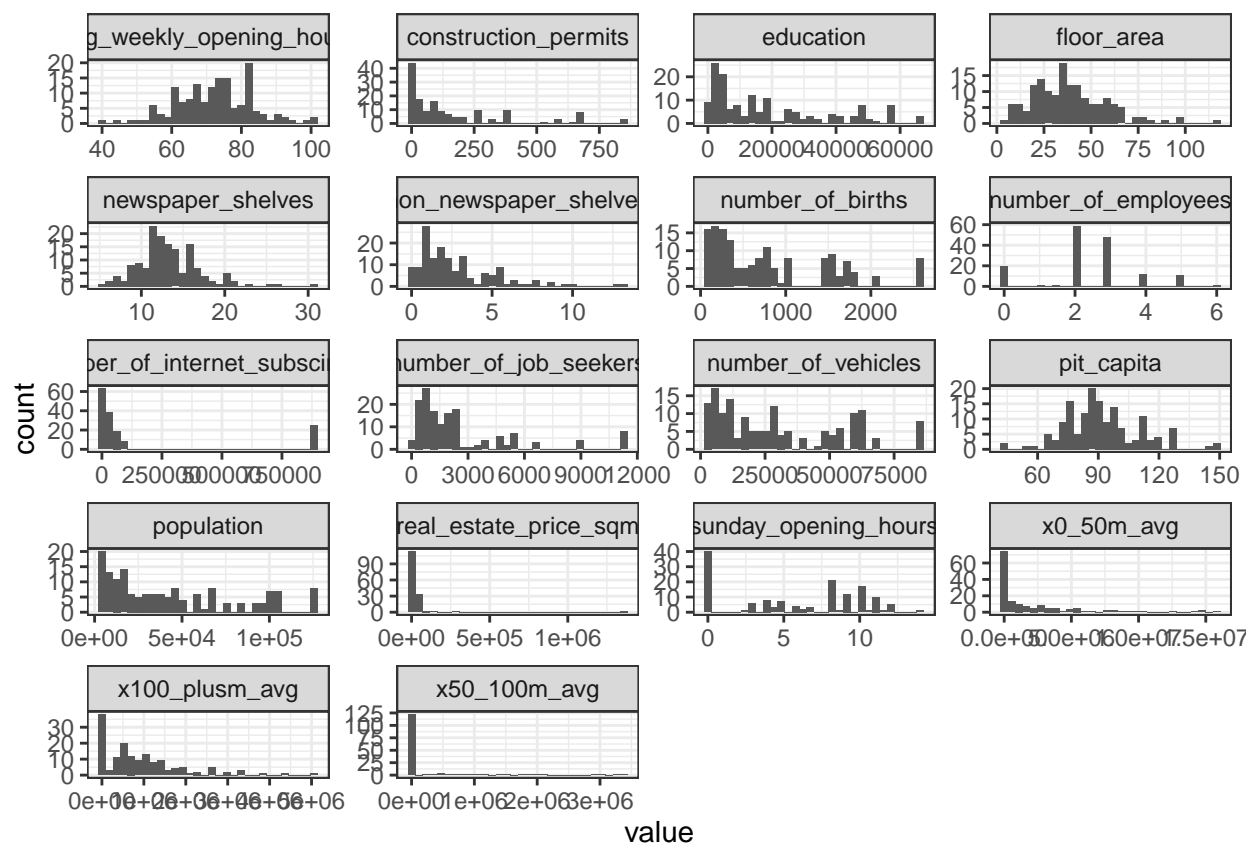
Data Wrangling

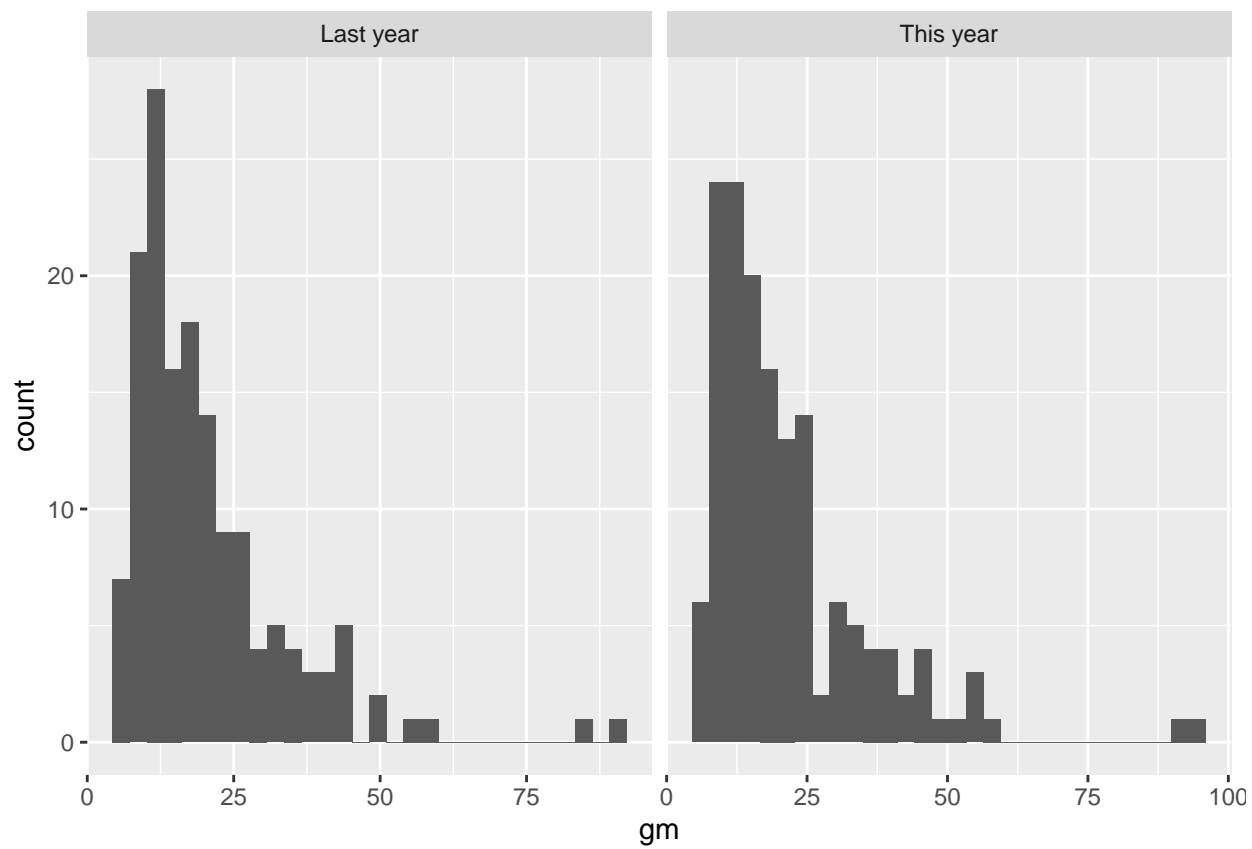
We created 3 major data tables with different aggregation level/granularity: 1. ys - one observation represents one store for one year

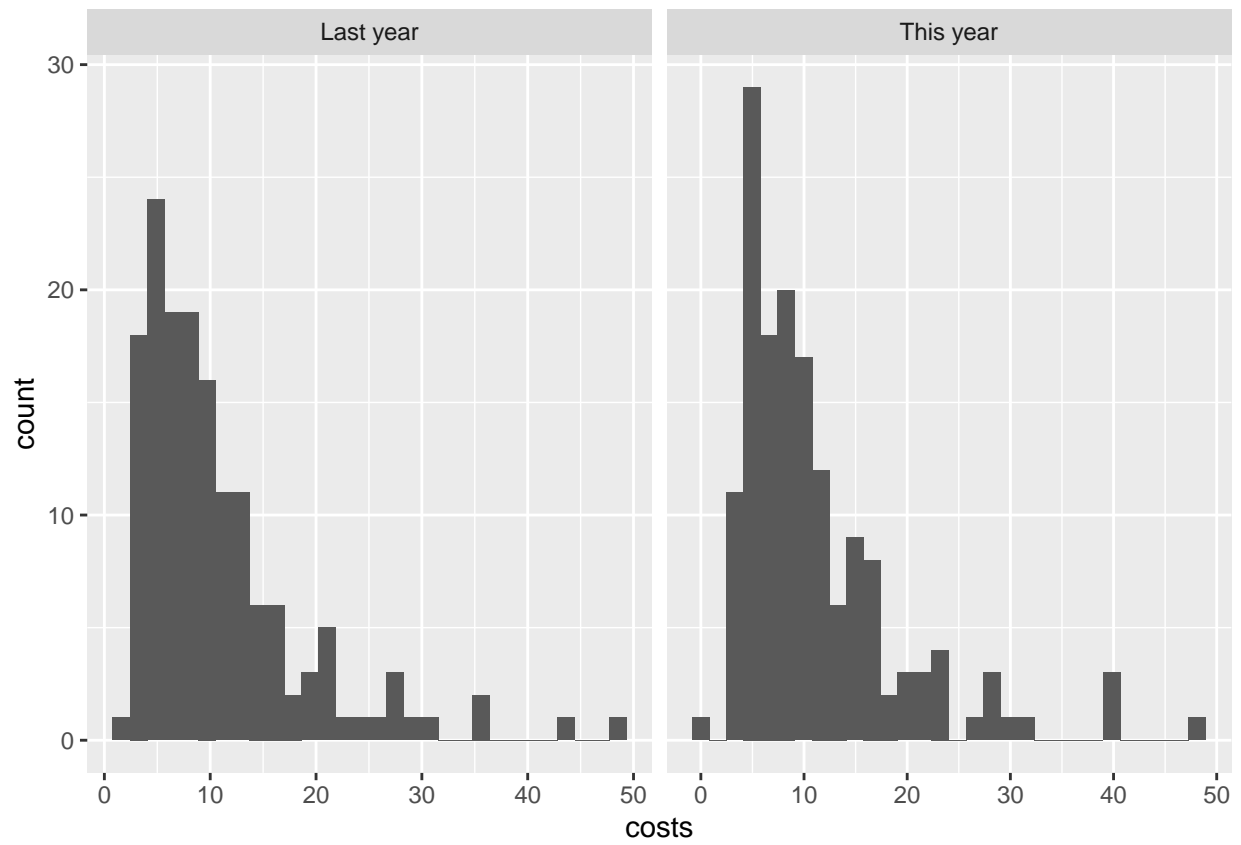
2. ysp - one observation represents one product category for one store for one year

3. yspm - one observation represents one month for one product category for one store for one year

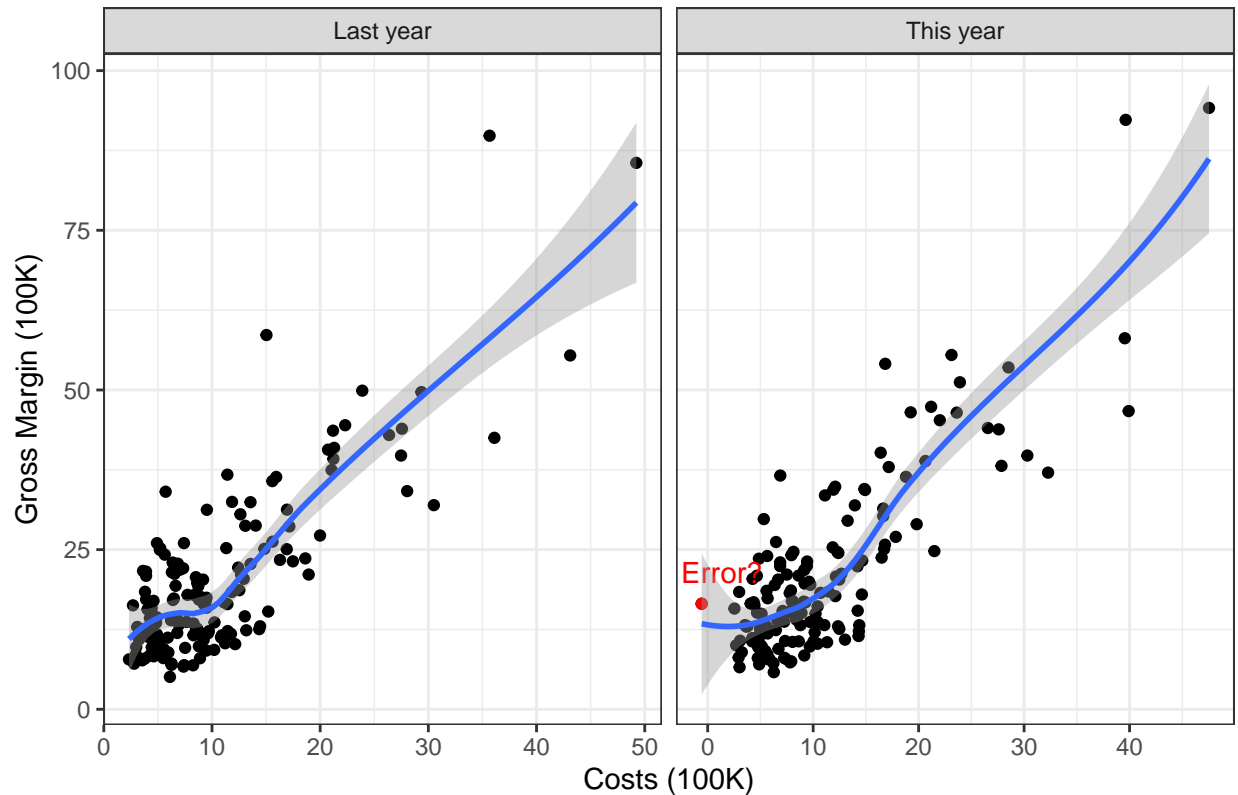
General Descriptive Statistics and Exploratory Data Analysis







Relationship between Non-COGS Costs and Gross Margin



We spotted an irregularity in the data which is very likely to be a measurement error. As it can be seen on the right-hand-side chart there is an observation with -0.57 (100K) non-COGS cost. It is unnatural for a cost to be negative.

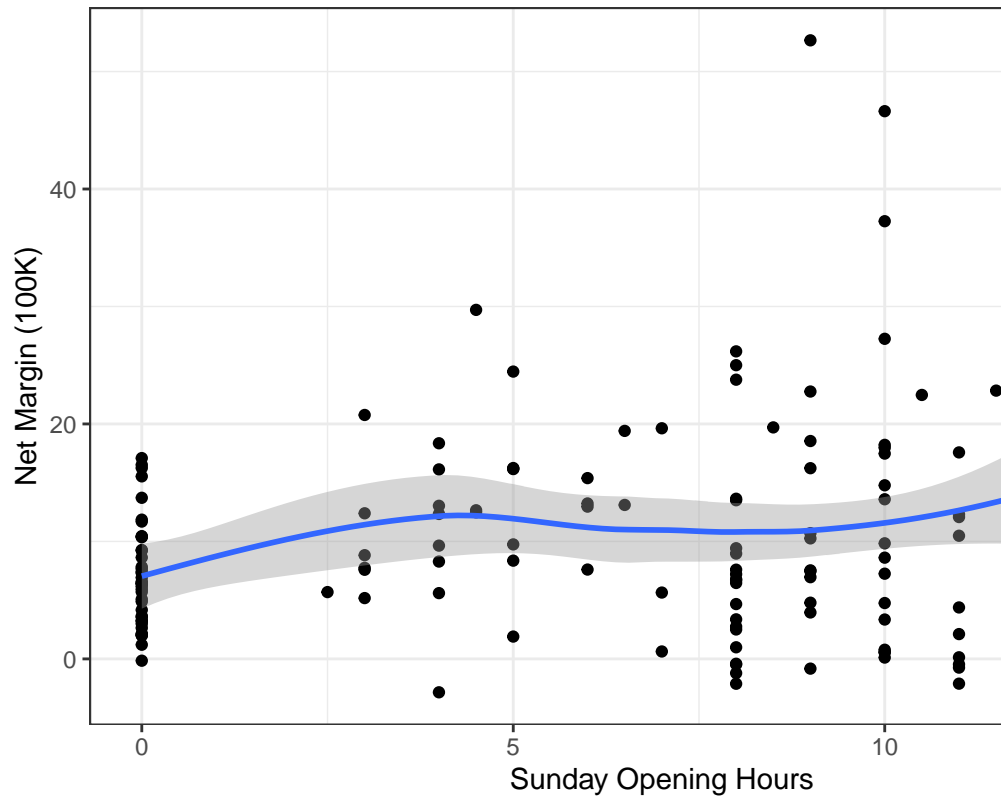
Revenue Improvement Strategies | EDA

To focus our attention to store-related patterns rather than stability in time we limited ourselves to the observation belonging to “This year”. However, it is advisable to check this stability in a later phase of the analysis.

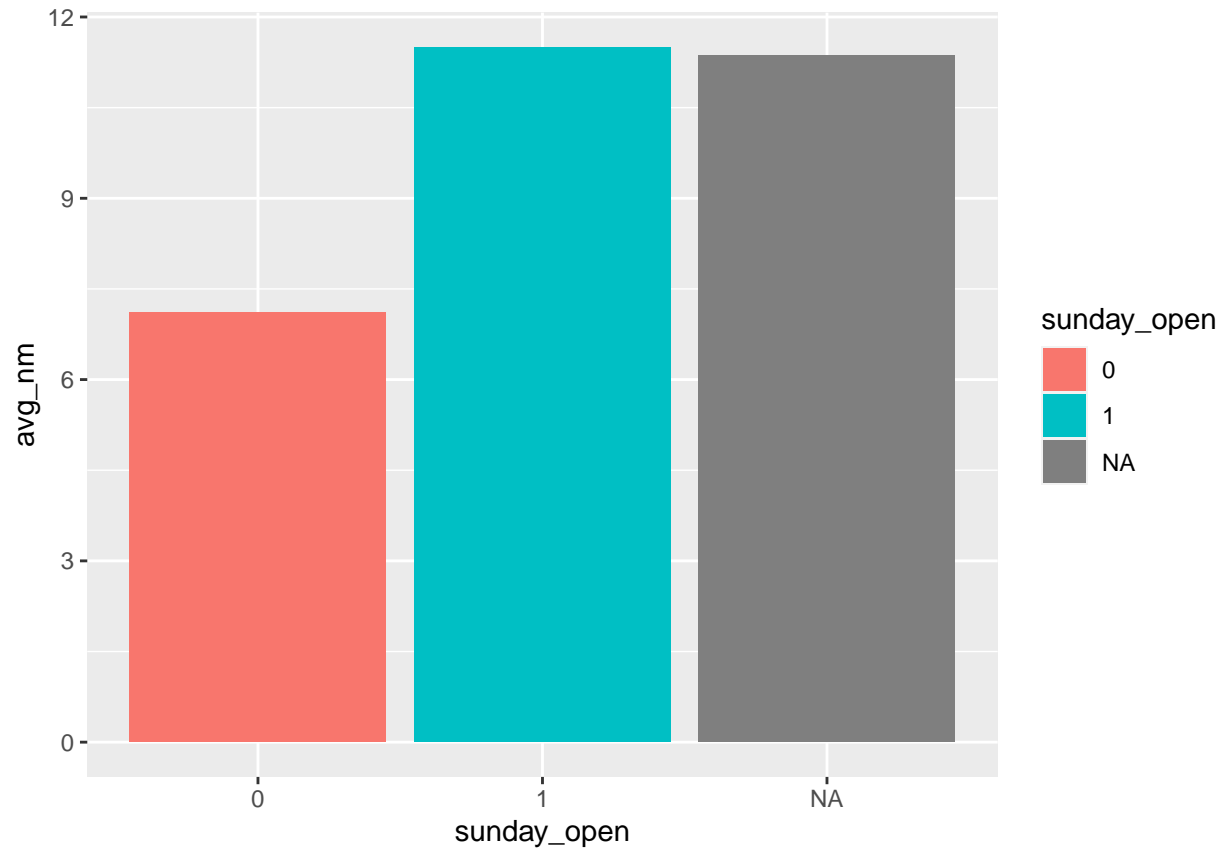
1. Weekend Opening Hours

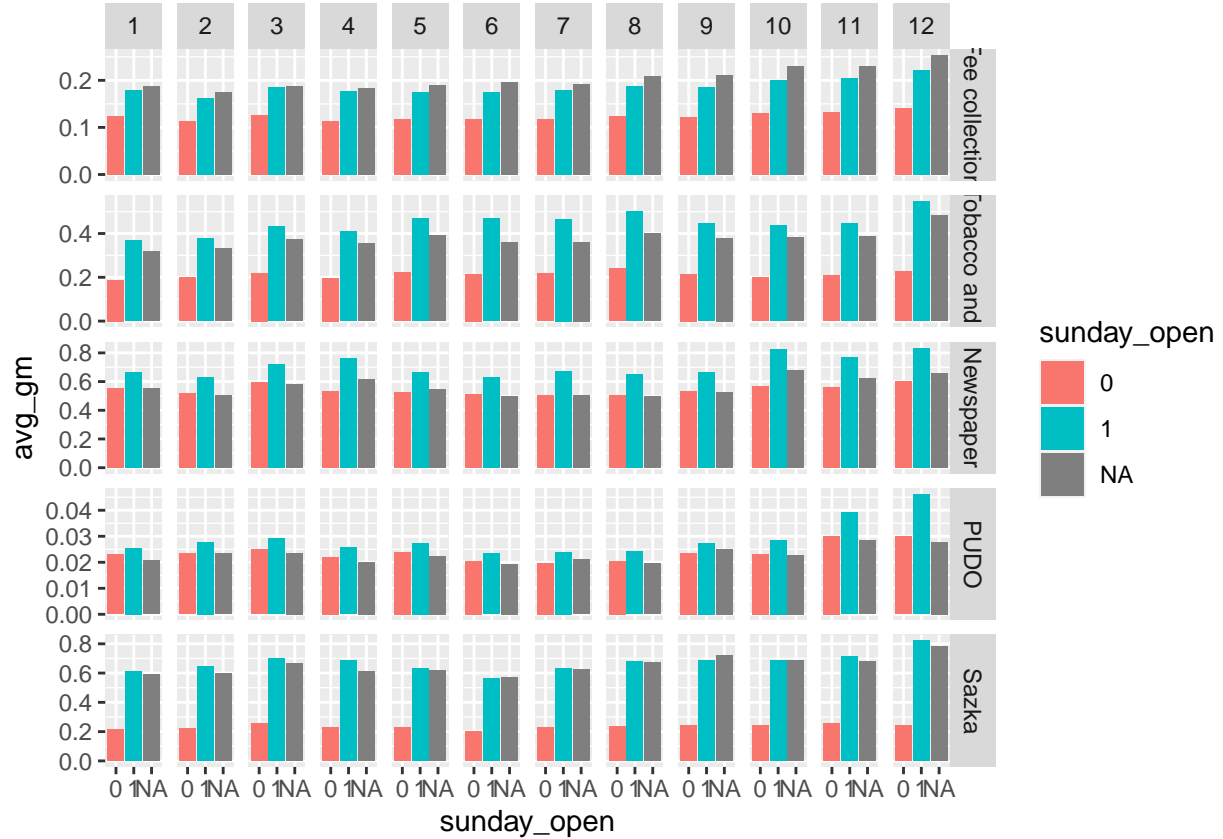
Our first strategic approach to increase revenue is to optimize which stores should be open on Sundays and which should not. In order to get closer to the solution for this we may want to find out whether being open on

Pattern of association between Net Margin and Sunday Opening Hours



Sundays results in higher Net Margin.



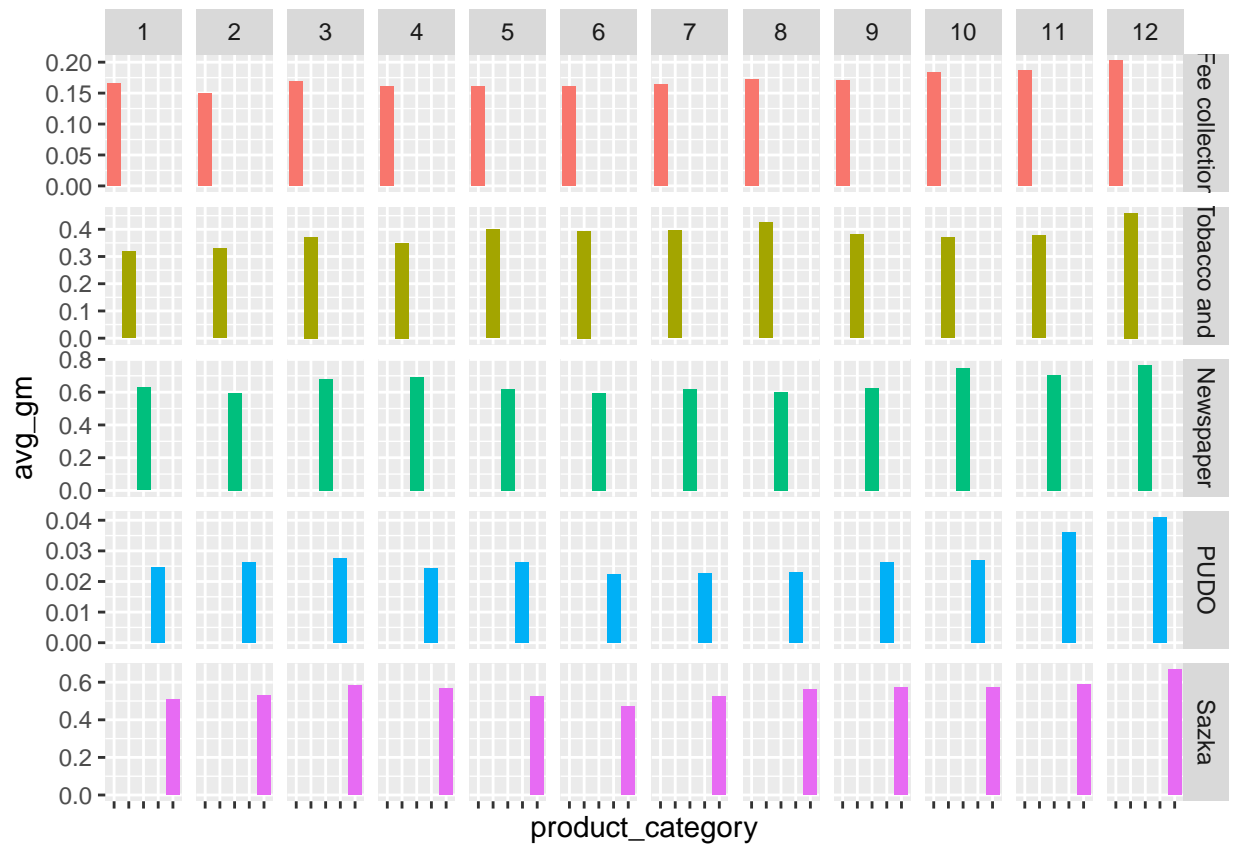


We can see based on the charts that average Net Margin is higher for stores that are open on Sundays. However we cannot be sure that this is only due to the fact that they are open on Sundays too. Stores may differ in many other aspects as well. A linear multiple regression would be helpful to say more on this.

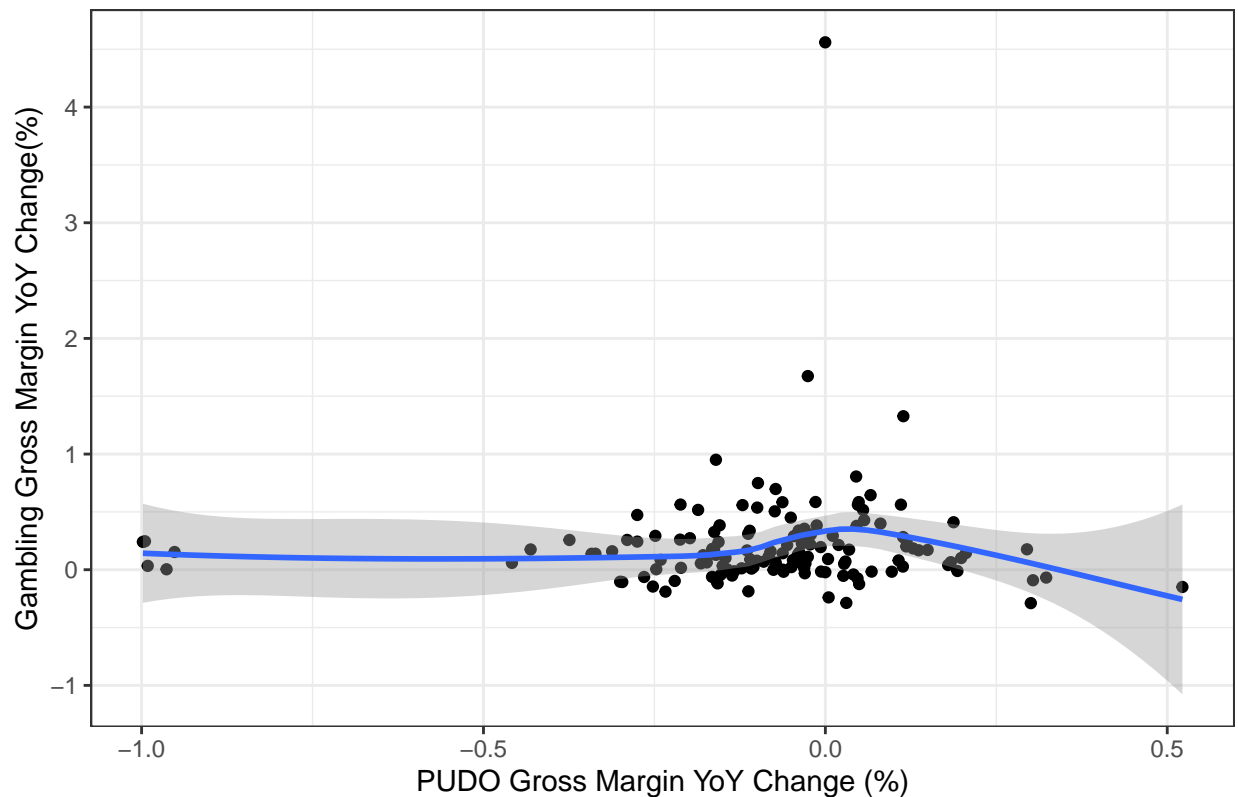
On the last chart we can see that there are significant differences in the effect of being open on Sundays to GM between various product categories which effect also changes meaningfully during the year in some cases. For example there are huge differences in gambling gross margins between stores that are open during Sunday and those that are not. In addition, this difference is even bigger in some months (Nov-Dec). This is useful to identify which stores should be open on Sundays and which should not using their revenue distribution along product categories.

2. Potential benefits of the abolition of PUDO Services

Our second proposed strategy considers the usage of PUDO services to boost revenue through the development of cross-sales. The ultimate goal would be to identify those stores where PUDO has a high cross-sales potential. Although 88% of the stores already offers PUDO services, in case of the rest it can be even more valuable to choose only those stores to offer this that are more likely to have a higher cross-sales. It also needs to be kept in mind that PUDO may decrease sales of other product categories. For instance, it is not advisable to have heavily-used PUDO services where gambling is intensive since both require quite much place and the staffs' time.



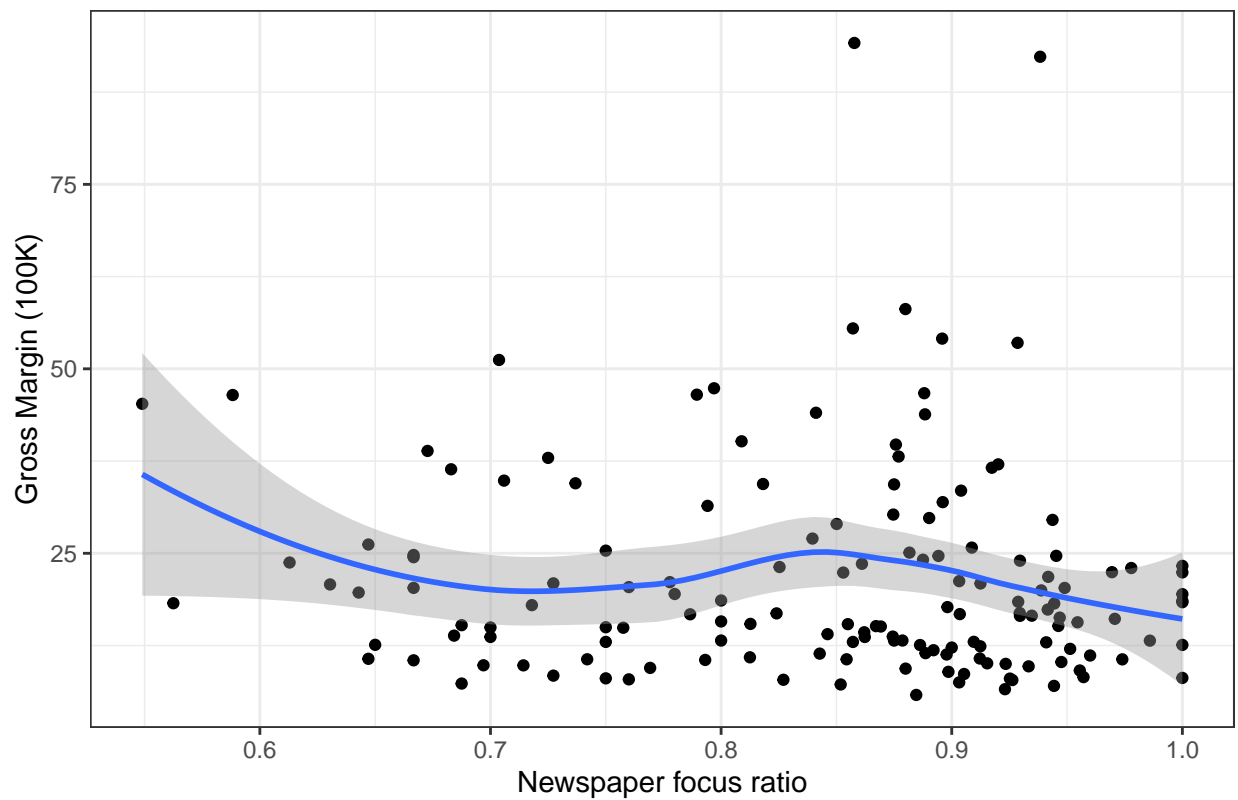
Pattern of association between PUDO and Gambling Gross Margin Change



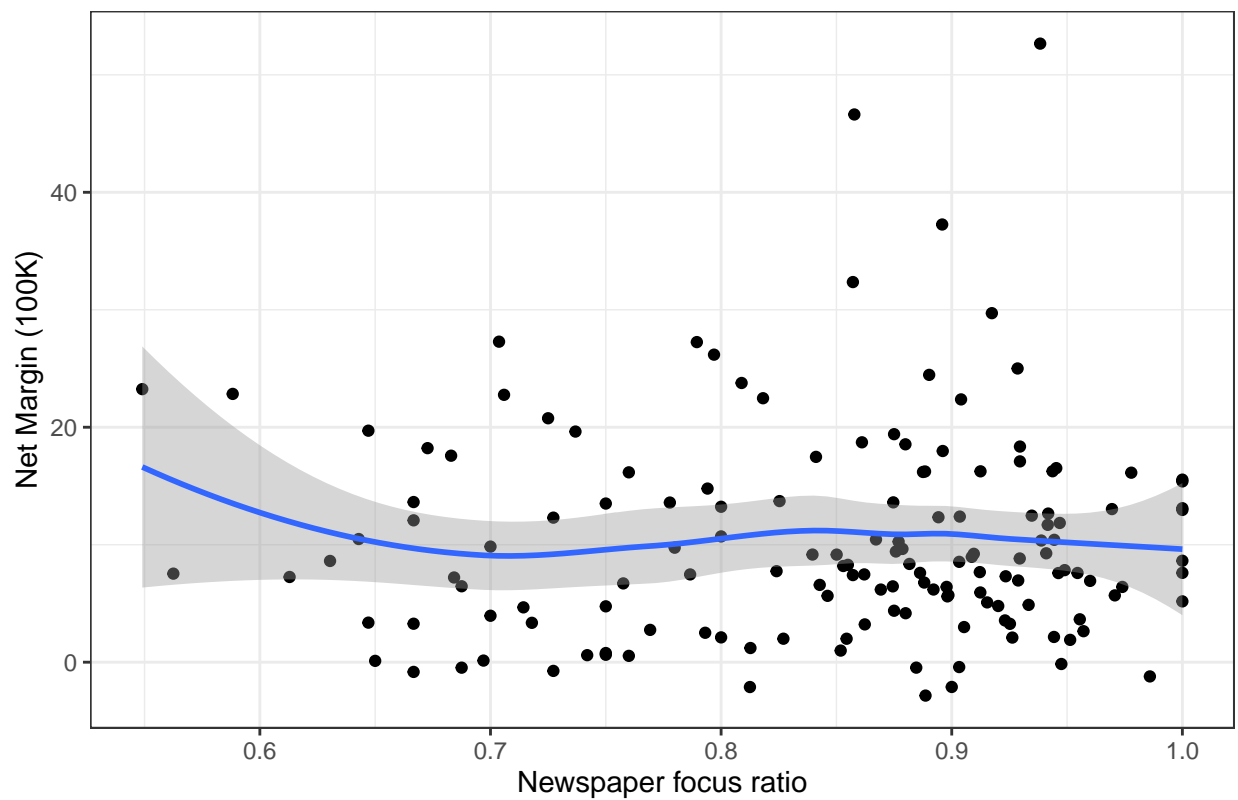
On the first graph we can see that there is significant amount of seasonality present in PUDO sales. It may help in picking the ideal date to launch a marketing campaign aimed at increasing cross-sales. On the second graph we try to explore the relationship between gambling and PUDO sales. We are interested whether there is a negative relationship between the two. In order to filter out the popularity of the stores we used relative change in Gross Margins from last year to this year. We cannot see a meaningful connection between the two but it may worth running some regression models to make sure there is no such pattern.

3. Finding better product combos/mixes

Pattern of association between Gross Margin and Newspaper focus ratio



Pattern of association between Gross Margin and Newspaper focus ratio



Preliminary exploratory analysis suggests that using this strategy will not increase revenue generation. It seems like there is no strong relationship between the store's newspaper focus and Gross Margin/Net Margin. A multiple regression would give a better idea whether there is indeed no pattern of association between the two.

4.

5.

6.

Notes