Measuring Merchant Growth

Lilly Nikkel

WHEATON COLLEGE
For Christ & His Kingdom

Mathematics and Computer Science Department

Introduction

Traditional trade merchants in Indonesia are a significant part of Indonesia's economy; 82% of Indonesia's population buy their groceries from local small and micro merchants. These merchants purchase goods from suppliers and in turn sell them to local consumers. Many of these merchants face challenges as they lack access to safe financing to stock up on supplies and as they lack resources to manage their transactions and their data digitally. The fintech company AwanTunai seeks to address some of these issues by providing safe loans and digitization resources to micro and small merchants in Indonesia. Their vision is to strengthen and prosper micro, small, and medium enterprises [1].

Problem Statement

How effective has AwanTunai been in helping merchants grow? Specifically, are merchants spending more through AwanTunai over time?

Methods

We use a dataset containing transaction records of merchants' inventory purchases from suppliers, extracted from AwanTunai's Point of Service system. The data includes the following attributes:
- order id, merchant id, supplier id
- sku id, sku name
- quantity, price (in Indonesian Rupiah [IDR])
- time stamp

We limit the data to the 58 merchants who transacted through AwanTunai during each month between August 2021 and July 2022. We then calculate the amount 'spent' on each transaction, which is the product of price and quantity; we use this measure to analyze how overall purchasing habits change over time.

We then mechanistically run an Ordinary Least Squares (OLS) regression of the 'spent' variable on 'time', for each of the 58 merchants as follows

$$\hat{y} = \alpha + \beta x_i + \varepsilon$$

where $\hat{y}$ is the predicted value of the 'spent' variable,[1] α is the intercept,[2] β is the regression coefficient, x_i is the 'time' variable, and ε is the error term. We are concerned with the β values for each of the merchants. We run this regression in 4 slightly different variations; first, we run a regression of 'spent' on time by 'day' and another regression of 'spent' on time by 'month;' we then replicate these two regressions taking the natural log of 'spent' to produce a level-log model for easier interpretation.

[1] It should be noted that this model currently is not intended to forecast future trends, but rather to analyze past trends
[2] i.e., what the 'spent' variable is at time = 0, based on the trend

Results

We take the regression coefficients (the β values) from the 58 OLS regressions, for each regression variation, and visualize the distributions of the β values. Below we have the distribution of the β values from the regression of 'spent' on 'day', next to the distribution of the beta values from the regression of 'spent' on 'month.'
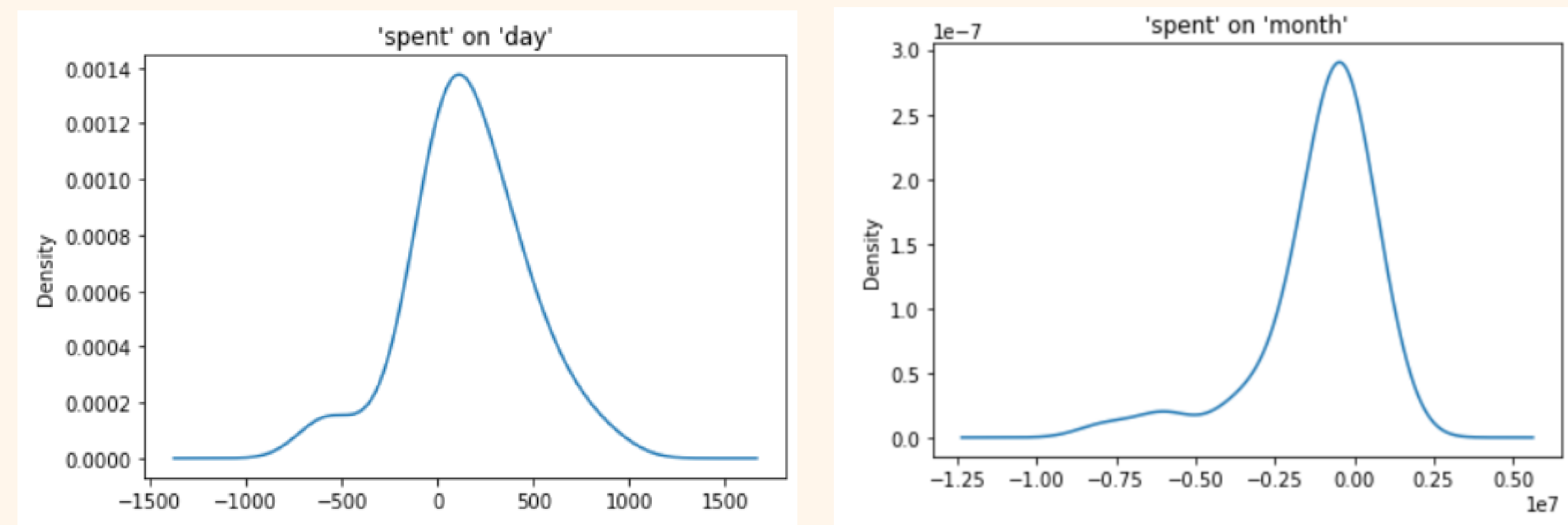
The mean of the distribution for the β values from the regression on 'day' is positive, while the mean of the distribution for the β values from the regression on 'month' is negative (to be exact, the means are 660.97 and -351416.28, respectively). The result is analogous in the regressions which use the natural log of the spent variable, denoted as 'lnspent,' shown below.
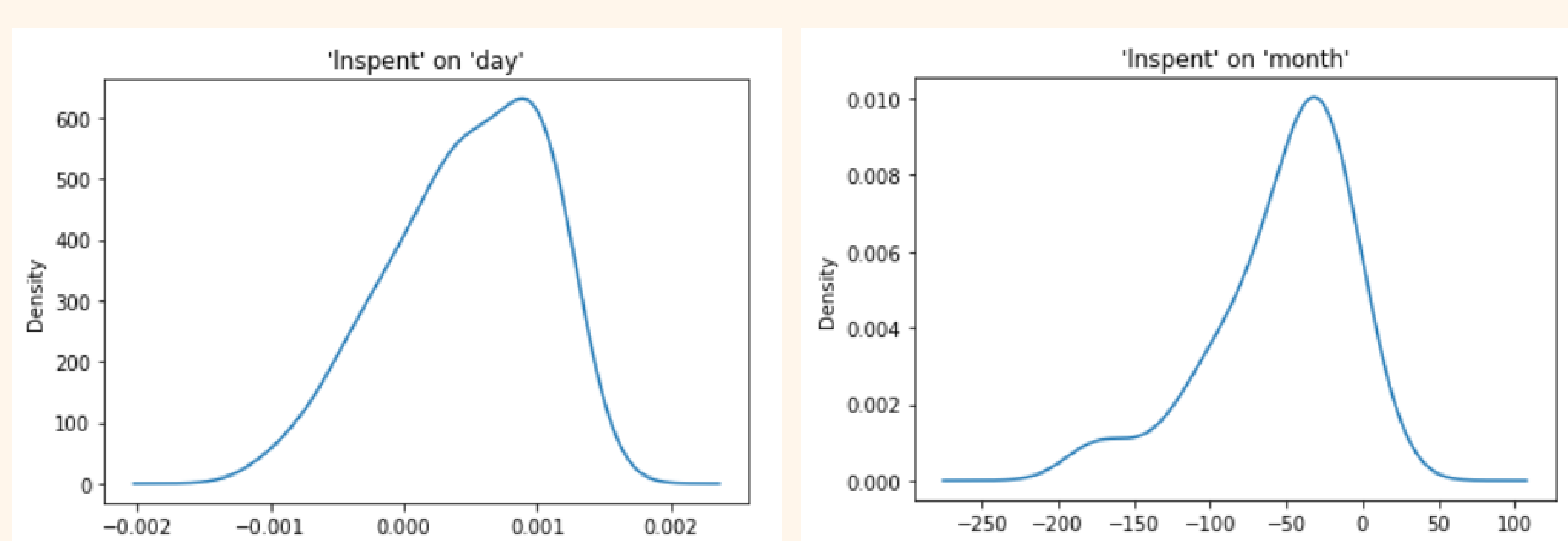
The respective means are 0.00079 and -35.3373.

Conclusion

Depending on how one groups the data, it is not super clear whether merchants that transacted through AwanTunai over the course of the August 2021-July 2022 spent more money through AwanTunai as the year progressed, based on this analysis alone. This analysis is only a first step; much more investigation is needed to understand the complex Fast-Moving Consumer Goods market.

Further analysis could explore why aggregating the 'spent' variable by month suggests a different spending trend over time; it could also investigate how external factors might have impacted transactions, such as promotions. Additionally, more analyses could be done to check assumptions regarding OLS regressions on time series data.

Case Example

Below, the individual regressions for Merchant 107 are visualized. The first is the regression of 'spent' on 'day;' the regression coefficient is -83.6195; we can interpret this as: for each additional day, Merchant 107 is likely to spend 83 IDR less. The second is the regression of 'spent' on 'month;' the coefficient is -5.512e+05; we can interpret this as: for each additional month, Merchant 107 is likely to spend 551,200 IDR less.
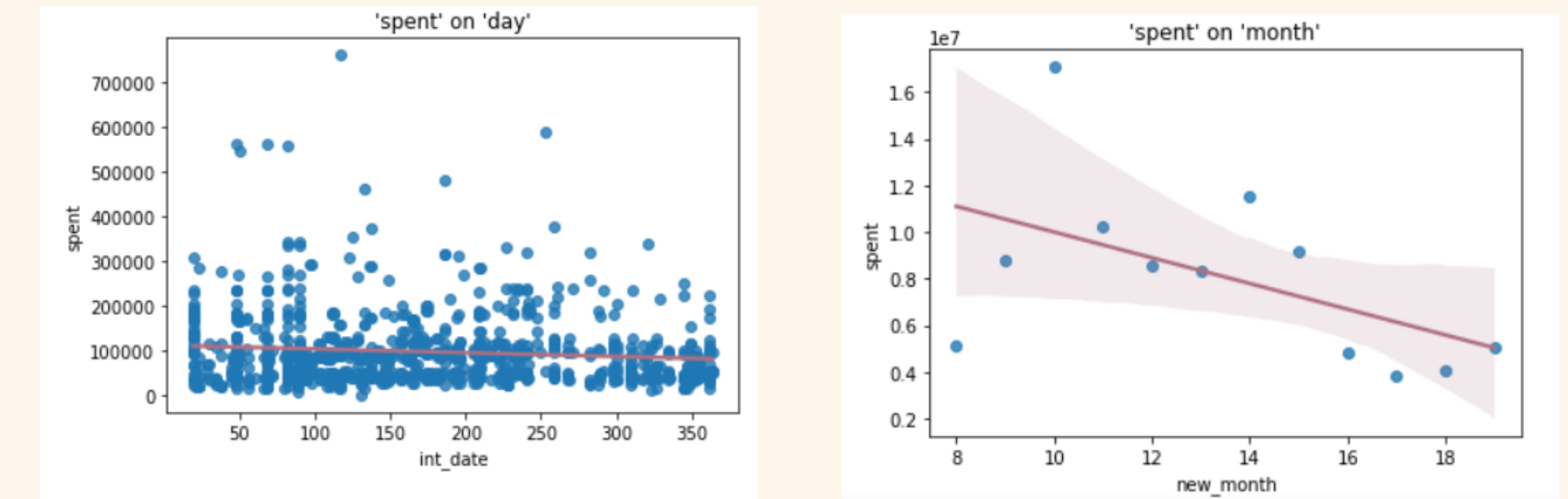
We also show the analogous regressions for the natural log of 'spent' below.
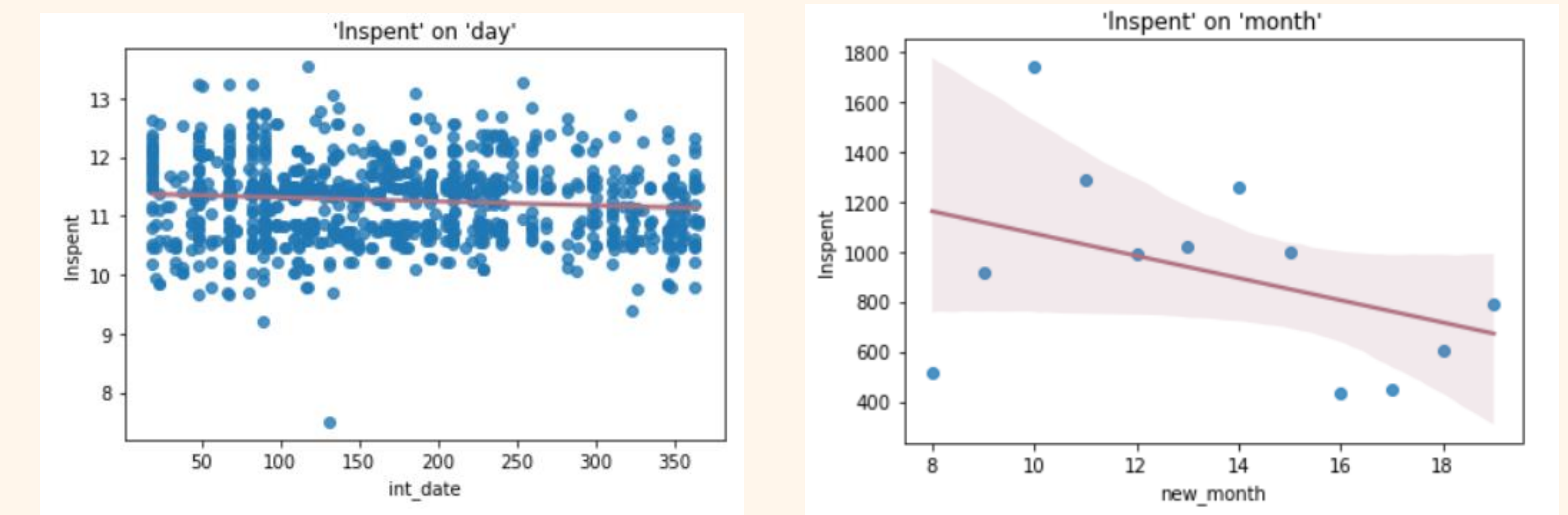
The respective coefficients are -0.0007 and -44.4896 for the regressions on 'day' and 'month' respectively. We interpret these as: for each additional day, Merchant 107 is likely to spend .07% less; and, for each additional month, Merchant 107 is likely to spend 4449% less.

In this case, there is consistency in the signs of the coefficients; the data suggest that Merchant 107 is spending less through AwanTunai over time.

In both daily regressions, the wide spread of the data means that we take the regression results cautiously; further work could include identifying more advanced methodologies to more effectively handle wide-spread data.

Caveats

It should be noted that in using 'spent' to measure growth, we make the imperfect assumption that merchants are not purchasing inventory from other suppliers. Additionally, each variation of the regression is related, but deals with different transformations of the data. Therefore, decision-making based purely on direct comparison between the resulting coefficients should be done cautiously.

References and Acknowledgements

[1] Li, Y., Robani, M. D., Suciu, V., & He-Yueya, J. (2022, September 27). *Inventory Purchase Recommendation for Merchants in Traditional FMCG Retail Business*