Article

Open Data Policy, E-Commerce Connectivity and Portfolio Size Predict a Global Brand's Online Popularity

Martijn J. Hoogeveen[1]

[1] Department Technical Sciences & Environment, Open University, the Netherlands.
* Correspondence: Martijn J. Hoogeveen; martijn.hoogeveen@icecat.com

Abstract: Background: Content marketing is increasingly important for online branding. Brand popularity can be more easily determined online than sales-based measures but is not yet well-explained from a content marketing perspective. Promising predictors of online brand popularity are open data syndication policies, connectivity to e-commerce platforms, product reviews, data health, and the depth and width of a brands product portfolio. A predictive content marketing model can help brand owners to understand their e-commerce potential. **Methods:** We used brand popularity (Brand Popularity Rank) and catalog data in combination with product reviews from an independent content aggregator. For all datasets, we selected the overlapping dataset for brand popularity and brand reviews based on a period of 90 days from June 10, 2022, till September 24, 2022 ($n = 333$ manufacturing brands). Backward stepwise multiple linear regression is used to determine a predictive content marketing model of the Brand Popularity Rank. **Results:** Through stepwise backward multiple linear regression five highly significant ($p < 0.01$) predictive factors for brand rank are selected in our content marketing model: the brand's data syndication policy, the number of connected e-commerce platforms, a brand's number of products, its number of products per category, and the number of product categories in which it is active. Our model explains 78% of the variance of Brand Popularity Rank and has a good and highly significant fit: $F(5, 327) = 233.5$, $p < 0.00001$. **Conclusions:** A content marketing model can adequately predict a Brand Popularity Rank based on online popularity. In this model an open content syndication policy, more connected e-commerce platforms, and catalog size, i.e., presence in more categories and more products per category are each related to a better (lower) Brand Popularity Rank score or online success for a manufacturing brand.

Highlights

- Brand Popularity Rank is a measure based on online product data-sheet downloads per brand.
- Brand Popularity Rank is improved when manufacturing brands adopt an open content syndication policy, or have more connections to e-commerce platforms, more products per category or presence in more product categories.
- A content marketing model based on these predictors, explains 78% of the variance of the Brand Popularity Rank.
- The added value of a brand's product review score or sustainability index as reliable predictors in our model is not yet clear.

Keywords: Brand rank; content marketing; predictive model; open data policy; e-commerce

1. Introduction

Content marketing is becoming increasingly important for online branding, both business-to-business (Bamm et al., 2018) and business-to-consumer. Sales based measures for consumer choices of brands are in the online space often less accessible than brand popularity based measures (Kim et al., 2019). Brand popularity is often estimated on the basis of product usage (Qian et al., 2022), but predicting it is still a challenge.

Product reviews, product completeness and content syndication are all considered important elements in the content marketing mix contributing to brand popularity, especially in the e-commerce era. Online reviews on product usage experiences are seen as an effective way to build a strong brand image, which in turn helps to reduce consumer's buying uncertainty (Chakraborty & Bhat, 2018). Further, product completeness or data health is believed to influence consumer purchase decisions as well (Amanah et al., 2018), although adding multimedia to complete product communication is only helpful if it is used congruently (Hoogeveen, 1997). And, finally content syndication has become a prominent element of content marketing in, for example, the publishing domain (Edo et al., 2019). Content syndication policies can be "open" (Bruzzone et al., 2020) or "exclusive" (Chang & Jhang, 2020).

Finally a brand's sustainability reporting (Loh & Tan, 2020) is increasingly seen as impactful regarding corporate performance (Cowan & Guzman, 2020), and a clear relation between e-commerce and sustainability is established (Reijnders & Hoogeveen, 2001).

Apart from evaluating the use of brand popularity ranking as an indicator of brand popularity instead of crude product usage metrics, the novelty of this research is in the aim to almost fully explain brand popularity from easily obtainable open data regarding a brand's online marketing mix. We aim to do this by including key content marketing parameters in a single predictive model. In this model, we adopt online content marketing variables such as data health or data completeness, product reviews score, content syndication policy in addition to more classical factors related to product usage and the size, depth and width of a brand's product catalog. The ability to predict a brand's popularity based on easily obtainable open content marketing variables is a powerful and modern tool for brand owners that want to quickly understand their e-commerce potential. Regarding modelling, we follow a similar statistical approach as we recently applied in the development of an adequate predictive model for COVID-19 seasonality (Hoogeveen et al, 2022).

This study's hypothesis is that a content marketing model, combining multiple digital marketing factors, improves the prediction of the global popularity rank of successful online manufacturing brands (Brand Popularity Rank) compared to each factor alone. Therefore, the main objective of this study is to assess whether a comprehensive model improves the predictive modelling of a brand's global online popularity ranking. A secondary objective is to determine the usefulness of Brand Popularity Rank instead of a crude usage measure like product content downloads.

2. Methods

2.1. Data

For this analysis, we selected datasets for an initial sample of 500 manufacturing brands, which is reduced to 333 brands (n = 333) by taking into account only the overlapping brands from each dataset.

For Brand Popularity Rank, we used the respective Icecat Rank dataset (2022) from June 10, 2022, till September 24, 2022, and selected the top 500 most successful brands. Icecat uses a standard method to recalculate a brand's ranking on the basis of a brand's online product data-sheet downloads over the trailing 90 days. A Brand Popularity Rank of 1 is the best score, and a Brand Popularity Rank of 500 is the lowest score in the dataset. As the open catalog is used by 10,000s of e-commerce companies over the world on a daily basis for updating their own catalogs, it is seen as a sufficiently representative and extensive statistical sample. Selected alternative online predictors for Brand Popularity Rank are the brands review score, based on the Testseek aggregated reviews database, and further from the Icecat dataset are obtained: data health score, product downloads, the number of connected e-commerce platforms, the content syndication policy of a brand, the number of products of a brand as an indication of portfolio size, and the number of cate-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 September 2022 doi:10.20944/preprints202209.0470.v1

gories in which a brand has at least 1 product as an indication of portfolio width. Additionally, we calculate the average number of products per category per brand as an indication of portfolio depth.

The brand review score is calculated as the average of the review scores of all a brand's individual products and results in a figure between 0 and 100% (or 0 and 1). There are for the selected brands 333 brands that have product reviews in the database to calculate the average score from. Data health is an average based on the completeness of the product data-sheets which describe a brand's online products. Product downloads is the number of times that a brand's products are downloaded by its e-commerce platforms or end-users in the trailing 90 days. The number of user platforms are the number of different e-commerce users that connect to the database to download product data-sheets of a brand in the trailing 90 days. The syndication policy can be 'open', all of a brand's product data are made available as open data, or 'closed', the brand's product data is only available for selected or authorized e-commerce users.

The lack of standardized sustainability data makes it still hard to obtain a dataset with sufficient coverage of the brands in our samples, so we could not include such a dataset in our present analysis.

2.2. Sensitivity analysis

As explained above, we selected for the brand ranking and reviews datasets the brands that were present in both datasets in the overlapping period.

For sensitivity analyses, we extended the Brand Popularity Rank dataset to include all 500 brands, initially selected. And, we perform multiple linear regression both with and without product data-sheet downloads as predictor given that this variable is not truly independent as it is used to determine the Brand Popularity Rank. Therefore, it is good test to see if our predictive model is sufficiently powerful.

2.3. Statistical analysis

Variables are presented with their sample sizes (n), means (M), and standard deviations (SD). Correlation coefficients are calculated to assess the strength and direction of relations of each independent variable with Brand Popularity Rank, and with each other.

Stepwise backward multiple linear regression for all independent variables on Brand Popularity Rank was used to keep only candidate predictors that are significant ($p < 0.05$) in the model and remove insignificant ones. During each step, per independent variable the (standard) coefficient, t-stat and its 95% confidence interval (CI), probability, and the variance inflation factor (VIF) value are calculated as well. Next, we removed the predictors that were multicollinear whereby we used an acceptable VIF value of at most 5 as a threshold, whereby we see a VIF value below 2.5 as ideal.. With the remaining independent variables the F-value, standard deviations and errors, degrees of freedom (DF), and significance level, are calculated to test the goodness of fit hypothesis for our predictive model for Brand Popularity Rank. Finally, the multiple R, Multiple R squared (R^2) and adjusted R^2 correlation coefficients are calculated to estimate the predictive power of the model. On the basis of the regression outcomes, the algebraic equation to predict Brand Popularity Rank is given, and we provide a dot plot to visualize the fit of the model's predicted versus the observed Brand Popularity Rank values.

Additionally, as linear regression assumes normality of residuals, the Shapiro-Wilk test is applied, and to test the homoscedasticity requirement – homogeneity of the variance of residuals – the White test is applied. Additionally, the priori power is calculated for each predictor separately and is compared with the outcomes of the predictive model.

For selected independent variables with a $p < 0.05$ and VIF score < 5, standard $\log_{10}$, and square root data transformations are applied to reduce non-linearity in relations between variables which helps to reduce skewness, and, especially, meet the normality and homoscedasticity requirements. Such data transformations do not change the nature and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 September 2022 doi:10.20944/preprints202209.0470.v1

direction of relations between independent variables and Brand Popularity Rank. As the variables in the datasets only contained positive numbers no further data transformations were necessary.

We reported the results in APA style.

All statistical analyses were done with Stats Kingdom 2022, which we benchmarked before on R version 3.5 (Hoogeveen, 2022). Key outcomes are calculated in Excel as well as a check.

3. Results

3.1. Variables and their correlations

The sample size (N), means, and SDs of the independent and dependent variables as used in our multiple linear regression models are given in Table 1. If applicable, the values are given for the data sets after applied data transformations.

Table 1: overview means (M), standard deviations (SDs) and skewness values.

Variable	**N**	**Mean**	**SD**
Downloads	333	9,237,924	3,130,1091
Log_{10}(Platforms)	333	1.83	0.46
Open data policy	333	0.46	0.50
Log_{10}(Products/category)	333	2.10	0.61
Log_{10}(Categories)	333	1.85	0.38
$Temperature^2$	333	221	142
Dew point temperature	333	8.56	5.70
Sqrt(Products)	333	130	155
Sqrt(Brand Popularity Rank)	333	14.0	5.37

Table 1: Overview of mean (M), and standard deviation (SD) per independent variable as used in the multiple linear regression models. The function Sqrt returns the square root of the variable.

The independent variables that inversely correlate (highly) significantly with Brand Popularity Rank (see Table 2), are in order of strength: platforms, open data policy, categories, products, and reviews score. The inverse correlations are logical as an improvement on these factors coincides with an improved, i.e., lower, Brand Popularity Rank. For example, an open data syndication policy, more connections to e-commerce platforms, and a bigger catalog size (more products in more categories and more products per category) lead to a better, i.e., lower, Brand Popularity Rank. The highly significant inverse correlation with downloads speaks for itself given that Brand Popularity Rank is directly based on it. Unsurprisingly, the same factors correlate (highly) significantly with downloads, except for reviews score. Worth to highlight is that data health correlates highly significantly with review score ($r(500) = 0.21$ ($p < 0.00001$), especially given that these factors are fully independent measures.

Table 2. Overview of correlations between variables before transformations and limitations of the dataset (n = 500).

	Brand Popularity Rank	Data Health	Downloads	Platforms	Review score	Open data policy	Products	Categories	PPC
Brand Popularity Rank	1.000	0.054	**-0.351**	**-0.538**	**-0.102**	**-0.321**	**-0.261**	**-0.320**	-0.047

						-			
Data Health	0.054	1.000	-0.029	-0.035	**0.207**	0.060	-0.002	-0.064	**-0.112**
Downloads	**-0.351**	-0.029	1.000	**0.669**	0.018	**0.176**	**0.648**	**0.432**	0.038
Platforms	**-0.538**	-0.035	**0.669**	1.000	0.075	**0.522**	**0.313**	**0.438**	-0.067
Reviews score	**-0.102**	**0.207**	0.018	0.075	1.000	0.045	0.033	**-0.090**	0.054
Open data policy	**-0.321**	-0.060	**0.176**	**0.522**	**0.045**	1.000	-0.012	0.070	-0.084
						-			
Products	**-0.261**	-0.002	**0.648**	**0.313**	0.033	0.012	1.000	**0.225**	**0.297**
Categories	**-0.320**	-0.064	**0.432**	**0.438**	**-0.090**	0.070	**0.225**	1.000	**-0.104**
						-			
Products/Category	-0.047	**-0.112**	0.038	-0.067	0.054	0.084	**0.297**	**-0.104**	1.000

Table 2: Overview of correlations between variables before transformations and limitations of the dataset (n = 500). Bold: significant at p < 0.05. Bold+Underlined: highly significant at p < 0.0.1

After transformations and dataset restriction to brands with a review score, the correlations with Brand Popularity Rank are more or less the same. Only the correlation with products per category has become highly significant as well ($r(333) = -0.46$, $p < 0.00001$).

In both situations there are no collinears where independent variables do not have a strong correlation ($r \geq 0.8$) with each other.

3.2. Outcomes predictive model

After multiple iterations during stepwise backward multiple linear regression, five independent variables were selected from the set of online marketing variables that are both significant ($p < 0.05$) and have an acceptable VIF value below 5. These selected predictors are: open syndication policy, the number of connected e-commerce platforms, the number of products, the products/category and the number of categories in which a brand has products (see Table 3). In the total datasets, Data Health did not have a significant correlation with Brand Popularity Rank ($r(500) = 0.06$, $p = 0.23$) and was deselected for our model. Also, Product Reviews was deselected, despite a significant correlation with Brand Popularity Rank ($r(500) = -0.10$, $p = 0.022$) and given that it did not add explanatory power to our model.

Table 3: multiple linear regression for predictors excluding downloads.

	Coeff.	***SE***	***t-stat***	***lower*** $t_{0.025(327)}$	***upper*** $t_{0.975(327)}$	***Stand. Coeff.***	***P***	***VIF***
b	*51.7*	*1.63*	*31.6*	*48.5*	*54.9*	*0*	*<0.00001*	
Log_{10}(Categories)	*-7.26*	*0.55*	*-13.2*	*-8.34*	*-6.18*	*-0.51*	*<0.00001*	*2.20*
Log_{10}(Platforms)	*-4.64*	*0.42*	*-11.2*	*-5.45*	*-3.82*	*-0.40*	*<0.00001*	*1.91*
Open data policy	*-2.47*	*0.35*	*-7.14*	*-3.16*	*-1.79*	*-0.23*	*<0.00001*	*1.55*
Sqrt(Products)	*0.00727*	*0.00166*	*4.37*	*0.00399*	*0.0105*	*0.209*	*0.000017*	*3.44*
Log_{10}(Products/Category)	*-7.40*	*0.42*	*-17.5*	*-8.23*	*-6.57*	*-0.84*	*<0.00001*	*3.43*

Table 3: Overview of outcomes per predictor after multiple linear regression without downloads resulting in an adjusted R-squared = 0.778. Selection of predictors is based on being (highly) significant and having multicollinearity (VIF) score below 5. The function Sqrt returns the square root of the variable.

 doi:10.20944/preprints202209.0470.v1

On the basis of this test, we can reject the null-hypothesis (H_0) that our predictive content marketing model with the five selected factors does not provide a good fit: $F_{(5, 327)} = 233.5$, $p < 0.00001$. R^2 equals 0.781, which means that our predictors explain 78.1% of the variance of Brand Popularity Rank (Adjusted R square = 0.778 and R = 0.884) (see Fig. 1).

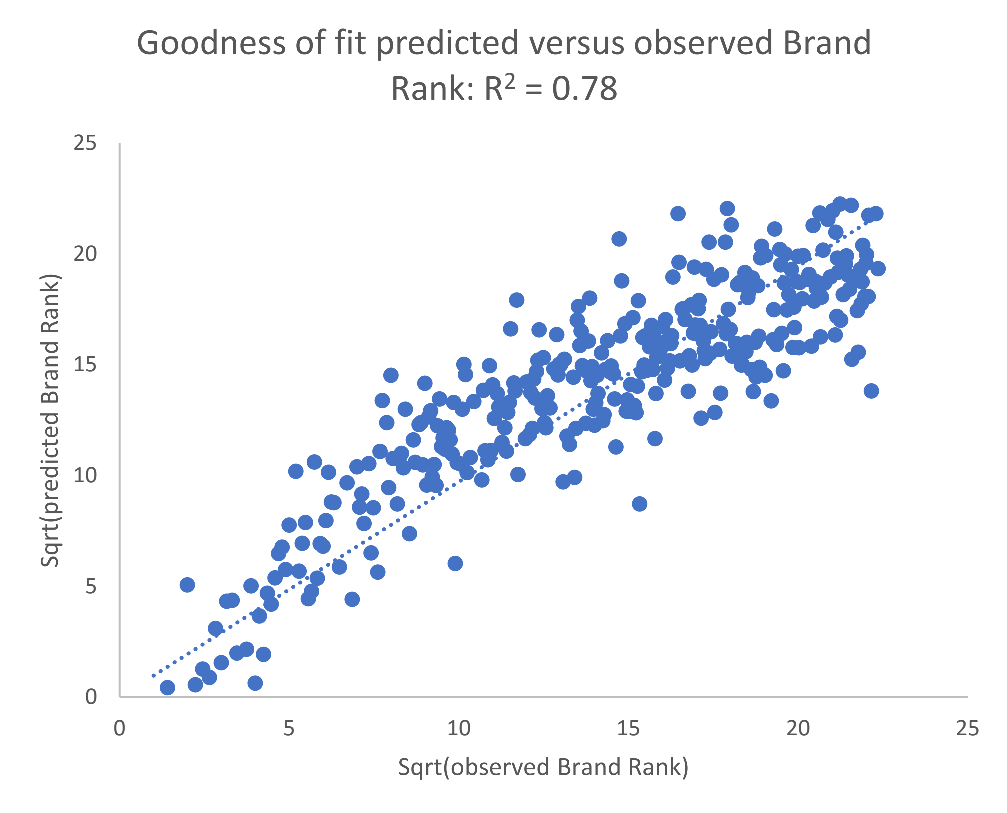

Fig. 1: The predictive content marketing model is explaining 78.1% of the variance of the observed Brand Popularity Rank.

Below, the predictive content marketing model's regression formula for $\widehat{GBR}$:

$$\left(51.7 - 7.25\log_{10} C - 4.64\log_{10} Pl - 2.47Po + 0.00727\sqrt{Pr} - 7.40\log_{10} PPC\right)^2$$

Where $\widehat{GBR}$ is the predicted Brand Popularity Rank, C is the number of categories in which a brand has products, Pl is the number of e-commerce platforms that are downloading a brand's product data, Po indicates whether a brand has an open data ("1") syndication policy or an exclusive one ("0"), Pr is the number of products that a brand has in total in its catalog, and PPC is the number of products that a brand has per category in its catalog. *C, Pr, and PPC* are all variables describing the size (width and depth) of a brand's product portfolio.

Given the somewhat raised multicollinearity scores, *Products* has to be interpreted in our content marketing model as a correction on *PPC.*

3.3. Statistical outcomes sensitivity analyses

As a sensitivity analysis we added downloads of a brand's product data-sheets to the multiple linear regression test. In effect, downloads replaces the products variable (see Table 4), the other predictors and the outcomes are more or less the same: also on the basis of this sensitivy test, we reject the H_0 that our predictive content marketing model does not provide a good fit: $F_{(5, 327)} = 225.2$, $p < 0.00001$, and R^2 equals 0.775, which means that our content marketing predictors still explain 77.5% of the variance of Brand Popularity Rank (Adjusted $R^2 = 0.771$ and $R = 0.880$). The outcomes are only slightly lower. The only difference is that the VIF scores are slightly better (< 2.5).

Table 4: multiple linear regression for all predictors including downloads

	Coeff.	SE	t-stat	lower $t_{0.025(327)}$	upper $t_{0.975(327)}$	Stand. Coeff.	P	VIF
b	48.6	1.33	36.6	46.0	51.3	0	<0.00001	
Downloads	1.79e-8	5.87e-9	3.06	6.40e-9	2.95e-8	0.105	0.00241	1.70
Log_{10}(Platforms)	-4.65	0.428	-10.9	-5.49	-3.81	-0.400	<0.00001	1.96
Open data policy	-2.55	0.352	-7.24	-3.25	-1.86	-0.237	<0.00001	1.56
Log_{10}(Products/category)	-6.35	0.288	-22.0	-6.92	-5.78	-0.720	<0.00001	1.55
Log_{10}(Categories)	-6.35	-0.478	-13.3	-7.29	-5.41	-0.444	<0.00001	1.62

Table 4: Overview of outcomes for selected predictors including downloads after multiple linear regression resulting in an adjusted R-squared = 0.771. Selection of predictors is based on being (highly) significant and having multicollinearity (VIF) score below 5.

In another sensitivy test, using downloads as a dependent variable instead of Brand Popularity Rank does not lead to satisfactory outcomes because of concerns related to the homoscedasticity of residuals and normality requirements. When including the whole dataset (n = 500 brands) despite having quite some missings does lead to the inclusion of Review Score in the predictive model but again homoscedasticity or normality requirements cannot be met.

4. Discussion

The predictive power of the content marketing model including a brands connected platforms, its content syndication policy, the number of products, products per category and the number of categories in which it is active is good (77.4%). For example, an open syndication policy, more connected e-commerce platforms, and more products per category and more categories in which a brand is active, are related to a better (lower) Brand Popularity Rank score. The number of products can be seen as a correction on the product per category variable in our model.

When we test our model by including the number of downloads these products generate as *independent* variable, the predictive power of our model remains more or less the same. Given that the downloads factor was used to determine the Brand Popularity Rank, it implies that our content marketing model contains a robust – equally powerful - set of independent predictors. The transformation of downloads into Brand Productivity Rank is superior to other transformations in the sense that homoscedasticity and normality requirements are fully met, which is not the case when downloads is used as *dependent* variable instead of Brand Productivity Rank.

Assuming that the same online marketing factors are relevant as predictors for brands outside our sample, we could hypothesize that our model has a similar predictive power for brands in other sectors or categories outside our sample, or, to a certain degree, for non-manufacturing brands.

Methodological concerns

The Brand Popularity Rank panel is relatively strong in Western countries, which might lead to a sampling bias. It would be interesting to see if the derived predictive content marketing model would fundamentally change if the panel is strengthened in East-Asian and African countries.

Further, it would of interest to see the completeness of the product reviews dataset in terms of brand coverage being improved. It might lead to the inclusion of product reviews score as a significant predictor, as shown in a sensitivity analysis, leading to a slightly improved predictive power of our model.

Finally, it would still be of interest to see if a brand's sustainability score can be obtained with sufficient coverage of our brands sample, and whether its inclusion has added value for the predictive model.

Conclusion

A predictive content marketing model explains 78.1% of the variance of the Brand Popularity Rank of a sample of 333 successful manufacturing brands, and has a good fit ($F_{(5, 327)} = 233.5$, $p < 0.00001$), as the predicted and observed Brand Popularity Rank correlate strongly and highly significantly. The significant predictors in the content marketing model are a brand's content syndication policy, the number of connect e-commerce platforms, a brand's number of products, products per category and the number of categories in which it is active. The mentioned factors are inversely correlated with Brand Popularity Rank, i.e., an open syndication policy, more connected e-commerce platforms, more products per category and being active in more categories all lead to a better (lower) Brand Popularity Rank. In our model, the number of products, another attribute of a brand's catalog size, should be seen as a correction on the product per category variable.

Using Brand Popularity Rank as a dependent variable leads to more statistically robust models than when product data-sheet downloads is used a brand popularity indicator.

It is yet to be determined whether a brand's product review or sustainability score have added value for a predictive content marketing model, due to the lack of standardized values per matching brand.

Martijn Hoogeveen:

Conceptualization, Methodology, Data Curation, Formal Analysis, Writing, Investigation, Visualization, Resources.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest

Corresponding other is founder of data science bureau Icecat, the provider of the used open data set regarding Brand Popularity Rank. But, there is no conflict of interest regarding any potential outcome of the data science project reported in this manuscript.

Data statement

The links to downloadable datasets are provided in the reference list or references to the sources are included. Upon request, the data used for this manuscript is available for inspection, but for other purposes we kindly refer to the respective copyright-holder(s).

Corresponding author

Martijn Hoogeveen, PhD
Attn. Marga Winnubst
Department of Technology & Environment
Open Universiteit Nederland
Valkenburgerweg 177
6419 AT Heerlen

 doi:10.20944/preprints202209.0470.v1

the Netherlands
Tel: +31 (6) 54256476
Email: martijn.hoogeveen@icecat.com

References

Amanah, D., Hurriyati, R., Gaffar, V., Layla, A. A., & Harahap, D. A. (2018, December). Effect of price and product completeness to consumer purchase decision at Tokopedia. com. In Proceedings of the 2nd Global Conference on Business, Management and Entrepreneurship 2017 (Vol. 1, pp. 34-37).

Bamm, R., Helbling, M. and Joukanen, K. (2018), "Online Branding and the B2B Context", Koporcic, N., Ivanova-Gongne, M., Nyström, A.-G. and Törnroos, J.-Å. (Ed.) Developing Insights on Branding in the B2B Context, Emerald Publishing Limited, Bingley, pp. 163-176. https://doi.org/10.1108/978-1-78756-275-220181009.

Bruzzone, A. G., Agresta, M., & Hsu, J. H. (2020). Word of Mouth, Viral Marketing and Open Data: A Large-Scale Simulation for Predicting Opinion Diffusion on Ethical Food Consumption. International Journal of Food Engineering, 16(5-6).

Chakraborty, U. and Bhat, S. (2018), "Credibility of online reviews and its impact on brand image", Management Research Review, Vol. 41 No. 1, pp. 148-164. https://doi.org/10.1108/MRR-06-2017-0173.

Chiang, I. R., & Jhang-Li, J. H. (2020). Competition through exclusivity in digital content distribution. Production and Operations Management, 29(5), 1270-1286.

Cowan, K., & Guzman, F. (2020). How CSR reputation, sustainability signals, and country-of-origin sustainability reputation contribute to corporate brand performance: An exploratory study. Journal of business research, 117, 683-693.

Edo, C., Yunquera, J., & Bastos, H. (2019). Content syndication in news aggregators: towards devaluation of professional journalistic criteria= La sindicación de contenidos en los agregadores de noticias: Hacia la devaluación de los criterios profesionales periodísticos.

Hoogeveen, M. (1997). Toward a theory of the effectiveness of multimedia systems. International journal of human-computer interaction, 9(2), 151-168.

Hoogeveen, M. J., Kroes, A. C., & Hoogeveen, E. K. (2022). Environmental factors and mobility predict COVID-19 seasonality in the Netherlands. Environmental Research, 211, 113030.

[dataset] Icecat. Brand rank search and overview page. . Accessed September 25, **2022**.

Loh, L., & Tan, S. (2020). Impact of sustainability reporting on brand value: an examination of 100 leading brands in Singapore. Sustainability, 12(18), 7392.

Kim, M.-Y., Moon, S., & Iacobucci, D. (2019). "The Influence of Global Brand Distribution on Brand Popularity on Social Media". Journal of International Marketing, 27(4), 22–38. https://doi.org/10.1177/1069031X19863307.

Qian, L., Dou, Y., Xu, X., Ma, Y., Wang, S., & Yang, Y. (2022, August). Product Success Evaluation Model Based on Star Ratings, Reviews and Product Popularity. In 2022 8th International Conference on Big Data and Information Analytics (BigDIA) (pp. 295-302). IEEE.

Reijnders, L., & Hoogeveen, M. J. (2001). Energy effects associated with e-commerce: A case-study concerning online sales of personal computers in The Netherlands. Journal of Environmental Management, 62(3), 317-321.