Journal of Digital Educational Technology
2025, 5(1), ep2501
e-ISSN: 2752-5503
https://www.jdet.net/
Research Article
OPEN ACCESS

Autonomous learning behaviors in an online coding community: A comparison between project viewing/playing and code remixing in Scratch using Benford's law

Ray Y. Shen [1*]

[1] Independent Researcher, Potomac, MD, USA
***Corresponding Author:** rayshen2718@gmail.com

Citation: Shen, R. Y. (2025). Autonomous learning behaviors in an online coding community: A comparison between project viewing/playing and code remixing in Scratch using Benford's law. *Journal of Digital Educational Technology, 5*(1), ep2501. https://doi.org/10.30935/jdet/15808

ARTICLE INFO

Received: 23 Jun. 2024

Accepted: 21 Nov. 2024

ABSTRACT

Previous studies of code-learning behaviors have been conducted in structured educational settings, utilizing student engagement metrics such as homework submission, task completion, and interactions with instructors. These types of metrics, however, are absent in open online coding platforms. To characterize autonomous code-learning behaviors in an online community, this work applied Benford's law to analyze user engagement metrics of trending projects on Scratch, the world's largest online coding platform for young learners. Statistical analysis revealed that the extent of conformity to Benford's law is independent of the project categories. Of all four user engagement metrics, the views metric exhibited the strongest conformity to Benford's law, while the remixes metric–the metric most closely associated with code-learning behaviors–showed the greatest deviation from Benford's law. This was confirmed by Pearson's χ^2 test, Nigrini's (2012) mean absolute deviation test, and an evaluation of the mantissas of the user engagement metrics. This study demonstrates that the extent of conformity to Benford's law can be used as novel features for characterizing autonomous code-learning behaviors in unsupervised online settings. The results of this work pave the way for future studies to correlate the extent of conformity to Benford's law with specific elements of code that attract autonomous learning, providing opportunities to optimize the content and design of online coding platforms.

Keywords: computer science education, computational thinking skills, code-learning behaviors, data mining, Benford's law

INTRODUCTION

Previously, studies on code-learning behaviors have been conducted in classroom settings or instructor-led online environments, where conventional student engagement metrics are assessed. For example, Chen et al. (2020) used machine learning to study the code-learning behaviors of 69 students enrolled in a Java programming course, analyzing engagement metrics like homework effort (e.g., number of specific failures, number of on-time submissions, and total submissions) and homework time (e.g., time to first submission, time to last submission, and duration in a specific state of failures). Stewart et al. (2021) examined the code-learning behaviors of ten middle school female students enrolled in a virtual coding camp using engagement metrics such as responses to prompts via chat, student presentations, and code-sharing. Sun et al. (2021) studied two instructional modes for teaching coding to 31 students–instructor-directed lecturing and learner-centered unplugged programming–using classroom audio and video recordings, student surveys, and interviews to assess code-learning behavioral patterns. Alebaikan et al. (2022) studied the online code-learning experience of 16 female middle schoolers, generating data from interviews, the teacher's diary, and content analysis of activities recorded in student progress reports. While many studies have utilized conventional student engagement metrics–such as attendance, help-seeking, homework submission, task completion, interactions with instructors, and logs of various study activities–to study code-learning behaviors in well-structured educational settings, these types of metrics are absent in unsupervised online coding platforms, where millions of users engage in autonomous learning. In contrast to the numerous studies of student code-learning behaviors in formal educational settings, to the best of our knowledge, there have been no parallel investigations of autonomous learning behaviors in online coding communities.

Benford's law, also known as the first-digit law or the law of anomalous numbers, describes a common statistical phenomenon in the distribution of leading digits of many naturally occurring datasets. First discovered by Newcomb

(1881) and later popularized by Benford (1938), the first-digit law can be used to analyze population sizes, stock values, scientific measurements, and tax records (Nigrini, 2012, p. 25). The law's fundamental principle dictates that in many non-synthetic datasets, a leading digit of "1" appears significantly more frequently than any other digit, at an expected 30.1% frequency. The expected frequency of occurrence decreases as the leading digit increases, to 17.6% for "2" as a leading digit and down to an expected mere 4.6% appearance rate for "9" as a leading digit (Benford, 1938). Benford's law can provide insights into unexpected patterns, anomalies, or even fraud that a dataset might contain based on conformity of the data. The first-digit law's numerous applications include assessing the quality of synthetic images without the need for any reference images (Varga, 2021), identifying dynamic transitions in cardiac models (Seenivasan et al., 2016), and evaluating the conformity of cancer registry data to expected patterns (Crocetti & Randi, 2016). Of particular interest, the extent of conformity to Benford's law has recently been used as novel features to train artificial intelligence models (for example, Caffarini et al., 2022). Considering the large number of datasets on learning behaviors already accumulated by educators and researchers, applying Benford's law to study learning behaviors could uncover new insights and potentially enhance educational practices.

Due to the vast amounts of data voluntarily provided by its users, social media contains many opportunities for studying trends in data, user tendencies, and human behaviors. Datasets created from user interactions on popular social media platforms such as Facebook, Instagram, Twitter, or TikTok can be used by researchers to track shifts in internet trends, monitor changes in public opinion over various topics, or analyze internet culture (Felt, 2016). Social media data is also useful in studying the impact of the spreading of information, political discourse, and the dynamics of various online communities (Abkenar et al., 2021). Many datasets gathered from mainstream social media conform to Benford's law, and those that don't could suggest unknown or unexpected user behaviors, data irregularities, or manipulation. The distributions of different user engagement metrics, such as views or reposts, have been found to closely follow Benford's law, and can therefore be used to discern patterns. Datasets mined from mainstream social media can both be validated or discredited by making statistical comparisons to the expected distribution derived from Benford's law (e.g., Bhosale & Di Troia, 2022).

Scratch (https://scratch.mit.edu) is a visual programming language and online coding platform designed to introduce programming concepts to young learners, with over 120 million registered users and more than 150 million published projects. Developed by the Lifelong Kindergarten Group at the MIT Media Lab, Scratch provides a user-friendly environment for creating various projects such as games, animations, and stories. Scratch's intuitive visual interface, drag-and-drop code blocks, and simple tutorials appeal to young learners, as the need for traditional text-based programming is eliminated. Acting as an integrated coding and social media platform, Scratch encourages its users to publish their projects, remix the creations of others, leave comments, and interact with the online community. While social media data has proven valuable for understanding student learning (Su & Lai, 2021) and studies have used Scratch to teach computational thinking to small groups in classroom settings (e.g., Erümit & Şahin, 2020; Kwon & Cheon, 2019; Kwon et al., 2018), online coding communities, including Scratch, have not yet been studied and analyzed utilizing data mining and data analytic techniques to extract valuable statistics related to the social interactions and autonomous learning behaviors of their users.

To address these gaps, this work takes the first steps in investigating the vast amount of social interaction and engagement data on Scratch by collecting and examining the projects featured on its trending page to extract statistical information on user engagement metrics in terms of their conformity to Benford's law. To the best of our knowledge, this study is the first to investigate learning behaviors using Benford's law. Our study unequivocally demonstrates that while the views metric of Scratch, like similar metrics in mainstream social media, obeys Benford's law, the remixes metric, which involves active coding, deviates dramatically from this law. These findings indicate that the extent of conformity to Benford's law can serve as a highly sensitive measure for characterizing code-learning in a social network dominated by autonomous learning activities, where conventional user engagement metrics are non-existent. Further studies can, for example, correlate the extent of conformity to Benford's law with the level of computational thinking skills exhibited in Scratch projects. These lines of research are expected to offer additional insights into specific elements in Scratch code that attract autonomous learning. Such insights could potentially lead to the optimization of online coding platforms, benefiting both autonomous learning within online communities and instructor-led learning in educational settings that utilizes popular programming interfaces such as Scratch.

METHODS

Benford's Law

Let D_i be the ith digit of an arbitrary number n. Benford's law denotes that for every positive integer $k \in Z^+ \sim \{1\}$, all $d_1 \in \{1, 2, 3 \dots, 9\}$ and all $d_j \in \{0, 1, 2, 3, \dots, 9\}$ with $j \in \{2, 3, 4, \dots, k\}$, the probability of n naturally occurring with its first k significant digits given by $d_1, \dots, d_k$ is (Nigrini, 2012, p. 13):

$$P(D_1 = d_1, \cdots, D_k = d_k) = log_{10}\left(1 + \left(\sum_{i=1}^{k} d_i * 10^{k-i}\right)^{-1}\right). \quad (1)$$

Note that since $k \in Z^+ \sim \{1\}$, we can calculate the probability of a specific leading digit by summing the probabilities of all numbers starting with that digit in Eq. (1) (Benford, 1938):

$$P(d) = log_{10}\left(1 + \frac{1}{d}\right), \quad (2)$$

where d is the leading digit.

Data Collection and Analysis

Scratch's trending page consists of six project categories: animations, art, games, music, tutorials, and stories. Each Scratch project has four user engagement metrics: loves, favorites, remixes and views. Users can browse through projects in each category and interact with them by viewing/playing the projects and leaving loves and favorites. Users are also able to remix a project, where they can customize, add to, or fully revamp another user's project and republish it as their own while automatically crediting the original author. In contrast to views, loves, and favorites, the remixes metric is associated with active coding.

This work utilizes Google's free web scraping tool, Web Scraper (https://webscraper.io), to automatically extract the counts of all user engagement metrics of each trending project from all six project categories and exports them to Microsoft Excel files. The data collection does not require ethical approval since the information is freely available in the public domain. The Web Scraper, which was installed as a Chrome Extension, had over 600,000 users as of 2023. We utilized a recursive pagination handler that allows verification of successful data collection by comparing the last scraped records with those on the initial pages. The selector graph used for data collection in this study is provided in **Appendix A** (**Figure A1**).

While Pearson's Chi-square (χ^2) test (Bock et al., 2019, p. 621) is the go-to test when comparing an experimental distribution with a theoretical one and some authors are still using it to test conformity to Benford's law, its limitations are well-documented due to its over-sensitivity to minor fluctuation in digits when the size of the dataset is large (Nigrini, 2012, p. 158), such as in the current study. To overcome this limitation, the mean absolute deviation (MAD) test has been developed and widely accepted for testing conformity to Benford's law (e.g., da Silva et al., 2020; da Silva Azevedo et al., 2021; Nigrini, 2012, p. 158). Additionally, we used the strong Benford's law (Berger & Twelves, 2018), which applies to the entire number rather than only its leading digit, to address the limitation of focusing solely on leading digits.

Let O_i represent the observed occurrence of leading digit i, let N represent the total number of records, and let P_i represent the frequency of the occurrence of leading digit i expected from Benford's law. The standardized residual of a leading digit can be found by subtracting the expected occurrence from the observed occurrence and dividing the difference by the square root of the expected occurrence (Bock et al., 2019, p. 624):

$$Standardized\ residual = \frac{O_i - NP_i}{\sqrt{NP_i}}. \tag{3}$$

Pearson's Chi-square statistic, denoted by χ^2, is found by summing the squared standardized residuals between observed and expected occurrences (Eq. [3]) for all leading digits 1 through 9 (Bock et al., 2019, p. 621):

$$\chi^2 = \sum_{i=1}^{9} \frac{(O_i - NP_i)^2}{NP_i}. \tag{4}$$

While null hypothesis significance tests, such as Pearson's Chi-square test, are prone to oversensitivity to insignificant spikes with large sample sizes, the MAD test is independent of the sample size and is known for its well-tested ability to assess conformity to Benford's law (Nigrini, 2012, p. 158). The MAD score can be calculated by taking the average of the absolute values of the differences between the observed frequency and expected frequency from Benford's law for each leading digit 1 through 9 (Nigrini, 2012, p. 158):

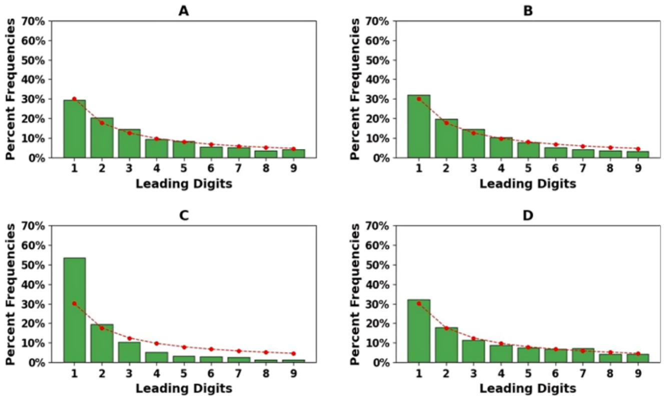

Figure 1. First digit distributions of the user engagement metrics in the animations category: (A) loves, (B) favorites, (C) remixes, & (D) views (green bars represent the observed percentage frequencies of leading digits, while the red circles depict the percentage frequencies of leading digits predicted by Benford's law) (Source: Author)

$$MAD = \frac{1}{9} \sum_{i=1}^{9} \frac{|O_i - NP_i|}{N}. \tag{5}$$

For the theoretical Benford distribution, the MAD score is zero. A higher MAD score for a distribution of leading digits indicates greater deviation from Benford's law, and vice versa. The ranges of critical MAD scores for conformity to Benford's law, as described by Nigrini (2012, p. 160), are, as follows: 0–0.006 for close conformity, 0.006–0.012 for acceptable conformity, 0.012–0.015 for marginally acceptable conformity, and > 0.015 for nonconformity.

All data analysis was performed using computer programs written in Python 3.12 with the following imported libraries: os, pandas, matplotlib, math, numpy, scipy.stats, seaborn, and scikit_posthocs. The nonparametric Kruskal-Wallis test, followed by Dunn's post-hoc test, was employed to compare MAD scores across both different project categories and different user engagement metrics.

RESULTS

Data mining located 1,285 projects from the animations category, 6,431 from the art category, 2,536 from the games category, 2,315 from the music category, 2,138 from the tutorials category, and 6,768 from the stories category. The successful project data retrieval rates ranged from 99.5% to 100% for all six categories and all sample sizes in the study are well above the sample sizes used in many studies regarding Benford behavior (e.g., Druică et al., 2018).

We first assess the extent of conformity to Benford's law by the four user engagement metrics in all project categories. The distribution of the percentage frequencies of leading digits for the animations project category as well as the theoretical Benford distribution (Eq. [2]) are shown in **Figure 1**. The percentage frequencies of leading digits for all six project categories are given in **Table 1**. Rather than a uniform distribution of leading digits, the datasets exhibit Benford-like behavior via a trend of decreasing percentage frequencies from a leading digit of 1 to 9 as displayed in **Figure 1** and **Table 1**.

Table 1. Percentage frequencies for each leading digit

Category	User engagement	1	2	3	4	5	6	7	8	9
Animations	Loves	29.56	20.20	14.59	9.20	8.35	5.54	5.07	3.51	3.98
	Favorites	32.06	19.65	14.32	10.34	7.64	5.09	4.14	3.50	3.26
	Remixes	53.72	19.42	10.55	5.04	3.12	2.88	2.64	1.44	1.20
	Views	32.06	17.78	11.47	8.66	7.49	6.79	7.10	4.29	4.37
Art	Loves	26.82	19.17	15.61	10.73	8.66	6.76	4.98	4.31	2.95
	Favorites	31.55	22.62	14.85	9.44	6.51	5.06	4.45	3.15	2.36
	Remixes	56.56	15.95	8.71	5.46	3.46	2.94	3.67	1.89	1.36
	Views	32.90	16.67	10.96	8.55	7.40	6.59	5.91	5.77	5.24
Games	Loves	32.78	23.37	13.52	9.37	6.56	4.98	3.99	3.08	2.33
	Favorites	34.37	23.96	13.18	8.90	6.78	4.33	3.31	2.86	2.33
	Remixes	51.77	19.11	8.73	6.58	3.80	3.54	2.28	2.28	1.90
	Views	29.27	18.60	12.12	9.12	7.50	6.67	6.24	5.77	4.70
Music	Loves	26.65	18.33	15.51	11.09	8.71	6.28	5.55	4.12	3.77
	Favorites	27.10	20.50	15.61	10.90	8.70	6.55	4.80	3.14	2.69
	Remixes	50.14	17.99	11.05	6.09	4.25	2.83	2.97	2.69	1.98
	Views	30.36	17.37	12.21	9.74	6.89	6.19	6.97	6.15	4.11
Tutorials	Loves	48.67	20.61	10.63	6.51	4.45	3.00	3.00	1.83	1.31
	Favorites	49.79	20.97	10.43	6.68	3.60	3.03	2.47	1.75	1.28
	Remixes	64.67	15.27	9.28	3.29	2.99	1.80	2.10	0.30	0.30
	Views	34.05	18.48	12.25	8.89	7.20	5.75	4.72	4.68	3.98
Stories	Loves	45.76	18.29	11.06	7.70	5.12	3.56	3.64	2.71	2.17
	Favorites	45.26	18.97	11.53	6.97	5.29	4.15	2.84	2.51	2.48
	Remixes	51.10	16.75	10.31	5.86	5.03	3.22	3.87	2.13	1.74
	Views	33.16	18.37	12.93	9.20	7.02	6.22	5.40	4.21	3.49

Table 2. Goodness-of-fit tests

Category	User engagement	Pearson's χ^2 test		MAD test	
		χ^2	p	MAD score	Conformity
Animations	Loves	21.3	6.3×10^{-3}	1.1×10^{-2}	Acceptable conformity
	Favorites	30.6	1.7×10^{-4}	1.4×10^{-2}	Marginally acceptable conformity
	Remixes	138.4	$< 1.0 \times 10^{-4}$	5.6×10^{-2}	Nonconformity
	Views	10.0	2.6×10^{-1}	7.7×10^{-3}	Acceptable conformity
Art	Loves	146.3	$< 1.0 \times 10^{-4}$	1.5×10^{-2}	Marginally acceptable conformity
	Favorites	285.7	$< 1.0 \times 10^{-4}$	2.0×10^{-2}	Nonconformity
	Remixes	343.9	$< 1.0 \times 10^{-4}$	5.9×10^{-2}	Nonconformity
	Views	54.6	$< 1.0 \times 10^{-4}$	9.3×10^{-3}	Acceptable conformity
Games	Loves	135.4	$< 1.0 \times 10^{-4}$	2.1×10^{-2}	Nonconformity
	Favorites	175.8	$< 1.0 \times 10^{-4}$	2.5×10^{-2}	Nonconformity
	Remixes	211.4	$< 1.0 \times 10^{-4}$	5.2×10^{-2}	Nonconformity
	Views	6.7	5.7×10^{-1}	4.9×10^{-3}	Close conformity
Music	Loves	41.8	$< 1.0 \times 10^{-4}$	1.3×10^{-2}	Marginally acceptable conformity
	Favorites	77.9	$< 1.0 \times 10^{-4}$	1.8×10^{-2}	Nonconformity
	Remixes	160.8	$< 1.0 \times 10^{-4}$	4.5×10^{-2}	Nonconformity
	Views	15.6	4.8×10^{-2}	5.6×10^{-3}	Close conformity
Tutorials	Loves	483.3	$< 1.0 \times 10^{-4}$	4.9×10^{-2}	Nonconformity
	Favorites	499.3	$< 1.0 \times 10^{-4}$	5.2×10^{-2}	Nonconformity
	Remixes	209.1	$< 1.0 \times 10^{-4}$	7.6×10^{-2}	Nonconformity
	Views	24.5	1.9×10^{-3}	1.1×10^{-2}	Acceptable conformity
Stories	Loves	931.5	$< 1.0 \times 10^{-4}$	3.6×10^{-2}	Nonconformity
	Favorites	870.0	$< 1.0 \times 10^{-4}$	3.7×10^{-2}	Nonconformity
	Remixes	366.0	$< 1.0 \times 10^{-4}$	4.7×10^{-2}	Nonconformity
	Views	64.9	$< 1.0 \times 10^{-4}$	9.3×10^{-3}	Acceptable conformity

Shown in **Table 2** are the results of Pearson's Chi-square test and the MAD test, performed to categorize levels of conformity to Benford's law for each user engagement metric under all six project categories. Standardized residuals for Pearson's Chi-square test (Eq. [3]) are provided in **Appendix A** (**Table A1**).

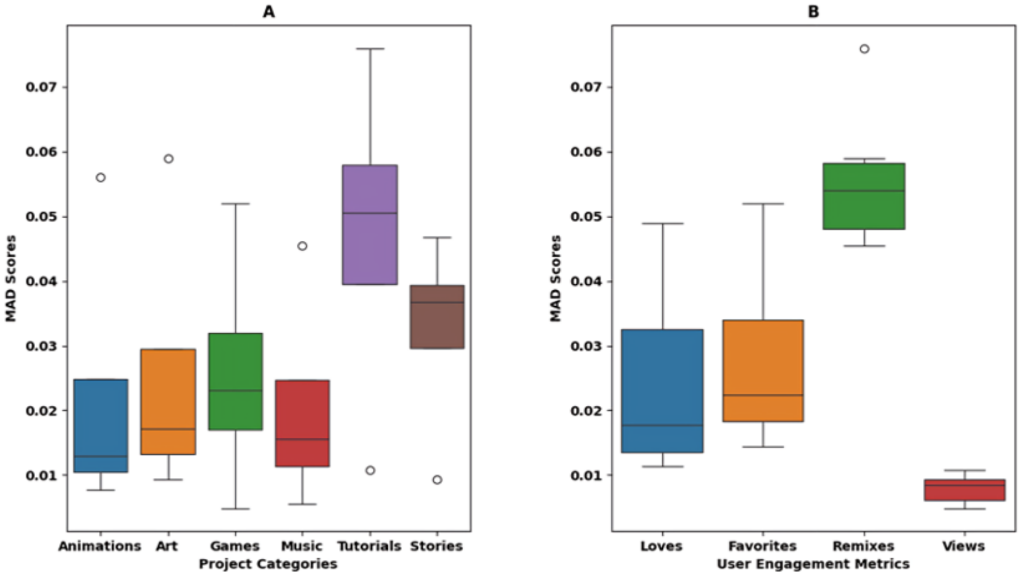

Figure 2. Box plots depicting the MAD scores across different project categories (A) and different user engagement metrics (B) (Source: Author)

The results listed in **Table 2** indicate a strong negative correlation between the p-values from Pearson's Chi-square test (Eq. [4]) and the MAD scores (Eq. [5]). The results of **Table 2** also confirm the observation made by Nigrini (2012, p. 153) that Pearson's Chi-square test can be overly sensitive for large sample sizes, such as the ones in this study. Consequently, for assessing differences in conformity to Benford's law across the six project categories and the four user engagement metrics, only the MAD test results are utilized. To compare Benfordness across different project categories, the MAD scores of the four user engagement metrics of each project category were averaged. Part A in **Figure 2** compares the mean MAD scores of the animations, art, games, music, tutorials, and stories project categories using a box plot, with each box in the plot representing a distinct project category. The Kruskal-Wallis test on the six groups of MAD scores outputs a Kruskal-Wallis statistic of 3.1 and a p-value of 0.68, demonstrating that there is no statistically significant difference in Benfordness across the different project categories. The test concludes that rather than being category-specific, the extent of conformity to Benford's law follows a similar pattern for all six project categories.

Next, we test whether different user engagement metrics averaged across different project categories demonstrate statistically different levels of conformity to Benford's law since loves, favorites, remixes, and views exhibit varying degrees of association with learning behaviors. As shown in **Table 2**, the views metric consistently has the largest p-values across all project categories. The p-values for the views of the animations and games project categories are well above the commonly used significance level of 0.05, despite Pearson's Chi-square test being overly stringent for our purpose. **Table 2** indicates that for all project categories, the views metric falls into the close conformity or acceptable conformity range by the MAD test. However, for the remaining loves, favorites, and remixes metrics, adherence to Benford's law mostly falls into the nonconformity range. In particular, all remixes metrics lie in the nonconformity range. Only the favorites metric for the animations category and the loves metric for the art and music categories are placed in the marginally acceptable conformity range, while the loves metric for the animations category was the only non-views metric categorized into the acceptable conformity range, as shown in **Table 2**. For comparing conformity to Benford's law across different user engagement metrics, the MAD score of each user engagement metric averaged over the six project categories was computed. The comparison is shown in part B in **Figure 2** via another box plot. The Kruskal-Wallis test on the four groups of user engagement metrics generated a Kruskal-Wallis statistic of 18.1 and a p-value of 0.00042, indicating statistically significant differences among the mean MAD scores of the user engagement metrics. When comparing loves with favorites, loves with remixes, loves with views, favorites with remixes, favorites with views, and remixes with views, Dunn's post-hoc test yielded the following p-values: 0.668, 0.032, 0.037, 0.086, 0.012, and 0.000024, respectively. The loves-favorites comparison indicated the least difference in conformity to Benford's law while the remixes-views comparison indicated the greatest difference.

Lastly, we examine the density distributions of the mantissas of the logarithms of the user engagement metrics, as the stringent form of Benford's law requires a uniform distribution of the fractional part of the logarithms of a dataset (Berger & Twelves, 2018). This stronger form of Benford's law is fundamental and more rigorous than Eq. (2). Displayed in **Figure 3** are the density distributions for the animations project category. Note the large difference in uniformity between the densities of the mantissas of the views and remixes metrics shown in **Figure 3**. The standard deviations of the binned densities for the mantissas for the animations category are 6.1%, 6.9%, 14.5%, and 1.5% for loves, favorites, remixes, and views, respectively. The density distributions of the mantissas for all project categories are listed in **Table 3**. All six project categories follow the same general pattern for the standard deviations of the densities of the mantissas: views < loves < favorites < remixes. This pattern corroborates results presented in **Table 2** and part B in **Figure 2**.

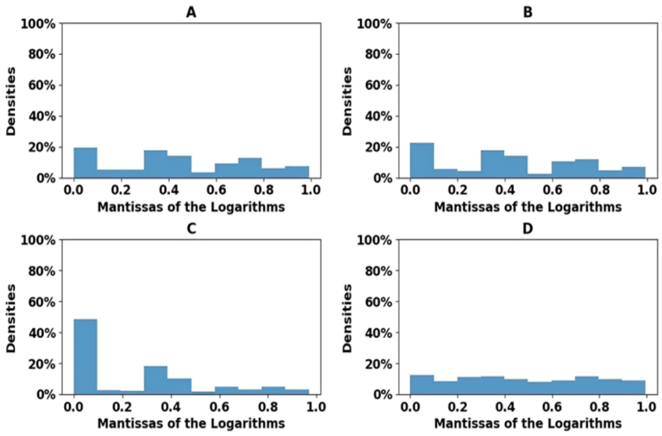

Figure 3. Density distributions for the mantissas of the logarithms of different user engagement metrics in the animations category: (A) loves, (B) favorites, (C) remixes, & (D) views. (Source: Author)

DISCUSSION

Though many studies have been conducted on user interaction data mined from mainstream social media platforms, to the best of our knowledge, this work is the first to extract and analyze user engagement metrics of an online coding community, Scratch, which combines both social and learning components. This work found that while all views metrics closely conform to Benford's law, approximately 58% of user engagement metrics across the six project categories were categorized into the non-conformity range with a MAD score above 0.015. It was determined that there is no statistically significant difference in Benfordness when comparing the six project categories (Kruskal-Wallis p-value = 0.68 >> 0.05; also see part A in **Figure 2**), but a great distinction among user engagement metrics averaged across the six project categories was found (Kruskal-Wallis p-value = 0.00042 << 0.05; also see part B in **Figure 2**).

One potential influence on these results can be attributed to Scratch's unique role as both a code-learning platform and a form of social media for young learners. A Scratch user may run a project as many times as they like, contributing to multiple recorded views on the project despite them all from one unique viewer. The total view counts of a project are the result of multiplying the number of viewers by the number of views per viewer, and as explained by Scott and Fasli (2001), mathematically, datasets adhere more closely to Benford's law when they are formed by multiplying two or more independent variables. Therefore, it is not surprising that the views metrics of all six project categories were found to both consistently and closely adhere to Benford's law. Although a Scratch user can view a project multiple times, projects are limited to at most one recorded love or favorite per user. As a result, the mechanism of enhancing Benfordness via the multiplication effect is not applicable to the loves and favorites metrics.

As previously stated, the remixes user engagement metric stands out from the other three as it is the metric most strongly associated with the code-learning behaviors of Scratchers. Mainstream social media platforms have components similar to Scratch's views, loves, and favorites metrics but do not have any features resembling the ability to remix projects. When a Scratch user plans to remix a project, they may have to study the project in-depth to understand where and how they can modify the project and add their own code. This technical experience offers valuable insights into programming logic, design principles, and problem-solving techniques, especially if the original project is complex. Unlike viewing and playing, remixing a project is associated with active code learning, often going beyond any built-in Scratch tutorials and more advanced online guides. By building upon existing projects, a user can experiment with new ideas, functionalities, and designs that are difficult to create without a template. Interestingly, the data shows that the remixes metric deviates the furthest from Benford's law, indicated by the Chi-square and MAD tests (see **Table 2** and **Figure 2**), as well as the standard deviations of the mantissas of the user engagement metrics (see **Table 3** and **Figure 3**).

Table 3. Standard deviations of densities (%) for mantissas of the logarithms

Category	User engagement	Standard deviations
Animations	Loves	6.1
	Favorites	6.9
	Remixes	14.5
	Views	1.5
Art	Loves	6.3
	Favorites	8.4
	Remixes	15.0
	Views	2.0
Games	Loves	8.2
	Favorites	9.0
	Remixes	13.9
	Views	2.1
Music	Loves	5.5
	Favorites	6.6
	Remixes	13.6
	Views	1.3
Tutorials	Loves	13.9
	Favorites	14.3
	Remixes	18.9
	Views	5.3
Stories	Loves	11.1
	Favorites	11.3
	Remixes	12.7
	Views	4.3

Other than the project-level user engagement metrics analyzed in this work, additional sources of information at the user level on the Scratch website could be studied, including follower count, following count, number of comments, and age of users. Scratch also has many other features, such as project studios where users can collaborate to build a collage of projects with a common theme or purpose. Additionally, Scratch has featured project groups aside from the trending page, such as the "What the Community is Loving" section and the "What the Community is Remixing" section which could be analyzed for Benford behavior. Data on these other features could be gathered and studied through data mining to provide further insights into peer interactions and learning experiences of Scratch users. Scratch also offers a comment section under each project, which can be analyzed using techniques such as natural language processing.

A straightforward extension of the current study to most online coding platforms is currently not feasible. For example, SNAP! (https://snap.berkeley.edu/) is closely modeled after Scratch and provides a very similar coding platform. It offers users options to view, embed, and download the code of published projects. But SNAP! does not report any user engagement data. There are no counts of views, loves, favorites, or remixes for any projects on SNAP!. Like Scratch, Thunkable (https://thunkable.com/solutions/education/) also uses a drag-and-drop coding environment, allowing learners to build video games by adding various components such as buttons, images, and voices, thus turning a passive online experience into an engaging educational one. However, Thunkable does not publish any user projects. Code.org offers a diverse array of instructional courses for teachers and students alike, with over 210 million projects created on the platform. However, like SNAP!, Code.org does not provide any user engagement metrics (https://studio.code.org/projects/public). On the other hand, GitHub provides public counts of stars and forks, akin to favorites and remixes in Scratch. However, no Benford analysis has been conducted on GitHub. In light of the numerous recent efforts in studying code-learning behaviors, it would be greatly beneficial for all educational coding platforms to provide user engagement data as these data can provide considerable insights to autonomous code-learning behaviors in online communities. Understanding these autonomous learning behaviors has many potential benefits, such as informing improved content and platform design, thereby aiding millions of online learners.

Although this study is limited to the projects featured on the trending page of Scratch (N = 21,473), the strategy used in this study can be extended to the entirety of Scratch. In particular, computational thinking skills, defined by abstraction and problem decomposition, logical thinking, synchronization, parallelism, algorithmic notions of flow control, user interactivity, and data representation, can be quantified (Moreno-León et al., 2019) and correlated with the extent of conformity to (or deviation from) Benford's law by user engagement metrics. The extent of conformity to Benford's law, in turn, can be used to gauge the development of computational thinking skills in an open online coding environment with peer support networks such as Scratch. A related study of academic publishing networks using Benford's law has been reported recently (Tošić & Vičič, 2021). Such studies may not only provide insights into the effectiveness of online platform design features but also offer opportunities for optimizing code-teaching strategies. As Benford's law is being applied to an ever-increasing array of diverse research areas, from detecting weak seismic signals (Díaz et al., 2015) to analyzing complex networks including social media networks (Morzy et al., 2016), the extent of conformity to Benford's law, including the complex patterns of mantissas of user engagement metrics as exemplified by **Figure 3**, may also serve as novel features in artificial intelligence models to study social interactions and autonomous learning behaviors in online coding communities.

In addition to the quantitative Benford analysis described here, incorporating qualitative data from users' comments is expected to provide a richer, more nuanced understanding of autonomous learning behaviors on the Scratch platform. Users often share a variety of insights in the Scratch comments section, including specific aspects of the projects they found engaging, innovative, or enjoyable. These comments can reveal users' emotional responses, preferences, and the educational value they perceive in the projects, offering important context and depth to the numerical data. By examining these comments, we can better understand why certain projects are more popular, shedding light on user preferences and the impact of specific project attributes. Furthermore, comments may include feedback on learning outcomes, such as how users apply the knowledge gained from others' projects to their own coding activities. Integrating these qualitative insights with the quantitative leading digit distributions can provide a more comprehensive view of user engagement, linking the observed Benford statistics to the underlying motivations behind autonomous code-learning behaviors.

Finally, the Benford analysis highlighted in this study can be applied to other online coding platforms such as GitHub and online learning environments other than computer programming. Outside of Scratch, other sources of qualitative data, particularly interviews, have already proven to be highly useful for gaining a deeper understanding of online coding communities (Dabbish et al., 2012). By analyzing qualitative data, educators can identify recurring themes and motivations that drive engagement, such as features that particularly resonate with students or elements of online educational platforms that foster learning. Therefore, combining quantitative Benford analysis with conventional student engagement metrics including qualitative data, for example, from student surveys and interviews is expected to be a fruitful approach to broadening our understanding of online learning activities beyond the Scratch community. This approach will not only enhance the interpretation of statistical analyses, including leading digit distributions, but also inform the design of various educational platforms to facilitate more effective and engaging learning experiences on a variety of educational platforms.

CONCLUSION

Our study using Benford's law revealed a highly significant difference between project viewing/playing and code remixing in the online Scratch community. This distinction in conformity to Benford's law is validated by multiple independent statistical analyses. Given that the extent of conformity to Benford's law has been utilized in various artificial intelligence applications (e.g., Caffarini et al., 2022; Hsu & Berisha, 2022), and artificial intelligence has been employed to study code-learning behaviors in supervised settings (e.g., Lin et al., 2024), our findings suggest that combining Benford's law and artificial intelligence presents a promising approach to advancing our understanding of autonomous code-learning behaviors and peer interactions in open online coding communities. These avenues of research are anticipated to foster new opportunities for optimizing the content and design of online coding platforms to benefit millions of users engaged in autonomous learning.

Funding: No external funding is received for this article.

Acknowledgments: The author would like to thank Mr. Steven Walfish for valuable discussions and comments. The author would also like to thank the anonymous reviewer for providing numerous helpful comments and suggestions, which have greatly improved and strengthened this study.

Ethics declaration: The author declares that this study does not require ethical approval since the information is freely available in the public domain (https://scratch.mit.edu/).

Declaration of interest: The author declares no competing interests.

Availability of data and materials: All data generated or analyzed during this study are available for sharing when appropriate request is directed to corresponding author.

REFERENCES

Abkenar, S. B., Kashani, M. H., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics,* 57, Article 101517. https://doi.org/10.1016/j.tele.2020.101517

Alebaikan, R., Alajlan, H., Almassaad, A., Alshamri, N., & Bain, Y. (2022). Experiences of middle school programming in an online learning environment. *Behavioral Science, 12*(11), Article 466. https://doi.org/10.3390/bs12110466

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society, 78*(4), 551–572.

Berger, A., & Twelves, I. (2018). On the significands of uniform random variables. *Journal of Applied Probability, 55*(2), 353–367. https://doi.org/10.1017/jpr.2018.23

Bhosale, S., & Di Troia, F. (2022). Twitter bots' detection with Benford's law and machine learning. In L. Bathen, G. Saldamli, X. Sun, T. H. Austin, & A. J. Nelson (Eds.), *Silicon Valley Cybersecurity Conference. SVCC 2022. Communications in computer and information science, vol 1683* (pp. 38–54). Springer. https://doi.org/10.1007/978-3-031-24049-2_3

Bock, D. E., Velleman, P. F., & De Veaux, R. D. (2019). *Stats: Modeling the world*. Pearson.

Caffarini, J., Gjini, K., Sevak, B., Waleffe, R., Kalkach-Aparicio, M., Boly, M., & Struck, A. F. (2022). Engineering nonlinear epileptic biomarkers using deep learning and Benford's law. *Scientific Reports, 12*, Article 5397. https://doi.org/10.1038/s41598-022-09429-w

Chen, H. M., Nguyen, B. A., Yan, Y. X., & Dow, C. R. (2020). Analysis of learning behavior in an automated programming assessment environment: A code quality perspective. *IEEE Access, 8*, 167341–167354. https://doi.org/10.1109/ACCESS.2020.3024102

Crocetti, E., & Randi, G. (2016). Using the Benford's law as a first step to assess the quality of the cancer registry data. *Frontiers in Public Health, 4*, Article 225. https://doi.org/10.3389/fpubh.2016.00225

da Silva Azevedo, C., Gonçalves, R. F., Gava, V. L., & de Mesquita Spinola, M. (2021). Benford's law based method for fraud detection using R library. *MethodsX, 8*, Article 101575. https://doi.org/10.1016/j.mex.2021.101575

da Silva, A. J., Floquet, S., Santos, D. O. C., & Lima, R. F. (2020). On the validation of the Newcomb-Benford law and the Weibull distribution in neuromuscular transmission. *Physica A: Statistical Mechanics and its Applications, 552*, Article 124606. https://doi.org/10.1016/j.physa.2020.124606

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 1277–1286). ACM. https://doi.org/10.1145/2145204.2145396

Díaz, J., Gallart. J., & Ruiz. M. (2015). On the ability of the Benford's law to detect earthquakes and discriminate seismic signals. *Seismological Research Letters, 86*(1), 192–201. https://doi.org/10.1785/0220140131

Druică, E., Oancea, B., & Valsan, C. (2018). Benford's law and the limits of digit analysis. *International Journal of Accounting Information System, 31*, 75–82. https://doi.org/10.1016/j.accinf.2018.09.004

Erümit, A. K., & Şahin, G. (2020). Plugged or unplugged teaching: A case study of students' preferences for the teaching of programming. *International Journal of Computer Science Education in Schools,* 4(1), 3–32. https://doi.org/10.21585/ijcses.v4i1.82

Felt, M. (2016). Social media and the social sciences: How researchers employ big data analytics. *Big Data & Society*, 3(1). https://doi.org/10.1177/2053951716645828

Hsu, L., & Berisha, V. (2022). Does human speech follow Benford's law? *arXiv*. https://doi.org/10.48550/arXiv.2203.13352

Kwon, K., & Cheon, J. (2019). Exploring problem decomposition and program development through block-based programs. *International Journal of Computer Science Education in Schools, 3*(1), 3–16. https://doi.org/10.21585/ijcses.v3i1.54

Kwon, K., Lee, S. J., & Chung, J. (2018). Computational concepts reflected on Scratch programs. *International Journal of Computer Science Education in Schools, 2*(3). https://doi.org/10.21585/ijcses.v2i3.33

Lin, C.-C., Cheng, E. S. J., Y.Q. Huang, A. Y. Q., & Yang, S. J. H. (2024). DNA of learning behaviors: A novel approach of learning performance prediction by NLP. *Computers and Education: Artificial Intelligence, 6*, Article 100227. https://doi.org/10.1016/j.caeai.2024.100227

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic analysis of scratch projects to assess and foster computational thinking. *Revista de Educación a Distancia, 46*.

Morzy, M, Kajdanowicz, T., & Szymański. B. (2016). Benford's distribution in complex networks. *Scientific Reports, 6*, Article 34917. https://doi.org/10.1038/srep34917

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics, 4*(1), 39–40. https://doi.org/10.2307/2369148

Nigrini, M. J. (2012). *Benford's law*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119203094

Scott, P. D., & Fasli, M. (2001). Benford's law: An empirical investigation and a novel explanation. *University of Essex*. https://repository.essex.ac.uk/8664/1/CSM-349.pdf

Seenivasan, P., Easwaran, S., Sridhar, S., & Sinha, S. (2016). Using skewness and the first-digit phenomenon to identify dynamical transitions in cardiac models. *Frontiers in Physiology, 6*, Article 390. https://doi.org/10.3389/fphys.2015.00390

Stewart, A. E. B., Solyst, J., Buddemeyer, A., Hatley, L., Henderson-Singer, Scott, K., Walker, E., & Ogan, A. (2021). Explaining engagement: Learner behaviors in a virtual coding camp. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *Artificial Intelligence in Education. AIED 2021. Lecture Notes in Computer Science(), vol 12749* (pp. 338–343). Springer. https://doi.org/10.1007/978-3-030-78270-2_60

Su, Y.-S., & Lai, C.-F. (2021). Applying educational data mining to explore viewing behaviors and performance with flipped classrooms on the social media platform Facebook. *Frontiers in Psychology, 12*, Article 653018. https://doi.org/10.3389/fpsyg.2021.653018

Sun, D., Ouyang, F., & Li, Y. (2021). Comparing learners' knowledge, behaviors, and attitudes between two instructional modes of computer programming in secondary education. *International Journal of STEM Education, 8*, Article 54. https://doi.org/10.1186/s40594-021-00311-1

Tošić, A., & Vičič, J. (2021). Use of Benford's law on academic publishing networks. *Journal of Informetrics, 15*, Article 101163. https://doi.org/10.1016/j.joi.2021.101163

Varga, D. (2021). Analysis of Benford's law for no-reference quality assessment of natural, screen-content, and synthetic images. *Electronics, 10*(19), Article 2378. https://doi.org/10.3390/electronics10192378

APPENDIX A

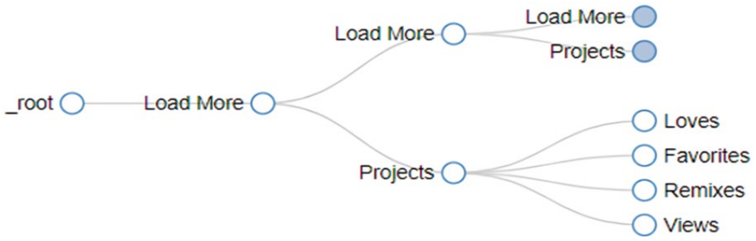

Figure A1. Selector graph for recursive pagination (the filled circles indicate nodes for recursive extension of the algorithmic paths)

Table A1. Standardized residuals of each leading digit from Pearson's χ^2 test

Category	User engagement	1	2	3	4	5	6	7	8	9
Animations	Loves	−0.35	2.21	2.12	−0.56	0.54	−1.60	−1.08	−2.54	−1.00
	Favorites	1.26	1.72	1.83	0.74	−0.35	−2.20	−2.45	−2.53	−2.18
	Remixes	8.79	0.88	−1.12	−3.05	−3.48	−3.01	−2.68	−3.32	−3.22
	Views	1.28	0.15	−1.04	−1.19	−0.55	0.13	1.93	−1.31	−0.35
Art	Loves	−4.79	2.99	7.07	2.67	2.12	0.22	−2.74	−2.86	−6.08
	Favorites	2.05	9.27	5.18	−0.62	−3.89	−4.90	−4.35	−6.73	−8.05
	Remixes	14.89	−1.22	−3.31	−4.20	−4.89	−4.48	−2.73	−4.40	−4.63
	Views	4.09	−1.80	−3.47	−2.93	−1.47	−0.31	0.37	2.32	2.49
Games	Loves	2.45	6.90	1.46	−0.52	−2.42	−3.33	−3.77	−4.52	−5.27
	Favorites	3.85	7.49	0.97	−1.26	−2.01	−4.53	−5.12	−4.94	−5.20
	Remixes	11.10	1.01	−2.99	−2.81	−4.12	−3.42	−4.11	−3.53	−3.52
	Views	−0.77	1.19	−0.53	−0.92	−0.74	−0.04	0.92	1.45	0.29
Music	Loves	−3.03	0.82	4.10	2.16	1.35	−0.77	−0.51	−2.12	−1.81
	Favorites	−2.59	3.26	4.17	1.84	1.32	−0.26	−1.96	−4.12	−4.16
	Remixes	9.70	0.24	−1.09	−3.07	−3.46	−3.97	−3.12	−2.85	−3.22
	Views	0.22	−0.28	−0.38	0.08	−1.76	−0.93	2.34	2.20	−1.04
Tutorials	Loves	15.63	3.30	−2.43	−4.72	−5.70	−6.60	−5.38	−6.72	−7.05
	Favorites	15.83	3.53	−2.57	−4.27	−6.77	−6.24	−6.10	−6.57	−6.79
	Remixes	11.51	−1.02	−1.66	−3.76	−3.20	−3.46	−2.81	−3.89	−3.65
	Views	3.33	0.95	−0.31	−1.19	−1.18	−1.68	−2.06	−0.90	−1.30
Stories	Loves	22.94	1.30	−3.26	−5.13	−7.99	−9.75	−7.22	−8.56	−9.06
	Favorites	21.62	2.54	−2.14	−6.84	−7.31	−7.71	−9.61	−9.00	−7.66
	Remixes	15.07	−0.80	−2.43	−4.84	−4.05	−5.29	−3.16	−5.21	−5.22
	Views	4.57	1.50	1.01	−1.29	−2.62	−1.51	−1.35	−3.27	−4.18