

Programming with ScratchJr: a review of the first year of user analytics

Kaitlyn D. Leidl^{1*}, Marina Umaschi Bers², Claudia Mihm³
 The DevTech Research Group^{1,2,3}
 Tufts University, United States^{1,2,3}

Kaitlyn.Leidl@tufts.edu, Marina.Bers@tufts.edu, Claudia.Mihm@tufts.edu

ABSTRACT

ScratchJr is a free programming application for young children ages 5-7, available for most tablet devices. This programming environment, developed by the DevTech Research Group at Tufts University, the Lifelong Kindergarten Group at MIT, and the Playful Invention Company, was launched in July, 2014. During the first year after the app's launch, no information was collected regarding usage other than informal communication with local educators and parents. Starting in January 2016, the ScratchJr team began to use the tool Google Analytics to gain a deeper insight into user behavior, and began to investigate the learning analytics data that could shed light on computational thinking in early childhood. This paper presents the first year of user data collection of ScratchJr.

KEYWORDS

Computational thinking, programming, early childhood, educational technology, analytics.

1. INTRODUCTION

ScratchJr is a free tablet app that provides an introductory programming environment for young children ages 5-7. It was developed as a collaboration between the DevTech Research Group at Tufts University, the MIT Lifelong Kindergarten Group, and the Playful Invention Company, with funding from the National Science Foundation (DRL-1118664). ScratchJr was first launched as a freely downloadable app on iPads in July, 2014, and has since been released for use on several other platforms including Android tablets, Amazon tablets, and Chromebooks. Used in classrooms and homes worldwide, ScratchJr enables children to create interactive stories and games by snapping together graphical programming blocks to make characters move, jump, dance, and sing. As shown in Figure 1, the ScratchJr interface allows children to use blocks that control motion, looks, sound, character communication, and more. Through these programming blocks, young children learn the basic concepts and powerful ideas of coding while creating personally meaningful projects (Bers, 2017). The programming app has been widely available for over two years, and in that time, educators, parents, and children around the world have used it to expand the range of creative programming projects and to connect coding to traditional school subjects such as science, mathematics, literacy, history, and more (Bers & Resnick, 2015).



Figure 1. ScratchJr programming app interface

In January, 2016, the ScratchJr team was able to integrate Google Analytics to examine how people are using the programming app. This tool has allowed the team to better understand when and where ScratchJr is being used, which programming blocks are most popular, how many projects are being created, and how long users spend during sessions with ScratchJr. Data collected for a year provides insights into early coding and computational thinking.

2. COMPUTATIONAL THINKING IN EARLY CHILDHOOD

ScratchJr was developed to encourage all young children to engage in computational thinking while coding. Within the open-ended programming environment, children learn the basic powerful ideas of computer science, such as algorithms, debugging, and modularity by snapping together programming blocks. While programming in ScratchJr, children think creatively, logically, and sequentially (Bers, 2008, 2012, 2017). Computational thinking has the potential to benefit all individuals as it involves understanding sequencing and order, as well as logical thinking. This type of thinking is involved in many everyday tasks, such as learning the steps to ride a bike, following a recipe, or editing and rewriting a research paper (Bers, 2017; Wing, 2006).

When computational thinking is supported at a young age by teaching children about coding, it has the potential to supplement and solidify many other social and behavioral skills, which will be valuable to society whether or not the child becomes an engineer or a computer scientist in the future (Wing, 2006). Therefore, we designed ScratchJr to be a developmentally appropriate programming language to engage children in computational thinking, and to provide a

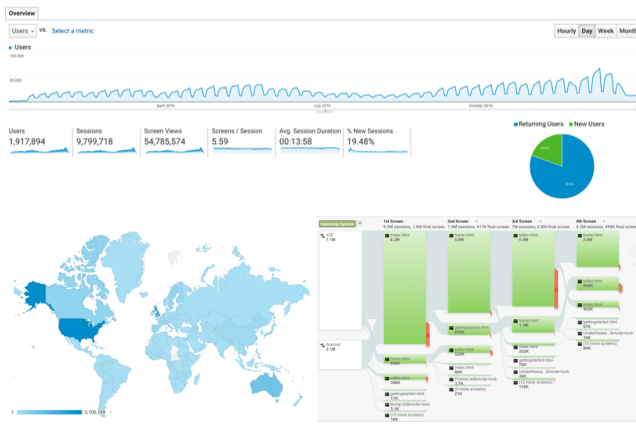


Figure 3. Google Analytics visualization data for ScratchJr

4.2. Using Analytics in Education

Data analytics tools can be used for a myriad of reasons. Large companies and small businesses alike often turn to data collection tools to redefine marketing strategies, increase revenue, and utilize user behavior patterns to improve overall user experience (Luo et al., 2015; Martin et al., 2015).

However, more recently in research, examples of studies around analytics data have emerged in the realm of education. In this context, the practice is known as learning analytics, and focuses on learners, the process of learning over time, and the context in which learning takes place (Baker & Inventado, 2014; Berland, Martin, Benton, Petrick Smith, & Davis, 2013; Luo et al., 2015).

As more data about learners becomes available due to the increased amount of and access to online courses, educational traffic on the web, and educational software and technology, more opportunities to expand educational research have subsequently emerged (Greller & Drachsler, 2012). Being able to transform the abstract progression of learning into tangible numbers and visual data gives researchers insight into learning patterns that could have major implications on the way educators teach core subjects in schools (Greller & Drachsler, 2012).

There have been several recent studies that use learning analytics to track how individuals learn how to program (Baker & Inventado, 2014; Berland et al., 2013; Blikstein et al., 2014). A common method of gathering data is taking screenshots of students' code generation over a period of time (Berland et al., 2013; Blikstein et al., 2014). Researchers can use computer algorithms to categorize these screenshots in terms of programming development by asking questions such as: how did the code change in complexity, length, and content over time? How did these changes impact the effectiveness of the programs overall? Did the later programs indicate growth in programming knowledge? (Berland et al., 2013; Blikstein et al., 2014). By using computer programs to quantify learning curves among students while they learn programming languages, researchers are uncovering learning patterns that could

have major implications on how we teach computer science in educational institutions (Blikstein et al., 2014).

Using data analytics in ScratchJr, we have gained insight into how the number of users, sessions, and locations has evolved over the course of one year. In this paper, we report these results.

4.3. ScratchJr in Google Analytics

Since January, 2016, the ScratchJr team has utilized the Google Analytics program to gain a better understanding of how ScratchJr is used across the globe. Although tools like Google Analytics are often used by businesses to track revenue and improve marketing strategy (Google Analytics Solutions, 2017; Luo et al., 2015), in the case of ScratchJr, our focus is on user behavior, location, patterns in new user acquisition, and the app features themselves. To protect the privacy of our young users, we do not collect personally identifying information, such as unique project content. Therefore, using Google Analytics alone, we cannot track the progression of project content and programming behaviors of individual users over time. Instead, we focused on data points presented and defined in Table 1:

Table 1. Data points and definitions (Google Inc., 2016)

Name of Data Point	Definition of Data Point
Session	The period time a user is actively engaged with the website, app, etc.
Users	Users that have had at least one session within the selected date range. Includes both new and returning users.
Returning Users	A user with existing Google Analytics cookies from a previous visit.
New Users	The number of first-time users during the selected date range. A new user is one who did not have Google Analytics cookies when they first opened the app. If a user deletes their cookies and re-opens the app, they will be counted as a new user.
Average Session Duration	The average length of a session.
Events	The categories that were assigned to triggered events.
Language	The language settings in the users' browsers. Analytics uses ISO codes.
Location	The location from which the session originated.
Real-Time	Data updates continuously and each pageview is reported seconds after it occurs. Shows the number of people on the app right now, their geographic locations, etc.

Through Google Analytics, the team collects data on where users click or tap within the app, which parts of the app

users use, and geographic location of where the app is being used based on device IP addresses and network location. The “click data” helps the team determine ways to improve both the app interface and available learning and teaching resources. The geographic location data helps to understand where ScratchJr is and is not being used. ScratchJr does not share any specific user information it collects with Google, and Google does not collect any personally identifying information about users.

Since January, 2016, the ScratchJr team has been gaining insight into how users, both adults and children alike, use the app. Google Analytics allows teams to see reports from any date range, and view data in terms of hours, days, weeks, and months. This allows the ScratchJr team to refine data regarding time in meaningful ways. For example, it is possible to determine which month, week, day, or hour is the most popular time to use ScratchJr, in terms of both how many users are active at those times, and how many sessions occur in those times. This has been especially useful to gauge the impact of computer science and programming education events that occur around the world in which ScratchJr is present. In observing the hourly and weekly data patterns of when ScratchJr is used, we can infer if the majority of children are programming with ScratchJr in classrooms with educators or in their homes with family.

Throughout 2016, the ScratchJr team discovered several notable patterns in behavior of ScratchJr users, and has subsequently begun to make steps towards improving the app and its resources.

5. FINDINGS

Overall, the average amount of sessions and number of users increased as the year progressed, yet other data points such as average session duration, percentage of new users per week, and users per week remained consistent. These are all telling data points regarding user loyalty to the programming app. The analytics highlights from 2016 are described in the following subsections.

5.1. Users & Sessions

There were nearly 2 million total ScratchJr users in 2016. There were more than 104,000 average active users per week, and nearly 27,000 average users on Thursdays alone, the most popular day to use ScratchJr in 2016. Only slightly more sessions occurred on Thursdays in 2016 than on Fridays (see Figure 4).

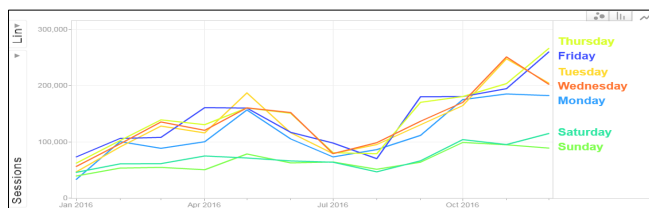


Figure 4. The most sessions occurred on Thursdays

The number of sessions on each of those days of the week came to over 1.7 million. The time of day that saw the most sessions in ScratchJr was 9:00 AM EST (7.83% of the total sessions occurred during this hour). Spikes and patterns in weekly and hourly users are shown in the graphs in Figures 5 and 6. Consistently, 20% of users each week were new to ScratchJr, and 80% were returning users.

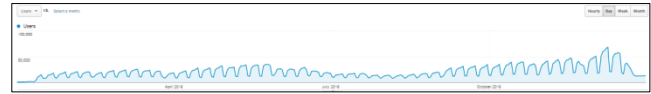


Figure 5. Google Analytics ScratchJr users daily view: peaks tend to be Thursdays and Fridays, low points tend to be Saturdays and Sundays.



Figure 6. Google Analytics ScratchJr users hourly view: peaks tend to be at 9:00 AM EST and 2:00 PM EST; low points tend to be at 11:00 PM EST and 12:00 AM EST.

There was an average of nearly 37,000 new users to ScratchJr each week in 2016. The week that recorded the most new users was December 4-10, 2016, with nearly 97,000 (see spike on right side of graph in Figure 7). This week was “Computer Science Education Week” in the United States, in which government officials encouraged engagement in programming in classrooms, and websites like Code.org provided numerous resources for learning how to code, including ScratchJr lesson plans (Code.org, 2016; Computer Science Education Week, 2016; The White House, 2016). Furthermore, the DevTech Research Group at Tufts University created ScratchJr videos teaching pillars of computational thinking, or “powerful ideas” (Bers, 2017; Papert, 1980), which were viewed hundreds of times, indicating a definite presence of the programming app in the United States throughout the week (DevTech Research Group, 2016).



Figure 7. New Users per week in 2016; spike on right side of graph indicates Computer Science Education week in the U.S., which brought many new users to ScratchJr.

In 2016, there were nearly 9.8 million recorded sessions in ScratchJr. The average session duration was 13 minutes and 58 seconds. Users averaged viewing 5.6 screens per session. The most common flow of screens for both iOS and Android operating systems began with the Index screen that appears when users first open the app, followed by the Home Lobby screen, then the Editor to create programs,

followed by the Home Lobby screen again, and then the Editor again. A smaller percentage of users went from the Index screen to the “Getting Started” screen to learn how to use ScratchJr.

5.2. Programming Projects Content

The year 2016 saw over 7.5 million projects created in ScratchJr. Furthermore, there were over 9 million existing projects edited, showing that users tend to go back into projects to work on them. There were 254,000 ScratchJr projects shared via either email or Apple AirDrop in 2016.

In 2016, there were nearly 148 million ScratchJr programming blocks added by users to the programming area in the app. The ten most popular programming blocks added were the Forward block (25 million added), Start on Green Flag, Move Up, Move Back, Say (a speech bubble block that allows characters to converse), Record Block (allows users to record their own sounds and add them into their program), Move Down, Shrink, Turn Right, and Grow. The least popular blocks were Reset Size, Send Message, Start on Message, Start on Bump, and Stop (Figure 8).



Figure 8. Most and least popular ScratchJr blocks in 2016

The most popular characters used by children in 2016 were those self-created or edited by the children in the Paint Editor, the Child, the Teen, Tac, and the Dragon (Figure 9). Users entered the Paint Editor to customize their characters and backgrounds over 23 million times.



Figure 9. Most popular characters in ScratchJr in 2016

The ScratchJr app includes eight sample projects to provide examples of programs users can make. These sample projects were viewed 1.6 million times in 2016.

5.3. Location & Language

In 2016, ScratchJr was used in all 50 states in the United States of America, and in all but five countries worldwide. The top 10 countries using ScratchJr based on the number of sessions recorded are displayed in Table 2.

The top language codes on devices using ScratchJr include English-US, English-Great Britain, English, English-Australia, Spanish-Spain, English-Canada, Swedish-Sweden, French-France, Korean-Korea, and Finnish-Finland.

Country	% of Total Sessions
United States	31.65%
United Kingdom	17.35%
Australia	10.32%
Canada	4.33%
Sweden	3.30%
Spain	3.16%
Finland	2.52%
France	2.28%
South Korea	2.24%
China	2.10%

Table 2. Top nations using ScratchJr

6. CONCLUSION

In using Google Analytics, the ScratchJr team is able to understand user behavior in a quantitative way. The team has been able to better comprehend the global reach of ScratchJr, using location and language statistics to determine the best methods for localization of ScratchJr. In learning that ScratchJr was used in 191 of 196 registered countries worldwide in 2016, the importance of and demand for computer science education across the globe became clear.

Furthermore, the tremendous growth in numbers during Computer Science Education Week in December, 2016 is an indication that ScratchJr was a popular vessel for learning about computer science and programming when classrooms reserved the time to teach the topics. This gives the team reason to continue making resources available for educators and parents, particularly during national and global initiatives to promote computer science.

Data that the ScratchJr team has gathered about when ScratchJr is used also gives a unique insight into how to support users. Thursdays and Fridays were the two most popular days for ScratchJr in 2016, and the most popular time of day was around 9:00 AM EST. This suggests teachers are using ScratchJr on a weekly basis towards the end of the week and in the mornings. The ScratchJr team could use this information to promote educational resources, tips, and ideas for ScratchJr at these times.

Although the ScratchJr team does not collect individual projects and thus cannot currently see the learning progression of users programming in ScratchJr, there are many insights we can still gain by having visual and numerical data. Based on the data we collected in 2016, it is clear that educators, parents, and children are finding

ways to learn programming, and ScratchJr has the potential to be one of the leading platforms young children use to engage in computational thinking.

7. FUTURE WORK

Moving forward, the ScratchJr team will continue to use Google Analytics to better understand user behavior. The team will use the data gathered to optimize localization efforts, and provide resources based on project content trends. Using data regarding the most popular days and times of day ScratchJr is used, the team will use social media outlets to support educators who may be teaching with the programming app at those times, and continue to build a ScratchJr community for users to share their ideas and experiences. Furthermore, the team will develop surveys to gather data that is not currently collected to be able to start inferring learning trajectories.

8. REFERENCES

- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75). Springer New York.
- Berland, M., Martin, T., Benton, T., Petrick Smith, C., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences*, 22(4), 564-599.
- Bers, M. (2008). *Blocks to Robots: Learning with Technology in the Early Childhood Classroom*. Teachers College Press, NY, NY.
- Bers, M.U. (2012). *Designing digital experiences for positive youth development: From playpen to playground*. Cary, NC: Oxford.
- Bers, M.U. (2017). *Coding as a Playground: Programming and Computational Thinking in the Early Childhood Classroom*. Routledge Press.
- Bers, M.U. & Resnick, M. (2015). *The Official ScratchJr Book*. San Francisco, CA: No Starch Press.
- Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *Journal of the Learning Sciences*, 23(4), 561-599.
- Clark, D. J., Nicholas, D., & Jamali, H. R. (2014). Evaluating information seeking and use in the changing virtual world: the emerging role of Google Analytics. *Learned Publishing*, 27(3), 185-194.
- Code.org (2016). *Hour of Code*. Retrieved from <https://code.org/learn>.
- Computer Science Education Week (2016). *Computer Science Education Week*. Retrieved from <https://csedweek.org/>.
- DevTech Research Group (2016). *DevTech Celebrated Computer Science Education Week!* Retrieved from <http://ase.tufts.edu/DevTech/CSEdWeek2016.html>.
- Flannery, L.P., Kazakoff, E.R., Bonta, P., Silverman, B., Bers, M.U., & Resnick, M. (2013). Designing ScratchJr: Support for early childhood learning through computer programming. In *Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13)*. ACM, New York, NY, USA, 1-10. DOI=10.1145/2485760.2485785
- Google Analytics Solutions (2017). *Success Stories*. Retrieved from https://www.google.com/analytics/success-stories/?product=analytics#?modal_active=none.
- Google, Inc. (2016). Google Analytics Solutions. Retrieved from: <https://www.google.com/analytics/>.
- Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, 15 (3), 42-57.
- Luo, H., Rocco, S., & Schaad, C. (2015, October). Using Google Analytics to Understand Online Learning: A Case Study of a Graduate-Level Online Course. In *2015 International Conference of Educational Innovation through Technology (EITT)* (pp. 264-268). IEEE.
- Martin, T., Petrick Smith, C., Forsgren, N., Aghababayan, A., Janisiewicz, P., & Baker, S. (2015). Learning fractions by splitting: Using learning analytics to illuminate the development of mathematical understanding. *Journal of the Learning Sciences*, 24(4), 593-637.
- Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York, Basic Books.
- ScratchJr (2017). *Learn Page*. Retrieved from <http://scratchjr.org/>.
- Strawhacker, A., Lee, M., Caine, C., & Bers, M.U. (2015). ScratchJr Demo: A coding language for Kindergarten. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC '15)*. ACM, Boston, MA, USA.
- The White House (2016 December 5). FACT SHEET: A Year of Action Supporting Computer Science for All. *The White House Office of the Press Secretary*. Retrieved from <https://obamawhitehouse.archives.gov/the-press-office/2016/12/05/fact-sheet-year-action-supporting-computer-science-all>.
- Wing, J. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-36.