Habits of Highly Successful Professional Learners and the Corresponding Online Curriculum

Rob Rubin, Alicia Redmond, Gregory Weber, Gulrez Khan

Microsoft Corporation

rorubin@microsoft.com, alboaz@microsoft.com,
gregwe@microsoft.com, gulkh@microsoft.com

ABSTRACT. Microsoft introduced the Microsoft Professional Program (MPP) in Data Science on edX -- a fourteen course online offering of which passing nine is necessary for successful completion, including a mandatory capstone project, to address the growing gap between the number of technology jobs available and the number of candidates qualified to fill those roles. In this paper, we take a first look at the descriptive analytics of a highly-motivated cohort who demonstrated success in this program to understand 1) the components to create an engaging course and curriculum, as measured by high levels of completion and learned content, and 2) the attributes contributing to learner success in completing a course and the entire curriculum.

Keywords: Online learning; MOOC; professional learning; effective habits of learning.

1 Introduction

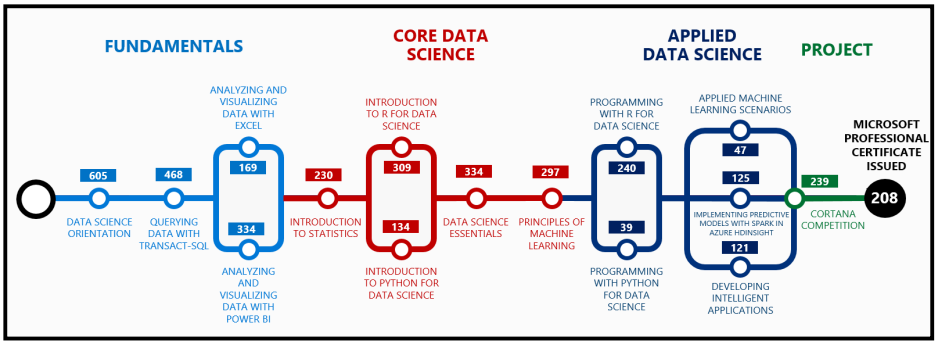
A significant challenge in industry and academia that can be addressed by MOOC's is training employees and to take on new roles in the workforce. Per the McKinsey Global Institute's research [1] by 2018 the United States will experience a shortage of 190,000 skilled data scientists, and 1.5 million managers and analysts capable of actionable insights. In 2016, Microsoft introduced the Microsoft Professional Program (MPP) in Data Science. The Microsoft Professional Program offers learners an employer-endorsed credential who have completed a rigorous curriculum that focuses on both conceptual knowledge, product capabilities, and application of their newly acquired skills to real world labs and problems. These curricula are designed to be delivered online with a 4-8 hour per week commitment, over 40 weeks, where videos are provided by working professionals and experts alike [2]. Each course includes videos, online assessments and labs with a discussion forum to seek help from fellow classmates and

Workshop on Integrated Learning Analytics of MOOC Post-Course Development. Workshop at the 7th International Learning Analytics and Knowledge Conference. Simon Fraser University, Vancouver, Canada, 13-17 March 2017, p. 112-121.

teaching assistants, including both formative and summative assessments [3]. A capstone project is included as part of the curriculum that provides a real-world exercise for the learner to demonstrate the skills they've learned through the program.

Our research agenda is focused on examining a multi-course offering to solve an in-demand industrial need. By contrast there has been a large focus on individual MOOC's [4]. We think multi-course offerings are an area ripe for investigation. While our research is early in development, we share descriptive analytics on the self-selecting cohort of successful learners who elected to participate in the MPP and hope this stimulates research into multi-course offerings such as the edX X-Series or online programs and Micro-degrees. We begin our study with a focus on the visualization of the program and course structure [5] to understand the impact of both program and course organization and design on learner performance [6], and we look at the behavior of the successful learners.

Fig. 1. MPP Data Science curriculum, including overarching concepts, course names and number of course completions by pilot participants (as of February 2017). At four points in the learning path, the learner has the option to choose one course out of two or three options.



Our target demographics for the MPP were informed by a focus on the growing millennial workforce, working professionals in existing job roles, and early analysis of the gateway courses to Data Science. The demographics we established are:

1. Employed technology professionals seeking to change roles, solve more meaningful problems, and/or stay current on technology.
2. Millennials in STEM seeking entry into a technology field.
3. Women professionals in role seeking to transition into STEM roles

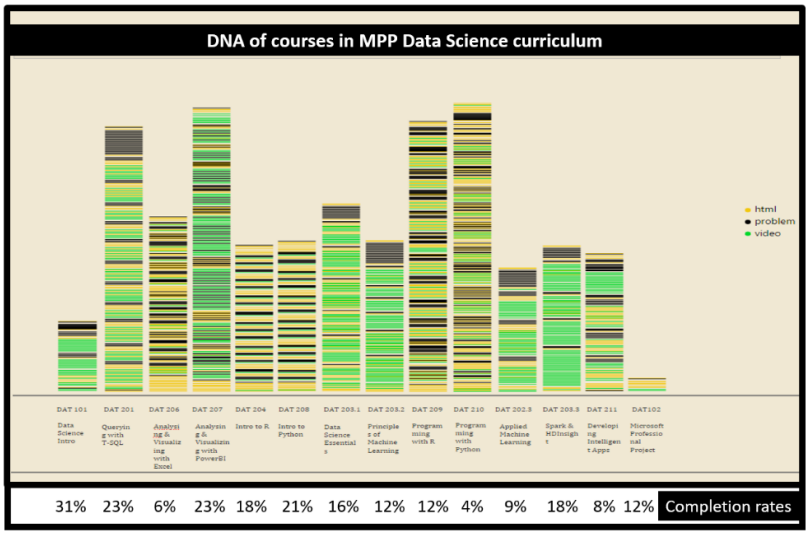
The first cohort was a pilot that consisted of full time employees. The curriculum was launched in May 2016 and participants were instructed to complete the curriculum by either September 2016 or January 2017. In admitting learners to the two cohort tracks, we did not control for level of experience or skills. Participation in the pilot cohorts was voluntary and free, but support from their manager was encouraged.

1.1 Pilot Curriculum Performance vs MOOC Performance

One critical measure for the pilot was course completions because learners were required to complete 9 of 14 courses in a curriculum to qualify for a credential. Course completion is measured as a learner earning a passing grade on the course, which is 70% on average for all courses. Figure 2 contains an outline of the first curriculum published in Data Science. Completion rates are calculated by dividing the number of learners who earn a passing grade on the course divided by the number of learners who started the course by accessing the course content. While most MOOC platforms cite completion rates between 2-7%, the publicly-offered courses included in the MPP Data Science curriculum on edX are experiencing an average completion rate of 14% with a range of 4-31% [4].

The learners are required to complete the entire curriculum to be eligible for the degree. Given that edX cites a 20% attrition rate course over course in its X-series courses, which commonly contain 3-4 courses in a series, we were concerned about the attrition rate through a curriculum that required completion of 9 courses from 14 available [7]. To date, the pilot has experienced an average of 12% attrition rate between courses offered in the curriculum. The engagement level of the courses comprising the curriculum is critical to drive the performance of these key measures. Therefore, the analysis in progress includes identifying attributes of courses with high completion rates and determine the areas within a course in which intervention is required to support the learner to persevere to completion.

Fig. 2. Course DNA for 14 courses included in the MPP Data Science curriculum. Includes completion rates from July-February 2017 for courses open to MPP participants and a public audience on edX.



1.2 Methodology

For this analysis, we are leveraging a total data set of 782 learners who agreed to participate in a pilot run of a Data Science curriculum in the Microsoft Professional Program, which we intend to make available as an anonymized data set, as researchers elsewhere have benefitted enormously [8]. The participants resided in 45 different countries - 15% were female, 75% were male and 10% did not disclose their gender. The age distribution is 41% under 35 years old and 59% are 35 years old and over. The educational background of the participants was 5% held a PhD or doctorate, 43% held a Master's degree, 45% held a Bachelor's degree, and 7% held other education.

Within the pilot period between July 2016 and January 2017, a subset of learners completed the curriculum of 9 courses, who we refer to as "Highly Successful Learners", while another subset completed between 4-8 courses, who we refer to as "Successful Learners". Our initial analysis focused on these two subsets of learners to better understand what behaviours they exhibited to be successful in the MPP Data Science online curriculum. These two groups are compared against each other to determine whether there are behaviours exhibited by learners who completed the curriculum within 6 months of the pilot launch that made them highly successful in the program compared to those that made it partially through the curriculum.

- 208 Highly Successful Learners who completed the 9-course curriculum
- 135 Successful Learners who completed between 4-8 courses in the curriculum.
- 439 Learners are considered in progress as of February 2017 and are not included in the following analysis on leaner behaviours and strategies

2 Analysis Areas

The following areas describe the analysis in progress to evaluate the learner behaviour through the MPP curriculum for highly successful learners who complete the curriculum compared to the learners who only complete a subset of the curriculum. Additional analysis areas will be evaluated and included as the research continues.

3 Impact of Course Design on Learner Performance

To better understand the impact of course design on performance, we developed a visualization of the course structure. Figure 2 shows that the two highest performing courses had completion rates of 21% and 25% respectively. These two courses shared essentially the same course structure. This structure consisted of brief video's, textual challenge descriptions, and programming problems. Only the introductory course had a higher completion rate than these two courses.

4 Session Time

The time spent in a learning session is based upon the difference in logged event times continuing until there is at least 30 minutes of no logged activity [7]. Note that the total lab time may not be accurately reflected since labs were performed outside of the learning platform. The analysis compares the session time between Highly Successful Learners who completed all 9 courses in the MPP curriculum and Successful Learners who completed between 4-8 courses in the curriculum.

Table 1. Median hours in learning session time per course completed, comparing Highly Successful and Successful Learners.

Median hours per course	Total time	Video time	Problem time	Forum time
Highly Successful	7.1	4.5	1.7	0
Successful	7.8	5.1	1.9	0

There was less time spent per course for Highly Successful Learners (7.1 hours median per course) compared to Successful Learners (7.6 hours median per course). There was not a substantial amount of course forum time by either group.

The total time spent in a learning session was analysed across courses completed by the pilot participants across 14 courses in the MPP Data Science curriculum across the dimensions of Highly Successful vs Successful Learners, age, and gender. Programming in Python stands out as having much more time spent in all categories for Highly Successful Learners as compared to Successful Learners. Forum posts indicated that this course was exceptionally challenging to complete.

When looking at the time spent per day in the learning session, the Highly Successful Learners tended to be more active early in the pilot program [9]. Successful Learners showed daily activity that was consistent throughout the pilot program. The hours spent in learning session per day ranged from under 30 minutes to over 10 hours. This broad distribution demonstrates the wide range of learning styles used in the course curriculum [10].

5 Video Length and Time on Task

We have augmented the Course DNA with time on task data for video and length of video. The result shows a high correlation between length of the video and percentage of video watched, as shown in Figure 3 below. The correlation amplifies some of the basic published research around video length and engagement [11], with a much higher percentage of videos viewed when there is a lower median length; We see this effect quite dramatically for the entire MPP curriculum.

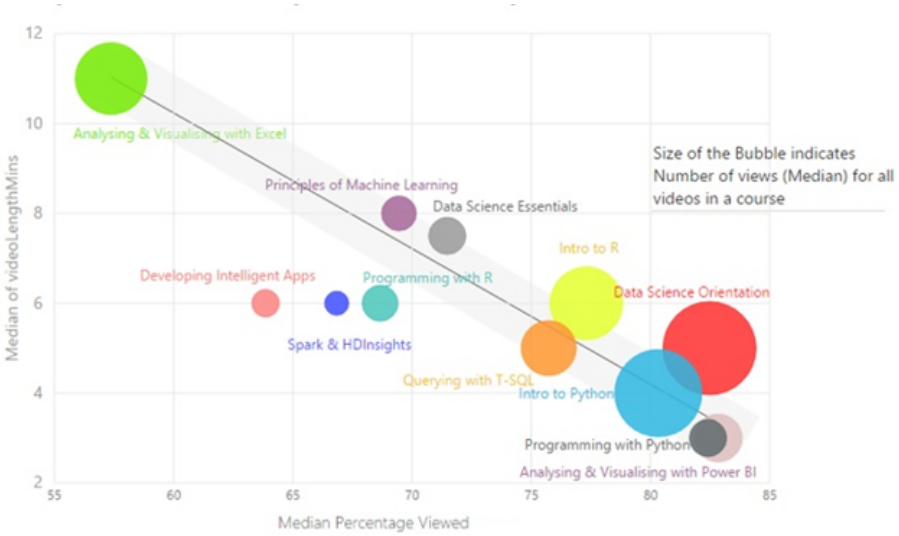


Fig. 3. Graph showing the median percentage of video viewed by the median video length per course in the MPP curriculum with a trend line showing a negative correlation.

6 Learning Strategies

Several learning strategies stood out between Highly Successful Learners and Successful Learners as learners progressed through the Data Science curriculum. The courses are intended to become more challenging as the learner is expected to learn more to build on their previous knowledge as they progress through the curriculum. Highly Successful Learners displayed "grit" or persistence [12], measured by learners who have taken every problem they possibly can, and every attempt needed to get the best possible score. Both groups of learners attempt similar numbers of problems throughout the first half of the curriculum. As courses became more difficult, Highly Successful Learners attempted more problems to achieve a passing score in the course.

The Capstone Project based course was where the forums were utilized mostly by Highly Successful Learners, was one of the most rewarding for learners, and heavily clustered student attempts [13]. The notable exception was the Programming Python course. By analyzing the forum posts, we determined the Programming Python increased forum usage was due to lab challenges. By contrast, the Project forum content was used for student collaboration.

Fig. 4. Total number of Problem events fired by course in the MPP curriculum, by Highly Successful Learner and Successful Learner

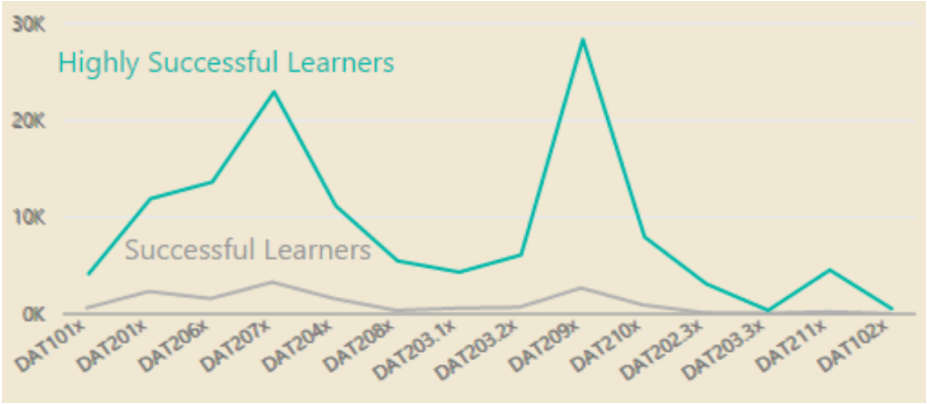
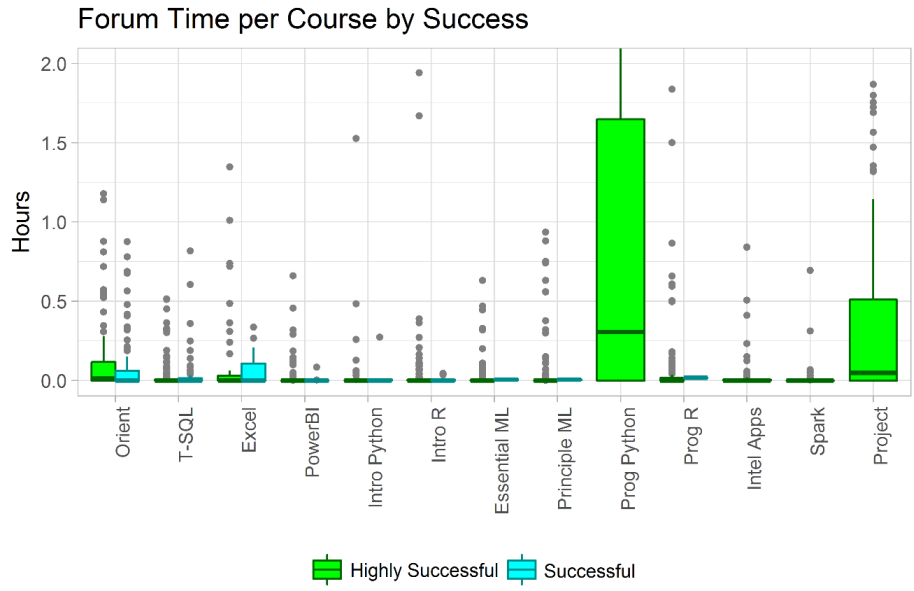


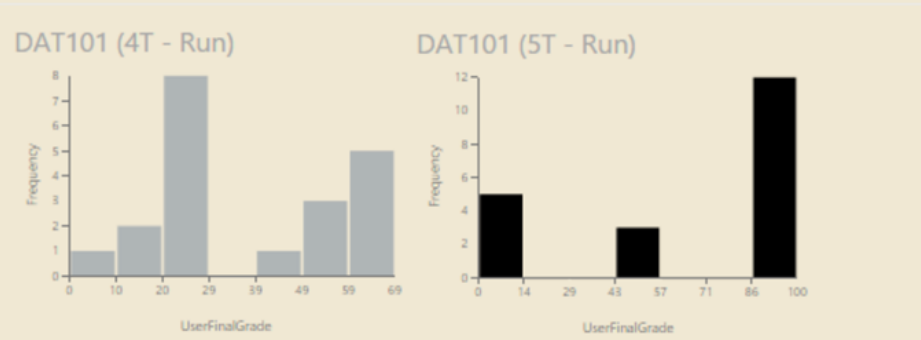
Fig. 5. Box plot distribution of time spent in course forums by course for Highly Successful and Successful leaners.



Another way to measure "grit" or persistence is whether a learner decides to try a course again after they have failed it the first time. The courses in the MPP Data Science

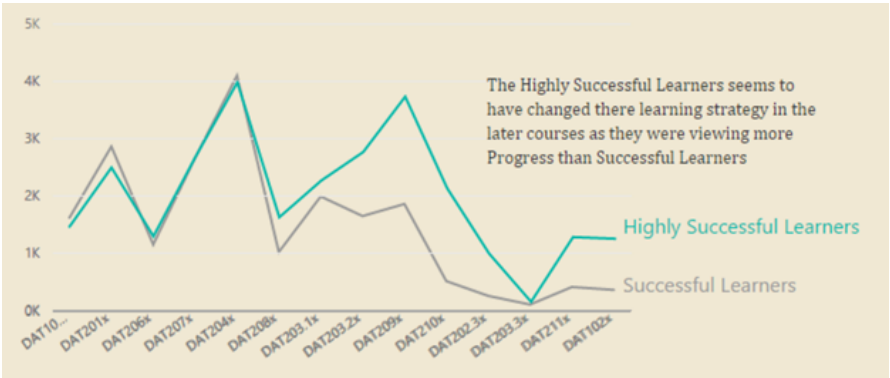
curriculum were offered multiple times during the pilot program, so learners had the opportunity to take a course again if they initially failed. In a comparison of one course, Data Science Orientation, where we see learners who earned a grade under 70% during the first run of the course either earned a passing grade during the next run of the course or they abandoned the program.

Fig. 6. Grade distribution for second run (5T) of the Orientation course for pilot participants who initially earned a failing grade < 70% for the first run (4T) of the Orientation course



Learners can view the score and progress through the course to see if they have earned a passing grade. They can also view the remaining problems/assignments to complete [14]. Highly successful learners are seen to be looking more to the progress page especially as courses become more difficult during the second half of the curriculum.

Fig. 7. Number of views on the course grade page per course for Highly Successful and Successful learners



7 Conclusion

Our study of an employer-created professional program offered online in a bounded time frame offers a starting point into understanding the learning behavior of professional learners. Highly Successful learners demonstrate "Grit" in multiple ways: They both attempt more assessments and 15% of them explore and complete more than the necessary courses to complete the program. They appear to both minimize the time on task by watching fewer minutes of video, as well as maximizing the number of assessments they attempt. They manage their session times with a heavier upfront commitment, and maintain that commitment consistently through the program. Highly Successful learners appear motivated to explore and learn as much as possible of the learning pathways and check their progress regularly. The capstone project based course was where the forums were utilized the most heavily, and post-completion interviews revealed that this course was one of the most rewarding for learners.

There are significant learning behaviors and analytics yet to be explored. The DNA of the courses appears to matter in student comments, but we have not yet quantified performance by course DNA, nor have we completed the qualitative survey analysis [15] or the predictive behavioral analytics [16] to understand better how we might help committed learners become highly successful learners. Finally, we plan to explore the learning pattern of those who fail more attempts at questions.

8 References

1. Lund, S., Manyika, J., Nyquist, S., Mendonca, L., Ramaswamy, S. (2013) Game changers: Five opportunities for US growth and renewal, , McKinsey Global Institute July 2013 Retrieved from http://www.mckinsey.com/global-themes/americas/us-game-changers
2. A. Fox, "Viewpoint: From MOOCs to SPOCs: How MOOCs Can Strengthen Academia," Communications of the ACM, vol. 56, Dec. 2013.
3. Bajzek, D., Brooks, J., Jerome, W., Lovett, M., Rinderle, J., Rule, G. & Thille, C. (2008). Assessment and Instruction: Two Sides of the Same Coin. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008 (pp. 560-565). Chesapeake, VA: AACE.
4. Ho, A. D., Chuang, I. Reich, J., Coleman, C.A., Whitehill, J., Northcutt, C.G., Williams, J.J., Hansen, J.D., Lopez, G., Petersen, R. (2015). HarvardX and MITx: Two Years of Open Online Courses Fall 2012-Summer 2014. (HarvardX Working Paper No. 10). Retrieved from Social Science Research Network: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2586847
5. Seaton, D. (2016) "Exploring Course Structure at HarvardX: A New Year's Resolution for MOOC Research". Retrieved from http://vpal.harvard.edu/blog/exploring-course-structure-harvardx-new-year%E2%80%99s-resolution-mooc-research
6. Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased Course Structure Improves Performance in Introductory Biology. CBE Life Sciences Education, 10(2), 175–186.
7. Reich, J. (2014). MOOC completion and retention in the context of student intent. Educause Review Online. Retrieved from http://www.educause.edu/ero/article/mooc-completion-and-retention-context-student-intent

8. Hansen, J.* & Reich, J. (2015). Demographics in MOOCs: Exploiting public datasets for estimates and comparisons. Proceedings of the Fifth Learning Analytics and Knowledge 2015 Conference, 59-63
9. D. T. Seaton, G. Kortemeyer, Y. Bergner, S. Rayyan, and D. E. Pritchard. eText Use in Blended Introductory Physics Courses: Interpreting Meaningful Interactions and the Effects of Course Structure. American Journal of Physics, 82(12):1186–1197, 2014.
10. J. Champaign, K. Colvin, A. Liu, C. Fredericks, D. Seaton, and D. E. Pritchard. Correlating skill and improvement in 2 MOOCs with a student's time on tasks. Proceedings of the First ACM Conference on Learning at Scale, Atlanta, GA, March 4-5 2014
11. P. Guo, J. Kim, and R. Rubin. How video production affects student engagement: an empirical study of MOOC videos. Proceedings of the First ACM Conference on Learning at Scale, Atlanta, GA, March 4-5 2014
12. Dweck, C. S. (1986). "Motivational processes affecting learning". American Psychologist. 41 (10): 1040–1048
13. H. Yin and A. Fox, "Clustering Student Programming Assignments to Multiply Instructor Leverage," in 2nd ACM Conference on Learning at Scale, 2015.
14. D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros, and D. E. Pritchard. Who Does What in a Massive Open Online Course? Communications of the ACM, 57(4):58, 2014.
15. Breslow, L. B., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. Research & Practice in Assessment, 8 13-25.
16. C. A. Coleman, D. T. Seaton, and I. Chuang. Probabilistic Use Cases: Discovering Behavioral Patterns for Predicting Certification. Proceedings of the Second ACM Conference on Learning at Scale, Vancouver, BC, March 14-18 2015