Procrastination in Online Exams: What Data Analytics Can Tell Us?

Yair Levy
Graduate School of Computer and Information Sciences
Nova Southeastern University, USA
levyy@nova.edu

Michelle M. Ramim
Huizenga School of Business and Entrepreneurship
Nova Southeastern University, USA
ramim@nova.edu

Abstract

Procrastination is an inevitable part of daily life, especially when it comes to activities that are bounded by deadlines. It has implications on performance and is known to be linked to poor personal time management. Although research related to procrastination in general behavior has been studied, assessing procrastination in the context of online learning activities is scarce. This study was set out as an exploratory investigation using advanced data analytics techniques about online exams. The dataset used for this study included 1,629 online exam records over a period of five terms in an academic institution in the southeastern United States. The online exams were provided during a weeklong timeframe where students were asked to take it based on material that they studied the previous week. The task performance time and task performance window were fixed on all records extracted. Results of this study indicates that when it comes to measuring online exams, over half (58%) of the students tend to procrastinate, while the rest (42%) do stage their work to avoid procrastination. However, those who procrastinated appear to perform significantly lower than those who stage their work. Clear trends were also observed based on whether the students work in the morning or the evening, their academic level, and gender.

Keywords: data analytics, procrastination, online exams procrastination, online exams data analytics, data-mining techniques in online learning, business intelligence (BI) in Web-mining.

Introduction

Procrastination is a prevalent phenomenon in modern life (Díaz-Morales, Ferrari, & Cohen, 2008; Steel, 2007). According to Gafni and Geri (2010), procrastination is defined as "the tendency to postpone an activity under one's control to the last possible minute, or even not to perform it at all" (p. 115). Various studies classified a segment of 15%-29% of the adult population as chronic procrastinators (Ferrari, 2010; Harriott & Ferrari, 1996; McCown, Johnson, & Petzel, 1989; Pychyl, 2010; Sigall, Kruglanski, & Fyock, 2000). Furthermore, according to Ferrari (2010), procrastination is prevalent at almost the same levels amongst western societies. Ancient societies viewed procrastination in positive terms, believing that it helped to avoid unnecessary work, and reduce impulsive behaviors. While there is acceptable pauses and delays in performing online tasks (Kalman, 2008; Kalman & Rafaeli, in press), studies associate procrastination found it related to personality characteristics, emotional disposition, and performance outcome (Ackerman & Gross, 2005; Van Eerde, 2003). However, procrastination may result from time pressure, rather than specific personality related characteristics (Freedman & Edwards, 1988). Moreover, Díaz-Morales et al. (2008) noted that time is the essential component of procrastination investigations. Additionally Beaudoin, Kurtz, and Eden (2009) indicated that "time management issues emerged as far and away the most dominant issue for these [online] learners" (p. 284). While better time management appears to reduce procrastination somewhat, increased volume of tasks among individuals in modern society still remains an unsolved challenge.

Previous research on procrastination provides mixed results on gender differences (Díaz-Morales et al., 2008; Gafni & Geri, 2010; Sarid & Peled, 2010). Procrastination may be task dependent and additional research to uncover such differences is warranted. Specifically, Díaz-Morales et al. (2008) noted that "future researchers should consider differentiating profiles of procrastinators, because distinctive profiles may lead to better diagnoses, treatments, and research on procrastination" (p. 238).

Nowadays, information systems are used to facilitate or aid in most daily tasks (Gefen, Ragowsky, Licker, & Stern, 2011). Such systems produce huge data sets of all task logs (Geri & Geri, 2011; Leventhal, 2010). Aside from the archival and retrieval problems, it's intriguing to understand what can be learned from such data sets? For example, companies are realizing the central importance of their customers' data that leads to increase interest in customer relationship management (CRM) systems. Moreover, organizations are faced with the need to make quick decisions about complex situation that have important consequences towards their success (Grummon, 2009). Few examples in the context of universities include enrollment, schedules, learning strategies, and retention that may be difficult to analyze, likely to be misinterpreted, and often exhibit some level of uncertainty. Consequently, universities seek to use approaches based on data analytics to make quick decisions. Data analytics is an emerging technique that dives into a data set without prior set of hypotheses, while letting the data derive meaningful trends or intriguing findings that were not previously seen or empirically validated (Leventhal, 2010). Data analytics in the context of businesses is known as Business Intelligence (BI), or Business Analytics (BA) (Turban, Sharda, Delen, & King, 2011). It studies the accumulation of raw data captured from various sources (i.e. discussion boards, emails, exam logs, chat logs in the context of online learning) to identify patterns, and relationships (Bose, 2009). According to Shmueli and Koppius (2011) exploratory data analytics "is used in a freeform fashion to support capturing relationships that are perhaps unknown or at least less formally formulated. This type of exploration is called exploratory visualization, as opposed to the more restricted and theory-driven confirmatory visualization" (p. 564). Thus, data analytics is distinguished from plain statistical analysis in that it does not initiate from a theoretical foundation or seek to identify a significant level to address hypotheses, rather data analytics uses data mining techniques to "find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner" (Leventhal, 2010, p. 138).

In the context of universities, data analytics can provide decision makers with various descriptive models, forecasting, predictive models, and simulations about students' behavior, level of academic performance pertaining to exams, retention, and more (Bose, 2009; Burke, 2009; Leventhal, 2010). The central aim of this study was to uncover any trends using data analytics techniques of procrastination in online exams. Specifically, this study examined a data set related to procrastination in online exams. Moreover, this study attempt to address the current trends in terms of task completion time, scores, gender, and academic level as time progress during the submission window. The main contribution of this study to the body of knowledge includes empirical results about procrastination from a data set of task completed in the context of online exams using data analytics techniques. The results can help support actions necessary to improve students' score on online exams, learning strategies, academic performance, and overall learning experience.

Methodology

Majority of procrastination related studies are based on surveys where information was collected post-experience (Gafni & Geri, 2010). In this study, we have extracted a data set of 1,629 online exam records. The data set was compiled from a period of five terms in an academic institution in the southeastern United States. The unit of analysis for this study is the task completed (i.e. an online exam) where each record indicated an instance of online exam completed. On average, there were about 35 students in each course, taking six online exams during each term, in a total of 10 courses distributed over the five terms. The 10 courses were in the context of information systems for sophomore, junior, and senior academic levels. All courses were structured similarly with the same instructor and the six online exams where staged throughout the term to cover the fundamental concepts discussed in the course. Two main time related measures were extracted: task completion window and task completion time. The task completion window was a weeklong timeframe (Monday 12am to Sunday 12pm), where the content tested was based on course content that was studied in the previous week. Procrastination was measured based on the proximity to due time and was measured in hours before the due time. The task completion time is the time that it took each student to complete the online exam and was measured in minutes. The time allocated for each online exam was fixed at 30 minutes, but students were informed that they can slightly exceed that time if they wish to do so. Additionally, the number of questions was the same for all exams.

According to Leventhal (2010), the process of data analytics begins with data preparation entailing data aggregation, organization of the data into a single file containing the records extracted and variables. We have followed the guidance provided by Leventhal (2010) and the data extracted was first organized into a single data set, then it was reviewed for any abnormalities and anomalies. There are instances where students log into an exam and due to technical issues, either lost connectivity or were not able to save their answers to the server. Such cases were removed prior to analysis. As indicated by Bose (2009), data analytics "is not a technology in and of itself, but rather, a group of tools that are used in combination with one another to gain information, analyze that information" (p. 156). Accordingly, our data set was inputted into IBM's SPSS 19 for the purpose of knowledge discovery and visualization using the SPSS syntax, while some meta-data from SPSS was inputted to Excel to augment the SPSS visualization process. Descriptive analysis was conducted on all variables, followed by specific data analytics slicing and visualization graphs. Once trends were observed in the data, Mann-Whitney U tests were conducted to test for statistical significance of the trends observed.

Results

Out of the 1,629 online exam records extracted, 56% (913) were completed by females and 44% (716) were completed by males. Interestingly enough, these courses had on average of slightly more male (52%) enrollment than females (48%), which suggests that there is a group of males (~8%) who don't complete all the required tasks throughout the term although it's part of their final grade. Table 1 provides summary of gender and academic level summaries.

Table 1: Descriptive Statistics and Demographics of Learners (n=1,629)

Item	Frequency	Percentage (%)
Gender		
Female	913	56.0%
Male	716	44.0%
Academic Level		
Sophomore	443	27.2%
Junior	938	57.6%
Senior	248	15.2%

Visualization of the data is a significant part of data analytics (Bose, 2009). However, given the space limitation, and in the effort of provide brief results, Figure 1 provides the summary results of the data analytics performed, aggregated into a single figure. The left side of Figure 1 provides a histogram of the procrastination time (in hrs) prior to the due time that is indicated by the red line. A close review of the data indicates clear distributions of the weekly days, while several cases were observed after the due time indicating some exceptions provided for medical or other special exceptions. The right side of Figure 1 provides a histogram of the average scores (out of 100) per day of the week, marked by the number of tasks completed per day of the week.

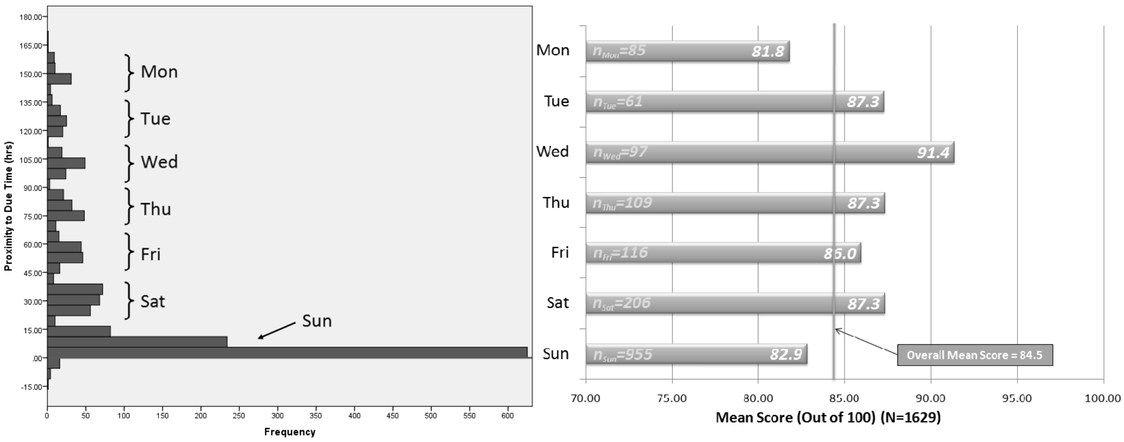
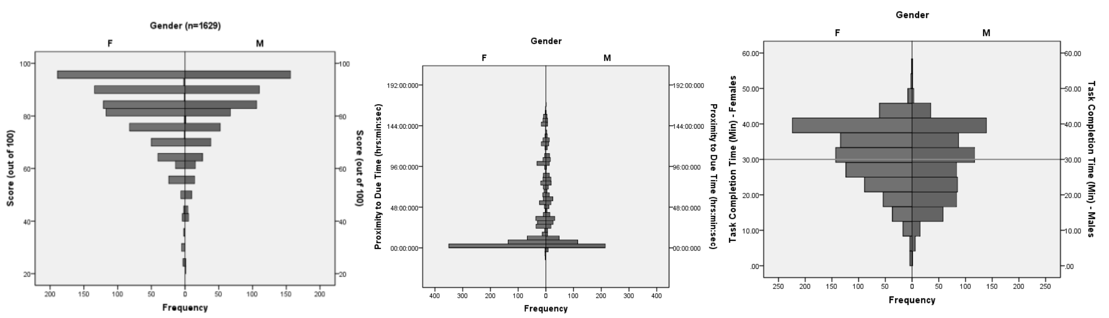


Figure 1: Descriptive Histograms of Procrastination (hrs) and Scores (n=1,629)

Figure 2 shows the distribution of scores, procrastination (measured in hours before the due time), and task completion time based on gender. Clear significant ($p<0.01$, using non-parametric Mann-Whitney U test (Mertler & Vannatta, 2010)) differences were observed. An interesting trend appears to emerge from the task completion time where it appears that males take significantly less time to complete the test prior to the allocated 30 minutes, where females tend to take longer to complete the task even going over the allocated 30 minutes. We suspect that given the fact that students were told they will not be penalized for going over the time, females appears to take advantage of it, while males appears to wish to complete the task quickly.



2a. Scores per Gender | 2b. Procrastination per Gender | 2c. Task Completion Time per Gender

Figure 2: Distribution of Scores, Procrastination, and Task Completion Time based on Gender (n=1629)

Due to the inconsistencies in the Monday records resulting from medical exceptions, along with the fact that weekday task completion frequencies are relatively smaller compare to those on Sunday, we decided to concentrate our further analysis on Tuesday through Sunday. For initial analytics trend visualization, Figure 3 provides another aggregated data analytics summary of exam scores, task completion time, gender, academic level, and aggregated procrastination (i.e. Sunday vs. Tuesday-Saturday). Results indicate some observed differences about scores, academic level, gender distribution, and task completion time. Specifically, scores for procrastinated online exams (i.e. those completed on Sunday) were found to be significantly ($p<0.001$, using non-parametric Mann-Whitney U test) lower than those completed during Tuesday to Saturday. Moreover, a significant gender difference ($p<0.01$) was observed where percentage-wise, 11% more females appeared to delay their task completion. Academic level also appears to provide some interesting trends, where percentage-wise, there was among difference on the seniors–between Tuesday-Saturday and Sunday. However, 12% more online exams were completed by sophomores (i.e. younger students academic-wise) when it comes to Sunday compared to Tuesday-Saturday. The shift in such numbers appears to come on the balance of juniors indicating that as students mature in the university, they tend to procrastinate less. In terms of task completion time, online exams completed during Sunday tend to take longer time compared to those completed during Tuesday-Saturday time frame indicates that procrastinating students might engage in scavenger hunt for answers to the exam questions, rather than demonstrating true knowledge of the material studied. Moreover, there were no observed differences of the scores based on academic level, course, or term taken.

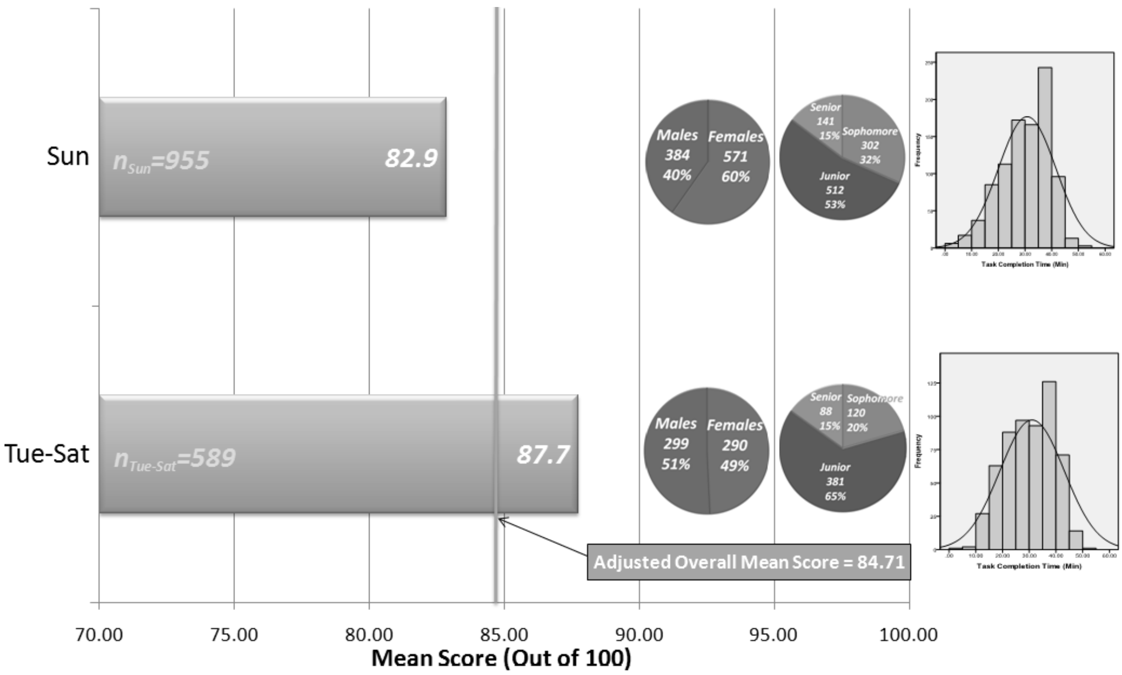


Figure 3: Data Analytics Plots of Scores, Procrastination, Task Completion Time, Gender, and Academic Level of Sunday vs. Tuesday-Saturday

Given that procrastination was measured in hours before the due time, another observed trend emerged from the data was the comparison between the morning hours (Ante-Meridiem (AM)) and evening hours (Post-Meridiem (PM)). According to Díaz-Morales et al. (2008), morningness–eveningness refers to "an individual's preference for specific times during the day" (p. 229). Figure 4 provides the trends in the data extracted from Tuesday to Sunday (i.e. 144 hrs to 0 hrs) regarding scores and procrastination based on morningness-eveningness, while the size of the bubbles represents the number of tasks completed. Interesting trend emerges that, aside from Friday, majority of the weekdays demonstrated clear pattern where online exams completed during the AM provided significantly ($p<.005$) higher scores then those completed during the PM hours. Moreover, it was observed that Friday appears to be a puzzling result and additional dichotomization of the data is needed to explore plausible explanation for such intriguing results.

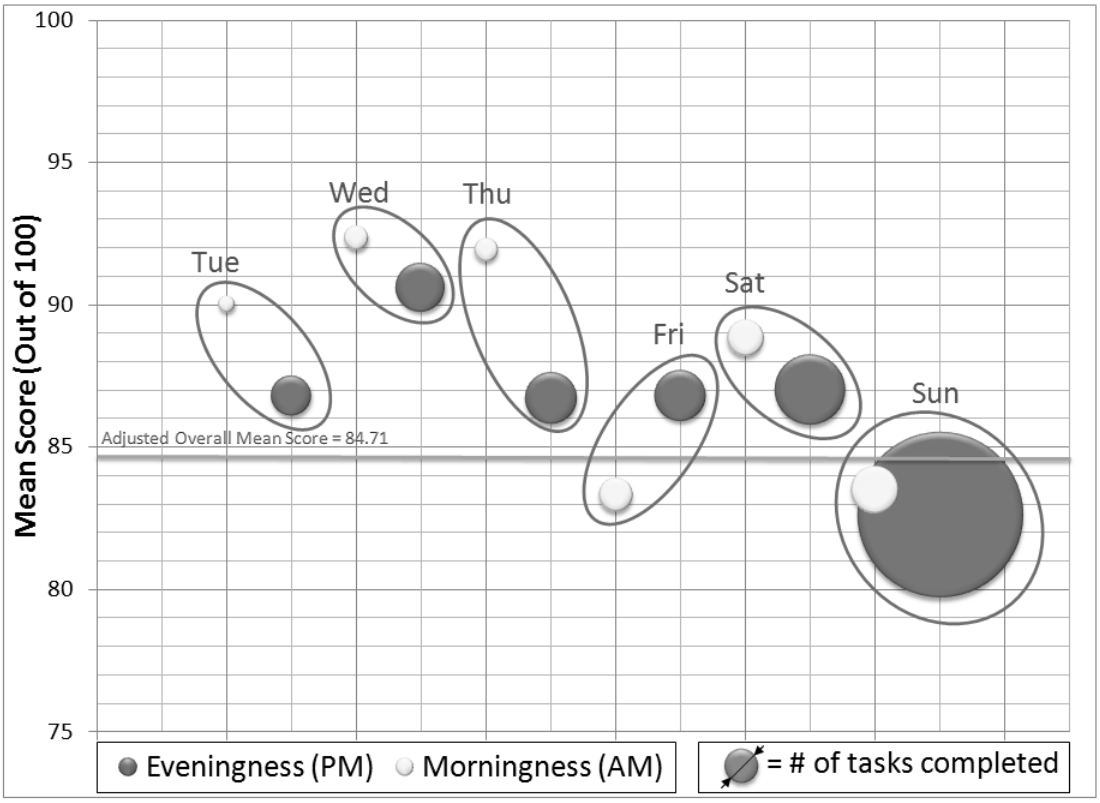


Figure 4: Distribution of Scores and Procrastination (Day) based on Morningness-Eveningness

Given the gender indicator that existed in the dataset, a further analytics exploration took place to separate the morningness-eveningness also based on gender. Results of such analytics are presented in Figure 5. Our findings indicate that aside from Wednesdays, males tend to outperform females during the AM. It is also evident that during Tuesdays-Thursday females tend to outperform males during the PM, while during Fridays-Sundays, males tend to outperform females slightly in the PM. We have also learned that Friday AM appears to be an out of the general female AM trend when it comes to performance. Further investigation of such female records indicated that the 18 records were mostly (73%) junior, majority (78%) scored below the mean, and almost all (90%) went above the 30 minutes allocated time. Further dichotomization of these records revealed that all but most of these females were mothers to school age kids and working full-time. We speculate that their out of trend results are due to the enormous demand that is placed on them balancing work during the weekdays and family obligations during the weekends.

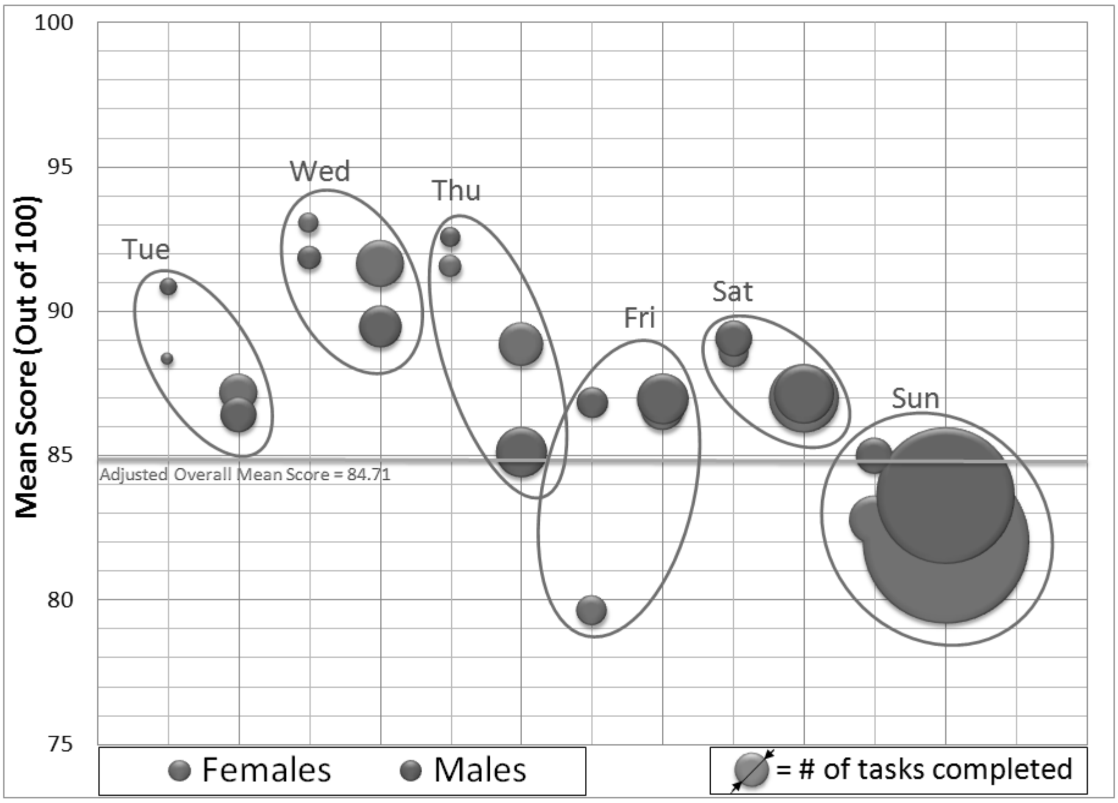


Figure 5: Distribution of Scores and Procrastination (Day) based on Morningness-Eveningness and Gender

Discussions and Conclusions

In this research, we have set out to use data analytics techniques to explore the problem of procrastination using a data set of historical records extracted from online learning systems. While to our knowledge there is much research done on the use of data analytics in the context of business organizations, also known as business intelligence (BI), in the past decade or more, there is very little research documented that incorporate such advanced data visualization and slicing techniques in the context of online learning systems. The extracted data set included a total of 1,629 historical records of online exams that were later analyzed for any observed trends. Given the main focus of the exploration was in the context of procrastination, our data analytics results actually nicely support the claim made by a participant in Beaudoin et al. (2009) study that said "the freedom to do work when you want is the best part of online learning, but also its biggest challenge" (p. 281). Our results indicated that although nearly 42% the students staged their workdays before the due time (Mon-Sat), over 58% procrastinated to the last day, or more precisely, about 40% procrastinated to the last 12 hours of the weeklong task completion window. We found that there although academic level among seniors was similar throughout the procrastination time, percentage-wise, significantly more young students procrastinate. Additionally, we have found that percentage-wise, more females procrastinated. We suspect that such results are due to the fact that enormous demand is placed on females, some are working mothers, which require them to balance work during the weekdays and family obligations during the weekends. Clearly such results are profound for researchers as it rises the opportunity for additional empirical research in the area of procrastination using data analytics to help validate these results in other institutions and cultures. Moreover, we find that our results should also provide an indication for practitioners and administrators of online programs to further understand the pressure that females, especially working mothers, are under in an attempt to help mitigate the unfavorable performances observed.

Our results also indicate that procrastination during the week doesn’t pay off, given that, by-in-large, there was an observed downwards trend during the weekdays where, in general trend, the longer the procrastination (i.e. closer to the due time), the lower the grade. This same trend also appears to hold true when it come to morningness-eveningness. In fact, morningness-eveningness can be considered as a daily procrastination, where students may delay completing their task throughout the early hours and eventually complete the task during the later part of the day. Also here, the data analytics trends show that it doesn’t pay off given that, by-in-large, those who completed their tasks later in the day have underperformed those who completed the task during the earlier part of the day. While any empirical research exhibits limitations (Ellis & Levy, 2009), this study wasn’t immune from it either. The central limitation of this study is with the use of data from a single institution and a single type of course. Future research using data analytics techniques is fruitful on data available in online learning systems in order to better understand current challenges and in an attempt to identify ‘at-risk’ groups in other institutions as well as other cultures.

References

Ackerman, D., & Gross, B. L. (2005). My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing Education, 27*(1), 5-13.

Beaudoin, M. F., Kurtz, G., & Eden, S. (2009). Experiences and opinions of e-learners: What works, what are the challenges, and what competencies ensure successful online learning. *Interdisciplinary Journal of E-Learning and Learning Objects, 5*, 275-289.

Bose, R. (2009). Advanced analytics: Opportunities and challenges. *Industrial Management & Data Systems, 109*(2), 155 - 172. doi: 10.1108/02635570910930073

Burke, E. (2009). Preserving the integrity of online testing. *Industrial and Organizational Psychology, 2*, 35–38.

Díaz-Morales, J. F., Ferrari, J. R., & Cohen, J. R. (2008). Indecision and avoidant procrastination: The role of morningness-eveningness and time perspective in chronic delay lifestyles. *The Journal of General Psychology, 135*(3), 228-228-240.

Ellis, T. J., & Levy, Y. (2009). Towards a guide for novice researchers on research methodology: Review and proposed methods. *Issues in Informing Science and Information Technology, 6*, 323-337.

Ferrari, J. R. (2010). *Still procrastinating? The no-regrets guide to getting it done.* Hoboken, NJ: John Wiley & Sons, Inc.

Freedman, J. L., & Edwards, D. R. (1988). Time pressure, task performance, and enjoyment. In J. E. McGrath (Ed.), *The social psychology of time* (pp. 11-33). Beverly Hills, CA: Sage Publications.

Gafni, R., & Geri, N. (2010). Time management: Procrastination tendency in individual and collaborative tasks. *Interdisciplinary Journal of Information, Knowledge, and Management, 5*, 115-125.

Gefen, D., Ragowsky, A., Licker, P., & Stern, M. (2011). The changing role of the cio in the world of outsourcing: Lessons learned from a cio roundtable. *Communications of the Association for Information Systems, 28*(1), 233-242.

Geri, N., & Geri, Y. (2011). The information age measurement paradox: Collecting too much data. *Informing Science Journal, 14*, 47 - 59.

Harriott, J. S., & Ferrari, J. R. (1996). Prevalence of procrastination among samples of adults. *Psychological Reports, 78*, 611-616.

Kalman, Y. M. (2008). *Silence in online education: The invisible component.* Paper presented at the Chais conference on instructional technologies research, Raanana: The Open University of Israel.

Kalman, Y. M., & Rafaeli, S. (in press). Online pauses and silence: Chronemic expectancy violations in written computer-mediated communication. *Communication Research*.

Leventhal, B. (2010). An introduction to data mining and other techniques for advanced analytics. *Journal of Direct, Data and Digital Marketing Practice, 12*(2), 137-153. doi: 10.1057/dddmp.2010.35

McCown, W., Johnson, J., & Petzel, T. (1989). Procrastination, a principle components analysis. *Personality and Individual Differences, 10*(2), 197-202.

Mertler, C. A., & Vannatta, R. A. (2010). *Advanced and multivariate statistical methods: Practical application and interpretation* (4th ed.). Los Angeles, CA: Pyrczak Publishing.

Pychyl, T. A. (2010). *The procrastinator's digest: A concise guide to solving the procrastination puzzle*. Bloomington, IN: Xlibris corporation.

Sarid, M., & Peled, Y. (2010). The effect of procrastination on multi-drafting in a web-based learning content management environment. *Interdisciplinary Journal of E-Learning and Learning Objects, 6*, 345-354.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly, 35*(3), 553-572.

Sigall, H., Kruglanski, A., & Fyock, J. (2000). Wishful thinking and procrastination. *Journal of Social Behavior and Personality, 15*, 283-296.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self regularity failure. *Psychological Bulletin, 133*(1), 65-94.

Turban, E., Sharda, R., Delen, D., & King, D. (2011). *Business intelligence* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Van Eerde, W. (2003). Procrastination at work and time management training. *Journal of Psychology, 137*(5), 421-434.