Incorporating Experiential Learning into Big Data Analytic Classes

Full Paper

Samuel H. Goh
Northern Kentucky University
samuel.goh@nku.edu

Xiaoni Zhang
Northern Kentucky University
zhangx@nku.edu

Abstract

Many data analytics classes utilize canned data, already cleansed, and with simplified scenarios for students to perform analysis. To better train students to tackle real-world analytics problems, using big data from a real company builds on prior research on the effectiveness of experiential learning. In this study, we introduced experiential learning into a data analytics class by working on an analytics project with a large Fortune 500 company. Our live case allowed students the opportunity to see firsthand the complex nature of analytics being more than just number crunching. Based on a student reflective assignment, we found that student learning outcomes and motivation were greatly enhanced. However, there were numerous challenges presented through the course of the study, including unclear expectations, ambiguity of the research project, and data quality issues. Our suggestions are that aligning expectations, effective communications, and managing and encouraging failure can be drivers of success.

Keywords (Required)

Big data, analytics, experiential learning, class project

Introduction

International Data Corporation (IDC) reports that the big data technology and services market will grow at a 27% compound annual growth rate to $32.4 billion through 2017 (IDC, 2013). In addition, McKinsey Global Institute forecasts that by 2020, the wider adoption of big-data analytics could increase the annual GDP in retailing and manufacturing by up to $325 billion and save as much as $285 billion in the cost of health care and government services (www.mckinsey.com, 2013). The potential for this area might however face a key bottleneck. ComputerWorld reports business intelligence/analytics is one of the 10 hottest IT skills for 2015 (ComputerWorld, 2015). Yet, supply has yet to meet demand as a survey by IDC notes that one of their big data challenges is the lack of sufficiently skilled big data and analytics IT staff (Press, 2013).

Given the demand in the market, more and more universities have added data analytics as a concentration/specialization to their curriculum. As educators, our job is to train students with the right skills that the market demands. The typical way of teaching data analysis is to use lectures and then to let students complete problem solving questions with hypothetical data. Usually the hypothetical data is small, cleansed, and ready for analysis. A common challenge for faculty is to obtain big data sample sets for teaching purposes. In reality companies need people to analyze their big data. In order for students to gain skills in analyzing big data, they must use big data for analysis. Collaboration between a university and a company is an ideal way to explore big data in an analytics class. Instructors can elaborate concepts and let students explore data. Such collaboration falls into the teaching pedagogy of experiential learning.

Experiential learning is a process by which the learner creates meaning from direct experience (Dewey, 1938). Kolb defines experiential learning as a "process whereby knowledge is created through the transformation of experience" (1984, p.38). Experiential learning takes place when a person is involved in an activity, looks back and evaluates it, determines what was useful or important to remember, and uses this information to perform another activity (Kolb, 1984).

Prior research provides evidence on the effectiveness of experiential learning on learning in higher education. The documented evidence includes improving grades (Reitmeier, 2000), improving attitudes
towards challenging material (Pugsley & Clayton, 2003), better motivation and attitudes towards learning as students were shown how the knowledge can be applied to their lives (Briers, 2005).

Statistics, business intelligence, data analytics and data mining classes are perceived to be hard classes by students. However, these skills are in demand and make students more marketable. There are many factors that influence the pedagogical approach in teaching analytics and big data. Data analytics class may be more interested to students if the learning process is self-directed. Students solve real-world problems and this learning process allows them to interact with industry and have the opportunities to impress potential employers.

Given the success of experiential learning used in marketing (Burgess, 2012), international business (Chavan, 2011), business communications (Goby & Lewis, 2000) Capstone MIS course (Chilton, 2012), and management (Gitsham, 2012) in higher education, we believe experiential learning is equally effective in delivering the learning outcomes of big data analysis. Experiential learning is especially helpful in understanding the complexity of data, practical business problems and decision making with the analysis results. With real data, students go through the complete process of data analysis from formulating research questions, data cleaning and integration, model formation, interpretation, and conveyance of results.

In this study, we introduced experiential learning into the classroom by partnering with one of the largest Fortune 500 companies on an analytics project. By using real data, students have the opportunities to see the complexities in data, and they have an up-and-close view about reality. This study intends to achieve the following objectives: 1) investigate the effect of experiential learning in a data analytics class; 2) assess students’ attitudes on the experiential learning method; 3) identify practical issues when analyzing and big data.

**Literature Review**

**Kolb’s Experience Learning Cycle**

Borzak (1981) suggested experiential learning as an instructional model which allows the learners to engage in the direct experience, reflect, discuss, analyze and evaluate the experience. Kolb (1984) developed the experiential learning theory in 1984 and this theory states that learning is a cycle and consists of the four learning phrases: concrete experience, reflective observation, abstract conceptualization and active experimentation. Learning can occur at any step in the learning cycle. When the entire cycle is completed learning is thought to be most effective.

Mintzberg (2004) noted that the best way to learn is by reflecting and learning from one’s own experience. He redefines experiential learning as a self-directed learning process (2004) since individuals continuously gain knowledge through their own experiences. Typically experiential learning activities include internships, student teaching, cooperative education, consulting, service learning etc. Real-time experiences produce incomparable benefits to other teaching pedagogies. Carver (2008) identify building and improving skills related to critical thinking, problem solving, discussion, and decision making as major benefits of experiential learning.
Applications of Experiential Learning

Kolb’s learning theory provides a conceptual model and practical framework for designing, implementing and evaluating many courses. For example, in a business school experiential learning takes the form of team-building exercises, simulations, guest speakers and internships. These types of activities facilitate interactions among students, industry and train students some essential skills such as communication skills, interpersonal skills, problem solving skills etc. In the meantime, students need to understand concepts and reflect on what they are taught.

Baden and Parkes (2013) applied experience learning to an entrepreneurship and social responsibility class. They compared two approaches when teaching the concepts of sustainable development. One group of students used experiential learning by working with social enterprises and the other group of students were taught with the traditional case study method. Students in both groups were required to write a reflective essay about their learning of the materials. The result of the content analysis of the reflective essay showed that the opportunity to work with social entrepreneurs and/or “responsible” business professionals provided the business students with inspirational role models and positive social learning opportunities. Experiential learning is an effective way of integrating ethics, responsibility and sustainability into the curriculum (Baden, 2013).

In teaching a junior level retail promotion class, Burgess (2012) used the experiential learning approach and let students be retailers operating pop-up stores. The project consisted of designing, producing, operating, and analyzing the outcomes of a pop-up retail consignment store on campus. Through hands-on operation and management of the pop-up store, students gained better knowledge about reality and how businesses work in practice. They learned to better equip themselves for the job market. Both the instructor and students enjoyed this experiential learning project.

In summary, prior studies offer practical suggestions and tips in the successful execution of the experiential learning projects. For example, Birchfield (2010) recommend that experiential learning should employ the whole learning approach from goal setting to experimenting, observing, reviewing, and action planning. Certo and Peter (1976) suggest a comprehensive exercise should be designed to evaluate experiential learning components. Baldwin & Baldwin (2000) conclude clear and specific feedback is important for effective learning.
Methods

We used a live case project to design and implement the big data analytics class as the live case is regarded as one of the most highly effective experiential learning pedagogies (Farazmand et al., 2010; Gupta et al., 2010; Karns, 2005).

Course Design

To frame the steps involved with big data analysis, the course incorporated the concepts introduced from the Spreadsheet Analytic Value Chain (SAVC) by Grossman (2006). The SAVC embraces the idea that real-world problems are inherently “messy” and that successful analysts must be able to perform a series of activities including modeling, spreadsheet engineering, analysis, articulation of insight, and managerial communication. Each of the steps is equally important and interrelated. As a general framework, the SAVC provides a very interesting contrast to what typical students expect from an analytics class. Many students assume analytics courses are about learning analysis methods and techniques, rather than considering that real-life applications require much more than data crunching.

A stated goal was to give students a large breadth in analytical tools in order to be successful on the project with a variety of analytics tools and topics that focused on how to manipulate and analyze data sets. Sample topics covered included sensitivity analysis, what-if analysis, simulations, advanced Scenario Manager and Solver in Excel, multiple regression, and one- and two-way ANOVA in SPSS.

Results and Discussions

Reflective Essay

After completion of the project and project presentations to the company, a reflective essay was collected from 16 graduate students who took the data analytics class for Fall 2014. Students were asked to reflect on how their initial understanding of the SAVC had changed after the project and how their learning of business analytics concepts were enhanced from working on a hands-on project with a real-life company.

Our preliminary work here has been to organize the project and essays into a basic narrative, to identify themes, and to provide our interpretation of the practical and pedagogical implications of our study.

Study Narrative

A large Fortune 500 Company (named G&P for this paper) expressed their interest in working with a mid-west university in collaboration on a big data project. Several meetings between senior-level managers and faculty were held in summer and fall of 2014. These meetings discussed initial details about project scope and deliverables.

Our observations from these meetings suggest that companies are very protective of their data and that trust takes time to build. Also, a company that is willing to share its data has high expectations of the resulting analyses. Since the course had yet to begin and there were many variables (e.g. quality & number of students) unknown, this highlights the substantial risks involved with trying to run a live case study.

Big Data Project Initiation

The original design of the class was to spend approximately half the class learning the process of big data analysis and to introduce the project after a sufficient number of analytical tools were learned. Class activities were to include traditional lectures, hands-on assignments, labs, podcasts, and case studies. The course included a face-to-face cohort and an online-only cohort, split into 4 self-selected project teams of 4 students each.

What was highly unusual was that even the instructor was unaware initially of what direction(s) the project would take. A meeting with G&P was planned for 5 weeks into the course. Prior to this, students were given a brief overview of G&P and some general ideas of the deliverables of the project. For G&P, a final written report and presentation to senior-level managers was expected. For the class, deliverables
that followed the SAVC were expected, for example research questions, hypotheses, data cleaning, data analysis, results, a group paper, and an individual reflection paper.

Also, a preliminary data dictionary and several tables of sample data were provided by G&P. Students were encouraged to examine these documents and to generate a list of questions that they might have for the first meeting. In retrospect, more emphasis should have been given to the problem understanding phase in the beginning.

**Breaking Assumptions**

The first meeting was difficult. G&P representatives provided a brief overview of the data and of what they were looking for. However, there were still clear disconnects. The data dictionary provided was geared towards an internal G&P audience. Students had difficulty understanding variable names, how the data was organized in the tables, how data was rolled up, and even what basic units of measurement were. Much time was spent on clarifying what the data was.

One student commented:

“I personally would have gotten more out of the time that G&P spent presenting their data to us if we had spent more time preparing for that briefing session. I was certainly not as prepared as I would have liked because I was still trying to grasp the documentation that G&P had provided despite having their raw data, powerpoint and word documents.”

“I [thought I was] was prepared to run tests on the data right out of the gate and wrap up the project in a week.”

Other students reflected:

“It is important to understand the product and the customers. The knowledge helps you in forming boundary conditions from business and the analysis. Understanding where the business concentrates and what their struggles are is important. The expectations and business strategy needs to be understood.”

“Looking back now, it’s easier to see that we should have spent more time initially as a group reviewing their questions and determining what we thought was important and doable.”

“This does require making intelligent assumptions about business context, considering data issues, defining key outputs and determining out how to compute outputs from inputs. These ideas are fairly straightforward, but I would say that I had unintelligent assumptions going into this phase.”

Our observations here are that initially, the company and the students’ expectations were very different. The students were expecting a very structured process: to have everything explained to them, clear research questions already formulated along with the variables of interest, and the analyses that were expected of them. This was in stark contrast to G&P’s expectations. Their expectations were much less structured where they expected the student groups to delve into the data, develop their own research questions, tests, and conclusions. From a teaching perspective, our observations suggest that aligning these perspectives earlier on would have helped alleviate student stress later on.

**Analysis Paralysis**

Following this meeting, students experienced a tremendous amount of pressure. There was a great deal of uncertainty as students began to realize this was not a canned class exercise and that much of the work would be self-directed. This was compounded by information overload. G&P provided 4 raw tables of data of substantial amounts. This data included multinational data points that pertained to marketing efforts across brands. The data included internal and external sources and were in a raw format needing to be cleaned and joined together. The data sets were very extensive with several tables of over 16,000 records each. Later as the class progressed, we requested additional datasets and ended up with 6 large tables of data.

Two students commented:
“For our project, I think we started out struggling with which variables to use and trying to formulate a question that would provide any insight into what G&P wanted answered. It was hard to see because of the large amounts of data and all the missing data. We were just worried about having enough data to show results.”

“Another key take away [is] how scope creep can truly paralyze the teams’ ability to move through a project. For instance when we came up with our research question on the G&P data we keep finding more and more things that maybe relevant to the project, with each new way of analyzing the data we were actually doing ourselves a disservice because we were not actually refining the data to see if the question was being proven or nullified.”

Data quality issues were the big concern during this phase. Students had to clean the data and decide how to deal with issues such as null entries, data inconsistencies and data redundancy. Further, the data had to be converted into a database with proper primary/foreign keys selected in order to join the tables together. Several students commented:

“I was a bit surprised first that big industry giants like G&P can have really disorganized and incomplete data. Then we quickly understood that in real-life the data is truly messy, and it is up to analysts to sort and analyze it in order to come up with strategic insights.”

“As with the G&P project data, it came very messy and ambiguous, it took our team weeks to clean out the data and it requires us to make a lot of intelligent assumptions about the business context, considering the fiscal year of the that and how inconsistence the data is to the fiscal year and months within each year.”

“To be completely honest, I was extremely frustrated with the whole G&P project initially. I remember thinking to myself it would have been nice if we could have canned data, or if we could have picked something that was familiar to us. I remember asking this question to myself: “Wouldn’t it be easier to learn this process if we also did not have to learn the data?”

After the arduous cleaning process, many of the groups progressed quickly into running analyses. However, we observed that while students loved jumping directly into the data analysis in order to show quick results, it became challenging to interpret results, as they had no a priori way to establish cause and effect. This was in direct contradiction to what the instructor had suggested in beginning to develop their own research questions and hypotheses (the SVAC suggests that modeling comes before analysis). These groups did learn some from the experience however:

“I could remember that we dropped or adjusted our research question because of not initially performing one or more of the task stated earlier. As easy as I thought it would be, it took us about two weeks to get this properly set. In some occasions we had started the merging of the data when it was discovered that some proposition in the hypothesis could not be met hence, we altered it again. Even further till analysis stage, we still altered our hypotheses to fit the context of the data we had.”

“Throughout the modeling phase of this, I learned a lot; first, the problem is not always clear. Also, it will not be apparent what issues really are there. It comes to interpreting at the end, but we found issues that we didn’t think should have existed. I learned that expectations, even though they seem logical may not be met.”

“Theory does not always translate to practice as seamlessly as you would like. This was realized in the Modeling of the G&P project. The difficult part of the Modeling for G&P was truly understanding the problem that was being presented and how to frame the data. The data had many inputs from different types of data sources and how those speak to each other in the Marketing Mix was something you couldn’t just do by looking at the raw data. To create a testable model our team had to understand how those pieces speak to the other and how our outputs could possibly produce valuable actionable results for G&P.”

**Breaking the Paralysis**

We observed several ways that the student groups got back on track. One group utilized feedback from the instructor and reflected:
“We did not struggle to divide up the work, we struggled to identify what we were going to solve, or what question we asking to derive hypothesis and insight. After a couple meetings of mostly silence by our team, we just picked something, and began to think about what tests we could run to determine some insight about our questions. We failed a few times, and then finally invited you [the instructor] to the group to help us to the next level. After a couple calls with you, we determined that we were never on the right track, and that we were always starting over. Again, the learning changed for us, as we all decided that we could start over 1000 times, and that is ok and normal.”

Other groups used the fact that it was a real-life company to motivate them to continue:

“In developing research questions and hypotheses, having a live company created more of a sense of urgency than using canned data. Rather than saying “OK, let’s just pick this”, it was more of a need to help them find some answers. What do they want to get out of this? Wendy was a live person who obviously lives this question every day. It wasn’t just a scenario out of a book.”

“Again, doing this for G&P brought a sense of purpose to this project. It wasn’t being done simply to learn and to get a grade.”

“What we did was something that even a huge business such as G&P need us to do. They wanted our data and our results to see what they can dig up on their end.”

Many of the groups used other team members as a resource and also as a source of support when things got tough:

“Working with a group in which everyone is concerned about putting out a quality product is key for great team work. At times when we felt like we had hit a brick wall, we were able to come together as a group and give encouragement to pick ourselves back up and try again.”

“Furthermore, I want to point out that I was very glad to be in a group of people where we could rely on each other. Although we had some arguments, we never crossed a point of getting into conflict. We helped each other and divided the work according to the strong sides of each team member. But all of us wanted to learn some things that others didn’t know and support was always there. We did have to put many hours into the work and each member was very dedicated to complete the task given no matter how much sleep they had to lose. In sum, I was very encouraged to have a strong team dynamics in one of the most interesting and challenging classes of our MBI program.”

For the instructor, the reflection essays suggested that feedback, positive reinforcement, and being a good facilitator and coach were also motivators to break analysis paralysis:

“In addition, we had meetings with our instructor who was always very open and very insightful in helping us.”

“I liked how Dr. [xxx] asked us questions that might come from the G&P folks to prepare us for their questions. It put our team on the spot and got us thinking about how to better organize our presentation. I feel that our final product was much superior to where we started from because we sought our feedback and we incorporated that into our presentation – and because we were challenged with these questions on what our data yielded and how our research questions were important to the company.

**Learning From Doing**

Throughout the analysis portion, we observed students enacting experiential learning. Unlike a class exercise with a clearly defined outcome, students were experimenting, failing, and learning in reiterative cycles. Some students thought:

“As we worked through our research questions and hypotheses, we changed our models at least three times during the process, scrapped tests that we had run and analyzed, re-analyzed, ran different models in different formats, and then did it all again.”

“Several models that we ran had no impact or didn’t make sense. This forced us to go back and look at the original model to figure out what was going wrong. We did this several times.”
“Again, the learning changed for us, as we all decided that we could start over 1000 times, and that is ok and normal.”

“In most class application you are given the question to answer and the data to use to obtain that answer, in work applications you are usually given the question but need to identify what data to use and figure out where to find it; in this case we were given many possible questions and many sources of data which created extra complexity.”

The reiterative process necessitated a significant amount of meetings with the instructor because of the necessity of feedback. During the project period, upwards of 20-30 hours a week of meeting time was recorded. We warn that it could be time prohibitive to do in larger classes but such engagement is absolutely essential.

A major learning component about big data analysis came in the final presentation to G&P. After being buried in analyses, students were challenged to articulate their findings in a clear and concise manner for a business audience. Some students found the presentations a way to “close the loop” in terms of their understanding of the project.

“Our final presentation was a good way for us to communicate with G&P managers and show our results and insights. It was helpful to hear their perspectives on the same questions. We could assume some things from the beginning and carried our own agenda throughout project. Finding their points of views was crucial so that we could possible change our project to something that they would like to know more in depth.”

“I learned during the course of the project how important it is to know your subject matter inside and out to avoid being perceived as inauthentic or worse, ignorant. The message is only as impactful as the messenger. Overly complex, poorly conceived illustrations combined with overly verbose explanations don’t effectively lead the audience as witnesses to the conclusions you want them to both understand and embrace.”

“Communicating the findings proved to be more difficult than I anticipated. From working on this project, I appreciate the communication process and how complex it can be. As an analyst, it is important to understand the research questions from the client’s perspective and to be able to communicate the finding to them in a manner that they understand. You have to think about how they might see the “big picture” of the project and what outputs they might be looking for. It was so easy for us to get focused on the numbers that we had a hard time stating in lay language what the findings were.”

As the final question in the reflection essays, we asked students to reflect on if and how their learning of the SVAC and big data analysis had been enhanced. Overwhelmingly, the feedback was very positive. Some sample comments include:

“This project highlighted how these analytical concepts rely on each other in a real way and what it involves to undertake a project of this magnitude. G&P themselves had problems gathering the necessary data, much of which was held in separate systems, in separate departments and in separate company units. While students may be amazed that the uncoordinated workings of large corporations, this separation is a commonplace thing in any large multinational company. It also highlighted the challenges that face analysts working with data from multiple sources. Uncommon naming conventions and mismatched data sets are again common things in large companies. Students who have had the opportunity to work in these large companies understand this as a natural phenomenon. This project enhanced my understanding of analytics not just through the ability to work with live data, I am lucky enough to be able to work with data in my current job, but through the opportunity to create a model of inception through the analysis and to the final presentation, with the feel back along the way. It is rare, in my experience that the formulation of the question is given to the role of the analyst to see through to the end. In most cases management has already seen the result of a problem that needs addressing and then asks the analyst to uncover the cause and then to recommend actions. The ability to understand and work through the entire process was a wonderful and appreciated opportunity.”

“I would have derived descriptive statistics, and reported basic facts about the data, and never gotten to the point that I could provide insight or direction to the client. This process provided an opportunity to
understand what the data was telling us, and to help define what the results of the data mean to the question we were asking.”

“This class was certainly out of my comfort zone. It was challenging, frustrating, and time consuming. However, now that I have done some analytical work, if I drill down further beneath the layers, I see that this class was also interesting, informative, and insightful. I have learned many things that I can take with me in my career path that I would otherwise never have stumbled upon.”

Conclusion

This study used an experiential learning method in conjunction with a live case study in order to teach a graduate level data analytics class. This approach allowed students to better experience what real-world big data analysis was about by conducting a research project for a Fortune 500 company. Based upon student feedback, student learning outcomes and motivation were greatly improved over more traditional class assignments. However, we did identify some challenges that were experienced and offer several prescriptions.

First, student expectations about analytics have to be properly aligned. Qualitative feedback from employers suggest that while skills in data analytics tools are important, analytics is more than just number crunching. Students should be made aware of this important distinction. Many of our students were initially overwhelmed by the statistics concepts to the point where they could not grasp that analytics is a process. In our live case study, the SAVC proved to be a useful framework in which to help better structure and teach the process.

Second, communications of all types is essential. Internal communications within the groups helped build good cohesion, reduced uncertainty, and helped build group resilience. Feedback from the instructor was a useful mechanism to improve the quality project deliverables, provide guidance and mentoring, and reassurance. In particular, two-way communication between G&P and the class was also necessary in order to produce outcomes which met the client’s needs. A significant amount of time was spent in face-to-face and electronic communications with G&P throughout the span of the project. Educators seeking to practice experiential learning should be aware of the high time commitment and high risk nature of such a project as compared to using canned data.

Lastly, we observed that experiential learning involves failing a lot. Managing this process proved challenging. We found that students look to faculty to provide an example. Our prescriptions include remaining positive, providing encouragement, and stressing that even the best planned study can fail or produce unexpected results. In a more encouraging light, helping students reflect and learn from past iterations helped future iterations become better. Working with a real-world data also helped students better understand the complexities associated with business analytics.

REFERENCES


