**Coursera Student Education Analysis**

Surendar N. Reddy

DeVos Graduate School, Northwood University

MTH 650: Introduction to Data Analytics

Dr. Itauma Itauma

March 2nd, 2025

**The Impact of Course Difficulty and Certificate Type on Student Ratings in Online Education**

**Introduction:**

Online education is growing at a very fast pace and is becoming very important for learning in these modern times. How much students like a course are very important because it decides whether the students will join and stay in courses or not. Course ratings measure this liking and represent the quality of learning. This study looks at how much the level of a course influences (Advanced, Beginner, Intermediate, Mixed) and the type of certificate given (COURSE, PROFESSIONAL CERTIFICATE, SPECIALIZATION) have on these ratings. Knowing this correlation will help teachers in creating improved courses and make students happier. This study will give beneficial information to enhance online learning experiences for all parties involved.  
  
**Problem Statement:**

Online learning platforms struggle to achieve high student satisfaction, which can have a strong impact on course completion rates, enrollment, and platform reputation. This study looks to find how levels of course difficulty and types of certificates impact student ratings, thus guiding course design improvement and student satisfaction.  
  
**Dataset Description:**

The data set has course ratings collected via an online learning platform, categorized by course difficulty levels (Advanced, Beginner, Intermediate, Mixed) and certificate types (COURSE, PROFESSIONAL CERTIFICATE, SPECIALIZATION). They are rated on a scale of 1 to 5, with 5 being the highest. Variables employed are:  
  
**Independent Variables:**

Course Difficulty (Advanced, Beginner, Intermediate, Mixed), Certificate Type (COURSE, PROFESSIONAL CERTIFICATE, SPECIALIZATION)  
  
**Dependent Variable:** Course Rating

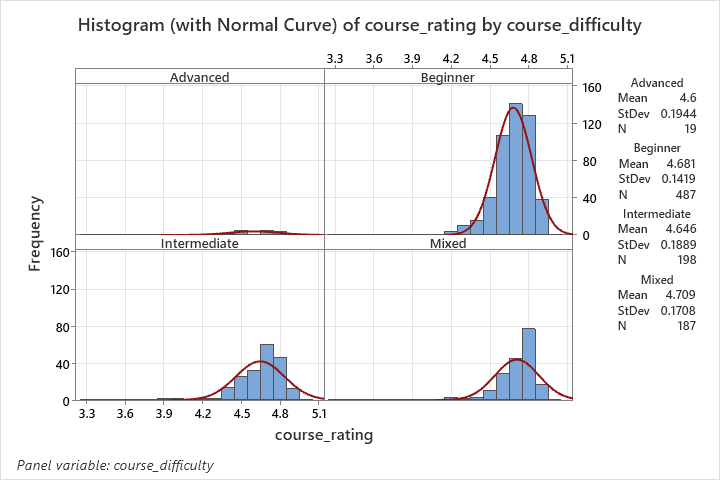
**Research Steps:**

**Step 1:** The Data was collected from web portal, including course ratings, difficulty levels, and certificate types.

**Step 2:** Data Cleaning and Preprocessing The data was checked for accuracy and completeness. Ratings were confirmed, and outliers or missing values were cleaned to ensure data quality.  
  
**Step 3:** Exploratory Data Analysis (EDA)

**Statistics**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **course\_difficulty** | **N\*** | **Mean** | **SE Mean** | **StDev** | **Minimum** | **Median** | **Maximum** |
| course\_rating | Advanced | 0 | 4.6 | 0.0445904 | 0.194365 | 4.2 | 4.7 | 4.9 |
|  | Beginner | 0 | 4.68090 | 0.0064322 | 0.141947 | 4.1 | 4.7 | 5 |
|  | Intermediate | 0 | 4.64646 | 0.0134273 | 0.188939 | 3.7 | 4.7 | 5 |
|  | Mixed | 0 | 4.70856 | 0.0124904 | 0.170803 | 3.3 | 4.8 | 4.9 |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **Table 1** |  |  |  |  |  |

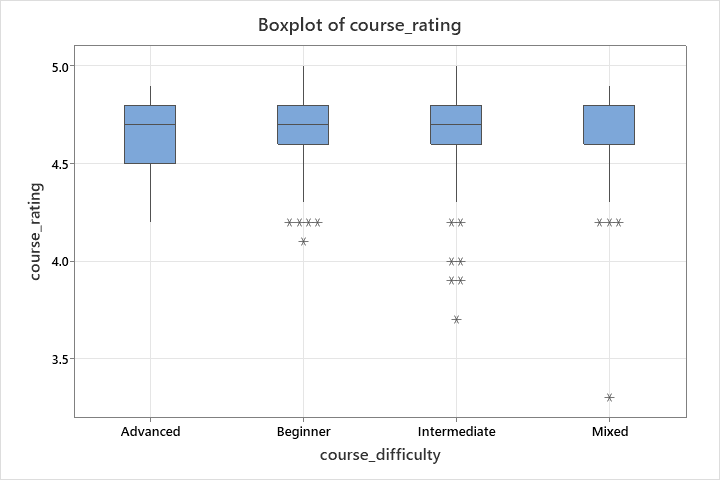


**Figure 1**

**Histogram of Course Ratings by Difficulty**

Histograms showed Beginner and Mixed courses had higher concentration around top ratings, while Advanced courses showed more variation in ratings.

**Interpretation:**  
  
 This data compares course ratings by level (Advanced, Beginner, Intermediate, Mixed).  
  
Highest average ratings belong to Beginner and Mixed courses (around 4.7). Intermediate courses slightly below (4.65). Advanced courses have an average of 4.6 too, but most spread range of ratings.  
  
The histograms show how spread out the ratings are. Mixed and Beginner courses are more spread out around the high mean. Advanced courses are more widespread, with very high and lower ratings.  
  
Easy courses generally get higher and more consistent ratings, while advanced courses are more split in opinion.



**Figure 2**

**Boxplot of Course Ratings by Difficulty**

Boxplots showed similar median ratings (~4.75) across difficulties, with Beginner and Mixed courses having a broader range, suggesting more variation.

**Interpretation**

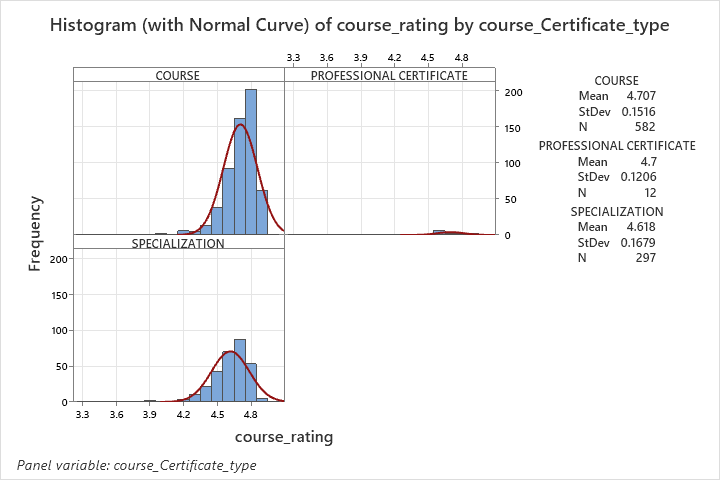
This boxplot displays the range of course ratings across four difficulty levels: Advanced, Beginner, Intermediate, and Mixed.  
  
There is a median rating of roughly 4.75 for each difficulty level, showing generally good reviews. There is varying dispersion in the range of ratings, however. The Beginner and Mixed courses have a broader range with lower ratings (outliers). The Advanced courses have the most skewed distribution, which means stable ratings.  
  
Essentially, while all course types are well-rated overall, Beginner and Mixed courses have greater variability in ratings, while Advanced courses are more uniformly high.

**Step 5:** Descriptive Statistics by Course Difficulty. The analysis revealed that Beginner and Mixed courses had slightly higher average ratings (4.68 and 4.71) compared to Advanced (4.6) and Intermediate (4.65) courses.

**Statistics**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **course\_Certificate\_type** | **N\*** | **Mean** | | **SE Mean** | **StDev** | **Minimum** | **Median** |
| course\_rating | COURSE | 0 | 4.70704 | | 0.0062860 | 0.151647 | 3.3 | 4.7 |
|  | PROFESSIONAL CERTIFICATE | 0 | 4.7 | | 0.0348155 | 0.120605 | 4.6 | 4.65 |
|  | SPECIALIZATION | 0 | 4.61818 | | 0.0097411 | 0.167874 | 3.7 | 4.6 |
| **Variable** | **course\_Certificate\_type** | **Maximum** | |
| course\_rating | COURSE | 5 | |
|  | PROFESSIONAL CERTIFICATE | 4.9 | |
|  | SPECIALIZATION | 5 | |

**Table 2**

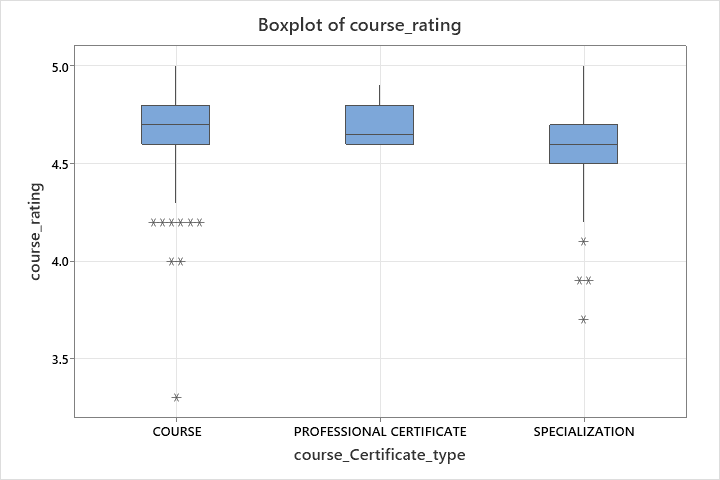


**Figure 3**

**Histogram of Course Ratings by Certificate Type**

Histograms showed that "COURSE" had a normal distribution, "SPECIALIZATION" had slightly lower ratings, and "PROFESSIONAL CERTIFICATE" had limited data, making conclusions tentative.

**Interpretation:**  
  
 This data is a comparison of ratings on courses of three categories: COURSE, PROFESSIONAL CERTIFICATE, and SPECIALIZATION.  
  
COURSE has the highest average rating (4.71) and highest number of data points (582). PROFESSIONAL CERTIFICATE is lower on average (4.7) but has incredibly few data points (12). SPECIALIZATION has the lowest average rating (4.62) but has an amazing number of data points (297).  
  
Histograms visually confirm these findings, showing each certificate type's distribution of ratings. COURSE and SPECIALIZATION have a reasonably normal distribution, while that of PROFESSIONAL CERTIFICATE is not so easily readable due to having very little data.



**Figure 4**

**Boxplot of Course Ratings by Certificate Type**

Boxplots showed similar median ratings (~4.5). However, "COURSE" showed more outliers, suggesting varied student experiences.

**Interpretation:**  
  
 This boxplot compares course ratings for three certificate types: "COURSE", "PROFESSIONAL CERTIFICATE", and "SPECIALIZATION". The boxes stand for the middle 50% of ratings (the interquartile range), and the line inside each box is the median rating.

Graphically, "COURSE" and "PROFESSIONAL CERTIFICATE" are also nearly equal in median rating, at just over 4.5, and "SPECIALIZATION" is slightly lower. The "COURSE" and "PROFESSIONAL CERTIFICATE" boxes are also higher, showing more spread within those categories of certificates.

The whiskers that project from the boxes signify the range of typical ratings. Outliers, denoted by asterisks, are ratings that lie outside this range of typical ratings. As one sees, "COURSE" holds several low outliers, which means that there are some courses that receive significantly lower ratings than the majority.

The median ratings are relatively close, but "COURSE" and "PROFESSIONAL CERTIFICATE" have a larger range of ratings, and "COURSE" has a greater tendency toward low ratings.

**Discussion of Findings:**

* Difficulty Level: Easier courses (Beginner, Mixed) consistently rated higher, suggesting this appeal more broadly.
* Certificate Type: Standalone "COURSE" and "PROFESSIONAL CERTIFICATE" are rated higher than "SPECIALIZATION," possibly due to perceived complexity or commitment.

**Recommendations:**

* Provide more support for advanced courses.
* Enhance engagement in specialization courses.
* Future studies should expand on demographic analysis to deepen understanding.

**Conclusion**

This study shows students prefer easier courses and simpler certificate types, valuing manageable workloads and clear outcomes. Advanced and specialized courses need targeted improvements to increase their attractiveness. Addressing these issues can greatly enhance student satisfaction and completion rates.

**Reference**  
<https://www.kaggle.com/datasets/siddharthm1698/coursera-course-dataset>

|  |
| --- |
|  |

Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. Computers & Education, 98, 157-168.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. Learning at Scale (L@S), 2013.

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers & Education, 50(4), 1183-1202.

Jordan, K. (2015). Massive Open Online Course Completion Rates Revisited: Assessment, Length, and Attrition. International Review of Research in Open and Distributed Learning, 16(3).