

COMP60411 January 2020

Exam Performance Feedback

This feedback is based for the **COMP60411** exam that took place in **January 2020 online**.

This report is based on Blackboard Item analysis report that can be found in your course unit (Course Tools > Tests, Survey and Pools – right click on the name of the test/exam and select Item Analysis).

Overview of student performance

Average Exam score: **57.98% (31.31/54) STD DEV 9.63%**

The exam had **26 objective questions** (Multiple Choice Questions, Multiple Answer Questions, True/False Questions) and **6 short essay questions**. On average, students **scored 65% on the objective questions** and **48.3% on the short essay questions**.

Average Exam Time: **1 hour 58 mins**

Difficulty

Below is a summary of the questions' difficulty.

Number of questions	Difficulty Bin	Difficulty (The percentage of students who answered the question correctly)
9	Easy Questions	greater than 80%
21	Medium Questions	30% to 80%
2	Hard Questions	less than 30%

Discrimination

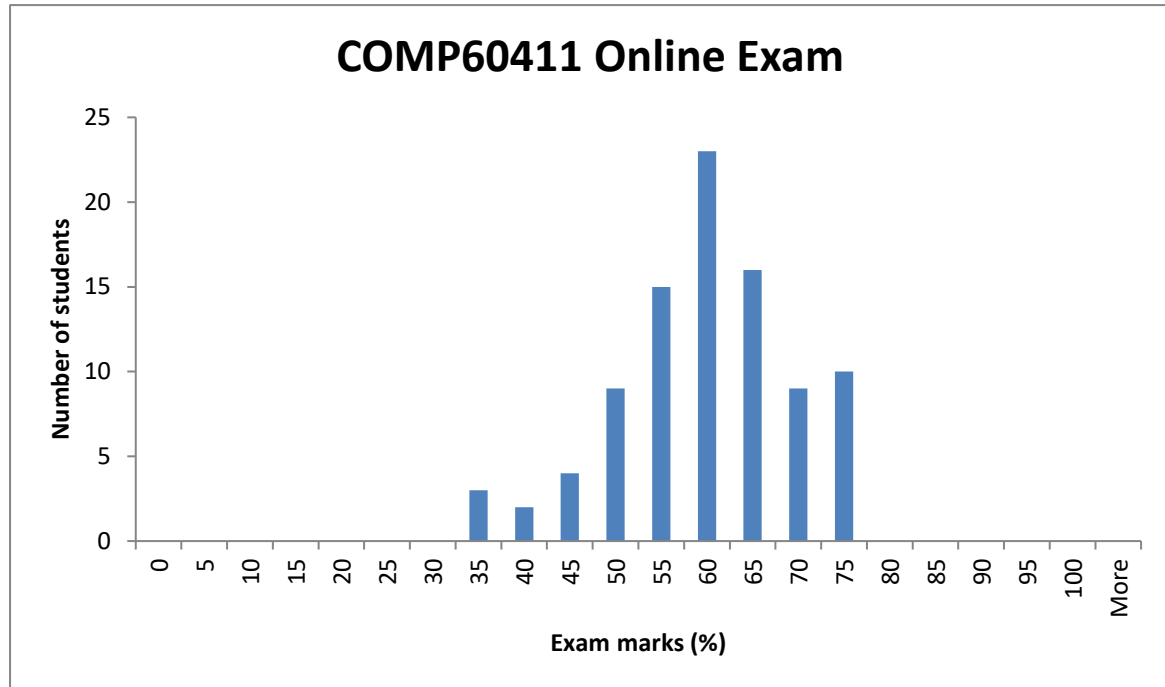
Next is the number of questions per discrimination: this is a metric that indicates how well a question differentiates between students who perform well and those who don't, as indicated by their performance on other questions. A question is a good discriminator when students who answer this question correctly also do well on the other questions. Values can range from -1.0 to +1.0. Questions are flagged for review if their discrimination value is less than 0.1 or is negative. Discrimination values can't be calculated when the question's difficulty score is 100% or when all students receive the same score on a question.

We reviewed all questions, and 'cancelled' 1 of the poor questions - a 1 mark MCQ that had a typo and confused students - we gave all students the full 1 mark for it. The other questions were reasonable.

15	Good questions
14	Fair Questions
2	Poor Questions
1	Cannot Calculate

Distribution of Marks

Overall, students performance shows a normal but low distribution, with only 19/91 students getting at least 70% of marks. Performance on objective question was considerably better than on subjective ones, as detailed above.



Detailed Observations on Subjective Questions

Question 27: In 4-5 sentences, explain the main similarities and differences between CSVs, JSON, ...[5 marks, average mark achieved: 2.37/5] This was a question of medium difficulty. Most marks were lost by incorrect claims related to which of these data models were (not) self-describing and how schemas can be used to make them 'more' self-describing. Another area of confusion related to what kind of data can be shared (how easily) via which of these models.

Question 28: Sketch out a RelaxNG schema for XML documents ...[4 marks, average mark achieved: 3.26/4] This was an easy question - as expected - and most students performed well.

Question 29: Consider again the XML-based format [...]: sketch out a similar JSON-based format for a recipe ...[3 marks, average mark achieved: 1.92/3] This was another easy question, and students performed well, as expected. I had expected them to *describe* the JSON-based format, but almost all translated the given snippet into JSON - which was also fine.

Question 30: Consider again the XML-based format for documents [...]: in 2-3 sentences, explain where and how XML Schema (XSD) datatypes ...[3 marks, average mark achieved: 1.09/4] This was a question of medium difficulty. Most students noticed that atomic XSD data types can be used to restrict the value range for attributes like "isVegan" and "caloriesPerUnitOfMeasure", but only few noticed that we can use enumerations for the value range of "unitOfMeasure". Finally, almost all

students failed to spot that we use the 'name' attribute on ingredient elements as an ID for ingredients, and on "usesIngredient" elements as an ID-reference, and that we can use XSD data type to ensure uniqueness of the former, and presence of the latter.

Question 31: Consider again the XML-based format for documents [...] in 4-5 sentences, explain two major weaknesses of this format and how to improve it... [4 marks, average mark achieved: 0.8/4] This was one of the most difficult questions. We saw many wild claims and many answers related to schema usage rather than the format given via this example. Many students claimed that having a list of ingredients separate from the recipes was bad (sometimes without explaining why and sometimes via a confused "they are not linked to recipes") but then did not discuss how "rolling" the ingredient information into the recipes would risk inconsistencies in their recorded features (e.g., eggs could be vegan in one recipe and non-vegan in another one). Many students also (wrongly) claimed that the 'step' elements aren't ordered and that having "step1", "step2",...would be better. Other students claimed that having so many attributes was bad, but did not explain why (and there are no clear reasons why that should be the case). Many students correctly pointed out that (and where) the format contains redundancy, why this is bad, and how that can be fixed.

Question 32: In 2-4 sentences, describe how you can use schemas to follow Postel's law in a robust application that produces and consumes data shared on the web... [4 marks, average mark achieved: 1.36/4] This was a question of medium difficulty. Almost all students could correctly state Postel's law, and most were able to explain that we can use different schemas, of varying strictness, to follow it. Most failed, however, to also mention other aspects of this usage, in particular schemas with good error reporting and schemas with good usage of default values to follow it.