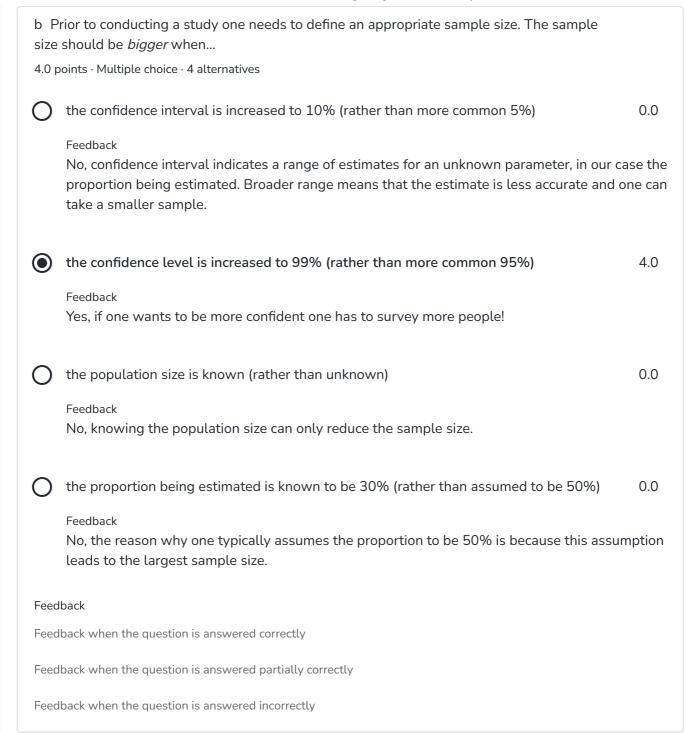# January 2023

2IMP40 Empirical Methods in Software Engineering 22/23 · 4 exercises · 100.0 points

1   Part 1 (20 points)

20.0 points · 5 questions

The first part of the exam consists of **five** multiple choice questions. Each question has four answer options and exactly one is correct. Each correctly answered question gives four points. We estimate that "an average student" will not spend more than 18 minutes answering these questions.

Text

---

a   Laboratory experiments are appropriate when the researchers aim at

4.0 points · Multiple choice · 4 alternatives

◯ building a new theory based on the existing ones	0.0
◉ establishing causal relations	4.0
◯ gaining profound understanding of the study participants' experiences or opinions	0.0
◯ understanding the phenomenon under study in the most realistic context	0.0

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

b Prior to conducting a study one needs to define an appropriate sample size. The sample size should be *bigger* when...

4.0 points · Multiple choice · 4 alternatives

○   the confidence interval is increased to 10% (rather than more common 5%)   0.0

      Feedback
      No, confidence interval indicates a range of estimates for an unknown parameter, in our case the proportion being estimated. Broader range means that the estimate is less accurate and one can take a smaller sample.

◉   **the confidence level is increased to 99% (rather than more common 95%)**   4.0

      Feedback
      Yes, if one wants to be more confident one has to survey more people!

○   the population size is known (rather than unknown)   0.0

      Feedback
      No, knowing the population size can only reduce the sample size.

○   the proportion being estimated is known to be 30% (rather than assumed to be 50%)   0.0

      Feedback
      No, the reason why one typically assumes the proportion to be 50% is because this assumption leads to the largest sample size.

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

c A *p-value* is

4.0 points · Multiple choice · 4 alternatives

○   the probability that the null-hypothesis is true   0.0

○   the probability that the alternative hypothesis is true   0.0

◉   the probability of obtaining the results observed (or more extreme ones) if the null-hypothesis is true.   4.0

○   the probability of obtaining the results observed (or more extreme ones) if the alternative hypothesis is true.   0.0

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

d When conducting *grounded theory* research it is crucial to

4.0 points · Multiple choice · 4 alternatives

○   use appropriate statistical tests   0.0

   Feedback
   No, grounded theory is a qualitative analysis technique

◉   interleave data collection and data analysis steps   4.0

   Feedback
   Yes, this is one of the main features of grounded theory building.

○   use interviews as the data collection method   0.0

   Feedback
   No. While interviews are indeed often used for data collection, this is not necessarily the case.

○   conduct a very detailed literature study prior to starting collecting and analysing the data   0.0

   Feedback
   No. Lightweight literature study might be a good idea to ensure that a new theory is need but a
   very detailed literature study is usually seen as a bad idea as it might bias the data collection
   process.

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

e Before conducting interviews the researcher must

4.0 points · Multiple choice · 4 alternatives

○   follow a training on interview techniques   0.0

> Feedback
> While such a training can be beneficial it is not a formal requirement.

○   contact the interviewee's manager to ensure the manager's consent   0.0

> Feedback
> This might be problematic as this might put the interviewee in an uncomfortable position, threaten
> their job or lead to untrustworthy answers. Furthermore not everyone might have a manager such
> as students, volunteer developers, or freelancers.

○   mine a software repository (such as GitHub or Stack Overflow) to identify potential
       interviewees   0.0

> Feedback
> While this is often done, interviewees can be identified in multiple ways.

⊙   seek approval of the Ethical Review Board   4.0

> Feedback
> Yes, any study involving human subjects should be approved by the ERB

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

2   Part 2 (40 points)

40.0 points · 8 questions

The italic text included below is a summary of an introduction of a recent empirical paper. This summary concludes with a research question. For this part of the exam, we ask you to sketch an empirical study that answers the listed research question. This part is divided over several sub-questions which ask you to indicate your choice for research strategy and to describe the data collection, and data analysis methods. In particular, for your data collection we ask you to explicitly describe and discuss your sampling. In addition to sketching your study design, we also ask you to motivate each part.

To grade your submissions we will consider the alignment between your chosen strategy and sampling, data collection and data analysis methods. Secondly, we are looking for detailed and succinct descriptions of your proposed data collection, sampling and data analysis that sketch a study that can be used to address the listed research question. In addition to that, we will also consider whether your motivation contains enough details to argue that the chosen solution is appropriate for the research question.

We estimate that "an average student" will not spend more than 36 minutes answering these questions.

Context and Research Question:

Self-Admitted Technical Debt (SATD) refers to source code comments (as well as other annotations elsewhere) indicating that the corresponding source code is (temporarily) inadequate, e.g., because the implementation is incomplete, buggy, or "smelly" (i.e., implemented in a suboptimal way).

Different authors have studied the sentiment and emotions expressed by developers. In particular, Mantyla et al. (2016) and Murgia et al. (2014) studied emotions expressed in the context of issue reports, finding that there appears to be a link between issue priority and complexity and negative emotions present in issue reports. When describing SATD, developers could express the same concept in neutral or in a rather negative fashion. However, nobody has studied how priority is expressed in different kinds of software development issue reports—and in particular SATD-related issues—and whether developers use negative sentiment to indicate priority. This leads us to ask the following research question:

RQ: How do developers annotate SATD that they believe requires extra priority?

Text

Research Strategies

Text

a Please indicate your choice for research strategy used to answer the research question.
This is merely an indicative question, answering this question is mandatory but it does not grant points.

Multiple choice · 8 alternatives

○   Sample Study   0.0

○   Judgement Study   0.0

○   Lab Experiment   0.0

○   Experimental Simulation   0.0

○   Field Experiment   0.0

○   Field Study   0.0

○   Computer Simulation   0.0

○   Formal theory   0.0

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

b Motivate your choice for the research strategy indicated in your answer to Question 2a.
Explain why your choice is appropriate for the research question stated above.

10.0 points · Open · 3/10 Page

+5 points
The answer gives a convincing argument on why the chosen research strategy is appropriate for the research question.

+5 points
The arguments are based on specific properties of the chosen research strategy. Showcasing that there is a detailed understanding of the research strategy.

Data Collection
Text

c Please describe a data collection method that can be used to answer the research question, given your research strategy.

5.0 points · Open · 2/5 Page

+2.5 points

The chosen data collection method is described in sufficient detail to understand the proposed study.

+1 point

The chosen data collection method is described in general terms only.

+2.5 points

The chosen data collection method uses the best practices as discussed during the course.

0 points

The chosen data collection method does not adhere to the best practices or it is not clear whether it adheres to the best practices.

d Motivate your choice of data collection method and explain how it aligns with your research strategy and the research question.

5.0 points · Open · 2/5 Page

+2.5 points

The data collection method described aligns with the chosen research strategy.

+2.5 points

The motivation puts forward a strong argument as to why the data collection method is appropriate for the research question.

Sampling

Text

e Describe, where needed in your data collection process, how you plan to sample.

5.0 points · Open · 2/5 Page

+4 points

The sampling strategy is described in sufficient detail to be reproduced by an independent researcher.

+2 points

The main idea of the sampling strategy is clear and valid but the description is not detailed enough to be reproduced.

0 points

The sampling strategy is invalid or unclear

+1 point

It is clear where the sampling strategy fits in the data collection process.

f Motivate your sampling choice(s) and explain why they are appropriate for your chosen research strategy, data collection method and the research question.

5.0 points · Open · 2/5 Page

+2 points
The sampling choice is well-motivated

+1 point
The sampling choices are clearly aligned with the research strategy chosen.

+1 point
The sampling choices are clearly aligned with the data collection method chosen.

+1 point
The sampling choices are clearly aligned with the research question.

Data Analysis
Text

g Describe a data analysis method that you propose to use to answer the research question, given your data collection method and sampling strategy(s).

5.0 points · Open · 2/5 Page

+5 points
Data analysis method is described in sufficient detail to be reproduced by an independent researcher

+2.5 points
The main idea of the data analysis method is clear and valid but the description is not detailed enough to be reproduced.

h Motivate your choice of data analysis method and explain how it aligns with your data collection, sampling, research strategy and the research question.

5.0 points · Open · 2/5 Page

+1 point

The data analysis method can be used to answer the research question.

+1 point

The data analysis method is aligned with the research strategy.

+0.5 points

The data analysis method is aligned with the data collection method.

+0.5 points

The data analysis method is aligned with the sampling strategy.

+2 points

The choice of the data analysis method is well-motivated.

3　Part 3 (40 points)

40.0 points · 6 questions

For the third part we ask you to read the accompanying paper. Please read the paper and come-up with two viable and distinct threats to validity. For each threat to validity please use the answer fields below to describe and classify it. The threats you describe should describe **two different types of threats to validity** (So if the first threat to validity you describe belongs to the category external, the second threat you describe cannot belong to the category external).

You should describe your first threat to validity in 3.a, describe the conclusion it invalidates in 3.b and classify it in 3.c, and you should describe your second threat to validity in 3.d, describe the conclusion it invalidates in 3.e and classify it in 3.f. We estimate that "an average student" will not spend more than 36 minutes answering these questions.

Text

---

a　Describe your first threat to validity

9.0 points · Open · 1/2 Page

+9 points

The answer describes a threat to validity that could indeed threaten the validity of the study

---

b　Explain which conclusions of the study would be invalidated by the threat and how

9.0 points · Open · 1/2 Page

+9 points

The answer points to a specific conclusion of the study that might be impacted by the threat to validity.

---

c　Classify your first described threat as

2.0 points · Multiple choice · 4 alternatives

⦿　Conclusion validity　　　　　　　　　　　　　　　　2.0

⦿　Construct validity　　　　　　　　　　　　　　　　2.0

⦿　External validity　　　　　　　　　　　　　　　　2.0

⦿　Internal validity　　　　　　　　　　　　　　　　2.0

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

d Describe your second threat to validity

9.0 points · Open · 1/2 Page

+9 points

The answer describes a threat to validity that could indeed threaten the validity of the study

e Explain which conclusions of the study would be invalidated and how

9.0 points · Open · 1/2 Page

+9 points

The answer points to a specific conclusion of the study that might be impacted by the threat to validity.

f Classify your second described threat as

2.0 points · Multiple choice · 4 alternatives

⦿ Conclusion validity   2.0

⦿ Construct validity   2.0

⦿ External validity   2.0

⦿ Internal validity   2.0

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

4 Extra page

0.0 points · 1 question

Extra space for answers. Please indicate what question each answer belongs to.

Open · 1 1/2 Page