**Evaluation form.**

**Course: Preparation for the RM thesis (MSBBSS12)**

**Title of report: Multilevel Multivariate Imputation by Chained Equations through Bayesian Additive Regression Trees**

**Student name: Heleen Brüggen**

**Student ID:**

**Supervisor name and title: Thom Volker, Msc., Hanne Oberman, Msc., Gerko Vink, Dr.**

**Research institute or university: Utrecht University / M&S**

**Date: 09/01/2023**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Grade:** | **Less than 5.5**  **Unsatisfactory:** | **Between 6 and 7.5**  **Satisfactory to good** | **Between 8 and 10**  **Good to Excellent** | **Weight** | **Grade** |
| Part A: Student’s performance | | | | | |
| **Amount of work done** | The amount of work done was less than would have been expected on the basis of the amount of ECTS. | The amount of work done was what would have been expected on the basis of the amount of ECTS. | The amount of work done was more than would have been expected on the basis of the amount of ECTS. | 0.25 | 7.5 |
| **Work independency** | Student’s progress heavily depended on supervisor’s ideas and input; many feedback sessions were required. The student made many errors that had to be corrected by the supervisor. | Supervisor did not need to provide excessive input; regular feedback sessions were held as scheduled. The student made few errors that had to be corrected by the supervisor. | Minimal input from supervisor was required; the amount of feedback needed was minimal. he student made hardly any errors that had to be corrected by the supervisor. | 0.25 | 9 |
| Part B: Quality of the report | | | | | |
| **Problem definition** | The purpose or the research problem is not clearly articulated, or its component elements are not identified or described. | Adequately identifies and describes (or sketches out) the purpose and the research problem and its components. | Effectively formulates a clear description of the purpose and the research problem and specifies major elements to be examined; | 0.20 | 7,5 |
| **Academic quality** | Information is poorly organized and/or superficially examined; information is often inaccurate or incomplete; presents little if any analysis or interpretation; inaccurately and/or inappropriately applies theories, frameworks, research methods, techniques, and/or models, to the analysis; presents few conclusions or recommendations; conclusions or recommendations are often not well supported, inaccurate, and/or inconsistent, and are presented in a vague or rudimentary manner; reflection/discussion is missing or lacks depth. | Gathers and examines information relating to the research problem; satisfactorily presents and appraises information with only minor inconsistencies, irrelevancies, or omissions; generally applies appropriate theories, frameworks, research methods, techniques, and/or models, with a few minor inaccuracies; outlines conclusions and recommendations that are logical and consistent with the analysis and evidence; identifies and/or lists conclusions and recommendations in a clear manner; sufficiently reflects on the results and the methods used. | Selects and prioritizes information appropriate to addressing the research problem; accurately and appropriately analyses and interprets relevant information; precisely and effectively applies appropriate theories, frameworks, research methods, techniques, and/or models, in developing and justifying multiple conclusions and recommendations; conclusions and recommendations are insightful, coherent, well supported, logically consistent, and complete; demonstrates; critically analyses the strengths and weaknesses of the project. | 0.20 | 8 |
| **References** | Most of the references are from sources that are not professional, and have uncertain reliability, or are mostly outdated; few if any appropriate citations are provided; reader doubts the validity of much of the material; frequent and significant errors in formatting. | Professional and up-to-date references are generally used; clear and fair citations are presented in most cases; most of the information/ evidence comes from sources that are reliable; formatting is employed appropriately in the research project with a few minor errors. | Presents compelling evidence from multiple professionally legitimate and up-to-date sources; attribution is clear and accurate; references are primarily professional journals or other approved sources; formatting is done accurately and consistently throughout the project. | 0.10 | 9 |
| **General comments** | Multiple advanced topics came together in this report: multiple imputation, multilevel modelling and non-parametric statistics (and setting up a data-generation scheme under this complex structure that can be used almost one-on-one in the actual thesis). At the same time, there was no missing data involved yet, and it would have been a nice addition if this was already included in this pilot.  Heleen is very independent in collecting the necessary information, writing up the initial manuscript and performing the statistical simulations. All work was done in an independent way, feedback was dealt with professionally and independently. Heleen required only textual feedback to arrive at a concise and well-structured Research Report.  Heleen has a clear picture of the necessary goals and expectations and has properly identified the purpose of the research and formulated this into a sufficient problem statement. All components of the problem description are there, but not all aspects fit perfectly together in the current form, so structure-wise there are still some improvements possible. Nevertheless, I think a good description of the problem is given, and the proposed strategy suits this description nicely.  The quality of the work performed is good and the statistical simulation is well-executed and well-presented. The provided background gives the right amount of introduction to the problem and the important literature is there and well summarized. The illustrative simulation is well thought off and executed in an appropriate manner, although I think that the data-generation section contains a lot of details that are relatively challenging to consume, perhaps a different way of formulating these details would have made it easier to follow (e.g., writing it out in matrix form would reduce the number of variables in text to keep track of tremendously). Also, the evaluation scheme is not entirely clear from text (evaluating with respect to predictions or parameters; if the former, is it within-sample or out-of-sample). All conclusions are provided with the necessary caution due to the limited set-up, but are a bit superficial at times  References are well chosen and summarized; all important background info is there. | | | | |
| **Overall grade** | 8.125 | | | | |

The deadline for handing in this form will be communicated by Mirjam Moerbeek during the course of the fall semester.