

Report title, e.g., Impact of weather on sales of cabbage
Bayesian statistics, EBM805C05, Group no

Claire Gun (s123456) and Clark Rifle (s654321)

Include also a phone number of one you so
that we (= avv and nvf) can discuss the proposal with you.

May 2, 2021

Abstract

This is the template for the report.

GUIDE LINES

General info

1. Students in groups of three write a report. We rely on you forming groups.
2. Each group has to choose a method (from the primary book, or from one of the other sources below) which the members of the group are interested in learning. A case based on a company at which one of the student is doing a master thesis would be an excellent choice. Otherwise, the group has to find a problem on the internet with data to which to apply their method of choice.
3. Each group has to review an intermediate and near to final version of the report of an other group.
4. The core of the report is at most 10 pages. Appendices are not included, but when long, we (=avv and nvf) will skip them.
5. Each group chooses another group to exchange feedback. Each group gives and receives feedback twice, on an intermediate and near final version of the report.

0.1 *Methods*

It is sufficient to understand the basic methods, and, as a general principle, it is best to start with simple methods based on what you have learned during the course. Once you understand the simple methods, you can move on to more difficult ones. If you use such harder methods, do not forget to explain in your report the simple methods you studied first, and what motivated you to dive into more challenging methods.

We do not give grades based on the performance of your methods; the grading is more done based on your explanation of the methods and *why* you think they perform good or bad. Or, after you did the analysis, why your methods does, or doesn't work, and what tests you used to convince yourself about your claims.

If (part of) your report is based on an project for a company, don't feel obliged that your method(s) 'must work'. The best advice is an honest advice. If you cannot find good methods to predict what they would like to see, just say so. (Let's reverse it, suppose your predictor works lousy, but you say it gives fantastic results. This is a recipe for disaster.) You already 'deliver' by saying that some things cannot be done.

You don't have to explain each method you use in your report. However, you should explain just one such method in considerable detail. The motivation is this. If you can explain one method well, we trust that you can also explain other methods well. So why do we mind about details? Well, after finishing your master, you have to be able to explain such methods at various level of detail to a board of managers. When there are managers with a math, computer science, engineering background, they might want to understand in very much detail what you built. For instance, my (=nvf) manager at Quintiq was Victor Allis. (You can find him on wikipedia.) Whenever somebody proposed some algorithm, he always wanted to really understand things in detail. And if you could not convince him during a meeting, you could go back to your desk, and start thinking again. There are also other types of managers that ask nasty questions, just to check whether you did your work thoroughly. People do not invest several millions of euros, just

because you say things are ok. You have to convince them, and show that you know what you are talking about. And if you fail at that skill a few times in row, you will have a problem, if you want to keep your job at least.

Your report should show that you know what you are talking about. We realize that this is a hard challenge. But both in business and academia, people don't like to read long reports. So, train and learn. Again, there is no way around for you (unless you decide to become a shepherd, for instance).

Time lines

Here is a time line to help you organize your work.

1. Friday 23/4, 17h: topic choice uploaded to nestor. Keep it short, and use this as a start for your intro. Please include some motivation about why you chose this topic, why you think it is suitable, and feasible within the course. The idea is that we (= avv and nvf) can give you some feedback about whether the topic is good, too simple or too hard. Include a phone number on the first page of your report of at least one of you so that we can contact you to give verbal feedback.
 2. Friday 23/4: Aim to have Section 2 written. No upload to nestor.
 3. Friday 30/4, 17h: research proposal uploaded to nestor. Section 3 written. Send it to the other group for feedback.
 4. Tuesday 4/5, 17h: intermediate feedback uploaded to nestor. Feedback on intermediate report sent back to the other group.
 5. Friday 7/5 17h: Mutual feedback discussed.
 6. Friday 28/5 17h: Final version 0.1 uploaded to nestor. Methods implemented and tested, and report nearly finished. Ensure you dealt with the feedback. Send report to feedback group.
 7. Tuesday 1/6, 17h: Final feedback report uploaded to nestor. Plan a feedback session.
 8. Friday 4/6, 23h59: Final report uploaded to on nestor. Ensure again that you dealt with the final feedback.
1. Include the tex files in your uploads to nestor.
 2. Also include the feedback, in the appendix. We don't include your feedback in the page count of your own report.

1 INTRODUCTION

1. What is the problem you want to analyze? State it in clear terms.
2. background/relevance: why is it a problem? Where does it come from?
3. Type of problem, e.g., classification, regression
4. We encourage you to use a case of the company where you are presently writing your master thesis.

The report has the following structure. After having introduced the data set in Section 2 we describe in Section 2 the methods to analyze the data and validate the analysis. ETC.

2 DATA DESCRIPTION

1. Why this data, how does it help to answer your question?
2. Describe the data. Where did you get it from? Include dates, and the like. Ensure that if people want to repeat your work, they can find your dataset.
3. Just mention simple filters you applied
4. Explain difficult filters, if this was necessary.
5. Include in any case some samples of the data so that people can see how things are organized.
6. Use graphs, and tables to provide overview, statistical insights such as size of the data.

3 METHODS AND VALIDATION

1. What method are you going to use? Where does it come from: ref to the literature, paper, book(s)?
2. Why is that method(s) suitable for your problem? To what extent, and what are its limitations?
3. How are you going to validate your analysis? e.g., cross validation? Why this method?
4. You are not bound to using the methods we discuss in the course. If you like to use deep learning, for instance, then that's fine too. However, in case you want to use a method that is not part of the course, mail us (= avv or nvf) and ask whether we agree. (Prevent surprises at the end.)

4 ANALYSIS AND INSIGHTS

1. What can we learn/see from the data?
2. How confident are you about your claims?
3. Can you come up with an estimate of the expected gain (monetary, or other KPIs) by using your methods?
4. Include graphs (mostly) and tables (sometimes). Some(many?) people tend to skip tables. Include a short description in the caption of the figures. You should know that people often start with glancing over the figures, and skip most (or all) of the text. Hence, figures and captions are your main initial selling points.

5 CONCLUSION AND SUMMARY

1. Main findings, quality/reliability of findings,
2. evaluation of chosen methods, and extensions.

A APPENDIX

1. Include here the core parts of your code, and explain how it works.
2. Don't include long tables with data, because we (=avv and nvf) will just skip them.

Python Code

```
1 a = 10
2 b = 5
3 print(a*b)
```

R Code

```
1 a <- 5
2 b <- 10
3 b*a
```

B FEEDBACK ON INTERMEDIATE REPORT

Review the report as if it is a research proposal. The main question is whether you are confident about:

1. the relevance of the problem,
2. feasibility of answering the problem? Is it too simple or too hard?
3. is the chosen method scientifically suitable to answer the question? How hard to you think it is to get it working?

Guidelines:

1. At most one page feedback.
2. Be respectful. It's your job to help the authors improve their work. So, you give suggestions, but it's their responsibility to use them or not. It is not your job to solve the problems of the authors, neither to make them unhappy.
3. When reading, ensure you really understand each and every step.
4. If you particularly (dis)like certain aspects of the other's report, use that to improve your own report. Try to figure out why you (dis)like that aspect.
5. Give a grade at the end: bad, average, good, excellent (< 4, 6, 8, 10).

B.1 *Problem and method*

1. Do you understand the problem, motivation, chosen method?
2. Include advice on what to improve, and why.

B.2 *Style*

1. Quality of writing?
2. Presentation: use of graphs and tables.

C FEEDBACK ON FINAL REPORT

Review the report as if you are a customer who hired the authors for a project. The main question is whether you are happy about your investment. You paid

$$3 \times 60h \times 50\text{€}/h \approx 10K\text{€}.$$

Is it well spent? Would you hire them again?

The rest of the feedback below is to help the authors improve their work. Guidelines:

1. At most one page feedback.
2. Be respectful.
3. When reading, ensure you really understand each and every step.
4. If you particularly (dis)like certain aspects of the other's report, use that to improve your own report. Try to figure out why you (dis)like that aspect.
5. Give a grade at the end: bad, average, good, excellent (< 4, 6, 8, 10).

C.1 *Problem and method*

1. Do you understand the problem, motivation, chosen method?
2. Include advice on what to improve, and why. Better method perhaps?

C.2 *Analysis*

1. Do you understand the problem, motivation, chosen method?
2. Include advice on what to improve, and why. Better method perhaps?
3. Ideas on to improve the type statistical analysis, quality of analysis.

C.3 *Style*

1. Quality of writing?
2. Presentation: use of graphs and tables.