

Preface

This is a case collection in data analysis containing cases mostly relevant for business studies. The first part of ten cases are mini-cases, that is, with a limited issue and a small dataset. The second part contains wider issues and most often larger datasets. The cases within each part are numbered in the order topics typically will occur in an introductory courses in statistics. For some topics that are not commonly treated in textbooks, we have supplied some theory. All cases are based on problems and data collected over many years. A few data sets are slightly modified in order to bring forward an issue more clearly and/or to bring anonymity to the data. In some cases the years are changed to “bring the problem into the current century”.

The name of the case may indicate the problem area. Each case starts with a topic statement, indicating the data type and modes of analysis. Then the context of the case is given and the problem to be solved, finalized as a task statement. For some cases this is detailed in two versions, giving the teacher and student a choice between two levels of challenge. The A-version is open with respect to approach to the problem, while the B-version is structured, with specific and itemized questioning, often with indication of recommended method of analysis. One possibility is that the students read the A-version first, and make up their minds on how the problem may be attacked and then take a look at the B-version to see whether this contains additional elements not thought of. The solution may then be done according to the B-version. This gives the opportunity for the students to reflect on the choice of method, an important ability in practice. The students may alternatively be given the challenge of attacking the problem, without help of the B-version (as in business practice), or go directly to the B-version (as often final exams are structured). The fourth possibility, of going direct to our solution, is not recommended.

Many problems have no unique best solution, and there are often several roads leading to the same solution or an equally good one. Our solution is maybe just one of the possible ways to attack the problem, and student creativity should not be restrained by this. The ability to attack new problems is vital in business practice, and an approach which deprives the student of such training is inferior. Some cases can be attacked on different level of theoretical knowledge and sophistication, and the one chosen in practice have to match the knowledge at hand. In such cases it is worthwhile to judge whether spending time on learning a more sophisticated method is worth the effort. In some cases we present a solution which may not seem completely satisfactory from a strict statistician’s viewpoint. This may give a statistics teacher an opportunity to comment and indicate a better way.

This collection of cases can be used in conjunction with any elementary or intermediate textbook in statistics, typically one with some emphasis of probability. It can also be used alone, if the reader is willing to consult available resources when needed, for instance on the internet. Suitable software is needed for actually doing the cases. A lot can be accomplished in spreadsheet programs like Excel, but they have limited analytical capabilities. Add-ons like XLSTAT may remedy this, but a sound statistical package like Minitab is recommended. Such packages often have good help functions and tutorials, which may be used to learn some theory behind the methods as they are applied. The output shown in the solutions are taken from Minitab exclusively. If students are given access to solutions, a possibility is that they are required to repeat the analysis in Excel (if possible) or an add-on (if available). The data are supplied for each case as an Excel worksheet file.

I owe thanks to many individuals for sharing their problems and data. In order to secure anonymity no one is mentioned here. I also acknowledge financial support from the “Faglitterære fond”.