Introduction to the Course

Quantititive Methods

Dr. Philippe J.S. De Brouwer
Honorary Consul of Belgium in Kraków
guest professor at the UJ, AGH, UEK and UW
board member of AGH and ISK
SVP at HSBC in Kraków

2024-10-01

Table of Contents

Contents

1 Your Teacher

Dr. Philippe De Brouwer studied theoretical physics and later acquired a second Master –Business Engineer– while working full time. Finishing this Master he solved the "fallacy of large numbers puzzle" that was formulated by P.A. Samuelson 38 years earlier. In this Ph.D. he successfully challenged the assumptions of the Noble price winning "Mean Variance Theory" of H. Markovitz that dominated our thinking about suitability of investments for more than 60 years.

Early in his career he moved from insurance to banking and from IT to asset management. For Fortis (BNP) he helped the young investment management company grow, stood at the cradle of one of the first capital guaranteed funds and got promoted to director in 2000.

In 2002 he joned KBC and relocated to Poland for whom he merged 4 companies and became CEO of the merged entity in 2005. Under his direction the company climbed from number 11 to number 5 on the market. In the aftermath of the crisis he helped creating a new investment management company for KBC in Ireland that soon accommodated the management of ca. 1000 investment funds and had about 24 Bln Euro under management. In 2012 he widened his scope to financial risk management and specializing in statistics, analytics, data and numeri- cal methods.

In 2015 Philippe was head of Analytics Development for the Royal Bank of Scotland Group. Now he is again in Poland and is SVP at HSBC where he oversees the Model Review Centre of Excellence in Krakow (as well as additional functions such as coordinator of collaboration with universities)

Philippe De Brouwer

url: http://www.de-brouwer.com
LinkedIn: https://www.linkedin.com/in/philippedebrouwer
email: philippe@de-brouwer.com
mobile: +48 790 715 002

Figure 1: Scan this QR-code to obtain Dr. De Brouwer's business card and connect via LinkedIn

Who are you

Quick introduction
Who are you? What do you expect from this program? What do you want from this program? What should we focus on?

2 Practical Information

Availability of Slides and other materials
Course materials:

A. url: http://www.de-brouwer.com
B. select "Courses" and then your program
C. read materials there and download resources that you prefer (slides, handouts, or booklet)

Materials from "the Big R-Book" (De Brouwer, 2020):

A. videos and code: http://www.de-brouwer.com/publications/r-book/index.html
B. code and slides: http://www.de-brouwer.com/publications/r-book/18901229-for-teachers.html

3 The Program Quantitative Methods

3.1 Objectives

Objectives of the program

know	the basics of statistics and data manipulation
know	at least one analytical tool (R)
understand	the importance of data in decision making
understand	uses and limits of various methods
apply	understand limits of models
apply	make informed decisions
apply	write a technical paper
apply	write a presentation and present it

3.2 The content of the program

The content of the program

A. (optional) Getting started with R and its use

B. SQL Databases and importing data in R

C. Data wrangling (preparing data to build a model)

D. Building powerful models (linear regressions, generalised linear regression, non-linear regression, decision tree, random forest, SVN, neural network, etc.) and model validation

E. (optional) Introduction to companies and financial markets

F. Automating presentations, documents, etc.

G. (optional) Big Data

H. (optional) Code performance (speeding up R)

I. (extra/optional) Quantum Computers

J. (extra/optional) Crowdfunding and Fintech

The Big R-Book
From data science to learning machines and big data
De Brouwer, 2020

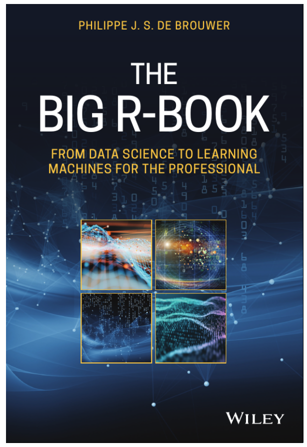

Figure 2: http://www.de-brouwer.com/publications/r-book/

3.3 The grading of the program

Grading of the course

- 40% presence and collaboration in classroom (including in-class assignments)
- 60% selected assignment (details see Chapter *Assignments*, page 5)

3.4 Assignments

Assignment

Definition 1 (the assignment)**.** Students are expected to gather data, analyse it and report the results in

A. a paper (between 5 and 50 pages)

B. a presentation (life in the classroom – during last course)

The format of the presentation (during the last lesson) is a "10 minutes elevator pitch" + 10 minutes questions

The Grading of the assignment

The assignment is 60% of the total grade. It splits as follows.

- 50% for the idea, logic, coherence, and conclusions
- 25% for the written materials (paper and/or slides)
- 25% for the presentation itself (quality of slides if used + oratorical qualities)

The assignments are teamwork.

Ideas for the content of the Assignment

A. solve a problem by using data and rely on one or more of the methods studied (eg. regression, MCDA), etc.

B. use an existing project/document/essay and improve the decisions proposed by using data

4 Levels of Measurement

Levels of Measurement

Introduction

It is customary to refer to the theory of scales as having been developed by Stevens, 1946. In that paper he argues that all measurement is done by assuming a certain scale type. He distinguished four different types of scale: nominal, ordinal, interval, and ratio scales.

4.1 Nominal Scale

Nominal Scale

The nominal scale is the simplest form of classification: only labels (with no order). Examples include asset classes, first names, countries, days of the month, etc.

Scale Type	Nominal
Characterization	labels (e.g. asset classes, stock exchanges)
Permissible Statistics	mode (not median or average), chi-square
Permissible Scale Transformation	equality
Structure	unordered set

Table 1: Characterization of the Nominal Scale of Measurement.

It is not possible to use statistics such as average or median. We can measure which label occurs the most (modus of mode).

Note that it is possible to use numbers as labels, but that this is very misleading. When using an nominal scale, none of the traditional metrics (such as averages) can be used.

4.2 Ordinal Scale

Ordinal Scale

This scale type assumes a certain order. An example is a set of labels such as very safe, moderate, risky, very risky. Bond rating such as AAA, BB+, etc. also are ordinal scales: they indicate a certain order, but there is no way to determine if the distance between, say, AAA and AA- is similar to the distance between BBB and BB-. It may make sense to talk about a median, but it does not make any sense to calculate an average (as is sometimes done in the industry and even in regulations)

Scale Type	Ordinal Scale
Characterization	ranked labels (e.g. ratings for bonds from rating agencies)
Permissible Statistics	median, percentile
Permissible Scale Transformation	order
Structure	(strictly) ordered set

Table 2: Characterization of the Ordinal Scale of Measurement.

Ordinal labels can be replaced by others if the strict order is conserved (by a strict increasing or decreasing function). For example AAA, AA-, and BBB+ can be replaced by 1, 2 and, 3 or even by -501, -500, and 500,000. The information content is the same, the average will have no meaningful interpretation.

4.3 Interval Scale

Interval Scale

This scale can be used for many quantifiable variables: temperature (in degrees Celsius). In this case, the difference between 1 and 2 degrees is the same as the difference between 100 and 101 degrees, and the average has a meaningful interpretation. Note that the zero point has only an arbitrary meaning, just like using a number for an ordinal scale: it can be used as a name, but it is only a name.

Scale Type	Interval Scale
Characterization	difference between labels is meaningful (e.g. the Celsius scale for temperature)
Permissible Statistics	mean, standard deviation, correlation, regression, analysis of variance
Permissible Scale Transformation	affine
Structure	affine line

Table 3: Characterization of the Interval Scale of Measurement.

Rescaling is possible and remains meaningful. For example, a conversion from Celsius to Fahrenheit is possible via the following formula, $T_f = \frac{9}{5}T_c + 32$, with T_c the temperature in Celsius and T_f the temperature in Fahrenheit.

An affine transformation is a linear transformation of the form $\boldsymbol{y} = \boldsymbol{A}.\boldsymbol{x} + \boldsymbol{b}$. In Euclidean space an affine transformation will preserve collinearity (so that lines that lie on a line remain on a line) and ratios of distances along a line (for distinct collinear points p_1, p_2, p_3, the ratio $||p_2 - p_1||/||p_3 - p_2||$ is preserved).

In general, an affine transformation is composed of linear transformations (rotation, scaling and/or shear) and a translation (or "shift"). An affine transformation is an internal operation and several linear transformations can be combined into one transformation.

4.4 Ratio Scale

Ratio Scale

Using the Kelvin scale for temperature allows us to use a ratio scale: here not only the distances between the degrees but also the zero point is meaningful. Among the many examples are profit, loss, value, price, etc. Also a coherent risk measure is a ratio scale, because of the property translational invariance implies the existence of a true zero point.

Scale Type	**Ratio Scale**
Characterization	a true zero point exists (e.g. VAR, VaR, ES)
Permissible Statistics	geometric mean, harmonic mean, coefficient of variation, logarithms, etc.
Permissible Scale Transformation	multiplication
Structure	field

Table 4: Characterization of the Ratio Scale of Measurement.

References

De Brouwer, Philippe J. S. (2020). *The Big R-Book: From Data Science to Learning Machines and Big Data*. New York: John Wiley & Sons, Ltd. ISBN: 978-1-119-63272-6.

Stevens, Stanley Smith (1946). 'On the theory of scales of measurement'. In: *Science* 103.2684, pp. 677–680.

Nomenclature

MCDA Multi Criteria Decision Analysis