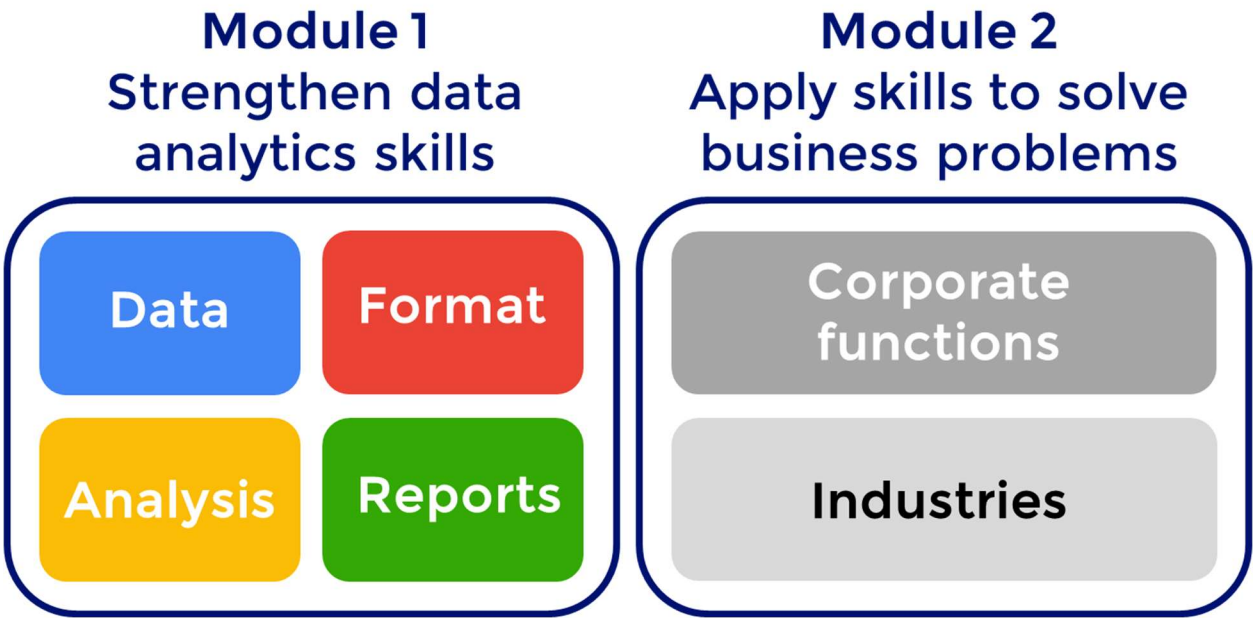
Solving Business Cases Using Applied Data Analytics [ZADA]
Summer 2023
OTH Regensburg
As of March 15, 2023

Professor: **Jonathan Whitaker**
E-mail: **jonathan.whitaker@external.oth-regensburg.de**
Class Room: **K220** (for on campus class meetings from June 5 - July 3)

Course Overview

This course will use a case-based approach, in which students apply SAS® data analytics software to solve business problems in multiple industries and corporate functions. SAS software is used by leading European firms such as Allianz, HSBC, ING, Lufthansa, Nestle, Rabobank, Vodafone, and Volvo.

The course will be structured into two modules. Module 1 will focus on basic SAS programming skills, and Module 2 will focus on applying SAS programming skills in a business case setting.



Mode of Instruction

Module 1 is designed to be delivered remotely through asynchronous video from March to May, and Module 2 is designed to be delivered in-person on the OTH RE campus during June and early July.

Grading

Components of the course grade will be as follows:

Component	Timing	Percentage of grade
Homework exercises	During Modules 1 and 2	30%
Problem sets	During Module 1	20%
Project	At end of the semester	50%

Homework

Homework exercises are drawn from the course textbooks. Each set of homework exercises is designed to require about 2-3 hours of work. SAS code for homework assignments will be submitted through the course Moodle site. Homework exercises will generally be assigned on a weekly basis.

I explicitly allow and encourage students to collaborate with each other on homework exercises. If students collaborate on a homework exercise, each student must be on their own computer, each student must write their own code, and each student must submit their own code on Moodle. I consider it to be cheating when a student cuts and pastes code from any other student or source, and/or when a student copies down code from another student or source without actively contributing to the solution.

Note on filenames for assignments. I will receive several hundred file submissions during the course, and I need your help to keep track of the files. Accordingly, I request four items for every submission:

1. Name your SAS file as *YourlastnameTwodigitchapternumberTwodigitexercisenumber*. The term 'two digit' means to put a leading zero in front of a one-digit number. As an example, for the first homework assignment (chapter 3 exercise 2, chapter 8 exercise 12, and chapter 8 exercise 14), my three file names would be *Whitaker0302*, *Whitaker0812* and *Whitaker0814*.

2. The first line of code for each homework assignment must contain a comment ** Firstname Lastname DD/MM/YY;* For example, for the first homework assignment due on March 27, the first line of my code would be ** Jonathan Whitaker 27/03/22;*

3. The second line of code for each homework assignment must contain a comment to list any students you collaborated with on the assignment. For example, if I collaborated on the first homework assignment with Steve Oakley and Sue Toomey, the second line of my code would be ** With Steve Oakley and Sue Toomey;* If I performed the assignment on my own, the second line of my code would be * *On my own;*

4. If the SAS code refers to a data file (or any type of file), write your SAS code such that the code would refer to the library name */home/yoursasID/sasuser.v94/Data/* for the file location. Please see details on next page for SAS® OnDemand for Academics.

I will deduct one point (out of six points) from each exercise if any of these four conventions are not followed.

Problem sets and final project

Problem sets will be more comprehensive than homework exercises. Each problem set is designed to require 4-6 hours of work. There will be a problem set midway through Module 1 and a problem set at the end of Module 1.

The final project will be distributed in class on June 26, and will be due on July 28. The final project will be designed to require about 20-25 hours of work. SAS code for problem sets and the final project will be submitted through the course Moodle site.

I request that students work individually on the problem sets and final project, because the intent is to evaluate the performance of each student at the end of each module and the end of the course. Students may refer to the following resources when working on the problem sets and final project - course textbooks, course notes, homework exercises, and SAS documentation website (documentation.sas.com).

Course Materials

Every student will be required to have a computer available for use during the course, and will be required to load the SAS® OnDemand for Academics (SAS ODA) on their computer. More details on SAS ODA can be found at https://www.sas.com/en_us/software/on-demand-for-academics.html.

The textbook for Module 1 will be *Learning SAS by Example: A Programmer's Guide, Second Edition,* by Ron Cody. 13-digit ISBN code 978-1635266597. If you wish to practice additional homework exercises from the textbook, you can install datasets from this textbook onto your personal computer. Page xxii of the textbook provides the website to download the data sets.

The textbook for Module 2 will be *Applying Data Science: Business Case Studies Using SAS*. Gerhard Svolba. 13-digit ISBN code 978-1607648895.

I recommend that you purchase or rent both textbooks (in physical or digital form) so that you have full access to the textbooks during the entire semester. A limited number of the textbooks have been placed in the OTH Regensburg library. Students are expected to be familiar with content in the textbooks for purposes of in-class meetings, homework assignments, problem sets and the final project. More details on class readings on pages 4-5 of this syllabus.

Module 1 Schedule

Notes: *Readings in this module are from the book "Learning SAS by Example"*
The day column indicates the date by which videos will be posted. I request that students watch the videos and follow-along with in-class exercises by Friday of each week.

	Topic	Day	Reading Chapter	Pages
1	**Introduce course**	Thu Mar 16	Chapter 1 What is SAS? Chapter 2 Writing your first SAS program Chapter 8 Performing iterative processing	4-7 9-18 106-112
2	**Load data I**	Mon Mar 20	Chapter 3 Reading raw data from external files Chapter 4 Creating permanent data sets	26-31 45-52
3	**Load data II**	Thu Mar 23	Chapter 5 Creating labels and formats Chapter 7 Performing conditional processing	57-64 87-90
4	**Format data I**	Mon Mar 27	Chapter 11 Working with numeric functions	161-177
5	**Format data II**	Thu Mar 30	Chapter 9 Working with dates	121-127
6	**Format data III**	Mon Apr 3	Chapter 12 Working with character functions	182-199
7	**Format data IV**	Thu Apr 13	Chapter 10 Subsetting and combining SAS data sets Chapter 7 Performing conditional processing	137-154 91-96
8	**Format data V**	Mon Apr 17	Chapter 26 Introducing structured query language	429-436
9	**Work on problem set 1**	Thu Apr 20	–	
10	**Analyze data I**	Mon Apr 24	Chapter 11 Working with numeric functions Chapter 16 Summarizing your data	164 261-276
11	**Analyze data II**	Thu Apr 27	Chapter 17 Counting frequencies	281-296
12	**Analyze data III**	Thu May 4	*Will post reading on Moodle*	
13	**Analyze data IV**	Mon May 8	Chapter 20 Creating charts and graphs	337-347
14	**Analyze data V**	Thu May 11	*Will post reading on Moodle*	
15	**Analyze data VI**	Mon May 15	*Will post reading on Moodle*	
16	**Report results I**	Mon May 22	Chapter 14 Displaying your data Chapter 18 Creating tabular reports	220-234 300-314
17	**Report results II**	Thu May 25	Chapter 15 Creating customized reports Chapter 19 Introducing output delivery system	238-255 329-333
18	**Work on problem set 2**	Thu Jun 1	–	

Module 2 Class Meeting Schedule

Notes: *Readings in this module are from the book "Applying Data Science"*
Module 2 is designed to be delivered in-person on the OTH campus.

	Topic	Day	Reading Chapter	Pages
19	**Manage an airline, such as Lufthansa**	Mon Jun 5	Chapter 6 Detecting structural changes in data Chapter 7 Detecting outliers and level shifts in data Chapter 9 Analyzing variability of data	89-103 105-118 119-128
20	**Forensic accounting for a bank, such as Commerzbank**	Mon Jun 12	Chapter 18 Checking accounting data Chapter 19 Checking for multiple accounts Chapter 20 Checking different patterns in data	266-280 281-289 291-304
21	**Forecast demand for consumer products, such as Adidas**	Thu Jun 15	Chapter 15 Performing demand forecasting Chapter 16 Using Poisson regression to forecast Chapter 17 Using similarity search to forecast	227-240 241-251 253-261
22	**Manage HR for a technical firm, such as Siemens**	Mon Jun 19	Chapter 1 Using survival analysis Chapter 2 Analyzing effect of influential factors Chapter 3 Visualizing employee retention data	3-21 23-42 59-70
23	**Perform analysis for investment bank, such as Deutsche Bank, and distribute final project**	Thu Jun 22	–	
24	**Identify data set for final project**	Mon Jun 26	–	
25	**Work on final project in class**	Thu Jun 29	–	
26	**Work on final project in class**	Mon Jul 3	–	

Due Dates for Homework Exercises, Problem Sets, and Final Project

Notes: *Homework exercises for Mar 28 - May 23 are from the book "Learning SAS by Example"*
Homework exercises for Jun 20 - Jun 27 will be based on the book "Applying Data Science"

Due date	Time	Assessment	Comments (if applicable)
Mon Mar 27	17:00	**Module 1 homework A** Chapter 3 problem 2 parts ab Chapter 8 problem 12 Chapter 8 problem 14	 – – –
Mon Apr 3	17:00	**Module 1 homework B** Chapter 3 problem 10 Chapter 5 problem 2 Chapter 7 problem 4	 – – –
Mon Apr 17	17:00	**Module 1 homework C** Chapter 9 problem 10 Chapter 9 problem 12 Chapter 11 problem 4	 – – –
Mon Apr 24	17:00	**Problem set 1**	*Will distribute problem set 1 on April 17*
Mon May 8	17:00	**Module 1 homework D** Chapter 10 problem 10 Chapter 12 problem 16 Chapter 26 problem 4	 – – –
Mon May 15	17:00	**Module 1 homework E** Chapter 16 problem 6 Chapter 17 problem 4 Chapter 17 problem 6	 – – –
Mon May 22	17:00	**Module 1 homework F** Chapter 15 problem 6 Chapter 18 problem 2 Chapter 20 problem 6	 – – –
Mon Jun 5	17:00	**Problem set 2**	*Will distribute problem set 2 on May 25*
Mon Jun 19	17:00	**Module 2 homework A**	*Will distribute homework 2A on June 12*
Mon Jun 26	17:00	**Module 2 homework B**	*Will distribute homework 2B on June 19*
Fri Jul 28	17:00	**Final project**	*Will distribute final project on June 22*