

Spring 2022 Graduate Course

OR531

**Analytics for Decision Analysis**

Tuesday: 4:30-7:10, Planetary 224

# Syllabus

Kenneth W. Comer, PhD

kcomer@gmu.edu

Department of Systems Engineering Operations Research

Volgenau School of Engineering

Office hours: Wednesday, 4:30 to 6:30 by scheduled remote teleconference

## Overview

Greetings! I’m excited to offer this course on analytics for decision analysis. This course is designed to provide a solid grounding in the theory and practice of data analytics using the most common tools available: Microsoft Excel spreadsheets.

In this class you will learn that most important analytic methods can be implemented in a spreadsheet. Spreadsheets have been criticized as too simplistic or inadequate for the data and decisions of a large enterprise. Excel becomes unwieldy with large data sets. And, even if a spreadsheet can handle the data (Excel worksheets have 1,000,000 lines by default), the process of manipulating, visualizing, and evaluating the data can become quite cumbersome. This creates a market for large-scale statistical packages (SAS, SPSS, MiniTab, S-plus, etc.) in the commercial and government sector and purpose-built open-source statistical software (R) in the academic world.

Excel has one advantage, however, over the purpose-built analytic packages: it is installed on every computer that has Microsoft Office. Thus, every laptop and desktop across the institution can load and manipulate an analytic program in Excel. And, even if an enterprise has licensed the large-scale packages, the licenses may be limited and may not be available to you as an analyst in a reasonable time frame. An answer to the decision-maker’s question is far more valuable before the decision than after it!

The most important reason to build this course around spreadsheets (and not line programs like Python, Matlab, or R) is the visibility they provide to the processes we will learn. With a spreadsheet the formulas and the results are displayed on the screen. You can watch as the output results are populated, and it’s often easy to tell when you make a mistake. Thus, while you may use one of the other tools when you need more capacity, it is important that you learn the methods and processed as well as develop the ability to interpret the results before you begin full-scale work.

The focus of this course is the *decision.* It is important to understand, even in the field of data analysis and data engineering, that the underlying purpose of advanced data analytic techniques is to support and justify a decision. Many practitioners of data analysis lose sight of this endgame. Every problem we will work will be focused on the allocation of resources or some other important decision that might be experienced in the course of business, industry, or government operations. Some examples are:

* If I’m offered additional resources, how much should I pay for them?
* What is the mix of production decisions that maximizes profit?
* What sequence should I use for a multi-step process?
* How much should I save to ensure an 80% chance of having a set sum at retirement?

I will also expect you to do more than Excel calculations. In order to receive proper credit, you will need to answer the question at hand. Your boss or your customer cannot be expected to search around your spreadsheet for the numerical answer. You can help yourself by re-reading the question before you submit your answer. Otherwise, I have no confidence that you understood the purpose for the calculation or the model.

### Issues

*Modeling and Software.* This is a course in which you will practice building a variety of models in Excel spreadsheets. I will present various data sets and ask you to use these to answer questions that are highly typical of managerial problems. In the past, answering these questions required several courses in theory and extensive mathematical treatment. Nowadays, software (when properly applied!) can provide a short cut to getting the answers.

We will also be using a student edition of an Excel-based analytic package called ***Analytic Solver Platform for Education*** *or* ***ASPE***. In the past this has been referred to by the name of the commercial products, *Analytic Solver Platform* or *Risk Solver*. The educational and commercial packages are slightly different, and you will be working with the educational trial version of the Analytic Solver package. NOTE: FOLLOW THE DIRECTIONS CAREFULLY IN INSTALLING THE SOFTWARE. I have learned from experience that getting the wrong “trial version” is a complicated error to recover from!

You should also know that Analytic Solver is not the only data manipulation package. Your company might have licensed software from other vendors (MATLAB, SAS, or SPSS are common), or the company may have developed some of these analytic capabilities in a home-grown application. The *techniques,* however, remain universal.

**WARNING**

This course is based on the Microsoft Windows version of analytic solver. There is a Mac version of the software, as well as a web-based option. This is an inadequate substitute. **The Mac version will not work.** It can’t perform all the functions that we teach in this course, it will not complete many of the later homework assignments, and it will not be able to answer exam questions. I can’t be responsible for helping you if you purchase the Mac version.

I will also let you in on a little secret: Excel contains nearly all the functionality available through Analytic Solver. Even after your license has expired, Excel has a data solver capability that was written by the same company that created Analytic Solver. It turns out that the same company that wrote Analytic Solver also provides the Solver add-in for Excel (under the data tab, once you have installed it). I will show you how to use this for most of the problems in this class.

*Textbook.* I will use *Business Analytics, The Art of Modeling with Spreadsheets, 5th Edition*, (2017) by Stephen Powell and Kenneth Baker. You will need to purchase or rent a copy of the book to get access to the homework problems. If you have access to the fourth edition, the problems and software remain the same, but the page numbers are different.

### Course Organization and Grading

Students are expected to keep up with all assigned readings in advance of classroom discussions, and to explore the example problems before the class. Only when you work through the problems, view the videos, and experiment with your hands on the keyboard, make mistakes and discover misunderstandings will you be able to master the material.

There will be eight projects assigned during the course. You will have two weeks from the assignment date to complete the project. They are due in Blackboard by midnight on the class day (Tuesday). The projects count for six points each, and the lowest grade will be dropped. An optional project will be available in the last two weeks and can be used to replace your next lowest grade.

Points for projects: 42.

There will be a *mid-term exam* (29 points) and a *final exam* (29 points). These will be proctored exams in which you must complete all of your work during the class period.

A *final exam* will be administered on **May 3**. You must attend on that day to receive credit for the exam.

*Resubmission.* I will look at any homework you have submitted ahead of the due date and correct it., I will offer some comments and try to guide you to a solution. Provided you have made at least one submission before the due date, you are welcome to resubmit the homework for a grade up to full credit. If you have not submitted your first attempt by the due date, I will comment on your homework and guide you to the right answer, but you can’t improve your grade. Late homework will lose one point a week, timed from the due date.

NO HOMEWORKS WILL BE ACCEPTED AFTER THE HOMEWORK DEADLINE. THE BLACKBOARD SUBMISSION PAGE WILL DISAPPEAR ON APRIL 26, 2022.

Grades will be awarded according to the table:

|  |  |  |
| --- | --- | --- |
| From | To | Grade |
| 100 | 100 | A+ |
| 92 | 99 | A |
| 90 | 91 | A- |
| 89 | 89 | B+ |
| 82 | 88 | B |
| 80 | 81 | B- |
| 60 | 79 | C |
| 0 | 59 | F |

Your projects and exams will be graded based upon your final output and analysis in response to the questions. I will not investigate whether you “understood the process” or “*could* have found the right answer.” This is because, in my experience in the business world, these things are worthless. Half an answer will get your company in big trouble. Nobody will examine your work to see if you had some or even most of the knowledge to solve the problem. You either build a solution or you don’t. You either deliver, or you don’t.

### Video of Course Instruction

Each problem-solving session will be available by video after the lecture. These will be cleaner versions of my in-class demonstrations. In both class sessions and on the video I will take you from the statement of the problem all the way to the solution.

### Course Topics

There are three key concept areas presented in this course:

1. **Optimization problems**. As data scientists, you are likely to become involved in conducting advanced optimizations. Companies and institutions in the past have depend on other analysts, including operations research analysts, to conduct optimization. In the modern era, however, all of the data needed to complete these problems must be acquired, shaped, and validated by big data analytics. In short, optimization decisions often is the *reason* for your data analysis.
2. **Data mining**. The mainstream of data analytics is the extraction of important patterns from a large data set. You will practice some of these “machine learning” techniques such as genetic algorithms, logistic regressions, k-nearest neighbor, and others.
3. **Data farming**. Sometimes the answer is not in the data. To help decision-makers understand and forecast future trends and the impact of their decisions it is necessary to simulate the system. When simulations produce many runs of a random system, the tools of data analytics must be applied to help inform the decision. You will do data farming using Monte Carlo and queuing theory methods.

### Communication

I plan to make use of Blackboard for this course. About 70% of Mason’s courses use Blackboard to communicate academically. Please post your assignments to Blackboard and submit the final project documentation there as well. You will also be able to view interim and final grades (yours, of course!). *Please do not e-mail me your homework projects. I can only assign grades through Blackboard.*

### Inclement Weather

The university will not reschedule a class if only one class session is missed due to snow. But, if more than one session is canceled, the university typically extends final week. This can be very disruptive: in the spring of 2014 the university was holding final exams on the day before graduation!

I plan to avoid this mess by recording a video or the missed lectures. You can review this several times at your leisure. I can cover any misunderstandings at the next regular class session. This is not ideal, but it keeps us from being forced into make-up sessions.

### Office Hour Limitations

With forty students in the class, I am not in a position to provide individual instruction. Before you schedule an office hour session, I expect that you will have made at least two attempts to solve a problem on your own and submitting it for a grade. You should re-watch the videos, and you must read the grader’s comments. I find myself in these sessions merely repeating the advice I gave the students when I graded their homework. We will open your last attempt and view it together.

### Academic Honesty and Collaboration

The integrity of the University community is affected by the individual choices made by each of us. GMU has an Honor Code with clear guidelines regarding academic integrity. ***Unless otherwise noted, all assignments are to be completed on an individual basis, with no communication or discussion with any member of the class, or anyone else, other than the instructor or the teaching assistant.***

This course presents unique challenges for students to maintain academic integrity. You must learn to set up the problem, solve it, and interpret the results on your own. I need to be certain that I am grading you on your own work, and not the work of others in the class, tutors, friends, and colleagues. The limits on collaboration apply to the eight problems assigned. This is the reason we will correct and comment on your homework attempts. If you must collaborate, in the book there are dozens of problems that require the use of the same techniques, and you are free to work those in collaboration. Likewise, a tutor can help you with problems similar to the homework. To be clear: **it is a violation of academic integrity to ask a tutor to solve the homework problem**. Both the student and the tutor will be found in violation of the Honor Code.

Our grading process is designed to help you with this. You can submit your “best guess” for the homework at any time, and I ~~or the TA~~ will guide you to the correct solution. You can resubmit many times before the deadline with no penalty.

It is also a violation of the honor code to use a homework solution you have found on the web.

As a security precaution, I never reuse my examination questions. Thus, any web-based “library” of exam questions will not help you very much, and may prove misleading. You are reminded that I reserve the option to make the final exams a proctored, camera-monitored exam during the assigned exam time. Thus, you should be using the homework problems to understand how to set up and solve these problems so that you can do so on the day of the exam.

### Privacy Restrictions

The university, in order to conform to the requirements of federal law (FERPA), now requires that all communications with respect to an academic course be conducted to and from a George Mason e-mail account. Thus, I must ask that if you send me an e-mail, you send it from your GMU account. This can be done via the web-mail application. I will observe this policy as well.

In addition, I cannot share an e-mail address with anybody else in the class. Thus, if you want me to help you make contact with anybody else, you have to give me explicit permission.

### Disability Statement

If you are a student with a disability and you need academic accommodations, please contact the Office of Disability Services (993-2474; <http://www.gmu.edu/student/drc/>). All academic accommodations must be arranged through the ODS.