Spreadsheet Modeling & Analytics
Team Project

Manoj Thulasidas
Associate Professor of Information Systems (Education)
Singapore Management University, 80 Stamford Road, Singapore 178902

I. Introduction

As the largest assessment component with 30% of the final grade, the Team Project has to be taken seriously and worked on with diligence. The detailed information provided in this document will guide you in developing an impressive project idea, delivering impactful presentations and creating a worthy model, enabling you to score well in this component.

II. Project Timeline

Week	Events
1	Know that the course has a project component
2	Look for teammates, make informal teams
3	Team signup. Start contacting potential clients
4	Decide on the client and start finalizing project topics
5	**Due: Project team formation. Random assignment**
6	**Due: Project title:** Email prof the project title and **short** description
7	**Release**: Proposal, Presentation, Reports Rubrics
8	Recess: Work on Proposal Report and Presentation
9	**Due:** Proposal Report**, Project Proposal Presentations**
10 & 11	Project Consultations
12	**Due: Final Project Presentations**
13	**Due:** Final Project Report

Figure 1: Weekly timeline for the project action items.

You are to form a project team of **five** (or a maximum of **six**) members by **Week 3**, to conceptualize and construct a spreadsheet model of an *original* problem for a *real* client. The problem must be of sufficient depth with a structure that permits a suitable model to be built and whose solution requires some analyses. The modeling and analysis aspects must use techniques learned in class and where appropriate, but may include other software tools and add-ins, if necessary.

During **Weeks 5** to **7**, please inform the instructor or meet with him to ensure that your project topic is in line with the expectation, in scope and depth.

As a group effort (earning equal marks for each person), all team members are expected to participate fully in the project. In case of difficulties, please try to resolve your problems first before approaching the instructor to mediate. Each team member when queried must be ready to individually explain any modeling aspect of their project.

III. Project Topics

Before bringing a topic for the instructor's approval, your project team should ask and answer the following questions:

- Who is the client? This has to be a real company (preferably with a real person who will eventually use the models developed during the project).
- Is Excel the appropriate tool to use for the project topic? Is it too light or too heavy for the project work?
- What commercial and quantifiable value does your project add to your client?
- Can you get good project closure? Do you have clear deliverables, which the client can work with and act on?
- Can you articulate the extensibility of the project? Can you apply the same model to other similar situations? In other words, can you treat your client not as the only one, but rather as the first of many?

A good project is not one with the most complex and fancy use of spreadsheet features. Keep in mind:

- Inappropriate use of overly complex features will not be looked upon favorably.
- Excessive and unnecessary use of VBA will be frowned upon as well.
- Consult the project grading criteria stated in the rubric below for guidance.

IV. Types of Projects

Over the years, students have come up with three main types of spreadsheet projects for this course. They are:

Operational: Business operations systems, secure access to sensitive info etc.

Decision Optimization: Data analysis, scheduling, profit optimization etc.

Exploratory Modeling: Analysis, prediction, forecasting, handling uncertainties etc.

The main focus of this course is **exploratory modeling** and therefore you should be doing that type of projects. This is also possibly the most appropriate use of spreadsheets, in exploring the business problems, organizing the requisite data, and establishing their structures for what-if analyses.

The **decision optimization** type of projects are better suited for management science, statistics or business subject courses. Such projects assume the problem question is generally known and seeks the "best" solution approach and values. Your project may include this aspect as well, but it should not be the main focus.

The **operational** type of projects are appropriate for courses in information systems, where the concepts and techniques for system design, data management, etc. are learned. The application setting is normally one with the problem question known, best solution approach and values found, but needs

the workflow, data management, and decision processing automated, to simplify human effort and shorten processing time.

Spreadsheets as simple business operations systems are often what project sponsoring companies ask for. To secure them as your client, your team may have to include aspects of operational project as well. The main part of the project should still focus on the **exploratory modeling** of the inputs, outputs, decision choices etc. of the system. From these, it should allow managers to identify key concerns, appropriate responses, and give a-priori predictions or a-posteriori analyses of the effectiveness of the recommendations. In order to illustrate the implications and to establish the robustness of the model, several analyses should be performed and documented. Some analyses ideas are listed below in **Appendix II**.

V. Project Deliverables

Project Proposal Report: Softcopy of the proposal report is to be uploaded on eLearn by **Week 9**. It should be short: ≤ 2 pages and ≤ 500 words. The report typically gives a brief background and introduces the problem, elaborating the business environment and possible uncertainties. An Influence diagram and a Black-Box model drawn to depict the variables, and suggestions for data collection and problem analysis are required.

Proposal Presentation: The short (5-minute) proposal presentation should focus on explaining the proposal, rather than on showmanship. Presentations will take place in the same order as your group number.

Final Project Presentation: A final project presentation (15 minutes) is required in class on **Week 12**. The presentation can include a short demo of the model as well (either as a live demo or as a screen recording).

Final Report and Model File: The final report ($\leq 8+2$ pages, $\leq 2,000$ words, all inclusive; shorter preferred), and the model file (Excel book) are due in **Week 13**. In addition to the report, you can also include, if appropriate, a user guide explaining how to use your spreadsheet model. All the files in the final submission have to be zipped into one archive and uploaded, as you can upload only one file. Multiple uploads are permitted, but subsequent uploads will overwrite the previous ones.

Dos and Don'ts:

- Don't submit your presentation slides.
- Don't submit multiple copies (e.g, Word *and* PDF).
- Don't include a table of content in your report.
- Do remember to provide passwords, if you have password-protected your model.
- Do provide sample data (for inputs/decisions/parameters) when submitting your final model. It can be pre-filled in your model, or in a parallel Excel file with the same structure as your model in such a way as to enable the instructor to easily copy and paste.
- Don't link your model to external files. Opening the model on a new computer should not give warnings about missing external links.
- Don't use Monte Carlo unless necessary. It makes the model sluggish, and usually does not add any value.
- Do keep the reports within the page limit and the word-count. The latter is specified so as to discourage the use of tiny fonts to circumvent the page limit specified.

VI. Assessment Rubric

	Category	Scoring Criteria	Points
For both Proposal & Final Presentation/Report	Project Scope	Client is well introduced	10
		Problem is well defined and is original	
		Solving method is logical and clear	
		Benefits and/or significance of the solution are presented	
	Presentation	Confident presentation, clear articulation and good language	10
		Well prepared, informative slides with no gimmicks	
		Supported by Influence Diagram and Blackbox mode	
		Able to answer questions	
	Report	Good language skills and grammar used	10
		Prose flow, vocabulary appropriate to hold interest	
		Aesthetically pleasing layout of text and diagrams	
		Length of the report as specified (2+1 pages for proposal, 8+2 for final report, not counting the cover page and ToC)	
		Influence Diagram is logical and is well-explained	
Only for Final Report and Model	Model File	Model works as described in the report	20
		Evidence of positive client feedback	
		The logic in the model is clear, elegant and well-defended	
		Appropriate level of analyses (sensitivity, trade-off, back-testing) to ensure the validity and robustness of the model	
		The client can easily navigate through the model, with sparingly used sheets hidden. Not unduly long	
		Well formatted with logical color scheme	
		Formulas protected against accidental overwriting	
		Suppress spurious error messages or #N/A etc.	
	Penalties	Late submission. **(Tardiness)**	-10
		Wrong kind of problem - only operational (database, inventory, simple trend/forecast, etc.) **(Project Scope)**	-10
		Virtual client, fake or assumed data. **(No Client)**	-10
		Assumptions not verified, conclusions not defended, lacking tradeoff, sensitivity or back-testing. **(Insufficient Analyses)**	-10
		Report way too long (well beyond 2+1 or 8+2 pages, with reasonable word-count) **(Verbosity)**	-10
		Extreme complexity of the model (too many sheets, too much VBA etc.) **(Complexity)**	-10
	Score	Total Points	50

Figure 2: Assessment rubrics for the project reports (both proposal and final) as well as the final model file.

The project is worth 30% of the total course grade, comprising:

5% Proposal Presentation and Report.
25% Final Presentation, Report and Model.

In grading the project, the instructor will consider:

- Problem originality, and modeling elegance and innovativeness.
- Ability to reveal problem structure and apply analysis
- Appropriate use of spreadsheet techniques and tools
- Clarity, completeness, and accuracy of the report
- The ranking of your presentation and project provided by your classmates
- Good writing skills (e.g., see Williams' "Style" or Strunk's "Elements of Style")

Appendix I: Modeling Basics

A. Steps: Model Building

- Start by drawing the *Influence Diagram*.
  - Identify variables and represent them as boxes.
  - Use arrow connectors to link boxes to show how variables influence one another.
  - Use consistent font and colors to highlight *Input Variables* (*Decision Variables* and *Parameters*) as well as *Output Variables* (*Performance Measures* and *Consequence Variables*).
- Summarize the *Influence Diagram* as a *Black-Box model*.
- Build an Excel model based on these two diagrams.

  - Ensure that every box in the *Influence Diagram* has a corresponding cell in the final spreadsheet model.
  - Cells containing *Performance Measures* and *Consequence Variables* (outputs) have formulas in them.
  - Cells for *Decision Variables* and *Parameters* (inputs) should contain only numbers, never formulas.
  - For optimization models (where the *Decision Variable* is optimized against the most desirable value for the *Performance Measure* using **Solver**), show the dependency of the *Decision Variable* on the *Performance Measure* using a **dotted arrow**.
  - Influence Diagram is *not* the same as a flow chart: The former shows relationships among various entities in your model, such that each box in it becomes a cell in the Excel model. The latter shows interactions, processes or information flows, and is *not* mandatory to include in the report.
- Identify and validate the assumptions.
- Analyze Trade-offs by changing the values of key controllable inputs (*Decision Variables*) to derive recommendations.
- Perform Sensitivity Analysis by changing the values of uncontrollable parameters or assumptions.
- Try other appropriate analysis ideas from **Appendix II**.
- Evaluate, based on your analyses, whether recommendations are still valid and add contingency measures as needed.

B. Suggestions: Excel Model Layout and Format

- Remember that the spreadsheet model you build may not be so easily understood by others and possibly not even by yourself later on unless you put in some effort to make it user-friendly.
  - Present data together (grouped by type) and clearly label them.
  - Place the input cells (*Decision Variables*) near the top of the sheet.
  - Position output cells (*Performance Measures*) also near the top, next to the inputs.
  - Provide information or user guide to help the user.
  - Place working tables near the bottom so as not to crowd out the inputs and outputs. This placement also allows the table to be easily grow downward as needed.
  - Add some documentation to show formulas of key cells. While `Ctrl ‘` reveals all formulas, having the formulas in key cells documented will help you and others better understand the model and its logic.
- All input data should be entered only once. Multiple entry of data can lead to errors.

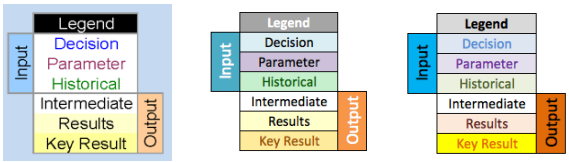

Figure 3: Our preferred color schemes.

- Store all data values in separate cells or sheets and refer to them in formulas by cell references. Avoid hard-coding of numbers. State units of measure, where appropriate.
- Use a consistent format and color scheme (Figure 3) to simplify model use. You can choose your own scheme, as long as you are consistent.
- Use Data Validation on inputs and protect formulas or parameters against accidental or intentional modifications:
  - Add `Data` → `Data Tools` → `Data Validation` to ensure only appropriate data can be entered.
  - Keep only the input cells unlocked: click Format in the Home tab and uncheck Locked Cell.
  - Protect Sheet to activate protection before releasing the model (Excel file) to others.
- Make the model less intimidating to the end user:
  - If your model has a lot of sheets within, provide a homepage and navigation bars.
  - Use VBA to hide the sheets that are not relevant to the current task.
  - Provide dashboard or illustrative charts. Consider providing a User Guide.
- Ensure Printing Support, if applicable: The choice of colors used must be such that printing or photocopying the sheet in black and white does not make it less legible.
  - Try to fit the complete model nicely within the computer screen, if possible, with minimal scrolling.
  - In the Page Layout tab, click Print Area/Set Print Area to specify the part of sheet to be printed.
  - Click Print Preview to see if sheet fits nicely on standard paper and adjust accordingly.
- Your Excel book in another computer may not perform as predicted if the application settings there are not the same as yours. Take the necessary actions to rectify.

Appendix II: Analysis Ideas

In addition to the Trade-off and Sensitivity analyses that you learned in class, here are a couple of more ideas.

A. Sensitivity and Trade-off Analyses

To recap, Sensitivity Analysis refers to how sensitive your output result is with respect to the values given to the input parameters or your assumptions. For example, your system may determine the optimal quantity to order, which is based on the holding cost and ordering cost. You may have assumed some values for these two costs. Do a sensitivity analysis of how the optimal quantity to order will change according to different possible values of holding cost and ordering cost. This result is best shown using a data table. Trade-off analysis is when you change the Decision Variable and study the Performance Measure. Example: Changing the price markup to optimize the profit in a retail business.

B. Back Testing

If your project contains any predictive element (e.g., revenue forecast based on historic data), you may be able to run back testing on it. For instance, if you have the data for the past 24 months, and you are making a prediction for the 25th month, you can use the first part of the data (23 months) and predict the (already known) value for the 24th month and compare it with the known value. If you have multiple methods for predicting, you may argue that the one that performs best in

back testing is the right one. Or, you may present the difference between the predicted and the realized values as a measure of the sensitivity.

C. Stress Testing

If you know the possible ranges of your input (decision) variables, you may want to test your model for the extreme values in the range, or even beyond in order to make sure that your model behaves reasonably.

D. Service Levels

If your model makes a computation based on a prediction (e.g., inventory level needed to satisfy the predicted demand), you may want to estimate the standard error on the prediction. Such error estimate will enable you to compute values for specific service level (e.g., the inventory level needed satisfy 90% of the customers for a given predicted demand).

In order to understand this better, consider the example of inventory levels based on predicted demands. Let's say you predict the demand for next month as d. It means 50% of the time, the realized value is going to be more than d, and 50% of the time less. If you also know the standard error in the prediction (σ_d), you can compute the inventory level needed to satisfy 90% of the customers on an average (or, in other words, to ensure a service level of 90%). This demand is given by $d_{90} = \texttt{NORM.INV}(90\%, d, \sigma_d)$, which depends on σ_d. If youre prediction is based on a trendline, you can see that σ_d is related to the coefficient of determination R^2: it is small when $R^2 \approx 1$ and large when $R^2 \approx 0$. Computing σ_d, however, is not straightforward. Here's a recipe: Suppose you are predicting the $n + 1^{\text{st}}$ value of d based on the previous n values, which have a standard deviation of σ. Then, for our purposes, the error σ_d, can be approximated as:

$$\sigma_d = \sigma\sqrt{\frac{1 - R^2}{n}} \tag{1}$$

Note that R^2 is easily computed in Excel using the function `RSQ()` for linear trends, much like the `SLOPE()` and `INTERCEPT()`.

E. Seasonality Effects

In most predictions (such as in the food and beverages industry), you may have a clear seasonality effect in the revenue or profit. If your data is not deep enough to compute the seasonality effects, try getting proxy data from similar businesses or industry-wide benchmarks. Once a seasonality effect is established or assumed, you may want to aggregate in order to improve the statistical power of your predictions. For example, if you are predicting the revenue of a restaurant on a monthly basis after recognizing a monthly seasonality effect, you can make a yearly prediction based on the historical data, and divide it up into months using the seasonality effect.

APPENDIX III: REPORT CONTENTS

The proposal and final reports should typically cover the following, in a management-directed format. Note that the bullet points listed below are for your reference as guidelines only. If a few of the points do not apply to your project, you need not include them. Ensure that your report is comprehensive, succinct, clear, convincing and appealing to read.

Client & Problem: Describe the business and the details of the problems you are planning to solver for them. List the objectives of the project. These objectives will translate to the features of the the system you will build for them.

System Scope & Features: Describe the areas covered by your proposed system and for each objective, describe the features planned. Do not write it like a user manual, but list the features as short description in point form to address the project objectives, highlighting the benefits to the client.

Performance Measures & Their Evaluations: *Performance Measures* are the main results of your system. They are also in your Influence Diagram and Black-Box model. Describe how your system evaluates/computes them. For example, if the performance measure is profit, then describe how costs and revenue are computed to arrive at the profit.

Data & Assumptions: Describe how the input data is collected. The data could be historical data from the client's point of sale or inventory, or accounting systems. Describe if there are any pre-processing steps before the data can be used by your system. In the absence of data, you may be forced to make assumptions, or use client's intuitions. You may also use industry-wide averages from other sources, for example, government statistics or other websites.

Decisions & Alternatives: Your project may help to make decisions. Describe what kinds of decisions are made and how. For example, the system may help to determine the optimal quantity to order for a certain inventory item. Describe how this optimal quantity is determined and how your system prompts the user to place order when the inventory item is low.

Model Sketches: You can have more than one set of Influence Diagram and Black-Box Model. For projects with more than one model/feature/sub-problem, each sub-problem should be supported by an Influence Diagram and a Black-Box Model. If appropriate, you may include a flow chart to show either the user interaction with your system or the data flow within the model. Output from one model can be the input to the second module.

Trade-offs, Scenarios & Sensitivity Analyses: The main results of your system could be determined by trading off different costs. For example, the optimal quantity to order is a trade-off between holding cost and ordering cost. Scenarios refer to the different possible results can be generated by your system. For example, your system may generate the triggers such as *invest* or *don't invest*.

Results & Implications: After analysis, your system will generate analysis results which are either in the form of tables or charts. Include them and explain the implications of these analysis results.

Model Limitations, Lessons Learnt & Conclusions: Going beyond the results, include some descriptions about what you learned from the project experience, how it can be improved and its larger implications. The larger implications (which are answers to the question, *"So what?"*) may form the basis of your conclusions.

Members' Roles, Client Feedback & Contacts: If appropriate, include a short description of each member's role in the project, as well as any feedback from the client, along with their contact details.