Increasing Real Estate Management Profits: Harnessing Data Analytics
by Daniel Egger and Jana Schaich-Borg

Best Practices for Setting up an Excel Spreadsheet using the Provided Templates

You can do all of the assignments for weeks three and four of this Capstone by following the instructions and entering your formulae and data into the appropriate Excel Template Spreadsheet. There are a total of six Template Spreadsheets. The first three are needed in week three:

1-First Best-Fit Line
2-Normalized Data and Model
3-Solver Rent Optimization

The second three will not be needed until week four:

4-Alternative to Solver
5-Forecasting Cash Flow + Profits
6-Sorting by Profitability

These Excel spreadsheet templates employ the "best practices for Excel" discussed below. We strongly recommended that you use them as structured to complete the Excel portion of the Capstone. You will have the opportunity to test your results, and verify that they are correct at each stage, before going on to the next stage.

Best Practices for Excel

➔When analyzing many different items, use only one row for all the information about each individual item.

➔Start with the most basic information in the left-most column. In Template 1, list the unique Watershed property IDs. Don't erase anything. For each new calculation, add a new column immediately to the right of the previous column.

➔In the spreadsheets, each one of the 244 relevant Watershed properties are to be assigned to one row. All further data and calculations related to that property are to be put in new columns in the same row – including all relevant short-term rental price and occupancy information for properties of the same type in the same location.

➔Within the spreadsheet, distinguish data imported into the spreadsheet unmodified, from interim and final results calculated using formulas. Color-coding the three types of numbers is recommended. In the spreadsheets, the following color scheme is used: unmodified database data are marked in blue, interim results are in orange, and final results are in green.

➔For ease of reading, limit digits displayed after decimal points to no more than 2, and add appropriate formatting to cells containing numbers that represent currency or percentages.

Note that when displaying numbers "rounded" in Excel, Excel continues to preserve the full original values, so no information is lost.

➔Use a small – and fixed - number of rows at the top of spreadsheets to document and explain exactly what each column contains. In the spreadsheets, the third row from the top gives a unique descriptive name of that column, the second row is used to explain the algorithm or Excel formula used to generate the numbers in that column, and any additional numbers needed as formula inputs are given in the top row.

➔As interim results build up, original database values may start to be "far away." When you need to re-use unmodified data or interim results from a column that is now many places to the left of the current calculation, simply copy the column again. This way most calculations involve columns that are visible immediately to the left.

➔When explaining your arithmetic or other Excel formulas (in the second row), refer to earlier columns by the names you have given them in the third row, *not* by their column letter. If you later change the column ordering the letter designations will be useless, but the names remain correct.

➔If you wish to generate a scatter plot – as you will do here to determine its best-fit line - put the y axis data immediately to the right of the column containing the x axis data. This way each ordered pair (x, y) is visible and easy to read. You can simply select both rows together and chose Chart/Scatter in Excel – Excel automatically assumes that data to the left of two columns goes on the x axis.

➔Divide headings from data using horizontal lines, or put boxes around the headings. When a certain stage of the analysis is complete, indicate so by putting a vertical line to the right of the final column.

➔When faced with a complicated formula where order of operations or other typographical errors are likely, break the formula up into several interim steps. Each column then becomes like a part of a formula in parenthesis, and an additional column is used to combine them.

➔When you have verified that certain steps of a complex analysis process are working correctly, copy the entire (correct) spreadsheet, rename the new spreadsheet, and add any additional columns only to the new spreadsheet. This way, if you find later that you have made a mistake, there are relatively few steps to check and debug, and if the mistake is really hard to find, you can always back to the last version that was fully correct, and start again from that point. *Note that the six Template spreadsheets are built up in exactly this way – each contains its complete predecessor*.