**Possible Errata and various comments on book or Excel files**

**for *Marketing Analytics* (1st Ed, 2nd printing) by Wayne L. Winston**

William V. Harper (Summer 2015), wharper@otterbein.edu (no guarantees of my accuracy implied)

These quick notes may have my own errors or misunderstandings. This is not a review of the full book. It is not even a line by line review of those sections I am covering in the course but just those items that seemed to pop out as unusual that I had time to dig into. I have found myself in an awkward yet exciting situation in which a course I have taught for many years in the Otterbein University’s MBA program is transitioning (with not a lot of notice) from an operations research class to a data analytics class. My background is a good match but I am not up to speed on all aspects of data analytics especially on the business end being an engineer/statistician.

I am only looking at the subset of the book I might cover this fall of 2015. I have not gone through all the sections yet I hope to cover; therefore, this is a work in progress. I do believe the book is a good book but being a first edition it has errata as likely an edition of any book.

WARNING on Excel files: I am working with the version I downloaded in late May 2015; therefore, any comments I have on the files may have since been fixed and no longer relevant.

Feel free to contact me if desired: Bill Harper, wharper@otterbein.edu. Keep in mind there are full sections of the book I am not touching as well as various sections here and there. I do not have my syllabus together yet.

I am sharing my comments with the author earlier but have added some new ones (they are dated July 21 2015 or later) and posted on my college webpage <http://faculty.otterbein.edu/wharper/> under [Some potential errata on the fine book Marketing Analytics by Wayne Winston, 1st ed, 2nd printing](http://faculty.otterbein.edu/wharper/PossibleErrataInBookOrExcelFilesForMarketingAnalytics_v12.docx). They will also be put in Blackboard for the students in my fall 2015 MBA 6600 class. The initial set of comments were emailed to Dr. Winston in June 2015; therefore, much of what follows may have been addressed in subsequent printings or a new edition.

 **Possible Errata and general comments**

Section 1: Chapter 3

Page 60, item 2 at bottom of page: book is correct but an error in cell S28:S31 with “=TRANSPOSE(N27:Q270)”, i.e., should have Q27 instead of Q270.

Section II Pricing entirely skipped by me

Section III: Chapter 9

Page 173, middle of page, item 1, end of 2nd line should have E4 instead of E5. Item 2 states that the selected area includes the labels of the variables; therefore, E4 (not E5) is the starting point of the selection.

Chapter 10

Page 190, paragraph starting “Car sales are highest”, 2nd line of paragraph should say fourth quarter (October through December) instead of “third quarter”.

Page 192, 2nd line has a superscript 1 instead of a subscript on X. Should be X1, not X1.

Chapter 11

Page 219, 2nd line should be K53:K57, not K52:K57.

Page 230, penultimate (next to last) bullet at bottom dealing with the upward trend should be 0.059 billion miles, not 59 billion.

Section IV: Chapter 16 Conjoint Analysis

Page 267, Figure 16-2. Maybe terminology or my ignorance (likely to be honest) but this design appears to me to be nearly but not fully orthogonal. When I worked with it I got non-zero cross-products with the last two terms as does the author with the value 1 shown in the Guarantee/Approval cell Q17. I am thinking fully orthogonal implies the resulting cross-product matrix would be a diagonal matrix with zeroes on all non-diagonal elements. Not a biggie, but thought I would share.

Page 270 & vicinity: In the file conjoint.xls (& perhaps others) there is a term alpha that is critical in the final solution but (as far as I can see) not covered in the text. For example you will see in row 1 of the last two tabs (worksheets Segments, GuaranteeApp), the same following label and value in cells L1, M1.



Alpha is used in the solution (M1 is a named cell) to cells O6:Q19 in the following type of equation (this one is from cell O6: =L6^Alpha/($L6^Alpha+$M6^Alpha+$N6^Alpha). Alpha is the cell changed my Solver as shown below to minimize the SSE (sum of squared errors). Maybe I need to re-read the material again but I did not notice written material on the role of alpha in the results shown in cells O6:Q19.

 

Excuse me for any oversights but I am rushing this summer of 2015 to use this fine book for an MBA class that used to be an OR (operations research) to now a data analytics course this fall with minimal notice. I am also constraining myself to do as much as possible in Excel since our students will not have general purpose stat packages at their work locations.

Chapter 17 Logistic Regression

Page 301, problem 8 has the incorrect temperature for the Challenger disaster with the O-rings. The temperature should be 31, not 36. It is correct in the Excel file.

Page 289 (just found this one July 21 2015): Below has two errors in equation (4). First this is an equation for 1-p and not p. Second the minus sign is missing after the e in the numerator.



Chapter 18 Discrete Choice Analysis

Page 312, 1st bullet: $180 is Xbox cost, not price. 2nd bullet is correct with price for other two competing game devices.



Page 315, Figure 18-9: image mismatch between Excel file and text. Not a biggie



Have not gotten any farther in book than this yet. In general I plan to cover aspects of sections (each has multiple chapters) I, III, IV, at least part of V, VI, X. There are some parts in these chapters I will skip – especially those that require other software such as things from Palisade that I know our IT dept would not deal with (think the free downloads are limited to 15 days anyway per the book). Palisade has lots of good things but my goal is to try to stay within Excel as much as I can even though some things can be done easier & better with other software including my beloved Minitab.

* - - -

**New errata after above emailed to author on 6/20/2015.**

Section V: Chapter 19 Calculating Lifetime Customer Value

Page 329, Figure 19-1, what is cell D17 doing? Under the value of 1834.14 is the formula =23.82\*77 but it is not obvious why this is here.

Page 329, middle of page, 2nd & 3rd bullets: remove “)” at end of line.

Page 329, last line error in formula: exponent of 1.10 should be .5, not 5. Excel file is correct.

Page 330, penultimate line in 1st para under new section: “two inputs (retention n rate…)” I think the n should not be here.

Page 332: near bottom of page, “Year n margin” equation is really end of year margin (called “end margin” in the Excel file). “begin margin” in Excel uses (year – 1) instead of year as part of the exponent to exp (e). What is labeled “margin” in Excel is the average of the end and begin margins (called “mid-point margin” in text at top of page 333. This is all given at the end of page 332 and start of page 333; however, the margin for year n depends on whether you are interested in the beginning, middle, or end of the year. Not a biggie but just worth clarifying since the way the text reads prior to giving the Excel implementation is that this formula is the year n margin while it is actually the end of year n margin.

Page 335 near top uses N and n for the number of purchases. Better to be consistent with either N or n.

Chapter 20 Using Customer Value to Value a Business

Pages 341-342: There are a fair number of issues with the Iron Pit problem. The Excel file as downloaded in May 2015 has errors; the text has errors; the text partially follows an updated corrected Excel file (shown in Figure 20-2) but sometimes uses incorrect results from the older incorrect Excel file that I downloaded from the book’s web page.

Pages 341-342: Excel file provided from web has incorrect formula for row 19. Excel provided does not have the 0.5 multiplier as shown in step 6 at the top of page 342. Once this is realized then one can see that the correct formulation was used to get Figure 20-2 on page 341.

Page 342, item 7 is incorrect in text but ok in Excel. One should not multiply the annual fixed cost by 0.5.

FYI page 342 item 8 that also has a 0.5 multiplier is correct in both text and Excel file provided.

Page 342, item 14 should state E27 rather than E28. The Excel file and Figure 10-2 show this in cell E27.

Page 342, items 16 & 17 grab the incorrect much larger answers from the older Excel version that I downloaded. The text lists the older incorrect values.

Old incorrect: ; new corrected: 

The above new corrected was corrected by me by multiplying row 19 values by 0.5 (as given in text). Note the corrected values match an obviously fixed (but not put on web) Excel file – see in the text the above values correct in cells F29:F30.

Page 343, last line and 1st line of page 344 incorrectly pull numbers out of the correct Figure 20-3 sensitivity analysis. As can be seen below going from a retention rate of 70% to 90% does not double the NPV as the text state, but almost quadruples the NPV. The IRR does increase from 17% but not to 33% for the retention of 90% - it increases to 30%.

|  |  |  |
| --- | --- | --- |
|  | **NPV** | **IRR** |
|  | **$937,113.25** | **23.14%** |
| **0.7** | **$423,919.73** | **17%** |
| **0.75** | **$660,397.29** | **20%** |
| **0.8** | **$937,113.25** | **23%** |
| **0.85** | **$1,262,399.22** | **26%** |
| **0.9** | **$1,646,340.76** | **30%** |
| **0.95** | **$2,101,112.89** | **33%** |

Hope to find time for other customer value chapters. Interesting material, not something I usually work with. But need to move along to other parts of the book I know I should cover. I wonder how this book is used by the author – is it a 2 semester class or what subset of the book is generally covered? Do try to find time to read and learn from chs 21-22.

**Material below is after the above was sent to Wayne on 6/22/2015**

Section VI, ch 23 Cluster Analysis

Page 382, Figure 23-5, Excel cells I9:I14 may warrant some explanation as students might not realize how these cells determine which column of the lookup table provide the return values from equations such as “=VLOOKUP($H5,Lookup,I$3)”.

Page 384, Figure 23-7 shows incorrect Solving method: web file and text below this figure correctly state Evolutionary as shown below.



Page 389, Figure 23-11: not a biggie but students may wonder why the full regression is not shown with the intercept. Maybe state it is not used in the subsequent steps.

Page 391 Note paragraph at top: good info on Evolutionary mode of Solver that perhaps could also be mentioned earlier in the book when used. Frontline web pages (they are the Solver vendor) has some good write-ups as well.

Page 391 Summary 5th bullet about increasing the Mutation Rate to 0.5 – was this mentioned before? If not it might be worthwhile to do so as a summary is typically reiterating highlights of what has already been covered.

Chapter 24 skipped for now.

Did glance at front end of chapter & I can see there is a lot of good material that I need to find time to read & learn from. Wow, for a 66 year old to be excited about so much new (both the subject & Excel) material is wonderful.

Chapter 25, Classification Trees

Page 405, equation (1) should have a negative sign in front of the summation. In paragraph immediately below this equation there is a minus sign missing on the 2nd term for a pure node.

Page 409, 410 section “Pruning Trees and CART” is sparse and not very useful. What the heck does one do with the rough 40% supposedly misclassified? This really seems like a pessimistic estimate. CART and related decision tree methods are likely best handled with products (as the author lists) like SPSS. But the front end (bulk) of the chapter is good stuff and some new jargon to moi.

Sections VII, VIII, IX are being skipped.

Section X

Chapter 37, Principal Component Analysis

Page 543, penultimate line “sign of the best-fitting line” should change to “sign of the slope of the best-fitting line” to be more clear and accurate as the simple y = a + b\*x straight-line model has two parameters (intercept a and slope b as shown here).

Pages 543-544: Errata/comment lists can appear to be so negative and that is not my intent on this book I am loving and learning from. One minor bit of praise on a topic dear to my heart (among many praises that can be bestowed on the book) is the author’s word descriptions of covariance and correlation: the use of unit-dependent (covariance) and unit-free (correlation) is one item. A second is that for both he clearly states that there are for linear association. I have battled correlation in particular being used as a screening tool (& I do it also) without considering such things as a perfect quadratic relationship between two variables might have a near zero linear correlation. Thank you!

Page 544 Equation 3 is obviously not completed for the sample correlation coefficient r between x and y.

Page 544, last line and equation (4) top line of page 545 cosmetic item: Suggest consistency with case for zi. I suggest both lower case since they are sample z values based on the observed data.

Page 545, near bottom, first matrix shown is not 2x3. It is 3x2 with 3 rows, 2 columns.

Page 547 somewhere on page: would not hurt to remind students we are back to Varcov.xlsx after the diversion to Matrixmult.xlsx of prior 1.5 pages. Minor issue.

Page 547 toward bottom with 2FE – HP. Students may not realize that this is the direct result of using C1. Since this vector is (2, -1, 0) it is used to get 2FE -1\*HP + 0\*PR.

Page 547 near bottom cosmetic issue: suggest increasing font size of first linear combination transpose in line immediately below equation (5) to match size shown in equation (5).

Page 549: when opening clusterfactors.xlxs, I got the Excel message (shown below) about it could not update some links – not sure if this created any problems in the material in this file. I choose Don’t Update.



Page 550, 1st line under Figure 37-6, “partial correlation matrix” might be changed to some other verbiage since partial correlation is a specific type of correlation and not what is being referred to here. Perhaps “lower half only correlation matrix” or “incomplete correlation matrix”.

Page 550, line above Figure 37-7, change B2:G7 to A1:G7.

Page 551: this may be my lack of knowledge on the GRG Nonlinear option, but the text in item 5 says use the GRG Multistart Engine. It was not checked for the Excel file provided. Also from the FrontlineSolvers <http://www.solver.com/excel-solver-change-options-grg-nonlinear-solving-method> it says (as shown below) that bounds are recommended but not required. For this problem it makes sense to limit the eigenvectors values to the [-1, 1] range as was done.



Chapter 38, Multidimensional Scaling (MDS)

Page 566: Diagonal Elements for breakfast set to 0; set to large value in prior city example. Need to have students think about why breakfast diagonals are as small as possible and very large for city example. Look at what is being optimized for a start.

Page 572: minor point, but it might be cleaner to say in points 3 and 4 that these are absolute differences and absolute deviations, respectively.

Chapter 39, Classiﬁcation Algorithms

Page 588, 2nd sentence under list of 6 items, “…Evolutionary Solver weights are needed” – change weights to constraints. For most people this would not be of importance but I wanted to mention it.

Chapter 40, one-way ANOVA

Page 596: last sentence prior to Example that says “you can conclude that the Null Hypothesis is true”. For likely the audience that may not be heavily statistical I can see this makes the explanation easier. And I know the difficulties of explaining hypothesis testing. But if nothing else maybe a footnote saying something along the lines that we do not show the null hypothesis Ho to be true. Instead we conclude at some confidence level that we either have sufficient evidence to reject Ho in favor of the alternative hypothesis Ha or we do not have sufficient evidence. I know this is confusing to those just getting into statistics.

Page 600, last line of 1st full paragraph that says “156”: not an error but it would be good (as it confused me at 1st) to remind readers that the input data is in hundreds.

Page 601, contrasts: Somewhere in this chapter it would be good to let the reader know that multiple Pacomparisons (e.g., Tukey) are commonly used in statistics packages. Contrasts are good, but do not always answer all the questions one may have.

Page 601, 2nd line of 2nd para under Contrasts: suggest making g a subscript to the mean mu (µg).

Page 601, 3rd line of 2nd para under Contrasts: reword to say the cis add to 0.

Chapter 41, two-way ANOVA

Page 610, first character of 3rd line below equation 1 should be deleted, i.e., “|”

Page 612, third line from bottom remarking about the large p-value of 0.79. It would be good to mention that p-values are probabilities and are bounded on the interval 0, 1. Not worth defining in detail for this material. But knowing that p-value has such bounds then helps one understand that 0.79 is a large p-value.

Page 613, bottom line: great on mentioning the parallelism.

Page 615: something is wrong, I need to look at it more. Near bottom was 35.333 is.

**Above sent to Wayne July 7, 2015**

Section XI was skipped.

**New notes subsequent to above (too tired to check them out now)**

Page 40, item 1: Need to check, but I think there is a typo and D1:D3 perhaps should be C1:D3.

Page 239, problem 1 may have a typo. Need to look into it. I have a note that says the data in Excel ends a year after the date given in the problem.