

Lecture Note #10: Data and Tables

BUSI 201: Business Data Analysis

Fall 2023

Topic 1. Conditional Formatting

Conditional formatting can be a powerful tool that can be used in initial rounds of data analysis, and also to produce a table that would be easier to read. Conditional formatting allows the user to alter the formatting of certain cells in the spreadsheet if certain conditions are met.

The range of conditions that the user can set is quite flexible, and the formatting options are similarly flexible. For instance, the user can change the background color of cells that include values that are among the top 10% of cells in a table. Or the user can change the font and font color of cells that have values that are less than or equal to the sample average.

This function can be accessed by navigating to **Home** > **Conditional Formatting** as displayed in Figure 1, and has various types of rules that we can apply. We will be covering four different types of conditional formatting options in this lecture: Highlighting Cell Rules, Top/Bottom Rules, Data Bars, and Icon Sets.

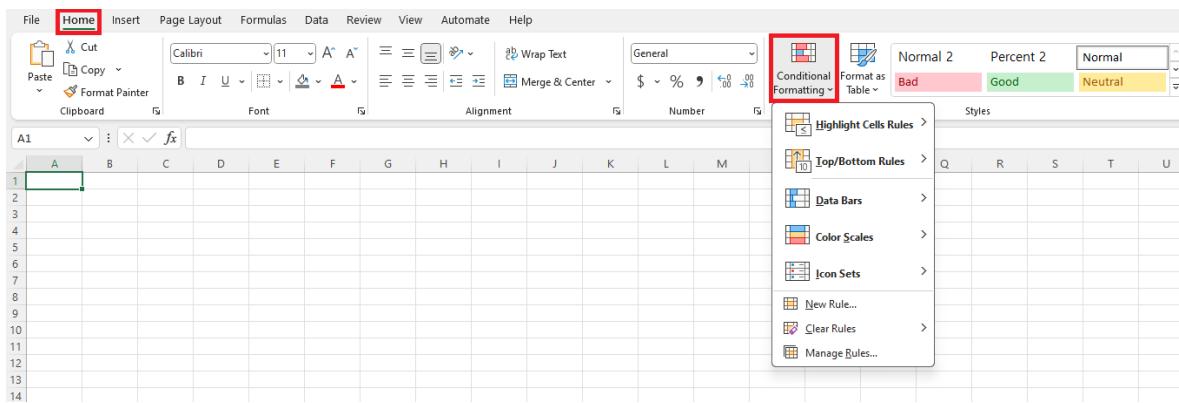


Figure 1: Conditional Formatting

Topic 2. Conditional Formatting: Greater Than...

Navigate to worksheet LAW in the workbook BUSI201-LEC10-Workbook.xlsx. This worksheet contains synthetic data on 26 applications received from a Law School. The data (B4:G30) includes the variables Applicant ID, LSAT, GPA, References, Statements, Experiences. Suppose that you are the administrative staff responsible of organizing this data according to the rules (I2:L30).

A	B	C	D	E	F	G	H	I	J	K	L											
1																						
LAW SCHOOL APPLICATION EVALUATIONS																						
EVALUATION RULES																						
4	Applicant ID	LSAT	GPA	References	Statements	Experiences	From	To	Verdict	Format	LSAT											
5	CL-20230001	146	3.88	8.3	3.4	6.9	170	180	Excellent	Score	From											
6	CL-20230002	180	3.22	7.8	3.3	7.2	155	169	Satisfactory	Score	To											
7	CL-20230003	152	3.47	1.8	1.3	9.1	145	154	Poor	Score	Verdict											
8	CL-20230004	154	3.4	9.5	7.4	9.6	0	144	Rejection	Score	Format											
9	CL-20230005	149	3.61	8.5	1.6	2.3	GPA															
10	CL-20230006	175	3.17	6.2	6.1	9.6	3.7	4.3	Excellent	Score	References											
11	CL-20230007	180	2.28	3.8	4	6.9	3.3	3.69	Satisfactory	Score	Criteria	Verdict	Format									
12	CL-20230008	176	4.25	7.2	1.7	9.3	3	3.29	Poor	Score	Top 25%	Excellent	Score									
13	CL-20230009	158	3.64	9.1	2.1	4.2	0	2.99	Rejection	Score	Top 50%	Satisfactory	Score									
14	CL-20230010	162	3.74	4	9.6	8.3	Statements															
15	CL-20230011	176	4.12	7.1	2.9	2.1	Top 75%	Top 20%	Excellent	Score	Criteria	Verdict	Format									
16	CL-20230012	167	4.14	9.8	7.4	10	Bottom 25%	Top 50%	Satisfactory	Score	Top 50%	Poor	Score									
17	CL-20230013	174	3.65	7.4	4.1	8.7	Top 80%	Top 75%	Rejection	Score	Top 80%	Rejection	Score									
18	CL-20230014	159	3.17	2.8	2	8.1	Bottom 20%	Bottom 25%	Excellent	Score	Bottom 20%	Rejection	Score									
19	CL-20230015	168	4.07	5.1	9.1	9	Score															
20	CL-20230016	147	3.18	1.2	8.7	8.2	Score															
21	CL-20230017	149	3.24	1.7	7.9	1.9	Score															
22	CL-20230018	162	3.3	6.8	3	2.6	Score															
23	CL-20230019	143	3.95	8.7	6.3	9.2	Score															
24	CL-20230020	138	3.24	1.7	9.8	5.1	Score															
25	CL-20230021	152	4.25	6	2.6	8.3	Score															
26	CL-20230022	159	3.45	9	9.9	5.8	Score															
27	CL-20230023	152	3.57	9.5	8.8	1.5	Score															
28	CL-20230024	176	4.1	3.7	7.3	2	Score															
29	CL-20230025	153	3.53	3	5.9	6.5	Score															
30	CL-20230026	128	3.83	4.3	2.4	9.2	Score															
31																						

Figure 2: The LAW Sheet

Formatting: Understanding LSAT Score Rules

The rules for the LSAT scores are given as the first rule in I2:L30. If the applicant has an LSAT score greater than or equal to 170, they are classified as “Excellent,” and the cells that stores their LSAT scores should have a green background, the font color should be white, and the text itself should be bolded.

Similarly, LSAT scores between 155 and 169 are classified as “Satisfactory,” and the cells should have a light green background. Those between 145 and 154 are classified as “Poor,” and have no special formatting. If the LSAT scores are lower than or equal to 144, this will trigger an automatic rejection, and the cells containing these scores should have a red background and white text color.

Formatting: Applying LSAT Score Rules

We will start off with the first conditional formatting task of marking the scores that are classified as “Excellent.”

1. Start off by selecting the portion of the data that stores LSAT scores (excluding the row that contains the variable name) as displayed by the red box in Figure 3.
2. Then head to **Home** > **Conditional Formatting** and select **Highlight Cells Rules** then **More Rules** to call up the New Formatting Rule window.
3. In the **New Formatting Rule** window, change the “greater than” to “greater than or equal to” in the green box.
4. In the purple box, choose I5, which contains the value 170 in the evaluation criteria table.
5. Then click on the **Format...** button... (continued on the next page)

The screenshot shows a Microsoft Excel spreadsheet titled "LAW SCHOOL APPLICATION EVALUATIONS". The data includes columns for Applicant ID, LSAT, GPA, References, Statements, Experiences, and Evaluation. A red box highlights the LSAT column from row 5 to 30. The Conditional Formatting ribbon tab is selected, and the "Highlight Cells Rules" option is chosen. The "More Rules..." button is highlighted with a blue box. The "Format only cells that contain" dropdown is open, showing options like "greater than", "between", etc., with "greater than or equal to" highlighted with a green box. The "Format..." button is also highlighted with a blue box. The "Format" dialog box is visible, showing color swatches for "Excellent" (green), "Satisfactory" (light green), "Poor" (yellow), and "Rejection" (red). The "Format" tab is selected in the dialog box.

Figure 3: Conditional Formatting: Excellent LSAT Part 1

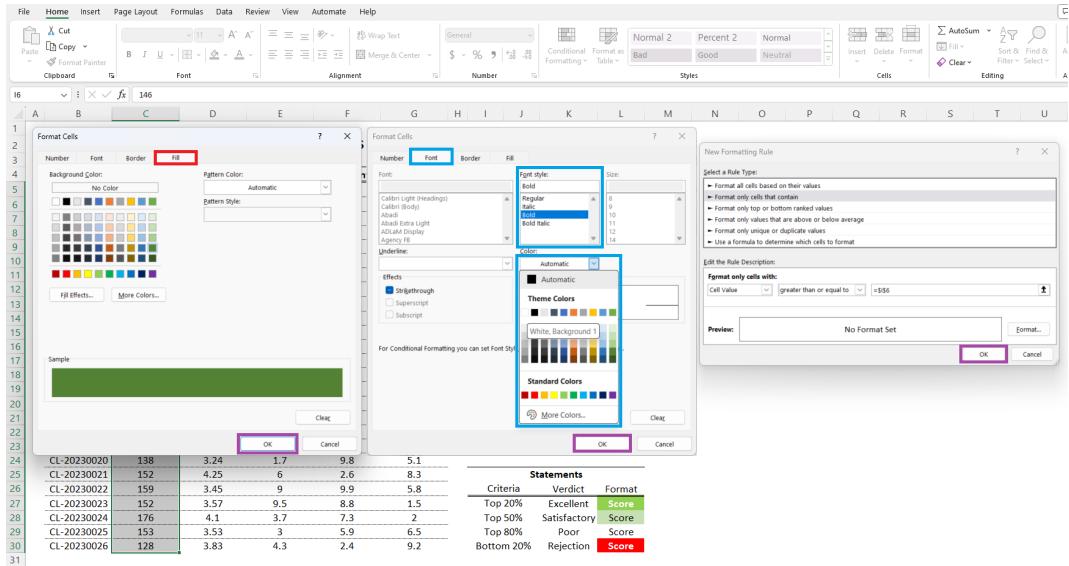


Figure 4: Conditional Formatting: Excellent LSAT Part 2

6. Under **Fill**, select a dark green background color for the cells with 170+ LSAT scores.
7. Under **Font**, select the bold text style, and white text color.
8. Click on **OK** to check the final result.

LAW SCHOOL APPLICATION EVALUATIONS						EVALUATION RULES			
Applicant ID	LSAT	GPA	References	Statements	Experiences				
CL-20230001	146	3.88	8.3	3.4	6.9				
CL-20230002	152	4.25	6	2.6	8.3				
CL-20230003	159	3.45	0	0.0	5.8				
CL-20230004	154	3.4	9.5	7.4	9.6				
CL-20230005	149	3.61	8.5	1.6	2.3				
CL-20230006	175	3.17	6.2	6.1	9.6				
CL-20230007	180	2.28	3.8	4	6.9				
CL-20230008	176	4.25	7.2	1.7	9.3				
CL-20230009	158	3.64	9.1	2.1	4.2				
CL-20230010	162	3.74	4	9.6	8.3				
CL-20230011	176	4.12	7.1	2.9	2.1				
CL-20230012	167	4.14	9.8	7.4	10				
CL-20230013	174	3.65	7.4	4.1	8.7				
CL-20230014	159	3.17	2.8	2	8.1				
CL-20230015	168	4.07	5.1	9.1	9				
CL-20230016	147	3.18	1.2	8.7	8.2				
CL-20230017	149	3.24	1.7	7.9	1.9				
CL-20230018	162	3.3	6.8	3	2.6				
CL-20230019	143	3.95	8.7	6.3	9.2				
CL-20230020	138	3.24	1.7	9.8	5.1				
CL-20230021	152	4.25	6	2.6	8.3				
CL-20230022	159	3.45	9	9.9	5.8				
CL-20230023	152	3.57	9.5	8.8	1.5				
CL-20230024	176	4.1	3.7	7.3	2				
CL-20230025	153	3.53	3	5.9	6.5				
CL-20230026	128	3.83	4.3	2.4	9.2				

Figure 5: Conditional Formatting: Excellent LSAT Part 3

Detour: Editing Conditional Formatting

Although we do have the “Excellent” LSAT scores formatted, it is not exactly the format we wanted. The background seems to be a bit darker than what we were aiming for, depicted in cell L5 in Figure 5. We have two choices to deal with this situation; removing the conditional format and trying again, or editing the rules.

The screenshot shows a Microsoft Excel spreadsheet titled "LAW SCHOOL APPLICATION EVALUATIONS". The data includes columns for Applicant ID, LSAT, GPA, References, Statements, Experiences, and Evaluations. A conditional format is applied to the LSAT column, specifically to cells C5 through C31. The rule is set to highlight cells where the LSAT value is greater than or equal to \$1\$6 (16). The background color for these cells is green. The "Conditional Formatting Rules Manager" dialog box is open, showing the rule applied to the range \$C\$5:\$C\$30. The "Edit Rule..." button is highlighted with a blue box. The "Manage Rules..." button at the bottom right of the dialog box is also highlighted with a red box.

Figure 6: Conditional Formatting: Editing and Removing

Select the LSAT scores, and navigate to Conditional Formatting > Manage Rules. Then we can either delete the rule and start over by selecting the rule we have and clicking Delete Rule , or selecting the rule that we have and clicking Edit Rule. Clicking Edit Rule will bring you back to Figure 4, where you can adjust the background color to fit the desired format.

Topic 3. Conditional Formatting: Between...

Now we have the “Excellent” scores formatted, we must work with the “Satisfactory” scores which are scores between 155 to 169. Follow the instructions below to set up the conditional formatting for the “Satisfactory” LSAT scores in the dataset.

The screenshot shows a Microsoft Excel spreadsheet titled "LAW SCHOOL APPLICATION EVALUATIONS". The data consists of columns for "Applicant ID" and "LSAT". The "LSAT" column contains values ranging from 146 to 175. A dropdown menu is open over the cell range C5:C30, showing the "Conditional Formatting" option. A sub-menu is displayed under "Highlight Cells Rules", specifically the "Between..." option, which is highlighted with a red box. This leads to the "Format Cells" dialog box, also shown with an orange border. In the "Format Cells" dialog, the "Fill" tab is selected, showing a color palette with various patterns. The "Background Color" section is set to "No Color". The "Pattern Color" section has "Automatic" selected. The "OK" button at the bottom right of the dialog is also highlighted with a red box.

Figure 7: Between Two Values

1. Select the LSAT scores C5 : C30, and
2. Select **Between** in the dropdown menu from Conditional Formatting.
3. Set up the **lower and upper bounds** in the popup Between window.
4. Select **Custom Format**.
5. Set up the desired format for the “Satisfactory” LSAT scores in the **Format Cells** window.

Topic 4. Conditional Formatting: Less than...

Now we have the “Excellent” and “Satisfactory” scores formatted, we must work with the last classification of “Rejection” scores which are scores less than or equal to 144. Following the steps listed below, we can finish formatting LSAT scores.

The screenshot shows a Microsoft Excel spreadsheet titled "LAW SCHOOL APPLICATION EVALUATIONS". The data includes columns for Applicant ID, LSAT, GPA, References, Statements, Experiences, and Evaluation scores. The "Evaluation" column contains scores ranging from 146 to 175. The "GPA" column contains values ranging from 3.17 to 4.25. The "Score" column at the bottom right contains three categories: Satisfactory (green), Poor (orange), and Rejection (red). The "Conditional Formatting" dropdown menu is open, and the "Less Than..." option is highlighted. A "Format Rule Description" dialog box is open, showing the formula `=C5 <= 144` in the "Format only cells with:" section. The "Format..." button is visible at the bottom right of this dialog.

Figure 8: Less than or Equal to

1. Select the LSAT scores C5 : C30, and
2. Select **More Rules** in the dropdown menu from Conditional Formatting.
3. Select **less than or equal to** in the dropdown.
4. Set up the threshold points in the **orange box**.
5. Set up the desired format for the “Rejection” LSAT scores by clicking the **Format** button.

Topic 5. Conditional Formatting: Bottom/Top Values

Scores for the references, statements, and experiences are graded relatively to the pool of applicants. Therefore, the top few applicants will be granted “Excellent” classification, the next few will be given “Satisfactory,” and so forth... We will be using conditional formatting with top/bottom rules to achieve this task.

The screenshot shows a Microsoft Excel spreadsheet titled "LAW SCHOOL APPLICATION EVALUATIONS". The data includes columns for Applicant ID, LSAT, GPA, References, Statements, and Experiences. The "References" column is selected. The "Conditional Formatting" ribbon tab is active, and the "Top/Bottom Rules" option is highlighted. A dropdown menu shows various ranking options like "Top 10 Items..." and "More Rules...". The "More Rules..." option is also highlighted. A "New Formatting Rule" dialog box is open, showing the "Edit the Rule Description" section where "Format values that rank in the" dropdown is set to "Top" and "25" is typed. The "Format..." button is highlighted with a red box. Below the dialog, a preview shows a green background for the top 25% of the "References" column values.

Figure 9: Top 25% in References

To access conditions based on rank...

1. Select the Reference scores E5 : E30, and
2. Select **Top/Bottom Rules > More Rules** in the dropdown menu from Conditional Formatting.
3. Select **Top**, Type in **25**, and Check the “% of the selected range” in the dropdown.
4. Set up the desired format for the “Excellent” Reference scores by clicking the **Format** button.

The Ordering of Rules

We now move on to the “Satisfactory” category of reference scores. You can repeat the same process as the previous part, but typing in 50 instead of 25 in step 3. But you will immediately run into a problem as soon as you apply this conditional formatting as we can see in Figure 10.

LAW SCHOOL APPLICATION EVALUATIONS

Applicant ID	LSAT	GPA	References	Statements	Experiences
CL-20230001	146	3.88	8.3	3.4	6.9
CL-20230002	180	3.22	7.8	3.3	7.2
CL-20230003	152	3.47	1.8	1.3	9.1
CL-20230004	154	3.4	9.5	7.4	9.6
CL-20230005	149	3.61	8.5	1.6	2.3
CL-20230006	175	3.17	6.2	6.1	9.6
CL-20230007	180	2.28	3.8	4	6.9
CL-20230008	176	4.25	7.2	1.7	9.3
CL-20230009	158	3.64	9.1	2.1	4.2
CL-20230010	162	3.74	4	9.6	8.3
CL-20230011	176	4.12	7.1	2.9	2.1
CL-20230012	167	4.14	9.8	7.4	10
CL-20230013	174	3.65	7.4	4.1	8.7
CL-20230014	159	3.17	2.8	2	8.1
CL-20230015	168	4.07	5.1	9.1	9
CL-20230016	147	3.18	1.2	8.7	8.2
CL-20230017	149	3.24	1.7	7.9	1.9
CL-20230018	162	3.3	6.8	3	2.6
CL-20230019	143	3.95	8.7	6.3	9.2
CL-20230020	138	3.24	1.7	9.8	5.1
CL-20230021	152	4.25	6	2.6	8.3
CL-20230022	159	3.45	9	9.9	5.8
CL-20230023	152	3.57	9.5	8.8	1.5
CL-20230024	176	4.1	3.7	7.3	2
CL-20230025	153	3.53	3	5.9	6.5
CL-20230026	128	3.83	4.3	2.4	9.2

EVALUATION RULES

LSAT			
From	To	Verdict	Format
170	180	Excellent	Score
155	169	Satisfactory	Score
145	154	Poor	Score
0	144	Rejection	Score

GPA			
From	To	Verdict	Format
3.7	4.3	Excellent	Score
3.3	3.69	Satisfactory	Score
3	3.29	Poor	Score
0	2.99	Rejection	Score

References			
Criteria	Verdict	Format	
Top 25%	Excellent	Score	
Top 50%	Satisfactory	Score	
Top 75%	Poor	Score	
Bottom 25%	Rejection	Score	

Statements			
Criteria	Verdict	Format	
Top 20%	Excellent	Score	
Top 50%	Satisfactory	Score	
Top 80%	Poor	Score	
Bottom 20%	Rejection	Score	

Figure 10: Formats Overwritten

To fix this issue, open up the conditional formatting rules manager following the steps in Figure 6. Then select the first rule to be applied in the **red box**, and move it up by clicking the button in the **orange box** in Figure 11.

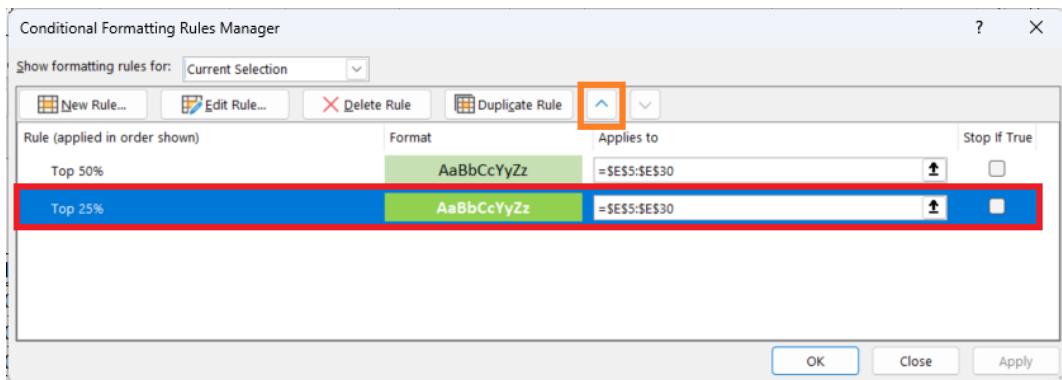


Figure 11: Conditional Formatting Rules Manager

Using these rules, fill out the remaining conditional formatting tasks in sheet LAW.

Topic 6. Formatting the Entire Row when it Meets a Condition

Navigate over to the next worksheet PUSHUP, which has data on a push-up trial. There are four criteria that determines who of the 14 people “passed” the test. In the text, we will talk about the first condition. Those who completed a total of 130 push-ups across their three attempts passes the test. The first task is to use the SUM function to find the number of push-ups that the individuals completed. In the previous task, we applied formats to the individual cells separately; so their LSAT, GPA, References, etc were all allowed to have different formats depending on how they performed in each aspect. Here, we want to check one condition for each person, and apply the format to the entire row for individuals who passed the test.

Figure 12: Pass Criteria #1

1. Select the data including the newly calculated total sum, B5 : F18.
2. Select **New Rule** in the dropdown menu from **Conditional Formatting**.
3. Select **use a formula to determine which cells to format**.
4. Type in the condition **$=\$F5>130$** .
5. Set up the desired format for the “Passing Push-ups” by clicking the **Format** button.

Push-up Scores										Evaluation Rules						
Participant	Attempt 1	Attempt 2	Attempt 3	Total	Highest	Lowest	Sum of Greatest Two									
PID-1	53	43	51	147												
PID-2	27	43	35	105												
PID-3	47	47	44	138												
PID-4	41	41	40	122												
PID-5	46	45	38	129												
PID-6	62	34	36	132												
PID-7	44	62	53	159												
PID-8	62	39	34	135												
PID-9	64	40	32	136												
PID-10	51	42	58	151												
PID-11	52	51	48	151												
PID-12	35	54	77	166												
PID-13	45	43	47	135												
PID-14	33	37	55	125												

Figure 13: Pass Criteria #1 Completed

Topic 7. Formatting the Entire Row when it Meets Many Conditions

Navigate over to the next worksheet TRIATHLON, which has data on a triathlon trial. We will learn how to format the entire row, while checking multiple conditions simultaneously. Suppose that we want to highlight the data entries for contestants that “passed” each element of the triathlon.

The screenshot shows an Excel spreadsheet titled "Triathlon". The data is organized into columns: "Contestant" (B5:B19) and "Swimming" (C5:C19). The "Swimming" column contains values "Pass" (rows 5-14) and "Fail" (rows 15-19). A conditional formatting rule is being applied to the range B5:E19. The "Format Rules" dialog box is open, specifically the "New Formatting Rule" section. Under "Select a Rule Type", the option "Use a formula to determine which cells to format" is selected. The formula entered is `=AND($C5="Pass",$D5="Pass",$E5="Pass")`. The "Preview" section shows the first few rows of the data with the first three columns highlighted in green, indicating they all contain "Pass". The "Format..." button is visible in the preview area.

Figure 14: Triathlon Pass Criteria

1. Select the data B5:E19.
2. Select New Rule under Conditional Formatting, and select “use a formula to determine which cells to format.”
3. Type in the condition `=AND($C5="Pass", $D5="Pass", $E5="Pass")`.
4. Set up the desired format.