**1-3– Tiles & Binning**

During data investigation and preparation, there are instances when we need to group rows of data. For example, it helps build a frequency distribution table which helps you analyze large datasets and determine the largest or smallest group of values in your source. Another use for grouping is to batch records based on a certain field value, such as sales. You can utilize groups to separate products that generates low, average or high sales. Alteryx offers multiple ways of grouping data using the tile tool & the multi-field binning tool.

**Tile Tool groups data into sets or “tiles” based on value ranges on a field. The tool does this based on the user specifying one of 5 methods.**

For this workflow, we are going to demonstrate some of the use cases of the Tile Tool. First, we are going to use the tile tool to create a header ID for this sales data source which has all of its records in a single column. The data should be separated into 4 different fields but we can’t simply feed it to a crosstab tool since it needs at least 2 columns, one for the header, and another for the value. In order to structure it properly, we will use the tile tool to create a numeric header ID for each 4 records and a record ID to create a distinction for each data row. From the Preparation tool tab, drag the “Tile Tool” and connect it to the Input anchor. In configuring the tool, the initial step is choosing an appropriate tile method**. The tile tool has 5 different range patterns; Equal Sum, Equal Records, Smart Tile, Unique Value, and Manual.** Each method has its own set of configurations. Since we have exactly 4 repeating lines of records in the dataset, we will use the method “Equal Records”. **Equal Records are used when input records are divided into the specified number of tiles so that each tile is assigned the same number of records.** The next setting is to indicate the number of tiles. The resulting tiles will be the record id for our dataset upon applying crosstab, so the first 4 records should be Tile 1 or Record 1, and the next 4 should be Tile 2 or Record 2. To get the number of tiles needed for this dataset, we simply need to divide the current total number of records by the number of intended columns, which would be 33,596 divided by 4. Type 8399 as the Number of tiles. Equal Records also has the option to disable split on a certain column, sort data, and group data. These settings are optional and we don’t need to use them since we are working on a single column. Run the workflow to view the results. The tile tool appended 2 fields to the original data**, Tile number** which is the assigned tile of the record, and **Tile Sequence Num** for the record position within the tile. We will use Tile number as the record ID and Tile Sequence Num as the Header ID. Drag a cross tab tool from the Transform tool tab and place it after the tile tool. Configure the cross tab to Group by “Tile\_Num”, then set the column headers to “Tile\_SequenceNum” and set the Values for New Columns as “Data”. Finally, choose “First” as the method for aggregation. Run the workflow to check the output. The crosstab’s result has 5 different columns, tile num, and 4 columns that we can rename appropriately using a select tool. From the Preparation tool tab, drag select tool and place it after the Cross Tab. Name Tile\_Num as “RecordID”, 1 into “Region”, 2 as “Product Category”, 3 as “Product SubCategory” and 4 as “Sales”. Change the Datatype of Sales from V\_String to a Double.

Next, we are going to group the records into 3 using the Sales field. Records with 500 sales or less will be marked as “Low”, 501 to 2000 will be marked as “Average” and 2001 above will be marked as “High”. We can create tiles or groups that has a specific range using **the Manual method of the Tile Tool. The Manual method lets the user specify the cutoffs for the tiles by typing a value on a new line for each range.** Insert a new tile tool into the canvas and connect it after the select tool. Set the Tile Method to “Manual” then choose “Sales” as the numeric column or the field to tile by. As the name suggests, the field must be numeric for it to show on the dropdown. Next, we need to indicate the tile cutoff. You only need to indicate each tiles upper limit on a new line. Since we need 3 group, we only need to input 2 cutoff points; 501 and 2001. The first tile will have 500 sales and below, 2nd tile will contain 501 to 2000 and third tile will have 2001 and above. Run the workflow to see the results. The Tile\_Num column contains the group number of either 1,2, or 3. Let’s add a formula tool to make the sales group distinguishable. From the Preparation Tool tab, drag “Formula” and place it after the 2nd tile tool. Create a new column called “Sales Group” and set its data type to V\_WString size 255. Then input the following expression “Switch([Tile\_Num],null(), 1,"Low", 2,"Average", 3,"High")”. This switch function tags the group name based on the number assigned in Tile\_Num. Run the workflow to check the output. Now we have the Sales Group for each record. Do note that the output for Manual mode is not sorted so you should add a sort tool if needed.

As our third example, we will demonstrate how the tile tool can be utilized when formatting a text table. We wanted to create a text table that has the Region, Product Category, Product Subcategory and Sales with a conditional format that sets an alternating background color based on the Region to create distinction between regions. The first region will have a white background, the second will be gray, the third will be white and so on. First, add a Summarize tool to aggregate the data. From the Transform tab, drag Summarize and connect it after the Formula tool. Highlight Region, Product Category, and Product SubCategory then choose “Group By” as the summarize action. Once done, select “Sales” from the fields and set “Sum” as its action. Now that we have the summarized data, add the Tile tool and connect it to the output anchor of the summarize tool. Since we are going to apply an alternating background on each unique region, we will use the tile method **“Unique Value”. In this method, a unique tile assigned for every unique value in a specified field or fields. If multiple fields are specified, a tile is assigned based on that combination of values.** In the tool’s configuration, select Region as the unique column. We already have the data sorted by Region and Product Category & SubCategory since we used the summarize tool so let’s enable the option to “Leave Unsorted”. If unchecked, the table will be automatically sorted by the field or fields you specified for the tile tool. Run the workflow to check the output. Now we have a unique Tile\_Num for each region. Next, navigate to the Reporting tool tab and insert a Table Tool after the tile tool. Uncheck the Tile\_Num and Tile\_SequenceNum on the column configuration since we don’t need them to be displayed in the output. **Next, click the button “Default Table Settings” to open the table style editor. By default, alteryx adds an alternate color for every other row.** Since we wanted to alternate it by region, set the alternate color to “White” in order to make the table background all white. Next, set the Header style to Bold and click OK to close the style editor window. To set the alternating background color, click the button “Create Row Rule”. Set the rule name as “Even Region Color” then choose to apply it using a Formula. In the formula textbox, type “Mod([Tile\_Num], 2) == 0”. This formula checks if the tile number of the record is even, if yes it will apply the formatting but if not it will follow the table’s default. Then tick the Background Color checkbox and choose the silver color. Once the table configuration is all done, add a browse tool to view the report, then run the workflow. Now we have a table with alternating colors per region.

**Multi-Field Binning Tool groups multiple numeric fields into tiles or bins. This tool is built primarily for the predictive toolset and only accepts numeric fields for binning unlike the Tile Tool.**

On our workflow, navigate to the Preparation Tool tab and drag multi-field binning tool into the canvas. Connect it to the output anchor of the summarize tool. In the tool’s configuration, the first option is to select the fields to be used for binning. In this case, we only have sum\_sales as the sole numeric field. Tick its checkbox to set it for grouping.

Next, the tool presents us 2 options on how the records will be binned. Set **Equal Records** if you want to divide the data into a specific number of bins with equal number of rows for each bin. This binning process is solely based on the records position. **Equal Intervals** will find the minimum and maximum values of the selected fields and the range is split into equal-sized sub-ranges. Records are assigned to bins based on these ranges. For this example, choose the method “Equal Records” and set the Number of Tiles to “4”. Run the workflow after finishing the configuration.

As we can observe on the output, the tool appended a single column to the dataset called “Sum\_Sales\_Tile\_Num”. This assigned our data into 4 bins where each individual bin has 17 rows.