**Alteryx User Guide**

**Team Data Wranglers**

**GMU DAEN 690 Section 4**

**Professor Robert E Kraig**

[**Table of Contents**](#_30j0zll)**: Page**

1 [Introduction  3](#_1fob9te)

[2](#_3znysh7)  Installation\Sign Up 3

[3](#_2et92p0)  Importing Data 5

4 Flow/Recipe Creation 8

5 Output 16

[6](#_4d34og8) Customer Support 17

[7](#_4d34og8) References 19

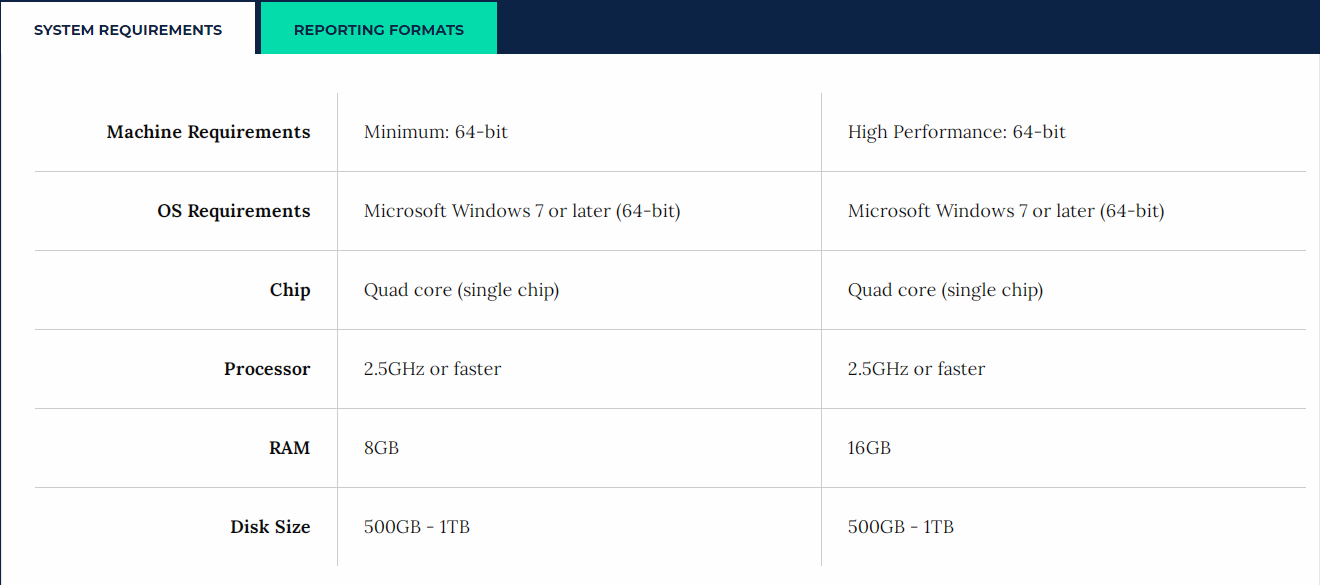
**1) Introduction:**

The contents of this document provide an insight to working on the data prep tool “Alteryx”. This document gives a step by step information from the installation of the tool to recipe creation and finally to get the output with the 64 standard attributes as required by the client “Share Our Strength”. The information in this document can be used by the client to understand the tool better and work on it without any difficulties in the recipe creation.

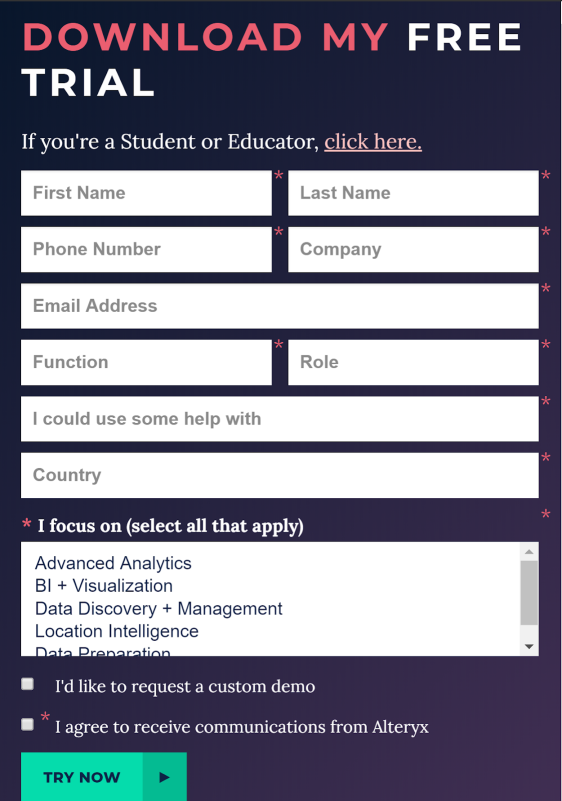
**2) Installation /Sign Up:**

Alteryx products, such as Alteryx designer, are available for download from the Alteryx Downloads and Licenses portal on their main website. It requires installation on the local machine, and free trials are available. Initially we have to sign up using an email address, license key, and by inputting other administrative information. You can then apply for the 14-day free trial or request an extended trial period by signing up as a student or educator.

Below is a table of the system requirements for installing 'Alteryx Designer'.



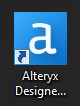
When installing the software, you will be asked to fill out information such as First Name, Last Name, Phone Number, Company, etc. This information is also collected for those that want to install the trial version.



Once we provide the details, it sends a verification to the email account in order to verify the email account provided. In addition, a license key will be sent to the registered email account as well.

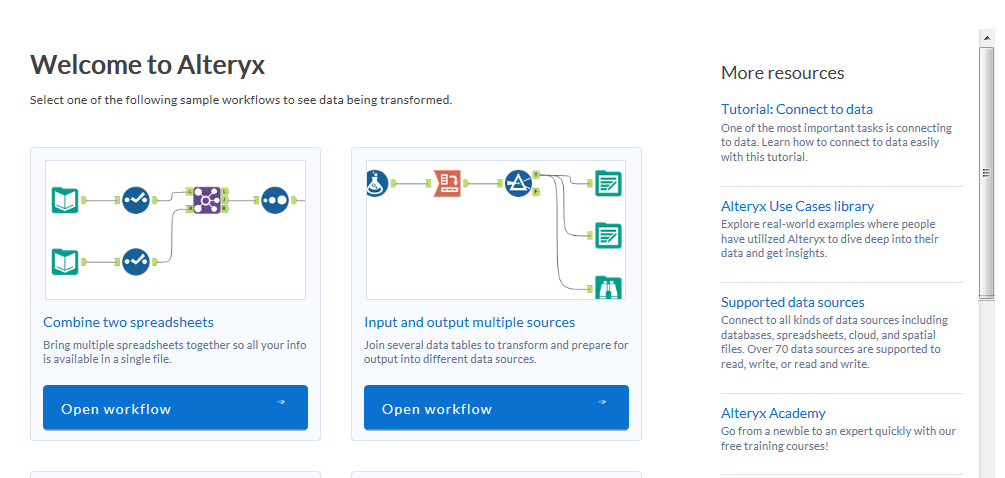
Then you proceed with the installation of the software, which, depending on the speed of your computer could take anywhere from 5-20 minutes.

After installation is complete, you should see the application icon saved somewhere on your computer.



As you would with any other software, double-click on it, and the program will be start up.

The opening screen will be displayed as is shown in the screenshot below. From here you can open up a previously saved project, start a new one, or learn about some more Alteryx tutorials. You can learn more about the learning resources available on Alteryx, where it gives a short but very informative glimpse of how Alteryx tools operate from importing data to analyzing results.



**Note:**

Please do not skip this tour as it can be very helpful in understanding the operation structure of the tool.

**3) Importing Data:**

Importing the data into the Alteryx is an important process which enables us to bring in data from our personal computer. As all the raw data files coming in are in an excel format for the 'No Kid Hungry' project, importing data is very simple.

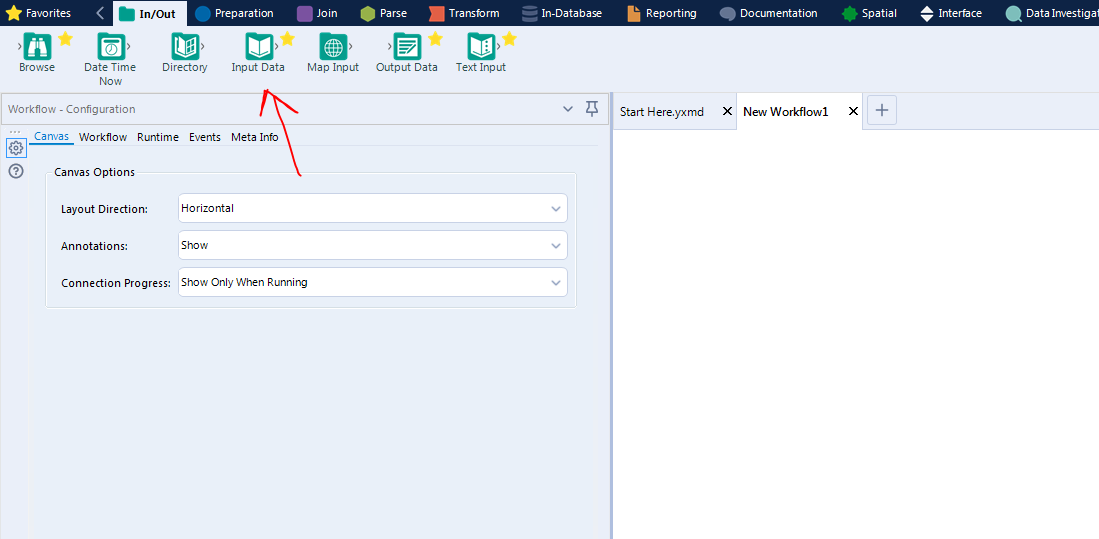
**Note:**

Alteryx considers each sheet in the Excel Workbook as a separate dataset. For Example: If we import an Excel workbook with three sheets, then Alteryx will consider it to have three datasets.

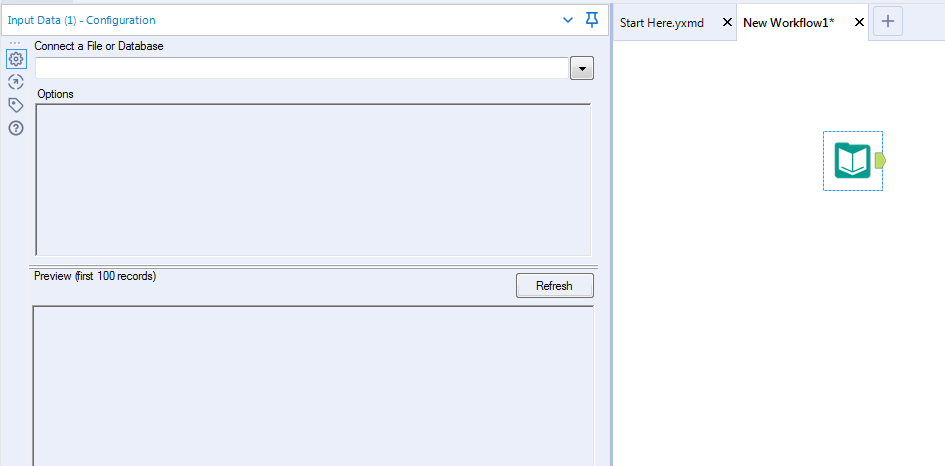
Alteryx can also handle different data sources other than from the Personal Computer. It can handle data sources such as “Hadoop File Distribution System”, Relational Databases (Microsoft SQL Server) and other databases such as ODBC, OleDB, Oracle OCI, and ESRI. A full list of all compatible file types can be seen on their website, or on the software itself.

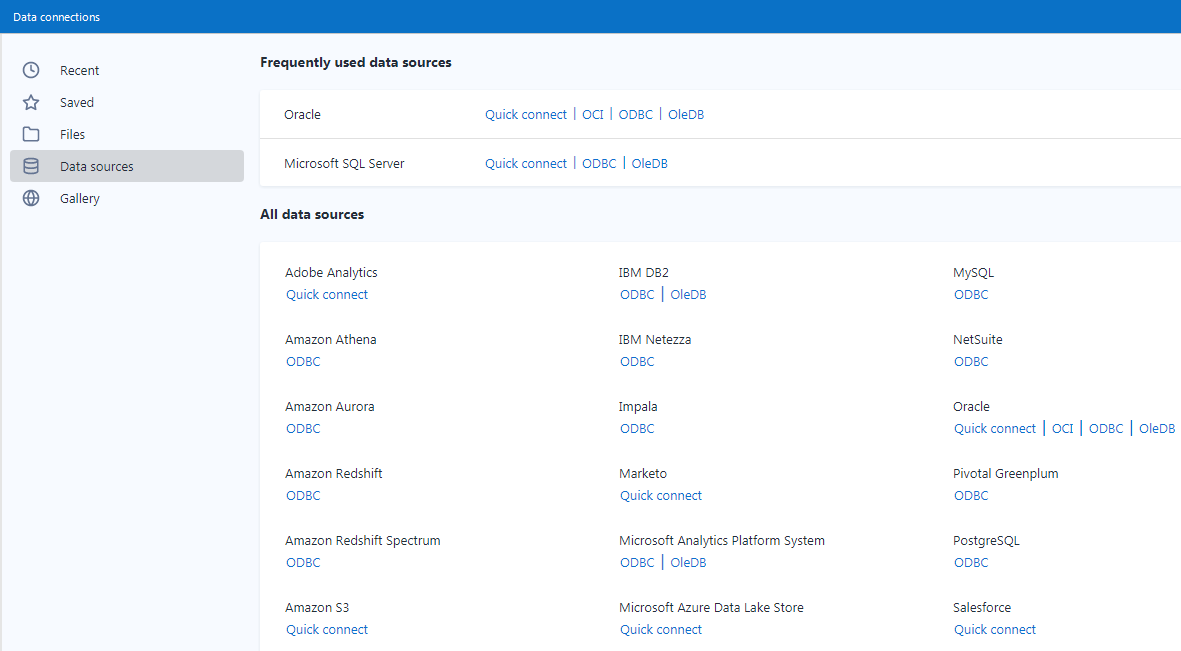
On the free version we used, the option to upload from a variety of data sources was available.

The process was as simple as dragging and dropping the 'Input Data' tool from the toolbar at the top of the page into the blank workflow space.

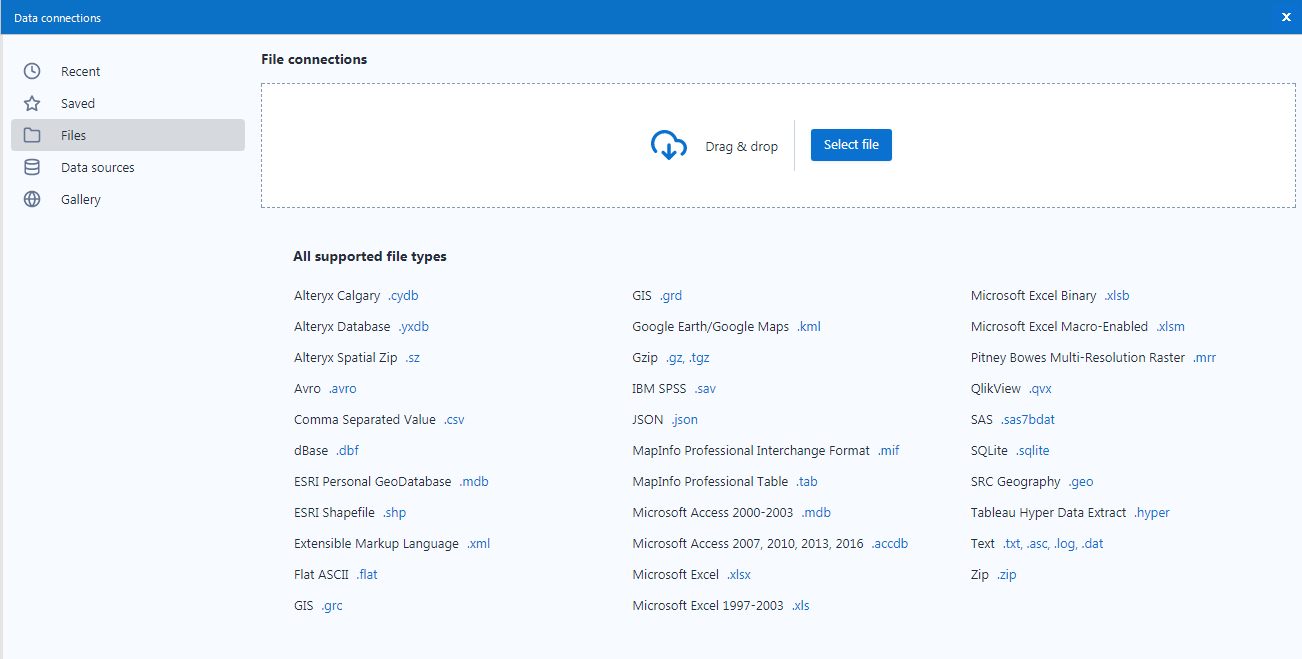


The 'Input Data' tool, as stated above, is simply dragged and dropped into the empty workflow. After placing it, an action window will appear on the left side. From there, the user will follow the step-by-step process to bring in the raw data to the workflow by selecting the drop-down arrow in the "Connect a File or Database" bar.



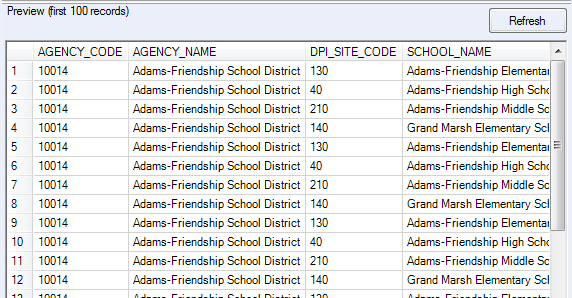


The above screenshot takes a closer look at connecting various different types of Data Sources to Alteryx Designer. Whereas the screenshot below looks to bring in a standard file to Alteryx. The supported file types are also mentioned below.



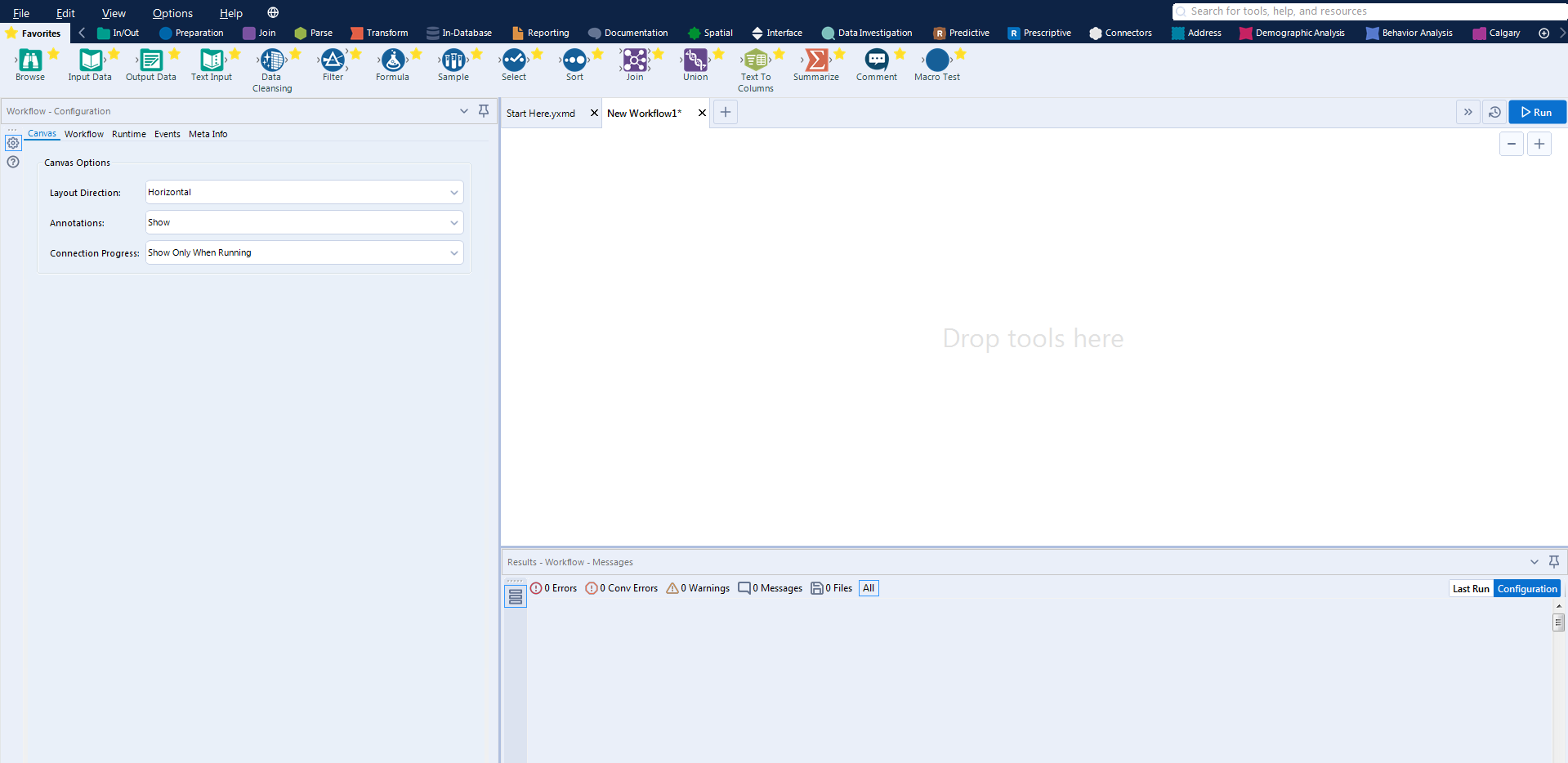
Once uploaded, some more actions might need to be taken based on the file type. For example, if the raw file is an excel file, Alteryx will ask to confirm which sheet in the excel file is being used, which row/columns are selected.

If you check just below the configuration window on the left side of the screen, we can notice the dataset has been successfully uploaded to the Alteryx workflow since we can see a quick preview of the dataset itself.



**4) Flow/Recipe Creation:**

Alteryx follows a drag-and-drop style for building workflows. The tools are dropped into the blank space and automatically form a connection when dropped next to each other.



Once the input tools have been set in place, the workflow begins, from there, the user can start configuring the settings according to their needs and start applying the required data transformation tools available.

Adding these transformation steps is as simple as dragging-and-dropping the icons from the toolbar at the top and placing them accordingly. The software automatically makes a connection between two adjacent tools and this will appear as a dashed line right before placement so that the user can make sure of their actions. If necessary, the tools can always be rearranged and altered as seen fit.

**Building a Recipe**

In Alteryx, building a recipe is made to be quite simple. A **recipe** is a sequence of transformation steps that is implemented to transform your source dataset. All the transformation steps are done with the use of the tools available to us. These tools are applied and can then be customized to fit the sequence as needed.

**Transforming Wisconsin Data by Building a Recipe**

The raw data coming in is converted into the clean data file by creating the 64 fields/columns as mentioned numerous times before. Alteryx performs these transformations through the implementation of the following tools:

* Input Data Tool
* Data Cleansing Tool
* Auto Field Tool
* Select Tool
* DateTime Parse Tool
* Formula Tool
* Browse Tool
* Join Multiple Tool
* Join Tool
* Output Data Tool

There are plenty more tools available to use, but the ones listed above are what were used to complete the Wisconsin data transformation on Alteryx. The input/output data tools are self-explanatory so there won’t be any further explanation there, but for the remainder of the tools, see below for more details.

**Data Cleansing Tool** 

The Data Cleansing tool is used to fix common data quality issues. It tidies up the data by performing actions such as replacing null values and removing whitespace. No steps are needed, just simply drag-and-drop the data cleansing tool and click on the actions from the menu that you want to apply.

**Auto Field Tool https://lh5.googleusercontent.com/ZXAwS2S0TdF43l2mrM6A9Me9zPzSJmF7jHGYuaTyzwcbIqBaTe13r3HP9I7aC8_4cxirGL_doXbAXvHc2qS6H4rBWvyBg4MhzCKRtISxRXqq6YZjsuuNKqLxM-GxhbdV57_0PGx3**

The Auto Field tool reads through the data being funneled into it and automatically sets the field type to the smallest possible size relative to the data. It does this automatically, and while this step is not necessary, it does help prepare the data for the upcoming transformations.

**Select Tool https://lh3.googleusercontent.com/aNa4CoOafVPBQuEd7LJJEzXPKB-3F5Qv3-xEHhmi7UIPOsRLmrq9xEnswVwsq5qwfGh5GBk6At0kxj9USO38wXKkTAhgTxK61Kk-Mx3CtTm5CmnexyEAhC5j9ic_DZzICbjIdPrq**

The Select tool is used to rename columns, rearrange them, and select/deselect them. Renaming columns is as simple as going to the target field in the popup window once the tool has been implemented and typing in the new column name. Rearranging the columns is also as simple as clicking on the field and then the up or down arrows at the top of the display to move them accordingly (the top = the left side of the spreadsheet, the bottom = the right side of the spreadsheet). Finally, selecting/deselecting columns means whether or not you want to keep it for the next transformation process and eventual output data (i.e. removing columns).

**DateTime Parse Tool https://lh3.googleusercontent.com/bYREHCtC-kOuGG-IFMS8iR9_bWCJZIESuw_I6QtGJzr4Gkoa-97b_3ksT_a2E7pPowfTBccUWmUf2baw82yqvKWcVJKN_GAQYgfR_jgC4OTT2C7MLSLX6dCHQrF1u-cYd8zGWsyl**

The DateTime Parse tool is performed to change a string type to a date format or vice-versa. This step is needed for the clean data file as the raw date format does not meet the requirements. The limitation here on Alteryx however, is that it only has one default date format which is yyyy-mm-dd. Any other format can only be altered if it comes back as a string type. Not a huge issue in the end, but certainly worth mentioning.

**Formula tool https://lh4.googleusercontent.com/tcbgVhl3TEV9pb-ffFA6cYrJap18UNo90L4ISEJvHYUO_nv5V6gzONGu8ue9qtnxV0InIcLrP-FmD2yWoMvx5fIoEzGzGZd7Erz2zEa0sGLT0bwE8mWFx1TlIoOMIbHi1ctiJyjR**

The formula tool is where most of the transformations will take place. This tool can create new columns, update existing columns, and use one or more expressions to perform a number of different calculations and operations. The tool can be used to do the following:

-        Apply conditional statements (if, then, else…)

-        Convert data types

-        Analyze spatial data

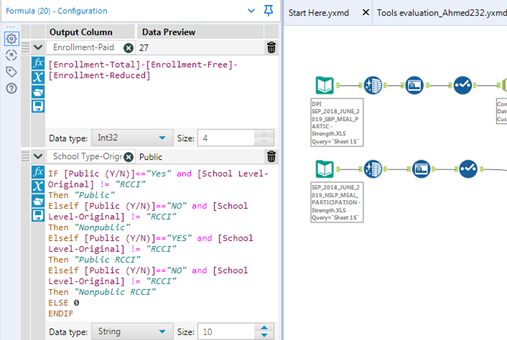
-        Apply financial algorithms

-        Apply mathematical operations

-        Create new columns (from pre-existing ones)

-        And more…

The expressions are unique to Alteryx but the syntax is similar to that of Tableau Prep. It might take time to figure out the proper functions of it but plenty of resources are available online for further assistance. Some formulas that were applied using the formula tool can be seen below.

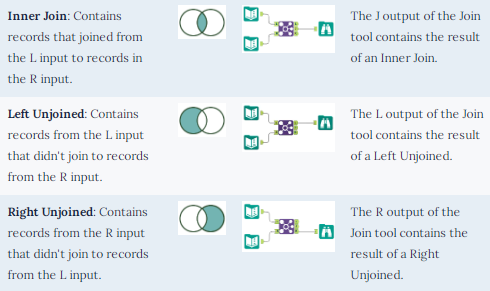


**Browse Tool https://lh5.googleusercontent.com/S1ZIk3HSzxL2O3GVZl5ut5FQo8IjQepvLyLsU1snEP5j2d17RSEsF0xpjAE257udH1VYULgY6zRXj3BJLu6_GSPmxNnZw1hdOdelBdVhqVGNb21fiS4FAln_SDVOjMhTMroJkftM**

The browse tool is used to display data from a connected tool. It displays the data profile information for multiple columns in a single view. The tool can be configured as needed to display what is required and as is with the rest of these tools, its implementation is simply done by drag-and-dropping into the workflow.

**Join Tool + Join Multiple Tool https://lh4.googleusercontent.com/VZFwmKmQpFk1qsYfApa_h0GOhI1uJNcz51IobQNWcb5ujvQOoVSIZUu-lr5xpSaTMPhZhWKjrxlWQlZlFyOkfEJpRlznF-oK5bsEBOEMO91soPOZ0RDNwcVtwxTnfBTVkSnAx8m0**

The Join tool, and the Join Multiple tool are pretty self-explanatory. The Join tool is used to combine two inputs based on a common field between the two datasets. The datasets can also be joined based on record position according to the user’s configuration. The join tool is dragged and dropped in the workflow and has two connections at the inputs (the two datasets). As stated, you can either join by record position or by a specific field. Once the selection has been made, the process should be complete baring no error messages. Important to note that the two input datasets can be joined by having one or more fields in common. The user can also select, deselect, and reorder columns in this step. Finally, there are multiple types of joins, each is used to obtain the desired outcome. Below is a table from the Alteryx website that displays some of the types of joins and how they are made.



The Join Multiple tool performs the exact same features as the Join tool with the only difference being that this tool allows for two or more inputs. In addition, by default, this tool outputs a full outer join, meaning all of the records from the inputs are kept in the output.

**Find & Replace Tool** find and replace tool.PNG**:**

The Find Replace tool searches for data in one text field from an input table and replaces it with specified text field table from a reference table. It resembles the VLOOKUP function in Microsoft Excel for those who are familiar. The tool has two inputs:

- F anchor: The left input links with the table that is to be updated with the results.

- R anchor: The right input is the lookup table "R", so this is the table containing data used to replace data in the original input.

The tool can be configured to match the workflow requirements. For the field value, the tool can be configured locate text fields in the following manner:

- Beginning of the field

- Any Part of Field

- Entire Field

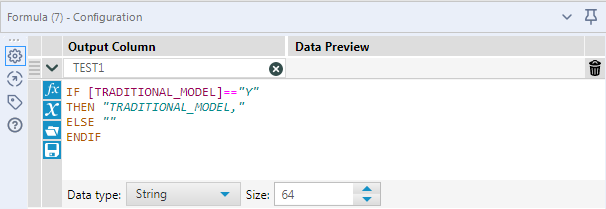
Other configurations can be made as can be seen in the screenshot below.



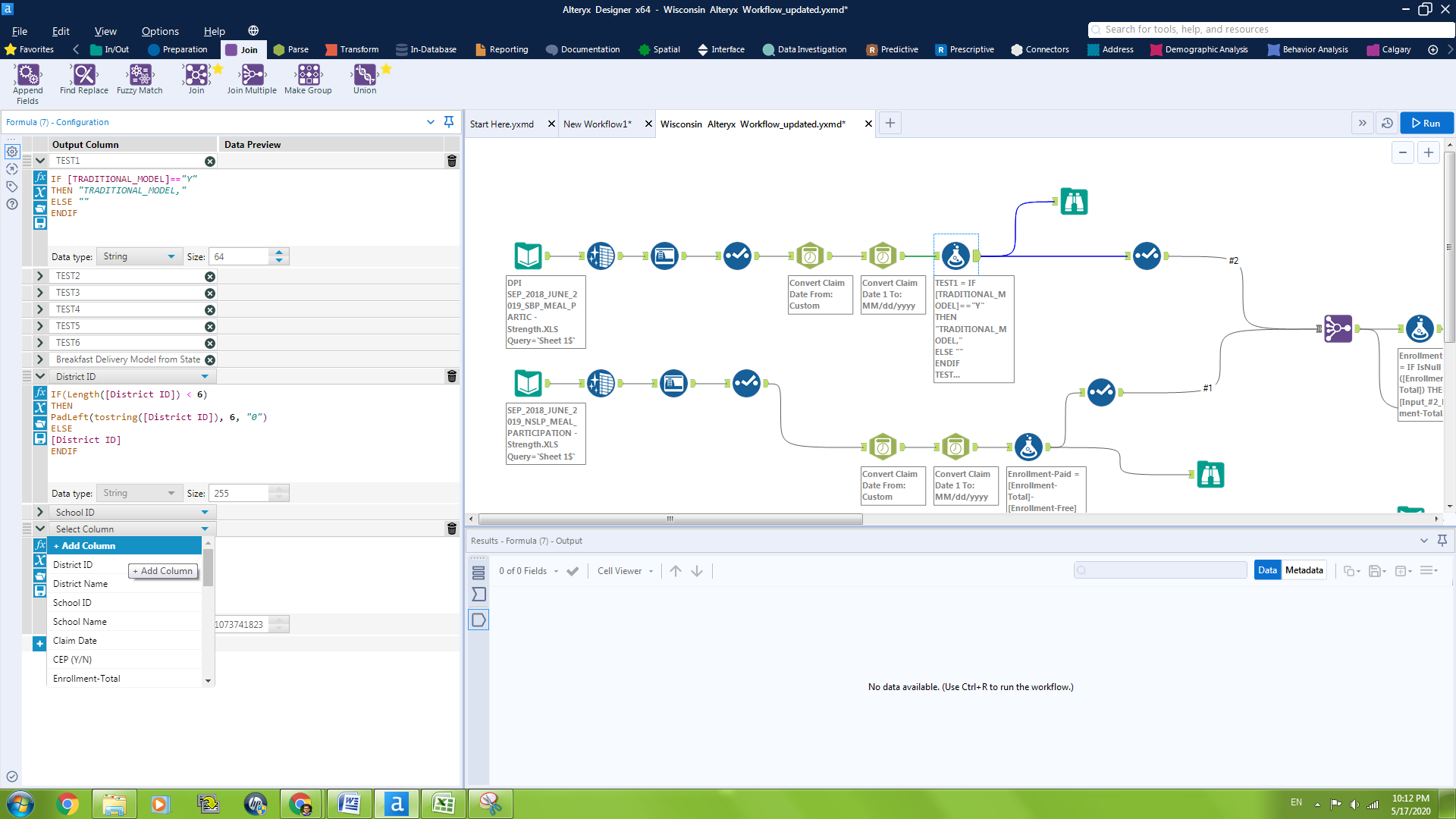
**Functions:**

Using the formula and select tool, a lot of the following functions can be performed, while these functions will differ depending on your requirements, a series of examples will be attached to each function to give some sort of guidance on the capabilities that can be achieved:

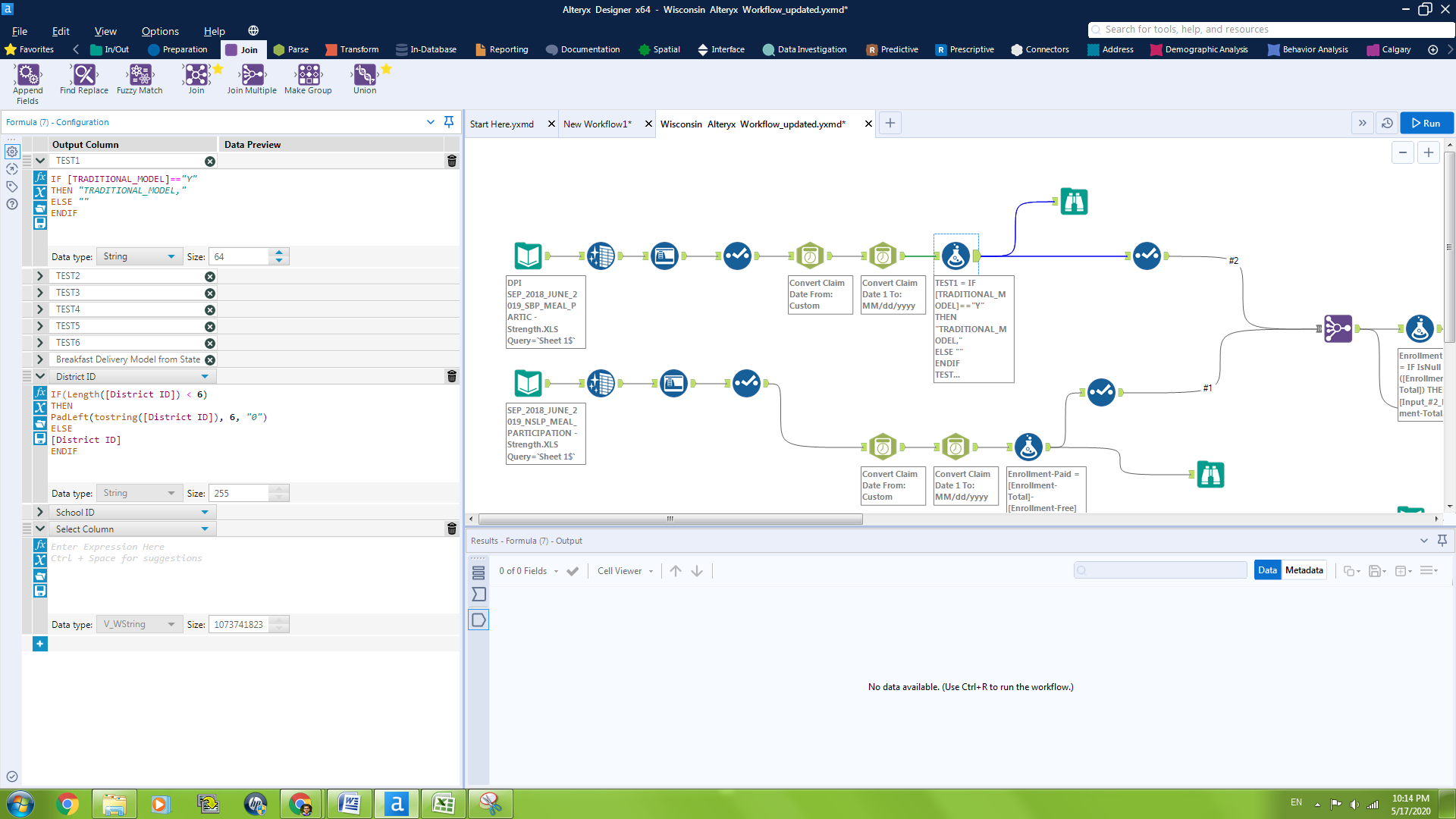
**IF Function:**



**Create a new Column:**



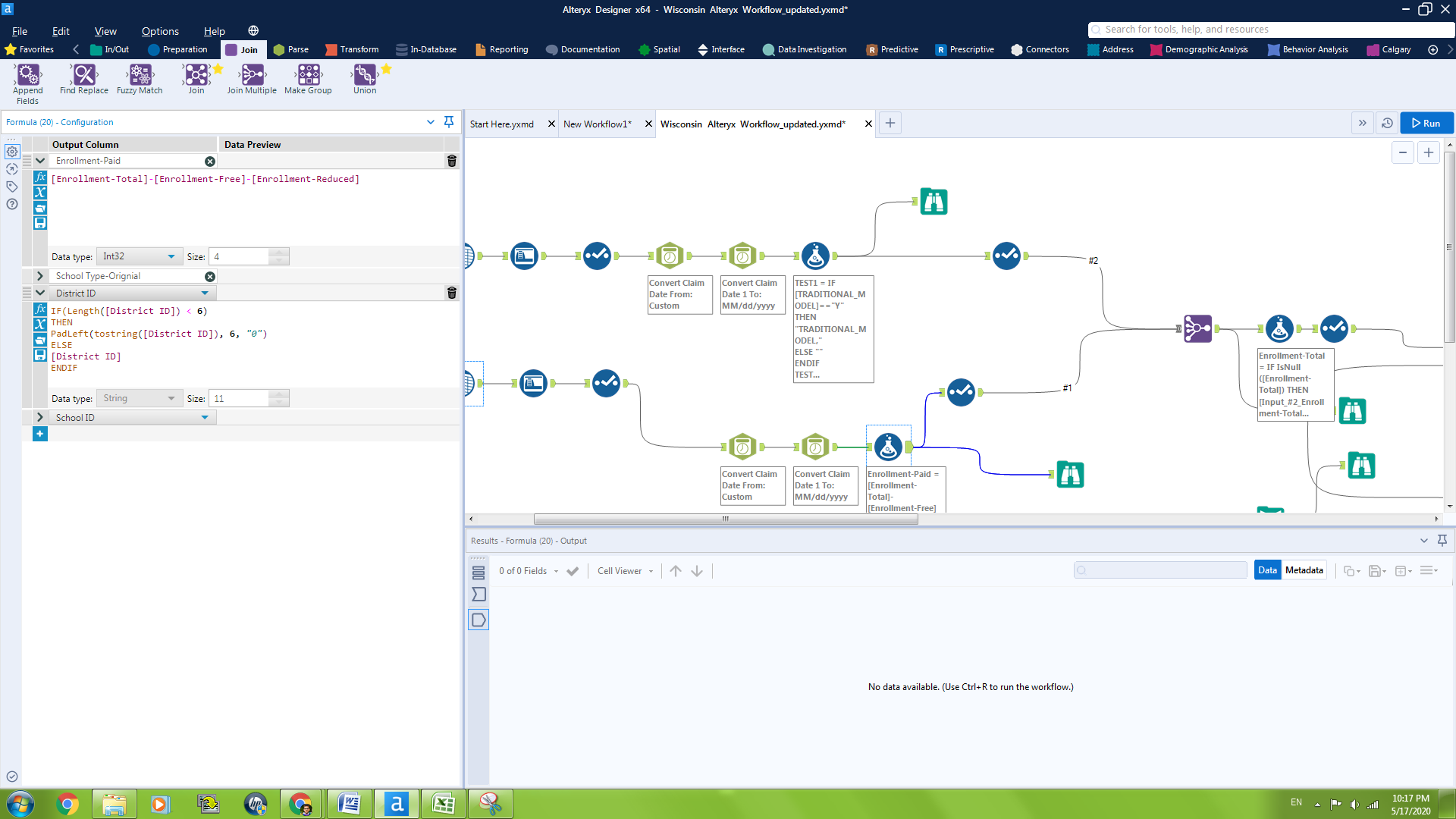
Creating a new column is simply done by clicking on add column in the drop down menu and inputting the proper syntax to populate it. Any function can now be performed here. If a new column is a simple modification of other columns, an IF function can be useful. An example is seen here below for creating a 'District ID' column:



**Add/Subtract Functions:**

The add function is performed when one column is generated by the combination of other columns. Usually this is left for numerical values but also be done for string values.

Example For Add Function:

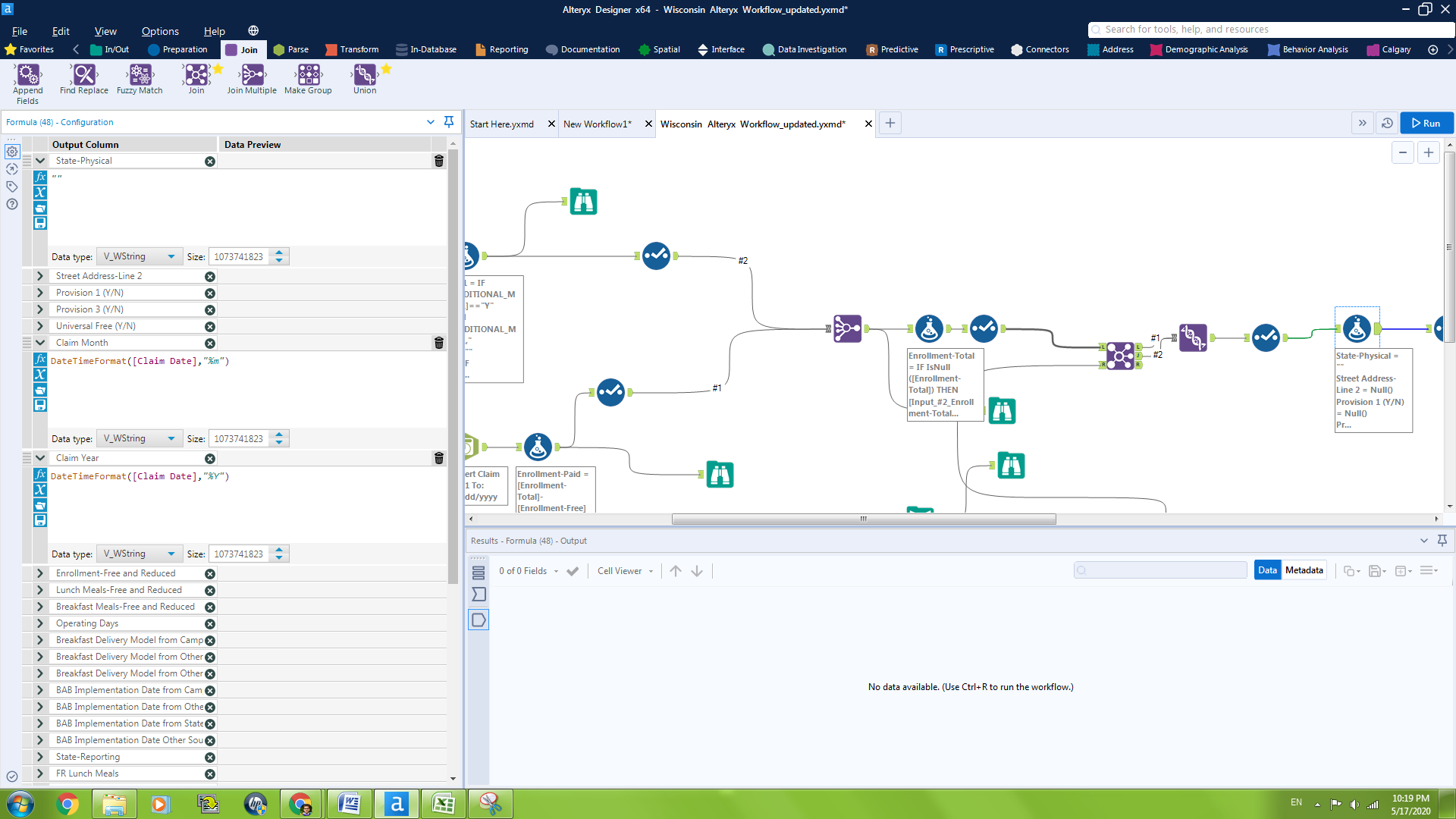
****

**DATE Functions:**

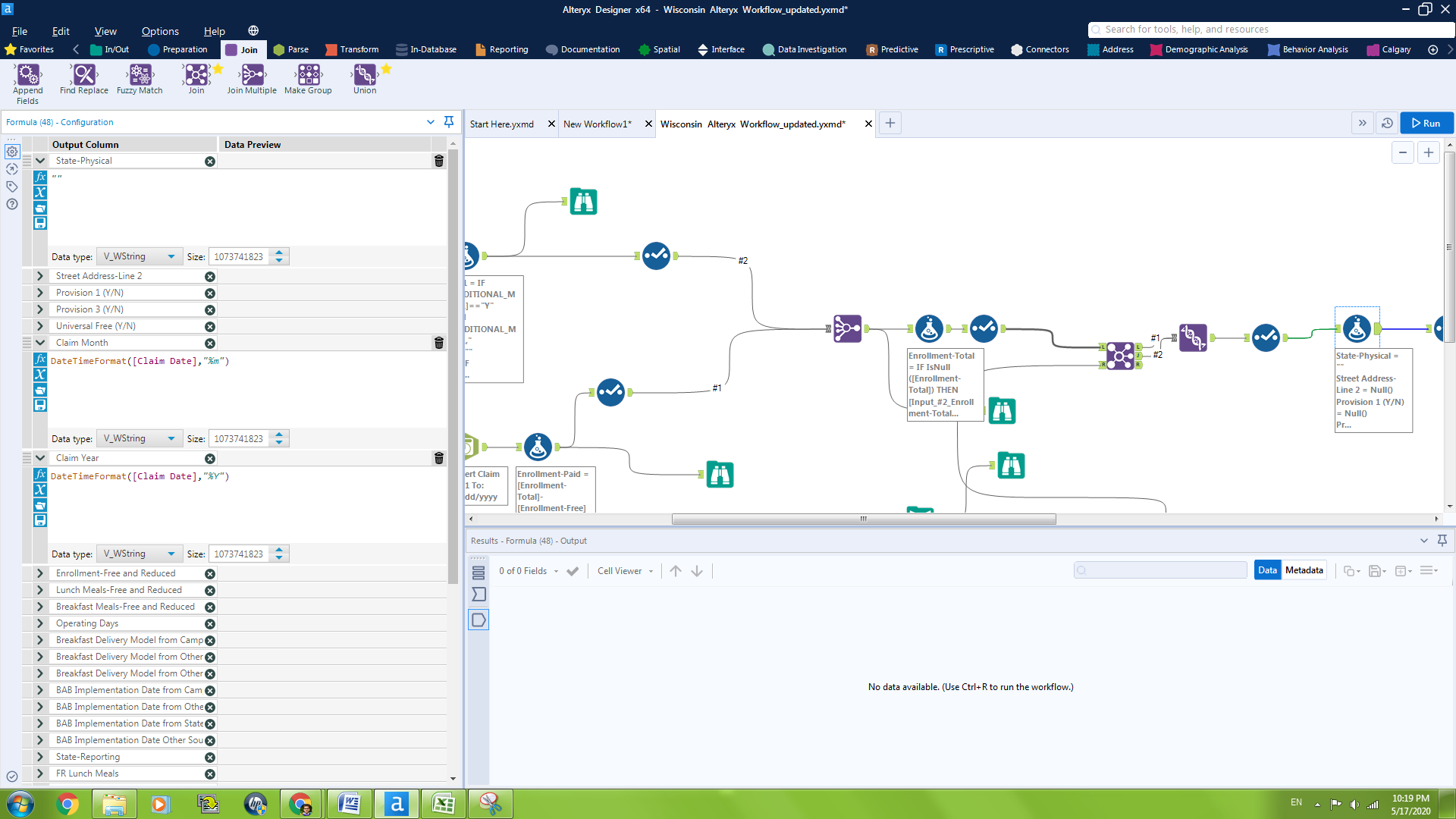
Date functions are inbuilt in the Alteryx tool, which can be used based on the information that is to be extracted from a date field or functions that can be implemented on a date field. Syntax is important for the date functions to be implemented.

**Example of Date Functions:**

Month Function:



Year Function:

****

**Some more examples of what can be achieved with the Formula Tool**

**Nested IF Functions:**

Derive School Type-Original

IF ({Public (Y/N)} == 'YES' && {School Type-Original} <> 'RCCI','Public', (IF({Public (Y/N)} == 'NO' && {School Type-Original} <> 'RCCI', 'Non-Public' ,(IF( {Public (Y/N)}=='YES' && {School Type-Original} =='RCCI', 'Public RCCI', (IF ({Public (Y/N)}=='NO'&& {School Type-Original} == 'RCCI', 'Non-Public RCCI')))))))

**Multiple IF Statements:**

Breakfast Delivery Model from State Agency Tracking-Original

IF(TRADITIONAL\_MODEL == 'Y', 'TRADITIONAL\_MODEL', '')

IF(MID\_MORNING\_MODEL == 'Y', 'MID\_MORNING\_MODEL', '')

IF(CLASSROOM\_MODEL == 'Y', 'CLASSROOM\_MODEL', '')

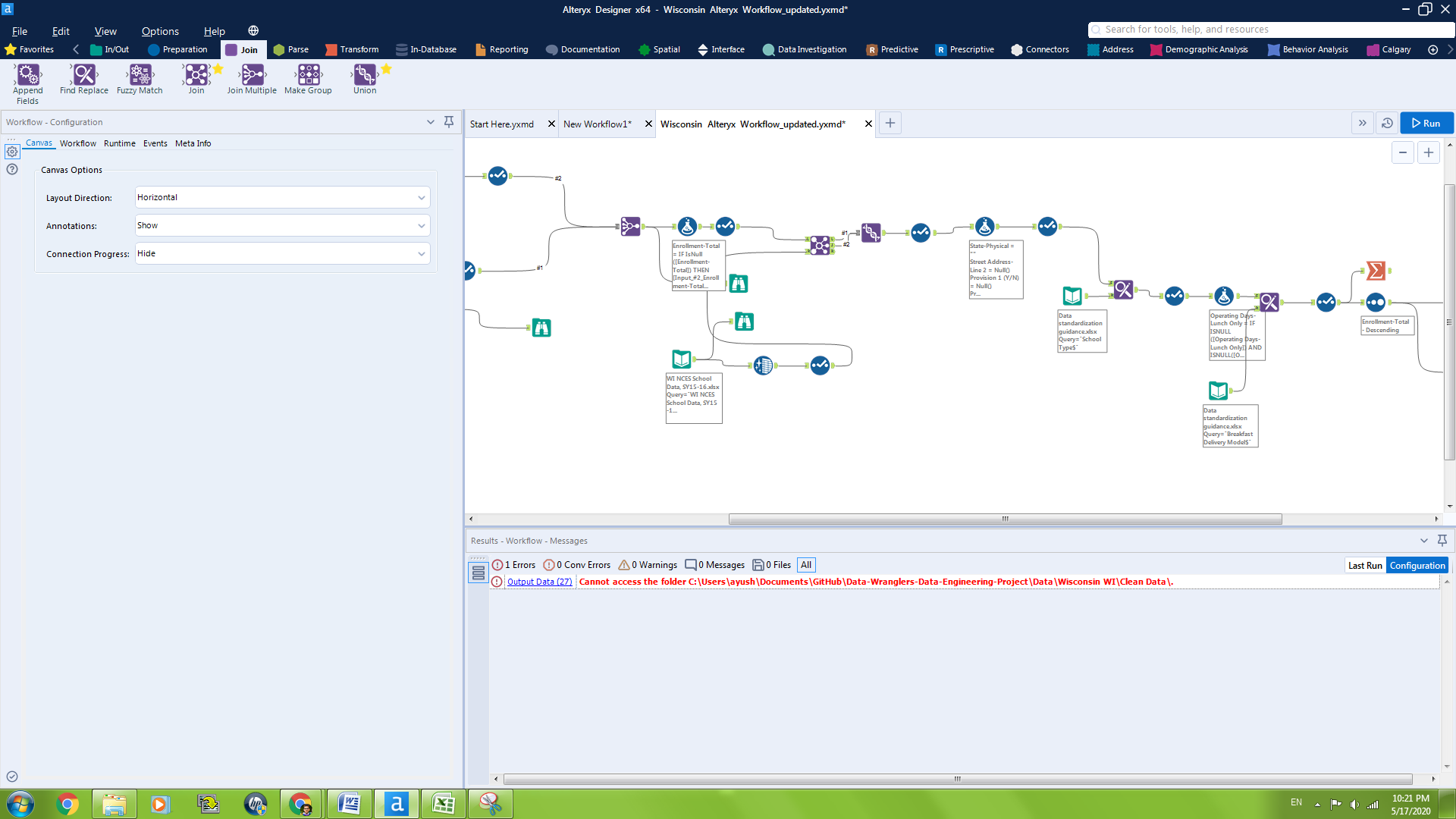
IF(REDUCED\_PRICE\_MODEL == 'Y', 'REDUCED\_PRICE\_MODEL', '')

IF(GRAB\_N\_GO\_MODEL == 'Y', 'GRAB\_N\_GO\_MODEL', '')

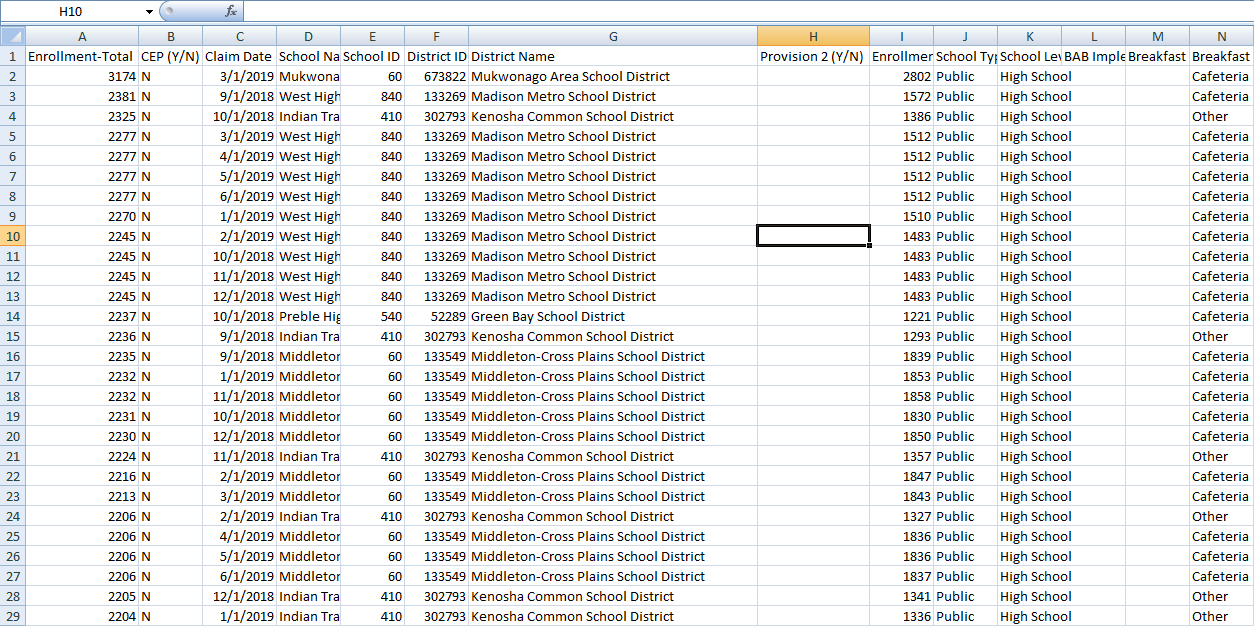
IF(FREE\_MODEL == 'Y', 'FREE\_MODEL', '')

**5) Output:**

The output is generated after the workflow has been executed. To do that, you simply click the 'Run' button at the top right of the workflow window. The workflow ends with the final connection being the 'Output Data' tool, which can be found in the toolbar above.



**Clean Data File Sample:**



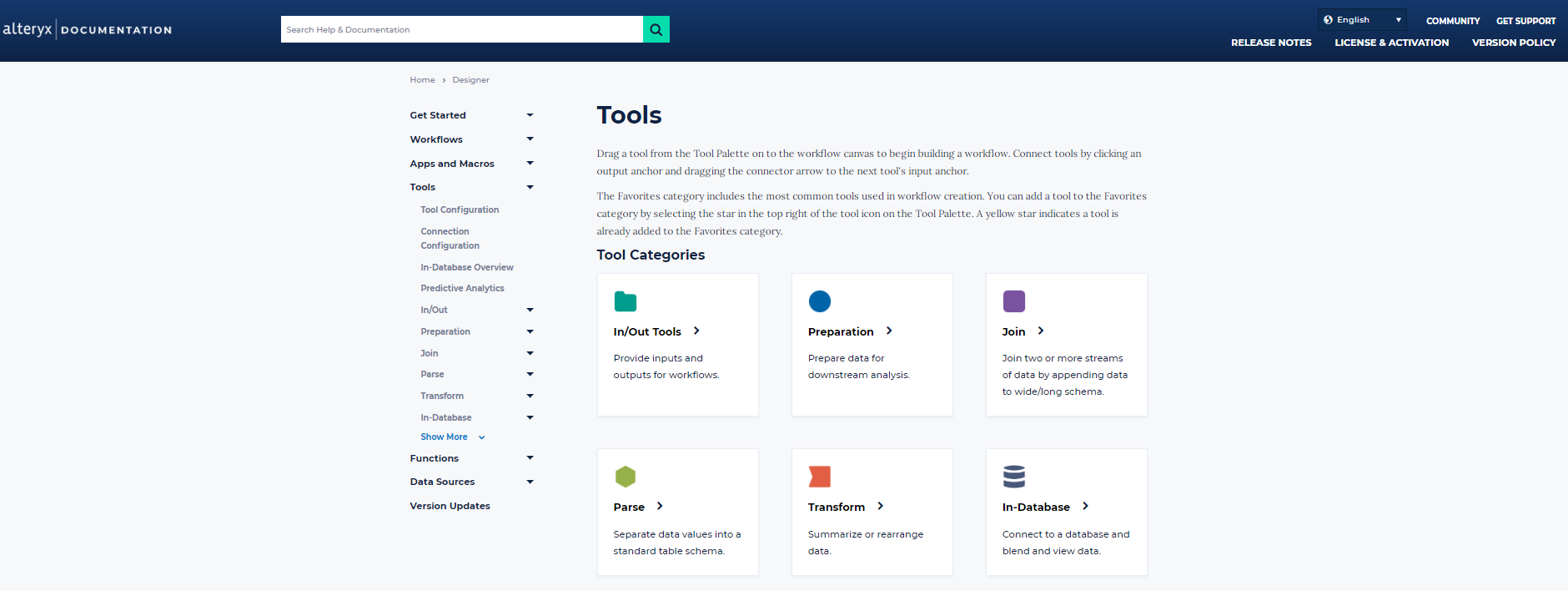
Clean Data file is in CSV format.

**Note:**

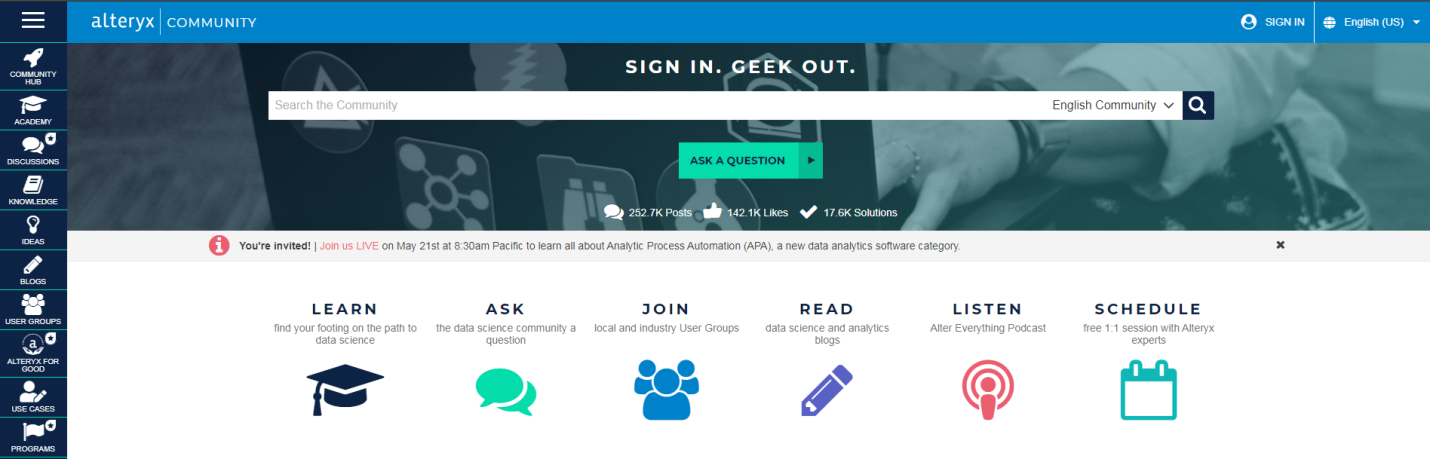
In case of any mismatches in the clean data file, changes can be made in the recipe at any point of time to alter the output.

**6) Customer Support:**

Alteryx provides support on three major levels: online Alteryx documentation, user forums on their website, and through their online technician support. The Alteryx documentation provides a comprehensive overview of everything Alteryx related. It breaks down all the tools, features, and capabilities of Alteryx Designer as well as their other products.



The forums are based on user questions that Alteryx technicians answer. In many cases, a solution to a specific issue that a user runs into can be found on these forums. The forums exist in the Alteryx community hub where discussions, use cases, and other ideas can be explored.



Finally, with their online technical support, a user can set up a meeting with a specialized technician to solve any issues they might be facing. The response time is usually quick, and a meeting can be set up within the next few business days.

Furthermore, there are additional resources available for the user in form of the 'Alteryx Academy'. The academy has learning paths, interactive lessons, videos, weekly challenges, and certifications that can be acquired.

Overall, excellent customer support.

For further information on using Alteryx, please visit:

Alteryx Community Hub: <https://community.alteryx.com/t5/Community-Hub/ct-p/community-hub>

Alteryx Help Page: <https://help.alteryx.com/current/designer>

**7) References:**

Alteryx Designer Help Page. (n.d.). Retrieved from https://help.alteryx.com/current/designer

Functions - Alteryx Documentation. (n.d.). Retrieved from <https://help.alteryx.com/current/designer/functions>

Gartner Peer Insights (2019).  [Data Preparation Tool Reviews](https://www.gartner.com/reviews/market/data-preparation-tools). Retrieved from <https://www.gartner.com/reviews/market/data-preparation-tools>

Koegelenberg, I. (2016, September 5). Introduction To Alteryx. Retrieved from<https://www.tridant.com/blog/introduction-to-alteryx> (\*for notes: TABLE in section 3)