Data Interpolation and Extraction

(ii)

This guide was created by the staff of the GIS/Data Center at Rice University and is to be used for individual educational purposes only.

The steps outlined in this quide require access to ArcGIS Pro software and data that is available both online and at Fondren Library.

- · Obtaining the Tutorial Data
 - OPTION 1: Accessing tutorial data from Fondren Library using the gistrain profile
 - OPTION 2: Accessing tutorial data online using a personal computer
 - Opening an Existing Project
- Scenario Description
- Preparing Data
 - Preparing Vector Data
 - Preparing Tabular Data
 - Mapping XY Data
- Interpolating Data
 - Point Interpolation
 - Inverse Distance Weighted Interpolation Method
 - Spline Interpolation Method
 - Natural Neighbors Interpolation Method
 - TIN Interpolation Method
 - Line Interpolation
- Extracting Data
 - Extracting Multi Values to Points
 - Raster Calculations
- (i)

The following text styles are used throughout the guide:

Explanatory text appears in a regular font.

- 1. Instruction text is numbered.
- Required actions are underlined.
- 3. Objects of the actions are in bold.

Folder and file names are in italics.

Names of Programs, Windows, Panes, Views, or Buttons are Capitalized.

'Names of windows or entry fields are in single quotation marks.'

"Text to be typed appears in double quotation marks."



The following step-by-step instructions and screenshots are based on the Windows 7 operating system with the Windows Classic desktop theme and ArcGIS Pro 2.1.3 software. If your personal system configuration varies, you may experience minor differences from the instructions and screenshots.

Obtaining the Tutorial Data

Before beginning the tutorial, you will copy all of the required tutorial data onto your *Desktop*. Option 1 is best if you are completing this tutorial in one of our short courses or from the GIS/Data Center and Option 2 is best if you are completing the tutorial from your own computer.

OPTION 1: Accessing tutorial data from Fondren Library using the gistrain profile

If you are completing this tutorial from a public computer in Fondren Library and are logged on using the gistrain profile, follow the instructions below:

- 1. On the Desktop, double-click the Computer icon > This PC > GISData (\lsmb.rdf.rice.edu\research\FondrenGDC) (O:) > GDCTraining > 1_Sh ort_Courses > Data_Interpolation_and_Extraction.
- 2. To create a personal copy of the tutorial data, drag the Interpolation folder onto the Desktop.
- 3. Close all windows.

OPTION 2: Accessing tutorial data online using a personal computer

If you are completing this tutorial from a personal computer, you will need to download the tutorial data online by following the instructions below:

(ii)

Tutorial Data Download

- Interpolation.zip
- 1. Click Interpolation.zip above to download the tutorial data.
- 2. Open the Downloads folder.
- 3. Right-click Interpolation.zip and select Extract All....
- 4. In the 'Extract Compressed (Zipped) Folders' window, accept the default location into the Downloads folder and click Extract.
- 5. Drag the unzipped Interpolation folder onto your Desktop.
- 6. Close all windows.

Opening an Existing Project

- 1. On the Desktop, double-click the Interpolation folder.
- 2. Double-click the Interpolation.aprx project file to open the project in ArcGIS Pro.

Scenario Description

Imagine that you work for an insurance agency and a hurricane has just hit Houston. You have gathered data on rain and wind from the National Oceanic and Atmospheric Agency (NOAA) in the form of rain gauge estimates of rainfall and a feature class depicting the peak average wind speeds when the hurricane landed. Your company has several policy holders claiming that damage was done to their homes during the course of the hurricane. Investigators have gone out to the houses to document the damage, but the company needs to know what the hurricane conditions were like at each house before awarding any claim money. However, none of the houses in question lie directly at a rain gauge or a hurricane wind speed line, so you must interpolate the data with the most accurate interpolation technique and extract the data so that it is in a convenient tabular format for your supervisor.

Preparing Data

Preparing Vector Data

- 1. In the Catalog pane on the right, expand the Folders folder.
- 2. Expand the Interpolation folder.
- 3. Expand the Interpolation.gdb geodatabase.
- Right-click the Hurricane feature class and select Add To New Map.

The rings represent the particular wind speed readings for the hurricane.

1. In the Contents pane, right-click the Hurricane layer name and select Attribute Table.

Notice that there is a *Windspeed* attribute that tells you the peak average wind speed along each ring when the hurricane landed. The wind speeds varied from 95 to 120 miles per hour.

1. Close the Hurricane table view.

Preparing Tabular Data

1. In the Catalog pane, double-click the Houses.xls Excel file to view its individual worksheets.

Remember that ArcGIS Pro can only add one worksheet within an Excel file at a time, so you will need to specify which worksheet you wish to add. In most cases, you will be interested in the first worksheet.

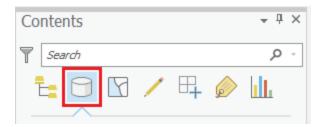
1. Under the Houses.xls Excel file, right-click the Sheet1\$ worksheet and select Add To Current Map.

Notice that a new section called 'Standalone Tables' has appeared at the bottom of the Contents pane, because the tabular data you just added does not have a spatial component, so it cannot yet be drawn and will not appear in the map view.

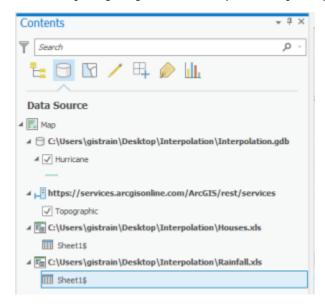
- 1. In the Catalog pane, double-click the Rainfall.xls Excel file to expand it and view its individual worksheets.
- 2. Under the Rainfall Excel file, right-click the Sheet1\$ worksheet and select Add To Current Map.

Because both of the individual Excel worksheets were not renamed, they are both listed as Sheet1\$ in the Table of Contents.

1. At the top of the Contents pane, click the List By Data Source button.



- 2. If you cannot see the entire file paths, hover over the right edge of the Contents pane until your cursor turns into a double-sided arrow.
- 3. Click and drag the right edge of the Contents pane to the right until you can see the full file paths.



Now you can see the name of the Excel file from which each *Sheet1\$* originated, but to avoid further confusion, you will rename the Excel worksheet layers. This action will not rename the worksheets in the actual Excel file, but only how they are listed in the Contents pane in this particular Map.

- 1. Underneath the *Houses.xls* file path, click the **Sheet1\$** table to select it.
- 2. Click directly on the Sheet1\$ table name again to enable renaming and type "Houses".
- 3. Underneath the Rainfall.xls file path, click the Sheet1\$ table to select it.
- 4. Click directly on the Sheet1\$ table name again to enable renaming and type "Rainfall".

Now you will examine what data is contained in each of the two tables.

1. Right-click the Houses table and select Open.

Note that the Houses table includes the longitude and latitude of each house that has filed an insurance claim.

- 1. Close the Houses table view.
- 2. Right-click the Rainfall table and select Open.

Note that the Rainfall table includes the longitude and latitude of each rain gauge in the hurricane area, as well as the inches of rainfall collected in each gauge during the hurricane.

1. Close the Rainfall table view.

Mapping XY Data

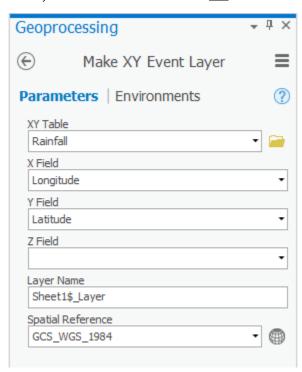
Both the *Houses* and *Rainfall* tables include coordinates for the longitude and latitude of each data point; however, in order to display this tabular data in a graphic format, you will need to map these points based on their XY coordinates.

1. In the Contents pane, right-click the Rainfall table and select Display XY Data.

Notice that the Geoprocessing pane has opened on the right as a new tab overlaying the Catalog pane. In the 'Make XY Event Layer' tool, notice that 'XY Table' is already set to the *Rainfall* table, since that was the table used to open the tool.

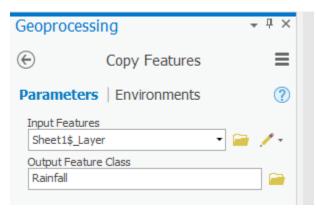
- 1. For 'X Field:' use the drop-down menu to select the Longitude field.
- 2. For 'Y Field:' use the drop-down menu to select the Latitude field.

3. Ensure your window matches that below and click Run.



Your map now shows the location of each rain gauge as a point feature, but the functionality of your data will be limited until you export the data to a new layer. The locations of the points in the *Rainfall* table are currently stored in a temporary layer called *Sheet1\$_Layer*. Since the points appear to be in reasonable locations, you will want to export the temporary layer into a new permanent feature class within your project geodatabase. Exporting to a feature class will allow you to manipulate these points and reuse them in other future maps without having to go through the display XY data process each time.

- 1. In the Contents pane, <u>right-click</u> the **Sheet1\$_Layer** layer and <u>select</u> **Data** > **Export Features**.
- 2. For 'Output Feature Class:', rename the feature class from Sheet1_Features_CopyFeatures to "Rainfall".
- 3. Ensure your 'Copy Features' window appears as shown below and click Run.



Since you have now created a permanent feature class, you may remove your temporary events layer and the corresponding Excel table.

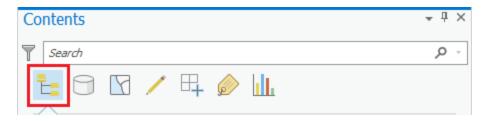
- 1. In the Contents pane, right-click the **Sheet1\$_Layer** layer and select **Remove**.
- 2. In the Contents pane, right-click the Rainfall table and select Remove.

Now you will repeat the same process to create a permanent point feature class for the Houses data table.

Repeat all steps in this Mapping XY Data section using the Houses table. Rename the exported feature class "Houses" and then remove the Sh eet1\$_Layer layer and Houses table from the Contents pane.

Your Map Display now shows the locations of all of the houses filing insurance claims due to the hurricane and you have all the layers necessary for analysis added into your map.

1. At the top of the Contents pane, click the List by Drawing Order button.



 Click and drag the right edge of the Contents pane back to the left, so that it is only wide enough to display the short layer names and the List by... icons.

Interpolating Data

Interpolation allows you to estimate the values of new data points based on the values of existing data points. In this case, you would like to estimate the rainfall values at each of the house locations based on the rainfall values recorded at each of the rain gauge locations. The following section will teach you several methods of interpolating both point and line vector data into raster datasets, which provide estimated values of the data variables across the entire region. You will see how these different interpolation techniques lead to variations in the resulting raster datasets. It is important to select an interpolation technique best suited to your existing data and your desired analysis.

Point Interpolation

First you will interpolate the rainfall data, because point data is usually easier to interpolate to raster data than line or polygon data. In this section, you will interpolate the rainfall data using the Inverse Distance Weighted, Spline, Natural Neighbors, and TIN interpolation methods.

Since you are only interpolating the rainfall data at this point, you will temporarily turn off some of the other layers.

1. In the Contents pane, uncheck the Houses and Hurricane layers so that they are no longer visible.

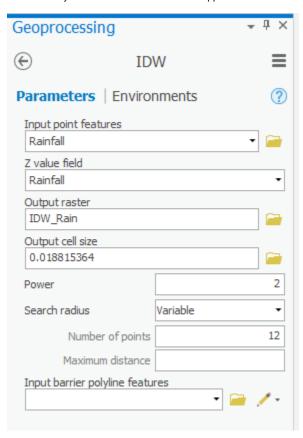
At this point, only the Rainfall and Topographic basemap layers should be visible.

Inverse Distance Weighted Interpolation Method

Typically, you would use the 'Find Tools' search box at the top of the Geoprocessing pane to search for the name of the tool you'd like to use, but, at times, especially when learning the software, it can be helpful to view the full hierarchy of all the tools available, because you will often discover related and helpful tools that you didn't know existed and wouldn't know to search for. You might also completely forget the name of a tool, but be able to locate it based on the hierarchy. For these reasons, we will be manually navigating the toolboxes throughout this tutorial. The more typical workflow of searching directly for a specific tool is covered at the end of the Introduction to Geoprocessing tutorial.

- 1. At the top of the Geoprocessing pane, click the Back arrow button.
- 2. At the top of the Geoprocessing pane, click the Toolboxes tab.
- 3. Click the Spatial Analyst Tools toolbox > Interpolation toolset > IDW (Inverse Distance Weighted) tool.
- 4. For 'Input point features', use the drop-down menu to select the Rainfall layer, which contains your rain gauge point locations.
- 5. For 'Z value field', use the drop-down menu to select the Rainfall field that contains the rainfall measurements you wish to interpolate.
- **6.** For 'Output raster', <u>rename</u> the exported raster from *Idw_Rainfall1* to "*IDW_Rain*".

7. Ensure that your 'IDW' tool Parameters tab appears as shown below and click Run.



It will take several seconds for the tool to begin running, at which point you will see the name of the tool along the bottom portion of the Geoprocessing pane. When the tool has finished running, a box will pop up in the same corner with the name of the tool and a green checkmark.

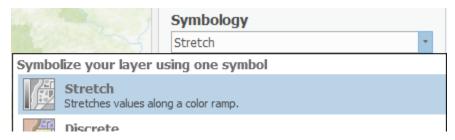
You can no longer see the *Topographic* basemap, because it is currently drawn beneath the new *IDW_Rain* raster layer. To make them both visible, you will make the raster layer transparent.

- 1. In the Contents pane, click the IDW_Rain layer to select it.
- 2. In the ribbon on the top, click the Appearance tab
- 3. In the Effects group, adjust the Layer Transparency slider to 30%.

Notice that you can now see the basemap beneath the raster layer.

The raster layer is currently symbolized using nine classes; however, the values of the cells in the raster are actually more continuous in nature, so you will symbolize them continuously, rather than in classes.

- 1. In the Contents pane, right-click the IDW_Rain layer and select Symbology.
- 2. For 'Symbology,' use the drop-down menu to select Stretch.



- 3. Click the drop-down Color scheme menu. Select the Show names check box at the bottom.
- 4. Scroll down the color scheme list and select the Blues (Continuous) color ramp.



You can now see the continuously interpolated raster containing rainfall values that were generated using the data collected at the specific rain gauge points. Notice how this IDW method results in distinct rings around each input data point, because it treats each data point as a local max or min.

1. In the Contents pane, uncheck the IDW_Rain layer.

Spline Interpolation Method

- 1. At the bottom of the Symbology pane, click the Geoprocessing tab.
- 2. At the top of the Geoprocessing pane, click the **Back** arrow button.
- 3. Within the Interpolation toolset, <u>click</u> the **Spline** tool.
- Beginning with step 5, repeat all steps in the 'Inverse Distance Weighted Interpolation Method' section above, but rename the output raster "Splin e Rain".

Compare the result of the spline interpolation to the result of the IDW interpolation by turning each layer on and off in the Contents pane. Since the layers are transparent, you need to ensure that only one raster is turned on at a time. Notice how spline interpolation tries to make the data surface as smooth as possible and does not exhibit the distinct rings around the input points, as the IDW method does. The spline interpolation process assumes that, overall, there is a smooth transition in the data and tries to mirror that effect in its output. Also, notice how the spline method extends any trends that have begun within the region covered by data points way out past the data points, causing extremely high rainfall values in the Gulf and near San Antonio that are far outside the numeric range of the original rain gauge data. The IDW and natural neighbors methods, do not generate values outside the range of the original data.

1. In the Contents pane, uncheck the Spline_Rain layer and any other raster layers you turned on.

Natural Neighbors Interpolation Method

- 1. At the bottom of the Symbology pane, click the Geoprocessing tab.
- 2. At the top of the Geoprocessing pane, click the Back arrow button.
- 3. Within the Interpolation toolset, click the Natural Neighbor tool.
- 4. Beginning with step 5, repeat all steps in the 'Inverse Distance Weighted Interpolation Method' section above, but rename the raster output "NN_Rain"

Compare the result of the natural neighbors interpolation to the results of the previous two interpolation methods. It is somewhat in between the two, in that it is smoother than IDW, but not as smooth as spline. While there are not distinct rings around each data point, as there are with IDW, the data points are clearly still creating local highs and lows. Note also that the interpolation is limited to the area covered by the input points, rather than a rectangle of the same extent.

1. In the Contents pane, uncheck the NN_Rain layer and any other raster layers you turned on.

TIN Interpolation Method

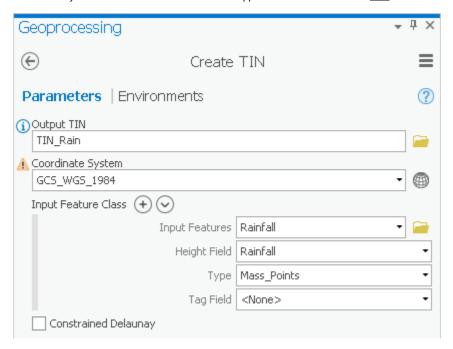
Using the Create TIN tool, you can create a TIN based on an existing feature class.

- 1. At the top of the Geoprocessing pane, <u>click</u> the **Back** arrow button.
- 2. Collapse the Spatial Analyst Tools toolbox.
- 3. Click the 3D Analyst Tools toolbox > Data Management toolset > TIN toolset > Create TIN tool.
- 4. For 'Output TIN', type "TIN_Rain".
- 5. For 'Coordinate System', use the drop-down menu to select the **Houses** layer, to ensure that the interpolated data will be in the same coordinate system as the *Houses* layer, which contains the points at which we need to extract the interpolated values.

As you will see in the warning, it is recommended to only create a TIN in a projected coordinate system, rather than a geographic coordinates system. The process of projecting a layer from one coordinate system to another is outside the scope of this course, but is covered in the Introduction to Coordinate Systems and Projections course.

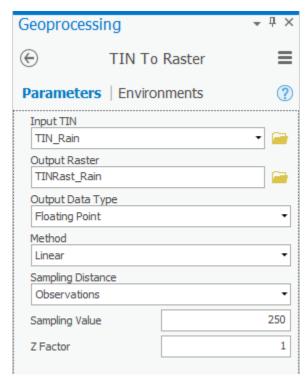
- 1. For 'Input Features', use the drop-down menu to select the *Rainfall* layer.
- 2. For 'Height Field', use the drop-down menu to select the *Rainfall* field.

3. Ensure that your 'Create TIN' tool Parameters tab appears as shown below and click Run.



This method interpolates the rainfall data points by defining the vertices of each triangle in the TIN and interpolating across each triangle. From here, it is very simple to convert the TIN to a raster, so that it is in the same format as your other interpolation layers.

- 1. At the top of the Geoprocessing pane, $\underline{\text{click}}$ the Back arrow button.
- 2. Within the 3D Analyst Tools toolbox, collapse the Data Management toolset.
- 3. Within the 3D Analyst Tools toolbox, click the Conversion toolset > From TIN toolset > TIN To Raster tool.
- 4. For 'Input TIN', use the drop-down box to select the TIN_Rain TIN.
- 5. For 'Output raster', rename the exported raster from tin_rain_tin to "TINRast_Rain".
- 6. Ensure that your 'TIN To Raster' tool Parameters tab appears as shown below and click Run.



- 7. In the Contents pane, $\underline{\text{uncheck}}$ the $\textit{TIN_Rain}$ layer.
- 8. Repeat the steps within the 'Inverse Distance Weighted Interpolation Method' section above to symbolize the TINRast_Rain layer as you did the other raster layers.

Compare the results of the TIN interpolation to the results of the previous three interpolation methods. Note that the triangulation method results in defined and continuous ridges and valleys, which makes sense, because TIN files are most commonly used to represent land elevation.

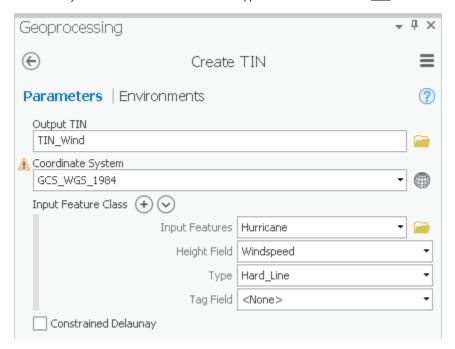
Now you have explored four methods for interpolating the rainfall data, so you will turn off the layers you have generated in this 'Point Interpolation' section.

1. In the Contents pane, uncheck all layers, except for the Topographic basemap layer.

Line Interpolation

You will now try interpolating the *Hurricane* line feature class using the same TIN method you just used to interpolate the rainfall point data. Recall that the *Hurricane* feature class contains a *Windspeed* field in its attribute table, which tells you the peak average wind speed along each line at the time the hurricane landed

- 1. In the Contents pane, check the Hurricane layer to make it visible.
- 2. At the top of the Geoprocessing pane, click the **Back** arrow button.
- 3. Within the 3D Analyst Tools toolbox, collapse the Conversion toolset.
- 4. Within the 3D Analyst Tools toolbox, click the Data Management toolset.
- 5. The TIN toolset should already be expanded, but, if not, click the TIN toolset.
- 6. Click the Create TIN tool.
- 7. For 'Output TIN', type "TIN_Wind".
- 8. For 'Coordinate System', use the drop-down menu to select the *Houses* layer.
- 9. For 'Input Features', use the drop-down menu to select the Hurricane feature class.
- 10. For 'Height Field', use the drop-down menu to select the Windspeed field.
- 11. Ensure that your 'Create TIN' tool Parameters tab appears as shown below and click Run.



Again, you need to convert the TIN to a raster, so that it is comparable to the other interpolation layers.

- 1. Repeat steps 1-8 regarding TIN to raster conversion in the 'TIN Interpolation Method' section above, but select *TIN_Wind* for the input TIN and name the output raster "*TINRast_Wind*" and uncheck the *TIN_Wind* layer.
- 2. In the Contents pane, click the arrow next to all of the raster and TIN layers to collapse their symbologies.

Extracting Data

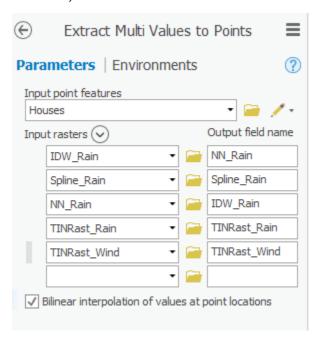
All of the raster interpolations you have created for wind and rain provide continuous value estimates over the entire hurricane area, but you are only interested in the data values at the particular locations of the individual houses filing insurance claims. You will extract point data for each of the house locations from each of the five raster layers you just created in a database-compatible format for the insurance company to use. Extracting data is one of the most important aspects of data manipulation, since interpolation is often useless if the results cannot be taken out of GIS and used for other procedures.

Extracting Multi Values to Points

- 1. At the top of the Geoprocessing pane, click the **Back** arrow button.
- 2. Collapse the 3D Analyst Tools toolbox.
- 3. Click the Spatial Analyst Tools toolbox > Extraction toolset > Extract Multi Values to Points tool.
- 4. For 'Input point features', use the drop-down box to select the *Houses* layer.
- 5. For 'Input rasters', use the drop-down box to <u>select</u> all five interpolation rasters you created earlier one at a time: *IDW_Rain*, *Spline_Rain*, *NN_Rain*, *TINRast_Rain* and *TINRast_Wind*.

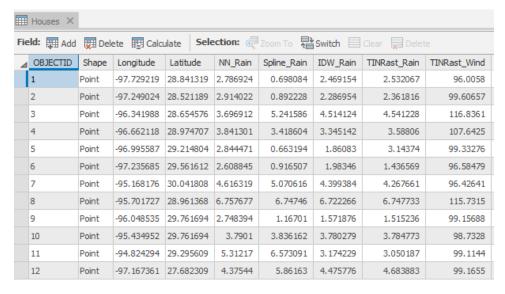
The order in which the rasters are listed is the same order in which the extracted value columns will appear in the output feature class attribute table.

- 1. Check Bilinear interpolation of values at point locations to further interpolate the value, if the house point does not lie exactly at the center of a raster cell
- 2. Ensure that your 'Extract Multi Values to Points' tool Parameters tab appears as shown below and click Run.



The Extract Multi Values to Points tool simply adds a field for the value of each raster to the end of your existing *Houses* attribute table, which is why you were not asked for an output location.

1. In the Contents pane, right-click the Houses layer and select Attribute Table.



Notice the five new fields added to the end of the existing attribute table which contain the values of each of the five rasters at each house location. Notice also the differences in rainfall estimates resulting from the four different interpolation techniques. At many house locations, the value of the spline interpolation is the most different from the other interpolation method values.

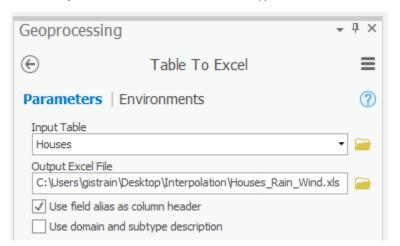
You are now ready to export the data table into a format that can be used by the insurance company, without having to use ArcGIS software.

- 1. At the top of the Geoprocessing pane, click the Back arrow button.
- 2. <u>Click</u> the Conversion Tools toolbox > Excel toolset > Table To Excel tool.
- 3. For 'Input Table', use the drop-down menu to select the Houses attribute table.

The output table location defaults to your project folder.

- 1. For 'Output Excel File', rename the Excel file from Houses_TableToExcel to "Houses_Rain_Wind".
- 2. Check Use field alias as column header.

3. Ensure that your 'Table To Excel' tool Parameters tab appears as shown below and click Run.



Now you have what the insurance company is looking for: a single text file containing the locations of the houses in question, as well as the corresponding rainfall and wind speed estimates to use in completing a damage quote for your customers.

1. Close the Houses table view.

Note that if you simply wanted a table containing the values of multiple rasters at each point and did not care about symbolizing or working with the points any further in the ArcGIS Pro environment, you could have used the Sample tool, which outputs a separate table, rather than adding the values to the end of your existing feature class attribute table.

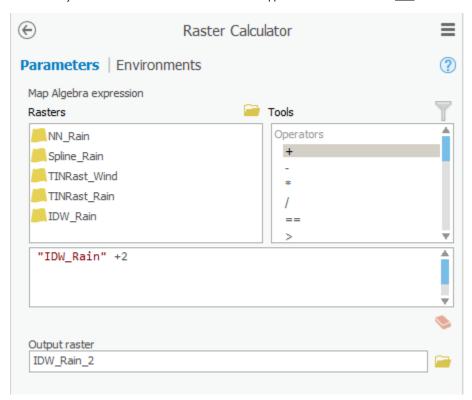
Raster Calculations

Your final project involves working with the raster calculator. The raster calculator performs calculations on a cell-by-cell basis and outputs the calculation results in a new output raster. You will perform a simple calculation using the raster calculator, but the raster calculator can perform many functions, such as adding or subtracting one raster from another, multiplying a raster by a constant, or removing a predetermined "null" value.

For this example, you have learned that the rainfall data collected from the rain gauges had a common discrepancy. Each gauge misrepresented the rainfall amount by two inches uniformly across the state. There are several ways you could approach this problem, but you will add two inches uniformly across the interpolated rain raster to practice using the raster calculator.

- 1. At the top of the Geoprocessing pane, <u>click</u> the **Back** arrow button.
- 2. Within the Spatial Analyst Tools toolbox, <u>click</u> the **Map Algebra** toolset > **Raster Calculator** tool.
- 3. Under 'Rasters', double-click the IDW_Rain layer.
- 4. Under 'Tools', double-click the + operator.
- 5. Type "2".
- **6.** For 'Output raster', <u>rename</u> the exported raster from *idw_rain_ras* to "*IDW_Rain_2*".

7. Ensure that your 'Raster Calculator' tool Parameters tab appears as shown below and click Run.



While the original *IDW_Rain* layer and the new *IDW_Rain_2* layer will look the same visually, notice in the Contents pane, next to the continuous color ramps for the two layers, that the max and min values for the *IDW_Rain_2* layer are now two higher than the max and min values for the original *IDW_Rain* layer.

You have now learned multiple methods of interpolating point and line data into raster data and extracting that raster data at corresponding point locations. You have also reviewed skills from other short courses, including Mapping Locations Using XY Coordinates.