Exercise 5: Georeferencing and Resampling

Occasionally you will need to use a scanned paper map as a background for digitizing or other work. First, you will need to scan your map into a suitable image file format, such as TIFF, JPEG, or IMG. Once the map exists as a digital file, you can load it into ArcGIS Pro as a raster layer. However, this map or image will not draw in the correct location relative to the map datum and/or projection because it does not have spatial reference information. In order to get the image to draw at the correct location, you will need to georeference it relative to reference data. This exercise shows you how to undertake this process. We will guide you through referencing a map of the West Virginia University downtown campus. You will then apply these techniques to georeference a historical map of Venice, Italy. You will validate your results using the root mean square error measure (RMSE).

Another common pre-processing task is changing the cell or pixel size of a raster grid, including aerial and satellite images. We will also show you how to perform resampling in ArcGIS Pro using a variety of methods.

Topics covered in this exercise include:

1. Calculate root means square error (RMSE)
2. Georeference an image using control points and reference data
3. Assess the quality of a georeferenced result using RMSE
4. Perform raster resampling

Step 1. Calculating RMSE

Root mean square error (RMSE) can be used to estimate the positional accuracy of geospatial data, including the results of a georeferencing process. To calculate this measure, the predicted X and Y coordinates are compared to the correct X and Y coordinates at point locations.

Here is the equation.

$$RMSE_x = \sqrt{} \ \Sigma(x_{predicted} - x_{actual})^2/n$$

$$RMSE_y = \sqrt{} \ \Sigma(y_{predicted} - y_{actual})^2/n$$

$$RMSE = \sqrt{} \ (RMSE_x^2 + RMSE_y^2)$$

Note that the RMSE in the X and Y directions are calculated separately and then combined to obtain the total RMSE. Calculating RMSE requires the following steps.

1. The predicted and actual values are differenced then squared.
2. The square differences are added together.
3. You then divide by the number of samples (n).
4. You take the square root to obtain the RMSE.
5. This process is completed for both the X and Y data.
6. To obtain the total RMSE you add the square of $RMSE_x$ and $RMSE_y$ then take the square root.

RMSE is in the units of the input measurements. For example, if X and Y coordinates represent UTM coordinates in meters as easting and northing then the final RMSE measure will be in meters.

☐ Calculate $RMSE_x$, $RMSE_y$, and total RMSE for the data provided below. You can do this by hand. If you'd prefer to perform the calculations using Microsoft Excel, you can download the **rmse_data.xlsx** file from http://www.wvview.org/. Under Education then "Labs and Data" tab. Click on "L11 Data" to start downloading the **Exercise_5** folder. Note that these represent easting and northing measurement relative to a UTM projection. The units are meters.

Reference X (easting)	Reference Y (northing)	Predicted X (easting)	Predicted Y (northing)
513,128.4	4,101,292.6	513,127.3	4,101,293.4
513,117.8	4,101,301.4	513,118.1	4,101,303.1
513,201.7	4,101,208.1	513,203.1	4,101,209.2
513,180.3	4,101,375.8	513,181.6	4,101,373.9
513,215.6	4,101,401.6	513,216.1	4,101,402.2

**Question 1.** What is the $RMSE_x$? (5 Points)

Question 2. What is the $RMSE_y$? (5 Points)

Question 3. What is the total RMSE? (5 Points)

Step 2. Copy Data

The two images that you will need to georeference are also provided in the **Exercise_5** folder downloadable from **http://www.wvview.org/**. The first image you will georeference is an image of the West Virginia University downtown campus (**DowntownCampusMap.jpg**). The second image you will georeference is a historic map of Venice, Italy (**venice2.jpg**). So that

the original files are not altered, you will georeferenced copies of the files as opposed to the originals. We recommend georeferencing from a copy.
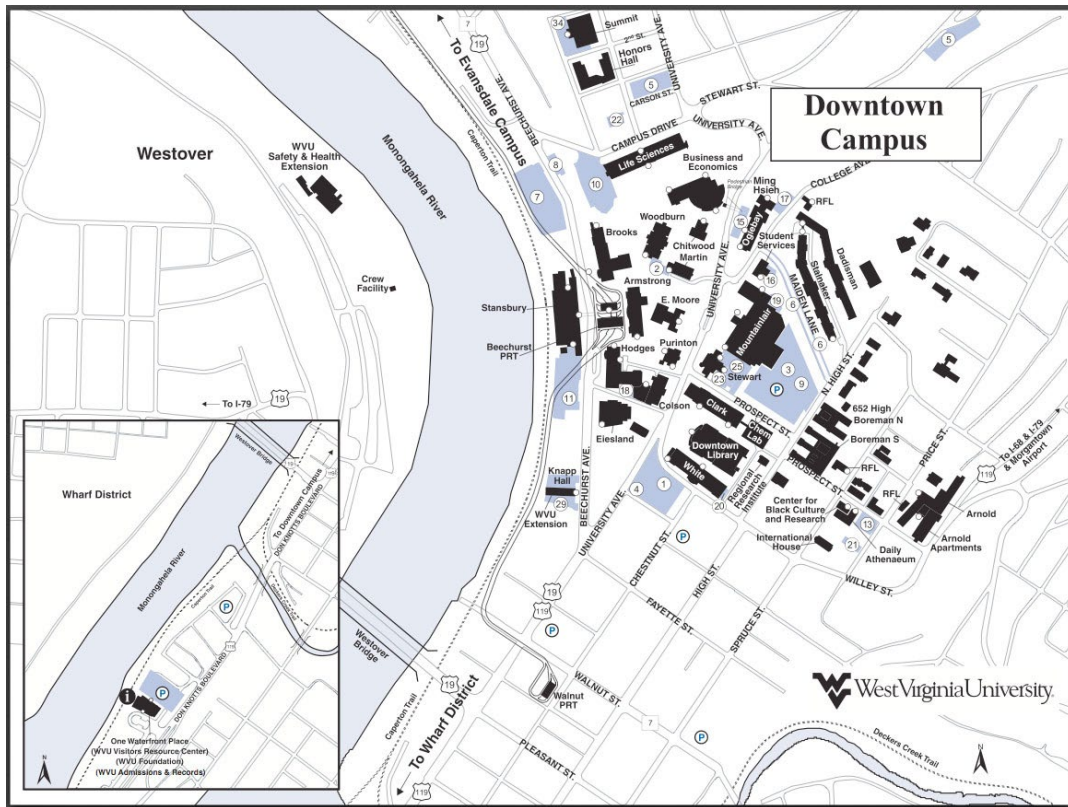
☐ Create a new folder within the **Exercise_5** folder and call it **Working.**
☐ Copy the two images in the **Exercise_5** folder to the newly created **Working** folder. Use these copies of the images in this exercise, not the originals. If something goes wrong you still have the originals.

Step 3. Open a Map Project

First, we need to download and open the **Exercise_5.aprx** file.

☐ Download the **Exercise_5** data from https://www.wvview.org/. All lab materials are available on the course webpage and linked to the exercise. You will need to extract the compressed files and save it to the location of your choosing.

☐ Open ArcGIS Pro. This can be done by navigating to All Apps followed by the ArcGIS Folder. Within the ArcGIS Folder, select ArcGIS Pro. Note that you can also use a Task Bar or Desktop shortcut if they are available on your machine.
☐ After ArcGIS Pro launches, select "Open another project."

☐ Navigate to the directory that houses the material for this course. The project files are in the **Exercise_5** folder where it was saved on your local machine.
☐ Select **Exercise_5.aprx**. Click OK to open the project.
☐ If necessary, navigate to the WV map.

Note: If you'd prefer, you can also just click on the **Exercise_5.aprx** file within the uncompressed folder directly to launch ArcGIS Pro.

This project should contain three maps: WVU, Venice, and Resample. The WVU map contains one layer, **morgantow_north_sw.sid**, which will serve as your reference data to perform the georeferencing. This is a 2 ft. spatial resolution image from 2003 made available by the West Virginia Statewide Addressing and Mapping Board (WVSAMB). The second map, Venice, contains one layer, **Venice_Bound**, which is a vector bounding box around the island. You will use the ESRI Basemap to reference this image. Lastly, the Resample map contains two layers: **charleston_points** and **charleston.tif**. You will use these data to investigate raster resampling.

☐ Navigate to the WVU map.

## Step 4. Georeference the WVU Campus Map

When performing a georeference, the goal is to use reference or source data and control points to reference an image relative to a datum and/or projection. Here, you will use the aerial photograph to reference the campus map. The first step is to add the unreferenced data.

☐ Use the Add Data button to add the **DowntownCampusMap.jpg** layer to the map. Make sure to use your copy of the data that is in the **Working** folder you created under the **Exercise_5** folder.
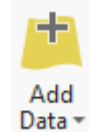
You will notice that the map does not draw at the correct location. This is because it has not been georeferenced yet.

☐ Right-click on the **DowntownCampusMap.jpg** layer in the Contents Pane then select Zoom to Layer. This will zoom the map to the extent of the unreferenced image. Your goal is to change the reference so the map draws at the correct location.
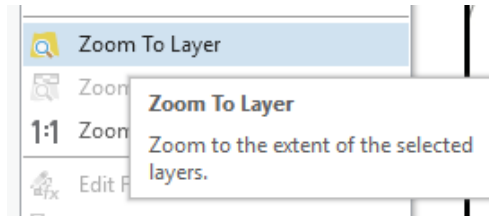
Here are some tips for georeferencing.

1. The reference data should be of high quality. High spatial resolution images are preferred, so that details can be distinguished. Positional errors in the reference image will impact the quality of the georeference. So, you should use reference data that has been referenced well.
2. You will need multiple control points. The number of control points needed depends on the algorithm being used. Here, you will need to produce at least four control points.
3. You must be able to find and identify the control points in both the reference data and the data being georeferenced. For example, you could use a road crossing or sidewalk intersection identifiable in both images.
4. Avoid using above ground features as control points, such as building corners. This is because buildings often lean in orthophotography due to vertical relief displacement. So, the true position is distorted.
5. The control points should be spread across the unreferenced image as opposed to clustered. This is so that the algorithm can do a decent job georeferencing the whole image.

Please keep these tips in mind as you select control points.

You will now prepare to collect control points.

- ☐ Make sure the **DowntownCampusMap.jpg** image is selected in the Contents Pane. This is so that the software will know that this is the image to be georeferenced.
- ☐ Under the Image Tab, select Georeference. This will activate the Georeferencing Tab.
- ☐ Click on the Control Point Table button to open the table. It will be empty, because you haven't added control points yet.
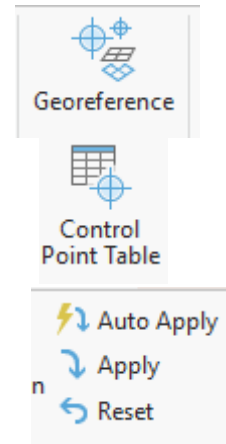
Note: If the Auto Apply option is selected, the image will be updated as you add control points. We would suggest not using this option. So, make sure that it is deselected. If you want to update the image as you add control points you can use the Apply button. If you want to remove the last control point, you can use the reset option.

In ArcGIS Pro, you must select the location in the unreferenced image first then select the same location in the reference data. The location in the unreferenced data is called the From Point and the associated point in the reference data is the To Point. Together, they create a control point.

Control points are added using the Add Control Point button.
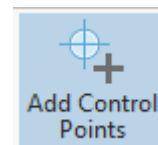
Below is an example of a From Point and To Point.

When georeferencing in ArcGIS Pro, you will need to navigate between the unreferenced data and the reference data. This can be accomplished by right-clicking on the layer in the Contents Pane then selecting Zoom to Layer. You can then navigate the map space using the mouse or the Explore tools under the Map Tab. Once you've established a control point, it will be added to the Control Point Table. If you would like to delete a control point, you can select it then use the Delete Selected button in the table.
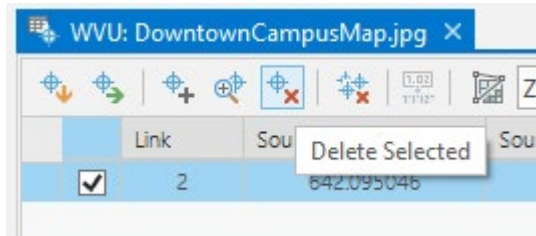
☐ Create at least four control points scattered around the map. If you would like to wait until you have created all of the control points for the image to be updated, make sure the Auto Apply option is not selected. If you select Apply, you can update the image. If you would like to not use one of your control points you can either delete it from the table or you can deselect it in the table. As you create your control points, make sure that you use the tips provided above. Also, make sure that they can be located in both images.

Below is an example of six control points displayed over the reference data and the associated table.
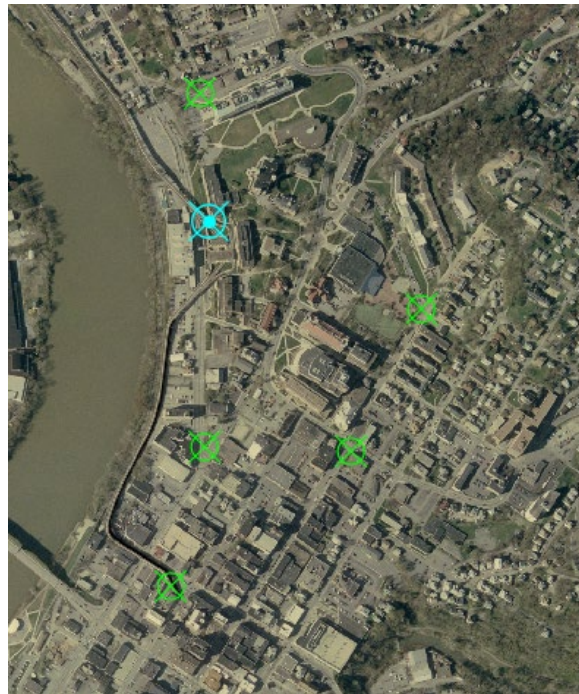
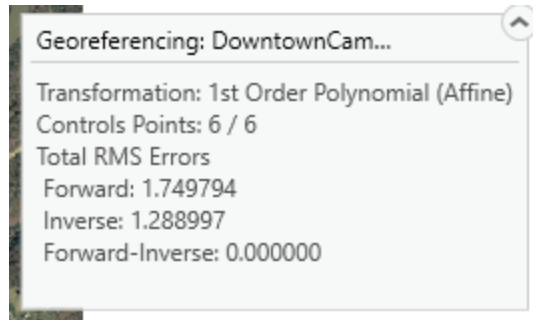|	Link	Source X	Source Y	X Map	YMap	Residual X	Residual Y	Residual
✓	2	642.095046	-304.056601	589,558.042166	4,387,806.527743	588,915.947120	4,388,110.584344	4,427,452.596381
✓	3	591.781440	-727.549348	589,497.725870	4,387,231.508718	588,905.944430	4,387,959.058066	4,427,301.086062
✓	4	889.943736	-409.666139	589,892.546175	4,387,669.548631	589,002.602439	4,388,079.214769	4,427,433.032895
✓	5	634.244094	-565.724327	589,550.818827	4,387,450.899573	588,916.574733	4,388,016.623899	4,427,359.554363
✓	6	636.229751	-155.838320	589,543.704963	4,388,008.172735	588,907.475212	4,388,164.011055	4,427,504.421486
✓	7	803.933326	-570.431343	589,781.692548	4,387,445.060220	588,977.759223	4,388,015.491563	4,427,366.571118

Once you have established control points, you are ready to perform the georeferencing. In this exercise, you will use the 1st Order Polynomial (Affline) method, which requires at least three control points.

☐ In the Georeferencing Tab and in the Adjust area, select the Transform Button then select 1st Order Polynomial (Affline) from the list.
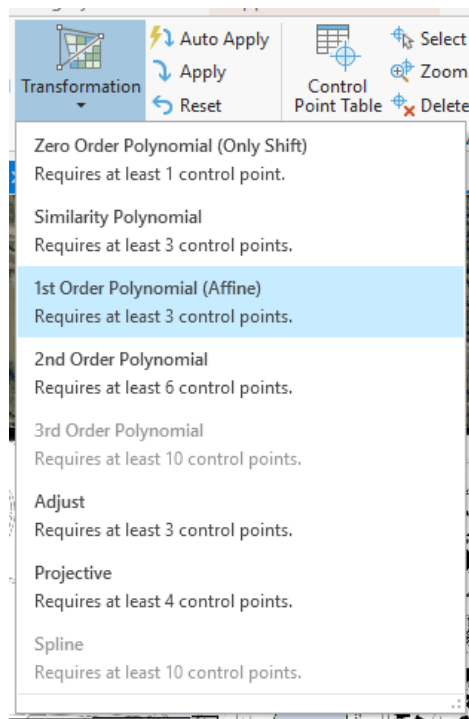
☐ In the Control Point Table, make sure all of the control points are checked. Select Apply in the Adjust area of the Georeferencing Tab.

The Georeferencing pop-up window provides an estimate of the RMSE using the control points as the forward RMSE.

Georeferencing: DowntownCam...

Transformation: 1st Order Polynomial (Affine)
Controls Points: 6 / 6
Total RMS Errors
  Forward: 1.749794
  Inverse: 1.288997
  Forward-Inverse: 0.000000

Question 4. What forward RMSE did you obtain for this georeferencing operation? (5 Points)

Let's save your georeferecning results using Save as New.

- ☐ Click the Save as New button in the Save area of the Georeferencing Tab. This will open the **Save As Tool**.
- ☐ Set the Output Location to the **Working** folder under the **Exercise_5** folder.
- ☐ Name the output **wvu_referenced.tif**.
- ☐ Make sure the Format is set to TIF.
- ☐ You do not need to change any of the other settings.
- ☐ Click OK to run this process.
- ☐ Once the tool executes, you will need to add the output to your map. Under the Map Tab, select Add Data. Navigate to the **Working** folder under the **Exercise_5** folder.
- ☐ Note, if you don't see the file, you may need to refresh the folder with the refresh button in the corner of the window.

Our result is shown below.
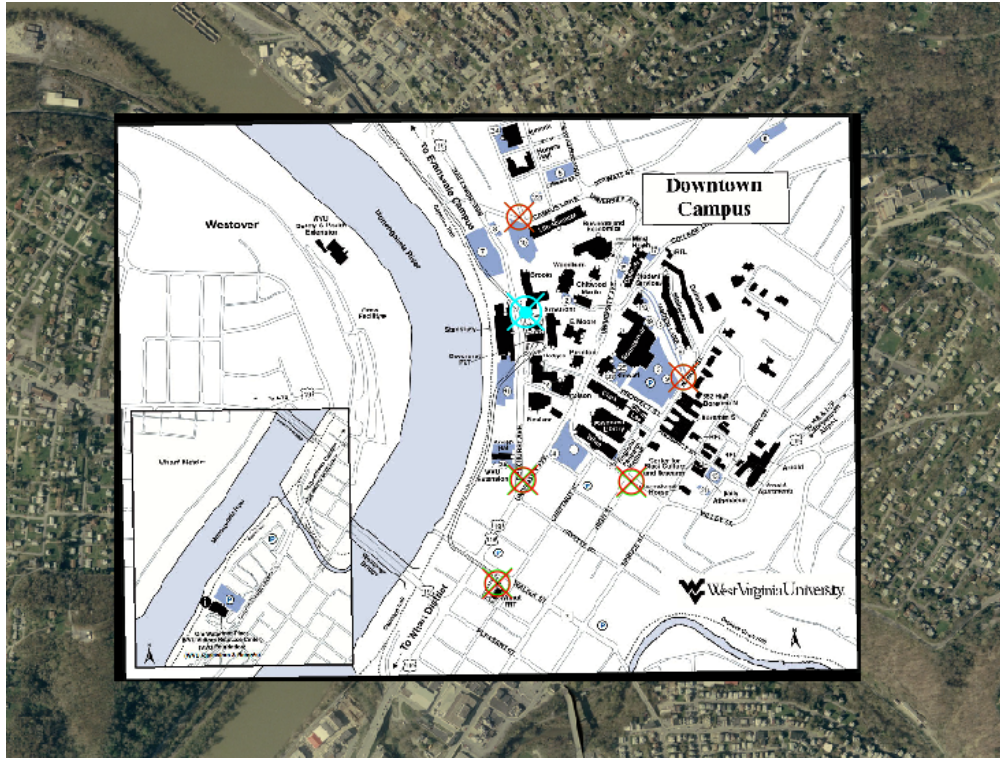
You are now done georeferencing this image. You now need to close the georeferencing session.

☐ In the Georeferencing Tab, click Save in the Save area followed by Close Georeferencing in the Close area.

Note: Clicking Save will apply the georeference to the **DowntownCampusMap.jpg** image that is stored in the **Working** folder under the **Exercise_5** folder. You can save your georeferenced this way as opposed to using Save as New as we did above. However, this will alter the original file. That is why it is best to work with a copy of the data.

**Step 5. Georeference the Venice Map**

You are now done working with the WVU Map.

☐ Navigate to the Venice Map.
☐ Add the **venice2.jpg** image from your **Working** folder in the **Exercise_5** folder. Similar to before, the image will not draw in the correct location because it has not been georeferenced yet.

You will now perform a georeferencing operation for the historic Venice map. Use the same process as that described above. You will need to create at least four control points. In this portion of the exercise, you will use the ESRI Image basemap as your source layer. Since this basemap covers the entire globe, The Zoom to Layer option will zoom to the entire globe. To alleviate this issue, we have provided a rectangular extent around the island as the **Venice_Bound** layer. You can zoom to the extent of this layer to return to the reference data.

☐ Perform the georeferencing. Make sure to collect at least four control points spread over the island. Use the 1st Order Polynomial (Affline) method.

Below is an example of the control points we created.
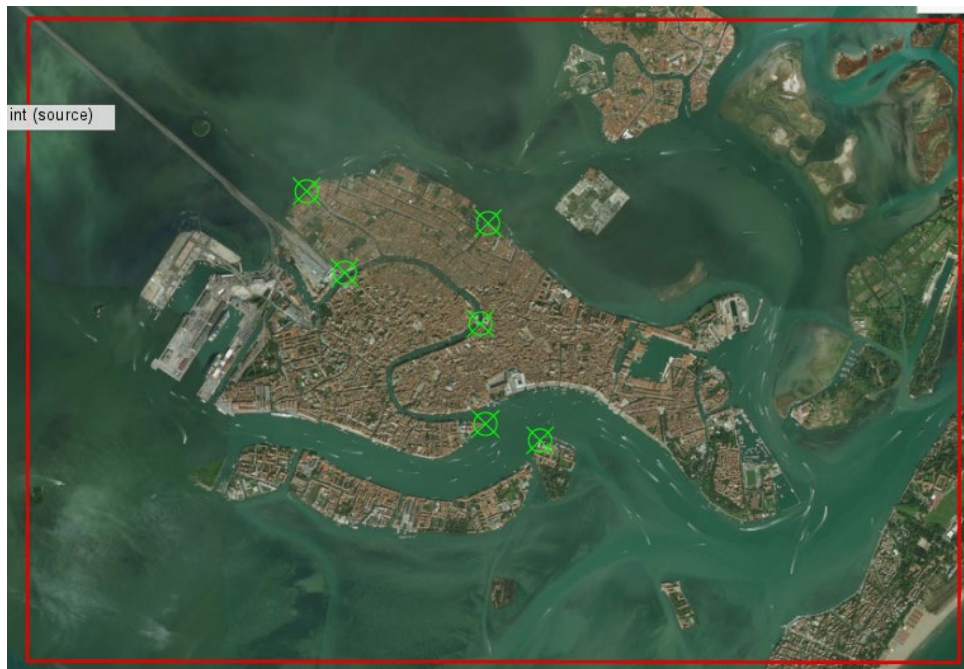
Deliverable 1. Provide a screen capture of your result that shows how the historic map has been referenced to the base imagery. (10 Points)

Question 5. What forward RMSE did you obtain for this georeferencing operation? (5 Points)

Step 6. Resampling Methods

Another common raster pre-processing task is resampling, or changing the cell size. Generally grids are coarsened as opposed to resampling to a smaller cell size. If you resample to a smaller cell size you aren't generally

gaining information but duplicating values. There are many reasons why you might resample a raster grid, such as to meet the needs of a mapping project, to align and have the same cell size as other raster grids in an analysis or model, to reduce processing time, and/or to decrease file size.

There are many methods available to resample a raster grid. Please answer some general questions about the resampling methods here.

Question 6. Why are bilinear interpolation and cubic convolution not appropriate for resampling categorical raster data, like land cover? (5 Points)

Question 7. In cubic convolution, how many original cell values are used to determine the new cell value? (2 Points)

Question 8. In bilinear interpolation, how many original cells values are used to determine the new cell value? (2 Points)

Question 9. Explain the concept of an inverse distance weighted average, as is used to perform bilinear interpolation. (5 Points)

Question 10. Explain how majority resampling works. (5 Points)

Let's explore different resampling methods in ArcGIS Pro.

☐ Navigate to the Resample Map.

In this portion of the exercise you will resample this aerial photograph using different resampling methods. This is a National Agriculture Imagery Program (NAIP) image for a portion of Charleston, West Virginia collected during the growing season of 2016. It is a true color image with red, green, and blue bands and with a 1 meter spatial resolution.
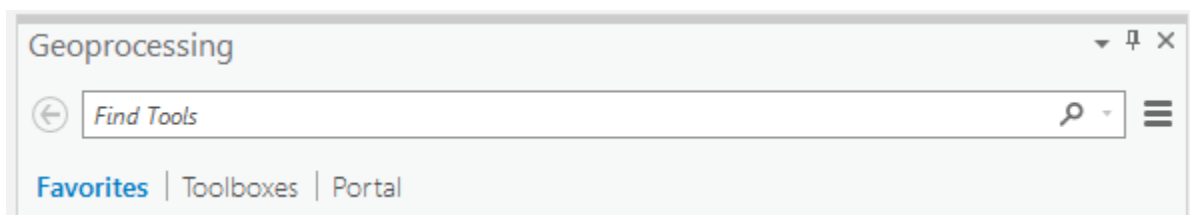
Resampling in ArcGIS Pro can be performed using the **Resample Tool**.
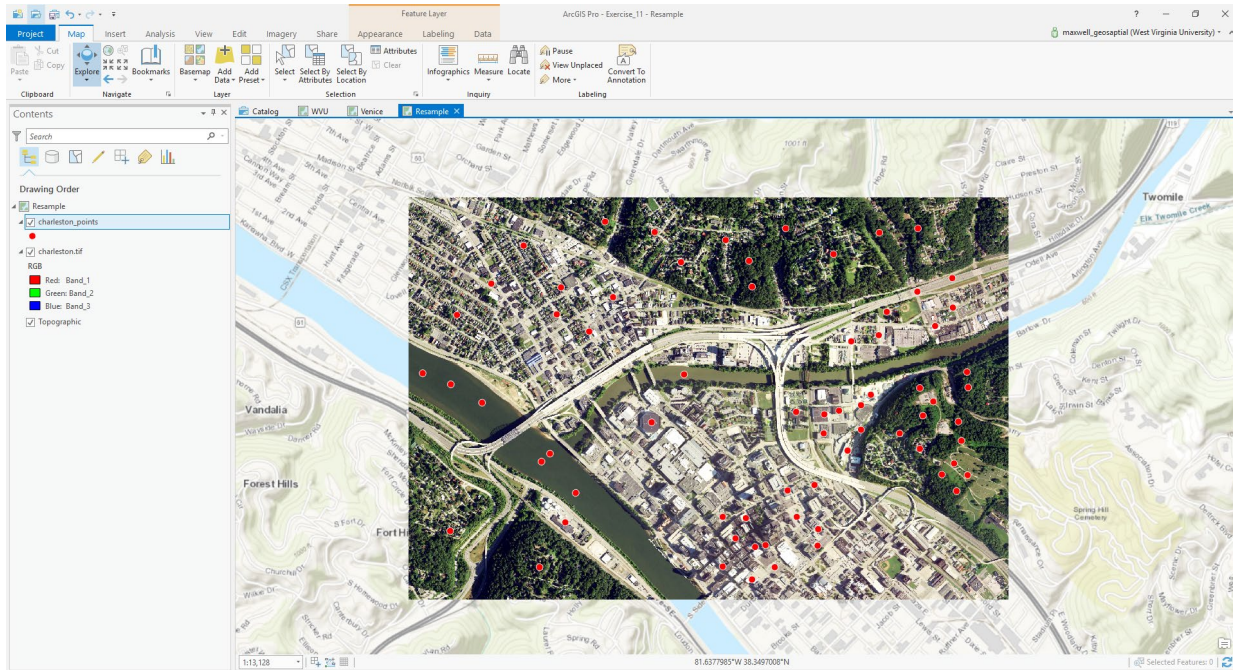
Note: In this lab, we will access tools from ArcToolbox. However, there are many ways to access tools in ArcGIS Pro. For example, some of the more common tools are provided in the Tools list in the Analysis Tab.
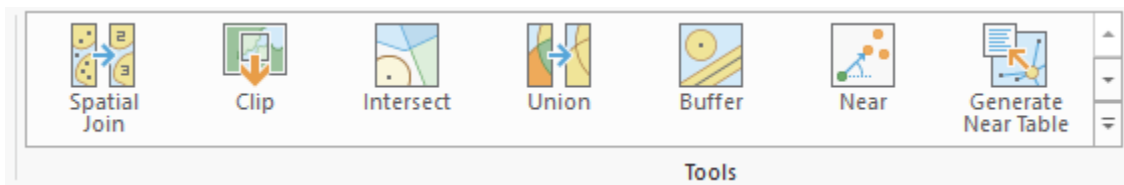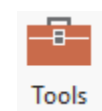

Once you open the Geoprocessing Pane, you can access favorite tools or search for Tools.

We have decided to demonstrate ArcToolbox here so that you get a sense of where the tools are located in the Toolbox directory.

☐ In the Analysis Tab, select Tools from the Geoprocessing Area. This should open the Geoprocessing Pane.
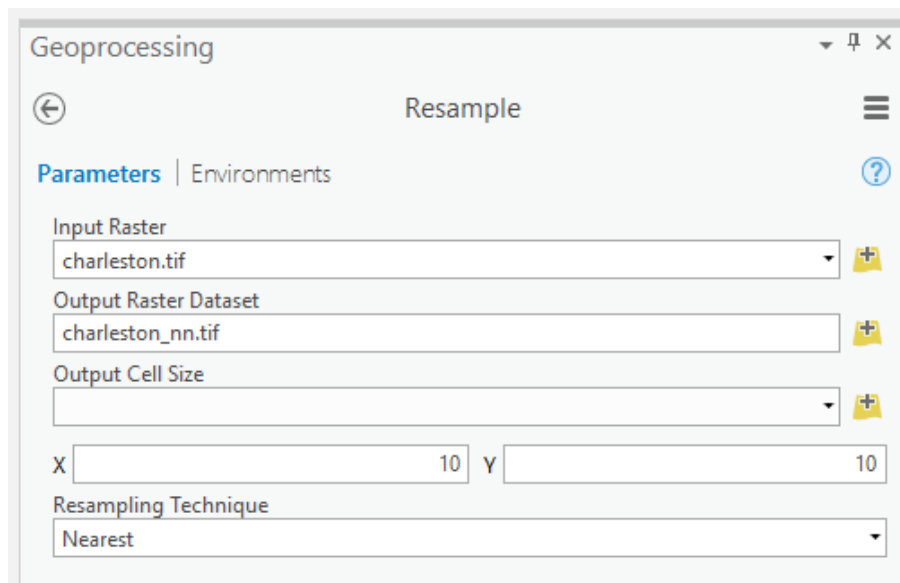
- In the Geoprocessing Pane, navigate to the Toolboxes.

Favorites | **Toolboxes** | Portal

Note: We will not provide these directions for accessing other tools. We will just tell you where to find them within ArcToolbox.

- Navigate to Data Management Tools followed by the Raster subtoolbox. The **Resample Tool** can be found under Raster Processing.
- Click on the **Resample Tool**.
- Set the Input Raster to the **charleston.tif** layer.
- Name the Output Raster Dataset **charleston_nn.tif** and save it in the **Working** folder.
- Set the Output Cell Size to 10 by 10 meters. The X value should be set to 10 and the Y value should be set to 10.
- Set the Resampling Technique to Nearest (this should be the default).
- Click Run to execute the tool.

```
Geoprocessing                               ▼ �📌 ✕
 ⬅                    Resample                    ☰

Parameters | Environments                         ⑦

  Input Raster
  [ charleston.tif                          ▼]  ➕
  Output Raster Dataset
  [ charleston_nn.tif                          ]  ➕
  Output Cell Size
  [                                         ▼]  ➕

  X [                    10]  Y [                10]
  Resampling Technique
  [ Nearest                                    ▼]
```

The result should automatically be added to the map. Note that this is a coarser representation of the data, as the data were resampled from a 1 m cell size to a 10 meter cell size.

Question 11. How many 1 meter cells would be contained within a 10 meter cell? (2 Points)

- Repeat this process using a 10 meter cell size and the bilinear interpolation method followed by the cubic convolution method. Name the bilinear interpolation output **charleston_bl.tif** and the cubic convolution result **charleston_cc.tif**.

You now have four versions of the image: the original image with a 1 m cell size and the three resampled images with a 10 meter cell size. Another method for resampling an image is pixel aggregation. This method calculates a statistic from the pixels or cells that fall within the new pixel or cell, such as the mean, sum, median, maximum, or minimum. You will now resample the image using this method as made available with the **Aggregate Tool**.

- ☐ Navigate to Spatial Analyst Tools followed by the Generalization subtoolbox.
- ☐ Click on the **Aggregate Tool**.
- ☐ Set the Input Raster to the **charleston.tif** layer.
- ☐ Name the Output Raster Dataset **charleston_agg.tif** and save it in the **Working** folder.
- ☐ Set the Cell Factor to 10. This means that the cell size will increase by a factor of 10. So, the output will have a 10 meter cell size as opposed to a 1 meter cell size.
- ☐ Set the Aggregating Technique to Mean. This means that the average pixel value from the 1 meter pixels that fall in the 10 meter pixel will be returned.
- ☐ You do not need to change any other settings.
- ☐ Click Run to execute the tool.
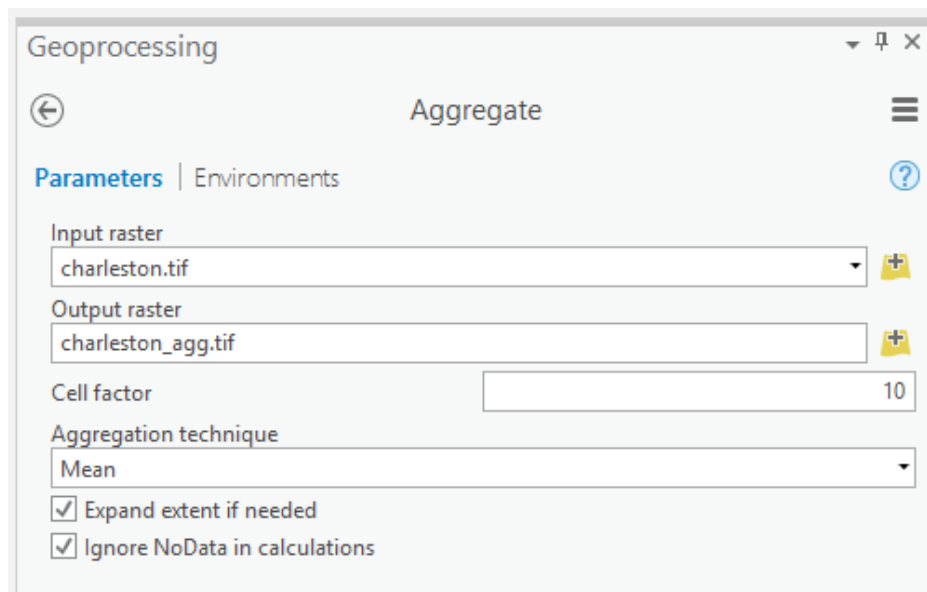
Note: This tool only works for a single band. So, only the first band in the stack is aggregated, in this case the red band. If you wanted to perform this operation for all of the bands, the tool would have to be executed separately for each band. You could then create a raster stack from them. This can be

accomplished using the **Composite Bands Tool**. We will not ask you to do so here.

To compare the data, you will extract the pixel values from each grid at the point locations provided in the **charleston_points** layer. You will do this using the **Extract Multi Values to Points Tool**.

☐ Navigate to Spatial Analyst Tools followed by the Extraction subtoolbox.
☐ Click on the **Extract Multi Values to Points Tool**.
☐ Set the Input Point Features to the **charleston_points** layer.
☐ Add all of the images to the Input Rasters list and define the Output Field Names as specified below:
  a. charleston.tif = orig
  b. charleston_nn.tif = nn
  c. charleston_bl.tif = bl
  d. charleston_cc.tif = cc
  e. charleston_agg.tif = agg
☐ Make sure "Bilinear interpolation of values at point locations" is not selected.
☐ Click Run to execute the tool.
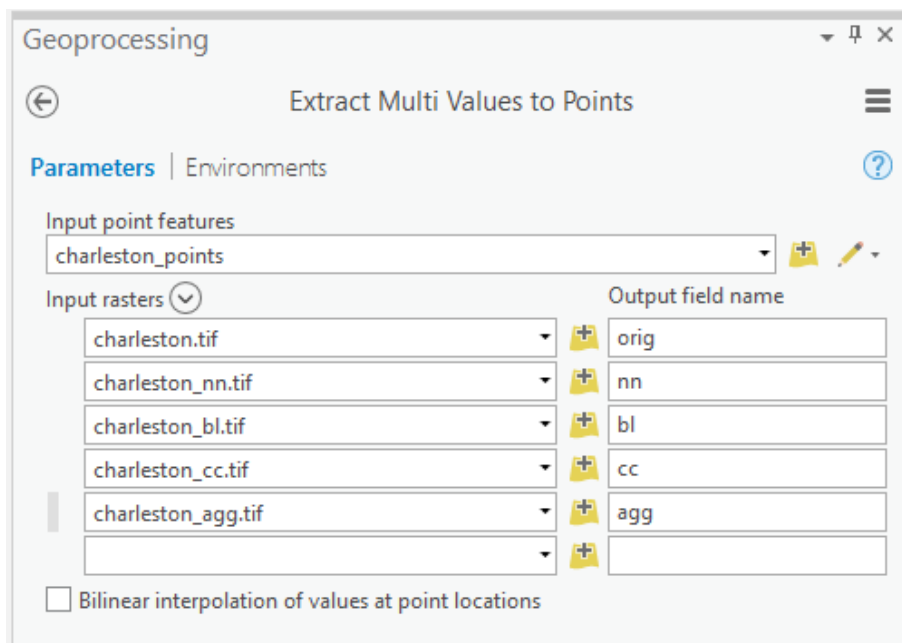
☐ Open the attribute table for the **charleston_points** layer. The new fields should have been added. The records correspond to the pixel values at that pixel or cell location.

Use the cell attribute table to answer the following question.

Question 12. Do these data suggest that resampling changes the brightness or digital number (DN) values at each location in the dataset? Why or why not? (5 Points)

Note: In the attribute table "b1" corresponds to the red band, "b2" corresponds to the green band, and "b3" corresponds to the blue band. The "agg" result is for the red band only.

END OF EXERCISE