# MapMasq<sup>®</sup> Documentation



### **Table of Contents**

About the Tools	5
Getting Started	7
Using the MapMasq <sup>™</sup> Tools	9
K-Anonymity	11
K Mean Neighbors	13
K Mean Angular Neighbors	15
Coordinate Rounding & Truncation	17
Coordinate Obfuscation	19
Additional Resources	21
K Mean Angular Neighbors Coordinate Rounding & Truncation Coordinate Obfuscation Additional Resources	13 15 17 19 21

### **About the Tools**

MapMasq<sup>™</sup> is a collection of tools to aid in the anonymizing of the geospatial component of data using a technique called geomasking. Geomasking put simply is modifying coordinate locations to protect privacy, and it introduces distortions that are much larger than measurement error.

The toolkit is a full range of geomasking techniques that hide true locations by displacement, allowing the user to select the statistical distribution, as well as its parameters, and any constraints to be applied to the distorted locations. With most methods there exists the possibility of a displacement of exactly zero, but the likely infrequency of such cases diminishes the danger of confidentiality loss.

The aim of this toolkit is to create a point file for analysis that is locationally representative while preserving anonymity.

### **Getting Started**

1. Save the MapMasq<sup>™</sup> Add-in package to the appropriate location and doubleclick the Add-in package to install.

	🚹 l 🗋 🗊 🖛 l	Compressed Fo	Ider Tools Mag	Masq.Pro.zip	
	File Home Share Vie	tw Extract			A. (
	Copy Paste Copy path	Move to - 🗙 D	elete • Interview Point	Properties	Select all Select none
	Coppears	A ManMara Pro sin	new C	Open Gearch ManMarco	Pro tio D
sri ArcGIS Add	This PC Desktop Documents In Installation Utility	↑ Name В Ma	∽ sMasq.Pro.esriAddinX	Type ESRI Add	linX File
nstallation succeed	ied.	v c			

When ArcPro launches, the MapMasq<sup>™</sup> tab will appear in the ribbon.

 Register the MapMasq<sup>™</sup> license key provided with the MapMasq<sup>™</sup> Add-In package.

副 副 副 ち・さ・・	ArcGIS Pro - MapMasq - Basemap	? – 🗆 🗙
Project Map Insert Analysis	View Edit Imagery Share MapMasq	amy (GISinc) • 🗘 🧄
MapMasq Help MapMasq B Register MapMasq	Register MapMasq	
	Enter your license key a1b2c3gisisfund4e5f6	
	OK Cancel	

Once the registration has validated, the MapMasq<sup>™</sup> button will be enabled.



\*Note – It is recommended that you set up your project directory before using  $MapMasq^{TM}$ . Most importantly, the geodatabase for exporting the new data should be established to help make the export process smooth.

### Using the MapMasq<sup>™</sup> Tools

1. Select input features (original point data set) with the browser.

MapMasq	View		* ů ×
	MapM	asq Tool	
Input Featu	res		
[	-	<b>2</b>	
Draw Ma	sking Area	]	
Masking Al	gorithm		•
Show Sta	ts		
Process	Export	Reset	Close
	Display	Options	
✓ Original I	Data Layer		
Mask Dat	ta Layer		
Result Da	ta Layer		

2. Define the Masking Area

The drawing toolbar will appear when the "Draw Masking Area" button is selected. Select the points and define the masking extent using the toolbar.



- After the masking area has been defined, points will draw in orange to show exactly which ones are to be processed. (As shown above)
- Choose the desired masking algorithm <u>K Mean Neighbors</u> or <u>K Mean</u> <u>Angular Neighbors</u> to calculate the <u>K-Anonymity</u> value. See here for coordinate rounding/truncation and <u>obfuscation</u> to further mask your data.
  - 4. Use the export button to save the processed data to the desired directory.

Tool Tip: The export does contain additional fields you may want to delete after your analysis. Likewise, if a small percentage of records do not meet your threshold you can filter and delete them using ArcGIS Pro functionality.

# **K-Anonymity**

The tools rely on a statistic called the K-Anonymity score. K-Anonymity is a quantifiable value that indicates how geomasked the output is in relation to the original location. While the score that is acceptable is dependent on the dataset and purpose, one case study indicates that a k-anonymity level of 5 was acceptable.

"The actual k-anonymity for each location was defined as the number of households from the E911 file that were closer to the original location than the distance of displacement." Link to study: <u>https://www.ncbi.nlm.nih.gov/pmc/articles/ PMC2952889/</u>

The basis for MapMasq<sup>™</sup> is code that estimates k-anonymity for each record in a database that is subjected to the K - Anonymity methods of geomasking. It considers the area in the masking, the total number of records, and location of overlapping circles. The "Possibilities" represent the number of displacement circles from the original record into which each record may fall.

Doughnut masking is the foundational concept behind geomasking. After determining the extent of your study area, displacement is based on the number of points in your dataset, their density, and distance from one another. Two circles will be drawn around each point, where the smaller circle represents the minimum distance that point needs to be displaced; the larger circle is the maximum displacement for your point. Your original point can potentially move anywhere between the two circles or doughnut.

The following sections give brief descriptions of each masking algorithm and step through the inputs: <u>K Mean Neighbors</u> and <u>K Mean Angular Neighbors</u>.

### **K Mean Neighbors**

MapMasq		т џ ×	This is the primary donut masking
MapMas Input Features	sq Tool		algorithm. The user determines the number of neighbors around each feature that are to be considered (K value). The tool
			creates the inner and outer circles
Draw Masking Area			of the deput based on the value
Masking Algorithm	K-Mean Neighbor	*	of K and the density / distance of
K-value	7		its neighbors. Each point is then
Specify average K Score range			shifted in a random compass
Minimum average	8		direction within the donut. Output
Maximum average	10		includes a <u>K-Anonymity</u> score to
✓ Show Stats			significance.
Process Export	Reset Clos	e	Step: Specify the K-value, and/or the K-score range, based on
Display C	Options		the level of anonymity.
<ul> <li>✓ Original Data Layer</li> <li>✓ Mask Data Layer</li> <li>✓ Result Data Layer</li> </ul>			Tool tip: K-value is the number of neighbors to consider when geomasking.

Tool tip: It is recommended to always specify the range for K Score. The process will run until the average <u>K-Anonymity</u> Score ends up between these values.

### **K Mean Angular Neighbors**

MapMasq	<b>→</b> ₽
MapMa	asq Tool
Input Features	
original points 🔹	<b>—</b>
Draw Masking Area	
Masking Algorithm	K-Mean Angular Neighbor
K-value	5
✓ Specify average K Score range	
Minimum average	7
Maximum average	9
Show Stats	
Process Export	Reset Close

×Same concept as <u>K Mean</u> Neighbors, but the direction that the feature is moved is based on the direction of the nearest neighbors. It builds a range in the centroid of the direction the nearest of neighbors. Although this decreases the area of the donut in which the point can move, it increases the number of circles that overlap. By bringing it closer to other neighbors, it can also add a layer of obfuscation. Nearer neighbors are more likely to

have overlapping donuts.

Step: Specify the K-value, and/or the K-score range, based on the level of anonymity.

Tool tip: K-value is the number of neighbors to consider when geomasking.

Tool tip: It is recommended to always specify the range for K Score. The process will run until the average <u>K-Anonymity</u> Score ends up between these values.

Once the process completes, you can choose to show statistics. Below is a sample outcome of statistics for either <u>K Mean</u> or K Angular Mean, and the graphic view for displacement.

✓ Show Stats	Step: Choose to show statistics or
- Minimum Possibilities: 1 - Average Possibilities: 10.93	not.
<ul> <li>Maximum Possibilities: 30</li> <li>Minimum K-Anonymity Score: 0 (4.93%)</li> <li>Average K-Anonymity Score: 8.52</li> <li>Maximum K-Anonymity Score: 403 (0.04%)</li> <li>Minimum Distance: 9.3 m</li> <li>Average Distance: 559.1 m</li> </ul>	As you can see to the left, many metrics are provided to validate the anonymity of the processed geomasked points.
- Maximum Distance: 34845.44 m	



### **Coordinate Rounding & Truncation**

These are two simple displacement algorithms with output that looks like point clustering. Users can input the number of spaces in the latitude and longitude coordinates for output. The output will indicate how the truncation coarsens the accuracy. It tends to aggregate data points to a single location, depending on the number of spaces past the decimal with which you are working.

- Average Distance: 76.58 m
- Maximum Distance: 143.09 m
- Minimum Distance: 3.28 m

Step (and Tool tip): Input the number of decimal places for either rounding or truncation.

Step: Click the process button to run.

The statistics derived from truncation and rounding look much different than K Mean as there is no user input for the level of anonymity.

## **Coordinate Obfuscation**

Random values of a distribution are created between a min and max value; that value is added/subtracted to the coordinates to displace the record. For example, random numbers might be generated from a Uniform distribution in the range [-0.01000 to +0.01000] to displace each latitude by up to 1/100 degree, or roughly 1km. At latitude 47, the meridians are only 68% as far apart as the parallels (0.68 is the cosine of 47 degrees), so random numbers in the range [-0.01466 to +0.01466] would be needed to achieve the same amount of displacement in the east-west direction.

### **Additional Resources**

Khaled El Emam, PhD and Fida Kamal Dankar, MSc. "Protecting Privacy Using K-Anonymity." Journal of American Medical Informatics Association Sep-Oct (2008): 627-637.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2528029/