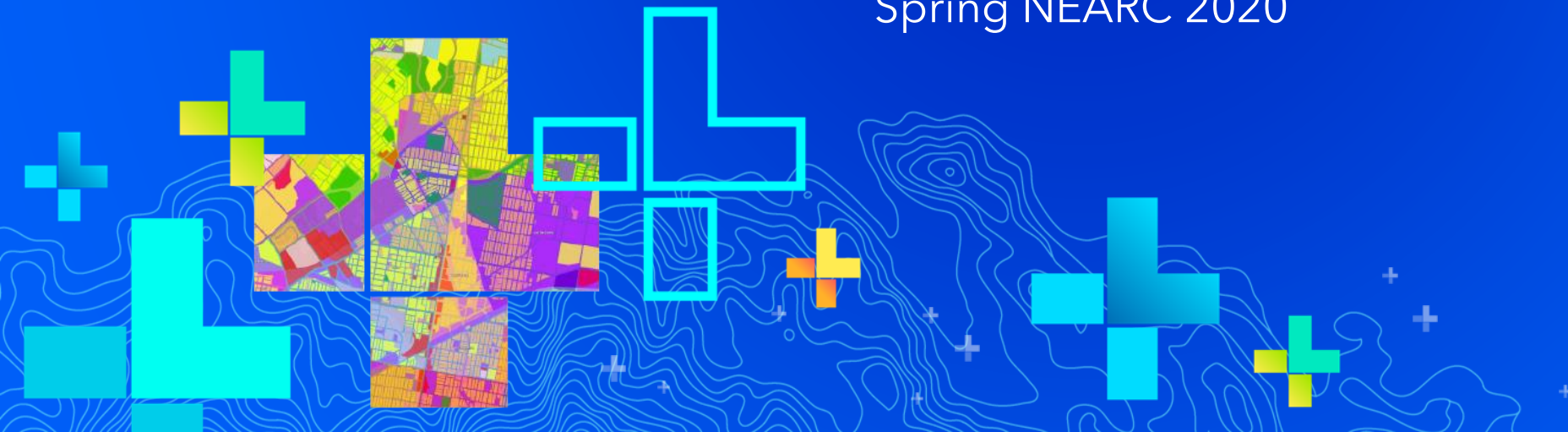


ArcGIS Spatial Analytics Tools and Techniques

Andrew Makowicki

Solution Engineer

Spring NEARC 2020



Spatial Analysis and Data Science

Data Engineering



Big Data Analytics



Machine Learning & AI



Visualization & Exploration



Sharing & Collaboration



Spatial Analysis

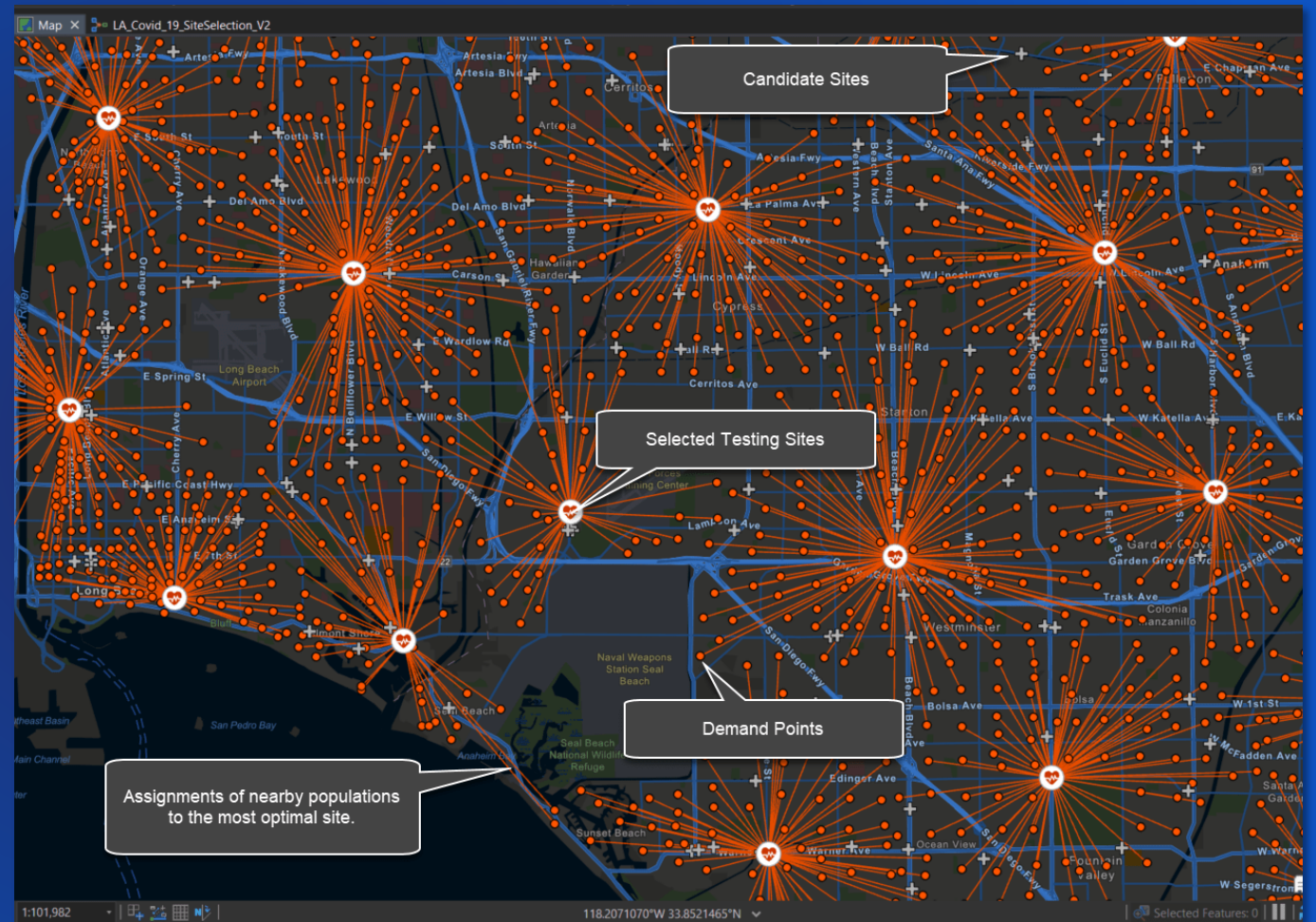


Modeling & Scripting



Why spatial analytics?

- Site selection
- Routing
- Territory Design
- Business decision support system
- Environmental modeling
 - Identifying critical source areas of pollutant emissions from non-point source pollution modeling
 - Spread of invasive species



Spatial Analysis Utilities

UI

- ArcGIS Pro
- ArcGIS Insights
- ArcGIS Online
- Enterprise Portal

Scripting

- ArcPy
- ArcGIS r-bridge
- ArcGIS API for Python
- ArcGIS Notebooks



Data Pre-processing

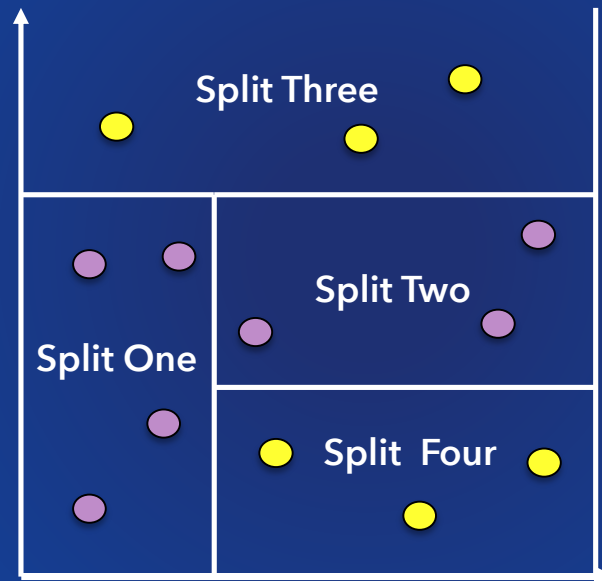
- **Cleaning data**
 - Evaluating missing data in fields
 - Encoding categorical data
- **Normalizing or standardizing data**
 - Normalize attribute of a data cut
 - Standardizing attribute data
- **Pivot tables**
 - Statistical summary of complex dataset (distill data)
- **Aggregate**
 - Bucket data into specific cuts for analysis
- **Feature extraction**
- **Summarize**

Data pre-processing
and preparation

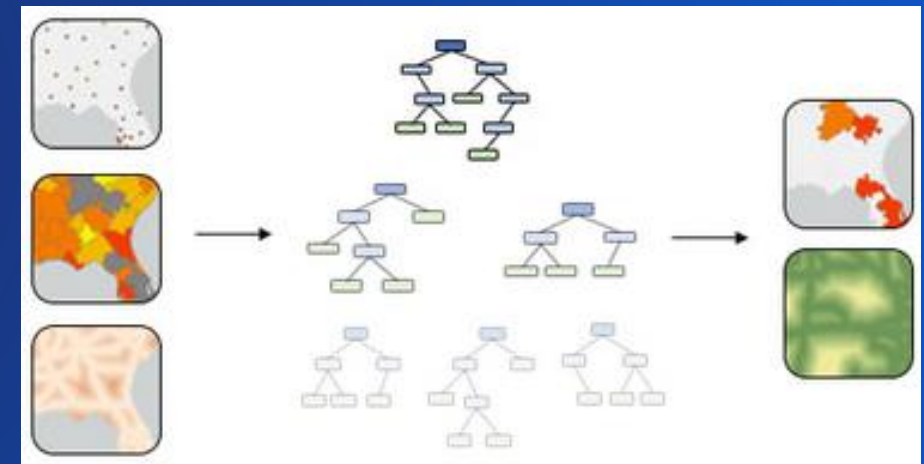


Data Engineering in ArcGIS Pro

- **Spatial statistics analysis**
 - Regression
 - Classification
 - Cluster mapping
- **Spatial Autocorrelation**
- **Geographically Weighted Regression**
- **Machine learning analysis**
 - Training/Testing
 - Decision Tree
 - Random Forest
 - Density based clustering



Random Forest graph representing splits of a dataset in a scatter plot.

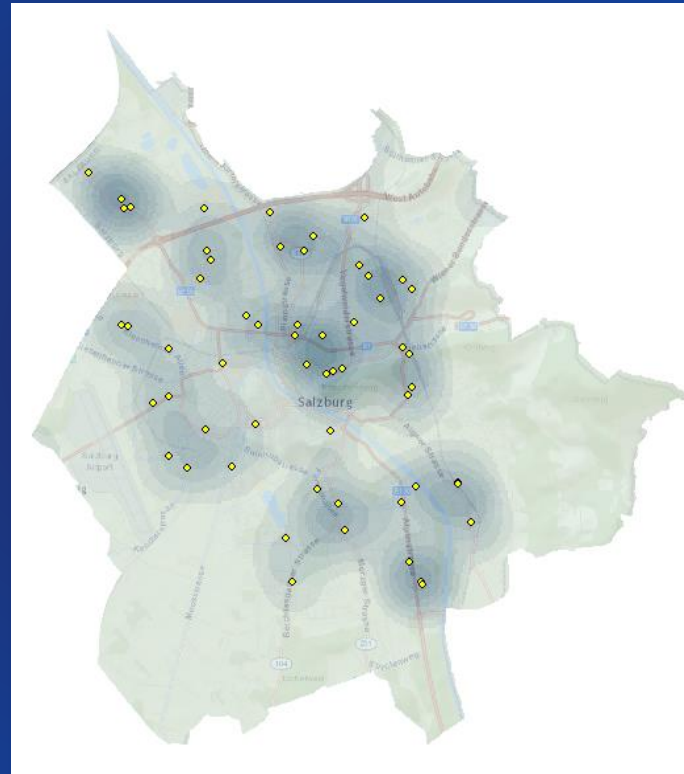


Random Forest high level overview workflow.

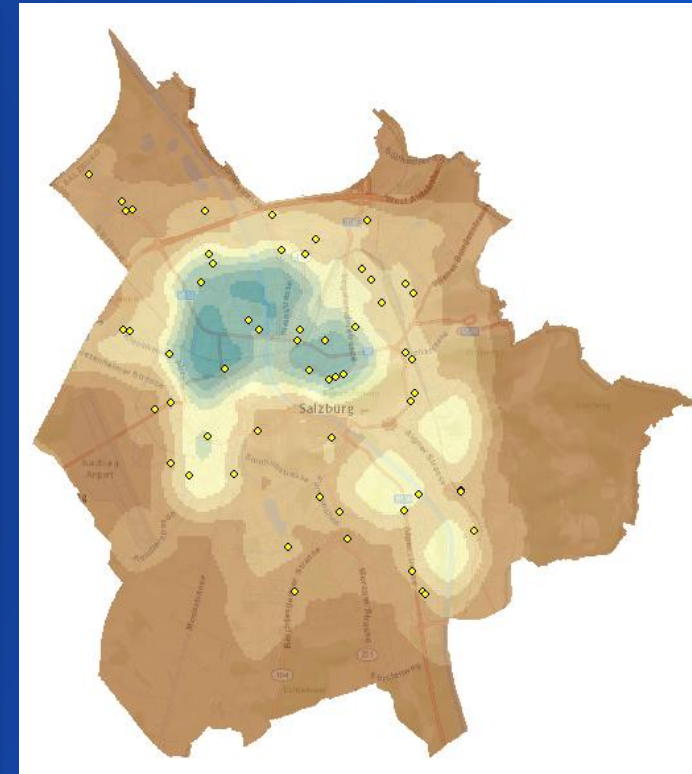
Spatial Relationships

Kernel Density Estimator

- **Patterns**
 - Hot spot analysis
 - Clustering
 - Linear Regression
- **Colocation analysis**
- **Kernel Density Estimator**
- **Interpolation methods**
 - Inverse distance weighted (IDW)
 - Kriging



Kernel density estimator for grocery store locations within a city



Kernel density estimator for populations within a city

Exploratory Data Analysis (EDA) Utilities

- Spatial Stats Toolbox
- GeoAnalytics or Geostatistical Toolbox
- ArcGIS R-bridge
- Python virtual environment
- ArcGIS Notebook
- Jupyter Notebooks Access

The screenshot displays the ArcGIS Pro interface with a Jupyter Notebook open. The notebook title is "random_forest_classification-Copy1". The code in the notebook is as follows:

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

In [3]: dataset = pd.read_csv('Social_Network_Ads.csv')
X = dataset.iloc[:, :-1].values
y = dataset.iloc[:, -1].values

In [4]: from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state

In [5]: print(X_train)
```

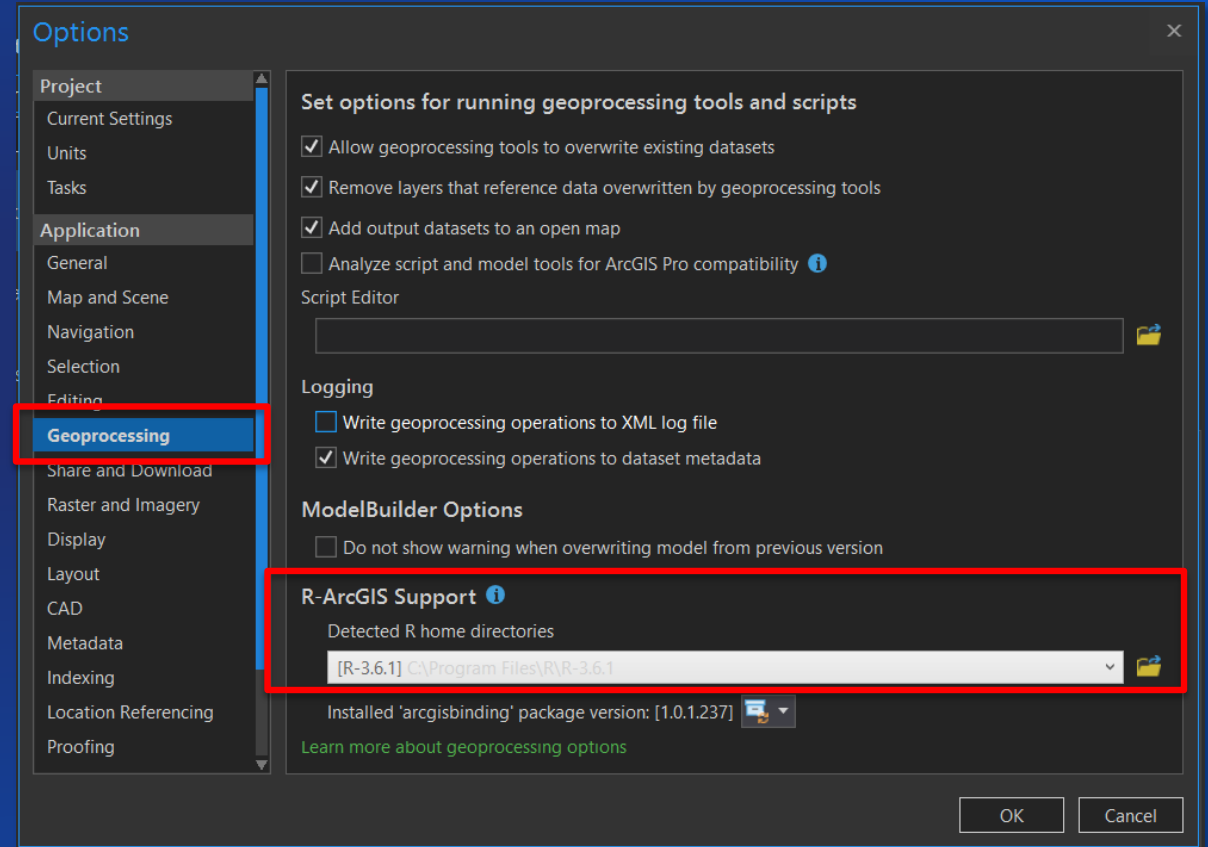
The output of the last cell shows a 5x2 array of data points:

```
[[ 44  39000]
 [ 32 120000]
 [ 38  50000]
 [ 32 135000]
```

The interface also shows a "Contents" pane on the left with a "Drawing Order" list. The "US_Prediction_Grid" layer is selected. The "Value" legend for this layer shows a color scale from -10.2449 (blue) to 41.6892 (red).

R-ArcGIS Bridge

- ArcGIS Pro → Options → Geoprocessing
 - Point to R executable program on local machine
 - Install arcgisbinding
- Jupyter Notebooks R analysis
 - Call package arcgisbinding
- Run notebook in ArcGIS Pro



R in Jupyter

Cell 1

```
In [ ]: install.packages("rgdal", type = "source")
library(rgdal)
library(arcgisbinding)
library(fda)
library(raster)
library(rts)
```

Install and initiate packages

Next, you will use the [R-ArcGIS Bridge](#) to read in the time-series mosaic that contains the predicted temperature profiles. Note that, you will need to convert a time field to a R time field.

Cell 2

```
In [ ]: ## INITIALIZE CONNECTION WITH ARCGIS PRO
arc.check_product()
## Define the location for GDB containing rasters
r <- arc.open('C:/ClimateAnalysis/p20/Default.gdb/temp_time_series')
## Read In Time Series Mosaic Times
forecast_times <- arc.select(r)
forecast_times <- as.Date(forecast_times$StdTime)
```

Check ArcGIS Pro connection and point to GDB

Next, using the `arc.raster` function of the R-ArcGIS Bridge, you will bring create a 3-D Raster Stack in R. Note that we are using `LockRaster` as the method to get one raster at a time. We read every time slice in a loop and create our Raster Stack.

Cell 3

```
In [ ]: ## DEFINE EXTENT FOR RASTERS
extent <- c(-125, 25, -70, 50)
nrow = 200
ncol = 100
## Read ALL the Rasters into One List
rList <- list()

for (i in 1:60) {
  ras <- arc.raster(r, extent = extent, nrow = nrow, ncol = ncol, nodata = -999, mosaic = list(method = "LockRaster", LockRast
  ras <- as.raster(ras)
  rList[[i]] <- ras
}
```

Define environment variables (raster extent)

Python

- **Python virtual environment**

- Creating virtual env in ArcGIS Pro
- Versioning control

- **Packages & Libraries**

- **Clone versioning**

- ArcGIS Pro environment with added libraries

- **ArcGIS Notebook Python**

Manage python environments

Manage multiple environments

Install packages for certain environments

Activate specific environments

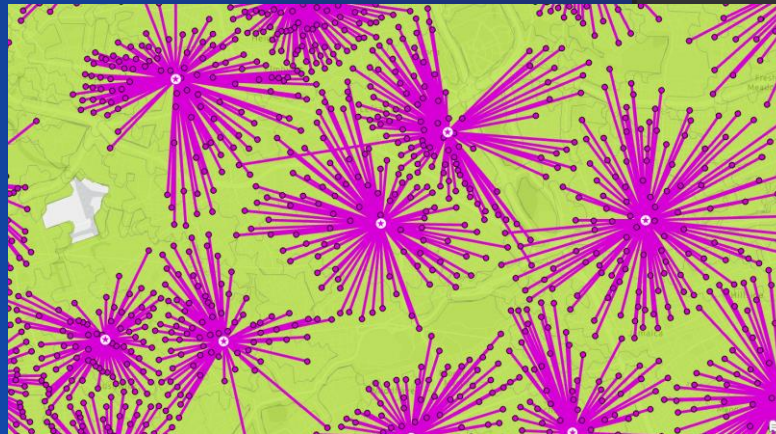
The screenshot shows the Python Package Manager interface in ArcGIS Pro. The main window is titled "Python Package Manager" and displays the "Project Environment" section. A red box highlights the "Manage Environments" button. Below it, there are buttons for "Installed Packages", "Update Packages", and "Add Packages". A "Python" button is highlighted in the left sidebar. A "Manage Environments" dialog box is open, showing a list of environments: "arcgispro-py3", "arcgispro-py3-clone", "climate_downscaling", and "with_Block_testing". A red box highlights the "arcgispro-py3" environment. Below the dialog, there is a table of installed packages. A red box highlights the "arcgis" package in the table. The table lists the following packages and versions:

Name	Version
arcgis	1.7.0
asn1crypto	0.24.0
atomicwrites	1.3.0
attrs	19.1.0
backcall	0.1.0
blas	1.0
bleach	3.1.0
ca-certificates	2019.8.28
certifi	2019.9.11
cffi	1.12.3
cftime	1.0.0b1
chardet	3.0.4
colorama	0.4.1
cryptography	2.7
cycler	0.10.0
decorator	4.4.0
defusedxml	0.6.0
despatch	0.1.0
entrypoints	0.3

At the bottom of the dialog, there is a "Description" section for the "arcgis" package, which states: "Script and automate ArcGIS Online and ArcGIS Enterprise, completing tasks ranging from performing big data analysis to content management and administration. The API integrates directly with the Jupyter Notebook and the SciPy stack."

Python EDA in ArcGIS Notebooks

- Geospatial data analysis
- Build raster datasets
- Linear and multi-linear regression analysis
 - Surface predictions for temporal data
- Network analysis and modeling



composition random_forest_classification-Copy1 multiple_linear_regression polynomial_regression-Copy1 location_allocation x

Edit View Insert Cell Help ArcGISPro

+ ↩ ↪ ↻ ↺ ↻ ↺ Run Markdown

Location Allocation Model

Create an Location - Allocation model for ArcPro
Python v 3.6.6
Current ArcPro 2.3.2 NA
Model type - Maximize Coverage
Candidate, Required and 1 Chosen DC Facility
Travel Time
Impedance cutoff = 320 min
Weight = Demand

```
In [ ]: try:
# Import libraries
import arcpy
from arcpy import env
import arcpy.na
import os
import json
import numpy as np

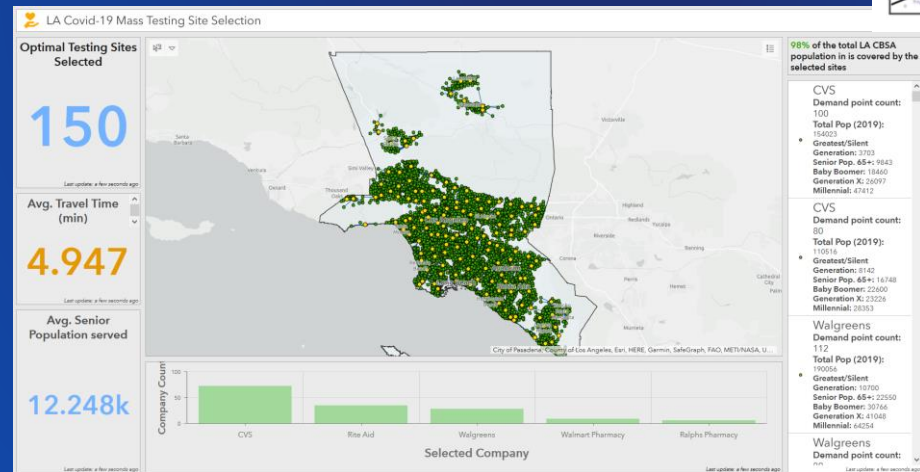
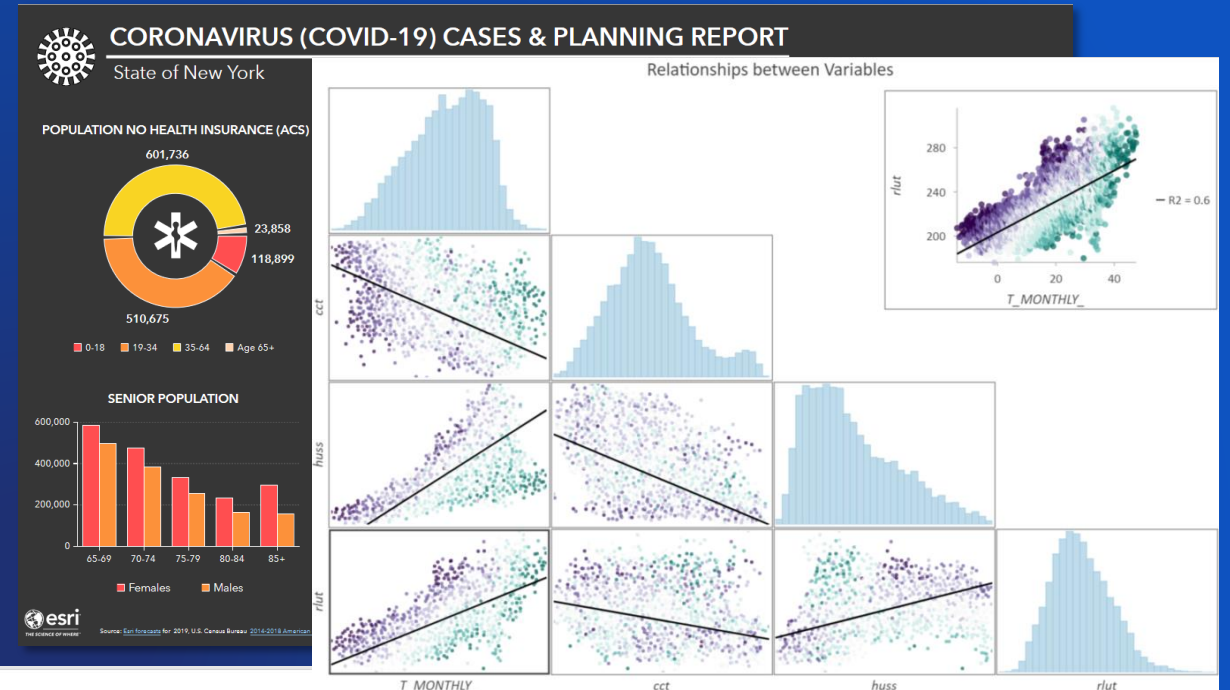
# Check for NA Extension
if arcpy.CheckExtension("network") == "Available":
    arcpy.CheckOutExtension("network")
else:
    raise arcpy.ExecuteError("Network Analyst Extension not available.")
```

Environment Settings

```
output_dir = r'D:\Python_Tools_Code\Location_Allocation' # switch out with your workspace
env.workspace = os.path.join(output_dir, "Loc_Alloc_py.gdb")
env.overwriteOutput = True
```

Results Visualization

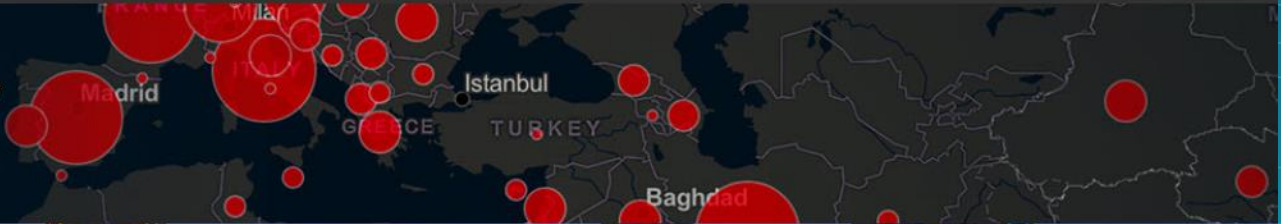
- Infographics
 - Business location summary
 - Demographic information
 - Tapestry segmentation information
- ArcGIS Pro
- Insights
- ArcGIS Dashboard



Esri Boston's Hub: <https://go.esri.com/bostonhub>

COVID-19: Get live maps, data, apps, and information

[View COVID-19 resources](#)



esri | THE SCIENCE OF WHERE™

Esri Boston Regional Office Hub

State & Federal	Local Governments	Fortune 100 Companies	Education
50 & 15	100%	90%	87
ArcGIS is used by all 50 States, and all 15 Federal Executive Departments	ArcGIS is used by all of the top 100 most populated Counties in the US	Percentage using ArcGIS in the Fortune 100	Number of Forbes top 100 colleges in the US using ArcGIS

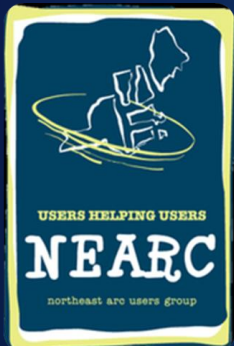
Esri is pioneering ArcGIS, the world's most powerful mapping and analytics software.

Esri Boston's Hub: <https://go.esri.com/bostonhub>

Sign Up and Follow Esri Boston

Sign up and follow Esri Boston to receive updates on local events and GIS stories from around Boston.

 Follow Our Office



11:30a.m.- 12:00 p.m.

Esri Session #5:

Rapid Field Data Acquisition with ArcGIS Quick Capture

Presenter: Mark Scott

Using ArcGIS Quick Capture, you can rapidly record field observations.

