Welcome to The Carpentries Etherpad!

This pad is synchronized as you type, so that everyone viewing this page sees the same text. This allows you to collaborate seamlessly on documents.

Use of this service is restricted to members of The Carpentries community; this is not for general purpose use (for that, try https://etherpad.wikimedia.org).

Users are expected to follow our code of conduct: https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html

----

Atlanta University Center Consortium Geospatial Data Carpentry Workshop Etherpad

Links

Workshop Website: https://bkmgit.github.io/2022-05-12-aucenter-online/
This Etherpad: https://pad.carpentries.org/2022-05-12-aucenter-online
Zoom link: https://us02web.zoom.us/j/89371843717?pwd=UHJENDFhcmtVZFltL1d5N05DYmVXZz09

Link to Curricula:

- https://datacarpentry.org/organization-geospatial/
- https://datacarpentry.org/r-intro-geospatial/
- https://datacarpentry.org/r-raster-vector-geospatial/

Minute Cards (to be filled out at the end of each daily session):
https://docs.google.com/forms/d/e/1FAIpQLSda9JAdBNY8iRaTIc_mvNmI60Osw_ImNXGNXDnFJ9q9w15_WQ/viewform

Setup

Log Into VCL

- VCL link: auc.apporto.com/
- If logging on for first time, look for your VCL account information from noreply@apporto.com in email
- Turn off your sound as we will be using Zoom for audio
- Make sure you can open RStudio in VCL
- Open up this Etherpad

Copy Over Geospatial Data for Workshop (when logged into VCL):
1. Navigate to AUC Shared Folders on your Desktop > Geospatial Summer Workshop >

Day 1 - Thursday

Sign in: Name, Pronouns, Institution, Role, Email (optional), Twitter (optional)
Please sign in so we can record your attendance.

- Angela Li, she/her, Princeton University, Data Carpentry Geospatial Workshop Co-Instructor, al49@princeton.edu
- Girmaye Misgna, he/him, University of Pennsylvania, gmisgna@upenn.edu
- Samuel Livingston, he/him, Morehouse College, Assoc. Prof, Africana Studies, samuel.livingston@morehouse.edu, @SamoryBa
- Courtney Massah Morehouse SOM graduate student clevingstonmassa@msm.edu
- Aaron Carter-Enyi, Morehouse College, research associate
- Lyrric Jackson, Lecturer at Spelman College (Department of Dance Performance and Choreography), she/her, ljacks@spelman.edu
- Chuang Peng, he/him, Morehoues College, Professor of Mathematics, chuang.peng@morehouse.edu
- jude louissaint, research assiatnt at morehouse school of meidicine
- Tameyah Mathis-Perry, she/her, Morehouse School of Medicine, Graduate Student, tmathisperry@gmail.com
- Blake Archer, he/him, Morehouse School of Medicine, Graduate Student
- Guanyu Huang, he/him, Spelman College, assistant professor, ghuang@spelman.edu
- Ertha Sefu Omba, she/her/hers, Morehouse School of Medicine, Graduate Student
- Fengxia Yan, she/her, Morehouse School of Medicine, associate professor in biostatistics, fyan@msm.edu
- Rob Brown, he/him.Spelman College, rbrown61@spelman.edu
- Nirajan Dhakal, he/him. Spelman College, ndhakal@spelman.edu
- Benson Muite, he/him, Kichakato Kizito, software developer, benson_muite@emailplus.org
- Lynne Patten, Associate Professor at Clark Atlanta University
- Tiffany Oliver
- Carry (CJ) Smith, PhD Student, Polticial Science, Clark Atlanta University

Please fill out the pre-training survey at https://carpentries.typeform.com/to/wi32rS?slug=2022-05-12-aucenter-online

How would you prefer the breaks for today's session? We expect to end around 1:30pm EST (Place an X next to your preferred option)
One 20-min break at ~11:30am EST -
Two 10-min breaks at ~11am and ~12:15 EST - XXXXxx
Three 6-min breaks at ~10:30am, ~11:30am, ~12:30am EST -

Are you using VCL for this workshop or running RStudio on your own computer?
I'm using VCL - XXxX
I'm running R and RStudio locally - xXXXXXX
I'm using VCL - Chuang Peng
i'm using VCL-Fengxia Yan

--------- **Introduction to Geospatial Concepts** -------------
https://datacarpentry.org/organization-geospatial/

- Introduction to Raster Data
- Introduction to Vector Data
- Coordinate Reference Systems
- The Geospatial Landscape

Brainstorm advantages and disadvantages for using raster data
resolution, may be an issue? -- for example how dense are your data points spatially? (i'm new, just throwing something out here).
Can allow comparison of two or more different data values geographically

Utilizing raster data can be advantageous beacause it can creater better resolution and clarity corresponding to the data

Perspective/interpretation of the data could vary across people and their style of data input. (not sure if everyone does it the same way)

Satellite imagery - generating raster data is easier this way
Better at representing continuous variables (i.e. elevation)

Reading a PROJ4 String
Here is a PROJ4 string for one of the datasets we will use in this workshop:
+proj=utm +zone=18 +datum=WGS84 +units=m +no_defs +ellps=WGS84 +towgs84=0,0,0

- What projection, zone, datum, and ellipsoid are used for this data?
- What are the units of the data?
- Using the map above, what part of the United States was this data collected from?

------------------------ **Morning Break** ------------------------

Break 11:05 EST - resume at 11:15pm

Let us know you're back by jotting down your name and telling us your favorite beverage
Benson - Orange Juice
Angela - Earl grey tea with milk
Sam - GreenTea, lemon and honey.... and Coffee.
Tiffany Oliver, coca cola
Lyrric-Tea and water :)
Courtney Massah Green tea or Cinnamon Chai tea

Fengxia Yan tea
Aaron Carter-Enyi, iced tea
Chuang Peng, ice tea
Rob Brown, coffee.

------ Introduction to R for Geospatial Data: Setup -------

https://datacarpentry.org/r-intro-geospatial/
Reference sheet: https://datacarpentry.org/r-intro-geospatial/reference.html

- Introduction to R and RStudio

```
mass <- 47.5
age <- 122
mass <- mass * 2.3
age <- age - 20
```

- Project Management With RStudio
  - **Link to download gapminder/nordic data**:
    https://datacarpentry.org/r-intro-geospatial/02-project-intro/index.html#save-the-data-in-the-data-directory
  - 1. For each link: Right click on it and select "Save link as.."
  - 2. Add ".csv" at the end of each File name
  - 3. For "Save as type", change the drop down from "Text Document" to "All Files"
  - 4. Click Save

------------------------ Morning Break ------------------------

Break 12:35 EST - resume at 12:45 EST

Further resources:

http://adv-r.had.co.nz/Data-structures.html
https://programminghistorian.org/en/lessons/

Let us know you're back by jotting down your name and one thing on your desk (or something that you can see from where you are)
jude-my anime themed keys
Angela - photo of my hometown
Lyrric- I am back and I have an hourglass with periwinkle sand in it with two photos of my mom. :)
Benson - back, a dictionary
Aaron Carter-Enyi, I'm back but don't have a desk
Chuang Peng, back, my phone is next to me on my desk
Courtney Massah lamp and easel

--------- Introduction to R for Geospatial Data -------------

#You can also try the following to download the nordic and gapminder data directly from r prompt.

download.file("https://raw.githubusercontent.com/datacarpentry/r-intro-geospatial/master/_episodes_rmd/data/gapminder_data.csv",
        "data/gapminder_data.csv")
download.file("https://raw.githubusercontent.com/datacarpentry/r-intro-geospatial/master/_episodes_rmd/data/nordic-data-2.csv",
        "data/nordic-data-2.csv")
download.file("https://raw.githubusercontent.com/datacarpentry/r-intro-geospatial/master/_episodes_rmd/data/nordic-data.csv",
        "data/nordic-data.csv")

- Data Structures
- Exploring Data Frames
- Subsetting Data

Code from Yesterday

1 + 1
1 + 100
1 + 50
3 + 5 * 2 # testing out R as calculator
(3 + 5) * 2 # using parentheses
2/10000
5e3

?str

1 == 1
1 != 1 # not equal
1 != 2
(5*9) > 8/4
< , <=, >, >=

# Variables and Assignment
x <- 100
x
x <- x + 1
y <- x * 2
mass <- 47.5
age <- 122
mass <- mass * 2.3
age <- age - 20
mass > age # Ctrl-Enter to Run
z = 20
score.4th <- 20

which ones are valid variable names?

```r
min_height <- 1
max.height <- 1
# _age <- 1
.mass <- 1 # hidden variable, not recommended
MaxLength <- 1
min-length <- 1
# 2widths <- 1
celsius2kelvin <- 1


# Installing Packages
install.packages("dplyr")
# Install "ggplot2" package
install.packages("ggplot2")
library(dplyr)
library(ggplot2)
```

Part 2

```r
getwd()

nordic <- read.csv("data/nordic-data.csv")

# What do each of these do?
str(nordic) # str(mtcars) - car data if not downloaded
summary(nordic)
ncol(nordic)
nrow(nordic)
dim(nordic)
View(nordic)

?summary

nordic$country
nordic$lifeExp
colnames(nordic)

nordic$lifeExp + 2 # mtcars$mpg + 10
nordic$lifeExp + nordic$country # does not work
```


Day 2 - Friday

Sam Livingston, Good Morning.

Sign in: Jot down your name and whether you are using VCL. If not, what is the operating system of your computer? (PC or Mac) (Note also that the code we wrote yesterday is pasted into the Etherpad between lines 183-250.)

NAME   VCL   Operating System if local

Benson Muite - Not using VCL, Linux

Angela Li - Yes, I am using VCL

Girmaye Misgna - VCL

Courtney Massah using VCL today

Robert Brown, Not using VCL, using my PC. locally.

Fengxia Yan not using VCL

Blake Archer using Mac

Chuang Peng, VCL,

Nirajan Dhakal, using Windows locally

jude louissaint using mac

Ertha Sefu Omba, I am using VCL

Aaron Carter-Enyi, not using VCL, Mac

Tameyah Mathis-Perry, using VCL

Lyrric Jackson

Break times:
11:00am

Back from break? Jot down your name and something you're looking forward to in the next week (or just say "I'm back")

Angela - traveling to Chicago to see some friends

jude louissaint

Rob Brown, I'm back!

Aaron Carter-Enyi

Fengxia Yan

Tameyah Mathis-Perry- seeing family

Chuang Peng - watching my daughter's music concert

Courtney Massah -I'm back, looking forward to some beach time next week

I'm back - Sam - applying raster and vector data to a couple of projects! ...and going home to the beach.

I am back, Lyrric Jackson- looking forward to seeing where and how the data we work with is plotted on maps. :)

Benson Muite

Reference links

https://support.rstudio.com/hc/en-us/articles/200711853-Keybreoard-Shortcuts-in-the-RStudio-IDE

Break -- come back at 12:50

Tiffany Oliver, The Cancer Genome Atlas - https://portal.gdc.cancer.gov/ Related GIS data
https://gis.cancer.gov/
Name, Interesting open dataset
Benson, IMF and Worldbank datasets

Fengxia Yan
Lyric Jackson, BACK :)
Courtney Massah, Back
Sam, I'm back.

Tameyah Mathis-Perry I'm back

Aaron Carter-Enyi, I'm back too
Ertha Sefu Omba-back
Rob Brown, Back

Exercise

Write a single command (which can span multiple lines and includes pipes) that will produce a dataframe that has the American values for lifeExp, country and year, but not for other Continents. How many rows does your dataframe have?

COMMANDS SO FAR

```
getwd()
(nordic <- read.csv("data/nordic-data.csv"))
# Subsetting
nordic[1,]
nordic[,3]
nordic[1,3]

# Get life expectancy for Norway
nordic[3,3]
# Get years
nordic[,2]
nordic$year
nordic["year"]
# Exercise
# Check the class
nordic[1, 1]
nordic[, 1]
nordic[1, ]
nordic[1]
nordic[[1]]
nordic$country
nordic["country"]

class(nordic)
class(nordic[1,1])
class(nordic[,1])
class(nordic[1])
```

```
class(nordic[[1]])
class(nordic$country)
class(nordic["country"])

nordic$year
nordic[[2]]
class(nordic$year)
class(nordic[[2]])
summary(nordic)
nordic[1:2,] # get rows 1 through 2
nordic[c(1,3),] # get rows 1 and 3

gapminder <- read.csv("data/gapminder_data.csv")
gapminder
head(gapminder)
head(gapminder, n = 10)
tail(gapminder)

str(gapminder)
summary(gapminder)

gapminder[gapminder$country == "Brazil",] # only look at Brazil
unique(gapminder$country) # what counties are represented in data?
# try with a country of your interest and save as dataframe
# with name of that country

gapminder[gapminder$year > 1982,] # look at years after 1982

rwanda <- gapminder[gapminder$country == "Rwanda",] # save

# Taking the average of numbers
numbers <- c(3, 4, 5)
mean(numbers)

mean(rwanda$pop)
max(rwanda$pop)
min(rwanda$pop)

# Get mean for multiple variables in one command?

mean(gapminder[gapminder$continent == "Africa",
         "gdpPercap"])
mean(gapminder[gapminder$continent == "Asia",
         "gdpPercap"])
# Use dplyr
install.packages('dplyr')
library(dplyr)
## Using select()
```

```r
year_country_gdp <- select(gapminder, year, country,
                  gdpPercap)
View(year_country_gdp)
## Using filter()
year_country_gdp_euro <- gapminder %>%
 filter(continent == "Europe") %>%
 select(year,country,gdpPercap)

View(year_country_gdp_euro)
## Breaking this up into two steps
year_europe <- gapminder %>%
 filter(continent == "Europe")
View(year_europe)

year_country_gdp_europe <- year_europe %>%
 select(year,country,gdpPercap)
View(year_country_gdp_europe)

# Exercise
year_country_lifeExp_americas <- gapminder %>%
 filter(continent == "Americas") %>%
 select(year,country,lifeExp)

## group_by
str(gapminder)
gapminder %>% group_by(continent) %>% str()
gapminder_continents <- gapminder %>%
 group_by(continent)
View(gapminder_continents)
?str
str(gapminder_continents)
## USe summarize to view differences
gdp_bycontinents <- gapminder %>%
 group_by(continent) %>%
 summarize(mean_gdpPercap = mean(gdpPercap))
gdp_bycontinents
lifeExp_bycountry <- gapminder %>%
 group_by(country) %>%
 summarize(mean_lifeExp = mean(lifeExp))
lifeExp_bycountry

## Mutate - create new columns
gdp_pop_bycontinents_byyear <- gapminder %>%
 mutate(gdp_billion = gdpPercap*pop/10^9) %>%
 group_by(continent, year) %>%
 summarize(mean_gdpPercap = mean(gdpPercap),
       mean_gdp_billion = mean(gdp_billion),
       sd_gdp_billion = sd(gdp_billion))
gdp_pop_bycontinents_byyear
```

```
View(gdp_pop_bycontinents_byyear)
# Visualization

library("ggplot2")
ggplot(data=gapminder, aes(x=lifeExp)) +
  geom_histogram()
ggplot(data=gapminder, aes(x=lifeExp))
names(gapminder)
# Exercise - plot a histogram of gdpPercap
```


Day 3 - Monday

Sign in: Write your name and a project you hope to use geospatial data for. What type of data will you use? Is it raster (continuous surface) or vector data (discrete boundaries, lines, or points)?

- Rob Brown, Conduct a data visualization of the effects a state's demography upon its politics, Excel dataset vector data.
- jude louissaint, future research
- Chuang Peng, this is the first time I had geo data, not sure now, but willing to try either/both
- Tameyah Mathis-Perry, tbd,
- Courtney Massah, conduct data visualization on the quality of care available in rural and inner city communities
- Blake Archer, unsure currently
- Guanyu Huang, air pollution at block level.
- Fengxia Yan, I ususally work with public health data. For geo data, I work with census data, enviromental data, most of the data are continuous.
- Ertha Sefu Omba, TBD
- Lyrric Jackson, Dance History Research
- Sam Livingston - Congo Curriculum, UN Mapping Report on the DRC; Mapping African Ethical texts; Other Africana Studies course related projects; I think it will be discreet/vector data.
- Benson Muite
- Tiffany Oliver

Questions

https://desktop.arcgis.com/en/arcmap/latest/manage-data/shapefiles/shapefile-file-extensions.htm

Breakout Rooms: Import Line and Point Shapefiles

Using the steps above, import the HARV_roads and HARVtower_UTM18N layers into R. Call the HARV_roads object lines_HARV and the HARVtower_UTM18N point_HARV.

Answer the following questions in each group (jot down notes in Etherpad):

- What type of R spatial object is created when you import each layer?
- What is the CRS and extent for each object?
- Do the files contain points, lines, or polygons?
- How many spatial objects are in each file?

Breakout Room 1

- What type of R spatial object is created when you import each layer?
- What is the CRS and extent for each object?
- Do the files contain points, lines, or polygons?
- How many spatial objects are in each file?

Breakout Room 2

- What type of R spatial object is created when you import each layer?
- What is the CRS and extent for each object?
- Do the files contain points, lines, or polygons?
- How many spatial objects are in each file?

Breakout Room 3

- What type of R spatial object is created when you import each layer?
- What is the CRS and extent for each object?
- Do the files contain points, lines, or polygons?
- How many spatial objects are in each file?

Breakout Room 4

- What type of R spatial object is created when you import each layer?   (line_harv-Multilinestring point_harv-point)
- What is the CRS and extent for each object?   (lines_harv:
- Coordinate Reference System:
-   User input: WGS 84 / UTM zone 18N
-
- point_harv:
- User input: WGS 84 / UTM zone 18N
-
- Do the files contain points, lines, or polygons?
    - POINTS for the tower file and MULTILINESTRING for the roads file
- How many spatial objects are in each file?

line_harv: 13
point_harv: 1

Breaks - 11:00am and 12:15am (10 min each)

Back from break? Let us know by jotting down what your favorite food is (or just type "back")

Aaron Carter-Enyi, curry

Lyrric Jackson- Tomatoes :)

jude louissaint, haitian patties

Blake Archer - Tacos

Ertha Sefu Omba-Jollof rice

Benson Muite - Shito on toast

CourtneyMASSAH sushi

Rob Brown, Indian cuisine(Curry Chicken)

Tameyah Mathis-Perry, french fries

Fengxia Yan: noodle

Sam - Ceebu Djen (Wolof Fish and Rice)

Chuang Peng, good soup

Code So Far - Day 3

```
# install.packages("sf")
library(sf)
library(ggplot2)

# as you're typing this, use "Tab" to check file path
extent_harv <- st_read("data/NEON-DS-Site-Layout-Files/NEON-DS-Site-Layout-
Files/HARV/HarClip_UTMZ18.shp")

View(extent_harv)
extent_harv
class(extent_harv)

st_geometry_type(extent_harv)
st_crs(extent_harv)
st_bbox(extent_harv)

ggplot() +
  geom_sf(data = extent_harv, fill = "green", size = 3) +
  labs(title = "Extent of HARV Data") +
  coord_sf()

nrow(extent_harv)
ncol(extent_harv)
names(extent_harv)

# Try reading HARV_roads and HARVtower
lines_harv <- st_read("data/NEON-DS-Site-Layout-Files/NEON-DS-Site-Layout-Files/HARV/
HARV_roads.shp")
```

```r
class(lines_harv)
lines_harv

point_harv <- st_read("data/NEON-DS-Site-Layout-Files/NEON-DS-Site-Layout-Files/HARV/
HARVtower_UTM18N.shp")

st_crs(point_harv)
st_geometry_type(point_harv)
st_bbox(point_harv)

ggplot() +
  geom_sf(data = lines_harv) +
  geom_sf(data = extent_harv) # plots box over lines

p <- ggplot() +
  geom_sf(data = extent_harv) +
  geom_sf(data = roads) +
  geom_sf(data = point_harv, size = 5, color = "red") +
  labs(title = "Roads and Tower in Harvard Forest")

p

ggsave(filename = "roads_map.jpg", plot = p)
ggsave("roads_map2.jpg")
?ggsave
?ggplot2
?ggplot


lines_harv
ncol(lines_harv)
names(lines_harv)
head(lines_harv) # try out these commands on the other two

lines_harv$TYPE

ggplot() +
  geom_sf(data = lines_harv, aes(color = TYPE)) +
  labs(color = "Type of Road")

library(dplyr)
lines_harv %>%
  filter(TYPE == "footpath")

# does the same thing in terms of calling filter from dplyr, not base R
footpath_harv <- lines_harv %>%
  dplyr::filter(TYPE == "footpath")
```
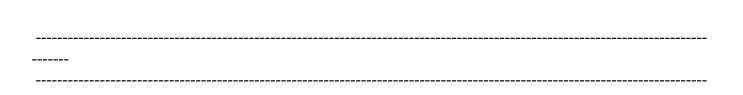
```
ggplot() +
  geom_sf(data = footpath_harv, aes(color = factor(OBJECTID))) +
  labs(title = "NEON Harvard Forest Field Site",
      subtitle = "Footpaths",
      color = "Footpath ID") +
  coord_sf()

names(footpath_harv)
str(footpath_harv)

# filter by boardwalk
boardwalk_harv <- lines_harv %>%
  dplyr::filter(TYPE == "boardwalk")

ggplot() +
  geom_sf(data = boardwalk_harv) +
  labs(title = "NEON Site",
      subtitle = "Boardwalks") +
  coord_sf()

  # filter by boardwalk
  _____ <- lines_harv %>%
  dplyr::filter(TYPE == "_____")

ggplot() +
  geom_sf(data = _____) +
  labs(title = "_____",
      subtitle = "_____") +
  coord_sf()
```

#Projection Systems

```
state_boundary_US <- st_read("data/NEON-DS-Site-Layout-Files/US-Boundary-Layers/US-State-
Boundaries-Census-2014.shp")
ggplot() +
  geom_sf(data = state_boundary_US) +
  ggtitle("Map of Contiguous US State Boundaries") +
  coord_sf()

#Visualizing maps using different projections.

#Display map using the following EPSG codes
#4326, 3857, 2955,2163
#How does the map look like in those various projections?
#Learn about EPSG codes and try to find out what projection each of those 4-5 digit EPSG codes stand
for? (Hint: st_crs() )

#"The EPSG codes are 4-5 digit numbers that represent a particular CRS definitions. The acronym EPSG,
comes from the European Petroleum Survey Group. "
```

```
# Mercator EPSG: 3857
proj4 <- st_crs(3857)
ggplot() +
  geom_sf(data = country_boundary_US, color = "gray18", size = 2) +
  ggtitle("Mercator, EPSG:3857") +
  coord_sf(crs=proj4)

# UTM Zone 11N
proj4 <- st_crs(2955)
ggplot() +
  geom_sf(data = country_boundary_US, color = "gray18", size = 2) +
  ggtitle("UTM Zone 11N, EPSG 2955") +
  coord_sf(crs=proj4)

# Equal Area EPSG:2163
proj4 <- st_crs(2163)
ggplot() +
  geom_sf(data = country_boundary_US, color = "gray18", size = 2) +
  ggtitle("US National Atlas, equal area, EPSG:2163") +
  coord_sf(crs=proj4)


Day 3 Continued:
library(raster)
library(rgdal)
library(ggplot2)
library(dplyr)
library(sf)

plot_location_harv <- read.csv("data/NEON-DS-Site-Layout-Files/NEON-DS-Site-Layout-Files/
HARV/HARV_PlotLocations.csv")
str(plot_location_harv)

plot_location_harv_sp <- st_as_sf()
```

Additional Resources

Tidy Census Package
https://walker-data.com/tidycensus/
https://cran.r-project.org/web/packages/tidycensus/index.html
https://r-spatial.github.io/sf/
https://ggplot2-book.org/
https://r-graphics.org/

```
download.file("https://ndownloader.figshare.com/articles/2009586/versions/10","data/2009586.zip")
unzip("data/2009586.zip")
```

https://stackoverflow.com/questions/70602170/is-there-a-way-to-autosave-my-work-in-r-studio

----

Day 4 - Tuesday

Resources:

Health service areas
https://data.dartmouthatlas.org/supplemental/#boundaries
https://seer.cancer.gov/seerstat/variables/countyattribs/hsa.html

RMarkdown:
https://kbroman.org/knitr_knutshell/pages/Rmarkdown.html

Mapview
https://r-spatial.github.io/mapview/
https://bookdown.org/nicohahn/making_maps_with_r5/docs/mapview.html

Sign in: Write your name and GPS coordinates of a location you like

jude louissaint
Benson Muite,   -4.0505200, 39.6671690
Courtney Massah 33.207630, -92.666267
Fengxia Yan
Rob Brown, Latitude: 33.837823 / N 33° 50' 16.162"
Longitude: -84.315678 / W 84° 18' 56.44", Spelman College.
Chuang Peng, 33.7473° N, 84.4155° W, Morehouse Collge
Blake Archer 39.0742° N, 21.8243° E (Greece)
Aaron Carter-Enyi 6.7944485,7.3084937 (Obimo), using my local machine (mac)
Samuel Livingston - 24.079571812653246, 32.88922854171067
The Nubian Museum in Aswan, Egypt.
Guanyu Huang, 33.837823 -84.315678 , Spelman College.
Angela Li, 41.8781° N, 87.6298° W, City of Chicago
Girmaye Misgna, 39.9526° N, 75.1652° W
Lyrric Jackson, 33°44'36.8"N 84°24'15.0"W, City of Atlanta
Ertha Sefu Omba **33° 55' 7.8996'' S(South Africa)**
Tameyah Mathis-Perry **10° 37' 50.0628'' N 85° 26' 21.6456'' W**

Code (Day 4, Tuesday)

```r
library(raster)
library(rgdal)
library(ggplot2)
library(dplyr)
library(sf)

plot_location_harv <- read.csv("data/NEON-DS-Site-Layout-Files/NEON-DS-Site-
Layout-Files/HARV/HARV_PlotLocations.csv")

# ways to look at your data
plot_location_harv
View(plot_location_harv)
head(plot_location_harv)
str(plot_location_harv)
names(plot_location_harv)

class(plot_location_harv)

plot_location_harv_sf <- st_as_sf(plot_location_harv, coords = c("easting",
"northing"))

class(plot_location_harv_sf)

# try to plot this
ggplot() +
  geom_sf(data = plot_location_harv_sf) +
  labs(title = "Location of Plots")

# check CRS
st_crs(plot_location_harv_sf)
st_crs(plot_location_harv_sf) <- 32618
st_crs(plot_location_harv_sf)

plot_location_harv_latlon <- st_transform(plot_location_harv_sf, 4326)
st_crs(plot_location_harv_latlon)

# plot latlon object
ggplot() +
  geom_sf(data = plot_location_harv_latlon) +
  labs(title = "Lat-Lon Location of Points")

# write out shapefile / XLSX / CSV
# I created a new "output" folder in here
st_write(plot_location_harv_latlon, "output/plot_location_harv_latlon.shp",
delete_layer = TRUE) # delete layer overwrites original output
```

```
st_write(plot_location_harv_latlon, "output/plot_location_harv_latlon.xlsx")
st_write(plot_location_harv_latlon, "output/plot_location_harv_latlon.csv")
```

Raster Data and RMarkdown

Introduction to Raster Data

Load libraries

````
```{r libraries}
library(raster)
library(rgdal)
library(ggplot2)
library(dplyr)
```
````

View raster file attributes

````
```{r view attribtues}
GDALinfo("data/NEON-DS-Airborne-Remote-Sensing/NEON-DS-Airborne-Remote-Sensing/HARV/
DSM/HARV_dsmCrop.tif")
```
````

Load a raster file

````
```{r load raster}
DSM_HARV <- raster("data/NEON-DS-Airborne-Remote-Sensing/NEON-DS-Airborne-Remote-
Sensing/HARV/DSM/HARV_dsmCrop.tif")
DSM_HARV
```
````

Plot, first convert to a dataframe
````
```{r plot DSM_HARV}
DSM_HARV_df <- as.data.frame(DSM_HARV, xy = TRUE)
str(DSM_HARV_df)
```
````
Plot using ggplot
````
```{r plot dataframe}
ggplot() +
  geom_raster(data = DSM_HARV_df,
         aes(x = x, y = y , fill = HARV_dsmCrop)) +
  scale_fill_viridis_c() +
  coord_quickmap()
```
````

Back from Break? Let us know what you like to do over the summer:

Courtney Massah kayaking or any water activity

Sam L. - Write, rest on a beach, a little travel.
Tameyah Mathis-Perry- going on trips
Lyrric Jackson- I like sleeping in the summer and going to the beach
Fengxia Yan beach vocation
Chuang Peng - send my daugher to college, if still have time, continue working on R-studio and geospatial data sets,

Angela Li - spending time at the beach, avoiding the humidity
Benson Muite, Travel
Rob Brown, Travel.

Name, Do you publish?

Benson, occasionally
Blake Archer, have not yet
Fengxia Yan, yes
Chuang Peng, have to, but try to catch up... I'm behind where I'd like to be
Courtney Massah, have not yet but soon.
Sam Livingston - Yes.
Rob Brown, Yes, I do, but I intend to publish much more.
Lyrric Jackson
Tameyah Mathis-Perry, tbd
Guanyu Huang, yes

Practice What You Learned Today!

Plot Neighborhood Planning Unit shapefle data for Atlanta, and then on top of this, plot some of the favorite locations in Atlanta
https://github.com/bkmgit/2022-05-12-aucenter-online/tree/gh-pages/data
If you make a nice plot, feel free to email the plot to us when you are done!

You can find more about the R Markdown at
https://rmarkdown.rstudio.com/lesson-1.html

Workshop Post-Surveys

- **AUC DSI:**
    - From Bettina: *"You can complete the AUC DSI post-workshop survey and upload your reflection report here: https://www.surveymonkey.com/r/aucdsiworkshop22 . The instructions for the reflection report are attached. Please go ahead and complete this ASAP. The overall deadline is August 1, 2022. For those who submit their reflection report and post-workshop surveys by this Thursday, stipends will be processed by May 31."*
- **Carpentries Post-Survey:**
    - https://carpentries.typeform.com/to/UgVdRQ?slug=2022-05-12-aucenter-online&typeform

Common Errors in R & RStudio

Errors are extremely common when coding - troubleshooting and debugging are skills that take practice over time. If you are stumped for a long time, a second pair of eyes (or taking a short break) can help. We also often enter our error messages into Google or StackOverflow to look for answers online.

The most common R error messages

- "could not find function" errors, usually caused by typos or not loading a required package.
    - You can double check that packages are installed in the Packages tab, or by running "library(package)".
- "cannot open the connection" Error
    - There can be two reasons for this error to show up when running an R script:
        - 1. A file/connection can't be opened because R can't find it (mostly due to an error in the path). Check your current working directory. Run **getwd()** at the console to see if your current working directory is what it's supposed to be. Folder containing your data should particularly be inside working directory.
            - You can use setwd() to set the correct working directory, or - our preferred method - **open up the .Rproj file** (check you're in the project by looking at the cube icon at the top right corner of RStudio) Make sure to open up the ".Rproj" file instead of just the R file so that R knows by default which working directory you are in.
        - 2. Failure in .onLoad() because a package can't find a system dependency
- **Spelling and syntax errors** - it is important to be precise and detail-oriented when typing code.
    - You've forgotten a comma, opened a bracket, but haven't closed it, misspelled character by mistake or something else R doesn't understand. Those are usually picked up by R and you will get error messages reminding you to proof-read your code and fix them.
    - One way to make less spelling errors is to use the **Tab key to autocomplete** file and variable names in RStudio.