Blanford, J.I. (2025) Teaching material. Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.

Title: Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.
Author: Justine I. Blanford
Objectives: To make statistics fun with geographic information and spatial analysis

Cite: *Blanford, J.I. (2025) Teaching material. Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.*
Further information: Blanford, J.I. (2025) Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.

GIS, Statistics and Data Science are important for analysing data. During this project you will learn about the different spatial analysis and statistical methods that can be used to analyse real-world data such as temperature. The project is linked to each of the different topics highlighted in Figure 1.

Figure 1. Overview of GIS, statistical and spatial analysis methods components covered during the course.

Data Capture, Create **Data Engineering** Transform, clean, enrich data	**Data Collection** Sampling & Spatial Sampling Point, line, area Systematic, Random, Stratified Stratified-systematic, Stratified-random Data Acquisition Finding data	**Data Reporting and Management** Data Fields and Attributes Projections – Map coordinate systems Managing spatial information Data structuring and organisation	**Data Engineering** Data fusion and integration Data wrangling Data conversion Data Grouping/Aggregation Working with data from various sources.
Visualisation and exploration	**Data visualisation** line graphs, bar graphs, histograms/ frequency graphs, divided bar graphs, pie graphs, flow diagrams, radial graphs and wind rose graphs, scatter graphs, triangular graphs Boxplots, violin plots	**Cartographic visualisations** Cartographic representation, symbology, Key design principles - legibility, visual contrast, figure-ground organization, hierarchical organization, and balance.	**Data Exploration** Interact with data Explore attributes Explore data ranges, data errors, data uncertainties, outliers, missing information Data summaries
Data Analysis Statistical Analysis of the data **Spatial Analysis** **Domain Area**	**Basic statistical methods** mean, median, mode, ranges, variance, standard deviation, confidence intervals; standard error of the mean Data distributions, outliers **Statistical inference & hypothesis** t-test Chi-square Correlation Regression (linear, multi-linear)	**Spatial Analysis** **Measurement** – distance between stations **Topological analysis –** adjacency, proximity, within, neighbour **Surface analysis** – Interpolation of temperature to create continuous temperature surface; compare and contrast temperatures **Spatial Statistical analysis** – Point Pattern Analysis of Samples, Kernel Density Estimates, Nearest Neighbour Analysis	**Trends in the data** **Comparisons between data** **Models: prediction, forecasting** Working with models Parton & Logan, 1981 Model for hourly temperature predictions

Temperature Class Project – Data Collection, Mapping and Analysis

Instructions to conduct the different analyses for the temperature class project are included in this document. Deadlines will be provided in class and posted in canvas.

Overview of the Project.

Part	Data	Description
1.0	**Data Collection**	Data sampling, collection and logger placement
		Data sampling and collection
2.0	**Data Management**	Data collection, documentation and management of logger data
		Data management, open science and reporting
		Map the location of the loggers
3.0	**Data analysis I**	Data preparation, exploration and analysis
		Data preparation
		Statistical Analysis of the temperature data for the locations your group collected
		Analyse the temperature data using all of the data
		Spatial Analysis of the temperature data using all of the data spatially
4.0	**Data Analysis II**	Comparing temperatures with other data logger sites, other sources and predicted temperatures
		Logger Sites
		Meteorological Station
		Comparing temperatures with other data sources – Remotely sensed data
		Exploring models to predict temperature
5.0	**Reflection**	Discussion and conclusion
		Reflect on the analysis, process and what you would differently

Table of Contents

Part 1 – Data planning, sampling and collection

By the end of this part of the project you will be familiar with utilising maps in different formats for planning sampling and data collection.

Software:

- Google Maps on mobile phones/laptop
- Paper maps if available. Paper maps can include tourist maps, walking/hiking/biking maps, ordinance survey maps, US geological survey maps, and other maps from national mapping agencies.

1. **Geography and temperature.** Review a map of the study area. Also examine the study area using Google Map. Use the different background maps (imagery, terrain, cartographic map) and streetview and the 3D view, if available. Discuss the geography of the study area. What features may be present and how, if at all, do these features affect temperature? Where to place the temperature loggers in the environment?
2. **Distribution and placement of temperature loggers.** Examine Figure 1. Following on from the discussions what sampling method is appropriate? What factors are you considering in your placement? How many loggers would you like to place?

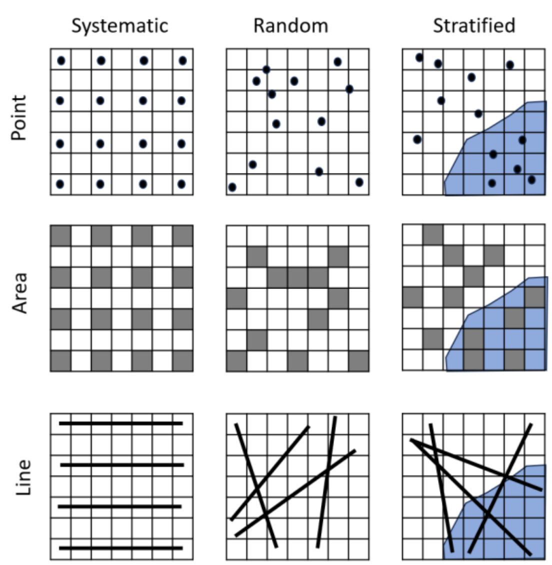


Figure 1: Vary of sampling strategies that can be considered based on points, areas or lines. Image created by Blanford (2024)

Part 2 – Data Management

Recording the logger locations and creating a map of the spatial distribution of the loggers.

By the end of this part of the project you will be familiar with data capture, data organisation and structuring, data definitions, data standards and data management and sharing.

Software:

- **Excel**
- **ArcGIS Pro**

If you are new to Excel and ArcGIS review the introductory videos and the information provided in the Appendix.

LINKs to Excel

Video	**Description**	**Total Time**	**Link**
Intro to Excel	Learn the essentials about Excel	2:27	https://youtu.be/Rogy-w0U_aQ

LINKs to ArcGIS

Video	**Description**	**Total Time**	**Link**
Intro to ArcGIS Pro	Learn the essentials about ArcGIS Pro	3:58 6:35 3:52	https://www.youtube.com/watch?v=bmjDQ2mvGIl https://www.youtube.com/watch?v=h3EozDlYHrI https://www.youtube.com/watch?v=JNtoAoiMmUU

EXCEL: Capturing the data in Excel and exporting the information as a CSV file.

In Excel, record the distribution of the loggers in the excel spreadsheet provided.

Group	Logger id	Description of location (or insert picture)	Sun / shade	Height above ground	elevation	Lat	Long	No days	Dates	Interval (mins)
										30
										30

Once all groups have added the logger information, export the sheet as a CSV file so that you can import this into ArcGIS.

Save the spreadsheet information in Excel to a CSV file (comma delimited file) – **station_loggers.csv**

ARCGIS: Create a map of the location of all of the loggers. Create a map of the spatial distribution of the loggers in the study area.

To create a new layer/geodatabase from the CSV file.

Add CSV file (station_loggers.csv) to ArcGIS

Once the CSV file has been added, drag and drop from ArcCatalog and place in the map area or the contents pane.

Right click on the CSV file. A menu will appear. From the list of options select – Display XY Data. Select the Lat and Long fields to map. Click OK. Once the layer has been added double check the mapped output. If the data is displaying correctly then right click the layer and export the layer and create a new permanent layer. This will save the data as a new geodatabase layer in the geodatabase. Save the layer as **station_loggers**.

Insert a map into your report of the study area and the locations of the loggers in this study.

Part 3 – Data analysis I

In this part you will analyse the data using statistical and spatial analysis methods.
By the end of this part of the project you will be familiar with various statistical and spatial analysis methods as well as visualisations.

Software:

- **Excel**
- **R**
- **ArcGIS**

If you are new to Excel and ArcGIS review the introductory videos and the information provided in the Appendix.

LINKs to Excel

Video	Description	Total Time	Link
Intro to Excel	Learn the essentials about Excel	2:27	https://youtu.be/Rogy-w0U_aQ

LINKs to ArcGIS

Video	Description	Total Time	Link
Intro to ArcGIS Pro	Learn the essentials about ArcGIS Pro	3:58 6:35 3:52	https://www.youtube.com/watch?v=bmjDQ2mvGII https://www.youtube.com/watch?v=h3EozDlYHrI https://www.youtube.com/watch?v=JNtoAoiMmUU

LINKs to R

Video	Description	Total Time	Link
Intro to R Studio	Learn the essentials about R	5:38	https://youtu.be/qL_38YRXJIE

Conduct descriptive statistical analysis of the data (Q1-Q2)

Q1. Descriptive statistical analysis of temperature. For each station you collected data for, calculate the average temperature recorded, examine the minimum and maximum temperature and the range in temperatures. Create a graph to show the temperatures and range of temperatures collected. How much did the temperature change, did you notice any trends in the data?

Q2. Variability. How much variability occurred during the time the data was collected for your sites?

Preparing the logger temperature data for statistical analysis.

During the part you will get the data ready for analysis. To do so all of the data will be reorganized, cleaned and set up in a single file.

Excel file: temperatureall.xlsx

EXCEL. Adding data to Sheet TEMPALL.

In the excel file (temperatureall.xlsx) add the temperature data your group collected into the columns that have been set up.

	A	B	C	D	E	F	G	H	I	J
1	recno	groupno	loggerid	date	ta	year	month	day	hour	min
2	1	1	10931484	########	19,76	###	9	1	0	0
3	2	1	10931484	########	19,66	###	9	1	0	30
4	3	1	10931484	########	20,14	###	9	1	1	0
5	4	1	10931484	########	20,04	###	9	1	1	30
6	5	1	10931484	########	19,66	###	9	1	2	0
7	6	1	10931484	########	19,66	###	9	1	2	30

1 – add your group number

2 - add the loggerid (you can find this at the top of the csv file)

3 – Copy the date and temperature from the csv file and paste it into columns D and E in the **temperatureall.xlsx** on the sheet named **tempall**. Now scroll to the end of where the data has been pasted and add the information from the next logger adding in the same information again.

4- Add in a sequential number for recno

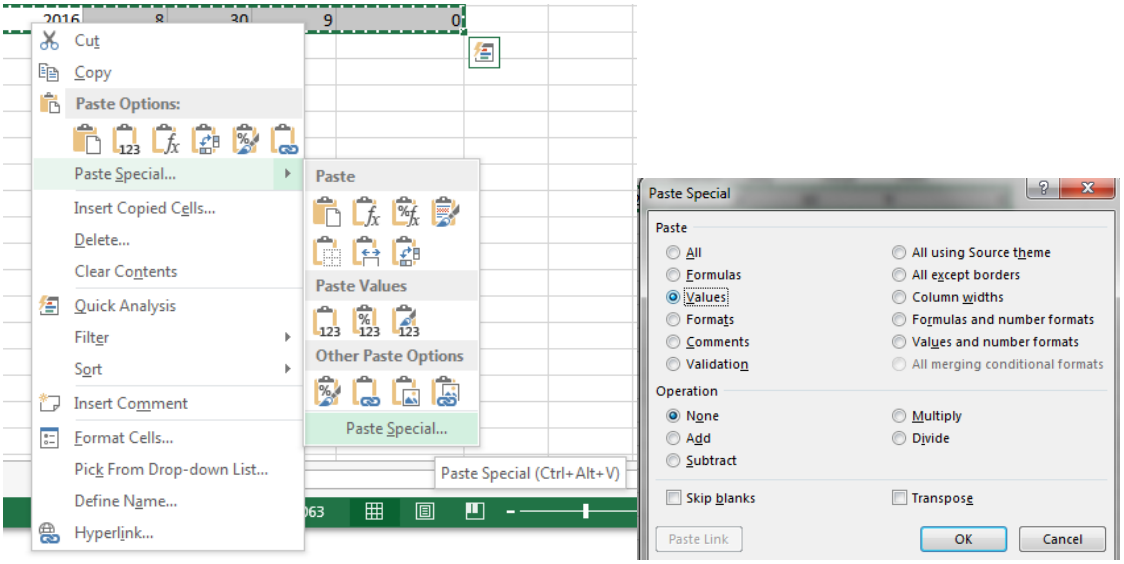
5 –Extract the year, month, day, hour and minute from the date field using the functions in Excel.

To do so in the following columns type

year =year(select cell)
month =month()
day =day()
hour =hour()
minute =minute()

ONLY do this for the first record, then select the record and double click the little black square in the lower right hand corner of the selection box, as shown in the image below. This will automatically apply the formulas for all records that contain data.

Remember that you have applied a formula. If you delete the data or move it around then the formula will recalculate. Select the columns and copy and paste the data **paste data as values**. Select Paste Special – Paste Special – Paste as Values.



EXCEL - Adding data to Sheet TEMP1.

In the excel file (temperatureall.xlsx) add the temperature data your group collected into the columns that have been set up.

Copy the date and temperature to the sheet named **temp1**.

Paste the data from each logger side by side and name the top of the column the loggerid number. Ensure that the time and date match for each row. This will be used for conducting the boxplot analysis and for summarizing the data in ArcGIS.

For ArcGIS add a letter in front of the logger id number otherwise this will not be considered a variable name.

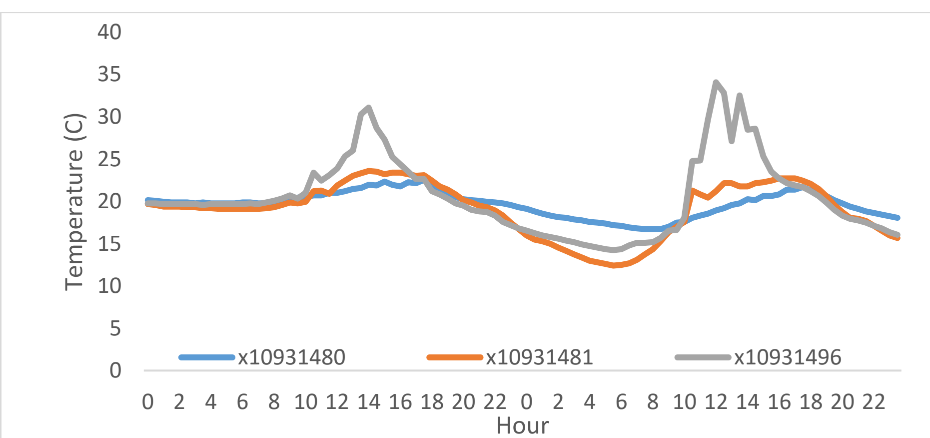
	A	B	C	D	E
1	recno	date	x10931480	x10931481	x10931496
2	1	#######	20,138	19,662	19,758
3	2	#######	20,043	19,567	19,758
4	3	#######	19,948	19,377	19,662
5	4	#######	19,853	19,377	19,662
6	5	#######	19,853	19,377	19,662

4- Add in a sequential number for recno. This will be a unique identifier so that you can always return to the original record in the data.

5 –Extract the year, month, day, hour and minute from the date field using the functions in Excel. See the instructions above. These will be used later for aggregating and summarizing the data by month and hour.

6-Visualise the data.

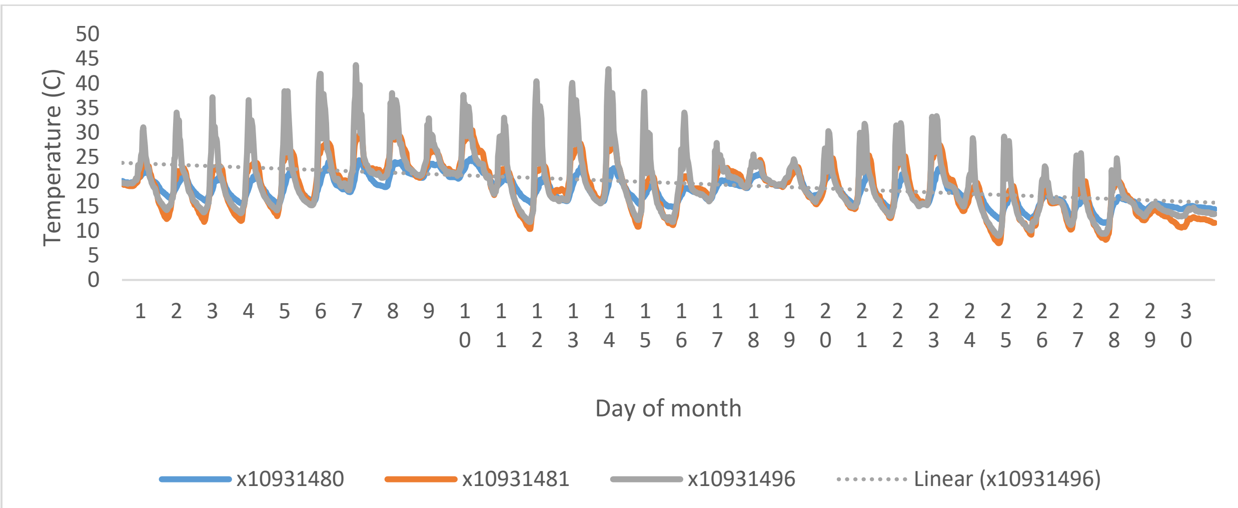
Create a graph of the data in sheet temp1 and examine how similar or different the temperatures were for your sites.



Examine the temperatures using a variety of visualisations.

Conditional formatting of cells creates a heatmap based on the temperatures.

x10931496	x10931481	x10931480
19,758	19,662	20,138
19,758	19,567	20,043
19,662	19,377	19,948
19,662	19,377	19,853
19,662	19,377	19,853
19,662	19,282	19,853
19,662	19,282	19,758
19,567	19,187	19,853
19,662	19,187	19,758
19,662	19,092	19,758
19,662	19,092	19,758
19,662	19,092	19,758
19,662	19,092	19,853
19,758	19,092	19,853
19,662	19,092	19,758
19,853	19,187	19,758
20,043	19,282	19,853
20,329	19,567	19,948
20,71	19,853	20,234
20,329	19,758	20,424
20,996	19,948	20,519
23,388	21,187	20,71
22,429	21,282	20,71
23,1	20,901	20,996
23,869	21,855	20,996
25,319	22,429	21,187
26	23,004	21,473
30,255	23,292	21,569
31,064	23,581	21,951
28,655	23,484	21,855
27,272	23,196	22,333
25,222	23,388	21,951
24,351	23,388	21,76
23,484	23,196	22,238
22,621	23,004	22,142
22,621	23,1	22,525
21,187	22,429	21,951
20,805	21,76	21,282
20,329	21,378	20,805
19,758	20,805	20,519
19,472	20,138	20,234
18,996	19,853	20,138
18,806	19,472	20,043
18,711	19,282	19,948
18,331	18,901	19,853
17,57	18,331	19,758
17,189	17,475	19,567
16,808	16,713	19,282
16,523	15,951	19,092
16,237	15,473	18,806
15,951	15,282	18,521
15,76	14,996	18,331
15,569	14,517	18,14
15,378	14,134	18,045
15,187	13,75	17,855
14,9	13,365	17,76
14,709	12,98	17,57



Summarize the data with descriptive statistics

In this section you will summarize the data in several ways

1- Monthly – used to create a temperature surface for the study area.
2- Daily – used to compare between sites and other data sources (Met Station; NOAA Satellite data)
3- Hourly – used to compare between sites, examine temperature variability and to predict temperatures and explore the Parton & Logan Model.

Summary statistics can be accomplished in Excel or in ArcGIS.

Software:

- Excel
- ArcGIS

EXCEL – Calculate Mean, Min, Max Monthly Temperature and Mean, Min, Max Daily Temperature

The average temperature for the month and each day can be calculated manually using the following functions

Calculating mean, min, max temperatures for the month
=average(select all the rows in the ta column for that month)
=min(select all the rows in the ta column for that month)
=max(select all the rows in the ta column for that month)
=stdev.s(select all the rows in the ta column for that month)
=count(select all the rows in the ta column for that month)

Calculating mean, min, max temperatures for a day
=average(select all the rows in the ta column for a day)
=min(select all the rows in the ta column for a day)
=max(select all the rows in the ta column for a day)
=stdev.s(select all the rows in the ta column for a day)
=count(select all the rows in the ta column for a day)

BUT this can get a little tedious and also after doing this a few times you will notice how easy it is to make mistakes.

ARCGIS – Summarising data for each station by month, day, hour

In ArcGIS you will create a number of different summaries.

Loggers	Statistics Type	Case Field	Output
GroupLoggers (N=3)	Mean, min, max, standard deviation, range	loggerid	**Summary_group_logger**
GroupLoggers (N=3)	Mean, min, max, standard deviation, range	loggerid, day	**Summary_group_logger_day**
GroupLoggers (N=3)	Mean, min, max, standard deviation, range	loggerid, hour	**Summary_group_logger_hour**
GroupLoggers (N=3)	Mean, min, max, standard deviation, range	loggerid, day, hour	**Summary_group_logger_dayhour**

- Select the **Tempall** sheet and save it as a CSV file. Save it in the directory where you have your temperature data. Close Excel.

	A	B	C	D	E	F	G	H	I	J
1	recno	groupno	loggerid	date	ta	year	month	day	hour	min
2	1	1	10931484	########	19,76	###	9	1	0	0
3	2	1	10931484	########	19,66	###	9	1	0	30
4	3	1	10931484	########	20,14	###	9	1	1	0
5	4	1	10931484	########	20,04	###	9	1	1	30
6	5	1	10931484	########	19,66	###	9	1	2	0
7	6	1	10931484	########	19,66	###	9	1	2	30
8	7	1	10931484	########	19,66	###	9	1	3	0

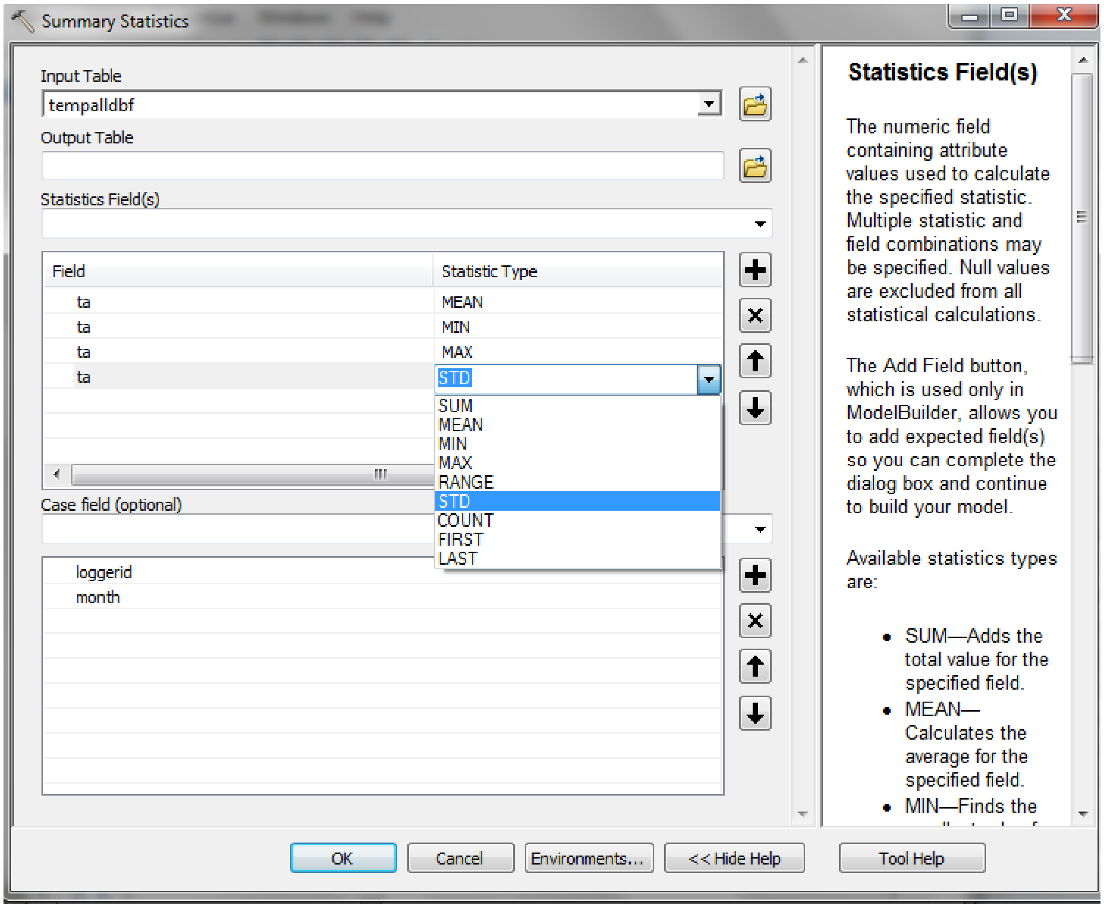
- Open ArcGIS – either as a new file or use the existing temperature project created earlier.
- In ArcCatalog - drag and drop the csv file from ArcCatalog and place it in the mapping area of ArcGIS or the Table of Contents.
- Once loaded, right click on Temp1.csv in the Table of Content and check to see that all of the data has been brought in correctly. All of the variables are good except for date. But we don't need the original date format since we have split the information into separate fields.
- If all looks good, then export the data to the default file format. The reason for this is ArcGIS will not perform an analysis on a csv file.

Now the data is ready for analysis in ArcGIS.

- In the ArcToolbox navigate to summary statistics and summarize the data. Go to Analysis Tools – Statistics – Summary Statistics.
- Create monthly, daily and hourly summaries of the temperature data. The instructions are on the following pages.

ARCGIS - MONTH: Summarizing the data by month

Output: **Summary_group_logger**



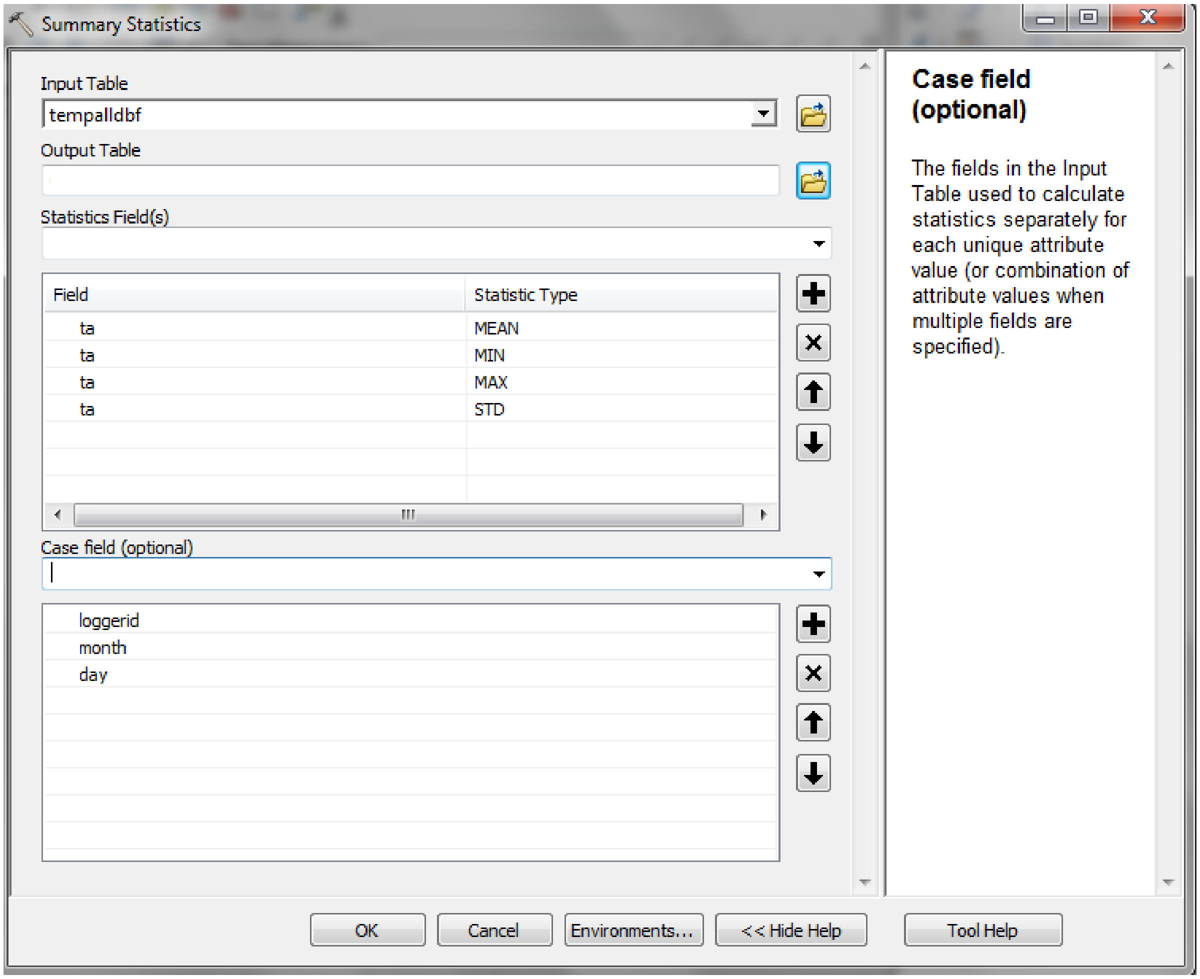
The output table will look something like this:

OBJECTID *	loggerid	month	FREQUENCY	MEAN_ta	MIN_ta	MAX_ta	STD_ta
1	10931482	8	78	21.441795	16.05	26.2	3.166455
2	10931482	9	1440	18.800264	6.37	30.86	5.028475
3	10931482	10	227	18.640969	12.3	28.16	3.571886

Now export the table as a txt file (which is really a CSV file so change the file type to csv) and open in excel and copy the information to the relevant spreadsheets.

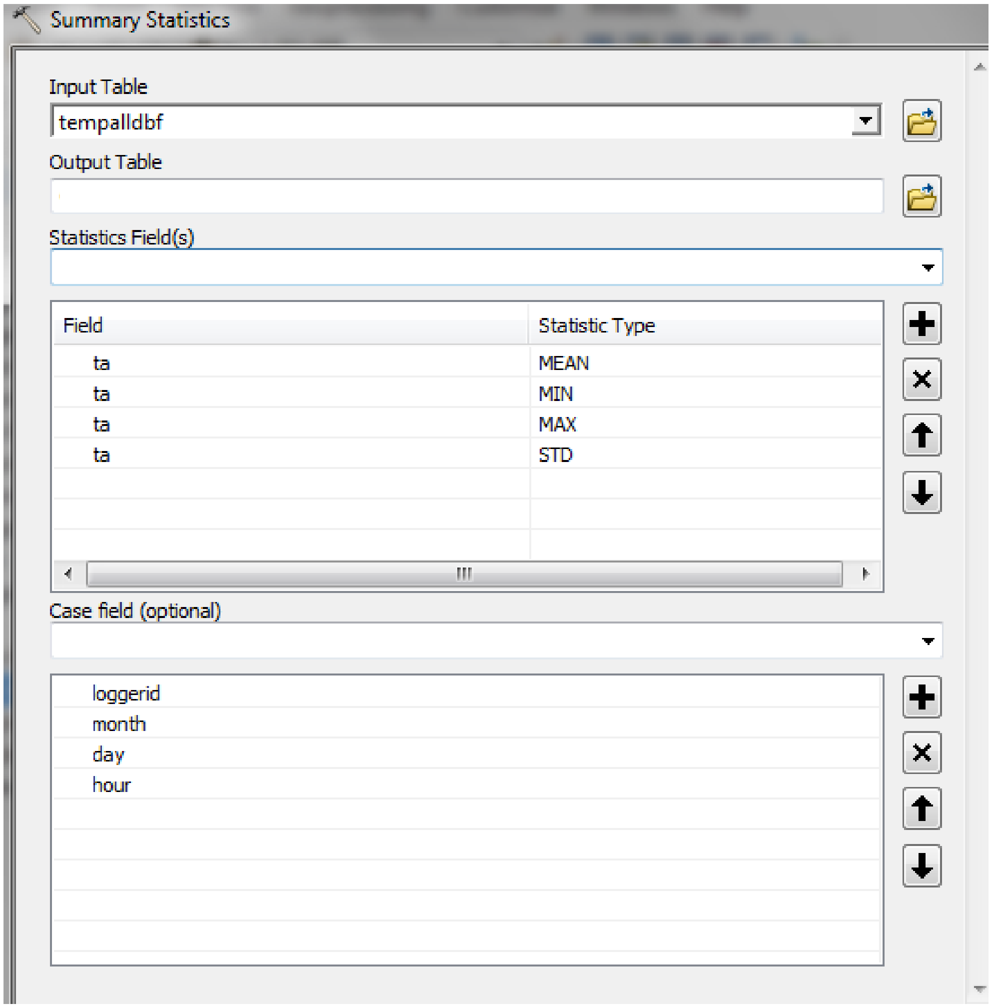
ARCGIS - DAILY: Summarizing the data by day.

Output: **Summary_group_logger_day**



ARCGIS - HOURLY: Summarizing the data for each hour of the day.

Output: **Summary_group_logger_dayhour**



Blanford, J.I. (2025) Teaching material. Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.

Visualize the data

Create a graph in Excel for at least one logger as shown below and add a trend line.

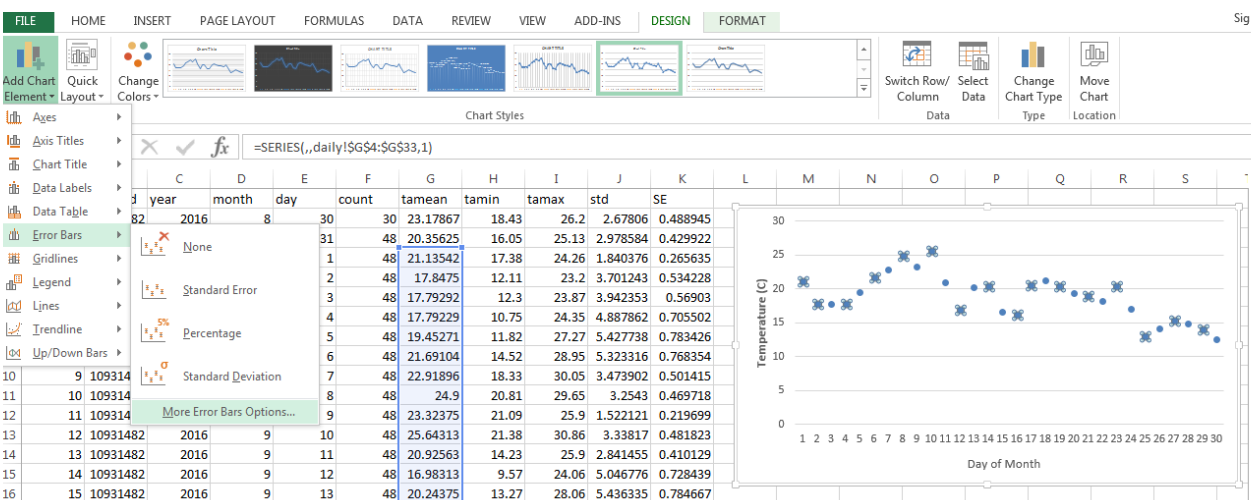
Excel - Calculating Standard Error of the Mean (SE) and adding Error bars to the mean daily temperature

To Calculate SE of the mean. In a new column using the daily temperature summary values use the following formula

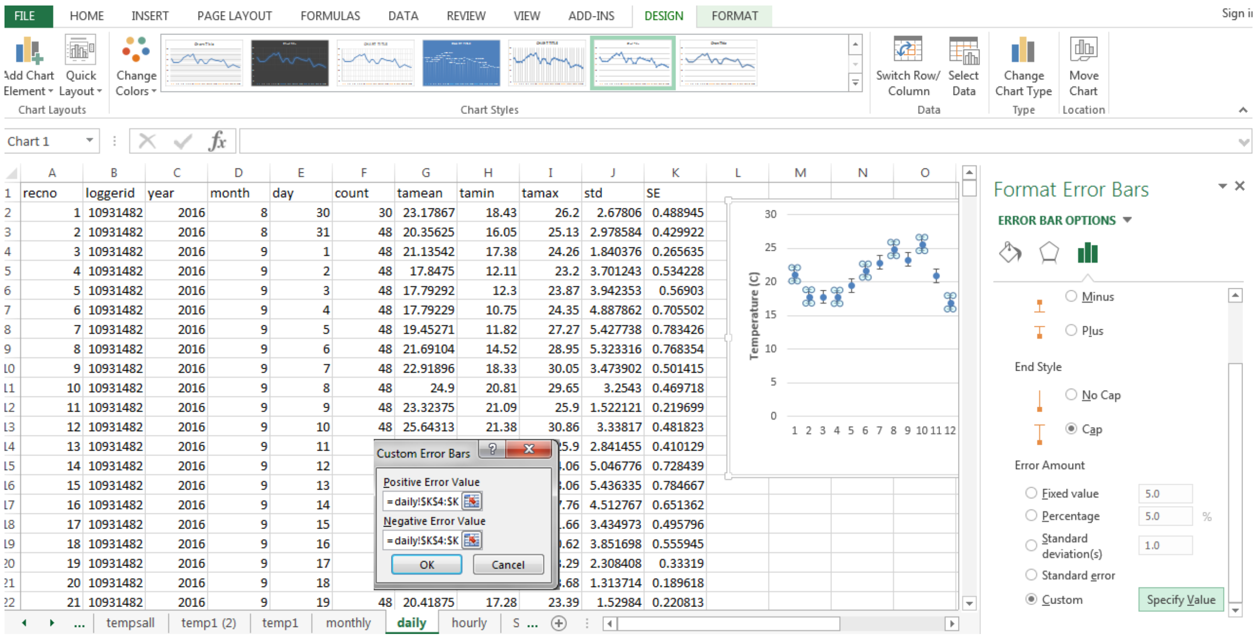
=STDEV(range) / SQRT(COUNT(range)) because you have already calculated the Standard deviation

(cell with standard deviation value)/sqrt (count or frequency)

Adding Error Bars



Select Custom (Specify Value) – For positive error bars select the rows for SE for Sep 1 to Sep 30 and then for the Negative error bars select the same values.



Create a figure with all of data from your loggers. Show data from September 1st to September 30th.

R /R Studio

Create a boxplot for each of the loggers.

Data – use the mean daily temperatures.

The temperature data will need to be set up as follows.

This can be done in Excel. Once the data has been organized as follows, save the sheet as a CSV file and complete the analysis in R.

Hour	x10931480	x10931481	x10931496
0	20.138	19.662	19.758
0	20.043	19.567	19.758
1	19.948	19.377	19.662
1	19.853	19.377	19.662
2	19.853	19.377	19.662

In R you use the following code.

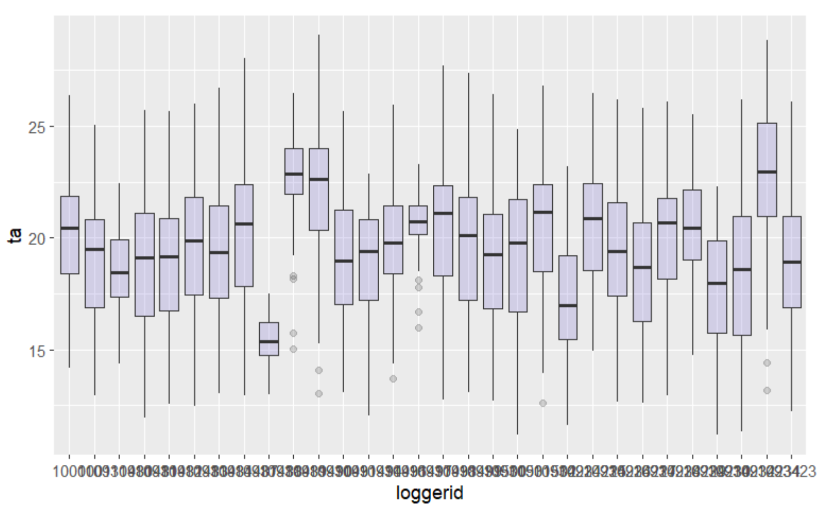
```
install.packages("ggplot2")
library(ggplot2)
tempsallfinaldaily_sep <- read.csv("C:/jb/teaching/class/tempsallfinaldaily_sep.csv")

ggplot(tempsallfinaldaily_sep, aes(x=as.factor(loggerid), y=ta)) +

  geom_boxplot(fill="slateblue", alpha=0.2) +

  xlab("loggerid")
```

The output will look as follows:



Spatial analysis of the temperature data using all of the data.

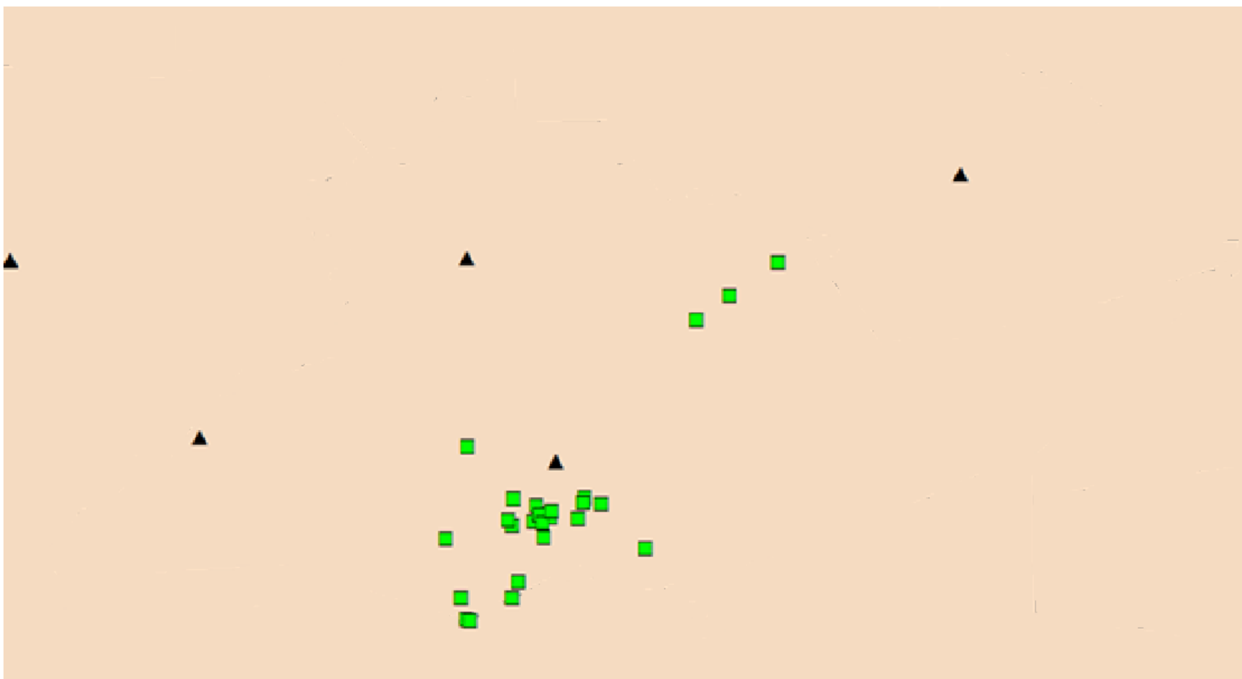
Q3.Temperature surface for the study area.
Create a continuous temperature surface of the mean temperature using all of the data loggers collected by the class? Where was the coldest/hottest location? Describe factors that might influence these differences?

Q4. Descriptive statistical analysis of temperatures.
What was the minimum/maximum and mean temperatures? What was the range in temperatures? How different were the temperatures between the hottest and coldest location?

Mapping and analysing the logger temperature data.

In part 2 you created a spatial dataset for the distribution of the loggers. This contained the following information. The loggerid is the unique identifier that will be used to connect the temperature data collected by the loggers with the location of the loggers.

Group	Logger id	Description of location (or insert picture)	Sun / shade	Height above ground	elevation	Lat	Long	No days	Dates	Interval (mins)
										30
										30



Distribution of all of the loggers (N= 30). Green squares = logger locations; black triangles = Met Stations

In part 2 you created monthly summaries of the temperature data for each of the loggers within your group. In this next step the monthly temperature for all loggers will be used.

ARCGIS - Calculate monthly temperatures for each logger location.

Repeat the data preparation that you conducted for the group loggers using all of the logger data

Loggers	Statistics Type	Case Field	Output
AllLoggers (N=28)	Mean, min, max, standard deviation, range	loggerid	**Summary_all_logger**
AllLoggers (N=28)	Mean, min, max, standard deviation, range	loggerid, day	**Summary_all_logger_day**
AllLoggers (N=28)	Mean, min, max, standard deviation, range	loggerid, hour	**Summary_all_logger_hour**
AllLoggers (N=28)	Mean, min, max, standard deviation, range	loggerid, day, hour	**Summary_all_logger_dayhour**
AllLoggers (N=30)	Mean, min, max	logger, day	**Summary_all_logger_metstn_day**

Summary_all_logger will be used to map the mean temperatures for the study area and serve as input for the interpolation analysis. This will contain the monthly temperature information.

Add **tempsallfinal_sep.csv** file to ArcGIS.

Export the CSV file to a geodatabase file in ArcGIS.

Use summary statistics to summarise the data for each loggerid (mean, min, max, stdev, count) **(summary_all_logger).**

Export the file as a CSV file for further analysis in R and Excel.

ARCGIS - Map the monthly temperature data

Create a temperature layer for each of the loggers

In ArcGIS join the **summary_all_logger** to the **station_loggers**.

Right click on the station_loggers layer in the table of contents. Select Joins and Relates – Join.

Join the datasets on the loggerid field, select the summary_all_logger data set, Keep all records.

For each loggerid in station_loggers the join function will match the loggerid based on the first record it finds with the same loggerid from the **summary_all_logger table.**

Check the information has joined correctly. Open the attribute table by right clicking on the logger locations layer. Check that each logger has a temperature value associated with it.

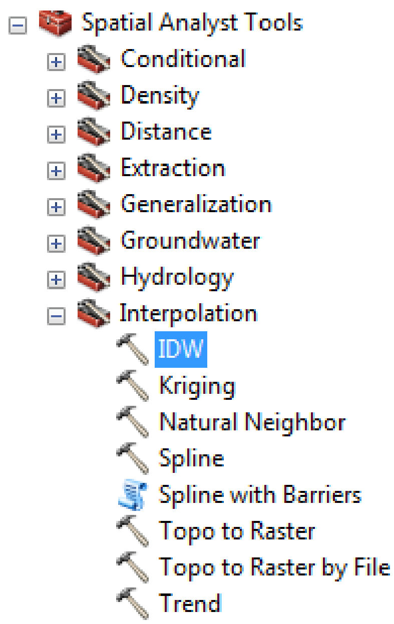
If the join is correct export the data and create a permanent dataset with the values attached. To do this right click on the station_loggers layer, find data – Export Data and then export data as a geodatabase feature class. To save this as a shapefile, add .shp to the end of the of the filename.

Save this as file: **loggers_mnth_tamean**

Create a continuous temperature surface for the study area

Input file: **loggers_mnth_tamean**

In ArcGIS in ArcGIS Toolbox navigate to the interpolation tools. Spatial Analyst Tools – Interpolation – IDW. A number of different interpolation methods are available (e.g. Natural Neighbors, Trend method, Inverse Distance Weighting, Spline and a variety of Geostatistical/Kriging methods (Source: (ESRI 2021); (Childs 2004)). For the purpose of this exercise we will use the inverse distance weighting (IDW) method.



Interpolation is a procedure used to predict the values of cells at locations that lack sampled points. It is based on the principle of spatial autocorrelation or spatial dependence (e.g. Tobler's First Law of Geography), which measures degree of relationship/dependence between near and distant objects (Childs 2004). Spatial autocorrelation determines if values are interrelated. If values are interrelated, it determines if there is a spatial pattern. This correlation is used to measure (Childs 2004):

- Similarity of objects within an area
- The degree to which a spatial phenomenon is correlated to itself in space
- The level of interdependence between the variables
- Nature and strength of the interdependence

Overview of Inverse Distance Weighting (Source: (ESRI 2021); (Childs 2004)).

Deterministic	**Inverse distance weighted (IDW)** interpolation makes the assumption that things that are close to one another are more alike than those that are further away. To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location (number of points). The measured values closest to the prediction location have more influence on the predicted value than those farther away (e.g. set using the power function). IDW is an exact interpolator and will never predict values above the maximum measured value or below the minimum measured value.

In ArcGIS, go to Spatial Analyst Tools – Interpolation - IDW. Start with the default settings and then try different power settings and set the number of points to use to calculate temperature values between the logger locations and see how the outputs change.

If you do use different settings, summarise the different setting you used.

	Power	**Number of Points**
Default	**2**	**12**
	1	12
	3	12

Conduct the IDW analysis and insert the IDW output you found captured the temperatures best for your study area. Include the parameters you used.

- What was the minimum / maximum and mean temperatures found in the study area.
- Where were the coldest / hottest locations?
- How different were the temperatures between the hottest and coldest locations? Describe what factors might have influenced these differences?
- Did you take any additional data into consideration when you examined the differences in temperature?

Part 4 – Data Analysis II

Comparison of local temperature data with other logger sites, other data sources and modeled predictions

For this part you will compare the temperature data you collected with the loggers with the temperature data collected from different sources.

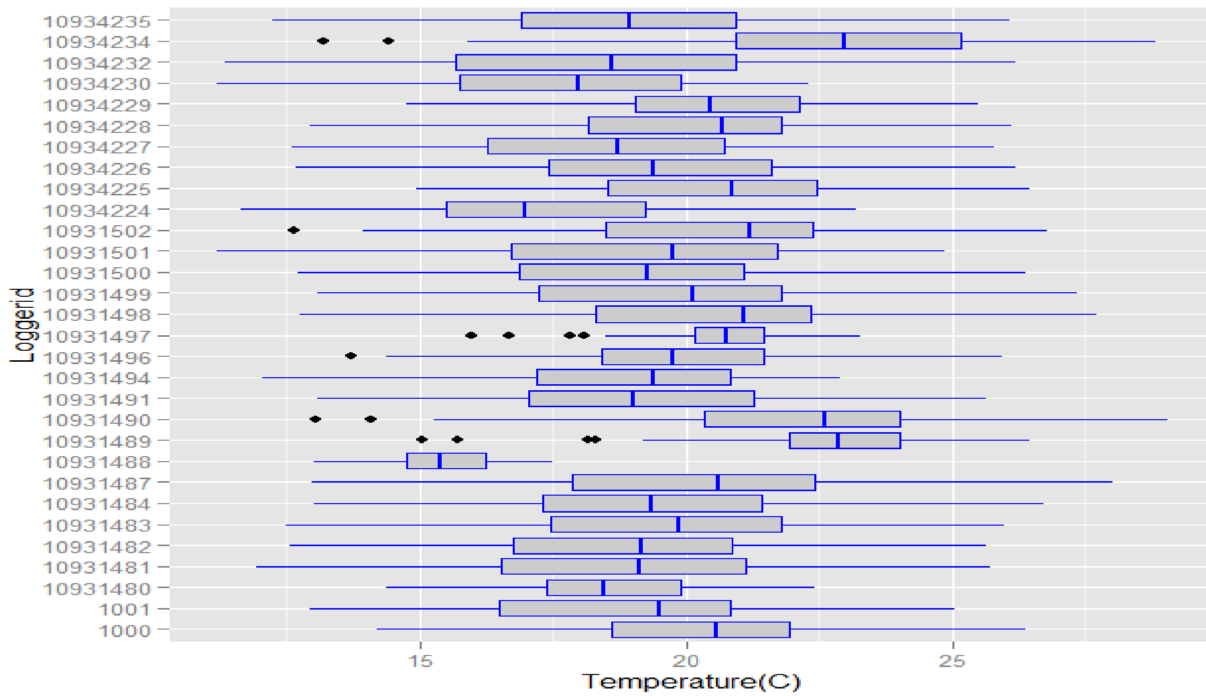
Q5. Comparison of logger and Meteorological Station temperature data.

Compare the temperatures collected using the loggers and the closest Met Station. Describe what you did to compare the logger data with the Met Station data.

- How different/similar were the temperatures you collected when compared with those from the nearest Met Station?
- Where was the closest Met Station (calculate distance).
- What was the overall temperature difference between the loggers and Met Station? (data will be provided and obtained from (https://www.ncei.noaa.gov/maps/daily/)

Comparison of temperatures between Logger locations

Create boxplots using **tempsallfinal_sepR.csv** and examine how variable temperatures were across the study area. Here are multiple boxplots of temperatures from across the different sites. This is only a subset of the data.



#create a boxplot of all of the loggers for the temperature data.

```
tempdata <- read.csv("C:/directory/ tempsallfinal_sepR.csv ", header = TRUE, sep = ",")
temp.df = tempdata
temp.df$loggerid = factor(temp.df$loggerid)
ggplot(temp.df, aes(x = loggerid, y = ta)) +
  geom_boxplot(fill = "grey80", colour = "blue") +
  scale_x_discrete() + xlab("Loggerid") +
  ylab("Temperature(C)")
```

#flip the plot of boxplots so that you can fit all of the loggers into a single plot

```
boxflip <- ggplot(temp.df, aes(x = loggerid, y = ta)) +
  geom_boxplot(fill = "grey80", colour = "blue") +
  scale_x_discrete() + xlab("Loggerid") +
  ylab("Temperature(C)")
boxflip + coord_flip()
```

#Save the plot to a jpeg file
jpeg('rboxplot.jpg')
boxflip + coord_flip()
dev.off()

Examine the boxplots. Do the error bars overlap. If they do then it shows that there are no real difference between-groups. The boxplots will show the spread of values for each site and allow you to see how similar the values are between the groups.

Comparison of temperatures between Logger and Met Station

Examine the boxplots. Do the error bars overlap. If they do then it shows that there are no real difference between-groups. The boxplots will show the spread of values for each site and allow you to see how similar the values are between the groups.

Q6. Comparison of logger and Land Surface Temperature (LST) data.

Compare the temperatures collected using the loggers with the LST. Describe what you did to compare the logger data with the LST.

- How different/similar were the temperatures you collected with the LST?

Input file: **loggers_mnth_tamean**

Add the land surface temperatures to the logger data in ArcGIS.
Spatial Analysis Tools – Extraction – Extract multi values to points
Select the logger shapefile you created that has the temperature data that you collected.

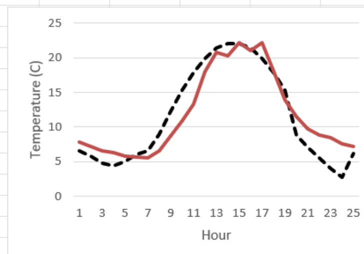
How different are the temperatures between the loggers and the LST?
What is the resolution of each grid cell for the LST?

Q7. Predicting hourly temperatures.

Collecting and sampling data from a variety of locations is costly therefore being able to predict temperatures can be useful. Parton & Logan (1981) developed a sine-exponential model that can be used to predict temperatures. During this exercise you will use this model to predict temperatures and compare how similar the predictions are to the logger data.

For this exercise you will use the Excel spreadsheet, as illustrated below, to explore the Parton & Logan (1981) model and adjust the parameters in the model and find the best fit you can with the model.

Cut and paste one day of temperatures into the column labeled Ta_Obs (Observed Temperature) Column (C9-C38)														
Adjust the parameters (A,B,C) Column (O4-O6)														
													Values	**Adjust A,B,C Values**
Summary	**MIN_Ta**	**MAX_Ta**	**Mean_Ta**		Daylength	**Sunrise_ti**	**Sunset_tir**	**MinTa_tin**	MaxTa_Tir	Tval	Hdecay		A = 1.5	1,5
	5,509667	22,18933	11,77288		12	6	18	6	13,5	2,513274	12		B = 2.8	2,8
													C = -0.1	-0,1
	Predicted_	Observed_Ta											Pival	3,141593
hour	Ta_Pred	**Ta_Obs**	Ta_Diff	Ta_Diff^2		TSNDN	TMINI	Arg	T	T_hr-sundown				
0	6,545271	7,857083	1,311813	1,720853		-10,4849	6,545271	-1,17647	5					
1	5,811425	7,246417	1,434991	2,0592		0,849109	5,811425	-1,41176	6					
2	4,800145	6,564083	1,763939	3,11148		16,46798	4,800145	-1,64706	7					
3	4,441195	6,267833	1,826638	3,336607		22,01183	4,441195	-1,88235	8					
4	5,064593	5,7915	0,726907	0,528393		12,38367	5,064593	-2,11765	9					
5	6,09719	5,6205	-0,47669	0,227233		-3,56442	6,09719	-2,35294	10					
6	6,58962	5,509667	-1,07995	1,1663		-11,1698	6,58962	**-2,58824**	11					
7	8,977564	6,605083	-2,37248	5,628666		8,977564	5,285129	-0,04928	0,20944					
8	12,2939	8,702	-3,5919	12,90173		12,2939	5,070406	-0,09856	0,418879					
9	15,31373	10,84492	-4,46881	19,97028		15,31373	4,87488	-0,14784	0,628319					
10	17,90507	13,308	-4,59707	21,1331		17,90507	4,707097	-0,19712	0,837758					
11	19,95468	17,99775	-1,95693	3,829582		19,95468	4,57439	-0,2464	1,047198					
12	21,37297	20,77442	-0,59856	0,358269		21,37297	4,48256	-0,29568	1,256637					
13	22,09796	20,32967	-1,76829	3,126863		22,09796	4,435619	-0,34496	1,466077					
14	22,09796	22,18933	0,091373	0,008349		22,09796	4,435619	-0,39424	1,675516					
15	21,37297	21,03008	-0,34289	0,117573		21,37297	4,48256	-0,44352	1,884956					
16	19,95468	22,1585	2,203818	4,856815		19,95468	4,57439	-0,4928	2,094395					



- Read the paper (link provided below) to learn about the model.
- Discuss how the model works in class
- Use the Parton & Logan (1981) model to predict hourly temperatures for 30 minutes (the time interval used to collect the temperature data at each logger). An excel spreadsheet will be provided. **LINK**
- Calculate RMSE

Explore the observed temperature values with the predicted temperature values

- Were you able to find a good fit for the temperature data?
- What were the final model parameters (A, B, C) that you settled on?
- Calculate the error between the observed and predicted data.
- What limitations are there with using this model?

Reference

Parton, W. J., and J. A. Logan. (1981) A model for diurnal variation in soil and air temperature. *Agricultural Meteorology* 23:205-216. doi: https://doi.org/10.1016/0002-1571(81)90105-9.

Finding a good model fit: The Parton & Logan Model

Using the Parton & Logan Model information provided (an excel spreadsheet and the Parton & Logan Model Paper will be provided).

- Open the PartonandLogan.xls file.
- Select one days worth of data and paste it into the column labeled "Observed"
- Adjust the parameters a,b,c to change the shape of the curve.
- Calculate the RMSE and assess how good a fit you achieved.

Were you able to find a good fit for the temperature data (yes or no)? What were the final model parameters that you settled on (what were the parameters for a, b and c)?

Calculating the RMSE (Root Mean Square Error)

Calculate the error between the observed and predicted data. One way to assess error is to compare the predicted value with the actual value and calculate the difference. Calculate the RMSE and assess how much error there is. A smaller value indicates a better fit.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}$$

Root mean square error – the ***root mean square error*** of a sample of *n* measurements is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - x_{\text{true}}}{x_{\text{true}}}\right)^2}$$

RMSE is a measure of average deviation, somewhat similar to standard deviation, but is concerned with deviations from the *true value* whereas *S* is concerned with deviations from the mean.

In excel calculate the difference between observed – expected
Then calculate the RMSE: using the following formula: =SQRT(SUMSQ(select range of cells with values)/COUNTA(select range of cells with values))

In R - RMSE <- sqrt(mean((y-y_pred)^2))

Part 5 – Reflection

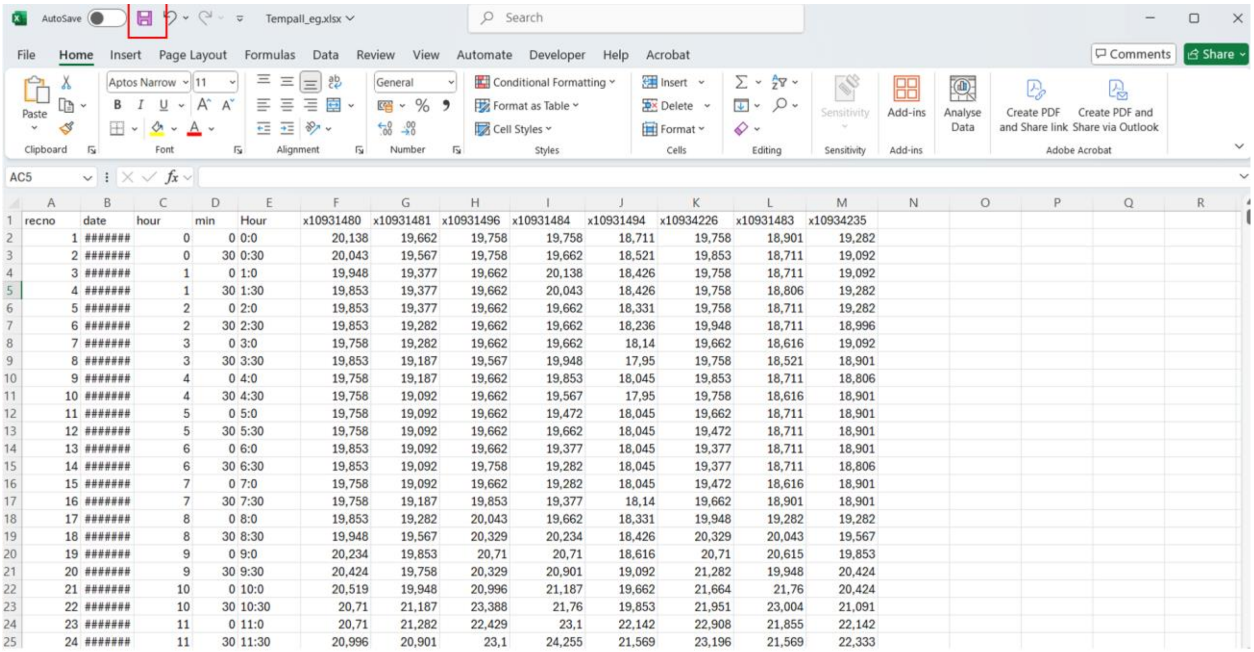
In this section discuss your findings and conclusions and reflect on the process, highlight limitations and challenges you encountered. Lastly reflect on the project itself and what you liked or found challenging.

Q8. Discussion and conclusion. Capture what you found out about temperatures in your study area.
Q9. If you had to do this again is there anything you would change? How might you place sensors in the environment in you had to do this again?
Q10. What challenges or issues did you encounter during the data collection and analysis, if any?
Q11. Were there any limitations associated with this study?
Q12. What did you like most about the project?
Q13. What did you like least / challenging about the project?

APPENDIX

New to Excel

Excel provides a rich data analysis environment with a variety of visualization and analytical tools and functions which can be expanded with a variety of add-ins.



recno	date	hour	min	Hour	x10931480	x10931481	x10931496	x10931484	x10931494	x10934226	x10931483	x10934235
1	#######	0	0	0:0	20,138	19,662	19,758	19,758	18,711	19,758	18,901	19,282
2	#######	0	30	0:30	20,043	19,567	19,758	19,662	18,521	19,853	18,711	19,092
3	#######	1	0	1:0	19,948	19,377	19,662	20,138	18,426	19,758	18,711	19,092
4	#######	1	30	1:30	19,853	19,377	19,662	20,043	18,426	19,758	18,806	19,282
5	#######	2	0	2:0	19,853	19,377	19,662	19,662	18,331	19,758	18,711	19,282
6	#######	2	30	2:30	19,853	19,282	19,662	19,662	18,236	19,948	18,711	18,996
7	#######	3	0	3:0	19,758	19,282	19,662	19,662	18,14	19,662	18,616	19,092
8	#######	3	30	3:30	19,853	19,187	19,567	19,948	17,95	19,758	18,521	18,901
9	#######	4	0	4:0	19,758	19,187	19,662	19,853	18,045	19,853	18,711	18,806
10	#######	4	30	4:30	19,758	19,092	19,662	19,567	17,95	19,758	18,616	18,901
11	#######	5	0	5:0	19,758	19,092	19,662	19,472	18,045	19,662	18,711	18,901
12	#######	5	30	5:30	19,758	19,092	19,662	19,662	18,045	19,472	18,711	18,901
13	#######	6	0	6:0	19,853	19,092	19,662	19,377	18,045	19,377	18,711	18,901
14	#######	6	30	6:30	19,853	19,092	19,758	19,282	18,045	19,377	18,711	18,806
15	#######	7	0	7:0	19,758	19,092	19,662	19,282	18,045	19,472	18,616	18,901
16	#######	7	30	7:30	19,758	19,187	19,853	19,377	18,14	19,662	18,901	18,901
17	#######	8	0	8:0	19,853	19,282	20,043	19,662	18,331	19,948	19,282	19,282
18	#######	8	30	8:30	19,948	19,567	20,329	20,234	18,426	20,329	20,043	19,567
19	#######	9	0	9:0	20,234	19,853	20,71	20,71	18,616	20,71	20,615	19,853
20	#######	9	30	9:30	20,424	19,758	20,329	20,901	19,092	21,282	19,948	20,424
21	#######	10	0	10:0	20,519	19,948	20,996	21,187	19,662	21,664	21,76	20,424
22	#######	10	30	10:30	20,71	21,187	23,388	21,76	19,853	21,951	23,004	21,091
23	#######	11	0	11:0	20,71	21,282	22,429	23,1	22,142	22,908	21,855	22,142
24	#######	11	30	11:30	20,996	20,901	23,1	24,255	21,569	23,196	21,569	22,333

Video series to learn Excel

Video	Description	Total Time	Link
Intro to Excel	Learn the essentials about Excel	2:27	https://youtu.be/Rogy-w0U_aQ

New to ArcGIS

Getting started with ArcGIS may seem overwhelming and can be since there is a lot of different pieces. In this section you will find a variety of materials that will help get you started and learn the essentials in 2-3 hours.

Getting the software.

If your institution does not have a license for you to use, then purchase a student or personal use licence from Esri.

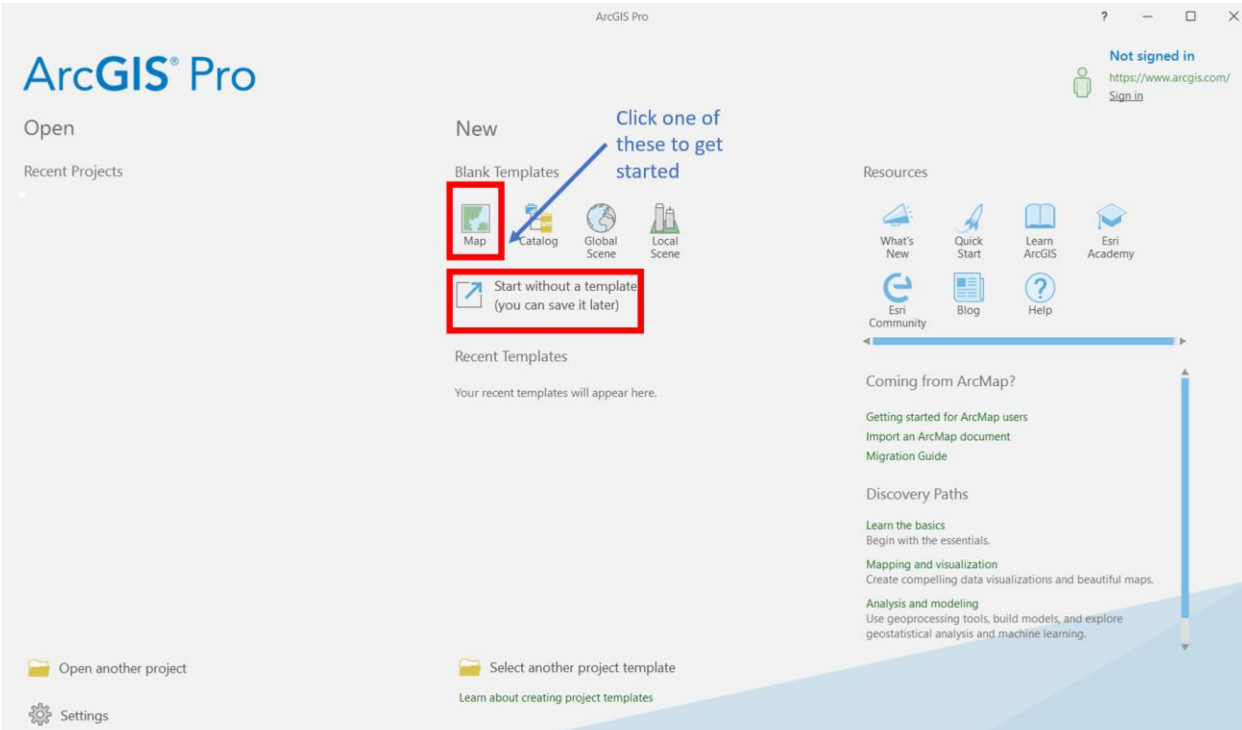
Video series to learn GIS and get familiar with ArcGIS

Video	Description	Total Time	Link
Overview of GIS	Module 100: Overview	5:20	https://www.youtube.com/watch?v=I7a42ZU3cXw
Intro to ArcGIS Pro	Learn the essentials about ArcGIS Pro	3:58 6:35 3:52	https://www.youtube.com/watch?v=bmjDQ2mvGIl https://www.youtube.com/watch?v=h3EozDlYHrI https://www.youtube.com/watch?v=JNtoAoiMmUU
Intro to Spatial Data 1-5	Module 110: Spatial Data 1-5	2:32 2:40 4:21 6:49 2:29	https://www.youtube.com/watch?v=WeAgkF-HX-s https://www.youtube.com/watch?v=JpNh9a8z5AY https://www.youtube.com/watch?v=lDFdc3QIpns https://www.youtube.com/watch?v=N7P4YltaIEs https://www.youtube.com/watch?v=N7P4YltaIEs
Intro to Visualisation and Cartographic Visualisations	Module 120: Visualisation Part 1	3:11	https://www.youtube.com/watch?v=XCtNMosgaAg
Intro to Spatial Analysis	Module 130: Spatial Analysis	4:34	https://www.youtube.com/watch?v=SlIKjV1bCZY

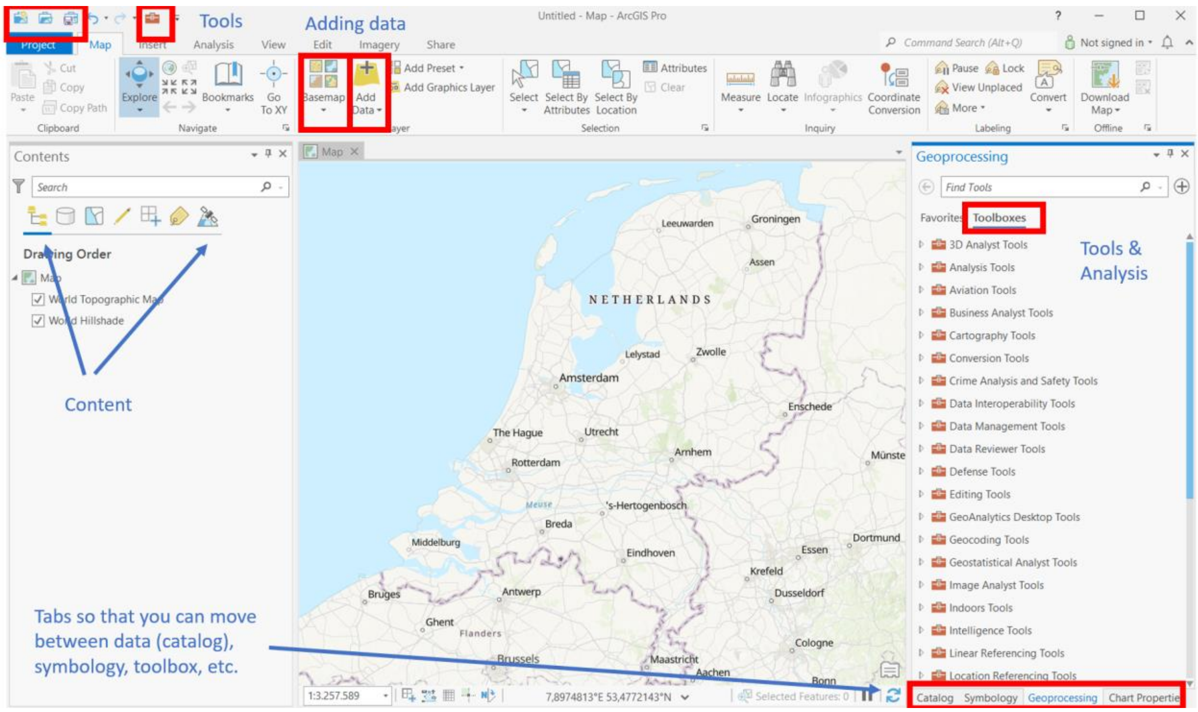
Overview of ArcGIS Pro

A quick overview of the software.

Start up page and getting into the software



Overview of ArcGIS Map Analysis Environment



Data catalog – connects to the data

By default files are stored in Documents/ArcGIS

Data are stored in Default.gdb

Projects are stored in Projects – projectname.aprx.

Any data created and saved to the project gdb can be found in the project specific project gdb file - e.g. Documents/ArcGIS/Projects/projectname/projectname.gdb

Symbology – allows you to change how the maps look

Toolboxes – the tools you will be using to manage or create data or perform analysis.

Familiarize yourself with different toolboxes as this is where the different analysis and functionality can be found.

During the course you will be using a variety of methods from different toolboxes. Details about each of the toolboxes and how the tools work can be found in the documentation. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/main/arcgis-pro-tool-reference.htm

Some key toolboxes to explore are:

Functionality	**Toolbox**
Analysis	Analysis Tools Geostatistical Analyst Tools Spatial Analyst Tool Spatial Statistics Tools
For data management	Conversion Tools Data Management Tools
Data editing	Editing Tools Geocoding Tools

Geographic Data files

There are a lot of different file types but here are some of the most popular ones. During this course we will use shapefiles mainly.

List of common GIS data files (source: geospatialsense)

Name	Extension	Description
Shapefiles	File.shp- contains the feature geometry file.shx – stores the shape index. File.dbf – stores attribute data. file.prj – contains projection information	Consist of a number of different files that contain information about the geography, typology, attributes, projections etc. Shapefiles consist of a number of files that MUST be kept together
geodatabase	.gdb	Personal geodatabase format. Similar to an MS Access database
Geojson	.geojson; .json	
GPS data files	.gpx	XML GPS data file, usually coming from a GPS device.
Google	.kml .kmz	XML plaintext file that may contain geometry, data, a web service Z files are compressed KML files
Sqlite/spatialite database	.sqlite	Spatialite is an extenion to sqlite to enable it spatially. Format often used with QGIS.
TIFF and TIFF World File, GeoTIFF	.tiff	Image data when associated with a tfw ("world file") of the same filename is a georeferenced image.
JPEG	.jpg	image data when associated with jpw (world file) is a georeferenced image. .jpx can contain additional metadata.
Ascii grid	.asc	Plaintext (ASCII) format
Grid files	Grid.bnd Grid.hdr Grid.sta	Compilation of files .bnd – boundary/extent metadata for the grid .hdr – metadata such as grid cell size .sta – grid cell statistics .vat – grid attribute table. Integers only
Comma delimited data files	.csv	Data stored in comma delimited format
Text data	.txt	can be .csv or tab delimited or just plain text
Tab-delimited data	.tab or .txt	Data stored in tab delimited format
Standard database file	.dbf	Data stored in database file format
Spreadsheet	.xls .ods	Data stored in spreadsheets by MS Excel or Open spreadsheets

Common project files for ArcGIS and ArcGIS Pro

Project files	File extension	source
ArcGIS	Mxd (.mxt, .lyr, .loc)	Esri
ArcGIS Pro	aprx	

By default files are stored in Documents/ArcGIS

Data are stored in Default.gdb

Projects are stored in Projects – projectname.aprx.

Any data created and saved to the project gdb can be found in the project specific project gdb file - e.g. Documents/ArcGIS/Projects/projectname/projectname.gdb

Blanford, J.I. (2025) Teaching material. Temperature variations across space and time: making statistics fun with geographic information and spatial analysis.

Additional Sources

For more in depth tutorials you can work through the quickstart tutorials provided by ESRI.

https://pro.arcgis.com/en/pro-app/latest/get-started/pro-quickstart-tutorials.htm

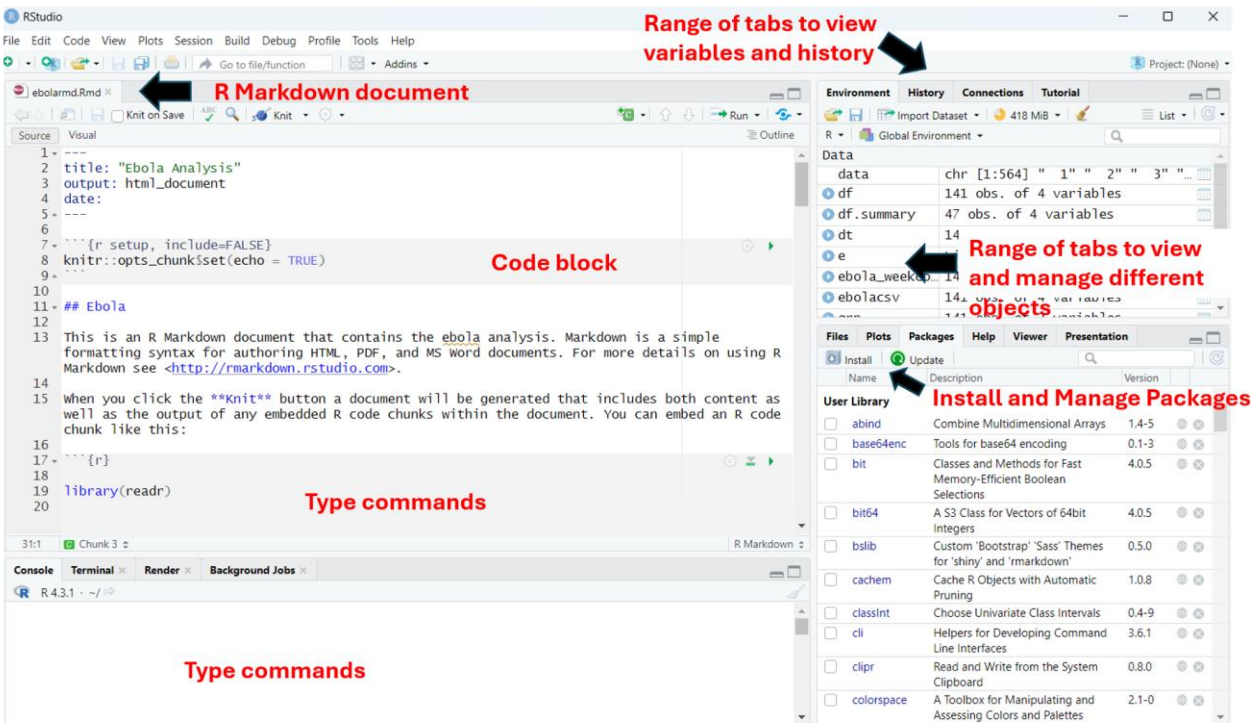
New to R /R Studio

Getting Started in R and RStudio

RStudio/R is a standard statistical analysis package that is free, is extensible and contains many new methods and analysis that are contributed by researchers across the globe.

You'll install R and RStudio at this link: https://www.rstudio.com/products/rstudio/download/ Versions are available for all the major platforms. Once installed, start the program up.

The command window appears as shown in the following figure.



Intro to R.

Video	Description	Total Time	Link
Intro to R Studio	Learn the essentials about R	5:38	https://youtu.be/qL_38YRXJIE

Obtaining and installing R packages

R contains many packages that researchers around the world have created. To learn more about what is available look through the following links

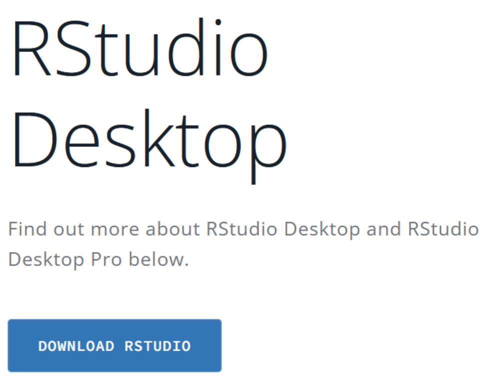
https://cran.r-project.org/web/packages/available_packages_by_name.html

https://www.rstudio.com/products/rpackages/

https://www.datacamp.com/community/tutorials/r-packages-guide

To download Rstudio once you have downloaded R.

https://posit.co/products/open-source/rstudio/



Loading Rmarkdown

Rmarkdown is a package and works in a similar manner to Jupyter Notebook, except it is specific to R. It allows you to save your code and for you to document what you did so that you can return to it without having to retype everything! Essentially a notepad.

Many different packages are available in R. To use them it is necessary to obtain the package and then install it. The packages needed for different analyses will be provided.

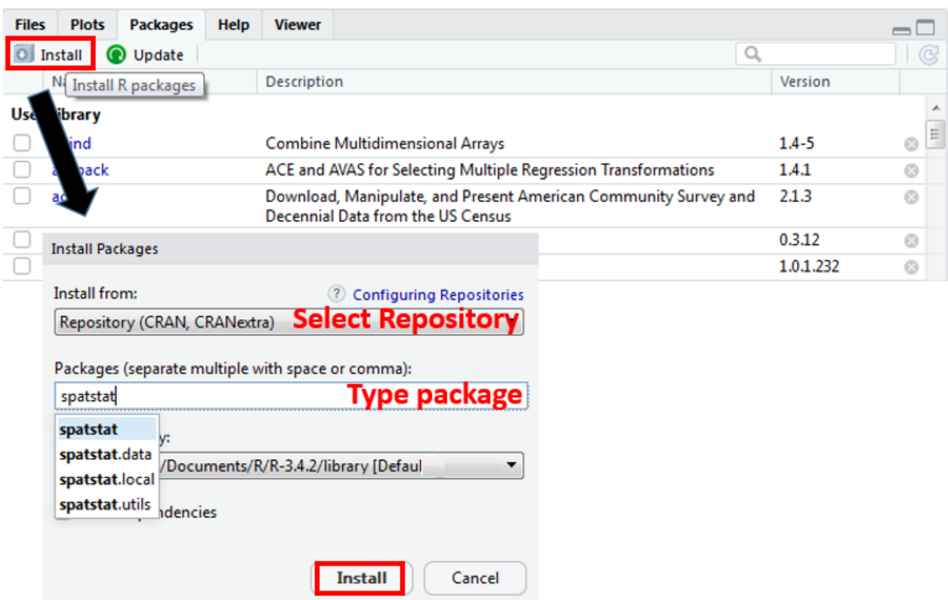
Installing packages by command line

> install.packages("ggplot2")

> library(ggplot2)

OR in RStudio

To do so click the **Packages** tab and click **Install**. A window will popup. Select repository to install from and then type in the package(s) you want. To install the package(s) click Install. Once the package has been installed you will see it listed in the packages library. To turn a package on or off, click on the box next to the package listed in the User Library on the packages tab. Packages are updated, so check for updates regularly.



View of the Packages tab and window showing how to install packages.

Install the following packages:

Notebook and interactive apps

- rmarkdown (book: https://bookdown.org/yihui/rmarkdown/)
- shiny: creating interactive apps https://www.rstudio.com/products/shinyapps/
- Learning tools in R
- swirl: Learning R in R - http://swirlstats.com/
- rcmdr: GUI for different statistical methods.

Various packages used for mapping, spatial analysis, statistics and visualizations.

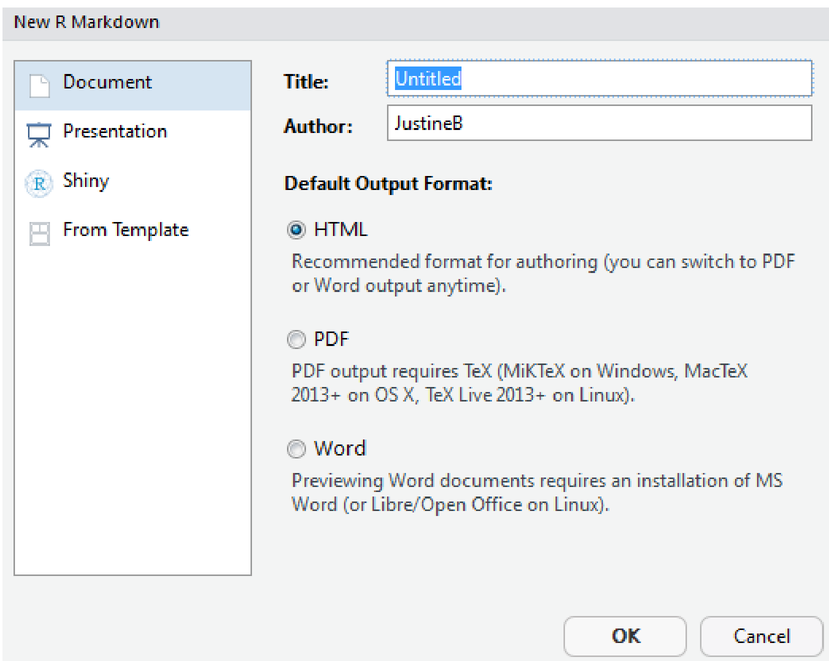
- leaflet: mapping elements - http://rstudio.github.io/leaflet/morefeatures.html
- maptools
- spatstat
- sp
- rgdal
- dplyr
- ggplot2
- doBy
- ggmap
- gridExtra
- RColorBrewer

R Markdown: a quick overview

Since typing commands into the console can get tedious after a while, so taking a notebooking approach is useful.

R Markdown is a package where you can executes R code and display the results inline in your file. Essential R Markdown acts as a notebook where you can integrate your notes with your analysis. When you are finished save the file and if you need to return to your analysis you can load the file into R and run your analysis again. Once you have completed the analysis and write-up create a final document.

This week you will be using a template that has been provided, but for future projects you will want to create a new file. To do so go to **File – New File – R Markdown**. Select the output format and document. Select html.



Window for creating a new R Markdown File.

Opening the RMD file and how to code the file.

Load the rmd file, **File – Open File – select the file.Rmd** you saved on your computer.

The file will open in a window in the top left quadrant of RStudio. This is essentially a text file where you will add your R code, run your analysis and write-up your results.

The template we have provided contains several pieces of information. The header where you should add a title, date, and your name as the author.

R code should be written in the grey areas between the ``` and ``` comments. {r, echo=TRUE) will include the r code into the output. If you do not want to include the r code then set echo=FALSE.

Once you have added what you need you can create a formatted text document, as shown below.

For additional information on formatting see the RStudio cheat sheet https://rmarkdown.rstudio.com/lesson-15.html

To execute the R code and create a formatted document, use **Knit**. **Knit** integrates the pieces in the file to create the output document, in this case the html_document. The **Run** button will execute the R code and integrate the results into the notebook.

A brief introduction to R: the essentials

Essential information about R with essential commands to get you going is provided in a separate pdf. This is not a definitive resource since it would end up as a book in its own right. Instead it provides a variety of commands that will allow you to load your data, explore and visualize your data.