Workshop #1
Water Quality and Summary Statistics

Appalachian A. I. Corps @ UTK

Think about the water you drink. Where is it from? What do you know about it?

Lesson Objective

In this lesson, you will learn how to use the Python programming language to calculate summary statistics and investigate nitrates and nitrates in water samples.

Materials Needed:

- Your computer
- A web browser (Chrome, Firefox, or Safari)
- A calculator

This lesson briefly touches on mean, median, range, and standard deviation. It assumes that students have seen these basic statistical concepts in the past and focuses on recalling them and using them in Python.

Workshop Structure

Workshop Structure

In the workshops, we will use:
- slides
- interactive modules
- handouts

Please have the module pulled up on your device and handout ready.

Navigate to: https://appalachianaicorps.org/ > Modules > Water Quality Monitoring > Lesson 1

Note: If you see a laptop icon in the slides, that means there is a corresponding activity in the module to complete!

Workshop Structure

There are multiple pages in each module. You navigate between them in the left-hand sidebar:

Lesson Contents

1. Overview

2. Intro to Water Quality

3. Nitrates Lab

This section is only included in lesson one to help familiarize students with the layout.We suggest having the link available to students in the LMS or as a tinyURL. Appalachian can be difficult to spell!This is the left-hand sidebar and how students will access different pages in the interactive module. Just click!

4. Intro to Statistics

5. Python: Lists

6. Center & Spread

7. Boxplots Activity

8. Boxplots & Outliers

9. Exit Ticket

Optional Exercise

Workshop Structure

▶ Each page contains **checkpoints**. The checkpoints are linked at the top of the page and include code blocks to run and questions to answer.

▶ They are numbered by the page they are on (number) and the

Let's Get Started!
What is Python?

▶ Python is a **programming language** — a way to give instructions to a computer.
▶ Python can:
  ▶ analyze data
  ▶ do math
  ▶ make graphs

In these workshops, we will learn Python together!

Let's Try Python Right Now!

▶ Python will print text. In programming, we call text **strings**.

Checkpoint 1.a: Print Statements

Click the Run Code button to run the block.

```python
# This is a comment - Python ignores lines that start with #
# Let's make Python print a message!

print("Hello, Water Scientist!")
```

What happens?

Let's Try Python Right Now!

The goal of this slide is to activate prior knowledge. What do students already know about computer science/Python? Importantly, our lessons assume no formal background.This is the one that doesn't render well.

```python
# This is a comment - Python ignores lines that start with #
# Let's make Python print a message!

print("Hello, Water Scientist!")
```

```
Hello, Water Scientist!
```

You just ran Python code! See how it printed "Hello, Water Scientist!" below the block?

Note that strings are always surrounded by quotation marks.

Python Does Math!

▶ Python can do math problems way faster than we can!

Checkpoint 1.b: Mathematical Operators

Click the Run Code button to run the blocks.

Addition:

```python
2 + 3
```

Subtraction:

```python
3 - 2
```

Intro to Water Quality
What is Water?

Question: What do you know about water?

▶ H2O is pure water!
  ▶ But water is almost always mixed with other things:
    ▶ minerals
    ▶ salts
    ▶ other chemicals

▶ Some of these things help water be safe to drink, while others make it unsafe.

Water Quality

Water Quality

▶ In the next several workshops, we will become water quality citizen scientists!

▶ The video discussed several indicators of water quality. We'll focus on **nitrogen** today. It can serve as both a **nutrient** and a **pollutant**.

Dissolved Oxygen	Salinity Levels	pH
Water Temperature	Nutrient Levels	Chlorophyll Concentrations
Bacteria Levels	Concentrations of Pollutants	Pesticides Herbicies Heavy Metals

Nitrogen as Nutrient

▶ Nitrogen is naturally occurring.

▶ When it combines with air and water, it forms ions: nitrates (NO3) and nitrites (NO2).

▶ Nitrogen within water is an important part of the nitrogen cycle, an important process necessary for life.

Nitrogen as Nutrient

Nitrogen as Pollutant

▶ Too much nitrogen in water can be a **pollutant**.

▶ Excess nitrogen–in the form of nitrates and nitrites–can result from:
  ▶ Agricultural operations (fertilizer runoff and livestock manure)
  ▶ Sewage and septic systems (human waste)

Nitrate Lab

We are going to conduct a brief water quality lab to practice gathering data about nitrates and nitrates in water samples.

Materials Needed:

- 4+ water quality samples per group
- 4+ nitrate/nitrate test strips per group
- 1 lab notesheet per person
- 1 computer with the submission form open for data entry
- link to form (provided by your teacher)
- pencils
- 4+ sets of multicolor sticky notes (one set per water sample)
- 1 Sharpie markers
- timing device (watch, clock, etc.)

Nitrate Lab: Roles

Assign roles to each person in your group:

Role 1: *Data Recorder (Lab Report)*
- ▶ 1 lab handout
- ▶ pencil

Role 2: *Data Recorder (Computer)*
- ▶ computer (everyone else can put theirs away temporarily)
- ▶ submission form pulled up in browser

Role 3: *Data Recorder (Sticky Notes)*
- ▶ sets of small sticky notes (4+ sets, one per water sample)
- ▶ Sharpie

Role 4: *Lab Technician*
- ▶ water samples
- ▶ test strips

Nitrate Lab: Procedure

Read the procedure on your lab sheet. Make sure everyone in your group understands the procedure and has the materials needed for their role.

LET'S BEGIN

When finished: With the help of the *Data Recorder (Lab Report)*, make sure everyone in your group has a copy of the data table on their own sheet.

Intro to Statistics

What is Statistics?

▶ How can we make sense of the data we just collected during the nitrates lab?
  ▶ We can use statistics!
▶ Statistics is a **collection of tools** that can be used to **analyze data**. Statistics helps us:
  ▶ Summarize lots of numbers into one useful number
  ▶ Find patterns
  ▶ Make decisions
  ▶ Spot things that don't fit

What is Statistics?

Measures of Center

We use **measures of center** to find the central value of a group of numbers. There are different types of central values. We'll use **mean** and **median** in this lesson.
Sample Dataset: 1, 2, 6, 5, 1

Mean

The *mean* is often called the "average". To find the mean, we add up all of the numbers of interest and divide by how many numbers there are.

$\frac{1+2+6+5+1}{5} = 3$

The mean value of this sample dataset is **3**.

Measures of Center

We use **measures of center** to find the central value of a group of numbers. There are different types of central values. We'll use **mean** and **median** in this lesson.
Sample Dataset: 1, 2, 6, 5, 1

Median

The *median* is a different kind of measure of center. To calculate the median, you first line up all the values in your data set, then you find the middle value by position.

$1 + 1 + 2 + 5 + 6$

Statistics: Your Turn

Materials Needed:

Sample Dataset: 1, 2, 6, 5, 1

Lists

In Python, we store multiple numbers in something called a **list**.

```python
# Our sample dataset from above
# (the square brackets [ ] make it a list)

sample_data = [1, 2, 6, 5, 1]

print("Our sample data:", sample_data)
```

Our sample data: [1, 2, 6, 5, 1]

Breaking it down:

- `sample_data` = the variable name we gave our list (you can name it anything!)
- `=` means "store this in the variable `sample_data`"
- `[1, 2, 6, 5, 1]` = the actual numbers, separated by commas

Python: Make Your Own Lists!

Checkpoint 5.a: Storing data as lists!

Replace the ??, ??, ??, and ?? placeholders with the nitrate and nitrite readings from your group!

Click the Run Code button to run the block.

```python
# Create a list of your group's nitrate and nitrite readings!
# Change the 10, 20, 30, 40 to your group's real values

group_nitrate = [10, 20, 30, 40]
group_nitrite = [10, 20, 30, 40]

print("My group's nitrate readings:", group_nitrate)
print("My group's nitrite readings:", group_nitrite)
```

My group's nitrate readings: [10, 20, 30, 40]
My group's nitrite readings: [10, 20, 30, 40]

Python: Make Your Own Lists!

Checkpoint 5.b: Basic functions with lists.

len()

How many items are in a list?

```python
print("Number of samples tested:", len(group_nitrate))
```

Center & Spread

Python: Measures of Center

▶ We can also use Python to calculate measures of center.
   ▶ We'll import a library of functions called NumPy (Numerical Python) to help us.

Mean

```python
import numpy as np

# Same sample dataset
sample_data = [1, 2, 6, 5, 1]

# Calculate mean the easy way!
mean_sample = np.mean(sample_data)

print("mean of sample data:", mean_sample)
```

```
mean of sample data: 3.0
```

Python: Measures of Center

Mean: Your Turn!

Checkpoint 6.a: Calculating group mean nitrate.
Replace the 10, 20, 30, and 40 placeholders with the nitrate readings from your group!
Click the Run Code button to run the block.

```python
# Recreate the list of your group's nitrate readings below

group_nitrate = [10, 20, 30, 40]

mean_nitrate = np.mean(group_nitrate)

print("Group nitrate readings (mg/L):", group_nitrate)
print("Group mean nitrate:", mean_nitrate, "mg/L")

# Is it safe? (Remember: the EPA's limit is less than 10 mg/L)
if mean_nitrate < 10:
    print("  Average nitrate is SAFE")
else:
    print("  Average nitrate is TOO HIGH")
```

```
Group nitrate readings (mg/L): [10, 20, 30, 40]
Group mean nitrate: 25.0 mg/L
```

Boxplots Activity

Boxplots: Visualizing Data

▶ A key practice in statistics is to visualize the data!

▶ One type of plot we can use to do this is a **boxplot**.

▶ To make a boxplot, we will order the data and then divide the data into four equal groups (by number of observations), called **quartiles**. **Attention** *Data Recorder (Sticky Notes)*

Data Recorders (Sticky Notes), sort all of your sticky notes by color. Pass them out to each group, matching the sticky's color to the group's tent color. This will result in each group having a full set of stickies representing the class nitrate data.

Boxplots: Activity

Checkpoint 7.a: As a group, construct a boxplot for the class nitrate data by hand.

Materials Needed:

- 1 Post-It grid paper sheet per group
- 1 poster marker per group
- Your group's full set of stickies (including those from other groups) from the nitrate lab
- Lab sheet handout (back)

Follow the steps on the back of your lab sheet to create a boxplot for the class nitrate data.

LET'S GO

Boxplots & Outliers

Class Data

Whew! That was a lot of work. Wouldn't it be great if Python could do it for us? Good news. It can! But first, let's re-import our data on this page.

Checkpoint 8.a: Import Class Data.

Once more, replace the placeholder (Line 4) with CSV URL from your teacher. Be sure the keep the quotation marks! This will pull the class nitrate and nitrite data from the CSV file so you can use it later down the page.

Click the Run Code button to run the block.

```python
# Replace the url with the one provided from your teacher.
# Make sure to keep the quotation marks!

csv_url = "replace_this_with_your_csv_url"

class_nitrate, class_nitrite = load_class_data(csv_url)
```

```python
print("Nitrate values:", class_nitrate)
print("Nitrite values:", class_nitrite)
```

```
Nitrate values: [0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5
Nitrite values: [0.0, 0.0, 0.1, 0.1, 0.15, 0.2, 0.2, 0.25, 0.3,
```

Class Data: Boxplots

Checkpoint 8.b: Create a boxplot for the class nitrate data.

Click the Run Code button to run the block and create a boxplot for the class nitrate data.

```python
import matplotlib.pyplot as plt

plt.figure(figsize=(6, 6))
plt.boxplot(class_nitrate, vert=True, patch_artist=True,
            boxprops=dict(facecolor='plum'))
plt.ylabel('Nitrate (mg/L)', fontsize=12)
plt.xticks([])
plt.title('Boxplot of Class Nitrate', fontsize=13)
plt.grid(True, alpha=0.3, axis='y')
plt.show()
```

Class Data: Boxplots

Checkpoint 8.b: Create a boxplot for the class nitrate data.

Click the Run Code button to run the block and create a boxplot

Exit Ticket

Exit Ticket

Great job! You've learned so much!Share what you've learned on the **Exit Ticket**.

Exercises

Exercises

Want to practice what we've learned?Try the **Exercises**.