**Lesson 4:** More on data: Python IDEs, Pandas, Numpy, and visualization in Python

**Estimated Time**: 8 hours

**Guiding Question:** What reproducibility tools are available in Python? How can we analyze data using Pandas, Numpy, and visualization packages?

**Concepts:** Jupyter Notebook, Google Colab, Spyder, VSCode, Virtual environments, requirements.txt, Pandas, Numpy, Matplotlib, Seaborn, Plotly

**Lesson Description**: Students are introduced to various Python IDEs to promote reproducibility of their work, to basic data analysis tools in Pandas, Numpy, Matplotlib, and Seaborn.

| **Materials and Resources** | **Learning Goals** |
| --- | --- |
| 1. Lesson slides 2. Projector and computer for teacher (if in-person) | * Learn various Python IDEs and their applicability for different tasks. * Learn basic data analysis utilizing Pandas, Numpy, and visualization tools. |

| **Time** | **Lesson Content** | **Instructor Notes** |
| --- | --- | --- |
| 9:00-11:00  10 min break | Python IDEs and reproducibility tools:   1. Jupyter Notebook 2. Google Colab 3. Spyder 4. VSCode 5. Virtual environments 6. requirements.txt | Ensure everyone gets the programs installed.    Suggest to uninstall the unnecessary files at the end of the class. |
| 11:10-12:10 | Numpy |  |
| 12:10-13:10 | LUNCH |  |
| 13:10-15:10  10 min break | Pandas |  |
| 15:20-16:20 | Visualization tools:   1. Matplotlib 2. Seaborn 3. Plotly |  |
| 40 min | **Assessment**  Assignment 4   1. Write code for creating and populating a database. 2. Run three queries using WHERE, GROUP BY, ORDER BY clauses | Allow 20 min for individual work.    Attend students' questions. |