LESSON PLAN

LESSON #3

Aim: Filtering data and handling missing data part 1

Objective: After a lesson on reading and selection of data, students will be able to use Python to perform the same procedures on two other datasets.

Do Now: What is a data science? https://www.youtube.com/watch?v=X3paOmcrTjQ

Standards: 9-12. CT.7 Computational Thinking, Algorithms, and Programming

9-12. DL.5 Digital Literacy, Digital Use

Mini-Lesson:

1. Filtering data
   a. Filtering by using Boolean indexing
   b. Code example: filtering values less than or equal to 6.5
2. Handling missing data
   a. Ways of handling missing data
   b. Code example: filtering missing data
   c. Subtle feature of NaN and the use of isnull( )
3. Conclusion and summary
   Discussion:
   - What are some of the problems or challenges you encountered?
   - How did you resolve them?
   - What did you learn from this lesson?
   - Do you have any lingering questions on today's lesson or data science in general?

CODE

Filtering by using Boolean indexing

```
edu[edu['Value'] > 6.5].tail()
```

Output

	TIME	GEO	Value
218	2002	Cyprus	6.60
281	2005	Malta	6.58
94	2010	Belgium	6.58
93	2009	Belgium	6.57
95	2011	Belgium	6.55

Handing of missing data

```
edu[edu["Value"].isnull()].head()
```

	TIME	GEO	Value
0	2000	European Union (28 countries)	NaN
1	2001	European Union (28 countries)	NaN
36	2000	Euro area (18 countries)	NaN
37	2001	Euro area (18 countries)	NaN
48	2000	Euro area (17 countries)	NaN