

Data Insight

Knowledge for Insight from Data

DATA SCIENTIST PROGRAM

Assignment - Pandas Techniques for Data Manipulation in Python

Real-world data is messy. That's why libraries like pandas are so valuable. Using pandas you can take the pain out of data manipulation by extracting, filtering, and transforming data in DataFrames, clearing a path for quick and reliable data.

In this assignment, you will give a tutorial on some useful pandas techniques that, in your opinion, are very important for dealing with data using python.

Explain in detail at least five pandas techniques. Give coding examples to illustrate your point. Your code should be written in a Jupyter notebook. Choose no more than two techniques from the list below and the others (not from the list) your choice.

1. Apply Function
2. Pivot Table
3. Crosstab
4. Sorting DataFrames
5. Iterating over rows of DataFrame
6. Boolean Indexing
7. Imputing missing values
8. Multi-Indexing
9. Merging DataFrames
10. Plotting

Upload your notebooks to your GitHub repository.

Using each notebook, write a blog post explaining your chosen pandas techniques. The blog post should be written [here](#) with code from your notebook embedded (use the code snippet option instead of taking screenshots of your code).

Submit the link to your notebook and blog post [here](#). Submission is due on or before Nov. 21, 2021.