Week 2 Mini-Project

← Go back

Week 2 Mini-Project

The goal of this project is to combine everything you have learned about data wrangling, cleaning, and manipulation with Pandas so you can see how it all works together. For this project, you will start with this messy data set [Shark Attack]. You will need to import it, use your data wrangling skills to clean it up, prepare it to be analyzed, and then export it as a clean CSV data file.

The dataset that we provide you is a significantly messy data set. Apply the different cleaning and manipulation techniques you have learned.

Examine the data for potential issues.

Produce a well documented Jupyter Notebook that shows the steps you took and the code you used to clean and transform your data set. Don't forget that all the data eliminated must be clearly justified.

Preferably use functions and classes to help with you cleaning flow.

Export a clean CSV version of your data.

Once you have a clean dataset think of one question your data can answer and show the answer with a analysis.

Necessary Deliverables

*A cleaned CSV data file** containing the results of your data wrangling work.

*A Jupyter Notebook ** containing all Python code and commands used in the importing, cleaning, manipulation, exporting and analysis of your data set.

*A documentation containing a detailed explanation of the process followed in the importing, cleaning, manipulation, exporting and analysis of your data as well as your

Week 2 Mini-Project

*Examine the data and try to understand what the fields mean before diving into data cleaning and manipulation methods.

*Break the project down into different steps - use the topics covered in the lessons to form a check list, add anything else you can think of that may be wrong with your data set, and then work through the check list.

*Use the tools in your tool kit - your knowledge of Python, data structures, Pandas, and data wrangling.

*Work through the lessons in class & ask questions when you need to!

*Commit early, commit often, don't be afraid of doing something incorrectly because you can always roll back to a previous version.

* Consult documentation and resources provided to better understand the tools you are using and how to accomplish what you want.

PREVIOUS LESSON

[CAREER SERVICES]
Produtividade

NEXT LESSON

Weekly Survey

←