

Pandas AI

The Generative AI Python Library

What is Pandas AI?

- Pandas AI is a Python library that uses generative AI models to supercharge pandas capabilities.
- It was created to complement the pandas library, a widely-used tool for data analysis and manipulation.
- Users can summarize pandas data frames data by using natural language. Moreover, you can use it to plot complex visualization, manipulate dataframes, and generate business insights.
- Pandas AI is beginner-friendly; even a person with little technical background can use it to perform complex data analytics tasks. Its utility helps analyze data faster and derive meaningful conclusions.

Getting Started with Pandas AI

In this section, we will learn how to install and set up Pandas AI for data analysis.

1. First, we will install Pandas AI using pip.
 - a. `pip install pandasai`
2. Second, we need to obtain an OpenAI API key and store it as an environment variable
 - a. <https://platform.openai.com/account/api-keys>

Demo

Open the code in your Google Collab

https://github.com/naashonomics/openai/blob/main/PandasAI_demo.ipynb

Pros and Cons of Pandas AI

Pros of Pandas AI

- Can easily perform simple tasks without having to remember any complex syntax
- Capable of giving conversational replies
- Easy report generation for quick analysis or data manipulation

Cons of Pandas AI

- Cannot perform complex tasks
- Cannot create or interact with variables other than the passed dataframe

FAQ

- Can PandasAI work without OpenAI's API?
 - Yes, other than ChatGPT you can also use Google's PaLM model, Open Assistant LLM and StarCoder LLM for code generation.
- Can PandasAI use numpy attributes or functions?
 - No, it does not have the ability to use numpy functions. All computations are performed either by using Pandas or in-built python functions in the backend.