

# pandas II


This function is one of Plugins Operation. You can find the movie in [ARGOS RPA+ video tutorial](#).

Need help?

**Technical** contact to [tech@argos-lab.com](mailto:tech@argos-lab.com)

May you search all operations,

- [Actions](#)
- [Verifications](#)
- [System Calls](#)
- [Interactives](#)



## pandas II

**Author: Jerry Chae**

This is the second release in our pandas plugin series. It will give you more flexibility and more agility as to the pre-process sequence of your data. Data-Scientists can try their cleansing and filtering in tools like Jupyter Notebook and apply it the flow immediately to this plugin for super quick and easy distribution and deployment in the field.


### Primary Features

This plugin runs python statement(s) on pandas with one (1) dataframe.

- [https://en.wikipedia.org/wiki/Pandas\\_\(software\)](https://en.wikipedia.org/wiki/Pandas_(software))
- [https://pandas.pydata.org/pandas-docs/stable/getting\\_started/10min.html](https://pandas.pydata.org/pandas-docs/stable/getting_started/10min.html)

### Prerequisite

**This plugin requires Python and Regular Expression skills.**

 **Please note** that the pandas solution is a large software using numerous Python machine learning sub-modules. The bot will take more than just a few minutes to download them to be ready. But this is just for the "first run". As to the second run on, the local VENV will be used to avoid downloading unless new pandas II version has been selected to replace what was in the bot originally.

And, **YES** --- if no pre-processing is performed, this plugin can act as **a file format converter**. Again, you will face a long first run download time because of the reason described above.

## Input, Output, Features, and Parameters.

### Required Input

1. Input File: One data file (dataframe).  
Supported input formats are **.xlsm, .xls, xlsx, .csv, .tsv, and .json**
2. Output File: One data file.  
Supported input formats are **.xlsm, .xls, xlsx, .csv, .tsv, and .json**

### Optional Input

3. Enter a **Python statement**, or multiple statements. Also a text file that contains a list of statements can be used as input.
4. When input file multiple **sheets**, you can select which sheet to be processed.

5. You can designate which row you can use as **header (variable)** for your processing.
6. You can specify a column to be used as the **index** of the dataframe.
7. You can specify which column(s) to be or not to be processed.
8. You can determine specific pandas **datatypes** for your column.
9. You can determine what character to use to **separate your data** (default is comma).
10. You can specify **encoding** technology of the input file (default is UTF-8).
11. You can select to either **show or hide the index** column in your output file.

## How to set parameters

pandas-II plugin parameters are 100% compatible to pandas read\_excel specifications

[https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.read\\_excel.html](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.read_excel.html)

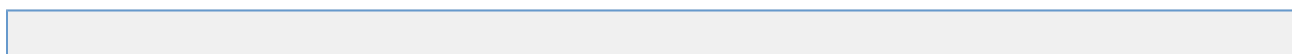
Please refer the parameters on the right in the pandas document above.

- Sheet Name      sheet\_name
- Header Row     header
- Index Col       index\_cols
- Use Col         usecols
- Data Type       dtypes

The screenshot shows the configuration interface for the 'pandas II' plugin. It includes fields for 'In file' and 'Out file', an 'Advanced' section with checkboxes for 'Statements', 'Statements File', 'Sheet Name', 'Header Row', 'Index Col', 'Use Cols', 'Data Type', 'CSV Sep', 'Encoding', and 'Show Index'. A 'Return value' section is at the bottom. Yellow callout boxes provide the following explanations:

- Input file in .xls .xlsx .xlsm .csv .tsv and .json.** (points to the 'In file' field)
- Output file (no need to exist) in .xls .xlsx .xlsm .csv .tsv and .json.** (points to the 'Out file' field)
- One Python statement or multiple python statements.** (points to the 'Statements' checkbox)
- You can use a text file that contains a series of Python statements as well.** (points to the 'Statements File' field)
- Header row default is 0 meaning the top row. You can change it here..** (points to the 'Header Row' field)
- If index column needs to be specified, use 0 base index here.** (points to the 'Index Col' field)
- Refer to pandas data type document.** (points to the 'Data Type' field)
- Data separation character can be changed here.** (points to the 'CSV Sep' field)
- You can decide either to show or hide the index column.** (points to the 'Show Index' checkbox)

Text from Image



## All Plugins

- AD LDAP
- Adv Send Email
- Arithmetic Op
- Attach Image
- AWS Textra Rekog
- Bot Collabo
- Chatwork Notification
- Clipboard
- Convert CharSet
- Create Newfile
- Detect CharSet
- Drag and Drop
- Email IMAP ReadMon
- Email Read Mon
- Env Check
- Env Var
- Excel Advanced
- Excel AdvII
- Excel AdvIII
- Excel Copy Paste
- Excel Formula
- Excel Macro
- Excel Newfile
- Fairy Devices mimi AI
- File Conv
- File Folder Op
- File Monitor
- Folder Structure
- Google Calendar
- Google Cloud Vision API
- Google Drive
- Google Sheets
- Google Token
- Google Translate
- Google TTS
- Html Extract
- IBM Speech to Text
- IBM Visual Recognition
- JSON Select
- LINE Notify
- MS Azure Text Analytics
- MS Word Extract
- NAVER OCR
- Newuser-SFDC
- PANDAS I
- pandas II
- pandas III
- PANDAS profiling
- Password Generate
- PDF2Doc
- PowerShell
- Regression
- REST API
- Rossum
- Scrapy Basic
- Simple SFDC
- Slack
- Speed Test
- SQL
- SSH Command
- SSH Copy
- String Manipulation
- Telegram
- Tesseract
- Time Stamp
- Web Extract
- Work Calendar
- XML Extract
- ZipUnzip

