



WPS Python procedure user guide and reference

*Interacting with Python from SAS
language programs*



Version: 4.0.3

Copyright © 2002-2018 World Programming Limited

www.worldprogramming.com



Contents

Introduction.....	3
Setup and configuration.....	4
Using Python with WPS.....	5
PROC PYTHON.....	6
EXPORT.....	6
IMPORT.....	7
SUBMIT.....	7
ENDSUBMIT.....	8
EXECUTE.....	8
Legal Notices.....	9

Introduction

The Python procedure enables SAS language programs to include code written in the Python language.

By combining Python and the SAS language you can:

- Use the SAS language to perform the bulk processing and preparation of data, and pass the processed data to Python.
- Use Python packages you have previously developed for data analysis.
- Use Python data analysis packages or solutions that may not be available in the SAS language.

Data is passed between WPS and Python using the `EXPORT` statement. Once data has been transferred, that data is made available as a pandas DataFrame to a Python program. On completion of the Python program data can, if required, be returned to the SAS language environment using the `IMPORT` statement.

Setup and configuration

The Python procedure is available with WPS version 3.3 and later. This section describes the setup of the Python environment for WPS.

When using Python with WPS:

- The 32-bit Python interpreter is required for the 32-bit version of WPS, and the 64-bit Python interpreter is required for the 64-bit version of WPS.
- The `pandas` package must be installed with the Python interpreter. This can be checked using the `pip` utility by running `pip freeze` on the command line.

The WPS distribution does not include either the Python interpreter or `pandas` package. If you do not have Python installed, you can obtain a copy of the Python interpreter from <https://www.python.org> or install a packaged Python environment that includes the necessary modules.

Supported platforms

The `PYTHON` procedure is currently supported in WPS on Microsoft Windows, Linux-based systems, and macOS.

The procedure is not currently supported on IBM mainframe.

Supported Python versions

For the WPS 3.3 release, the only Python version 3.5.x is supported.

Setting `PYTHONHOME`

You must set the `PYTHONHOME` environment variable for WPS to locate and use Python. This variable must reference the folder where the main Python library – for example, `python35.dll` on Microsoft Windows – is located.

Standard output and error streams

The Python standard output stream (`sys.stdout`) is redirected into WPS when the `PYTHON` procedure is run. When using WPS on the command line, the stream is directed to the WPS listing file; when WPS Workbench is used, the stream is directed to all selected output formats.

The Python standard error stream (`sys.stderr`) is directed to the WPS log file.

Using Python with WPS

Using Python in a SAS language program enables you to make use of specialist Python packages such as `Scikit-learn` or `pandas`.

The first time the Python procedure is invoked in a SAS language program, WPS automatically imports the `pandas` and `numpy` packages. You can access the functionality in the `pandas` and `numpy` packages in an in-line Python program – written between the `SUBMIT` and `ENDSUBMIT` statements – by referencing the fully-qualified package, class or function name, for example:

```
PROC PYTHON;
  SUBMIT;
content = pandas.read_csv('file:///C:/project/sourcedata/wps.csv')
...
  ENDSUBMIT;
RUN;
```

Other Python packages can be imported and used within the in-line Python code, for example:

```
PROC PYTHON;
  EXPORT DATA=source;
  SUBMIT;
import statsmodels.formula.api as lm
result = lm.ols(formula='x ~ y + z', data=source).fit()
...
  ENDSUBMIT;
RUN;
```

Each subsequent use of the Python procedure in a SAS language program uses the same Python environment. This means any global variables or imported packages used in a Python procedure invocation are available to all subsequent Python procedure invocations.

PROC PYTHON

The `PYTHON` procedure invokes the Python environment within which you execute Python programs.

```
» PROC PYTHON ; «
```

When the Python environment is invoked, the `pandas` and `numpy` packages are automatically loaded.

Datasets created in WPS can be made available to the Python program using the `EXPORT` statement, and a dataset imported from the Python program into WPS using the `IMPORT` statement.

A Python program can be either written in-line in the `PYTHON` procedure, or run from a separate file:

- To run an in-line Python program, use the `SUBMIT` and `ENDSUBMIT` statements.
- To run a Python program stored in an external file use the `EXECUTE` statement.

The Python environment is exited using a `RUN` statement.

Once the Python environment has been invoked, the same environment is used for any subsequent `PYTHON` procedure calls in the same SAS language program.

Supported statements

- [EXPORT](#) (page 6)
- [IMPORT](#) (page 7)
- [SUBMIT](#) (page 7)
- [ENDSUBMIT](#) (page 8)
- [EXECUTE](#) (page 8)

EXPORT

The `EXPORT` statement creates a pandas DataFrame from a WPS dataset.

```
» EXPORT export-options ; «
```

Export options

DATA

```
» DATA = dataset «
```

The name of the dataset to be made available to the Python program as a pandas DataFrame. The dataset name is also used as a global variable to identify the DataFrame within the Python program.

IMPORT

The `IMPORT` statement creates a WPS dataset from a pandas DataFrame.

```
IMPORT import-options ;
```

Import options

DATA

```
DATA = dataset
```

The name of the dataset to be created from the pandas DataFrame.

PYTHON

```
PYTHON = dataframe-name
```

The Python global variable name that identifies the pandas DataFrame to be returned to WPS.

SUBMIT

The `SUBMIT` statement identifies the statements following it, until an `ENDSUBMIT` statement is encountered as Python program code.

```
SUBMIT ;
```

Python program code must start on a new line. The first line of the Python program code must not start with any white space and any subsequent Python statements must follow the normal Python requirements for indentation, for example:

```
PROC PYTHON;
  SUBMIT;
  fruits = ['apple', 'banana', 'cherry', 'damson', 'elderberry', 'fig']
  for fruit in fruits:
    print(fruit)
  ENDSUBMIT;
RUN;
```

ENDSUBMIT

The `ENDSUBMIT` statement marks the end of in-line Python program code.

```
»» ENDSUBMIT ; ««
```

EXECUTE

The `EXECUTE` statement is used to run a Python program stored in a separate file.

```
»» EXECUTE filename ; ««
```

filename

The absolute path of the file containing the Python program. The path must be in the correct form for the operating system on which WPS and Python are run.

Example

Running an external file in the `PYTHON` procedure:

```
PROC PYTHON;  
    EXECUTE C:\temp\test.py;  
RUN;
```


Legal Notices

Copyright © 2002–2018 World Programming Limited.

All rights reserved. This information is confidential and subject to copyright. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system.

Trademarks

WPS and World Programming are registered trademarks or trademarks of World Programming Limited in the European Union and other countries. (r) or ® indicates a Community trademark.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

All other trademarks are the property of their respective owner.

General Notices

World Programming Limited is not associated in any way with the SAS Institute.

WPS is not the SAS System.

The phrases "SAS", "SAS language", and "language of SAS" used in this document are used to refer to the computer programming language often referred to in any of these ways.

The phrases "program", "SAS program", and "SAS language program" used in this document are used to refer to programs written in the SAS language. These may also be referred to as "scripts", "SAS scripts", or "SAS language scripts".

The phrases "IML", "IML language", "IML syntax", "Interactive Matrix Language", and "language of IML" used in this document are used to refer to the computer programming language often referred to in any of these ways.

WPS includes software developed by third parties. More information can be found in the THANKS or acknowledgments.txt file included in the WPS installation.