

[← BACK TO BLOG](#)

# Leverage the power of Python in Tableau with TabPy

[🔗 SHARE](#)

BORA BERAN

PRODUCT MANAGER, TABLEAU SOFTWARE

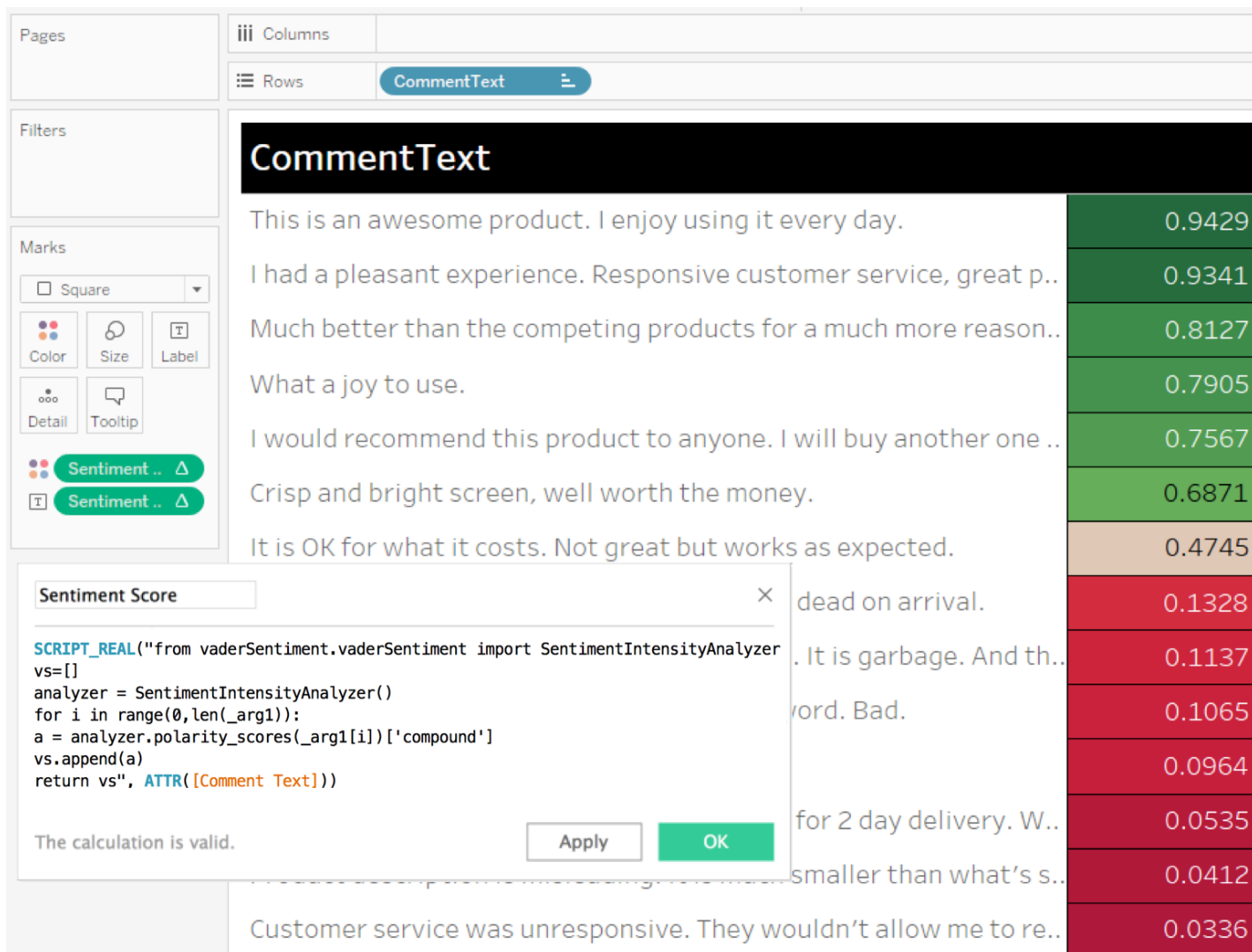
NOVEMBER 4, 2016

We're happy to announce the beta release of **TabPy**, a new API that enables evaluation of Python code from within a Tableau workbook.

When you use TabPy with Tableau, you can define calculated fields in Python, thereby leveraging the power of a large number of machine-learning libraries right from your visualizations.

This new Python integration in Tableau enables powerful scenarios. For example, it takes only a few lines of Python code to get the sentiment scores for reviews of products sold at an online retailer. Then you can explore the results in many ways in Tableau.

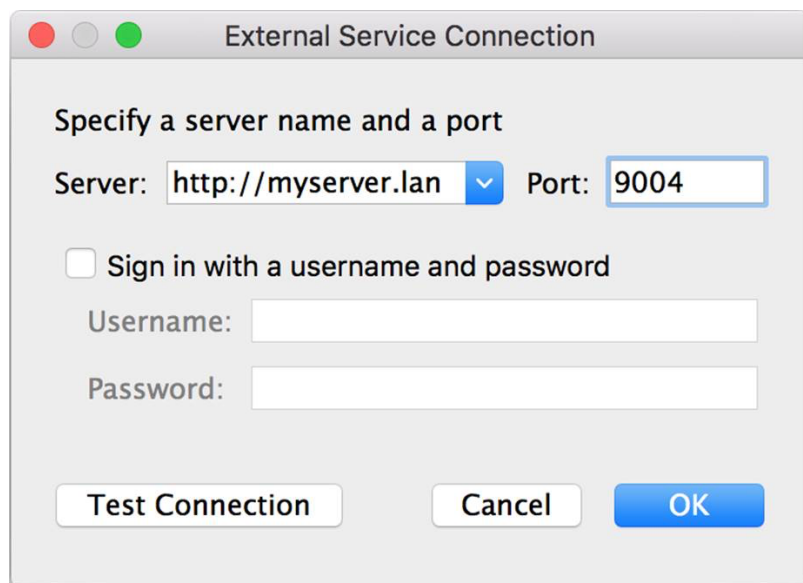
You might filter to see just the negative reviews and review their content to understand the reasons behind them. You might get a list of customers to reach out to. Or you might visualize overall sentiment changes over time.



Other common business scenarios include:

- **Lead scoring:** Create a more efficient conversion funnel by scoring your users' behavior with a predictive model.
- **Churn prediction:** Learn when and why users leave, and predict and prevent it from happening.

You can easily install the TabPy server on your computer or on a remote server. Configure Tableau to connect to this service by entering the service URL and port number under Help > Settings and Performance > Manage External Service Connection in Tableau Desktop. Then you can use Python scripts as part of your calculated fields in Tableau, just as you've been able to do with R since Tableau 8.1.



External Service Connection

Specify a server name and a port

Server:  Port:

☐ Sign in with a username and password

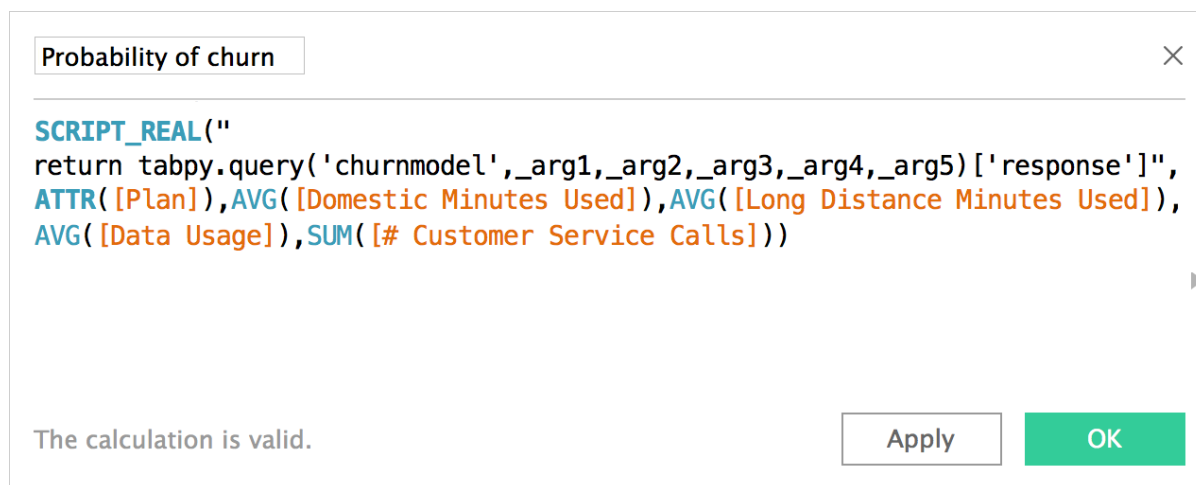
Username:

Password:

TabPy uses the popular Anaconda environment, which comes preinstalled and ready to use with many common Python packages including scipy, numpy, and scikit-learn. But you can install and use any Python library in your scripts.

If you have a team of data scientists developing custom models in your company, TabPy can also facilitate sharing those models with others who want to leverage them inside Tableau via published model.

Once published, all it takes to run a machine-learning model is a single line of Python code in Tableau regardless of model type or complexity. You can estimate the probability of customer churn using logistic regression, multi-layer perceptron neural network, or gradient boosted trees just as easily by simply passing new data to the model.



Probability of churn

```
SCRIPT_REAL("  
return tabpy.query('churnmodel',_arg1,_arg2,_arg3,_arg4,_arg5)['response']",  
ATTR([Plan]),AVG([Domestic Minutes Used]),AVG([Long Distance Minutes Used]),  
AVG([Data Usage]),SUM([# Customer Service Calls]))
```

The calculation is valid.

Using published models has several benefits. Complex functions become easier to maintain, share, and reuse as deployed methods in the predictive-service environment. You can improve and update the model and code behind the endpoint while the calculated field keeps working without any change. And a dashboard author does not need to know or worry about the complexities of the model behind this endpoint.

Together, Tableau and Python enable many more advanced-analytics scenarios, making your dashboards even more impactful. To learn more about TabPy and download a copy, please visit our [GitHub page](#). Give TabPy a try, and let us know what you think at [beta@tableau.com](mailto:beta@tableau.com).

If you're attending the [Tableau Conference](#), stop by our [session on TabPy](#) to learn more and see some cool demos.

## Do more with Tableau and Big Data

Find out how Tableau solves many of the problems Big Data can present to organizations of any size.

[Learn more.](#)

You might also be interested in...

## Comments

Submitted by Wilbur (not verified) on November 4, 2016 - 10:44am

This is really amazing. This opens up potential opportunities for Tableau to be used in semi-real-time applications. This can be used to call microservices or to trigger actions.

[Reply](#)

Submitted by Zach Leber on November 5, 2016 – 5:46pm

URL actions are another way of calling services to trigger actions.

[https://onlinehelp.tableau.com/current/pro/desktop/en-us/actions\\_url.html](https://onlinehelp.tableau.com/current/pro/desktop/en-us/actions_url.html)

We use Tableau URL actions from tooltips to send REST messages which control and change data.

[Reply](#)

Submitted by Parikshit Shinge on November 3, 2017 – 4:27am

Hi Wilbur,

Can you please let me know how can I call a service using Python in Tableau?

Regards,

Parikshit

[Reply](#)

Submitted by Prayson Daniel (not verified) on November 4, 2016 – 11:46am

Gospel for us Pythonic lovers! Thank you Tableau! This is amazing. I cannot wait to explore this feature.

[Reply](#)

Submitted by Melissa Torres (not verified) on November 4, 2016 – 2:09pm

Well... Just put my hair in a bun, started a fresh pot of coffee, am rolling up my sleeves, and checking out of reality and into TabPy.

[Reply](#)

Submitted by Chris Schultz on November 4, 2016 – 8:17pm

Not crazy about having to install a server to use this feature....how is performance of Python calculations when using large datasets compared to R...which seemed very slow when I tried it.

## Reply

Submitted by Wilbur (not verified) on November 4, 2016 – 10:58pm

Chris S.: not sure what you mean. Perhaps you can provide a specific example? Pure Python is generally comparable or faster than pure R on most benchmarks. Of course, in real-life use cases, it boils down to the specific libraries you use. Like R (with Rcpp), most of the data libraries in Python (Pandas, Numpy, etc.) call underlying code written in a lower-level language like C/Fortran (e.g. BLAS/Lapack) so the performance difference should be non-issue.

## Reply

Submitted by Alex Blakemore on November 5, 2016 – 11:56am

Awesome news.

Question though -- How do the Script\_XXX() functions determine whether it is a R script or a Python script? Does Tableau infer that the script somehow? Do you have to configure a workbook to either use R or Python, but not both? For example if the script is just the number 1, which is valid in Python and R, where does Tableau send that script to execute?

## Reply

Submitted by Bora Beran on November 5, 2016 – 7:30pm

Hi Alex,

Currently you can have only one active external service at a time (one per each running Tableau.exe) on desktop and one setting per server.

~ Bora

## Reply

Submitted by Ashish Singh (not verified) on November 5, 2016 – 10:25pm

Getting error: package not found via pip and conda

## Reply

Submitted by Bora Beran on November 6, 2016 – 12:44am

Hi Ashish,

Did you download the zip file from Github and follow the install instructions or did you try 'pip install tabpy'? There is no such package in PyPI at this point so the latter will fail. If you don't want to go through full install, there is install instructions for users who already have anaconda configured also in Github. The steps provided pip using the local package contained in the download from Github repository.

~ Bora

## Reply

Submitted by Tharashasank (not verified) on November 7, 2016 – 12:51am

This is amazing.

## Reply

Submitted by Ashish Singh (not verified) on November 7, 2016 – 5:17am

Hi Bora,

I was earlier using pip install TabPy / conda install TabPy

Gone through the Github, seems like I'll be able to install and have it running once I am home in evening. Thanks a ton for clarifying.

~ Ashish

## Reply

Submitted by Distributed Ene... (not verified) on November 7, 2016 – 11:11pm

Great to know the new feature of TabPy in Python code it make me more easy to work thank you.

## Reply

Submitted by Jenny (not verified) on November 7, 2016 - 11:15pm

Thank for the new Python Code,Extended Team Model

[Reply](#)

Submitted by Andrew Thacker on November 8, 2016 - 12:49am

Great !!!

[Reply](#)

Submitted by Kenneth Black (not verified) on November 10, 2016 - 10:06am

It was more than three years ago when I first wrote about my desire to have a feature like this in Tableau. At the time, R integration was a proposed and was an upcoming idea. With each passing year, I have to keep updating this article because Tableau keeps delivering new integrated technologies. It's becoming a full-time job trying to keep up with Tableau:

<https://3danim8.wordpress.com/2013/10/08/i-wish-tableau-had-these-two-capabilities/>

[Reply](#)

Submitted by Ashish Singh (not verified) on November 16, 2016 - 8:39pm

Hi Bora,

I am getting this error while trying to run setup.bat

Could not open requirements file: [Errno 2] No such file or directory: './tabpy-server/requirements.txt'

Invalid requirement: './tabpy-client'

Traceback (most recent call last):

File "c:\python27\lib\site-packages\pip\req\req\_install.py", line 82, in \_\_init\_\_  
req = Requirement(req)

File "c:\python27\lib\site-packages\pip\\_vendor\packaging\requirements.py", line 96,



```
in __init__
requirement_string[e.loc:e.loc + 8]))
InvalidRequirement: Invalid requirement, parse error at "'./tabpy-'"

Invalid requirement: './tabpy-server'
Traceback (most recent call last):
File "c:\python27\lib\site-packages\pip\req\req_install.py", line 82, in __init__
req = Requirement(req)
File "c:\python27\lib\site-packages\pip\_vendor\packaging\requirements.py", line 96,
in __init__
requirement_string[e.loc:e.loc + 8]))
InvalidRequirement: Invalid requirement, parse error at "'./tabpy-'"
```

Any guess why. Will appreciate any help,

Ashish

[Reply](#)

Submitted by Bora Beran (not verified) on November 21, 2016 – 5:54pm

Hi Ashish,

Did you download the zip using the green link in the upper left corner, unzip it and ran setup from inside tabpy folder that contains tabpy-server folder?

Thanks,

Bora

[Reply](#)

Submitted by Nathan Graham (not verified) on November 21, 2016 – 10:21am

I'm receiving this error in Tableau when trying to use the sentiment script shown above. Anyone have an idea what's causing this? Thanks

```
Error in base::parse(text = .cmd) : :1:6: unexpected symbol 1: from
vaderSentiment.vaderSentiment^
```

## Reply

Submitted by Bora Beran on November 21, 2016 - 6:12pm

Looking at the error, issue seems like that the script is being sent to R, instead of Python. Can you check if Manage External Services dialog is pointing to TabPy server? Also note that this will work with Tableau 10.1 or higher.

## Reply

Submitted by Bora Beran on November 21, 2016 - 6:12pm

Looking at the error, issue seems like that the script is being sent to R, instead of Python. Can you check if Manage External Services dialog is pointing to TabPy server? Also note that this will work with Tableau 10.1 or higher.

## Reply

Submitted by Nathan G (not verified) on November 22, 2016 - 10:24am

I was able to connect to the TabPy server, but now I'm receiving this error when running sentiment score script shown above.

Error processing script

Error when POST /evaluate: Traceback

Traceback (most recent call last):

File "C:\Users\ngrah008\Anaconda\envs\Tableau-Python-Server\Lib\site-packages\tabpy\_server\tabpy.py", line 467, in post

result = yield self.call\_subprocess(function\_to\_evaluate, arguments)

File "C:\Users\ngrah008\Anaconda\envs\Tableau-Python-Server\lib\site-packages\tornado\gen.py", line 1008, in run

value = future.result()

File "C:\Users\ngrah008\Anaconda\envs\Tableau-Python-Server\lib\site-packages\tornado\concurrent.py", line 232, in result

raise\_exc\_info(self.\_exc\_info)

File "C:\Users\ngrah008\Anaconda\envs\Tableau-Python-Server\lib\site-

```
packages\tornado\gen.py", line 282, in wrapper
```

```
yielded = next(result)
```

```
File "C:\Users\ngrah008\Anaconda\envs\Tableau-Python-Server\Lib\site-packages\tabpy_server\tabpy.py", line 482, in call_subprocess
```

```
exec(function__to__evaluate)
```

```
File "", line 2
```

```
from vaderSentiment.vaderSentiment import sentiment as vs=[]
```

```
^
```

SyntaxError: invalid syntax

Error type : SyntaxError

Error message : invalid syntax (, line 2)

[Reply](#)

Submitted by Dave Savell on December 17, 2016 - 10:37pm

I get the same error as Nathan G when trying to emulate the first sample shown (vaderSentiment). Any ideas - did you solve your issue Nathan?

[Reply](#)

Submitted by Bora Beran on December 17, 2016 - 11:12pm

Looks like the Python library used in the example got an update that has some breaking changes. I reinstalled from scratch and observed the same error. So please try the following instead. We will update the image in the blog post with the same script.

```
SCRIPT_REAL("from vaderSentiment.vaderSentiment import  
SentimentIntensityAnalyzer  
vs=[]  
analyzer = SentimentIntensityAnalyzer()  
for i in range(0,len(__arg1)):  
a = analyzer.polarity_scores(__arg1[i])['compound']
```

```
vs.append(a)  
return vs", ATTR([Comment Text]))
```

Reply

Submitted by Bora Beran on December 17, 2016 – 11:38pm

Also please don't forget to install the package

```
pip install VaderSentiment
```

If currently active Anaconda environment is Tableau-Python-Server this will make the package immediately accessible to your running TabPy server

You can find more info here:

<https://github.com/tableau/TabPy/blob/master/server.md#extending-the-env...>

Reply

Submitted by Dave Savell on December 18, 2016 – 1:50am

You superstar Bora.

Thanks for going above and beyond over a weekend to find a solution to this. I was struggling with very limited knowledge to get this working.

Dave.

Reply

Submitted by saranya (not verified) on January 20, 2017 – 4:51am

Python is a widely used high level, general purpose; interpreted, dynamic programming language supports multiple programming paradigms....

Reply

Submitted by vijaykumar h on July 16, 2017 – 11:13pm

Hello,

My company network is not allowing to installing and Run the TabPy server, I was suggested to configure the Proxy, I get ProtocolError('Connection aborted') error message.

Could you please suggest some steps to resolve this?

TY,

Vijay

[Reply](#)

Submitted by gordon lyn on August 7, 2017 - 4:28pm

hi,i'm getting this problem while using tabpy. using the sample - superstore datasets, and i want to cluster the sub-category using the sum([Profit]) and sum([Sales]), but with the error returns: ValueError : n\_samples=1 should be >= n\_clusters=2. And here's my script:

```
SCRIPT_str("  
from sklearn.cluster import KMeans  
kmeans = KMeans(n_clusters=2)  
tmp=[]  
for i in range(len(__arg1)):  
tmp.extend([[__arg1[i],__arg2[i]])]  
KMmodel = kmeans.fit(tmp)  
labels = KMmodel.labels_  
return labels",  
sum([Profit]),sum([Sales]))
```

[Reply](#)

Submitted by zhanju (not verified) on August 16, 2017 - 7:27pm

Python calls the Java code, and then calls the Python program using TabPy. The first calculation succeeds, and when it is re calculated, the table is computed, and then the Python will hang up.

```
SCRIPT_STR("import jpype
import os.path
jarpath = os.path.join(os.path.abspath('.'), 'E:/resource/test.jar')
if not jpype.isJVMStarted():
jpype.startJVM(jpype.getDefaultJVMPath(), '-ea', '-Djava.class.path=%s' % jarpath)

jprint = jpype.java.lang.System.out.println
JDClass = jpype.JClass('com.Utills')
a = JDClass()

text = a.encryptData__ECB(__arg[0])

return text",ATTR([细分]))
```

[Reply](#)

Submitted by rakesh Gupta (not verified) on September 21, 2017 - 4:42am

I have my tableau python server running . i open my code in jupyter and do run all . it does not give any errors . but it does not create the endpoint . i was looking under staging folder .

any idea why it is not creating end points . i don't see anything under query objects also

[Reply](#)

Add new comment

Your name Kristy M.

Comment \*

☐ Notify me when new comments are posted

SAVE

## Subscribe to our blog

Business E-mail \*

Kristy.McGee@tn.gov

SUBSCRIBE

WE RESPECT YOUR PRIVACY

English

[System Status](#) [Blog](#) [Academic](#) [Careers](#) [Contact Us](#)[LEGAL](#) [PRIVACY](#) [UNINSTALL](#)

© 2003-2018 TABLEAU SOFTWARE. ALL RIGHTS RESERVED