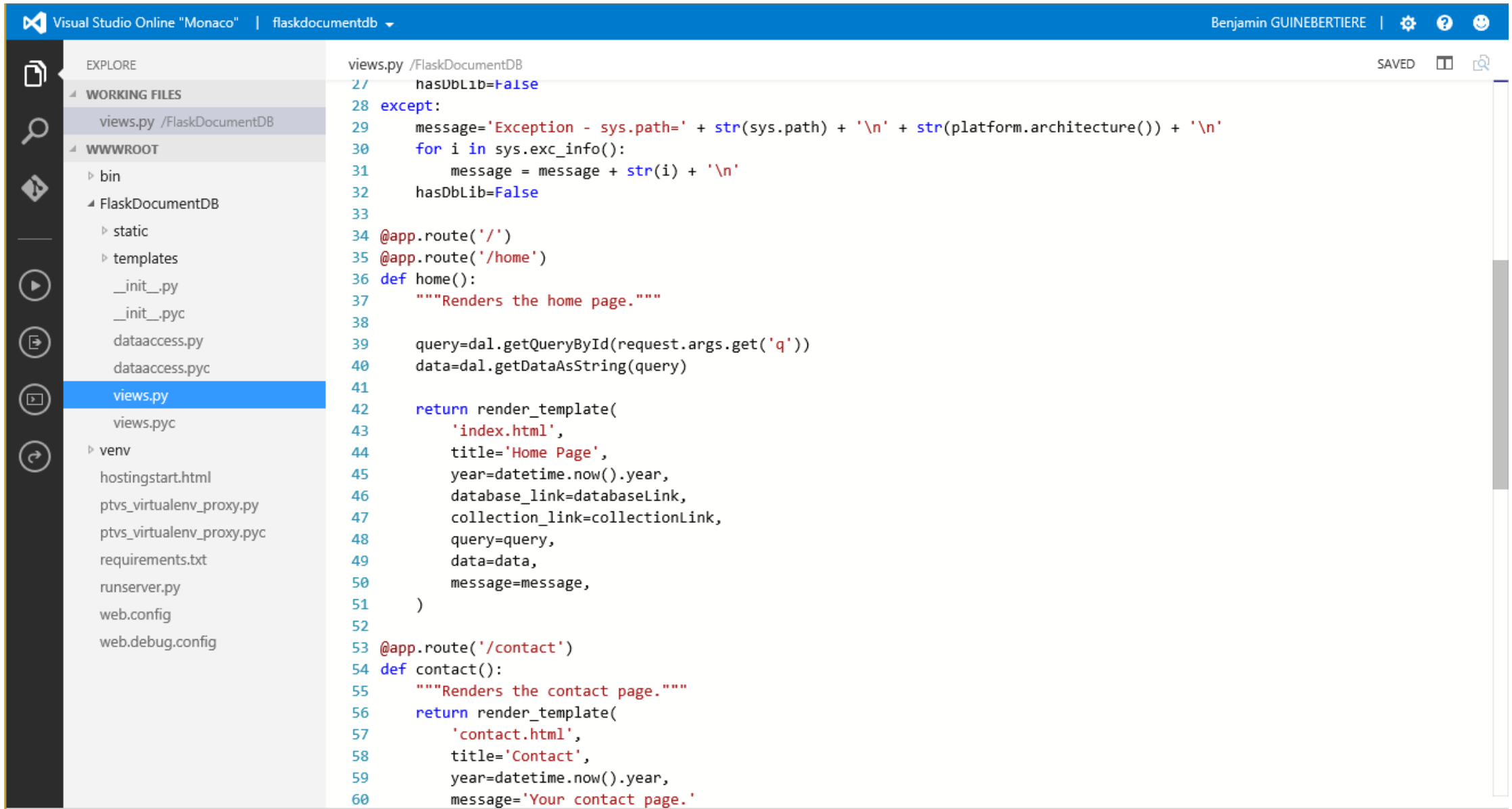
Benjamin Guinebertière

Technical Evangelist, Microsoft France
Azure, data insights, machine learning
@benjguin | http://3-4.fr

Web App


Visual Studio Online "Monaco" | flaskdocumentdb
Benjamin GUINEBERTIERE |
EXPLORE
WORKING FILES
views.py /FlaskDocumentDB
WWWROOT
bin
FlaskDocumentDB
static
templates
__init__.py
__init__.pyc
dataaccess.py
dataaccess.pyc
views.py
views.pyc
venv
hostingstart.html
ptvs_virtualenv_proxy.py
ptvs_virtualenv_proxy.pyc
requirements.txt
runserver.py
web.config
web.debug.config
views.py /FlaskDocumentDB
SAVED
hasDbLib=False
except:
message='Exception - sys.path=' + str(sys.path) + '\n' + str(platform.architecture()) + '\n'
for i in sys.exc_info():
message = message + str(i) + '\n'
hasDbLib=False
@app.route('/')
@app.route('/home')
def home():
"""Renders the home page."""
query=dal.getQueryById(request.args.get('q'))
data=dal.getDataAsString(query)
return render_template(
'index.html',
title='Home Page',
year=datetime.now().year,
database_link=databaseLink,
collection_link=collectionLink,
query=query,
data=data,
message=message,
)
@app.route('/contact')
def contact():
"""Renders the contact page."""
return render_template(
'contact.html',
title='Contact',
year=datetime.now().year,
message='Your contact page.'

Python on HDInsight

Python2.7 is installed by default on HDInsight 3.0 and later clusters. Hive can be used with this version of Python for stream processing (data is passed between Hive and Python using STDOUT/STDIN).

HDInsight also includes Jython, which is a Python implementation written in Java. Pig understands how to talk to Jython without having to resort to streaming, so it's preferable when using Pig.

Hive and Python

Python can be used as a UDF from Hive through the HiveQL **TRANSFORM** statement. For example, the following HiveQL invokes a Python script stored in the **streaming.py** file.

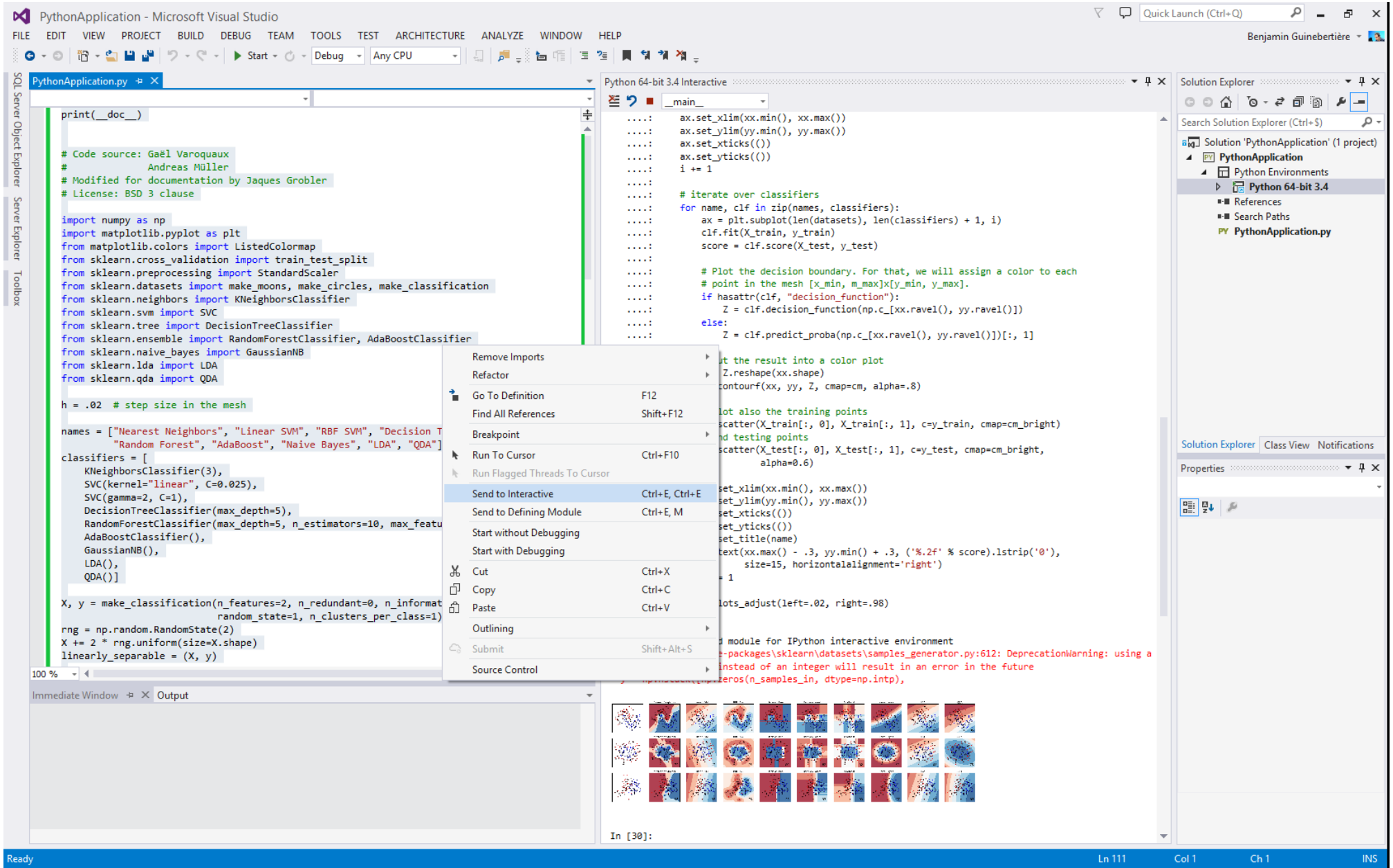
Linux-based HDInsight

```
add file wasb:///streaming.py;

SELECT TRANSFORM (clientid, devicemake, devicemodel)
  USING 'streaming.py' AS
  (clientid string, phoneLable string, phoneHash string)
FROM hivesampletable
ORDER BY clientid LIMIT 50;
```

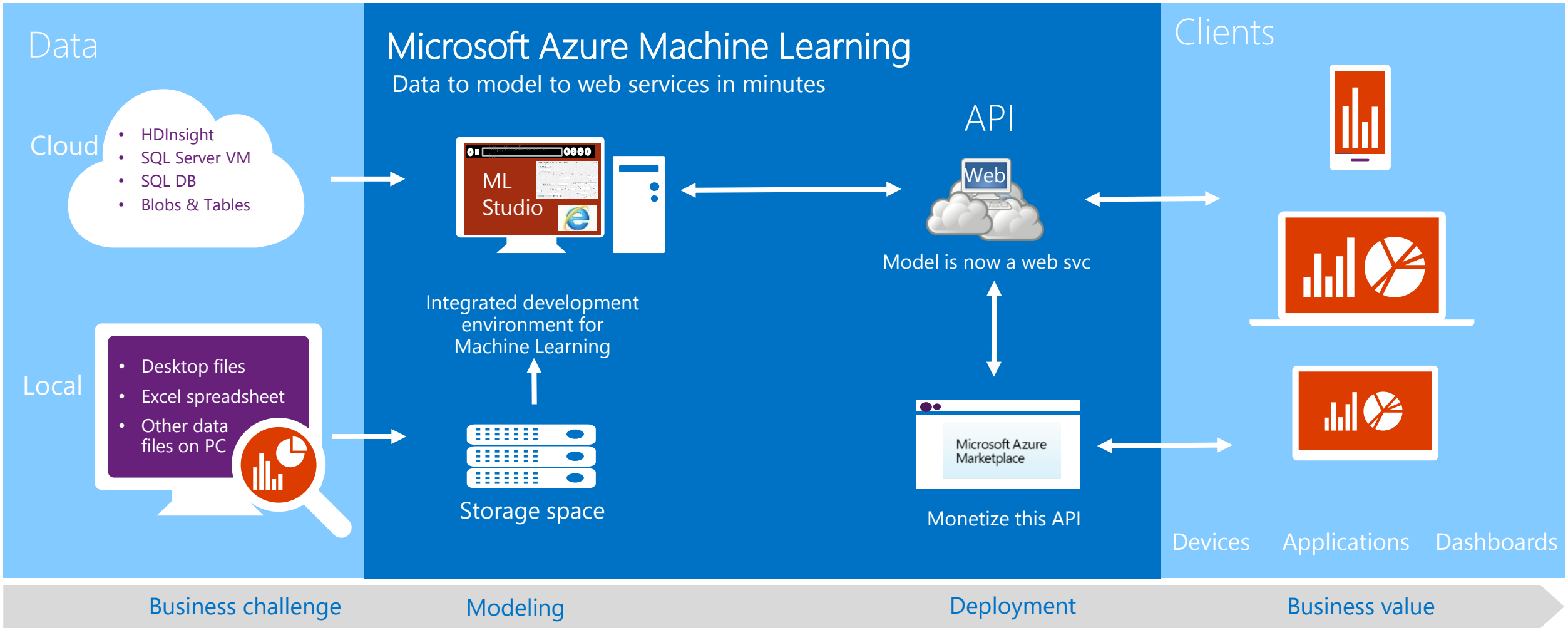
Windows-based HDInsight

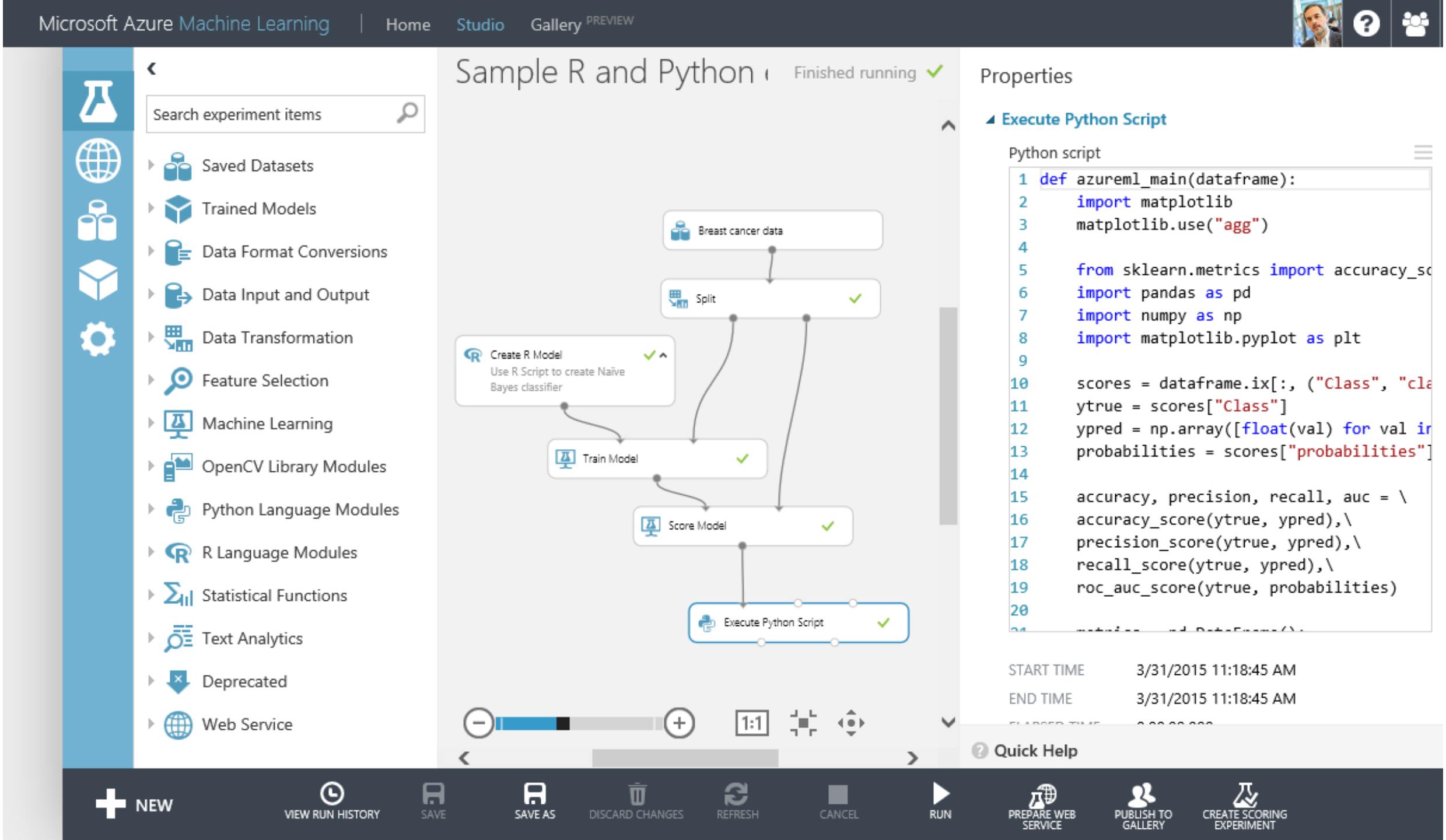
```
add file wasb:///streaming.py;

SELECT TRANSFORM (clientid, devicemake, devicemodel)
  USING 'D:\Python27\python.exe streaming.py' AS
  (clientid string, phoneLable string, phoneHash string)
FROM hivesampletable
ORDER BY clientid LIMIT 50;
```

Python Tools for Visual Studio



Azure ML


Data
Cloud
HDInsight
SQL Server VM
SQL DB
Blobs & Tables
Local
Desktop files
Excel spreadsheet
Other data files on PC
Microsoft Azure Machine Learning
Data to model to web services in minutes
ML Studio
Integrated development environment for Machine Learning
Storage space
API
Web
Model is now a web svc
Microsoft Azure Marketplace
Monetize this API
Clients
Devices
Applications
Dashboards
Business challenge
Modeling
Deployment
Business value


Microsoft Azure Machine Learning
Home
Studio
Gallery PREVIEW
Search experiment items
Saved Datasets
Trained Models
Data Format Conversions
Data Input and Output
Data Transformation
Feature Selection
Machine Learning
OpenCV Library Modules
Python Language Modules
R Language Modules
Statistical Functions
Text Analytics
Deprecated
Web Service
Sample R and Python
Finished running
Breast cancer data
Split
Create R Model
Use R Script to create Naïve Bayes classifier
Train Model
Score Model
Execute Python Script
1:1
Properties
Execute Python Script
Python script
1 def azureml_main(dataframe):
2 import matplotlib
3 matplotlib.use("agg")
4
5 from sklearn.metrics import accuracy_sc
6 import pandas as pd
7 import numpy as np
8 import matplotlib.pyplot as plt
9
10 scores = dataframe.ix[:, ("Class", "cla
11 ytrue = scores["Class"]
12 ypred = np.array([float(val) for val ir
13 probabilities = scores["probabilities"]
14
15 accuracy, precision, recall, auc = \
16 accuracy_score(ytrue, ypred),\
17 precision_score(ytrue, ypred),\
18 recall_score(ytrue, ypred),\
19 roc_auc_score(ytrue, probabilities)
20
START TIME 3/31/2015 11:18:45 AM
END TIME 3/31/2015 11:18:45 AM
Quick Help
NEW
VIEW RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
REFRESH
CANCEL
RUN
PREPARE WEB SERVICE
PUBLISH TO GALLERY
CREATE SCORING EXPERIMENT

▸ Standard Output

▾ Standard Error

Interpreter produced no output.

▾ Graphics