

# Analysis with Python

```
import pandas as pd
import matplotlib.pyplot as plt

df1 = pd.read_csv('/home/vedang/Documents/CS671/Programming_languages.csv')
df2 = pd.read_csv('/home/vedang/Documents/CS671/Technologies.csv')
```

FINISHED

	c++	haskell	java	julia	lua	matlab	octave	python	r	sas	scala	sql
0	1	0	1	0	0	1	0	1	1	1	0	0
1	0	0	0	0	0	0	0	1	0	0	0	0
2	1	0	0	0	0	0	0	1	1	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	1	0	0	0

Took 0 sec. Last updated by anonymous at November 05 2017, 4:51:51 PM. (outdated)

```
df2.head()
```

FINISHED

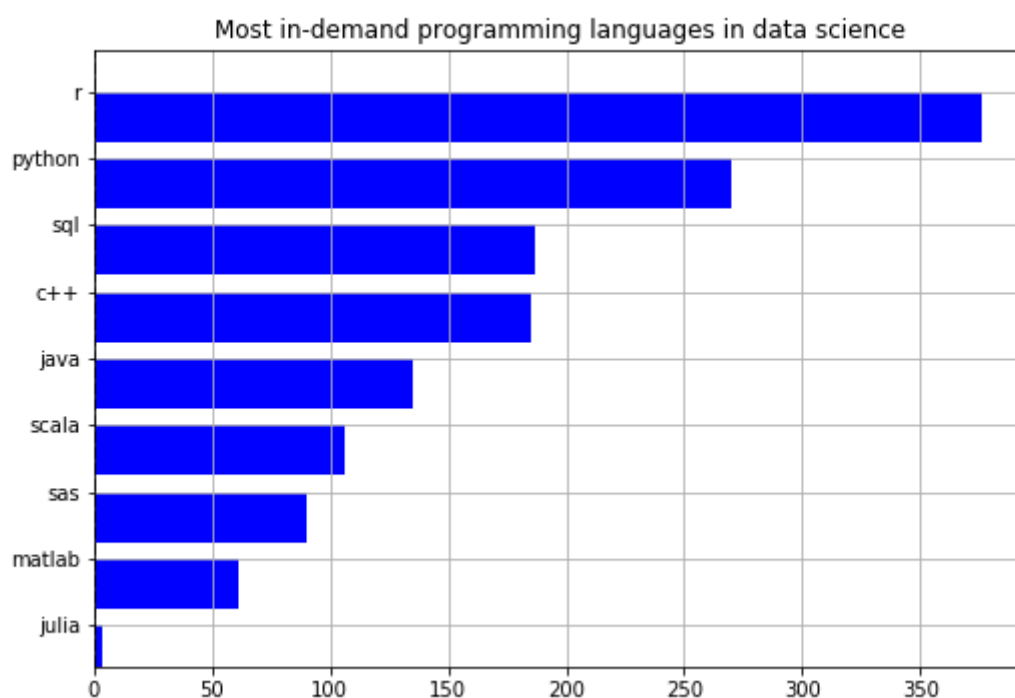
	aws	cassandra	hadoop	hbase	hive	mongodb	mysql	nltk	numpy	pandas	pig	spark	tensorflow
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	1	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0	0	1	0

Took 0 sec. Last updated by anonymous at November 05 2017, 4:52:15 PM.

```
df = pd.Series(df1.sum())
df.sort()
df = df.drop(['octave', 'lua', 'haskell'])
df.plot(kind='barh', title='Most in-demand programming languages in data science')
```

FINISHED

<matplotlib.axes.\_subplots.AxesSubplot object at 0x7f4a376d2ed0>

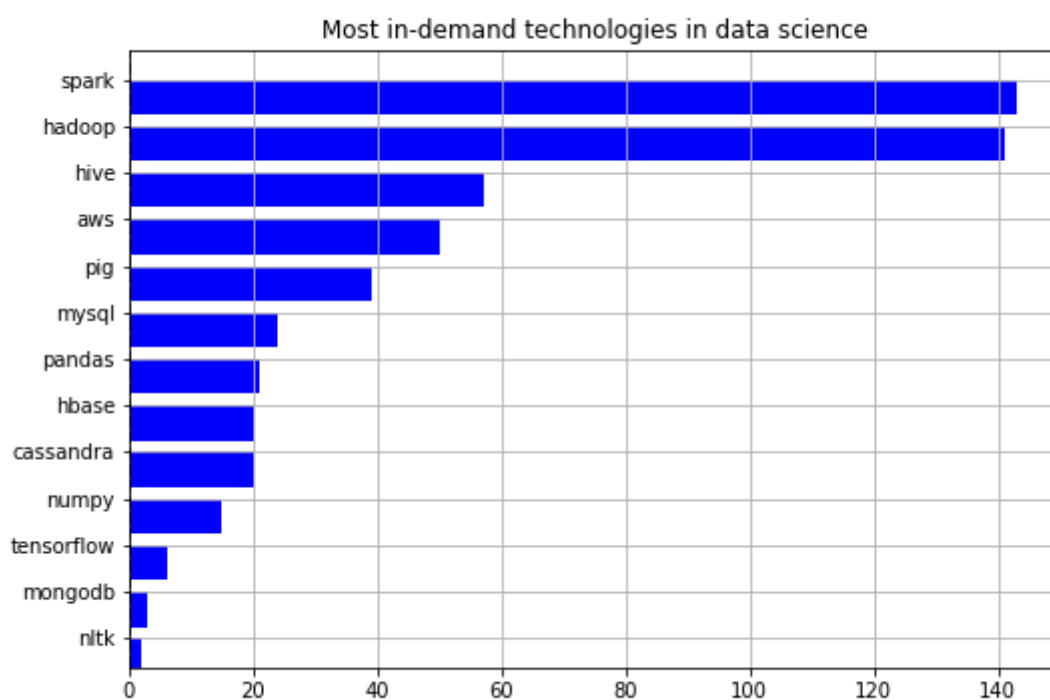


Took 1 sec. Last updated by anonymous at November 05 2017, 5:33:25 PM.

FINISHED

```
df = pd.Series(df2.sum())
df.sort()
#df = df.drop(['octave', 'lua', 'haskell'])
df.plot(kind='barh', title='Most in-demand technologies in data science')
```

<matplotlib.axes.\_subplots.AxesSubplot object at 0x7f4a3b683b50>



Took 2 sec. Last updated by anonymous at November 05 2017, 5:33:48 PM.

READY