

Fix the bug by encircling the part (maximum of one line) of the code snippet you wish to replace with your correction. If you want to insert a line instead, use an arrow to indicate where you want to insert your correction. Your correction is up to a line long only.

Numpy

1)

Code Snippet	Expected Output	Actual Output
<pre>def add_arrays(): a = range(10) b = np.linspace([10,100,10]) return np.array(a)+b</pre>	<pre>array([10., 21., 32., 43., 54., 65., 76., 87., 98., 109.])</pre>	<pre>STypeError: _linspace_dispatcher() missing 1 required positional argument: 'stop'</pre>

2)

Code Snippet	Expected Output	Actual Output
<pre>def matrix_multiply(): a = np.ones([3,5]) b = np.diag(a.shape) return np.matmul(a,b)</pre>	<pre>array([[5., 5., 5.], [5., 5., 5.], [5., 5., 5.]])</pre>	<pre>ValueError: matmul: Input operand 1 has a mismatch in its core dimension 0, with gufunc signature (n?,k),(k,m?)->(n? ,m?) (size 3 is different from 5)</pre>

3)

Code Snippet	Expected Output	Actual Output
<pre>def square_some_elements(): g = np.array([[1,2,3] , [4,5,6] , [7,8,9]]) g[g<=6 and g>=2] = g[g<=6 and g>=2]**2 return g</pre>	<pre>array([[1, 4, 9], [16, 25, 36], [7, 8, 9]])</pre>	ValueError: The truth value of an array with more than one element is ambiguous. Use a.any() or a.all()

4)

Code Snippet	Expected Output	Actual Output
<pre>def long_broadcast(): a = np.ones([2,2,1]) b = np.ones([2,3,2,3]) a = a[:, :, :, None] return a*b</pre>	< an array of shape (2,3,2,3) containing only 1 >	ValueError: operands could not be broadcast together with shapes (2,2,1,1) (2,3,2,3)

Pandas

5)

Code Snippet	
<pre>def read_file(): df = pd.read_csv("mnt/data/public/nyctaxi/trip_data/trip_data_1.csv" , nrows=100) return df</pre>	
Expected Output	Actual Output
< a DataFrame >	FileNotFoundError: [Errno 2] No such file or directory: 'mnt/data/public/nyctaxi/trip_data/trip_data_1.csv'

6)

Code Snippet	
<pre>def read_special_file_again(): df = pd.read_csv("/mnt/data/public/wikipedia/clickstream/" "clickstream/2017-11/clickstream-enwiki-2017-11.tsv.gz") return df</pre>	
Expected Output	Actual Output
< a DataFrame >	ParserError: Error tokenizing data. C error: Expected 1 fields in line 103, saw 2

7)

Code Snippet																	
<pre>def capitalize_words(): data = { "name": ["april", "may", "august"] , "job" : ["student", "teacher", "friend"] , "age" : [10, 25, 30] } df = pd.DataFrame(data) df['name'] = df['name'].upper() return df</pre>																	
Expected Output	Actual Output																
<table><tr><th></th><th>name</th><th>job</th><th>age</th></tr><tr><td>0</td><td>APRIL</td><td>student</td><td>10</td></tr><tr><td>1</td><td>MAY</td><td>teacher</td><td>25</td></tr><tr><td>2</td><td>AUGUST</td><td>friend</td><td>30</td></tr></table>		name	job	age	0	APRIL	student	10	1	MAY	teacher	25	2	AUGUST	friend	30	AttributeError: 'Series' object has no attribute 'upper'
	name	job	age														
0	APRIL	student	10														
1	MAY	teacher	25														
2	AUGUST	friend	30														