

Business Analytics Programming - Spring 2019, Practice Set # 5

	a	b	text
0	0	0	Thanks so much for driving me home.
1	1	1	Thanks so much for cooking dinner. I really appreciate it.
2	2	2	Excuse me sir, you dropped your wallet.
3	3	3	I'm sorry for the mess. I wasn't expecting anyone today.
4	4	4	My name is Sophie and I'm learning English.

1. Create a table df2, with just the columns 'text' and 'b'?
2. In table df2, create 2 columns named 'Polarity' and 'Subj'. Using the Textblob function, generate the polarity and subjectivity for each text and put it in the respective columns.
3. In table df2, create 1 column named 'textl' that is all the words from the text column in lowercase.
4. In table df2, create 1 column named 'textlist' that is the 'textl' values split into words into a list.
5. In table df2, create 1 column named 'textarray' that is the 'textlist' values converted into a numpy array.
6. In table df2, create 1 column named 'textadj' that is the 'textarray' arrays with the words 'I' and 'me', taken out.
7. In table df2, create 1 column named 'textadj2' that is the 'textadj' values with the word 'batman', inserted as the first item in the array.
8. In table df2, create 1 column named 'textadj3' that is the 'textadj2' arrays with any word with the string 'ing' removed from the arrays.
9. In table df2, create 1 column named 'textnew' that is the 'textadj3' arrays concatenated into each new string. (per row)
10. Create one giant text that is the concatenation of all the texts in column 'textnew'.
11. Get the frequency of all the words in the 'textadj3' column.