

CIS*4500 (Fall 2019) Marking Scheme for Assignment One		Marks	Comments
Sentence Splitting (12)	Correct splitting (12)		
Tokenization (48)	Labels (2)		
	Words (6)		
	Numbers (8)		
	Apostrophized (14)		
	Hyphenated (10)		
	Newlines (2)		
	Delimiters (4)		
	Punctuations (2)		
POS-Tagging (10)	Correct tagging (10)		
Data Analysis (15)	Number of documents (2)		
	Document lengths by sentences (4)		
	Document lengths by tokens (4)		
	Avg Sentence lengths (5)		
Style and Documentation (15)	Readme (4)		
	Makefile (3)		
	Testing Plan (5)		
	Coding Style (3)		
Total (100)			