te Topic	Developer	Tasks	Data Resources			
2/20 Math in scipy	Tian	-scipy.optimize.minimize -scipy.integrate.quad -conceptual question abou -Evaluate the definite integ -linear algebra: transpose, -exercise: give the standar	 Intermeidate Python - Scientific Programming: http://snowball.millersville.edu/-adecaria/ ESCl386P/esci386-lesson04-operators.pdf SciPy Reference Guide: https://docs.scipy.org/doc/scipy/reference/tutorial/general.html 		lide=id.g1fd457f1a5	_0_14
2/22 scikit-leam intro	Keeley	-test-train-validation split -OLS -LASSO -Ridge	https://github.com/ds-modules/core-resources/folob/master/machine-learning/examples/RidgeLassoEN.ipynb https://github.com/dlab-berkeley/python-machine-learning https://github.com/miberkeley/Data-Science-Decal-Fall-2017/blob/master/Day1-PythonBootcamp/Scikit-Learn_IntroStudenLipynb ds100fa17 lab 11: cross validation			
2/27 model selection	Aniket					
3/1 preprocessing text	Tian	-tokenizing -stemming -chunking	UN debate transcripts: <pre>c-preprocessing</pre> and chunking: <pre>https://github.com/dlab-berkeley/python-text-analysis/blob/master/Intro_to_TextAnalysis/Intro_to_TextAnalys</pre>			
3/6 intro to text analysis	Gibson	-bag of words -document-term frequency	-document-term frequency (you can probably use almost everything before the vector space model section): https://glithub.com/henchc/textxd-2017/blob-bag of words definition + code example: http://www.ds100.org/fa17/assets/notebooks/22-lec/Feature_Engineering.html UN debate transcripts: c-intro to text analysis (first section only): https://github.com/dlab-berkeley/python-text-analysis/blob/master/intro_to_TextAnalysis/intro_to_TextAnalysis.intro_to_TextAnaly	1.ipynb		
3/8 None- project work day	-	-				
3/13 parsing XML data	Jason Tina	-requests library -XML syntax (document-of-locating content with XPA				
3/15 regex/dictionary methods	Keiko	-search -match -findall -building dicitionaries	https://github.com/dlab-berkeley/python-text-analysis/blob/master/Intro_to_TextAnalysis/Intro_to_TextAnalysis.jpynb UN debate transcripts: https://github.com/henchotextxd-2017/blob/master/03-regex.jpynb			
3/20 text classification continued			https://qithub.com/henchc/textxd-2017/blob/master/07-Classification.jpynb			
3/22 topic modeling			https://github.com/dlab-berkeley/python-text-analysis/blob/master/Infro_to_TextAnalysis/Infro_to_TextAnalysis.jpynb			
4/3 word-to-vector models			https://github.com/dlab-berkeley/python-text-analysis/blob/master/Intro to TextAnalysis/Intro to TextAnalysis.jpynb			
4/5 word-to-vector models II						
4/10 word embedding models			https://github.com/henchc/textxd-2017/blob/master/08-Word-Embeddings.jpynb			
1/12 word embedding models II						
4/17 feature selection		-recursive feature eliminati	on .			
4/19 ensemble methods		-random forest -AdaBoost				