****

**Text Analytics**

**Assignment 2 – Word Frequency Analysis**

**Yanzhe Li**

**02/22/2017**

Word Frequency Analysis on Reviews from Vogue During 2015 Fashion Weeks

1. Simple bag-of-words approach

In this approach, a simple word frequency analysis was used. In addition, to improve the performance, the python model also allow user to set n for n-grams in case the user wants to see not only words but also phrases. In this case, we set n = 1.

Below chart “Words Frequency – Simple Bag of Words Approach” shows the top 30 frequent features

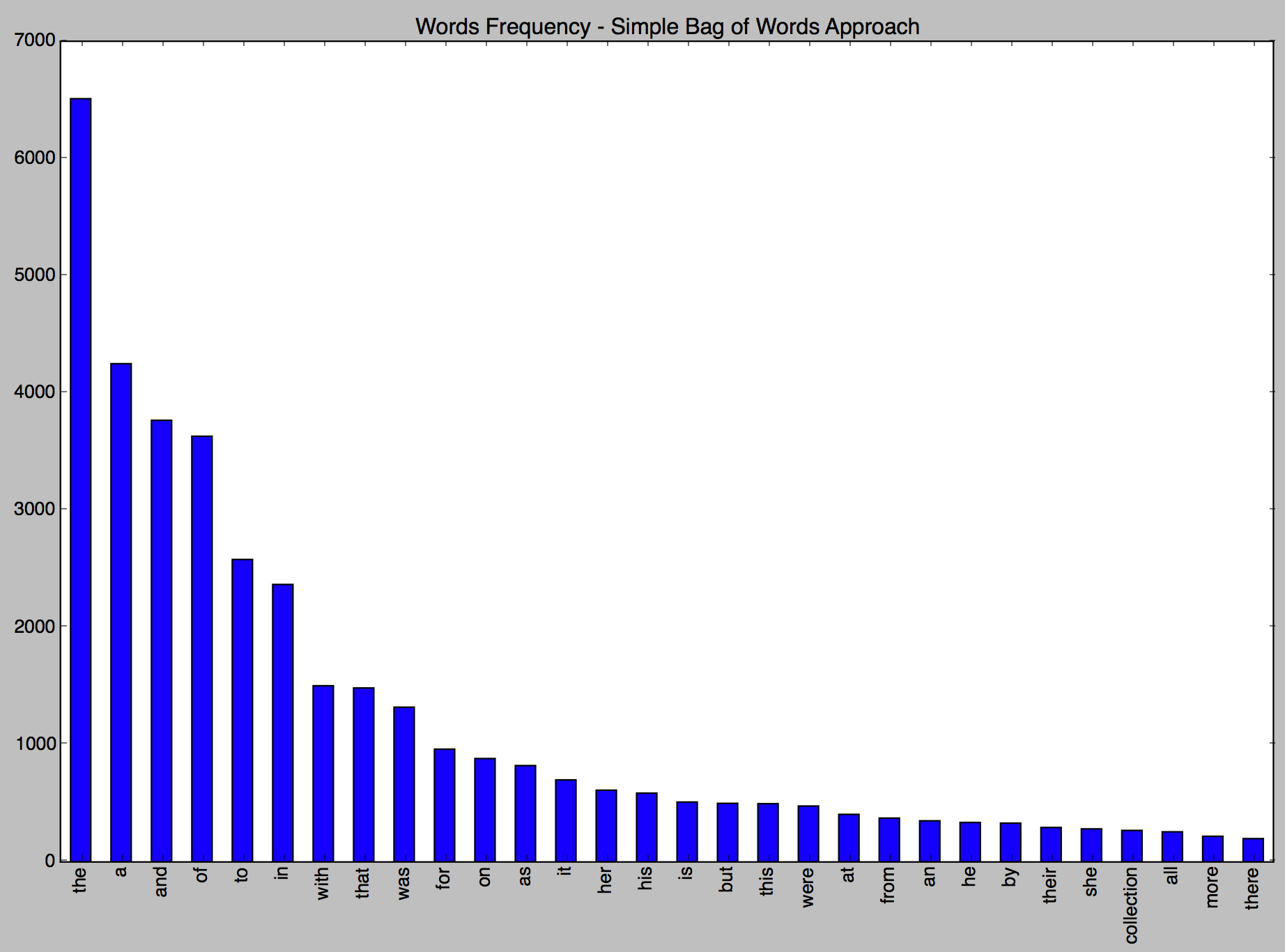
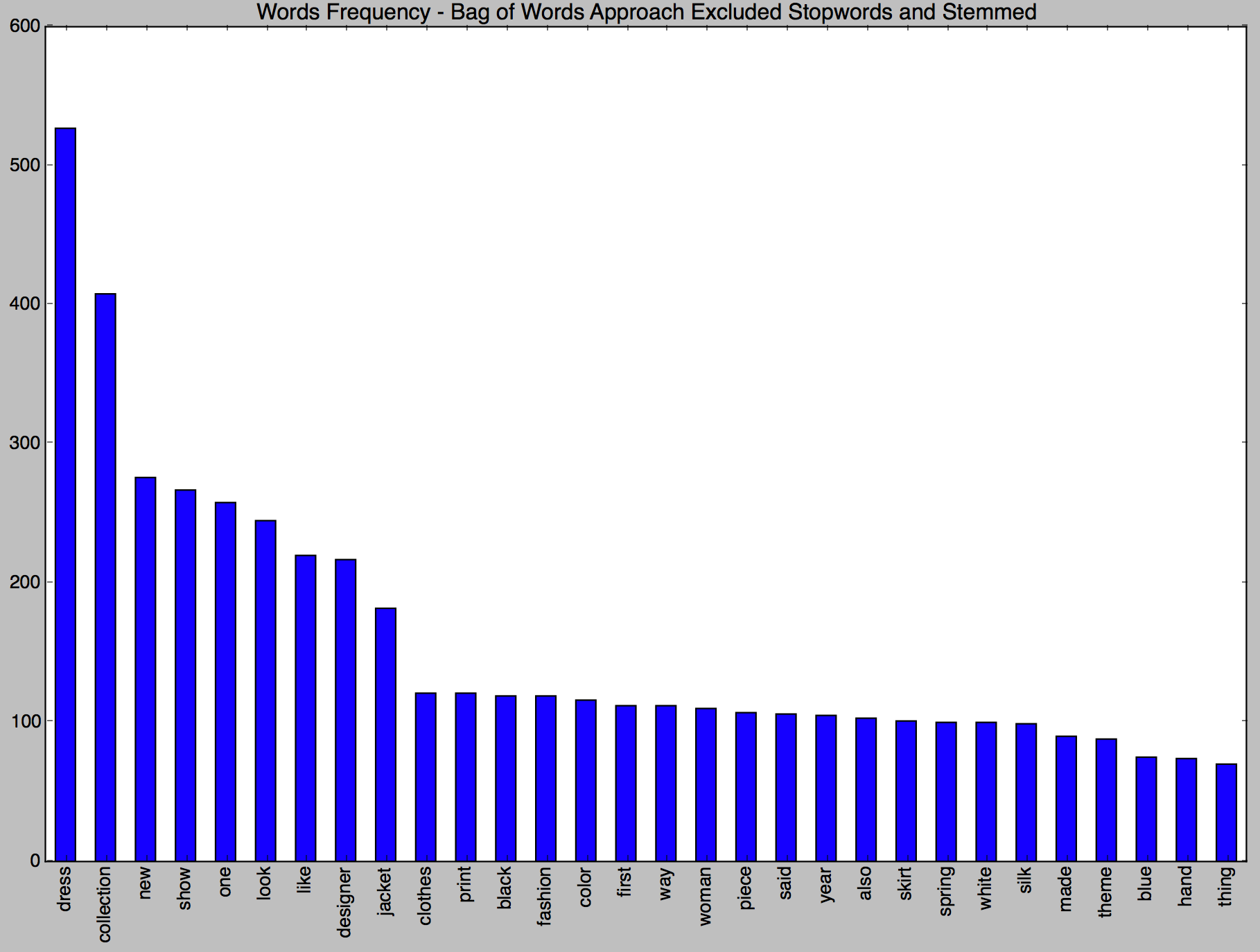
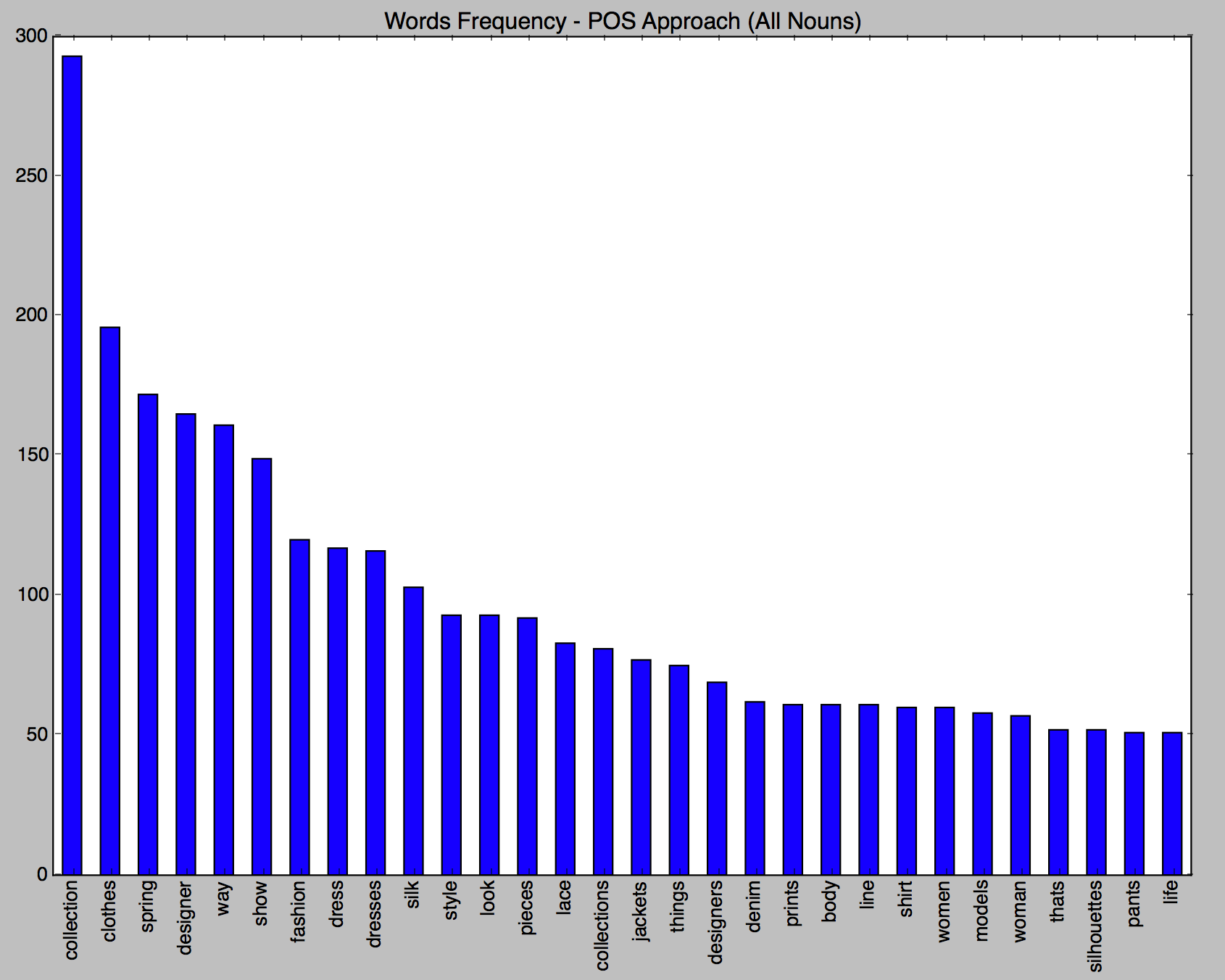


Figure1.1 Words Frequency – Simple Bag of Words Approach

1. Bag-of-words approach with stemming and stop words removal



1. POS approach and focus on all the noun forms (NN, NNP, NNS, NNPS)



1. POS approach and only focus on NNP

