Homework 1 Output

$ **python task-1.py**

task-1.py // print out the "text" part of the 300th review
I know live in Vegas and always wanted to try pinks. It's nothing special. I'd rather buy
a better tasting & more robust hotdog & a drink @ Costco for just $1.50 than spend $6-$7
for something less appealing @ pinks.

$ **python task-2.py**

task-2.py // print out the tokenized "text" part of the first 3 reviews
['pinks', 'at', 'planet', 'hollywood', 'totally', 'fucking', 'fails', 'they', 're',
'there', 'when', 'they', 'want', 'and', 'that', 's', 'it', 'from', 'what', 'the', 'info',
'desk', 'girls', 'said', 'they', 'come', 'by', 'like', '4', 'times', 'a', 'year', 'and',
'it', 's', 'a', 'craps', 'shoot', 'whether', 'or', 'not', 'you', 'll', 'be', 'around',
'when', 'they', 'are', 'now', 'i', 'have', 'to', 'go', 'to', 'hollywood', 'to', 'try',
'this', 'place', 'out', 'luckily', 'for', 'me', 'i', 'll', 'be', 'in', 'socal', 'next',
'weekend']
['i', 'had', 'my', 'hopes', 'of', 'trying', 'pink', 's', 'hot', 'dogs', 'for', 'the',
'longest', 'time', 'and', 'i', 'finally', 'tried', 'it', 'here', 'at', 'planet',
'hollywood', 's', 'race', 'sports', 'book', 'ooooooo', 'what', 'a', 'disappointment',
'it', 'was', 'just', 'a', 'small', 'little', 'stand', 'with', 'one', 'worker', 'who',
'looked', 'like', 'he', 'didn', 't', 'want', 'to', 'be', 'there', 'i', 'ordered', 'a',
'regular', 'hot', 'dog', 'combo', 'which', 'came', 'with', 'chips', 'that', 'cost', '5',
'95', 'the', 'chili', 'dog', 'combo', 'was', '6', '95', 'it', 'just', 'tasted', 'like',
'a', 'regular', 'hot', 'dog', 'except', 'that', 'it', 'had', 'a', 'nice', 'snap', 'to',
'it', 'which', 'made', 'me', 'give', 'the', 'place', 'an', 'extra', 'star', 'other',
'than', 'that', 'nothing', 'special', 'about', 'it', 'the', 'chili', 'tasted', 'like',
'crap', 'on', 'it', 'probably', 'because', 'it', 'was', 'old', 'and', 'in', 'the',
'container', 'all', 'day', 'ewww', 'actually', 'i', 'bet', 'all', 'the', 'food', 'they',
'served', 'there', 'was', 'old', 'ekkkk', 'o', 'i', 'had', 'asked', 'the', 'guy',
'working', 'there', 'when', 'they', 'will', 'be', 'opening', 'the', 'actual', 'official',
'stand', 'and', 'he', 'said', 'he', 'wasn', 't', 'sure', 'however', 'they', 'will', 'be',
'opening', 'one', 'near', 'the', 'hawaiian', 'tropic', 'entrance', 'of', 'planet',
'hollywood', 'sooo', 'look', 'out', 'hopefully', 'by', 'that', 'time', 'the', 'dogs',
'will', 'be', 'better']
['decent', 'if', 'you', 've', 'been', 'to', 'the', 'original', 'la', 'location',
'prepare', 'to', 'be', 'somewhat', 'disappointed', 'food', 'meh', 'the', 'dogs',
'seemed', 'precooked', 'the', 'buns', 'were', 'plain', 'and', 'the', 'mushrooms',
'looked', 'canned', 'the', 'planet', 'hollywood', 'dog', 'was', 'damn', 'near', '8',
'alone', 'i', 'had', 'to', 'ask', 'for', 'a', 'side', 'of', 'nacho', 'cheese', 'because',
'the', 'ph', 'dog', 'here', 'only', 'comes', 'with', 'a', 'slice', 'of', 'what',
'looked', 'like', 'kraft', 'cheese', 'the', 'prices', 'were', 'at', 'least', '2', 'more',
'than', 'the', 'la', 'location', 'but', 'i', 'guess', 'its', 'understandable', 'since',
'its', 'las', 'vegas', 'the', 'fries', 'weren', 't', 'freshly', 'prepared', 'like',
'the', 'la', 'location', 'either', 'and', 'to', 'top', 'it', 'off', 'they', 'didn', 't',
'have', 'grape', 'crush', 'soda', 'or', 'yoo', 'hoo', 'like', 'the', 'la', 'location',
'service', 'the', 'cashier', 'chicks', 'were', 'nice', 'the', 'dogs', 'and', 'sides',
'were', 'served', 'up', 'made', 'in', 'assembly', 'line', 'unlike', 'the', 'personal',
'attention', 'you', 'get', 'from', 'one', 'person', 'at', 'the', 'la', 'location',

'overall', 'it', 'was', 'just', 'meh', 'i', 'd', 'rather', 'drive', '4', 'hours', 'then',
'wait', 'in', 'line', 'for', 'another', 'hour', 'to', 'have', 'an', 'authentic', 'pink',
's', 'hotdog', 'i', 'probably', 'won', 't', 'come', 'back', 'here', 'unless', 'they',
'somehow', 'manage', 'to', 'replicate', 'everything', 'that', 'the', 'la', 'location',
'offers', 'this', 'place', 'seemed', 'like', 'the', 'mcdonalds', 'version', 'of', 'pink',
's', 'the', 'only', 'good', 'thing', 'is', 'that', 'they', 'accept', 'credit', 'cards',
'while', 'the', 'la', 'location', 'is', 'only', 'accepts', 'cash']

$ python task-3.py
task-3.py // print out the "text" part of the first review after stopword removal
['pinks', 'planet', 'hollywood', 'totally', 'fucking', 'fails', 'info', 'desk', 'girls',
'4', 'times', 'year', 'craps', 'shoot', 'hollywood', 'place', 'luckily', 'socal',
'weekend']

$ python task-4.py
task-4.py // print out the "text" part of the first review after stemming
['pink', 'planet', 'hollywood', 'total', 'fuck', 'fail', 'info', 'desk', 'girl', '4',
'time', 'year', 'crap', 'shoot', 'hollywood', 'place', 'luckili', 'socal', 'weekend']

$ python task-5.py
task-5.py //
i) Find the top twenty unigrams in all reviews without stemming:
[('dog', 1121), ('hot', 924), ('dogs', 570), ('pink', 425), ('chili', 421), ('good',
421), ('cheese', 351), ('vegas', 344), ('place', 298), ('fries', 257), ('la', 250),
('food', 215), ('pinks', 185), ('hollywood', 180), ('bacon', 153), ('strip', 153),
('pretty', 150), ('great', 147), ('eat', 145), ('ordered', 143)]

ii) Find the top twenty unigrams in all reviews with stemming:
[('dog', 1691), ('hot', 924), ('pink', 610), ('good', 431), ('chili', 421), ('chees',
351), ('vega', 345), ('place', 329), ('la', 317), ('fri', 263), ('order', 261), ('eat',
234), ('food', 220), ('onion', 219), ('hollywood', 180), ('time', 171), ('tast', 165),
('hotdog', 159), ('locat', 158), ('strip', 157)]

$ python task-6.py
task-6.py // find the top twenty bigrams in all reviews without stemming
[('hot dog', 511), ('hot dogs', 364), ('chili cheese', 178), ('planet hollywood', 134),
('cheese fries', 84), ('onion rings', 82), ('pink hot', 78), ('las vegas', 67), ('cheese
dog', 63), ('chili dog', 61), ('nacho cheese', 52), ('pretty good', 45), ('dog good',
45), ('hollywood dog', 45), ('polish sausage', 44), ('polish dog', 41), ('vegas strip',
38), ('bacon chili', 37), ('pink la', 36), ('outdoor seating', 35)]