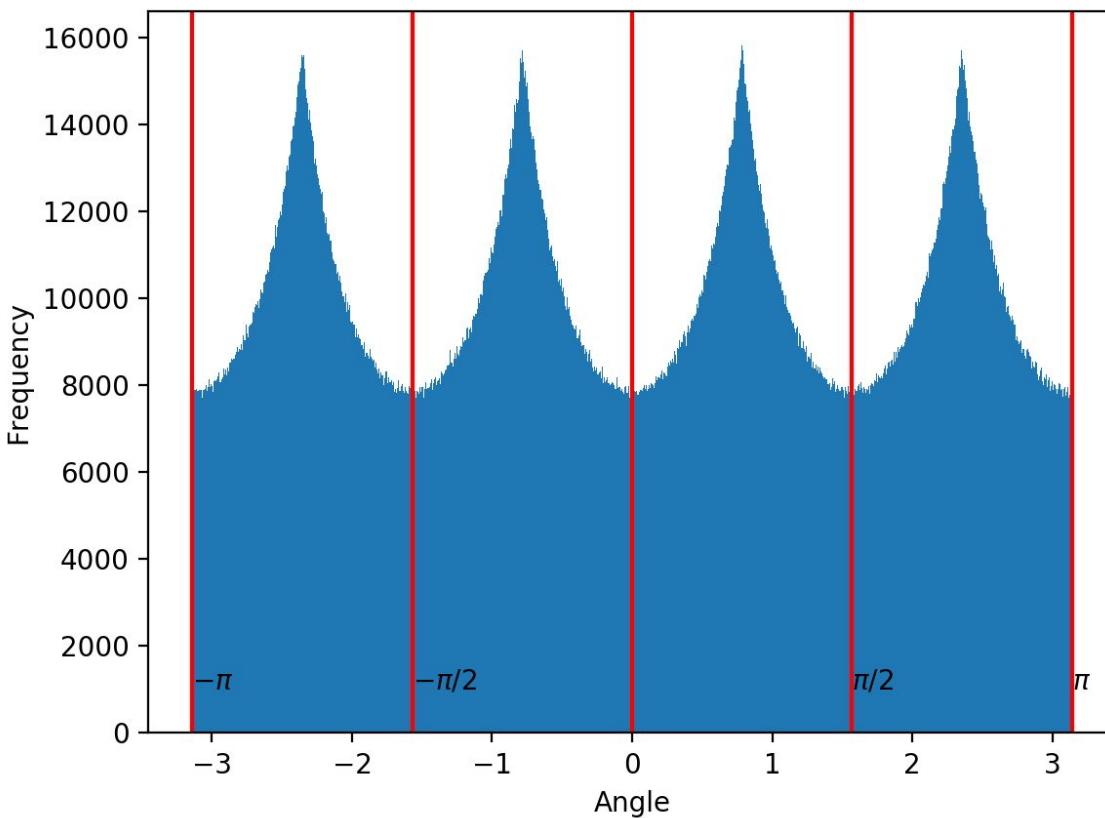


Problem 1

Please refer to [vectors.py](#)



Choosing $a, b \in [-1, 1]$ generates points on a 2×2 square with its center on the origin. The arctan function then produces the angle of the vector generated by connecting the random point (a, b) with the origin.

This produces the bias shown in the graph because there are more vectors possible at the corners of the square (at every $\pi/4$ interval) than there are elsewhere. At every $\pi/2$, this discrepancy is widest because those are the angles where the longest possible vector is shortest.

Problem 2

For the full code, please refer to [bigramcount.py](#)

```
def getBigrams(line):
    bigrams = []
    sentences = re.split('\.|\?|', line.strip().lower())
    for s in sentences:
        words = list( map(lambda w: w.strip(string.punctuation), s.split()) )
        bigrams += [(words[i], words[i+1]) for i in range(0, len(words)-1)]
    return bigrams
```

((u'i', u'am'), 1929)	((u'for', u'the'), 549)	((u'to', u'make'), 386)
((u'in', u'the'), 1755)	((u'of', u'this'), 548)	((u'do', u'you'), 380)
((u'i', u'have'), 1675)	((u'is', u'not'), 546)	((u'pray', u'you'), 378)
((u'my', u'lord'), 1670)	((u'will', u'not'), 541)	((u'thou', u'hast'), 375)
((u'of', u'the'), 1592)	((u'in', u'my'), 539)	((u'for', u'a'), 371)
((u'i', u'will'), 1583)	((u'in', u'his'), 536)	((u'give', u'me'), 370)
((u'to', u'the'), 1532)	((u'let', u'me'), 530)	((u'with', u'me'), 370)
((u'it', u'is'), 1089)	((u'of', u'your'), 508)	((u'and', u'his'), 365)
((u'to', u'be'), 1026)	((u'in', u'a'), 508)	((u'at', u'the'), 364)
((u'that', u'i'), 939)	((u'to', u'me'), 501)	((u'to', u'your'), 363)
((u'i', u'do'), 838)	((u'to', u'my'), 498)	((u'they', u'are'), 362)
((u'and', u'the'), 816)	((u'this', u'is'), 494)	((u'and', u'all'), 360)
((u'and', u'i'), 767)	((u'do', u'not'), 492)	((u'when', u'i'), 358)
((u'you', u'are'), 748)	((u'if', u'i'), 487)	((u'i', u'must'), 358)
((u'the', u'king'), 747)	((u'from', u'the'), 486)	((u'the', u'duke'), 355)
((u'of', u'my'), 746)	((u'on', u'the'), 484)	((u'i', u'had'), 353)
((u'is', u'the'), 685)	((u'such', u'a'), 471)	((u'for', u'my'), 352)
((u'i', u'would'), 678)	((u'out', u'of'), 468)	((u'what', u'is'), 351)
((u'he', u'is'), 659)	((u'in', u'this'), 464)	((u'and', u'a'), 351)
((u'you', u'have'), 654)	((u'a', u'man'), 453)	((u'if', u'thou'), 350)
((u'with', u'the'), 652)	((u'there', u'is'), 440)	((u'my', u'heart'), 346)
((u'i', u'know'), 637)	((u'and', u'so'), 427)	((u'he', u'hath'), 343)
((u'of', u'his'), 623)	((u'you', u'shall'), 427)	((u'will', u'be'), 343)
((u'with', u'a'), 603)	((u'the', u'world'), 425)	((u'of', u'our'), 342)
((u'all', u'the'), 579)	((u'me', u'to'), 420)	((u'i', u'was'), 338)
((u'of', u'a'), 577)	((u'that', u'you'), 418)	((u'of', u'all'), 337)
((u'thou', u'art'), 573)	((u'i', u'shall'), 401)	((u'of', u'it'), 335)
((u'is', u'a'), 572)	((u'and', u'in'), 398)	((u'with', u'him'), 332)
((u'as', u'i'), 570)	((u'king', u'henry'), 395)	((u'for', u'i'), 332)
((u'if', u'you'), 567)	((u'to', u'you'), 394)	((u'in', u'your'), 331)
((u'like', u'a'), 563)	((u'i', u'pray'), 390)	((u'and', u'my'), 330)
((u'shall', u'be'), 561)	((u'and', u'to'), 390)	((u'that', u'he'), 328)
((u'by', u'the'), 559)	((u'i', u'think'), 389)	
((u'no', u'more'), 550)	((u'but', u'i'), 386)	

Problem 3

Please refer to [movies.py](#) for details on implementation

```
=== Problem 3-1: The average number of users a movie is rated by ===  
Average number of users a movie is rated by: 11.0306640194
```

```
=== Problem 3-2: For each genre, the average rating of all movies in that genre ===  
(no genres listed) genre: 3.72058823529 rating on average  
Documentary genre: 3.68674379408 rating on average  
Film-Noir genre: 3.6690720174 rating on average  
War genre: 3.53402101912 rating on average  
Drama genre: 3.44741681831 rating on average  
Western genre: 3.42100145649 rating on average  
Animation genre: 3.41897434097 rating on average  
Mystery genre: 3.3831816398 rating on average  
Romance genre: 3.34458798254 rating on average  
Musical genre: 3.33402514932 rating on average  
Crime genre: 3.30161713817 rating on average  
Adventure genre: 3.22986158087 rating on average  
Comedy genre: 3.19035316365 rating on average  
Fantasy genre: 3.18382718117 rating on average  
Thriller genre: 3.1830728458 rating on average  
Sci-Fi genre: 3.16705222187 rating on average  
Children genre: 3.13867530631 rating on average  
Action genre: 3.09086299473 rating on average  
IMAX genre: 3.05902750997 rating on average  
Horror genre: 2.99193298202 rating on average
```

```
=== Problem 3-3: The names of the top three movies (by average user rating) in each  
genre ===
```

Mystery: Top 3 Movies

1. Mike's Murder (1984) (5.0)
2. Lured (1947) (5.0)
3. Past, The (Le passé) (2013) (5.0)

Drama: Top 3 Movies

1. O Lucky Man! (1973) (5.0)
2. Bleak House (2005) (5.0)
3. One Magic Christmas (1985) (5.0)

Horror: Top 3 Movies

1. Phantom of the Paradise (1974) (5.0)
2. Vampyros Lesbos (Vampirás, Las) (1971) (5.0)
3. Asylum (1972) (5.0)

Film-Noir: Top 3 Movies

1. Lured (1947) (5.0)
2. In a Lonely Place (1950) (4.83333333333)
3. Detour (1945) (4.75)

Romance: Top 3 Movies

1. Touch of Pink (2004) (5.0)
2. Rustom (2016) (5.0)
3. Law of Desire (Ley del deseo, La) (1987) (5.0)

Action: Top 3 Movies

1. American Ninja (1985) (5.0)
2. All-Star Superman (2011) (5.0)
3. Absolute Giganten (1999) (5.0)

Documentary: Top 3 Movies

1. Pervert's Guide to Cinema, The (2006) (5.0)
2. Side by Side (2012) (5.0)
3. Frozen Planet (2011) (5.0)

Thriller: Top 3 Movies

1. Funeral in Berlin (1966) (5.0)
2. Shining Through (1992) (5.0)
3. State of Siege (État de siège) (1972) (5.0)

Sci-Fi: Top 3 Movies

1. Ghost in the Shell 2.0 (2008) (5.0)
2. Holy Motors (2012) (5.0)
3. Phantasm II (1988) (5.0)

Crime: Top 3 Movies

1. New Police Story (Xin jing cha gu shi) (2004) (5.0)
2. Get Carter (1971) (5.0)
3. Manson Family, The (2003) (5.0)

Fantasy: Top 3 Movies

1. Holy Motors (2012) (5.0)
2. 16 Wishes (2010) (5.0)
3. Peter Pan (1960) (5.0)

Western: Top 3 Movies

1. Bone Tomahawk (2015) (5.0)
2. Walker (1987) (5.0)
3. Bite the Bullet (1975) (5.0)

Animation: Top 3 Movies

1. Wolf Children (Okami kodomo no ame to yuki) (2012) (5.0)
2. Disappearance of Haruhi Suzumiya, The (Suzumiya Haruhi no shôshitsu) (2010) (5.0)
3. One Piece Film: Strong World (2009) (5.0)

War: Top 3 Movies

1. Walker (1987) (5.0)
2. Generation Kill (2008) (5.0)
3. Andrei Rublev (Andrey Rublyov) (1969) (5.0)

IMAX: Top 3 Movies

1. Alaska: Spirit of the Wild (1997) (5.0)
2. Mission to Mir (1997) (4.5)
3. Dark Knight, The (2008) (4.23553719008)

Children: Top 3 Movies

1. Ice Age: The Great Egg-Scapade (2016) (5.0)
2. Peter Pan (1960) (5.0)
3. Ice Princess (2005) (5.0)

(no genres listed): Top 3 Movies

1. Survive and Advance (2013) (5.0)
2. Victoria (2015) (5.0)

3. Boy Crazy (2009) (5.0)

Comedy: Top 3 Movies

1. Hotel de Love (1996) (5.0)

2. Kidnapping, Caucasian Style (Kavkazskaya plennitsa) (1967) (5.0)

3. House Bunny, The (2008) (5.0)

Musical: Top 3 Movies

1. Tango (1998) (5.0)

2. Storefront Hitchcock (1997) (5.0)

3. This Is the Army (1943) (5.0)

Adventure: Top 3 Movies

1. Disappearance of Haruhi Suzumiya, The (Suzumiya Haruhi no shôshitsu) (2010) (5.0)

2. Bite the Bullet (1975) (5.0)

3. Shogun Assassin (1980) (5.0)

=== Problem 3-4: Top ten movie watchers ranked by the number of movies they have rated
===

User 547 watched 2391 movies

User 564 watched 1868 movies

User 624 watched 1735 movies

User 15 watched 1700 movies

User 73 watched 1610 movies

User 452 watched 1340 movies

User 468 watched 1291 movies

User 380 watched 1063 movies

User 311 watched 1019 movies

User 30 watched 1011 movies

=== Problem 3-5: Top ten pairs of users ranked by the number of movies both of them
has watched ===

671

Users 15 and 73 have watched 813 movies in common

Users 15 and 547 have watched 796 movies in common

Users 452 and 547 have watched 771 movies in common

Users 15 and 624 have watched 741 movies in common

Users 468 and 547 have watched 678 movies in common

Users 73 and 624 have watched 670 movies in common

Users 547 and 564 have watched 657 movies in common

Users 15 and 468 have watched 644 movies in common

Users 15 and 380 have watched 628 movies in common

Users 73 and 380 have watched 609 movies in common