```
In [1]: #import all the necessary packages.
        from PIL import Image
        import requests
        from io import BytesIO
        import matplotlib.pyplot as plt
        import numpy as np
        import pandas as pd
        import warnings
        from bs4 import BeautifulSoup
        from nltk.corpus import stopwords
        from nltk.tokenize import word tokenize
        import nltk
        import math
        import time
        import re
        import os
        import seaborn as sns
        from collections import Counter
        from sklearn.feature extraction.text import CountVectorizer
        from sklearn.feature extraction.text import TfidfVectorizer
        from sklearn.metrics.pairwise import cosine similarity
        from sklearn.metrics import pairwise distances
        from matplotlib import gridspec
        from scipy.sparse import hstack
        import plotly
        import plotly.figure factory as ff
        from plotly.graph objs import Scatter, Layout
        plotly.offline.init notebook mode(connected=True)
        warnings.filterwarnings("ignore")
```

Here I am loading the data instead of computing again due to my computer's memory and processing constriants

Loading the Image Vectors calculated from CNN

```
In [2]: bottleneck_features_train = np.load('16k_data_cnn_features.npy')
    asins = np.load('16k_data_cnn_feature_asins.npy')
```

Loading the Preprocessed Apparel data

The indices of bottleneck_features_train are not in right order. So, we have to order them. Or this will be problem when we calculate pairwise distances using the inbuilt function and add the weighted sum to our total pairwise distances.

One hot encoding for brand, product_type, color columns

```
In [8]: # some of the brand values are empty.
# Need to replace Null with string "NULL"
data['brand'].fillna(value="Not given", inplace=True )

# replace spaces with hypen
brands = [x.replace(" ", "-") for x in data['brand'].values]
types = [x.replace(" ", "-") for x in data['product_type_name'].values]
colors = [x.replace(" ", "-") for x in data['color'].values]

brand_vectorizer = CountVectorizer()
brand_features = brand_vectorizer.fit_transform(brands)

type_vectorizer = CountVectorizer()
type_features = type_vectorizer.fit_transform(types)

color_vectorizer = CountVectorizer()
color_features = color_vectorizer.fit_transform(colors)
extra_features = hstack((brand_features, type_features, color_features)).tocsr()
```

Idf weighted Word2Vec for title column.

```
In [12]: from tqdm import tqdm
         idf_w2v_titles = []
         for title in tqdm(data['title']):
             sent_vect = np.zeros(300)
             idf sum = 0
             for word in title.split():
                 if word in model.keys() and word in word_idfs.keys():
                     sent vect += (model[word] * word idfs[word])
                     idf_sum += word_idfs[word]
             if idf sum != 0:
                 sent_vect /= idf_sum
             idf_w2v_titles.append(sent_vect)
         idf_w2v_titles = np.array(idf_w2v_titles)
         print(idf_w2v_titles.shape)
         100%|
                                                                                         16042/16042 [00:01<00:00, 1331
         5.39it/s]
         (16042, 300)
In [14]: from IPython.display import display, Image, SVG, Math, YouTubeVideo
```

```
In [75]: def display img(url):
            # we get the url of the apparel and download it
              response = requests.get(url)
              img = Image.open(BytesIO(response.content))
            img = Image(url=url, embed=True)
            # we will display it in notebook
            display(img)
         def get distance(vec1, vec2):
            # vec1 = np.array(#number_of_words_title1 * 300), each row is a vector of length 300 corresponds to each word
            # vec2 = np.array(#number of words title2 * 300), each row is a vector of length 300 corresponds to each word
            final dist = []
            # for each vector in vec1 we caluclate the distance(euclidean) to all vectors in vec2
            for i in vec1:
                dist = []
                for j in vec2:
                     # np.linalq.norm(i-j) will result the euclidean distance between vectors i, j
                    dist.append(np.linalg.norm(i-j))
                 final dist.append(np.array(dist))
            # final dist = np.array(#number of words in title1 * #number of words in title2)
            # final dist[i,j] = euclidean distance between vectors i, j
            return np.array(final dist)
         def heat map w2v brand(sentance1, sentance2, url, doc id1, doc id2, df id1, df id2):
            # sentance1 : title1, input apparel
            # sentance2: title2, recommended apparel
            # url: apparel image url
            # doc id1: document id of input apparel
            # doc id2: document id of recommended apparel
            # df id1: index of document1 in the data frame
            # df id2: index of document2 in the data frame
            # model: it can have two values, 1. avg 2. weighted
            #s1 vec = np.array(#number of words title1 * 300), each row is a vector(weighted/avg) of length 300 correspor
            s1 words = [word for word in sentance1.split() if word in model.keys() and word in word idfs.keys()]
            s1 vec = [(model[word] * word idfs[word]) for word in s1 words]
            #s2 vec = np.array(#number of words title2 * 300), each row is a vector(weighted/avg) of length 300 correspor
            s2 words = [word for word in sentance2.split() if word in model.keys() and word in word idfs.keys()]
            s2 vec = [(model[word] * word idfs[word]) for word in s2 words]
```

```
# s1 s2 dist = np.array(#number of words in title1 * #number of words in title2)
# s1 s2 dist[i,j] = euclidean distance between words i, j
s1 s2 dist = get distance(s1 vec, s2 vec)
data matrix = [['Asin', 'Brand', 'Color', 'Product type'],
           [data['asin'].loc[df id1],brands[doc id1], colors[doc id1], types[doc id1]], # input apparel's fed
           [data['asin'].loc[df id2],brands[doc id2], colors[doc id2], types[doc id2]]] # recommonded appare
colorscale = [[0, '#1d004d'],[.5, '#f2e5ff'],[1, '#f2e5d1']] # to color the headings of each column
# we create a table with the data matrix
table = ff.create table(data matrix, index=True, colorscale=colorscale)
# plot it with plotly
plotly.offline.iplot(table, filename='simple table')
# devide whole figure space into 25 * 1:10 grids
  gs = gridspec.GridSpec(25, 15)
fig = plt.figure(figsize=(25,5))
# in first 25*10 grids we plot heatmap
 ax1 = plt.subplot(qs[:, :-5])
# ploting the heap map based on the pairwise distances
ax1 = sns.heatmap(np.round(s1 s2 dist,3), annot=True)
# set the x axis labels as recommended apparels title
ax1.set xticklabels(s2 words)
# set the y axis labels as input apparels title
ax1.set yticklabels(s1_words)
# set title as recommended apparels title
ax1.set title(sentance2)
plt.show()
# in last 25 * 10:15 grids we display image
 ax2 = plt.subplot(qs[:, 10:16])
 # we dont display grid lins and axis labels to images
  ax2.grid(False)
 ax2.set xticks([])
 ax2.set yticks([])
# pass the url it display it
display img(url)
```

```
In [76]: def idf w2v brand(doc id, w1, w2, w3, num results):
             # doc id: apparel's id in given corpus
             # w1: weight for w2v features
             # w2: weight for brand and color features
             # pairwise dist will store the distance from given input apparel to all remaining apparels
             # the metric we used here is cosine, the coside distance is mesured as K(X, Y) = \langle X, Y \rangle / (||X||^*||Y||)
             # http://scikit-learn.org/stable/modules/metrics.html#cosine-similarity
             idf w2v dist = pairwise distances(idf w2v titles, idf w2v titles[doc id].reshape(1,-1))
             ex feat dist = pairwise distances(extra features, extra features[doc id])
              image dist = pairwise distances(ordered image vect, ordered image vect[doc id].reshape(1,-1))
             pairwise dist = (w1 * idf w2v dist + w2 * ex feat dist + w3 * image dist)/float(w1 + w2 + w3)
             # np.argsort will return indices of 9 smallest distances
             indices = np.argsort(pairwise dist.flatten())[0:num results]
              #pdists will store the 9 smallest distances
              pdists = np.sort(pairwise dist.flatten())[0:num results]
             #data frame indices of the 9 smallest distace's
             df indices = list(data.index[indices])
             for i in range(0, len(indices)):
                   display(Image(url=data['medium image url'].loc[df indices[i]], embed=True))
                 heat map w2v brand(data['title'].loc[df indices[0]],data['title'].loc[df indices[i]], data['medium image
                 print('ASIN :',data['asin'].loc[df indices[i]])
                 print('Brand :',data['brand'].loc[df indices[i]])
                 print('euclidean distance from total data (including image) :', pdists[i])
                 print('='*125)
```

Let us give equal weight to title, extra features and image features and see our results first.

In [77]: # arguments for idf_w2v_brand function are row_number, title_weight, extra_feats_weight, image_feats_weight, num
idf_w2v_brand(12566, 5, 5, 5)

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQB5FQ	Si-Row	Brown	TOYS AND GAMES

				burnt umber tiger tshi	rt zebra stripes xl xxl					
Ħ.	0	35	35	30	36	34	30	29	_	- 32
ber b	35	0	33		30					-
tiger um	35	33	0	27		32				- 24
Shirt 6	30		27	0	26		11	8.9		
zebra ts	36	30		26	0	28	26	25	-	16
			32	21	28	0				
XIstripes	30		28	11	26	22	0	8.1	-	- 8
×	. 29			8.9	25		8.1	0		_
	burnt	umber	tiger	tshirt	zebra	stripes	x	xxl	_	0



ASIN : B00JXQB5FQ Brand : Si Row

euclidean distance from total data (including image) : 1.4216537238098681e-06

Acin	Buand	Color	Dradust tuno
Asin	Brand	Color	Product type

B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQASS6	Si-Row	Pink	TOYS AND GAME!

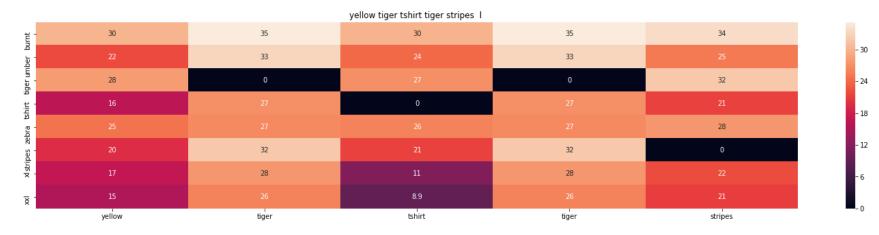
			pink t	tiger tshirt zebra stripes xl	xxl			
purut	29	35	30	36	34	30	29	
umber bu		33		30			22	
tiger um		0			32		26	
tshirt ti	11		0	26		11	8.9	
zebra ts			26	0			25	
	20	32		28	0	22	21	
xlstripes	12	28	11	26	22	0	8.1	
<u> </u>	9.9		8.9			8.1	0	
	pink	tiger	tshirt	zebra	stripes	xl	xxl	ı



ASIN: B00JXQASS6 Brand: Si Row

euclidean distance from total data (including image) : 10.747429163661844

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JXQCUIC	Si-Row	Yellow	TOYS AND GAME!





ASIN : B00JXQCUIC Brand : Si Row

euclidean distance from total data (including image) : 14.664675581507474

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQCWTO	Si-Row	Brown	TOYS AND GAME!

				brown white tiger tsh	irt tiger stripes xl xxl			
burnt	29	28	35	30	35	34	30	29
umber b			33		33			22
tiger um	28		0		0	32	28	26
shirt ti	16	11		0			11	8.9
zebra ts						28		25
			32		32	0		21
xlstripes	16	11	28	11	28		0	8.1
× -	15	8.3		8.9			8.1	0
	brown	white	tiger	tshirt	tiger	stripes	xl	xxI



ASIN : B00JXQCWTO Brand : Si Row

euclidean distance from total data (including image) : 14.972075478957672

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JV63QQE	Si-Row	Red	TOYS AND GAME!

red	pink floral	heel sleeveless	shirt xl	xxl
-----	-------------	-----------------	----------	-----

purut -			31	39	31			29
			24	34	25			22
tiger umber			29	36	29			26
tshirt ti	13	11	16		15	8.9	11	8.9
zebra ts				33				25
	19	20		30		20		21
xlstripes	13	12	16		16	12	0	8.1
×.	11	9.9	14		14	9.2	8.1	0
	red	pink	floral	heel	sleeveless	shirt	×	xxl

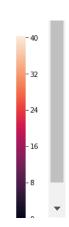


ASIN: B00JV63QQE Brand: Si Row

euclidean distance from total data (including image) : 17.035315252858016

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01CUPYBM0	Styleco	White	SHIRT

	edv cheetah run purple multi xl								
burnt		40	31		33	30			
nber bu		38				22			
tiger um						28			
shirt 6	10	32	20	14		11			
ebra ts						26			
es a		37	25			22			
xlstrip	9.8	33	20	15	17	0			
× .	5.9	31	18	13	16	8.1			



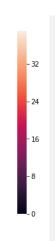


ASIN : B01CUPYBM0 Brand : Styleco

euclidean distance from total data (including image) : 17.354493654190243

Asin		Brand	Color	Product type
B00JXQB5I	EQ.	Si-Row	Brown	TOYS_AND_GAME:
B00JV63V0	8	Si-Row	Purple	TOYS AND GAME!

			purple	floral heel sleeveless shirt	xl xxl		
purut		31	39	31			29
umber b			34				22
tiger um			36				26
tshirt ti	14	16		15	8.9	11	8.9
zebra ts			33				25
	19		30		20		21
xlstripes	15	16		16	12	0	8.1
× -	13	14		14	9.2	8.1	0
	purple	floral	heel	sleeveless	shirt	xl	xxl



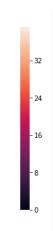


ASIN: B00JV63VC8 Brand: Si Row

euclidean distance from total data (including image) : 17.360399161602984

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01LX6H43D	ВОВОВ	Black	BOOKS 1973 AND

				moana logo adults hot	t v neck shirt black xxl			
purut -	32	33	39	31				29
			34					22
tiger umber			36					26
tshirt ti		16			14	8.9	11	8.9
zebra ts			36					25
		22	35			20	20	21
xlstripes	19	18		19	14	12	11	8.1
lx -	17	17		17	12	9.2	8.3	0
	moana	logo	adults	hot	neck	shirt	black	xxl





ASIN: B01LX6H43D Brand: B0B0B

euclidean distance from total data (including image) : 17.418471723870393

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B071FCWD97	Namnoi-Clothing-Store	As-Show	SHIRT







ASIN: B071FCWD97

Brand: Namnoi Clothing Store

euclidean distance from total data (including image) : 17.425361012427164

==========

Asin Brand	Color	Product type
------------	-------	--------------

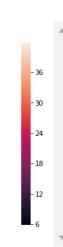
- 30

- 25

- 15

B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01CR57YY0	ABADAY	Multi	BOOKS 1973 AND

				abaday multico	lor cartoon cat pri	nt short sleeve long	gline shirt large			
burnt		35					29	42	30	29
		24			24	24	24		24	23
tiger umber							27	35		27
tshirt ti	10	20	20	21	13	13	11		8.9	13
zebra ts							26	35	26	26
	21						20		20	22
xlstripes	9.8	20		22	14	13	12		12	12
× -	5.9	19	20	20	11	10	9.8		9.2	9.3
	abaday	multicolor	cartoon	cat	print	short	sleeve	longline	shirt	large





ASIN : B01CR57YY0 Brand : ABADAY

euclidean distance from total data (including image) : 17.430129060318404

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B010NN9RXO	YICHUN	Multicoloured	SHIRT

	yichun womens tiger printed summer tshirts tops								
print	29	29	35	29	30	32	30		
ber bu			33	24			24		
tiger um		26	0	27	29	29	28		
shirt ti	10	10		13	15	11	13		
zebra ts					28	27	27		
		21	32	22			22		
Xistripes	9.8	9.9	28	14	15	14	13		
XX -	5.9	6.4		12	12	12	11		
	yichun	womens	tiger	printed	summer	tshirts	tops		



ASIN : B010NN9RXO Brand : YICHUN

euclidean distance from total data (including image) : 17.44386532744975

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01JUNHBRM	Namnoi-Clothing-Store	As-show	SHIRT

womens thin style tops tees pastel watermelon print

			****	mens unin style tops tee	s paster watermeron p	iiic		
ţ.		35	30		35	33	37	29
ber by		31		24		24	31	24
iger um		33			31	32	34	26
tshirt ti	10	25	15	13	19	20		13
bra ts	26	34	27	27	30		32	26
es ze	21	30	22	22		24	31	22
×lstrip	9.9	26	15	13	21	21	27	14



ASIN : B01JUNHBRM

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 17.47822748164476

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B01F7PHXY8	Danskin-Now	Pink-Hombre	SHIRT

			da	anskin womens vneck	loose performance	tee xsmall pink omb	re			
prt					32	31	29		38	- 36
nber b	23	23	23				23		28	
tiger um									35	- 30
shirt ti	10	10	10	16	17	15	10	11	24	- 24
zebra ts								24	31	10
10			21			24		20	28	- 18
×Istripes	9.8	9.9	9.8	16	16	16	9.8	12	24	- 12
XX -	5.9	6.4	5.9	13	14	14	5.9	9.9	23	-6
_	danskin	womens	vneck	loose	performance	tee	xsmall	pink	ombre	— -6



ASIN: B01F7PHXY8
Brand: Danskin Now

euclidean distance from total data (including image) : 17.51399213379263

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B071SBCY9W	Namnoi-Clothing-Store	As-Show	SHIRT

kawaii pastel tops tees baby blue flower design

				pp		•		
Ę-	34	33	30	35	34		33	33
ber bu		24	24		31	21		26
ger um	30	32		31			30	30
hirt ti	20	20	13	19	21	11	21	17
bra ts	30			30				28
뭐								





ASIN: B071SBCY9W

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 17.567494069971524

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B015H3W9BM	KINGDE	Pink	SHIRT

kingde star pink flower dog stamp sleeveless vestbqn24

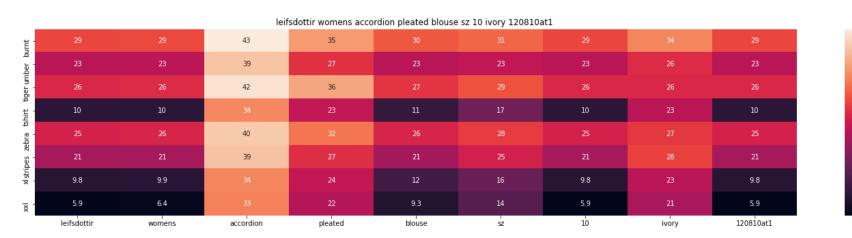
Ę.		33		33	35	38	31	29
er b	23	28	21	26	31	34	25	23



ASIN: B015H3W9BM Brand: KINGDE

euclidean distance from total data (including image) : 17.583547871985036

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01B6E0DX2	Leifsdottir	lvory	SHIRT



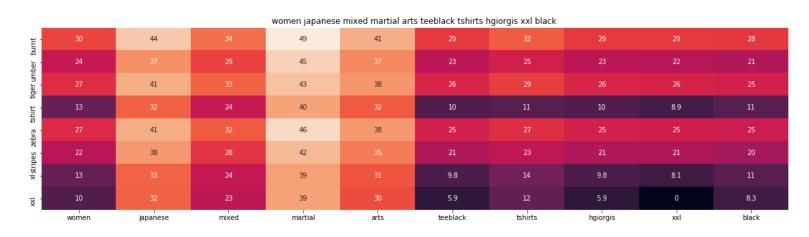


ASIN: B01B6E0DX2
Brand: Leifsdottir

euclidean distance from total data (including image) : 17.58946914406769

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01JQ096HW	Constanza	Black	BOOKS 1973 AND



- 40

- 30

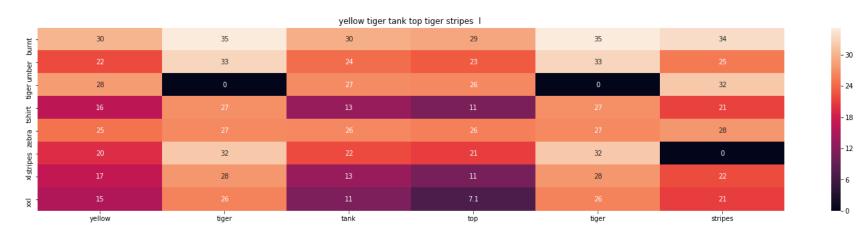
- 20



ASIN: B01JQ096HW Brand: Constanza

euclidean distance from total data (including image) : 17.61472419434818

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JXQAUWA	Si-Row	Yellow	TOYS AND GAMES





ASIN : B00JXQAUWA Brand : Si Row

euclidean distance from total data (including image) : 17.63771451979729

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01I80A93G	Namnoi-Clothing-Store	As-show	SHIRT



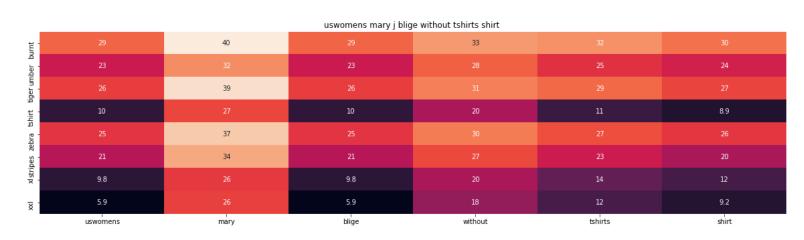


ASIN : B01I80A93G

Brand: Namnoi Clothing Store

euclidean distance from total data (including image) : 17.64385550266359

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME(
B01M0XXFKK	South-Dakota-A.	Black	BOOKS 1973 AND





ASIN : B01M0XXFKK

Brand : South Dakota A.

euclidean distance from total data (including image) : 17.661228759041066

==========

From above result we can see some new shirts that appeared which are not there in previous models (like MOANA shirt). This might be due to image of the shirt as the shape of the shirt is similar but I am not quite sure why the mdoel picked those shirts.

Now we try giving extra features more weight (5 times than other features) and see the results.

In [78]: idf_w2v_brand(12566, 5, 25, 5, 20)

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQB5FQ	Si-Row	Brown	TOYS AND GAME!

_				burnt umber tiger tshi	rt zebra stripes xl xxl			
purut	0	35	35	30	36	34	30	29
umber b	35	0	33		30			22
tiger um	35	33	0			32		26
tshirt ti	30		27	0			11	8.9
zebra ts	36	30			0			25
	34		32		28	0		21
xlstripes	30		28	11		22	0	8.1
xx	29			8.9			8.1	0
	burnt	umber	tiger	tshirt	zebra	stripes	хİ	xxl



ASIN : B00JXQB5FQ Brand : Si Row

euclidean distance from total data (including image) : 6.092801673470863e-07

Asin Brand Color Produc	type
-------------------------	------

B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQASS6	Si-Row	Pink	TOYS AND GAME!

			pink	tiger tshirt zebra stripes x	xxl			
burnt	29	35	30	36	34	30	29	- 32
umber b		33	24	30				
tiger um		0	27		32			- 24
tshirt ti	11	27	0	26	21	11	8.9	
zebra ts				0				- 16
	20	32	21	28	0	22		
xlstripes	12	28	11	26		0	8.1	- 8
lo -	9.9	26	8.9	25		8.1	0	
	pink	tiger	tshirt	zebra	stripes	×	xxl	— -0



ASIN : B00JXQASS6 Brand : Si Row

euclidean distance from total data (including image) : 5.414163105782559

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JXQCWTO	Si-Row	Brown	TOYS AND GAMES

				brown white tiger tsh	irt tiger stripes xl xxl			
pnrut	29	28	35	30	35	34	30	29
umber b			33		33			22
tiger um	28		0		0	32	28	26
tshirt ti	16	11	27	0	27		11	8.9
zebra ts						28		25
			32		32	0		21
xlstripes	16	11	28	11	28		0	8.1
× -	15	8.3	26	8.9			8.1	0
	brown	white	tiger	tshirt	tiger	stripes	xl	xxl



ASIN : B00JXQCWTO Brand : Si Row

euclidean distance from total data (including image) : 6.416603776696145

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQCUIC	Si-Row	Yellow	TOYS AND GAME!

			yellow tiger tshirt tiger stripes				
Ħ.	30	35	30	35	34		
iber b	- 22	33	24	33			30
tiger um	. 28	0	27	0	32	-	2
5hirt t	16		0	27	21		- 18
zebra t	25		26		28		
S.	20	32	21	32	0	-	- 12
xlstrip.	17	28	11	28		-	6
×.	15		8.9	26	21		٨
	yellow	tiger	tshirt	tiger	stripes		U



ASIN : B00JXQCUIC Brand : Si Row

euclidean distance from total data (including image) : 7.092982999144971

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JV63QQE	Si-Row	Red	TOYS AND GAME!

red	pink floral	heel sleeveless	shirt xl	xxl
-----	-------------	-----------------	----------	-----

				rea pink noral neers	ice veless simile XI XXI			
burnt			31	39	31			29
nber bu				34				22
tiger um				36				26
Shirt ti	13	11	16		15	8.9	11	8.9
zebra ts				33				25
	19	20		30	22	20		21
xlstripes	13	12	16		16	12	0	8.1
<u>×</u> -	11	9.9	14		14	9.2	8.1	0
	red	pink	floral	heel	sleeveless	shirt	x	xxl

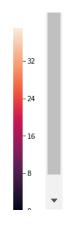


ASIN: B00JV63QQE Brand: Si Row

euclidean distance from total data (including image) : 8.108971429723777

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JV63VC8	Si-Row	Purple	TOYS AND GAME!

			purple	floral heel sleeveless shirt	xl xxl		
purut		31	39	31			29
umber b			34				22
tiger um			36				26
tshirt ti	14	16		15	8.9	11	8.9
zebra ts			33				25
	19		30		20		21
xlstripes	15	16		16	12	0	8.1
lx -	13	14		14	9.2	8.1	0



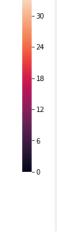


ASIN: B00JV63VC8 Brand: Si Row

euclidean distance from total data (including image) : 8.24829310490019

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQAUWA	Si-Row	Yellow	TOYS AND GAME!

			yellow tiger tank	top tiger stripes		
burnt	30	35	30	29	35	34
mber b		33			33	25
tiger um	28	0				32
tshirt ti			13	11		21
zebra ts						28
		32			32	0
xlstripes	17	28	13	11	28	22
lo .			11	7.1		21
	yellow	tiger	tank	top	tiger	stripes





ASIN: B00JXQAUWA Brand: Si Row

euclidean distance from total data (including image) : 8.367142544126322

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JXQAO94	Si-Row	White	TOYS AND GAMES



			black v	vhite tiger tank top tiger st	ripes I		
purut -	28	28	35	30	29	35	34
		21	33			33	25
tiger umber		26				0	32
tshirt ti	11	11		13	11		21
zebra ts							28
		19	32			32	0
xlstripes	11	11	28	13	11	28	22
lx -	8.3	8.3		11	7.1		21
	black	white	tiger	tank	top	tiger	stripes

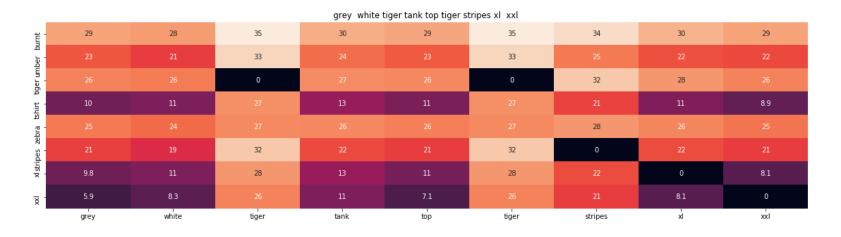


ASIN: B00JXQA094 Brand: Si Row

euclidean distance from total data (including image) : 8.383451211390259

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!

	B00JXQAFZ2	Si-Row	Grev	TOYS AND GAMES
--	------------	--------	------	----------------





ASIN : B00JXQAFZ2 Brand : Si Row

euclidean distance from total data (including image) : 8.444773165977129

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JV63CW2	Si-Row	Red	TOYS AND GAME!

- 30

red butterfly black white tank top xl xxl

ig -		37	28		30	29	30	29
nber b		30	21					
tiger um		31	25					
tshirt ti	13	25	11	11	13	11	11	8.9
zebra ts		30	25					
ss -		30	20					
xlstrip6	13	25	11	11	13	11	0	8.1
× -	11	24	8.3	8.3	11	7.1	8.1	0
	red	butterfly	black	white	tank	top	x	xxl



ASIN: B00JV63CW2 Brand: Si Row

euclidean distance from total data (including image) : 8.505606095943415

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQABB0	Si-Row	Red	TOYS AND GAME!

				red rose bud heel da	almatian tshirt xl xxl			
print -	28		42	39	38	30		29
umber bu	21		38	34	33	24		22
tiger um			40	36				26
tshirt ti	13		34			0	11	8.9
zebra ts	24		39	33				25
	19		39			21		21
xlstripes	13		34			11	0	8.1
× -	11	20	33			8.9	8.1	0
_	red	rose	bud	heel	dalmatian	tshirt	x	xxl



ASIN : B00JXQABB0 Brand : Si Row

euclidean distance from total data (including image) : 8.536867653352724

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQAX2C	Si-Row	Pink	TOYS AND GAME!

pink rose bud heel dalmatian tank top xl xxl

				p	a meer aanmaciam car				
ţ.			42	39	38	30			29
ber bu		28	38	34	33	24			22
iger um			40	36					26
ij.	11		34			13	11	11	8.9
bra tsh	24	31	39	33					25
es -	20	28	39	30	32	22	21	22	21

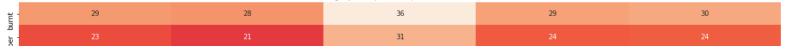


ASIN : B00JXQAX2C Brand : Si Row

euclidean distance from total data (including image) : 8.538338808806833

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00 XQCFRS	Si-Row	Grey	TOYS AND GAME!

grey red peacock print tshirt I







ASIN : B00JXQCFRS Brand : Si Row

euclidean distance from total data (including image) : 8.766389416144886

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQC0C8	Si-Row	Blue	TOYS AND GAME!

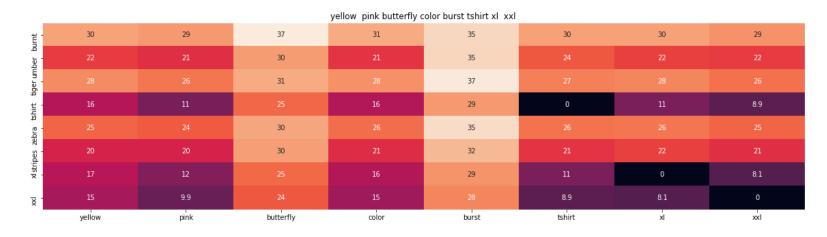
blue green butterfly color burst tshirt I



ASIN : B00JXQC0C8
Brand : Si Row

euclidean distance from total data (including image) : 8.778242366258077

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQBBMI	Si-Row	Yellow	TOYS AND GAME!





ASIN : B00JXQBBMI Brand : Si Row

euclidean distance from total data (including image) : 8.832777899079192

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01CUPYBM0	Styleco	White	SHIRT



- 32



ASIN : B01CUPYBM0 Brand : Styleco

euclidean distance from total data (including image) : 8.949498029547014

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME(
B01LX6H43D	вовов	Black	BOOKS 1973 AND

				moana logo adults ho	t v neck shirt black xxl			
burnt	32	33	39	31				29
		27	34					22
tiger umber		30	36					26
tshirt ti		16			14	8.9	11	8.9
zebra ts		28	36					25
		22	35			20	20	21
xistripes	19	18			14	12	11	8.1
× -	17	17		17	12	9.2	8.3	0
	moana	logo	adults	hot	neck	shirt	black	xxl



ASIN: B01LX6H43D Brand: B0B0B

euclidean distance from total data (including image) : 8.976917202267078

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01CR57YY0	ABADAY	Multi	BOOKS 1973 AND

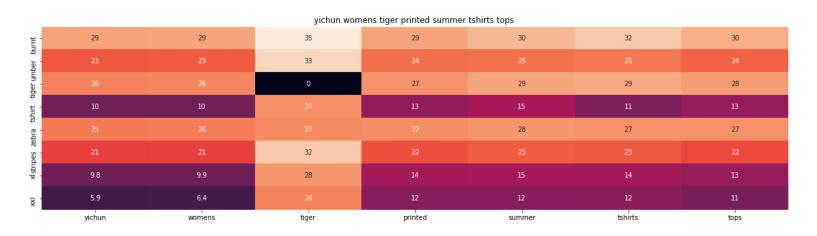
				abaday multico	lor cartoon cat pri	nt short sleeve lon	gline shirt large			
purut		35					29	42	30	29
	23	24			24	24	24		24	23
tiger umber				23				35	27	27
tshirt ti	10	20	20	21	13	13	11		8.9	13
zebra ts							26	35	26	26
	21				22	23	20		20	22
xlstripes	9.8	20		22	14	13	12		12	12
× -	5.9	19	20	20	11	10	9.8		9.2	9.3
	abaday	multicolor	cartoon	cat	print	short	sleeve	longline	shirt	large

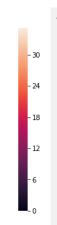




ASIN : B01CR57YY0 Brand : ABADAY euclidean distance from total data (including image) : 8.98191320360194

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B010NN9RXO	YICHUN	Multicoloured	SHIRT



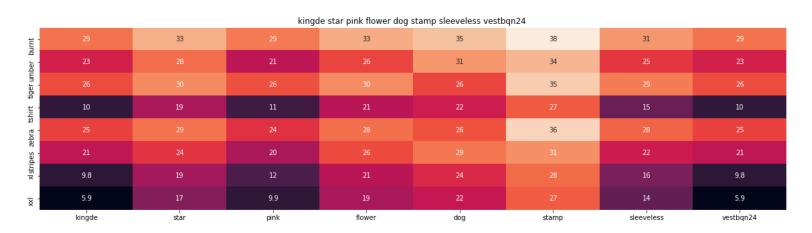




ASIN: B010NN9RXO Brand: YICHUN

euclidean distance from total data (including image) : 8.987800175229658

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B015H3W9BM	KINGDE	Pink	SHIRT





ASIN: B015H3W9BM Brand: KINGDE

euclidean distance from total data (including image): 9.047664122887639

=========

We can see that most of them are of same brand (which we already saw in previous models) and some shirts appeared even if they dont have same extra featured which might be due to high co-relation between images and title.

Now we try giving image features more weight (5 times than other features) and see the results.

In [79]: idf_w2v_brand(12566, 5, 5, 25, 20)

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JXQB5FQ	Si-Row	Brown	TOYS AND GAME!

_				burnt umber tiger tshi	rt zebra stripes xl xxl				
burnt	0	35	35	30	36	34	30	29	
umber b	35	0	33		30			22	
tiger um	35	33	0			32		26	
shirt ti	30		27	0	26		11	8.9	
zebra ts	36	30			0	28		25	
	34		32		28	0		21	
xlstripes	30		28	11	26	22	0	8.1	
<u> </u>	29			8.9	25		8.1	0	
	burnt	umber	tiger	tshirt	zebra	stripes	x	xxl	



ASIN : B00JXQB5FQ Brand : Si Row

euclidean distance from total data (including image) : 3.0464008367354315e-06

Asin Brand Color Produc	type
-------------------------	------

B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQASS6	Si-Row	Pink	TOYS AND GAMES

			pink	tiger tshirt zebra stripes x	xxl			
burnt	29	35	30	36	34	30	29	- 32
umber b		33	24	30				
tiger um		0	27		32			- 24
tshirt ti	11	27	0	26	21	11	8.9	
zebra ts				0				- 16
	20	32	21	28	0	22		
xlstripes	12	28	11	26		0	8.1	- 8
lo -	9.9	26	8.9	25		8.1	0	
	pink	tiger	tshirt	zebra	stripes	×	xxl	— -0



ASIN : B00JXQASS6 Brand : Si Row

euclidean distance from total data (including image) : 21.777567297819363

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B00JXQCUIC	Si-Row	Yellow	TOYS AND GAMES

			yellow tiger tshirt tiger stripes			
burnt	30	35	30	35	34	
nber b		33	24	33	25	- 30
tiger um	28	0	27	0	32	- 24
tshirt	16	27	0	27	21	- 18
zebra t	25		26		28	
10.		32	21	32	0	- 12
xlstripes	17	28	11	28		- 6
×	15	26	8.9	26	21	-0
	yellow	tiger	tshirt	tiger	stripes	- 0



ASIN : B00JXQCUIC Brand : Si Row

euclidean distance from total data (including image) : 29.862643294697396

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B00JXQCWTO	Si-Row	Brown	TOYS AND GAME!

				brown white tiger tsh	irt tiger stripes xl xxl			
prut -	29	28	35	30	35	34	30	29
umber b		21	33		33			22
tiger um	28	26	0	27	0	32	28	26
shirt ti	16	11		0			11	8.9
zebra ts						28		25
		19	32	21	32	0		21
xlstripes	16	11	28	11	28		0	8.1
× -	15	8.3		8.9			8.1	0
	brown	white	tiger	tshirt	tiger	stripes	×	xxI



ASIN : B00JXQCWTO Brand : Si Row

euclidean distance from total data (including image) : 31.559550292460653

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B071FCWD97	Namnoi-Clothing-Store	As-Show	SHIRT

kawaii pastel tops tees pink flower design								
print -	34	33	30	35	29	33	33	
nber bi		24	24		21		26	
tiger um	30	32		31		30	30	
	20	20	13	19	11	21	17	
Zebra -	30			30	24		28	
		24	22		20		24	
Xistripes	22	21	13	21	12	21	17	
- XX	20	19	11	20	9.9	19	15	
	kawaii	pastel	tops	tees	pink	flower	design	



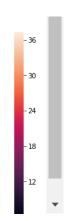
ASIN: B071FCWD97

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 34.54371507466186

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01JUNHBRM	Namnoi-Clothing-Store	As-show	SHIRT

			WO	mens thin style tops tee	es pastel watermelon p	print		
umber burnt		35	30		35	33	37	29
		31					31	24
tiger um		33			31	32	34	
tshirt ti	10	25	15	13		20		13
zebra ts		34					32	
		30	22	22			31	
xlstripes	9.9	26	15	13	21	21		14
× XX	6.4	24	12	11	20	19	26	11





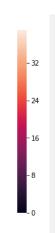
ASIN : B01JUNHBRM

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 34.75833641861227

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B00JV63QQE	Si-Row	Red	TOYS AND GAME!

				red pink floral heel s	leeveless shirt xl xxl			
print			31	39	31			29
				34	25			22
tiger umber				36				26
tshirt ti	13	11	16		15	8.9	11	8.9
zebra ts				33				25
	19	20		30	22	20		21
xlstripes	13	12	16		16	12		8.1
lx -	11	9.9	14		14	9.2	8.1	0
	red	pink	floral	heel	sleeveless	shirt	x	xxI





ASIN : B00JV63QQE Brand : Si Row

euclidean distance from total data (including image) : 34.82624830242464

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01CUPYBM0	Styleco	White	SHIRT

			edv cheetah rur	n purple multi xl		
burnt		40	31		33	30
tiger umber bu		38		21		22
						28
tshirt ti	10	32	20	14		11
zebra ts						26
		37				22
xlstripes	9.8	33	20	15	17	0
× -	5.9	31	18	13	16	8.1
	edv	cheetah	run	purple	multi	×l

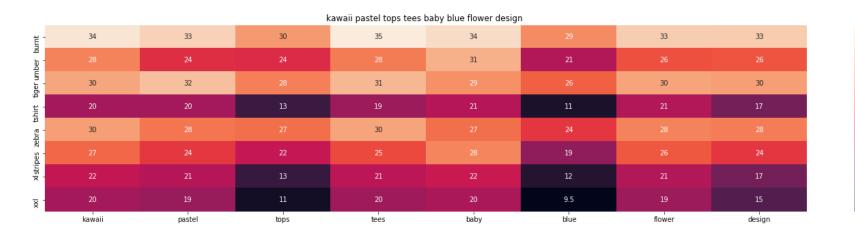


ASIN : B01CUPYBM0 Brand : Styleco

euclidean distance from total data (including image) : 34.89277450204693

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!

DOZ4CDCVOW	Name of Clashina Chana	A = Cl=	CLUDT
B071SBCY9W	Namnoi-Clothing-Store	As-Show	SHIRT





ASIN: B071SBCY9W

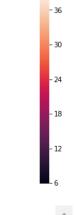
Brand: Namnoi Clothing Store

euclidean distance from total data (including image) : 34.90153170800119

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01F7PHXY8	Danskin-Now	Pink-Hombre	SHIRT

danskin womens vneck loose performance tee xsmall pink ombre

	danskin womens vneck loose performance tee xsmall plink ombre								
print					32	31		29	38
mber b	23	23	23				23		
tiger um								26	35
shirt ti	10	10	10	16	17	15	10	11	24
zebra ts								24	31
ss -			21			24		20	
XIstripe	9.8	9.9	9.8	16	16	16	9.8	12	24
<u>~</u> -	5.9	6.4	5.9	13	14	14	5.9	9.9	
_	danskin	womens	vneck	loose	performance	tee	xsmall	pink	ombre





ASIN: B01F7PHXY8
Brand: Danskin Now

euclidean distance from total data (including image) : 34.992759847239256

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01LX6H43D	ВОВОВ	Black	BOOKS 1973 AND

	moana logo adults hot v neck shirt black xxl							
purut	32	33	39	31				29
mber bu			34					22
tiger um			36					26
tshirt ti		16		19	14	8.9	11	8.9
zebra ts			36					25
			35			20	20	21
xistripes	19	18			14	12	11	8.1
X -	17	17		17	12	9.2	8.3	0
	moana	logo	adults	hot	neck	shirt	black	xxI



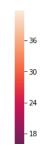
ASIN: B01LX6H43D Brand: B0B0B

euclidean distance from total data (including image) : 35.04036622987303

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B01CR57YY0	ABADAY	Multi	BOOKS 1973 AND

abaday multicolor cartoon cat print short sleeve longline shirt large

				abaday multico	or cartoon cat prin	t short sleeve lon	gille sillic large			
Ħ.		35						42	30	29
ber bu	23	24			24	24	24		24	23
ger um				23				35	27	27
hirt ti	10	20	20	21	13	13	11		8.9	13
bra ts				25	26			35	26	26
es -	21	25	26	28	22	23	20	34	20	22





ASIN : B01CR57YY0 Brand : ABADAY

euclidean distance from total data (including image) : 35.05044558779271

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01I80A93G	Namnoi-Clothing-Store	As-show	SHIRT

summer alpaca 3d pastel casual loose tops tee design

ŧ.	30	48	39			30	30		33
br br	25	43	31	24	24	25	24	25	26





ASIN : B01I80A93G

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 35.058148068384845

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B0177DM70S	MISS-CHIEVOUS/AROUND-THE-WORLD-A	PP Black	SHIRT

miss chievous juniors striped peplum tank top medium shadowpeach

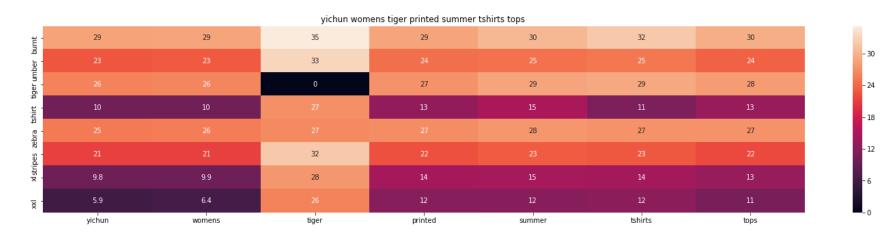


ASIN : B0177DM70S

Brand: MISS CHIEVOUS/AROUND THE WORLD APP

euclidean distance from total data (including image) : 35.08496467788137

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B010NN9RXO	YICHUN	Multicoloured	SHIRT





ASIN : B010NN9RXO Brand : YICHUN

euclidean distance from total data (including image) : 35.16128817814253

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAMES
B071WYLBZS	Namnoi-Clothing-Store	As-Show	SHIRT



- 35

- 30

- 25

- 20

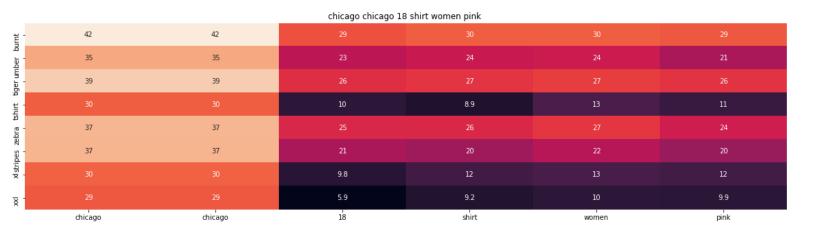


ASIN : B071WYLBZS

Brand : Namnoi Clothing Store

euclidean distance from total data (including image) : 35.22567998799676

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01GXAZTRY	Tony-Arden	Pink	BOOKS 1973 AND





ASIN: B01GXAZTRY Brand: Tony Arden

euclidean distance from total data (including image) : 35.25919432917166

==========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01B6E0DX2	Leifsdottir	lvory	SHIRT





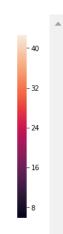
ASIN : B01B6E0DX2 Brand : Leifsdottir - 32

- 16

euclidean distance from total data (including image) : 35.323003604293525

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME:
B01MPX6IDX	Nancy-Lopez	White/Black/Lemon-Drop	SHIRT





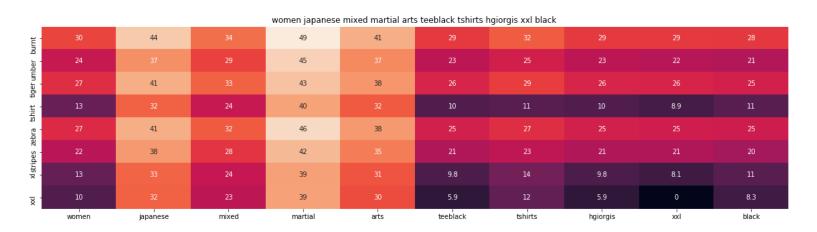


ASIN : B01MPX6IDX Brand : Nancy Lopez

euclidean distance from total data (including image) : 35.32758761946154

=========

Asin	Brand	Color	Product type
B00JXQB5FQ	Si-Row	Brown	TOYS_AND_GAME!
B01JQ096HW	Constanza	Black	BOOKS 1973 AND



- 30

- 20



ASIN: B01JQ096HW Brand: Constanza

euclidean distance from total data (including image) : 35.33444060059118

==========

This model didnt do well (as for my observation) as lot of shirts are not found to be co-related to the main shirt.

let us do analysis for one more shirt and stop as the file gets lenthy.

In [80]: idf_w2v_brand(931, 5, 5, 5, 10)

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT

_				ā	nnakaci sm fit	blue green polk	a dot tie front r	uffle trim blous	е			
. <u>.</u>	0	17	13	8.9	12	27	19	15	11	20	17	9.3
maka	17	0	21	18		32	23			26		
fitan	13	21	0	15	17	30						16
plue -	8.9			0	11	28	17		12		17	11
en bl	12	20	17	11	0	29	20				18	14
- gre	27	32	30	28	29	0	33	30	29	33	31	28
dotpolkar		23		17		33	0	24		26		
tie d		24	18			30	24	0	17			16
벋	11	20		12		29	20	17	0			14
flefro		26				33	26		21	0	22	
trimufflefront	17	23		17		31					0	
se tri	9.3	19	16	11	14	28	21	16	14		18	0
plour	annakaci	sm	fit	blue	green	poľka	dot	tie	front	ruffle	trim	blouse



ASIN : B00KLHUIBS Brand : Anna-Kaci

euclidean distance from total data (including image) : 1.5573436636865759e-06

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT

B0184RPZNM	Eternal-Child	Black,-White,-Lime	SHIRT

				etern	al child womens	black floral prin	nted vneck sho	rt sleeve tunic s	sz new				
ij -	29		4.9		14	12		8.5	9.2	16	16	8.5	
smakaci	33	30	17	18		20	17	20				18	- 32
- fitan	31	29	14	14		17	13				20	14	
plue -	29		9.6	7.7	15	13	8.9	12	11		16	11	
en bl	30		13	12	16	15	12	14	14	18	20	13	- 24
kagre	37	36	27		28	29				30	30		
dopolką	32	32	20	19		20	19						- 16
ë -	32	30	15					17	16			17	
늗 -	30		11	12	17	14	11	13	14			13	
timufflefront	35	33	20				20					20	- 8
<u>_</u>	34	30	17	16	20		17	18	18			17	
- N	30		9.4	10	14	13	9.3	12	11	13	16	13	- 0
plour	eternal	child	womens	black	floral	printed	vneck	short	sleeve	tunic	52	new	-0



ASIN : B0184RPZNM Brand : Eternal Child

euclidean distance from total data (including image) : 17.20011996587542

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B00KOBQEBO	Anna-Kaci	Pink	SHIRT

_					annakaci sm	fit pink bow ril	bbon back trim	racer detail s	hell style top				
ig -		17	13	10		26	9.6	17	24	18	18	12	4.4
arsnakaci	17	0	21			30	20		29				18
- Ift	13	21	0	16		29	15		26			16	14
plue -	8.9	18	15	8.3			12	17	24			14	9.5
등 1	12		17	12			14		27			16	13
e gre		32	30	28	32	35	28	31	33	33	33	28	27
dopolka						28	21		30				19
ë.			18	17		27	16		28			18	15
ont	11		16	14		26	12		25		20		10
lefto						30	21		30				19
trimuffleff	17			18		28	17	0	28				16
se tri	9.3	19	16	11	20	25	13	18	26	19	19	14	10
nop	annakaci	sm	fit	pink	bow	ribbon	back	trim	racer	detail	shell	style	top



ASIN : B00KOBQEBO Brand : Anna-Kaci

euclidean distance from total data (including image) : 17.283581044668928

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B00ROB8A8W	Bandolino	Aqua-Refelction	SHIRT

_				bandoli	no womens erir	round v neck ti	ie dyed shirt to	p small aqua re	felction			
		4.9	21	14	11	15		8.8	4.4	7.1		0
- Fig.	17	17					31	19	18	18		17
- It	13	14			17	18	28	15	14	14		13
e -	8.9	9.6		17	13	16		10	9.5	10	18	8.9
en pj	12	13		17	16	18		13	13	13	20	12
kagre			32	30	29	30	36	28			32	27
lo d -		20					30	20				19
ge-			26	17	17	0	29	16		16		15
벋	11	11	23	17	14	17	29	13	10	12		11
flefro	20		28	22			28	20				20
<u>_</u>	17	17	27	21			28	18	16	17		17
se th	9.3	9.4	21	16	12	16		8.3	10	11	20	9.3
nop	bandolino	womens	erin	round	neck	tie	dyed	shirt	top	small	aqua	refelction

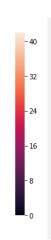


ASIN: B00ROB8A8W Brand: Bandolino

euclidean distance from total data (including image) : 17.369906725871893

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B00L87YPVY	Anna-Kaci	White	SHIRT

_				annaka	ci sm fit white s	heer floral em	broidered antiqu	e lace scallop h	em top			
		17	13	7.2	15	14	17		13	33	17	4.4
maka	17	0		18						36		18
ft au	13			14		19	20		17	36		14
plue -	8.9	18	15	6.8	17	15	16		13	33	18	9.5
en bl	12	20	17	12		16			16	34	20	13
- agre								33		42		27
dobolkar	19			18						36		19
ge -	15		18	16		20			17	36		15
펕 -	11	20	16	12	18	17			16	35		10
trimufflefront	20			20				33		36		19
흍-	17			16		20				35		16
se tr	9.3	19	16	10	16	14	16		11	33	14	10
poq	annakaci	sm	fit	white	sheer	floral	embroidered	antique	lace	scallop	hem	top





ASIN: B00L87YPVY Brand: Anna-Kaci

euclidean distance from total data (including image) : 17.433933853683165

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B0731PMQZ2	Fever	White-&-Green	SHIRT

	fever ladies size small rayon blouse white green										
kaci	30	10	7.7	7.1	20	9.3	7.2	12			
- mak	35	20	18	18			18	20			
- It	32	16	13	14		16	14	17			
plue -	30	13	11	10	20	11	6.8	11			
ig -	31	16	14	13		14	12	0			
- gre	38				32			29			
dopolka	35		19	19		21	18	20			
tie o	33	18	17	16		16	16	18			
out	32	14	12	12		14	12	15			
÷	35		20	20			20				
timuffle	34		17	17		18	16	18			
se tr	31	12	12	11	19	0	10	14			
nopq	fever	ladies	size	small	rayon	blouse	white	green			



ASIN: B0731PMQZ2 Brand: Fever

euclidean distance from total data (including image) : 17.457954237725403

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B0178WOJVS	Styleco	White	SHIRT

style co plus size new whit cuffed shortsleeve tee 2x 1498 dbfl

_				5	tyle co plus size	e new whit cull	ied shortsleeve	tee 2x 1498 dbi	l			
	12	15	7.7	7.7	8.5			0	14	16		
ana ka	20		17	18	18	33	34	17			17	17
it an	16	20	14	13	14	32		13	18		13	13
plue	14	17	11	11	11	30		8.9	15	18	8.9	8.9
en b	16		14	14	13	31	31	12	14	20	12	12
e gre		31			28	39	38	27		31		
dopolka			19	19	20	33	35	19			19	19
ė.	18		17	17	17	32	30	15	18			
벋	15	18	12	12	13	31		11	16		11	11
elefro	20			20	20	32	32	20			20	20
trimufflefront			17	17	17	34	32	17			17	17
Se tri	14	17	12	12	13	31		9.3	15	18	9.3	9.3
inold	style	0	plus	size	new	whit	cuffed	shortsleeve	tee	2x	1498	dbfl



ASIN: B0178WOJVS Brand: Styleco

euclidean distance from total data (including image) : 17.46974721771374

Asin	Brand	Color	Product type
BOOKLHUIRS	Δnna-Kaci	Rlue/Green	CHIRT

BOOKEHOIDS	Allia Naci	blue/ di ceri	ווווווו
B01KOHFVP0	namnoishop-Crop-Tops	As-show	SHIRT

			harajuku c	loud printed elastic crop to	pps women			
kaci	0		12		13	11	8.8	
maka-	17	29	20	29	22			- 32
fit ar	13	29	17		18	15	16	
e -	8.9		13		16	12	12	
en blu	12		15	27	17	15	15	- 24
kagre		36	29	35	30	29	28	
dopolka		30	20	31	23			- 16
ë -					20	16	17	
olit olit	11		14	26	17	14	13	
flefro -		30	21	29	23	20		-8
Timuf	17	30	19	28	20	17	18	
se tr	9.3		13	24	16	13	11	
poq	harajuku	doud	printed	elastic	crop	tops	women	0



ASIN : B01KQHFVP0

Brand : namnoishop Crop Tops

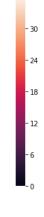
euclidean distance from total data (including image) : 17.52429944221097

.=========

Asin	Brand	Color	Product type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B010EH3B8Q	Anna-Kaci	White	SHIRT

anna kaci sm f	fit classic white should	ler hare short sleeve	nenlum flare blouse

									p				
. <u>.</u> -		0	17	13	14	7.2			8.5	9.2			9.3
snakaci		17	0	21		18		28	20			30	
- fit		13		0	18	14	18						16
g -		8.9	18	15	14	6.8	16		12	11	20		11
en blue		12		17	18	12	18		14	14			14
- gre	32		32	30			31	34	28	28	32	35	28
dobolkar		19				18						28	21
ė -		15	24	18		16			17	16			16
out -		11		16	16	12	16		13	14			14
- lefto	28	20	26			20							19
imufflefi		17				16			18	18		28	18
se tri		9.3		16	16	10	16		12	11	18		0
nopq	anna	kaci	sm	fit	dassic	white	shoulder	bare	short	sleeve	peplum	flare	blouse





ASIN: B010EH3B8Q Brand: Anna-Kaci

euclidean distance from total data (including image) : 17.530900492194565

===========

Asin Rrand Color Product type

Apili	Dranu	Color	rroduct type
B00KLHUIBS	Anna-Kaci	Blue/Green	SHIRT
B07552VG9B	Rbwinner	As-Shown	SHIRT

lady

sleeve

shirt

sxxxxl

				women	blouses turn col	lar floral tops lady	/ long sleeve shir	t sxxxxl			
aci-	8.8			17	14	11	17	6.7	9.2	8.8	0
share -		27	26								17
- fit	16					15		14	15		13
pine -	12			17		12		11	11	10	8.9
en bl	15				16	15		13	14	13	12
kagre	28	32	32	31	28	29	30	27	28	28	27
lodop		28					25				19
tie o	17	25				16		16	16	16	15
ŧ.	13				17	14		12	14	13	11
flefro											20
mof.		25				17		17			17



blouses

turn

ASIN: B07552VG9B Brand: Rbwinner

euclidean distance from total data (including image) : 17.56089272412314

collar

floral

The results for above example seems to be coming from different key words in the main products title like blouse and trim and some are recommended as they are same brand and remaining from image.

Conclusion

- I think most of the recommendations are due to the image feature similarities (even with equal weights) as our image features dimensions is high compared to dimensions of other features. Due to this most portion of the distance is taken from image and the values are already amplified even if you give same weights due to thier high dimentions. As you can see from above code 25088 dimensions are from image features and 6035 are from title and other features.
- This may lead to not-so-satisfactory results and previous models where image is not considered did well (as for my observations).

```
In [ ]:
```