

## trecs\_exp-user\_topic\_mapping

December 13, 2022

```
[ ]: import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import style
import numpy as np
from sklearn.decomposition import NMF
from importlib import reload
import wrapper

import sys
# caution: path[0] is reserved for script path (or '' in REPL)
sys.path.insert(1, '../t-recs/')
import trecs
import os
from scipy import sparse
from trecs.models import ContentFiltering
from trecs.metrics import MSEMeasurement, InteractionSpread, InteractionSpread,
    ↳InteractionSimilarity, RecSimilarity, RMSEMeasurement, InteractionMeasurement
from trecs.components import Users

from wrapper.models.bubble import BubbleBurster
# from src.utils import get_topic_clusters
from wrapper.metrics.evaluation_metrics import SerendipityMetric,
    ↳DiversityMetric, NoveltyMetric, TopicInteractionMeasurement,
    ↳MeanNumberOfTopics, UserMSEMeasurement

random_state = np.random.seed(42)
plt.style.use("seaborn-paper")
```

/var/folders/sm/hcy50x855gvf2b1qwkjstnv0000gn/T/ipykernel\_15680/1425888263.py:2  
4: MatplotlibDeprecationWarning: The seaborn styles shipped by Matplotlib are deprecated since 3.6, as they no longer correspond to the styles shipped by seaborn. However, they will remain available as 'seaborn-v0\_8-<style>'. Alternatively, directly use the seaborn API instead.  
plt.style.use("seaborn-paper")

**Ref:** <https://towardsdatascience.com/dimensionality-reduction-for-data-visualization-pca-vs-tsne-vs-umap-be4aa7b1cb29>

```
[ ]: from sklearn.cluster import KMeans

def get_topic_clusters(interaction_matrix, n_clusters:int=100, n_attrs:int=100,
↳max_iter:int=100):
    """
    Creates clusters of movies based on their genre.
    Inputs:
        binary_ratings_matrix: a binary matrix of users and movies
        n_attrs: number of attributes to use in NMF
        nmf_solver: solver to use in NMF
    Outputs:
        clusters: a list of cluster assignments
    """
    # Create topic clusters
    #create co-occurrence matrix from binary_interaction_matrix
    print('Calculating clusters...')
    co_occurrence_matrix = interaction_matrix.T @ interaction_matrix

    co_occurrence_matrix = interaction_matrix.T @ interaction_matrix

    # Matrix factorize co_occurrence_matrix to get embeddings
    nmf_cooc = NMF(n_components=n_attrs, max_iter=max_iter, verbose=1)
    W_topics = nmf_cooc.fit_transform(co_occurrence_matrix)

    # cluster W_topics
    cluster_ids = KMeans(n_clusters=n_clusters, max_iter=max_iter,
↳random_state=random_state, verbose=1).fit_predict(W_topics)
    # np.save(file_path, cluster_ids)

    print('Calculated clusters.')

    return cluster_ids
```

```
[ ]: # import warnings filter
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)
```

```
[ ]: ratings_df = pd.read_csv('data/ml-100k/u.data', sep="\t", names=['UserID',
↳'MovieID', 'Rating', 'Timestamp'])
```

```
[ ]: from sklearn.decomposition import NMF

binary_ratings_df = ratings_df.drop(columns=['Timestamp'])
binary_ratings_df.loc[binary_ratings_df['Rating'] > 0, 'Rating'] = 1

# turn dataframe into matrix where each movie is a column and each user is a row
```

```

binary_ratings_matrix = binary_ratings_df.pivot(index='UserID',
    ↪columns='MovieID', values='Rating').fillna(0).to_numpy()

from lightfm.cross_validation import random_train_test_split
from scipy import sparse

# split data into train and test sets
train_interactions, test_interactions = random_train_test_split(sparse.
    ↪csr_matrix(binary_ratings_matrix), test_percentage=0.2,
    ↪random_state=random_state)
train_interactions = train_interactions.toarray()
test_interactions = test_interactions.toarray()

n_attrs=100
nmf = NMF(n_components=n_attrs, solver="mu")
user_representation = nmf.fit_transform(binary_ratings_matrix)
item_representation = nmf.components_
print(user_representation.shape, item_representation.shape)

# nmf_solver='mu'
n_clusters=50
max_iter=500
n_attrs=100
item_topics = get_topic_clusters(binary_ratings_matrix, n_attrs=n_clusters,
    ↪max_iter=max_iter)#, nmf_solver=nmf_solver)

users = Users(size=(943,100), repeat_interactions=False)

```

```

/Users/madisonthantu/miniforge3/envs/fairRS/lib/python3.8/site-
packages/lightfm/_lightfm_fast.py:9: UserWarning: LightFM was compiled without
OpenMP support. Only a single thread will be used.
    warnings.warn(
/Users/madisonthantu/miniforge3/envs/fairRS/lib/python3.8/site-
packages/sklearn/decomposition/_nmf.py:1692: ConvergenceWarning: Maximum number
of iterations 200 reached. Increase it to improve convergence.
    warnings.warn(

(943, 100) (100, 1682)
Calculating clusters...
violation: 1.0
violation: 0.0008111012211140843
violation: 0.0018579971758544973
violation: 0.002348404969455981
violation: 0.001963674297444194
violation: 0.0015236872281129985
violation: 0.0012535307370826517
violation: 0.0010656736015344352
violation: 0.0009157574897872806

```

violation: 0.0007991225960594281  
violation: 0.000714636141551671  
violation: 0.0006490381082209844  
violation: 0.0005972169726710217  
violation: 0.0005549932878832688  
violation: 0.0005206685817139828  
violation: 0.0004915790573408767  
violation: 0.0004652150471906113  
violation: 0.00044215158985576287  
violation: 0.00042257032842413857  
violation: 0.00040465988180981  
violation: 0.00038809403110232906  
violation: 0.0003732735237999536  
violation: 0.00035958751936517996  
violation: 0.00034623336152060695  
violation: 0.00033458078658895094  
violation: 0.0003231686127826073  
violation: 0.00031319485990230906  
violation: 0.0003040539872361975  
violation: 0.000295455944650737  
violation: 0.00028767005864218845  
violation: 0.0002803228965203025  
violation: 0.0002727477995441988  
violation: 0.0002659940811304535  
violation: 0.0002591283217727306  
violation: 0.0002529716205280376  
violation: 0.0002472614811032203  
violation: 0.0002416809427548128  
violation: 0.0002366353774709594  
violation: 0.00023188267831263402  
violation: 0.0002273211363352962  
violation: 0.00022277371693806613  
violation: 0.00021874834772473648  
violation: 0.0002144331913096802  
violation: 0.00021083945802284306  
violation: 0.00020720496028834226  
violation: 0.00020380619951583227  
violation: 0.00020086767557508933  
violation: 0.00019787315562147986  
violation: 0.00019475156372002648  
violation: 0.0001917370730802528  
violation: 0.00018886868887122652  
violation: 0.0001859912197103872  
violation: 0.00018323580093702934  
violation: 0.00018062290626195358  
violation: 0.00017828193428424952  
violation: 0.00017620805406318927  
violation: 0.0001742345484770772

violation: 0.00017245644026093025  
violation: 0.00017049719087666428  
violation: 0.00016852268829433163  
violation: 0.00016630931100662804  
violation: 0.00016437812817536446  
violation: 0.0001624993397099876  
violation: 0.00016085606594463538  
violation: 0.0001590699135630679  
violation: 0.00015748793170812647  
violation: 0.000156056768668566  
violation: 0.00015461488500869694  
violation: 0.00015339804936877617  
violation: 0.00015206951623181157  
violation: 0.00015057571823102212  
violation: 0.0001492616140571389  
violation: 0.00014798536436320817  
violation: 0.00014667546143148487  
violation: 0.00014572591204239528  
violation: 0.00014468182427252872  
violation: 0.00014373682709311738  
violation: 0.0001426940209269569  
violation: 0.0001416867766560684  
violation: 0.0001406517657688315  
violation: 0.0001397758457086101  
violation: 0.00013907734142832513  
violation: 0.00013849462187389986  
violation: 0.00013772753215316225  
violation: 0.00013689198994673097  
violation: 0.00013609823909708756  
violation: 0.00013520129221992656  
violation: 0.00013429387141027088  
violation: 0.00013355703492046994  
violation: 0.00013270918822509817  
violation: 0.0001318869785264556  
violation: 0.0001309997350818319  
violation: 0.0001300957879009074  
violation: 0.00012931157491679333  
violation: 0.00012846390487191925  
violation: 0.00012775505093184374  
violation: 0.00012695874676038632  
violation: 0.00012626633434095858  
violation: 0.00012555779264843494  
violation: 0.000124941965222154  
violation: 0.00012434488540528533  
violation: 0.00012365930006005336  
violation: 0.0001231128331119413  
violation: 0.0001226249236900704  
violation: 0.0001222458218540237

violation: 0.00012183570637850564  
violation: 0.00012156473524481501  
violation: 0.00012126683561310604  
violation: 0.00012091007188147373  
violation: 0.00012068220515971065  
violation: 0.00012050852157940062  
violation: 0.000120280005609624  
violation: 0.00012005844604729106  
violation: 0.00011980599712798704  
violation: 0.00011964615289779544  
violation: 0.00011944671580841445  
violation: 0.00011916096337377208  
violation: 0.00011893054269861008  
violation: 0.0001187475781019576  
violation: 0.00011839229318121637  
violation: 0.00011799960669818226  
violation: 0.00011752254126307526  
violation: 0.0001171080948196955  
violation: 0.00011667759723436887  
violation: 0.00011621738822528687  
violation: 0.00011578483616821223  
violation: 0.00011539233096189993  
violation: 0.0001150391914142034  
violation: 0.00011469063757728474  
violation: 0.00011415785145315042  
violation: 0.0001137170754083149  
violation: 0.00011324478808189727  
violation: 0.00011280984386055837  
violation: 0.00011230808047380984  
violation: 0.00011170372047301133  
violation: 0.00011128154022196592  
violation: 0.00011076280424335234  
violation: 0.00011032562619120264  
violation: 0.00010989595146958646  
violation: 0.00010927287165412746  
violation: 0.00010863641576810569  
violation: 0.00010801692431622748  
violation: 0.0001073486761618235  
violation: 0.00010661347258346636  
violation: 0.00010581302430954739  
violation: 0.00010508941133671069  
violation: 0.00010435534617830357  
violation: 0.00010357589174540546  
violation: 0.00010286976853961824  
violation: 0.00010225929454772402  
violation: 0.0001015852154712779  
violation: 0.00010084513096403187  
violation: 0.00010011576501942235

```

violation: 9.940881036862153e-05
Converged at iteration 155
Initialization complete
Iteration 0, inertia 48.15432092266328.
Iteration 1, inertia 35.26550347136774.
Iteration 2, inertia 34.437288975092144.
Iteration 3, inertia 34.21929857699359.
Iteration 4, inertia 34.09290293216635.
Iteration 5, inertia 33.98012121826649.
Iteration 6, inertia 33.94793167438911.
Iteration 7, inertia 33.90728602110529.
Iteration 8, inertia 33.888443453013295.
Iteration 9, inertia 33.879798450818974.
Iteration 10, inertia 33.84723002422411.
Iteration 11, inertia 33.84245092302271.
Iteration 12, inertia 33.83652751863964.
Converged at iteration 12: strict convergence.
Initialization complete
Iteration 0, inertia 47.77610769502342.
Iteration 1, inertia 35.46903220989502.
Iteration 2, inertia 34.52495167045987.
Iteration 3, inertia 33.929671393784275.
Iteration 4, inertia 33.7669670689433.
Iteration 5, inertia 33.65685717683232.
Iteration 6, inertia 33.610838454179415.
Iteration 7, inertia 33.60099606512115.
Iteration 8, inertia 33.59455282492198.
Iteration 9, inertia 33.593515755496036.
Iteration 10, inertia 33.59223849332847.
Iteration 11, inertia 33.590234774685456.
Iteration 12, inertia 33.58982793446874.
Iteration 13, inertia 33.5860942699757.
Iteration 14, inertia 33.58459076143862.
Iteration 15, inertia 33.584029075082576.
Iteration 16, inertia 33.58319045906775.
Converged at iteration 16: center shift 1.2489345732429019e-07 within tolerance
1.453389046531099e-06.
Initialization complete
Iteration 0, inertia 49.39264107916929.
Iteration 1, inertia 36.049597750027225.
Iteration 2, inertia 34.97539863591017.
Iteration 3, inertia 34.56974483930776.
Iteration 4, inertia 34.321885921712344.
Iteration 5, inertia 34.22924972638986.
Iteration 6, inertia 34.16320404661296.
Iteration 7, inertia 34.14566016589062.
Iteration 8, inertia 34.12895034659863.
Iteration 9, inertia 34.11358925534266.

```

Iteration 10, inertia 34.10373745493676.  
 Iteration 11, inertia 34.097013178316644.  
 Iteration 12, inertia 34.07869937899624.  
 Iteration 13, inertia 34.069955423026805.  
 Iteration 14, inertia 34.06297065385411.  
 Iteration 15, inertia 34.06057135190548.  
 Iteration 16, inertia 34.06007117652773.  
 Iteration 17, inertia 34.05957346114912.  
 Iteration 18, inertia 34.059383496374956.  
 Converged at iteration 18: center shift 1.9702746403215421e-07 within tolerance 1.453389046531099e-06.  
 Initialization complete  
 Iteration 0, inertia 48.15707489671477.  
 Iteration 1, inertia 36.09804562795394.  
 Iteration 2, inertia 34.97081897990375.  
 Iteration 3, inertia 34.667956995527256.  
 Iteration 4, inertia 34.519456599144135.  
 Iteration 5, inertia 34.428996988027905.  
 Iteration 6, inertia 34.334163393113386.  
 Iteration 7, inertia 34.28018099157191.  
 Iteration 8, inertia 34.26577692468223.  
 Iteration 9, inertia 34.25018204006152.  
 Iteration 10, inertia 34.24631735016583.  
 Iteration 11, inertia 34.24503133789294.  
 Iteration 12, inertia 34.24482971618422.  
 Converged at iteration 12: center shift 8.177206990120169e-07 within tolerance 1.453389046531099e-06.  
 Initialization complete  
 Iteration 0, inertia 47.88841010462291.  
 Iteration 1, inertia 35.52704221373162.  
 Iteration 2, inertia 34.60367650382151.  
 Iteration 3, inertia 34.36151435348752.  
 Iteration 4, inertia 34.282846958348394.  
 Iteration 5, inertia 34.252369783992904.  
 Iteration 6, inertia 34.2302046245341.  
 Iteration 7, inertia 34.199282913224394.  
 Iteration 8, inertia 34.178554194170005.  
 Iteration 9, inertia 34.13108164583353.  
 Iteration 10, inertia 34.07464199608229.  
 Iteration 11, inertia 34.01870074582298.  
 Iteration 12, inertia 33.93754874127561.  
 Iteration 13, inertia 33.84242543133578.  
 Iteration 14, inertia 33.80849621256224.  
 Iteration 15, inertia 33.78955739444958.  
 Iteration 16, inertia 33.77715743094834.  
 Iteration 17, inertia 33.769774545599525.  
 Converged at iteration 17: strict convergence.  
 Initialization complete



Iteration 0, inertia 48.59302782872548.  
 Iteration 1, inertia 35.59681519855493.  
 Iteration 2, inertia 34.63531442979463.  
 Iteration 3, inertia 34.355103097936826.  
 Iteration 4, inertia 34.28547309542735.  
 Iteration 5, inertia 34.28181451428637.  
 Iteration 6, inertia 34.280847249144756.  
 Iteration 7, inertia 34.280720281037446.  
 Converged at iteration 7: strict convergence.  
 Initialization complete  
 Iteration 0, inertia 48.31913551771591.  
 Iteration 1, inertia 34.894035422795845.  
 Iteration 2, inertia 33.952475992020915.  
 Iteration 3, inertia 33.74336125662098.  
 Iteration 4, inertia 33.57025839962331.  
 Iteration 5, inertia 33.49065463214723.  
 Iteration 6, inertia 33.420631739380354.  
 Iteration 7, inertia 33.41199563339929.  
 Iteration 8, inertia 33.40421222606188.  
 Iteration 9, inertia 33.40026508656289.  
 Iteration 10, inertia 33.387565391836084.  
 Iteration 11, inertia 33.37581196133628.  
 Iteration 12, inertia 33.37235398522871.  
 Iteration 13, inertia 33.37002678696871.  
 Converged at iteration 13: strict convergence.  
 Initialization complete  
 Iteration 0, inertia 47.804014247676434.  
 Iteration 1, inertia 34.983172785440026.  
 Iteration 2, inertia 34.253716248611056.  
 Iteration 3, inertia 34.00490731425433.  
 Iteration 4, inertia 33.913891518411035.  
 Iteration 5, inertia 33.870776877665165.  
 Iteration 6, inertia 33.85201515351134.  
 Iteration 7, inertia 33.851630123457106.  
 Converged at iteration 7: center shift 4.5112980118090687e-07 within tolerance  
 1.453389046531099e-06.  
 Initialization complete  
 Iteration 0, inertia 47.530164254061624.  
 Iteration 1, inertia 34.65904628293155.  
 Iteration 2, inertia 33.545196270677565.  
 Iteration 3, inertia 33.14165928266153.  
 Iteration 4, inertia 33.025351013466704.  
 Iteration 5, inertia 32.91692018596355.  
 Iteration 6, inertia 32.8612023409233.  
 Iteration 7, inertia 32.856651531039695.  
 Iteration 8, inertia 32.828673449962665.  
 Iteration 9, inertia 32.826707111034494.  
 Iteration 10, inertia 32.81877318469584.

```

Iteration 11, inertia 32.81665779227108.
Iteration 12, inertia 32.81440589172527.
Iteration 13, inertia 32.813281923594275.
Converged at iteration 13: center shift 1.1317241037209851e-06 within tolerance
1.453389046531099e-06.
Initialization complete
Iteration 0, inertia 48.52701109722608.
Iteration 1, inertia 35.378678769572936.
Iteration 2, inertia 34.54186556532896.
Iteration 3, inertia 34.38200209475497.
Iteration 4, inertia 34.344062699600656.
Iteration 5, inertia 34.32014792548512.
Iteration 6, inertia 34.286042526517264.
Iteration 7, inertia 34.267768115316265.
Iteration 8, inertia 34.23511564448992.
Iteration 9, inertia 34.20369842646498.
Iteration 10, inertia 34.19133109168454.
Iteration 11, inertia 34.18655172785239.
Iteration 12, inertia 34.18355548866075.
Iteration 13, inertia 34.18277223067997.
Iteration 14, inertia 34.182148351182974.
Iteration 15, inertia 34.181869824240415.
Converged at iteration 15: strict convergence.
Calculated clusters.

```

## 1 Instantiating BubbleBurster

```

[ ]: sys.path.insert(1, '../')

from wrapper.models.bubble import BubbleBurster
from trecs.metrics import MSEMeasurement, InteractionSpread, InteractionSpread,
↪InteractionSimilarity, RecSimilarity, RMSEMeasurement, InteractionMeasurement

```

```

[ ]: bubble = BubbleBurster(
    # num_users=number_of_users,
    # num_items=num_items,
    # num_attributes=number_of_attributes,
    item_topics=item_topics,
    user_representation=user_representation,
    item_representation=item_representation,
    actual_user_representation=users,
    record_base_state=True,
)

```

```

[ ]: user_pairs = [(u_idx, v_idx) for u_idx in range(len(user_representation)) for
↪v_idx in range(len(user_representation))]

```

```

bubble.add_metrics(
    MSEMeasurement(),
    InteractionSpread(),
    InteractionSpread(),
    InteractionSimilarity(pairs=user_pairs),
    RecSimilarity(pairs=user_pairs),
    RMSEMeasurement(),
    InteractionMeasurement()
)

print("These are the current metrics:")
print(bubble.metrics)

```

These are the current metrics:

```

[<trecs.metrics.measurement.MSEMeasurement object at 0x137e89ee0>,
<trecs.metrics.measurement.InteractionSpread object at 0x137e0a310>,
<trecs.metrics.measurement.InteractionSpread object at 0x137e0a610>,
<trecs.metrics.measurement.InteractionSimilarity object at 0x137e0a940>,
<trecs.metrics.measurement.RecSimilarity object at 0x137e0a5e0>,
<trecs.metrics.measurement.RMSEMeasurement object at 0x137e0a160>,
<trecs.metrics.measurement.InteractionMeasurement object at 0x137e0ac10>]

```

## 2 Compute user-topic mappings

```

[ ]: from src.utils import user_topic_mapping

actual_user_profiles, actual_item_attributes = bubble.actual_user_profiles,
↳bubble.actual_item_attributes

actual_user_topic_mapping = user_topic_mapping(actual_user_profiles,
↳actual_item_attributes, item_topics)

```

/Users/madisonthantu/Desktop/COMS 6998/Project/fair-recommender-systems/src/utils.py:135: SyntaxWarning: assertion is always true, perhaps remove parentheses?

```

assert ((user_profiles.shape[1] == item_attributes.shape[0]),
/Users/madisonthantu/Desktop/COMS 6998/Project/fair-recommender-systems/src/utils.py:137: SyntaxWarning: assertion is always true, perhaps remove parentheses?
assert ((item_topics.shape == (item_attributes.shape[1],)),

```

```

[ ]: actual_user_topic_mapping.shape

```

```

[ ]: (943, 100)

```

```

[ ]: print(actual_user_topic_mapping[1,:])

```

```

[ 2.56964903e-01  1.25251968e-01  2.63593280e-01  4.79362095e-02

```

```

5.43423744e-02  1.98480664e-01  2.86978375e-01  2.35920670e-01
6.29925112e-02 -3.84059410e-04  2.78490848e-01  2.97822356e-01
2.03296318e-01  1.41966541e-01  5.63235221e-01  2.02339206e-01
2.01101158e-01  3.00633318e-01  1.16399742e-02  1.99466265e-02
7.48324102e-01  7.38473352e-01  1.54858094e-01  2.03099055e-01
6.12540796e-02  3.35930804e-01  3.03640859e-01  3.60386094e-02
2.14261032e-01  1.29881815e-01  4.46951415e-01  3.01742032e-02
1.26016683e-01  5.73982085e-01  8.62943310e-02  1.65594644e-02
1.10557985e-01  7.31949230e-03  7.02577000e-02  6.75177145e-02
4.17764818e-01  5.99169195e-01  3.29893134e-01  1.67902263e-01
-3.12085869e-01  1.48059449e-01  1.97184923e-01  2.08322268e-02
1.14066086e-01  1.57768681e-01  1.38855192e-01  3.89795335e-01
2.22731696e-01  3.70708166e-02 -2.20417165e-01  6.19282493e-01
4.03631358e-01  6.60432602e-01  2.17344508e-01  7.29723511e-01
4.17670047e-01  2.20886894e-01 -1.47363020e-01  6.56852408e-02
1.64408807e-01  5.95908914e-01  2.79115740e-02  3.37466043e-01
1.54310494e-01  4.86966161e-01  3.32688691e-03 -1.04852516e-02
7.39664328e-01  2.72695453e-01  6.38432344e-02  1.41291446e-02
1.87203876e-01  3.93263180e-01  5.65938292e-02  3.25403310e-01
1.19063538e-01  3.52088792e-01 -1.99822832e+00  1.11803875e-01
-3.17083872e-01  2.84724339e-01  3.41556765e-01  2.53051456e-02
1.23810432e-02  1.73522985e-01  4.27795036e-02  3.56214673e-02
9.91566409e-02  3.99785205e-02  8.50098410e-02 -1.42557576e-01
-9.69143015e-02  6.24352959e-01 -1.62096359e-02  2.97362350e-01]

```

```

[ ]: user_clusters = get_topic_clusters(binary_ratings_matrix.T, n_attrs=n_clusters,
    ↪max_iter=max_iter)#, nmf_solver=nmf_solver)
print(user_clusters.shape)
print(item_topics.shape)

```

Calculating clusters...

```

violation: 1.0
violation: 0.00023105663147641192
violation: 0.0008968555782884329
violation: 0.002797937212173178
violation: 0.00319759143004992
violation: 0.003113167357527142
violation: 0.0025264247993327677
violation: 0.002108708820660568
violation: 0.0017758168790870848
violation: 0.0015422410030184273
violation: 0.0013973782795763843
violation: 0.0013138559707382403
violation: 0.001232101569542869
violation: 0.00116297004761247
violation: 0.0010996678746988044
violation: 0.0010225828845818446
violation: 0.0009603342206569841

```

violation: 0.0009052762554306112  
violation: 0.0008578066670110475  
violation: 0.0008155155380737319  
violation: 0.0007934177895929073  
violation: 0.0007767893749273988  
violation: 0.000751243440723186  
violation: 0.0007313453770383489  
violation: 0.0007198664975299297  
violation: 0.0007083381562689952  
violation: 0.0006964311361125571  
violation: 0.0006875259660570504  
violation: 0.0006786099146629198  
violation: 0.0006679631629007681  
violation: 0.0006584014898705648  
violation: 0.0006484507432968698  
violation: 0.0006379619188683355  
violation: 0.0006270094492889831  
violation: 0.0006147416159439034  
violation: 0.0006034339191292861  
violation: 0.0005970503552810804  
violation: 0.0005912020033861382  
violation: 0.0005836785751015997  
violation: 0.0005772143633034478  
violation: 0.0005712103671903105  
violation: 0.0005622982533813432  
violation: 0.0005552852043687316  
violation: 0.0005501256404276743  
violation: 0.0005432499554931478  
violation: 0.0005372252743169817  
violation: 0.0005285410474037706  
violation: 0.0005217023598620187  
violation: 0.0005161706517748762  
violation: 0.0005071883546391862  
violation: 0.0004992470721596444  
violation: 0.0004924670143981365  
violation: 0.0004835285054936876  
violation: 0.00047580407198975154  
violation: 0.0004635585926472889  
violation: 0.0004549551648482233  
violation: 0.00044648249659430463  
violation: 0.0004365747106250542  
violation: 0.0004262311920258342  
violation: 0.0004167765561718009  
violation: 0.00040667656566726885  
violation: 0.00039667825938806466  
violation: 0.0003858842935319456  
violation: 0.00037773511514543456  
violation: 0.00036926722103784314

violation: 0.00036139377516849455  
violation: 0.0003541104215953919  
violation: 0.000346875301682097  
violation: 0.0003399573430193335  
violation: 0.0003336482920386987  
violation: 0.0003283308869897798  
violation: 0.00032373393031182295  
violation: 0.0003189648006183863  
violation: 0.0003148067171023186  
violation: 0.00030605872659556415  
violation: 0.00030086731966276877  
violation: 0.00029721510914602824  
violation: 0.0002929587578806185  
violation: 0.0002902233645633245  
violation: 0.0002878523913857171  
violation: 0.00028450160118520117  
violation: 0.00028144621499566957  
violation: 0.00027827151414167907  
violation: 0.00027581298865250747  
violation: 0.00027315889220076093  
violation: 0.00026861963246149187  
violation: 0.00026739946951989446  
violation: 0.0002660552511345655  
violation: 0.0002643516772822293  
violation: 0.0002619682942046704  
violation: 0.00025991756860185047  
violation: 0.0002582001127923777  
violation: 0.00025504186692200533  
violation: 0.0002519510636934312  
violation: 0.00024989317218474444  
violation: 0.0002473673635037919  
violation: 0.0002455020053750209  
violation: 0.00024456537591252255  
violation: 0.00024314922358259676  
violation: 0.0002416924877553776  
violation: 0.00024020293548653282  
violation: 0.0002387456441575096  
violation: 0.00023668015611098806  
violation: 0.00023500098434188463  
violation: 0.00023374018033196556  
violation: 0.00023269556249144082  
violation: 0.00023159034365932682  
violation: 0.00023053479018534154  
violation: 0.0002284268677119379  
violation: 0.0002275258744437841  
violation: 0.0002263655300523255  
violation: 0.00022534602329691992  
violation: 0.00022481091263331393

violation: 0.0002238468414342713  
violation: 0.00022204924991835978  
violation: 0.00022097204900989429  
violation: 0.0002195778813025691  
violation: 0.0002183085845205513  
violation: 0.00021693166135542996  
violation: 0.00021500737791679697  
violation: 0.00021383838885194052  
violation: 0.00021321828593981587  
violation: 0.00021186792248160306  
violation: 0.00021101809537301362  
violation: 0.00021015433023678393  
violation: 0.00020944380913302696  
violation: 0.00020754146620214421  
violation: 0.00020656716707540164  
violation: 0.0002059097563292371  
violation: 0.00020526514469925656  
violation: 0.00020459936817250846  
violation: 0.00020386703832190586  
violation: 0.00020344724832499345  
violation: 0.0002029299030209613  
violation: 0.00020255795323476027  
violation: 0.00020237071380110045  
violation: 0.00020220915630984835  
violation: 0.0002020183092529636  
violation: 0.00020197877159925103  
violation: 0.00020164909655069285  
violation: 0.00020098283617651108  
violation: 0.00020057000961555305  
violation: 0.00020031697055761114  
violation: 0.0002000956188192834  
violation: 0.00020017366196597435  
violation: 0.0002000404243481933  
violation: 0.0001996972988569716  
violation: 0.00019961873007258023  
violation: 0.00019760754210740804  
violation: 0.00019675990292734424  
violation: 0.00019604936637580557  
violation: 0.00019529326142802856  
violation: 0.0001951327814303422  
violation: 0.00019503595926075745  
violation: 0.00019486140039213324  
violation: 0.0001943855234788091  
violation: 0.00019430725243369944  
violation: 0.0001938048831416383  
violation: 0.00019388264817774706  
violation: 0.00019393774136990254  
violation: 0.00019394475531208936

violation: 0.00019444278985944642  
violation: 0.00019473566114418042  
violation: 0.000194941171524405  
violation: 0.00019492697278148305  
violation: 0.00019483312377529993  
violation: 0.00019458612250188224  
violation: 0.00019433762783132544  
violation: 0.00019399211956087553  
violation: 0.00019344280198663435  
violation: 0.0001927989992206531  
violation: 0.00019232659354775325  
violation: 0.00019171056459214248  
violation: 0.00019090458491266123  
violation: 0.00019035568917318834  
violation: 0.0001896651072209847  
violation: 0.0001890938966505794  
violation: 0.0001885632396417336  
violation: 0.0001880167432382361  
violation: 0.0001874621657097898  
violation: 0.00018681470229347933  
violation: 0.00018601614229075107  
violation: 0.00018533551165468348  
violation: 0.00018464004880998605  
violation: 0.00018335679872048322  
violation: 0.0001822727155298402  
violation: 0.0001815738788446468  
violation: 0.00018075608853581216  
violation: 0.0001796346958119595  
violation: 0.00017856733154394058  
violation: 0.00017752564932213606  
violation: 0.0001765122400093257  
violation: 0.00017597095288799397  
violation: 0.00017534459070859366  
violation: 0.0001743770315538584  
violation: 0.00017372349574700763  
violation: 0.0001726429171866598  
violation: 0.00017152608806192225  
violation: 0.00017027363534538806  
violation: 0.0001691368985609367  
violation: 0.00016806360651771193  
violation: 0.00016716379339870322  
violation: 0.00016626891974245327  
violation: 0.00016535595010234178  
violation: 0.00016436060837342757  
violation: 0.00016343978052803509  
violation: 0.00016233544771072068  
violation: 0.00016114151548567659  
violation: 0.0001596038562284963



violation: 0.00015848991707879224  
violation: 0.00015767968286166642  
violation: 0.0001567228941127434  
violation: 0.0001557327067357646  
violation: 0.00015484407616172093  
violation: 0.00015389126720684177  
violation: 0.00015299841557831194  
violation: 0.000152209944768211  
violation: 0.00015126424025098972  
violation: 0.00015051980433897504  
violation: 0.00014958135289866597  
violation: 0.00014873527645262627  
violation: 0.00014801414499280115  
violation: 0.00014711778941377626  
violation: 0.0001464561106564992  
violation: 0.00014524284000629688  
violation: 0.0001441993096410555  
violation: 0.00014317780377953745  
violation: 0.00014239111079633336  
violation: 0.00014171026660116304  
violation: 0.00014113164699706057  
violation: 0.00014065261903076914  
violation: 0.00014012376517708288  
violation: 0.00013947677647549618  
violation: 0.00013896547530245205  
violation: 0.00013856086096814156  
violation: 0.00013808735423045197  
violation: 0.00013754957360126438  
violation: 0.0001372881482269671  
violation: 0.00013693335252153128  
violation: 0.0001363426808063565  
violation: 0.00013579500697423228  
violation: 0.00013541767686159524  
violation: 0.00013504177097624885  
violation: 0.0001346695868760007  
violation: 0.00013425183498090633  
violation: 0.00013369849823037245  
violation: 0.0001334359364822611  
violation: 0.00013321899081252184  
violation: 0.00013314028081687444  
violation: 0.00013321983240100838  
violation: 0.00013257243091620256  
violation: 0.00013236952850350335  
violation: 0.0001322398959330609  
violation: 0.00013204175704052565  
violation: 0.00013201413233716804  
violation: 0.00013238766664525433  
violation: 0.00013221040287134954

violation: 0.00013166937957508667  
violation: 0.00013106844966827503  
violation: 0.00013060328335704136  
violation: 0.00013030166029076652  
violation: 0.00013011917338324704  
violation: 0.00013000752619452113  
violation: 0.00013037660989617038  
violation: 0.00013017770096779928  
violation: 0.00012992373573108982  
violation: 0.0001297390818815286  
violation: 0.00012931114910976987  
violation: 0.00012896839506065948  
violation: 0.00012910496809913108  
violation: 0.00012921094986824452  
violation: 0.00012921930702443243  
violation: 0.00012939542925042343  
violation: 0.00012959719411285593  
violation: 0.0001296594093977291  
violation: 0.00012947683370783454  
violation: 0.00012968366985201054  
violation: 0.00012972986472100837  
violation: 0.00012984050560588484  
violation: 0.00012966772885711833  
violation: 0.00012964790995819618  
violation: 0.00012965356118655466  
violation: 0.00012946190026018483  
violation: 0.00012935569738637963  
violation: 0.00012945510120999327  
violation: 0.0001297470250383389  
violation: 0.0001303000094355032  
violation: 0.00013046105406624016  
violation: 0.00013080791687858573  
violation: 0.0001315085145860125  
violation: 0.00013196868806787069  
violation: 0.00013240618436033863  
violation: 0.000132794810299209  
violation: 0.000133079343293987  
violation: 0.0001330410528994564  
violation: 0.0001333564795448265  
violation: 0.00013358400503839395  
violation: 0.0001336146736096596  
violation: 0.00013342694472733853  
violation: 0.00013336127846074995  
violation: 0.00013323007828411237  
violation: 0.00013320283856527854  
violation: 0.00013294915952880246  
violation: 0.0001330457633731801  
violation: 0.00013304502259985207

violation: 0.00013275571417004802  
violation: 0.0001318858108824086  
violation: 0.00013172520702353638  
violation: 0.0001316170323268093  
violation: 0.00013105563232161007  
violation: 0.00013032312772347658  
violation: 0.00013003131651412005  
violation: 0.00012973720378233806  
violation: 0.0001292845357547563  
violation: 0.00012891460292933397  
violation: 0.00012853040256351973  
violation: 0.0001284459256483586  
violation: 0.00012807553941161716  
violation: 0.00012760919708765153  
violation: 0.0001273168621004605  
violation: 0.00012700719616897188  
violation: 0.00012670330147512057  
violation: 0.00012635024744442133  
violation: 0.00012602577750548473  
violation: 0.00012567619090743447  
violation: 0.00012532967981231273  
violation: 0.0001250363237401324  
violation: 0.00012473430095218905  
violation: 0.00012437556370189658  
violation: 0.00012408027391922827  
violation: 0.00012377308678321205  
violation: 0.00012341253583037943  
violation: 0.00012297006530582376  
violation: 0.00012266886619811954  
violation: 0.0001223340060529698  
violation: 0.0001220284112972847  
violation: 0.00012178366664643479  
violation: 0.00012149790759675722  
violation: 0.00012128430192361708  
violation: 0.00012090926645668062  
violation: 0.00012065593557621982  
violation: 0.00012038125626971141  
violation: 0.0001202094631410268  
violation: 0.00011997504657180223  
violation: 0.00011965936960432531  
violation: 0.0001195817218766747  
violation: 0.00011959487722745453  
violation: 0.00011937500698397274  
violation: 0.00011917936174253856  
violation: 0.00011905133289352314  
violation: 0.0001189432256143155  
violation: 0.00011886247666019603  
violation: 0.0001187630745018924

violation: 0.00011863779878138066  
violation: 0.0001183246678776806  
violation: 0.00011831256067279968  
violation: 0.00011831506285814726  
violation: 0.0001182661699545628  
violation: 0.00011809976271528624  
violation: 0.00011802654432681158  
violation: 0.00011811066623921699  
violation: 0.00011802770842198514  
violation: 0.00011793181506194074  
violation: 0.00011787749907396532  
violation: 0.00011780704207763008  
violation: 0.00011763472425502143  
violation: 0.00011799223101350968  
violation: 0.00011787575701191278  
violation: 0.00011751829515210013  
violation: 0.00011743520979814978  
violation: 0.00011715314132311972  
violation: 0.0001169672407684028  
violation: 0.00011678302495303588  
violation: 0.00011662228016475324  
violation: 0.0001164575831922925  
violation: 0.00011626370494898773  
violation: 0.00011611282208943533  
violation: 0.00011601707036869638  
violation: 0.00011584123650940679  
violation: 0.0001157025447749587  
violation: 0.00011564814284117516  
violation: 0.00011565373830771702  
violation: 0.00011569010645751152  
violation: 0.00011560891087149276  
violation: 0.00011543728909743764  
violation: 0.00011527771893991003  
violation: 0.00011501481626344177  
violation: 0.00011489004089841334  
violation: 0.00011476051257431393  
violation: 0.00011459571917321296  
violation: 0.00011446156826506874  
violation: 0.0001142497328898823  
violation: 0.00011400453480961702  
violation: 0.00011376220476952039  
violation: 0.00011342339686067686  
violation: 0.0001131159706766515  
violation: 0.00011291647624163866  
violation: 0.00011243911690051634  
violation: 0.00011221967260460139  
violation: 0.00011195531732203191  
violation: 0.00011165853400010571

violation: 0.00011144299282344261  
violation: 0.00011123792562981561  
violation: 0.00011105980730652981  
violation: 0.00011092815323453345  
violation: 0.00011084912583780216  
violation: 0.00011066131303687335  
violation: 0.00011045968860964283  
violation: 0.00011017725333031826  
violation: 0.00010984981384256787  
violation: 0.00010962248662127491  
violation: 0.00010942155178979839  
violation: 0.00010923560958556218  
violation: 0.00010908075595858733  
violation: 0.00010895781318892518  
violation: 0.0001088796043847106  
violation: 0.00010878890955811974  
violation: 0.00010857174112761494  
violation: 0.00010835750351800646  
violation: 0.00010822595576165581  
violation: 0.00010809492851705961  
violation: 0.0001079512452615907  
violation: 0.00010767547166359753  
violation: 0.00010744078959630728  
violation: 0.0001071928457255301  
violation: 0.00010691299753713961  
violation: 0.00010661953265057088  
violation: 0.00010633693827490721  
violation: 0.00010605556984429274  
violation: 0.00010583023835191906  
violation: 0.00010561829646579356  
violation: 0.00010533195934997365  
violation: 0.00010497053829071802  
violation: 0.00010460131428085328  
violation: 0.00010425436708161393  
violation: 0.00010391435273988114  
violation: 0.00010349126300540587  
violation: 0.00010313128272605646  
violation: 0.00010283122207252228  
violation: 0.00010252369483327408  
violation: 0.00010222524380559169  
violation: 0.00010192794884596724  
violation: 0.00010165643006475742  
violation: 0.00010143617287753535  
violation: 0.00010124685845635797  
violation: 0.0001010214837559125  
violation: 0.00010073820546631556  
violation: 0.00010056556909730763  
violation: 0.00010042978430520254

violation: 0.00010024830973945044  
 violation: 0.00010001695788942446  
 violation: 9.977317881285299e-05  
 Converged at iteration 453  
 Initialization complete  
 Iteration 0, inertia 15.230433191612041.  
 Iteration 1, inertia 10.9258315256894.  
 Iteration 2, inertia 10.639470565383899.  
 Iteration 3, inertia 10.528642521125548.  
 Iteration 4, inertia 10.49028335070887.  
 Iteration 5, inertia 10.477188619846498.  
 Iteration 6, inertia 10.47458564596802.  
 Iteration 7, inertia 10.471592019888611.  
 Iteration 8, inertia 10.470625516653385.  
 Iteration 9, inertia 10.470470163609779.  
 Iteration 10, inertia 10.46966106578488.  
 Converged at iteration 10: strict convergence.  
 Initialization complete  
 Iteration 0, inertia 15.009376777337653.  
 Iteration 1, inertia 11.114539643525653.  
 Iteration 2, inertia 10.807156561335212.  
 Iteration 3, inertia 10.732447705374128.  
 Iteration 4, inertia 10.690543404298705.  
 Iteration 5, inertia 10.685056330699155.  
 Iteration 6, inertia 10.680159540909708.  
 Iteration 7, inertia 10.675850107280791.  
 Iteration 8, inertia 10.670904462728972.  
 Iteration 9, inertia 10.670686025168568.  
 Iteration 10, inertia 10.670592091054885.  
 Converged at iteration 10: strict convergence.  
 Initialization complete  
 Iteration 0, inertia 14.82893371288212.  
 Iteration 1, inertia 11.095889400341619.  
 Iteration 2, inertia 10.837225773680263.  
 Iteration 3, inertia 10.72533714195952.  
 Iteration 4, inertia 10.695704996499922.  
 Iteration 5, inertia 10.683764534125093.  
 Iteration 6, inertia 10.677641133131779.  
 Iteration 7, inertia 10.673018787027722.  
 Iteration 8, inertia 10.6678267310072.  
 Iteration 9, inertia 10.661600059548999.  
 Iteration 10, inertia 10.647241755416495.  
 Iteration 11, inertia 10.639391037872437.  
 Iteration 12, inertia 10.633035999037556.  
 Iteration 13, inertia 10.62475136760789.  
 Iteration 14, inertia 10.620813436743001.  
 Iteration 15, inertia 10.619613694648397.  
 Converged at iteration 15: strict convergence.

```

Initialization complete
Iteration 0, inertia 14.95949764661044.
Iteration 1, inertia 10.98969721183126.
Iteration 2, inertia 10.563722378040026.
Iteration 3, inertia 10.40408102179357.
Iteration 4, inertia 10.339857425313262.
Iteration 5, inertia 10.289604476170322.
Iteration 6, inertia 10.27520232269998.
Iteration 7, inertia 10.267634294723901.
Iteration 8, inertia 10.262889573548652.
Iteration 9, inertia 10.260045837611397.
Iteration 10, inertia 10.259552663113249.
Converged at iteration 10: strict convergence.
Initialization complete
Iteration 0, inertia 14.860359549222512.
Iteration 1, inertia 10.91219585474904.
Iteration 2, inertia 10.558240549042553.
Iteration 3, inertia 10.505001725355132.
Iteration 4, inertia 10.476938330746375.
Iteration 5, inertia 10.472157681024022.
Iteration 6, inertia 10.469283066850469.
Iteration 7, inertia 10.467748662348393.
Iteration 8, inertia 10.467531814803465.
Iteration 9, inertia 10.466008999748816.
Iteration 10, inertia 10.464508698658001.
Converged at iteration 10: strict convergence.
Initialization complete
Iteration 0, inertia 14.808833593086847.
Iteration 1, inertia 10.983820260862624.
Iteration 2, inertia 10.59569782418625.
Iteration 3, inertia 10.51377043996034.
Iteration 4, inertia 10.474704081679281.
Iteration 5, inertia 10.462952288419686.
Iteration 6, inertia 10.442177960035295.
Iteration 7, inertia 10.43471167756422.
Iteration 8, inertia 10.432569309032122.
Iteration 9, inertia 10.432056151906025.
Converged at iteration 9: strict convergence.
Initialization complete
Iteration 0, inertia 14.897559864886855.
Iteration 1, inertia 10.889128488131165.
Iteration 2, inertia 10.593726156645307.
Iteration 3, inertia 10.494076369760258.
Iteration 4, inertia 10.438482780030728.
Iteration 5, inertia 10.418009395133478.
Iteration 6, inertia 10.409718662059557.
Iteration 7, inertia 10.401633909205117.
Iteration 8, inertia 10.386820928009971.

```

Iteration 9, inertia 10.368835273050383.  
 Iteration 10, inertia 10.357220627630419.  
 Iteration 11, inertia 10.35490405198269.  
 Iteration 12, inertia 10.354026729838246.  
 Iteration 13, inertia 10.352431047699083.  
 Iteration 14, inertia 10.351805063139743.  
 Iteration 15, inertia 10.35165003043074.  
 Iteration 16, inertia 10.3515589360542.  
 Converged at iteration 16: strict convergence.  
 Initialization complete  
 Iteration 0, inertia 14.839412663296933.  
 Iteration 1, inertia 10.777320064909642.  
 Iteration 2, inertia 10.528384851177165.  
 Iteration 3, inertia 10.395741626951729.  
 Iteration 4, inertia 10.358186967091925.  
 Iteration 5, inertia 10.341677786872328.  
 Iteration 6, inertia 10.337641539436644.  
 Iteration 7, inertia 10.33713330430263.  
 Converged at iteration 7: center shift 2.066732268015102e-07 within tolerance 7.003378491781546e-07.  
 Initialization complete  
 Iteration 0, inertia 15.294135782968407.  
 Iteration 1, inertia 10.76025287346923.  
 Iteration 2, inertia 10.4053572354922.  
 Iteration 3, inertia 10.3242926392241.  
 Iteration 4, inertia 10.265629006656292.  
 Iteration 5, inertia 10.237240611214919.  
 Iteration 6, inertia 10.220941558503846.  
 Iteration 7, inertia 10.220843675435697.  
 Iteration 8, inertia 10.220139449777594.  
 Iteration 9, inertia 10.219365602519915.  
 Converged at iteration 9: center shift 1.4452889134366685e-07 within tolerance 7.003378491781546e-07.  
 Initialization complete  
 Iteration 0, inertia 15.219164941410767.  
 Iteration 1, inertia 10.925257269966284.  
 Iteration 2, inertia 10.583533300151208.  
 Iteration 3, inertia 10.464389367271789.  
 Iteration 4, inertia 10.395482811418859.  
 Iteration 5, inertia 10.372666301479178.  
 Iteration 6, inertia 10.365475948058915.  
 Iteration 7, inertia 10.364014655396417.  
 Iteration 8, inertia 10.360559937464938.  
 Iteration 9, inertia 10.35733934303313.  
 Converged at iteration 9: strict convergence.  
 Calculated clusters.  
 (943,)  
 (1682,)



### 3 Visualizing user-topic mappings

```
[ ]: import time

# For plotting
import plotly.io as plt_io
import plotly.graph_objects as go
```

```
[ ]: #PCA
from sklearn.decomposition import PCA
#TSNE
from sklearn.manifold import TSNE
#UMAP
import umap
#LDA
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA
```

/Users/madisonthantu/miniforge3/envs/fairRS/lib/python3.8/site-packages/tqdm/auto.py:22: TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. See [https://ipywidgets.readthedocs.io/en/stable/user\\_install.html](https://ipywidgets.readthedocs.io/en/stable/user_install.html)  
from .autonotebook import tqdm as notebook\_tqdm

```
[ ]: def plot_2d(component1, component2):

    fig = go.Figure(data=go.Scatter(
        x = component1,
        y = component2,
        mode='markers',
        marker=dict(
            size=20,
            color=user_clusters, #set color equal to a variable
            colorscale='Rainbow', # one of plotly colorscales
            showscale=True,
            line_width=1
        )
    ))
    fig.update_layout(margin=dict(l=100,r=100,b=100,t=100),width=750,height=450)
    fig.layout.template = 'plotly_dark'

    fig.show()
```

```
[ ]: def plot_3d(component1, component2, component3):
    fig = go.Figure(data=[go.Scatter3d(
        x=component1,
        y=component2,
        z=component3,
```

```

        mode='markers',
        marker=dict(
            size=10,
            color=user_clusters,          # set color to an array/list of
↳desired values
            colorscale='Rainbow',      # choose a colorscale
            opacity=1,
            line_width=1
        )
    ))
# tight layout
fig.update_layout(margin=dict(l=50,r=50,b=50,t=50),width=900,height=500)
fig.layout.template = 'plotly_dark'

fig.show()

```

```

[ ]: from sklearn.preprocessing import StandardScaler
    ## Standardizing the data
    x = StandardScaler().fit_transform(actual_user_topic_mapping)

```

### 3.0.1 PCA

```

[ ]: start = time.time()
    pca = PCA(n_components=3)
    principalComponents = pca.fit_transform(x)
    print('Duration: {} seconds'.format(time.time() - start))
    principal = pd.DataFrame(data = principalComponents
        , columns = ['principal component 1', 'principal component
↳2', 'principal component 3'])

```

Duration: 0.009582996368408203 seconds

```

[ ]: plot_2d(principalComponents[:, 0],principalComponents[:, 1])

```

```

[ ]: plot_3d(principalComponents[:, 0],principalComponents[:,
↳1],principalComponents[:, 2])

```

### 3.0.2 t-SNE

```

[ ]: from sklearn.decomposition import TruncatedSVD

    truncated_svd = TruncatedSVD(n_components=50)
    X_svd = truncated_svd.fit_transform(x)
    tsne = TSNE(random_state = 42, n_components=3,verbose=0, perplexity=40,
↳n_iter=400).fit_transform(X_svd)
    print('Duration: {} seconds'.format(time.time() - start))

```

Duration: 2.1017019748687744 seconds

```
[ ]: plot_2d(tsne[:, 0],tsne[:, 1])
```

```
[ ]: plot_3d(tsne[:, 0],tsne[:, 1],tsne[:, 2])
```

### 3.0.3 UMAP

```
[ ]: start = time.time()
reducer = umap.UMAP(random_state=42,n_components=3)
embedding = reducer.fit_transform(x)
print('Duration: {} seconds'.format(time.time() - start))
```

Duration: 3.2564289569854736 seconds

```
[ ]: plot_2d(reducer.embedding[:, 0],reducer.embedding[:, 1])
```

```
[ ]: plot_3d(reducer.embedding[:, 0],reducer.embedding[:, 1],reducer.embedding[:, 2])
```

### 3.0.4 LDA

```
[ ]: start = time.time()
X_LDA = LDA(n_components=3).fit_transform(x, y=user_clusters)
print('Duration: {} seconds'.format(time.time() - start))

# from sklearn.decomposition import LatentDirichletAllocation

# lda = LatentDirichletAllocation(n_components=3, random_state=0)
# lda_x = lda.fit_transform(x)
# print('Duration: {} seconds'.format(time.time() - start))
```

Duration: 0.07077908515930176 seconds

```
[ ]: plot_2d(X_LDA[:, 0],X_LDA[:, 1])
```

```
[ ]: plot_3d(X_LDA[:, 0],X_LDA[:, 1],X_LDA[:, 2])
```