Topic: Recommendation Engine

Instructions:

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

Name: _Nukala Ayyappa Bharthwaj_ Batch ID: <u>DSWDMCOS21012022</u>

Topic: Recommender Engine

Hints:

1. Business Problem

 1.1. What is the business objective?

 1.1. Are there any constraints?

2. Work on each feature of the dataset to create a data dictionary as displayed in the image below:

Name of Feature	Description	Type	Relevance
ID	Customer ID	Quantitative, Nominal	Irrelevant, ID does not provide useful information

3. Data Pre-processing

 2.1 Data Cleaning and Data Mining.

4. Exploratory Data Analysis (EDA):

 4.1. Summary.

 4.2. Univariate analysis.

 4.3. Bivariate analysis.

5. Model Building

 5.1 Build the Recommender Engine model on the given data sets.

6. Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?

Problem Statement: -

Q) Build a recommender system with the given data using UBCF.

This dataset is related to the video gaming industry and a survey was conducted to build a recommendation engine so that the store can improve the sales of its gaming DVDs. Snapshot of the dataset is given below. Build a Recommendation Engine and suggest top selling DVDs to the store customers.

userId	game	rating
3	The Legen	4
6	Tony Haw	5
8	Grand The	4
10	SoulCalibu	4
11	Grand The	4.5
12	Super Mar	4
13	Super Mar	4
14	Grand The	4.5
16	Grand The	3
19	Grand The	5
22	Tony Haw	3
23	The Legen	4
24	Tony Haw	4
31	Perfect Da	3
34	Grand The	5
39	Metroid P	5

Statement : Best Recommended and rated game to **"Super Mario Galaxy 2"** are **"The Legend of Zelda: Breath of the Wild"** and **"Grand Theft Auto IV"**

```
In [34]: get_recommendations("Super Mario Galaxy 2", topN = 5)
    ...: Gaming_index["Super Mario Galaxy 2"]
    index                                        name     Score
0       6                         Super Mario Galaxy 2  1.000000
1      11      The Legend of Zelda: Breath of the Wild  0.902898
2       2                           Grand Theft Auto IV  0.894961
3    2497  Mario vs. Donkey Kong: Minis March Again!    0.878403
4      85                                        Halo 3  0.750816
5      30                                      Tekken 3  0.748540
Out[34]: 6
```

Benefits and Impact Solution : If one customer is Playing **"Super Mario Galaxy 2",** he are she will be interested In playing **"The Legend of Zelda: Breath of the Wild"** and **"Grand Theft Auto IV"** as this games are as interesting as **"Super Mario Galaxy 2",**

Problem Statement: -

The Entertainment Company, which is an online movie watching platform, wants to improve its collection of movies and showcase those that are highly rated and recommend those movies to

its customer by their movie watching footprint. For this, the company has collected the data and shared it with you to provide some analytical insights and also to come up with a recommendation algorithm so that it can automate its process for effective recommendations. The ratings are between -9 and +9.

Id	Titles	Category	Reviews
6973	Toy Story (1995)	Drama, Romanc	-8.98
6778	Jumanji (1995)	Action, Adventur	8.88
9702	Grumpier Old M	Action, Comedy,	99
6769	Waiting to Exhal	Sci-Fi, Thriller	99
1123	Father of the Bri	Action, Comedy,	-0.44
9860	Heat (1995)	Comedy, Drama	-6.65
1803	Sabrina (1995)	Action, Adventur	99
9721	Tom and Huck (1	Drama, Military,	-5.19
6563	Sudden Death (1	Action, Comedy,	-7.86
1323	GoldenEye (199	Action, Comedy,	3.01
7889	American Presid	Drama, Fantasy,	99
7650	Dracula: Dead a	Drama, School, :	-7.23
5467	Balto (1995)	Action, Comedy,	99
7451	Nixon (1995)	Action, Drama, N	5.97
5898	Cutthroat Island	Comedy, Drama	-5.92
7355	Casino (1995)	Adventure, Dram	2.91
6791	Sense and Sens	Drama, Music, R	99
8574	Four Rooms (19	Adventure, Fanti	0.53

Statement : Fot the Given Statement as Example we took Movie **Babe (1995)** and Got Recommended movies are **Get Shorty (1995)** and **Pocahontas (1995).**

```
In [35]: get_recommendations("Babe (1995)", topN = 10)
    index                            name     Score
0    20               Get Shorty (1995)  1.000000
1    32                     Babe (1995)  1.000000
2    44               Pocahontas (1995)  1.000000
3     5                     Heat (1995)  0.862342
4    14           Cutthroat Island (1995)  0.862342
5    43       Seven (a.k.a. Se7en) (1995)  0.862342
6    39              Restoration (1995)  0.591565
7    30          Dangerous Minds (1995)  0.452648
8    11  Dracula: Dead and Loving It (1995)  0.439094
9    42  How to Make an American Quilt (1995)  0.412933
10   47          Mighty Aphrodite (1995)  0.377452
```

Benefits and Impact Solution : If one customer is watching a movie or looking for a movie to watch with particular gener he can get similar suggestion of another movies.