```
!pip install turicreate
```

```
from google.colab import drive
drive.mount('/content/gdrive')
```

        Mounted at /content/gdrive

▾ Build a song recommender system

```
import turicreate
```

▾ Load some music data

```
song_data = turicreate.SFrame('/content/gdrive/My Drive/Turicreate/Week 5/song_data.sframe/')
```

▾ Explore our data

```
song_data
```

user_id	song_id	listen_count	title
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SOAKIMP12A8C130995	1	The Cove
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SOBBMDR12A8C13253B	2	Entre Dos Aguas
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SOBXHDL12A81C204C0	1	Stronger
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SOBYHAJ12A6701BF1D	1	Constellations
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SODACBL12A8C13C273	1	Learn To Fly
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SODDNQT12A6D4F5F7E	5	Apuesta Por El Rock 'N' Roll ...
b80344d063b5ccb3212f76538 f3d9e43d87dca9e ...	SODXRTY12AB0180F3B	1	Paper Gangsta

Show the most popular songs in the dataset

| | f3d9e43d87dca9e ... | | | | | kosmisch |

```
song_data['song'].show()
```

| | | | | | | your |

Count the number of unique users in the data

| | The Cove - Jack Johnson | |
| | Entre Dos Aguas - Paco D... | |

```
users = song_data['user_id'].unique()
```

| | Stronger - Kanye West | |

```
len(users)
```

66346

| | Apuesta Por El Rock 'N' | |

Create a song recommender

| | f.......g...... | |

```
train_data,test_data = song_data.random_split(.8,seed=0)
```

| | EXE5 - lUlEXEl.. | |

Create a very simple popularity recommender

```
popularity_model = turicreate.popularity_recommender.create(train_data,
                                                user_id = 'user_id',
                                                item_id = 'song')
```

```
Warning: Ignoring columns song_id, listen_count, title, artist;
```

▼ Use the popularity model to make some predictions

```
Data prepared in: 1.41054s
```

```
popularity_model.recommend(users=[users[0]])
```

user_id	song	score	rank
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Sehr kosmisch - Harmonia	4754.0	1
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Undo - Björk	4227.0	2
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	You're The One - Dwight Yoakam …	3781.0	3
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Dog Days Are Over (Radio Edit) - Florence + The …	3633.0	4
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Revelry - Kings Of Leon	3527.0	5
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Horn Concerto No. 4 in E flat K495: II. Romance …	3161.0	6
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Secrets - OneRepublic	3148.0	7
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Fireflies - Charttraxx Karaoke …	2532.0	8
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Tive Sim - Cartola	2521.0	9
c66c10a9567f0d82ff31441a9 fd5063e5cd9dfe8 …	Drop The World - Lil Wayne / Eminem …	2053.0	10

[10 rows x 4 columns]

```
popularity_model.recommend(users=[users[1]])
```

user_id	song	score	rank
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Sehr kosmisch - Harmonia	4754.0	1
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Undo - Björk	4227.0	2
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	You're The One - Dwight Yoakam …	3781.0	3
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Dog Days Are Over (Radio Edit) - Florence + The …	3633.0	4
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Revelry - Kings Of Leon	3527.0	5
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Horn Concerto No. 4 in E flat K495: II. Romance …	3161.0	6
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Secrets - OneRepublic	3148.0	7
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Hey_ Soul Sister - Train	2538.0	8
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Fireflies - Charttraxx Karaoke …	2532.0	9
02f015d32ac2cd1e52d26e3ec 36048711dd5711b …	Tive Sim - Cartola	2521.0	10

[10 rows x 4 columns]

▾ Build a recommender with personalization

```
personalized_model = turicreate.item_similarity_recommender.create(train_data,
                                                                    user_id = 'user_id',
                                                                    item_id = 'song')
```

```
        Warning: Ignoring columns song_id, listen_count, title, artist;
         To use one of these as a rating column, specify the column name to be used as target
            and use a method that allows the use of a target.
        Preparing data set.
            Data has 893580 observations with 66085 users and 9952 items.
            Data prepared in: 1.49533s
        Training model from provided data.
        Gathering per-item and per-user statistics.
        +--------------------------------+------------+
        | Elapsed Time (Item Statistics) | % Complete |
        +--------------------------------+------------+
        | 1.407ms                        | 1.5        |
        | 55.57ms                        | 100        |
        +--------------------------------+------------+
        Setting up lookup tables.
        Processing data in one pass using dense lookup tables.
        +-----------------------------------+-----------------+-----------------+
        | Elapsed Time (Constructing Lookups) | Total % Complete | Items Processed |
        +-----------------------------------+-----------------+-----------------+
        | 256.656ms                         | 0               | 0               |
        | 2.01s                             | 100             | 9952            |
        +-----------------------------------+-----------------+-----------------+
        Finalizing lookup tables.
        Generating candidate set for working with new users.
        Finished training in 3.09063s
```

▾ Apply personalized model to make song recommendations

```
personalized_model.recommend(users=[users[0]])
```

user_id	song	score	rank
c66c10a9567f0d82ff31441a9fd5063e5cd9dfe8 …	Cuando Pase El Temblor - Soda Stereo …	0.019450453611520622	1
c66c10a9567f0d82ff31441a9fd5063e5cd9dfe8 …	Fireflies - Charttraxx Karaoke …	0.014473731701190654	2

```
personalized_model.recommend(users=[users[1]])
```

user_id	song	score	rank
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Where The Boat Leaves From (Album) - Zac Brown …	0.06153608858585358	1
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Different Kind Of Fine (Album) - Zac Brown Band …	0.06052837520837784	2
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Jolene (Album) - Zac Brown Band …	0.05786820501089096	3
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Sic 'Em On A Chicken (Album) - Zac Brown Band …	0.05518669635057449	4
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Who's Kissing You Tonight - Jason Aldean …	0.05306333303451538	5
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	What Country Is - Luke Bryan …	0.037490807473659515	6
02f015d32ac2cd1e52d26e3ec36048711dd5711b …	Highway 20 Ride (Album) - Zac Brown Band …	0.0373315736651 4206	7
02f015d32ac2cd1e52d26e3ec	Do I - Luke Bryan	0.033077359199523926	8

▾ Apply model to find similar songs in the data set

```
personalized_model.get_similar_items(['With Or Without You – U2'])
```

song	similar	score	rank
With Or Without You - U2	I Still Haven't Found What I'm Looking For …	0.04285717010498047	1
With Or Without You - U2	Hold Me_ Thrill Me_ Kiss Me_ Kill Me - U2 …	0.033734917640686035	2
With Or Without You - U2	Window In The Skies - U2	0.03283584117889404	3
With Or Without You - U2	Vertigo - U2	0.03007519245147705	4
With Or Without You - U2	Sunday Bloody Sunday - U2	0.02713179588317871	5
With Or Without You - U2	Bad - U2	0.02517986297607422	6
With Or Without You - U2	A Day Without Me - U2	0.023715436458587646	7
With Or Without You -	Another Time Another	0.020325183868408203	8

```
personalized_model.get_similar_items(['Chan Chan (Live) – Buena Vista Social Club'])
```

song	similar	score	rank
Chan Chan (Live) - Buena Vista Social Club …	Murmullo - Buena Vista Social Club …	0.1881188154220581	1
Chan Chan (Live) - Buena Vista Social Club …	La Bayamesa - Buena Vista Social Club …	0.1871921420097351	2
Chan Chan (Live) - Buena Vista Social Club …	Amor de Loca Juventud - Buena Vista Social Club …	0.18483412265777588	3
Chan Chan (Live) - Buena Vista Social Club …	Diferente - Gotan Project	0.021459221839904785	4
Chan Chan (Live) - Buena Vista Social Club …	Mistica - Orishas	0.020576119422912598	5
Chan Chan (Live) - Buena Vista Social Club …	Hotel California - Gipsy Kings …	0.019304990768432617	6
Chan Chan (Live) - Buena Vista Social Club …	Nací Orishas - Orishas	0.019157111644744873	7
Chan Chan (Live) -	Gitana - Willie Colon	0.018796980381011963	8

▾ Compare the models quantitatively

We now formally compare the popularity and the personalized models using precision-recall curves.

```
%matplotlib inline
model_performance = turicreate.recommender.util.compare_models(test_data, [popularity_model, person
```

↳

```
compare_models: using 2931 users to estimate model performance
PROGRESS: Evaluate model M0
recommendations finished on 1000/2931 queries. users per second: 2880.67
recommendations finished on 2000/2931 queries. users per second: 2827.94
WARNING:root:Model trained without a target. Skipping RMSE computation.


Precision and recall summary statistics by cutoff
```

cutoff	mean_precision	mean_recall
1	0.02388263391334016	0.006151279820675928
2	0.02200614124872059	0.011831884708036184
3	0.021835551006482472	0.018200163875701256
4	0.019617877857386577	0.022059848108978125
5	0.018287273967929066	0.025565896062313683
6	0.017229614466052538	0.02918379494482813
7	0.016376663254861864	0.032297225590018665
8	0.01586489252814739	0.03537798330230148
9	0.015201486030554625	0.037579592016192646
10	0.014602524735585164	0.03973713158417237

```
[10 rows x 3 columns]
```

The table shows that the personalized model provides much better performance.

```
recommendations finished on 2000/2931 queries. users per second: 2725.34
WARNING:root:Model trained without a target. Skipping RMSE computation.


Precision and recall summary statistics by cutoff
```

cutoff	mean_precision	mean_recall
1	0.1047424087342204	0.03105546520186173
2	0.08512453087683383	0.04995969537784218
3	0.07380871147503701	0.06378640659452275
4	0.0655066530194473	0.07525315547171171
5	0.05991129307403611	0.08455082357183623
6	0.0540202433754123	0.09211122064093345
7	0.049617390456694464	0.09838824378080202
8	0.04652848857045376	0.104057632203517
9	0.04367110201296487	0.10908059173005889
10	0.04152166496076429	0.11476113602370451

```
[10 rows x 3 columns]
```

✓ 3초   오후 11:12에 완료됨   ● ✕