OLAP and Data Cube

Setup

```
!pip install atoti
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
Collecting atoti
  Downloading atoti-0.7.4-0_cb607ee-py3-none-any.whl (127.0 MB)
                                            ━━━ 127.0/127.0 MB 7.8 MB/s eta 0:00
Requirement already satisfied: python-dateutil in /usr/local/lib/python3.9/di
Requirement already satisfied: py4j~=0.10.9 in /usr/local/lib/python3.9/dist-
Collecting jdk4py~=17.0
  Downloading jdk4py-17.0.3.0-0-py3-none-manylinux1_x86_64.whl (37.2 MB)
                                            ━━━ 37.2/37.2 MB 15.6 MB/s eta 0:00
Collecting atoti-core==0.7.4
  Downloading atoti_core-0.7.4-0_cb607ee-py3-none-any.whl (60 kB)
                                            ━━━ 60.4/60.4 KB 6.5 MB/s eta 0:00:
Collecting typeguard
  Downloading typeguard-3.0.2-py3-none-any.whl (30 kB)
Requirement already satisfied: pyarrow in /usr/local/lib/python3.9/dist-packa
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.9/
Collecting atoti-query==0.7.4
  Downloading atoti_query-0.7.4-0_cb607ee-py3-none-any.whl (39 kB)
Requirement already satisfied: pandas in /usr/local/lib/python3.9/dist-packag
Requirement already satisfied: numpy~=1.22 in /usr/local/lib/python3.9/dist-p
Collecting importlib-metadata~=4.10
  Downloading importlib_metadata-4.13.0-py3-none-any.whl (23 kB)
Collecting typeguard
  Downloading typeguard-2.13.3-py3-none-any.whl (17 kB)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.9/dist-
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.9/dist-pack
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.9/dist-pac
Installing collected packages: typeguard, jdk4py, importlib-metadata, atoti-c
  Attempting uninstall: importlib-metadata
    Found existing installation: importlib-metadata 6.1.0
    Uninstalling importlib-metadata-6.1.0:
      Successfully uninstalled importlib-metadata-6.1.0
Successfully installed atoti-0.7.4 atoti-core-0.7.4 atoti-query-0.7.4 importl
```

```
import pandas as pd
import atoti as tt
```

Welcome to atoti 0.7.4!

By using this community edition, you agree with the license available at http
Browse the official documentation at https://docs.atoti.io.
Join the community at https://www.atoti.io/register.

atoti collects telemetry data, which is used to help understand how to improv
If you don't wish to send usage data, you can apply for an evaluation license

You can hide this message by setting the `ATOTI_HIDE_EULA_MESSAGE` environmen

▾ Create Session

```
session = tt.Session()
```

```
df = pd.read_csv('/content/Spotify_Youtube_Sample.csv')
df.head()
```

	Artist	Track	Album	Album_type	Views	Likes	Comments	License
0	Gorillaz	Feel Good Inc.	Demon Days	album	693555221.0	6220896.0	169907.0	Tru
1	Gorillaz	Rhinestone Eyes	Plastic Beach	album	72011645.0	1079128.0	31003.0	Tru
2	Gorillaz	New Gold (feat. Tame Impala and Bootie Brown)	New Gold (feat. Tame Impala and Bootie Brown)	single	8435055.0	282142.0	7399.0	Tru
3	Gorillaz	On Melancholy Hill	Plastic Beach	album	211754952.0	1788577.0	55229.0	Tru
4	Gorillaz	Clint Eastwood	Gorillaz	album	618480958.0	6197318.0	155930.0	Tru

```
views = session.read_csv(
    '/content/Spotify_Youtube_Sample.csv',
    keys = ['Artist', 'Track', 'Album'],
)
views.head()
```

Artist	Track	Album	Album_type	Views	Likes	Comments	Licensed
Ryan Castro	Wasa Wasa	Wasa Wasa	single	115121545.0	761203.0	17238.0	True
	Avemaría	Avemaría	single	10838443.0	96423.0	6616.0	False
Omar Apollo	Invincible (feat. Daniel Caesar)	Ivory	album	1967236.0	38113.0	764.0	True
	Useless	Apolonio	album	469551.0	13611.0	405.0	True
	Endlessly	Ivory (Marfil)	album	210243.0	4704.0	123.0	True

▾ Create a Data Cube

```
cube = session.create_cube(views, name = 'Views')
cube
```

- Views
  - Dimensions
    - Spotify_youtube_sample
      - Album
        1. Album
      - Album_type
        1. Album_type
      - Artist
        1. Artist
      - Licensed
        1. Licensed
      - Track
        1. Track
      - official_video
        1. official_video
  - Measures
    - Comments.MEAN
      - formatter: DOUBLE[#,###.00]
    - Comments.SUM
      - formatter: DOUBLE[#,###.00]
    - Likes.MEAN
      - formatter: DOUBLE[#,###.00]
    - Likes.SUM
      - formatter: DOUBLE[#,###.00]
    - Stream.MEAN
      - formatter: DOUBLE[#,###.00]
    - Stream.SUM
      - formatter: DOUBLE[#,###.00]
    - Views.MEAN
      - formatter: DOUBLE[#,###.00]
    - Views.SUM
      - formatter: DOUBLE[#,###.00]
    - contributors.COUNT
      - formatter: INT[#,###]

```
cube = session.create_cube(views, name = 'Views')
cube
```

```
m= cube.measures
m
```

- Measures
  - Comments.MEAN
    - formatter: DOUBLE[#,###.00]
  - Comments.SUM
    - formatter: DOUBLE[#,###.00]
  - Likes.MEAN
    - formatter: DOUBLE[#,###.00]
  - Likes.SUM
    - formatter: DOUBLE[#,###.00]
  - Stream.MEAN
    - formatter: DOUBLE[#,###.00]
  - Stream.SUM
    - formatter: DOUBLE[#,###.00]
  - Views.MEAN
    - formatter: DOUBLE[#,###.00]
  - Views.SUM
    - formatter: DOUBLE[#,###.00]
  - contributors.COUNT
    - formatter: INT[#,###]

```
l = cube.levels
l
```

- Levels
  - Album (Spotify_youtube_sample/Album/Album)
    - dimension: Spotify_youtube_sample
    - hierarchy: Album
    - type: String
    - order: NaturalOrder
  - Album_type (Spotify_youtube_sample/Album_type/Album_type)
    - dimension: Spotify_youtube_sample
    - hierarchy: Album_type
    - type: String
    - order: NaturalOrder
  - Artist (Spotify_youtube_sample/Artist/Artist)
    - dimension: Spotify_youtube_sample
    - hierarchy: Artist
    - type: String
    - order: NaturalOrder
  - Licensed (Spotify_youtube_sample/Licensed/Licensed)
    - dimension: Spotify_youtube_sample
    - hierarchy: Licensed
    - type: boolean
    - order: NaturalOrder
  - Track (Spotify_youtube_sample/Track/Track)
    - dimension: Spotify_youtube_sample
    - hierarchy: Track
    - type: String
    - order: NaturalOrder
  - official_video (Spotify_youtube_sample/official_video/official_video)
    - dimension: Spotify_youtube_sample
    - hierarchy: official_video
    - type: boolean
    - order: NaturalOrder

▾ Play with the Cube

▾ Apex or 0-D

```
cube.query(m['Views.SUM'])
```

	Views.SUM
0	1,902,053,002,307.00

1D

```
cube.query(m['Views.SUM'], levels = [l['Artist']])
```

	Views.SUM
Artist	
$NOT	110,784,903.00
$uicideboy$	334,135,108.00
(G)I-DLE	1,754,953,941.00
*NSYNC	1,027,832,862.00
070 Shake	96,099,359.00
...	...
will.i.am	2,831,320,166.00
Ángela Aguilar	1,385,295,291.00
Ñejo	626,680,824.00
Ñengo Flow	812,726,315.00
Øneheart	34,623,310.00

2063 rows × 1 columns

```
cube.query(m['Views.SUM'], levels = [l['Album']])
```

	Views.SUM
Album	
!Volare! The Very Best of the Gipsy Kings	5,760,198.00
"Awaken, My Love!"	694,453,372.00
"Heroes" (2017 Remaster)	29,328,667.00
"Let Go" Dj Pack	56.00
"Let's Rock"	14,005,512.00
...	...
踊り子	52,395,337.00
離開地球表面	20,561,513.00
青と夏	34,397,912.00
黑·白·灰	5,961,450.00
龍が如く0 誓いの場所 オリジナルサウンドトラック (Side B)	748,694.00

11727 rows × 1 columns

▼ 2D

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Artist']])
```

Album	Artist	Views.SUM
!Volare! The Very Best of the Gipsy Kings	Gipsy Kings	5,760,198.00
"Awaken, My Love!"	Childish Gambino	694,453,372.00
"Heroes" (2017 Remaster)	David Bowie	29,328,667.00
"Let Go" Dj Pack	Dina Rae	56.00
"Let's Rock"	The Black Keys	14,005,512.00
...	...	...
踊り子	Vaundy	52,395,337.00
離開地球表面	Mayday	20,561,513.00
青と夏	Mrs. GREEN APPLE	34,397,912.00
黑·白·灰	Eason Chan	5,961,450.00
龍が如く0 誓いの場所 オリジナルサウンドトラック (Side B)	SEGA SOUND TEAM	748,694.00

13955 rows × 1 columns

▼ 2D with Slicing

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track']], filter=l['Artist'] =
```

		Views.SUM
Album	**Track**	
1 (Remastered)	**Hey Jude - Remastered 2015**	56,888,766.00
	I Want To Hold Your Hand - Remastered 2015	28,799,448.00
Abbey Road (Remastered)	**Come Together - Remastered 2009**	31,607,772.00
	Here Comes The Sun - Remastered 2009	114,971,785.00
	Something - Remastered 2009	96,849,734.00
Help! (Remastered)	**Yesterday - Remastered 2009**	42,856,675.00
Let It Be (Remastered)	**Let It Be - Remastered 2009**	131,925,109.00
Please Please Me (Remastered)	**Twist And Shout - Remastered 2009**	53,228,361.00
Rubber Soul (Remastered)	**In My Life - Remastered 2009**	63,223,118.00
The Beatles (Remastered)	**Blackbird - Remastered 2009**	33,852,949.00

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track']], filter=l['Artist'] =
```

		Views.SUM
Album	**Track**	
Bad (Remastered)	**Smooth Criminal - 2012 Remaster**	84,652,754.00
	The Way You Make Me Feel - 2012 Remaster	436,325,398.00
Dangerous	**Remember the Time**	504,057,332.00
HIStory - PAST, PRESENT AND FUTURE - BOOK I	**They Don't Care About Us**	978,680,025.00
Off the Wall	**Don't Stop 'Til You Get Enough**	349,102,050.00
	Rock with You - Single Version	301,845,823.00
Thriller	**Beat It**	905,692,299.00
	Billie Jean	1,300,732,161.00
XSCAPE	**Chicago**	545,655.00
	Love Never Felt So Good	343,328,494.00

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track']], filter=l['official_v
```

| | | Views.SUM |
Album	Track	
!Volare! The Very Best of the Gipsy Kings	Hotel California (Spanish Mix)	5,760,198.00
"Awaken, My Love!"	Me and Your Mama	86,929,807.00
	Redbone	607,523,565.00
"Heroes" (2017 Remaster)	Heroes - 2017 Remaster	29,328,667.00
"Let Go" Dj Pack	Let Go - Acapella	56.00
...	...	...
踊り子	踊り子	52,395,337.00
離開地球表面	私奔到月球	20,561,513.00
青と夏	点描の唄	34,397,912.00
黑·白·灰	十年	5,961,450.00
龍が如く0 誓いの場所 オリジナルサウンドトラック (Side B)	Friday Night	748,694.00

14327 rows × 1 columns

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track']], filter=l['official_v
```

		Views.SUM
Album	**Track**	
"Miguel"	**Te Amaré**	30,083,671.00
"和自己對話" 實驗專輯	**不為誰而作的歌**	156,510,054.00
#1s ... and then some	**Brand New Man**	33,246.00
$outh $ide $uicide	**Cold Turkey**	214,405.00
	Muddy Blunts	31,879.00
...	**...**	...
祝福	**祝福**	31,826,386.00
秋の惑星、ハートはナイトブルー。	**秋の惑星 -Opening-**	286,681.00
第二人生(明日版)	**我不願讓你一個人**	643,542.00
認了吧	**好久不見**	12,901,602.00
	淘汰	15,843,955.00

4237 rows × 1 columns

▼ 3D

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track'], l['official_video']])
```

			Views.SUM
Album	**Track**	**official_video**	
!Volare! The Very Best of the Gipsy Kings	**Hotel California (Spanish Mix)**	**true**	5,760,198.00
"Awaken, My Love!"	**Me and Your Mama**	**true**	86,929,807.00
	Redbone	**true**	607,523,565.00
"Heroes" (2017 Remaster)	**Heroes - 2017 Remaster**	**true**	29,328,667.00
"Let Go" Dj Pack	**Let Go - Acapella**	**true**	56.00
...	**...**	**...**	...
踊り子	**踊り子**	**true**	52,395,337.00
離開地球表面	**私奔到月球**	**true**	20,561,513.00
青と夏	**点描の唄**	**true**	34,397,912.00
黑·白·灰	**十年**	**true**	5,961,450.00
龍が如く0 誓いの場所 オリジナルサウンドトラック (Side B)	**Friday Night**	**true**	748,694.00

18564 rows × 1 columns

▾ 3D with Slicing

```
cube.query(m['Views.SUM'], levels = [l['Album'], l['Track'], l['official_video']],
```

			Views.SUM
Album	**Track**	**official_video**	
Bad (Remastered)	**Smooth Criminal - 2012 Remaster**	false	84,652,754.00
	The Way You Make Me Feel - 2012 Remaster	true	436,325,398.00
Dangerous	**Remember the Time**	true	504,057,332.00
HIStory - PAST, PRESENT AND FUTURE - BOOK I	**They Don't Care About Us**	true	978,680,025.00
Off the Wall	**Don't Stop 'Til You Get Enough**	true	349,102,050.00
	Rock with You - Single Version	true	301,845,823.00
Thriller	**Beat It**	true	905,692,299.00
	Billie Jean	true	1,300,732,161.00
XSCAPE	**Chicago**	false	545,655.00
	Love Never Felt So Good	true	343,328,494.00

Excecise

You can play with other interests in measure, and try different dimension of the cube.