

Loading package and connecting to database

In [2]: %load_ext sql

```
%sql mysql+pymysql://root:@localhost/chocolates #connecting to db names chocolates running on local mysql server
```

0 rows affected.

Out[2]: []

Tables content

In [13]: %SQL

```
SHOW TABLES;
```

* mysql+pymysql://root:***@localhost/chocolates
4 rows affected.

Out[13]:

Tables_in_chocolates
geo
people
products
sales

In [15]: %SQL

```
SELECT * FROM GEO LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
6 rows affected.

Out[15]:

GEOID	GEO	REGION
G1	India	APAC
G2	USA	Americas
G3	Canada	Americas
G4	New Zealand	APAC
G5	Australia	APAC
G6	UK	Europe

In [16]: %SQL

```
SELECT * FROM PRODUCTS LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[16]:

PID	PRODUCT	CATEGORY	SIZE	COST_PER_BOX
P01	Milk Bars	Bars	LARGE	1.52
P02	50% Dark Bites	Bites	LARGE	2.57
P03	Almond Choco	Bars	LARGE	9.6
P04	Raspberry Choco	Bars	LARGE	4.09
P05	Mint Chip Choco	Bars	LARGE	1.54
P06	Eclairs	Bites	LARGE	2.24
P07	Drinking Coco	Other	LARGE	1.62
P08	99% Dark & Pure	Bars	LARGE	2.47
P09	Orange Choco	Bars	LARGE	7.14
P10	Spicy Special Slims	Bites	LARGE	5.79

In [18]: %SQL

```
SELECT * FROM SALES LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[18]:

SPID	GEOID	PID	SALEDATE	AMOUNT	CUSTOMERS	BOXES
SP01	G4	P04	2021-01-01 00:00:00	8414	276	495
SP02	G3	P14	2021-01-01 00:00:00	532	317	54
SP12	G2	P08	2021-01-01 00:00:00	5376	178	269
SP01	G4	P15	2021-01-01 00:00:00	259	32	22
SP19	G2	P18	2021-01-01 00:00:00	5530	4	179
SP17	G1	P13	2021-01-01 00:00:00	2184	63	122
SP10	G6	P04	2021-01-01 00:00:00	1057	295	71
SP14	G5	P16	2021-01-01 00:00:00	1036	370	37
SP10	G5	P17	2021-01-01 00:00:00	4039	536	176
SP06	G4	P01	2021-01-01 00:00:00	12894	115	478

In [9]: %SQL

```
#WHERE
SELECT *
FROM SALES
WHERE AMOUNT > 10000
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[9]:

SPID	GEOID	PID	SALEDATE	AMOUNT	CUSTOMERS	BOXES
SP06	G4	P01	2021-01-01 00:00:00	12894	115	478
SP10	G1	P06	2021-01-01 00:00:00	15596	32	975
SP25	G6	P05	2021-01-04 00:00:00	14273	335	752
SP18	G2	P21	2021-01-04 00:00:00	19229	64	1013
SP23	G1	P16	2021-01-05 00:00:00	17248	163	664
SP05	G3	P02	2021-01-07 00:00:00	10451	155	1307
SP18	G1	P18	2021-01-07 00:00:00	11228	236	388
SP08	G3	P10	2021-01-08 00:00:00	12726	179	579
SP10	G2	P12	2021-01-12 00:00:00	11739	22	903
SP15	G2	P21	2021-01-13 00:00:00	12334	119	686

In [19]: %SQL

```
#ORDER BY
SELECT *
FROM SALES
WHERE AMOUNT > 12000
ORDER BY AMOUNT #DESC FOR DECREASING ORDER
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[19]:

SPID	GEOID	PID	SALEDATE	AMOUNT	CUSTOMERS	BOXES
SP20	G2	P13	2022-02-28 00:00:00	12005	33	1334
SP04	G5	P13	2021-11-11 00:00:00	12012	124	546
SP11	G4	P17	2021-03-15 00:00:00	12012	322	1092
SP11	G5	P21	2021-09-13 00:00:00	12019	111	752
SP02	G4	P19	2021-02-26 00:00:00	12026	166	926
SP06	G6	P22	2021-01-20 00:00:00	12033	373	482
SP05	G6	P17	2021-04-20 00:00:00	12040	255	416
SP11	G3	P08	2021-03-23 00:00:00	12040	53	860
SP05	G5	P14	2021-01-22 00:00:00	12047	295	927
SP11	G4	P10	2021-06-04 00:00:00	12047	202	603

In [20]: %SQL

```
#AND
SELECT *
FROM PRODUCTS
WHERE Category = "Bars"
AND SIZE = "LARGE"
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
7 rows affected.

Out[20]:

PID	PRODUCT	CATEGORY	SIZE	COST_PER_BOX
P01	Milk Bars	Bars	LARGE	1.52
P03	Almond Choco	Bars	LARGE	9.6
P04	Raspberry Choco	Bars	LARGE	4.09
P05	Mint Chip Choco	Bars	LARGE	1.54
P08	99% Dark & Pure	Bars	LARGE	2.47
P09	Orange Choco	Bars	LARGE	7.14
P12	Fruit & Nut Bars	Bars	LARGE	1.58

In [21]: %SQL

```
#OR
SELECT *
FROM PRODUCTS
WHERE Category = "Bars"
OR Category = "Bites"
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[21]:

PID	PRODUCT	CATEGORY	SIZE	COST_PER_BOX
P01	Milk Bars	Bars	LARGE	1.52
P02	50% Dark Bites	Bites	LARGE	2.57
P03	Almond Choco	Bars	LARGE	9.6
P04	Raspberry Choco	Bars	LARGE	4.09
P05	Mint Chip Choco	Bars	LARGE	1.54
P06	Eclairs	Bites	LARGE	2.24
P08	99% Dark & Pure	Bars	LARGE	2.47
P09	Orange Choco	Bars	LARGE	7.14
P10	Spicy Special Slims	Bites	LARGE	5.79
P11	After Nines	Bites	LARGE	4.43

In [24]: %SQL

```
#BETWEEN
SELECT *
FROM SALES
WHERE SALEDATE BETWEEN '2021-01-01' AND '2021-12-31'
ORDER BY BOXES DESC
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[24]:

SPID	GEOID	PID	SALEDATE	AMOUNT	CUSTOMERS	BOXES
SP18	G2	P16	2021-03-30 00:00:00	24367	272	3481
SP07	G4	P16	2021-01-21 00:00:00	19523	57	3254
SP06	G1	P09	2021-05-21 00:00:00	17451	135	2909
SP12	G1	P09	2021-02-04 00:00:00	18949	300	2707
SP22	G6	P18	2021-04-01 00:00:00	20832	78	2604
SP10	G4	P16	2021-02-22 00:00:00	15351	116	2559
SP24	G2	P16	2021-12-14 00:00:00	14959	211	2494
SP22	G1	P09	2021-10-13 00:00:00	20825	258	2314
SP08	G2	P16	2021-07-01 00:00:00	13720	209	2287
SP02	G1	P16	2021-07-22 00:00:00	15848	2	2264

In [34]: %SQL

```
#LIKE
SELECT *
FROM PEOPLE
WHERE SALESPERSON LIKE '%B%' #RETURNS ALL NAMES STARTING WITH B
LIMIT 10;
```

% REPRESENTS ANY NUMBER OF CHARACTERS
_ REPRESENTS A SINGLE CHARACTER
THESE CHARACTERS ARE CALLED WILD CARDS IN SQL

* mysql+pymysql://root:***@localhost/chocolates
4 rows affected.

Out[34]:

SPID	SALESPERSON	TEAM	LOCATION
SP01	Barr Faughny	Yummies	Hyderabad
SP10	Brien Boise	Jucies	Wellington
SP19	Beverie Moffet	Jucies	Seattle
SP32	Benny Karolovsky	Jucies	Paris

In [36]: %SQL

```
#LIKE
SELECT *
FROM PEOPLE
WHERE SALESPERSON LIKE '%B____%' #RETURNS ALL NAMES THAT INCLUDE B FOLLOWED BY AT LEAST 4 CHARACTERS
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[36]:

SPID	SALESPERSON	TEAM	LOCATION
SP01	Barr Faughny	Yummies	Hyderabad
SP05	Gigi Bohling	Delish	Hyderabad
SP08	Ches Bonnell	Hyderabad	
SP10	Brien Boise	Jucies	Wellington
SP16	Marney O'Brien	Jucies	Wellington
SP17	Rafaelita Blaksland		Wellington
SP19	Beverie Moffet	Jucies	Seattle
SP20	Oby Sorrel	Jucies	Seattle
SP28	Ebonee Roxburgh		Paris
SP32	Benny Karolovsky	Jucies	Paris

In [38]: %SQL

```
#AS
SELECT *, SALESPERSON AS NAME # AS USED TO ASSIGN A CUSTOM NAME TO COLUMN
FROM PEOPLE
WHERE SALESPERSON LIKE '%B____%'
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[38]:

SPID	SALESPERSON	TEAM	LOCATION	NAME
SP01	Barr Faughny	Yummies	Hyderabad	Barr Faughny
SP05	Gigi Bohling	Delish	Hyderabad	Gigi Bohling
SP08	Ches Bonnell		Hyderabad	Ches Bonnell
SP10	Brien Boise	Jucies	Wellington	Brien Boise
SP16	Marney O'Brien	Jucies	Wellington	Marney O'Brien
SP17	Rafaelita Blaksland		Wellington	Rafaelita Blaksland
SP19	Beverie Moffet	Jucies	Seattle	Beverie Moffet
SP20	Oby Sorrel	Jucies	Seattle	Oby Sorrel
SP28	Ebonee Roxburgh		Paris	Ebonee Roxburgh
SP32	Benny Karolovsky	Jucies	Paris	Benny Karolovsky

In [43]: %SQL

```
#DISTINCT - RETURNS DISTINCT VALUES IN RESULT SET
SELECT DISTINCT(LOCATION)
FROM PEOPLE
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
4 rows affected.

Out[43]:

LOCATION
Hyderabad
Wellington
Seattle
Paris

In [5]: %SQL

```
#GROUP BY
SELECT SUM(AMOUNT), SUM(BOXES)
FROM SALES
GROUP BY GEOID;
```

* mysql+pymysql://root:***@localhost/chocolates
6 rows affected.

Out[5]:

SUM(AMOUNT)	SUM(BOXES)
7310254	490374
7012523	473759
7350091	491482
7435918	493139
7263151	482536
7189609	470021

In [12]: %SQL

```
#GROUP BY USED TO CALCULATE REVENUE OF EACH SALESPERSON IN EACH GEO LOCATION
SELECT (AMOUNT * BOXES) AS REVENUE, GEOID , SPID
FROM SALES
GROUP BY SPID, GEOID
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[12]:

REVENUE	GEOID	SPID
228984	G1	SP01
389277	G2	SP01
17150	G3	SP01
4164930	G4	SP01
2279970	G5	SP01
43470	G6	SP01
419286	G1	SP02
14081508	G2	SP02
28728	G3	SP02
370888	G4	SP02

In [17]: %SQL

```
#GROUP BY USED TO GET REVENUE OF EACH TEAM
SELECT (S.AMOUNT * S.BOXES) AS REVENUE, P.TEAM
FROM SALES AS S
JOIN PEOPLE AS P ON S.SPID = P.SPID
WHERE TEAM != ""
GROUP BY P.TEAM;
```

* mysql+pymysql://root:***@localhost/chocolates
3 rows affected.

Out[17]:

REVENUE	TEAM
38332	Delish
989870	Jucies
4164930	Yummies

In [5]: %SQL

```
#having clause used to enforce condition after execution of group by clause
SELECT PID, SUM(AMOUNT) AS total_sales
FROM SALES
GROUP BY PID
HAVING SUM(AMOUNT) > 1000
LIMIT 10;
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[5]:

PID	total_sales
P01	1989078
P02	1991836
P03	2055627
P04	2090242
P05	1871751
P06	1955975
P07	1941562
P08	2023070
P09	1899212
P10	2004415

In [3]: %SQL

```
#NULL FUNCTIONS
SELECT * FROM PEOPLE WHERE TEAM IS NULL;
```

* mysql+pymysql://root:***@localhost/chocolates
5 rows affected.

Out[3]:

SPID	SALESPERSON	TEAM	LOCATION
SP08	Ches Bonnell	None	Hyderabad
SP17	Rafaelita Blaksland	None	Wellington
SP21	Mallorie Waber	None	Seattle
SP22	Jehu Rudeforth	None	Seattle
SP28	Ebonee Roxburgh	None	Paris

In [21]: %SQL

```
#DECLARING A PROCEDURE
CREATE PROCEDURE GetSalesByPID(IN product_id TEXT)
BEGIN
    SELECT SPID, GEOID, PID, SALEDATE, AMOUNT, CUSTOMER
    FROM SALES
    WHERE PID = product_id
    LIMIT 10;
END
```

* mysql+pymysql://root:***@localhost/chocolates
0 rows affected.

Out[21]: []

In [22]: %SQL

```
#CALLING THE PROCEDURE
CALL GetSalesByPID('P01');
```

* mysql+pymysql://root:***@localhost/chocolates
10 rows affected.

Out[22]:

SPID	GEOID	PID	SALEDATE	AMOUNT	CUSTOMERS	BOXES
SP06	G4	P01	2021-01-01 00:00:00	12894	115	478
SP18	G2	P01	2021-01-01 00:00:00	4669	121	180
SP03	G6	P01	2021-01-06 00:00:00	2709	27	101
SP10	G2	P01	2021-01-08 00:00:00	420	196	14
SP23	G4	P01	2021-01-11 00:00:00	2709	188	113
SP01	G5	P01	2021-01-11 00:00:00	7238	67	315
SP11	G6	P01	2021-01-12 00:00:00	189	123	8
SP12	G6	P01	2021-01-14 00:00:00	6881	10	222
SP17	G2	P01	2021-01-14 00:00:00	10339	207	357
SP10	G1	P01	2021-01-18 00:00:00	238	364	11

In []: %SQL