

mysqldtopython

리눅스 mysql 에서 workbench 및 파이썬으로 테이블 내보내기

최종으로 내보낼 테이블 정리

hdfs에서 받아온 pltest2

```
mysql> select * from pltest2
-> ;
```

crimeclassification	seoul	busan	daegu	incheon	gwangju	daejeon	ulsan	sejong	gyeonggido	gangwondo	chungcheongdo	Jeollado	gyeongsangdo	jeju	othercities	nonurbanareas
마약범죄	1360	759	393	557	69	78	67	2	1399	110	299	105	662	26	16	509
보건의범죄	3941	1832	417	654	760	204	230	15	3131	232	421	578	1140	168	36	843
환경범죄	53	198	19	69	8	13	53	2	1192	57	140	143	429	46	30	503
교통범죄	74958	35841	39725	31163	23412	13556	16600	1056	141506	18416	34076	33453	72263	14438	2325	43877
노동범죄	215	190	31	28	13	14	63	0	243	24	28	87	154	9	1	45
안보범죄	38	4	4	20	3	4	2	0	16	1	13	3	6	0	1	6
선거범죄	24	24	18	19	9	8	15	5	83	15	51	63	133	23	3	267
병역범죄	4591	959	800	1200	371	630	382	86	4595	265	1112	691	1839	242	155	808
기타범죄	41752	19868	9766	13677	5301	4602	4881	462	45409	4548	11530	11814	18988	4461	1341	17797
강력범죄	67	20	13	17	11	8	12	2	89	10	22	14	32	11	7	30
강력범죄	96	41	23	34	7	16	11	1	128	14	45	26	57	17	4	44
강력범죄	274	161	60	103	44	64	29	0	300	26	78	70	135	33	4	65
강력범죄	969	383	216	304	215	181	132	3	1134	108	278	300	503	78	18	328
강력범죄	118	24	21	35	17	14	10	1	118	22	32	21	39	13	0	33
강력범죄	4282	991	645	924	495	362	293	34	3507	274	729	587	947	197	28	764
강력범죄	94	37	20	31	22	12	9	0	123	23	32	26	60	12	2	55
강력범죄	315	117	86	94	48	41	36	5	416	27	80	75	146	35	4	148
절도범죄	55280	20136	12759	10679	8438	9495	5449	638	50522	4699	15457	11712	23962	4230	643	11754
폭력범죄	7174	4675	2736	2179	1162	1516	1353	117	8493	1858	2945	2614	6292	801	149	3941
폭력범죄	36904	8852	6607	9774	5386	4054	4213	139	37588	3095	6150	7188	10764	2489	363	5712
폭력범죄	186	67	39	81	26	36	20	3	217	32	59	66	93	12	4	87
폭력범죄	1844	682	443	433	297	286	175	29	2326	217	402	411	761	290	44	547

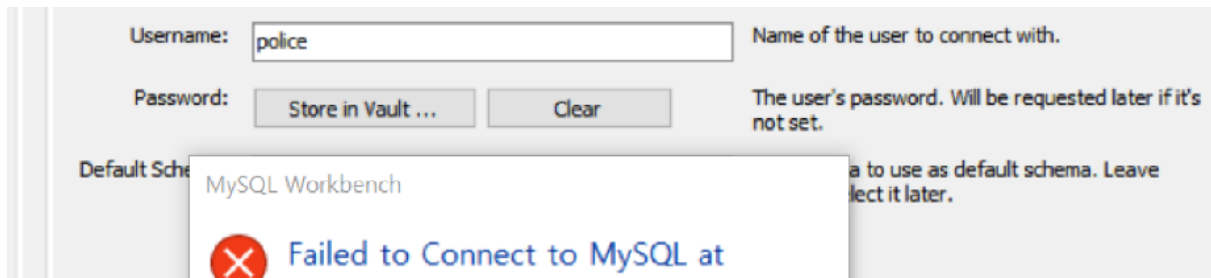
```
mysql> SELECT crimeclassification, SUM(seoul) AS seoul_total, SUM(busan)
AS busan_total, SUM(daegu) AS daegu_total, SUM(incheon) AS incheon_total,
SUM(gwangju) AS gwangju_total, SUM(daejeon) AS daejeon_total, SUM(ulsan)
AS ulsan_total, SUM(sejong) AS sejong_total, SUM(gyeonggido) AS
gyeonggido_total, SUM(gangwondo) AS gangwondo_total
```

- FROM pltest2
- GROUP BY crimeclassification

→ order by crimeclassification;

crimeclassification	seoul_total	busan_total	daegu_total	incheon_total	gwangju_total	daejeon_total	ulsan_total	sejong_total	gyeonggido_total	gangwondo_total
강력범죄	6215	1774	1084	1542	860	698	532	46	5815	504
교통범죄	74958	35841	39725	31163	23412	13556	16600	1056	141506	18416
기타범죄	41752	19868	9766	13677	5301	4602	4881	462	45409	4548
노동범죄	215	190	31	28	13	14	63	0	243	24
마약범죄	1360	759	393	557	69	78	67	2	1399	110
병역범죄	4591	959	800	1200	371	630	382	86	4595	265
보건의범죄	3941	1832	417	654	760	204	230	15	3131	232
선거범죄	24	24	18	19	9	8	15	5	83	15
안보범죄	38	4	4	20	3	4	2	0	16	1
절도범죄	55280	20136	12759	10679	8438	9495	5449	638	50522	4699
지능범죄	72851	25247	14195	16938	9724	9112	6898	680	71370	5655
특별경제범죄	22075	13868	3421	5751	2535	1635	1289	187	17679	1307
폭력범죄	65247	20928	14708	18054	9602	8839	8277	452	71930	8036
중속범죄	6741	2085	1418	1580	592	689	403	29	5306	309
환경범죄	53	198	19	69	8	13	53	2	1192	57

이 테이블을 사용할 예정... 이었으나



접근이 거부 당해서 권한 확인

```
mysql> show grants for police@localhost;
+-----+
| Grants for police@localhost |
+-----+
| GRANT USAGE ON *.* TO `police`@`localhost` |
| GRANT ALL PRIVILEGES ON `crime`.* TO `police`@`localhost` |
+-----+
```

부여되어 있어서 계정 재생성 리눅스의 ip인 192.168.111.100으로 생성

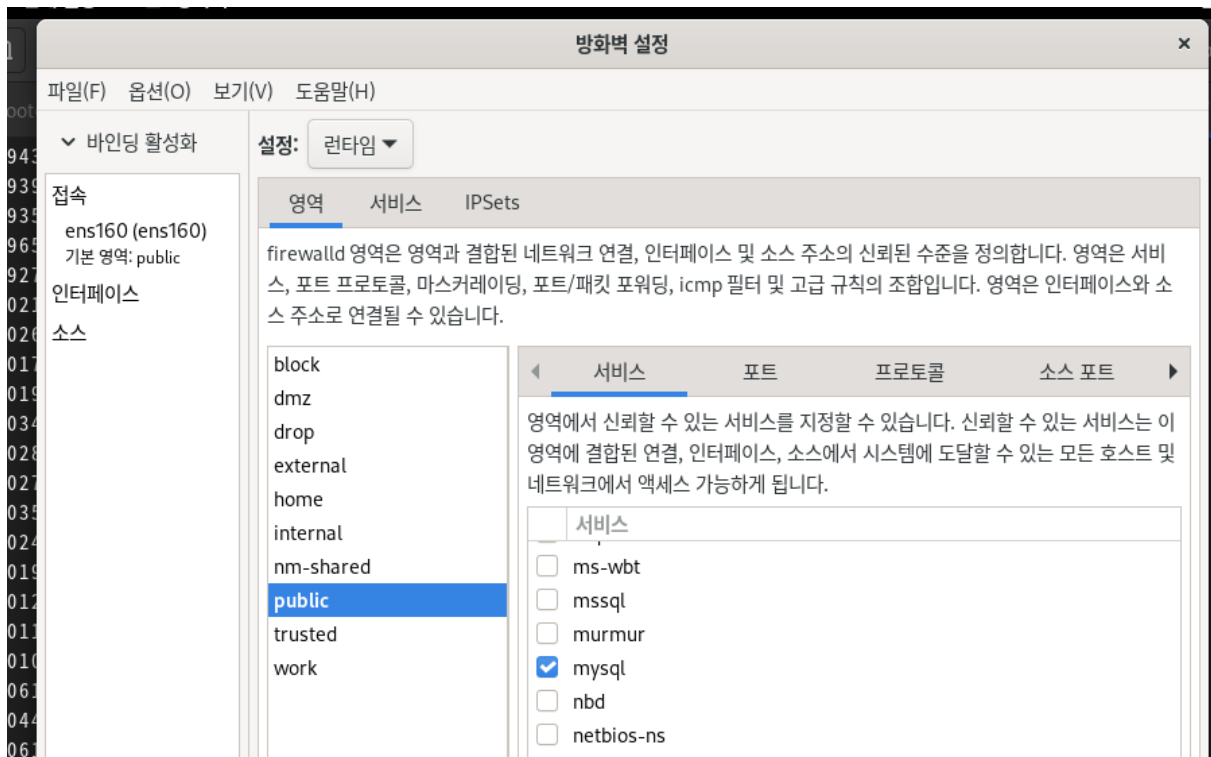
```
mysql> create user testuser@'192.168.111.%' identified by 'password';
Query OK, 0 rows affected (0.09 sec)

mysql> grant all privileges on *.* to testuser@'192.168.111.%';
Query OK, 0 rows affected (0.04 sec)
```

192.168.111.100 아이피의 입력을 하면 워크벤치가 멈춤

→ 방화벽 체크

firewall-config



같은 이유로 데이터를 불러올 수 없어서 시각화 할때는 데이터 프레임을 직접 입력했음

```
data = {
    "도시": ["서울특별시", "부산광역시", "대구광역시", "인천광역시", "광주광역시", "대전광역시", "울산광역시", "세종특별자치시", "경기도", "강원특별자치도", "충청북도", "충청남도", "전북특별자치도", "전라남도", "경상북도", "경상남도", "제주특별자치도"],
    "강력범죄": [6215, 1774, 1084, 1542, 860, 698, 532, 46, 5815, 504, 541, 755, 664, 455, 832, 1087, 396],
    "교통범죄": [74958, 35841, 39725, 31163, 23412, 13556, 16600, 1056, 141506, 18416, 14713, 19363, 16007, 17446, 32309, 39954, 14438],
    "기타범죄": [41752, 19868, 9766, 13677, 5301, 4602, 4881, 462, 45409, 4548, 3978, 7552, 6122, 5692, 7403, 11585, 4461],
    "노동범죄": [215, 190, 31, 28, 13, 14, 63, 0, 243, 24, 9, 19, 45, 42, 53, 101, 9],
    "마약범죄": [1360, 759, 393, 557, 69, 78, 67, 2, 1399, 110, 154, 145, 48, 57, 204, 458, 26],
    "병역범죄": [4591, 959, 800, 1200, 371, 630, 382, 86, 4595, 265, 391, 721, 381, 310, 772, 1067, 242],
}
```

```

"보건범죄": [3941, 1832, 417, 654, 760, 204, 230, 15, 3131, 232, 196, 225, 410,
168, 405, 735, 168],
"선거범죄": [24, 24, 18, 19, 9, 8, 15, 5, 83, 15, 18, 33, 34, 29, 63, 70, 23],
"안보범죄": [38, 4, 4, 20, 3, 4, 2, 0, 16, 1, 0, 13, 1, 2, 1, 5, 0],
"절도범죄": [55280, 20136, 12759, 10679, 8438, 9495, 5449, 638, 50522, 4699,
6140, 9317, 7448, 4264, 9456, 14506, 4230],
"지능범죄": [72851, 25247, 14195, 16938, 9724, 9112, 6898, 680, 71370, 5655,
5886, 9766, 8084, 6195, 11519, 17259, 4391],
"특별경제범죄": [22075, 13868, 3421, 5751, 2535, 1635, 1289, 187, 17679, 1307,
1045, 1776, 1417, 1063, 2400, 2972, 1251],
"폭력범죄": [65247, 20928, 14708, 18054, 9602, 8839, 8277, 452, 71930, 8036,
7160, 7993, 8181, 6840, 11317, 16055, 5518],
"풍속범죄": [6741, 2085, 1418, 1580, 592, 689, 403, 29, 5306, 309, 449, 648,
636, 413, 752, 865, 407],
"환경범죄": [53, 198, 19, 69, 8, 13, 53, 2, 1192, 57, 50, 90, 80, 63, 123, 306, 46]
}

df = pd.DataFrame(data)
df.set_index("도시", inplace=True)

```

이후 'root' 문제였던 권한부여 해결

GRANT ALL PRIVILEGES on . TO 'root'@'%' WITH GRANT OPTION

1 • select * from p1test;

	crimedassification	seoul	busan	daegu	incheon	gwangju	daejeon	ulsan	sejong	gyeonggido	gangwondo	chungcheongdo	Jeollado	gyeongsangdo	jeju	othercities	nonurbanareas
▶	미약범죄	1360	759	393	557	69	78	67	2	1399	110	299	105	662	26	16	509
	병역범죄	4591	959	800	1200	371	630	382	86	4595	265	1112	691	1839	242	155	808
	보안범죄	3941	1832	417	654	760	204	230	15	3131	232	421	578	1140	168	36	843
	언거범죄	24	24	18	19	9	8	15	5	83	15	51	63	133	23	3	267
	강력범죄	6215	1774	1084	1542	860	698	532	46	5815	504	1296	1119	1919	396	67	1467
	기타범죄	74958	35841	39725	31163	23412	13556	16600	1056	141506	18416	34076	33453	72263	14438	2325	43877
	교통범죄	41752	19868	9766	13677	5301	4602	4881	462	45409	4548	11530	11814	18988	4461	1341	17797
	노동범죄	215	190	31	28	13	14	63	0	243	24	28	87	154	9	1	45
	안보범죄	38	4	4	20	3	4	2	0	16	1	13	3	6	0	1	6
	절도범죄	55280	20136	12759	10679	8438	9495	5449	638	50522	4699	15457	11712	23962	4230	643	11754
	지능범죄	72851	25247	14195	16938	9724	9112	6898	680	71370	5655	15652	14279	28778	4391	2495	17856
	특별경제범죄	22075	13868	3421	5751	2535	1635	1289	187	17679	1307	2821	2480	5372	1251	1631	3027
	폭력범죄	65247	20928	14708	18054	9602	8839	8277	452	71930	8036	15153	15021	27372	5518	825	15985
	폭속범죄	6741	2085	1418	1580	592	689	403	29	5306	309	1097	1049	1617	407	119	1050
	환경범죄	53	198	19	69	8	13	53	2	1192	57	140	143	429	46	30	503

그런데도 colab python은 연동불가

```

pymysql.connect(db='crime', user='root', passwd='password', host='192.168.111.100', port=3306)

-----
TimeoutError                                Traceback (most recent call last)
/usr/local/lib/python3.10/dist-packages/pymysql/connections.py in connect(self, sock)
    643                                     try:
--> 644                                     sock = socket.create_connection(
    645                                     (self.host, self.port), self.connect_timeout, **kwargs)

-----
4 frames -----

TimeoutError: timed out

During handling of the above exception, another exception occurred:

OperationalError                                Traceback (most recent call last)
/usr/local/lib/python3.10/dist-packages/pymysql/connections.py in connect(self, sock)
    709                     if DEBUG:
    710                         print(exc.traceback)
--> 711                     raise exc
    712
    713                     # If e is neither DatabaseError or IOError, It's a bug.

OperationalError: (2008, "Can't connect to MySQL server on '192.168.111.100' (timed out)")

```

그런데 jupyter lab에서는 연동이 된다

```

[11]: pymysql.connect(db='crime', user='root', passwd='password', host='192.168.111.100', port=3306)

[3]: <pymysql.connections.Connection at 0x1e72778a810>

[12]: from sqlalchemy import create_engine, MetaData, Table
import pymysql
# MySQL 연결 정보 설정
db_user = 'root'
db_password = 'password'
db_host = '192.168.111.100'
db_port = '3306'
db_name = 'crime'

# MySQL 연결 엔진 생성
engine = create_engine(f'mysql+pymysql://{db_user}:{db_password}@{db_host}:{db_port}/{db_name}')

# 테이블 메타데이터 가져오기
metadata = MetaData()
metadata.reflect(engine)

# 테이블 객체 가져오기
pltest_table = Table('pltest', metadata, autoload=True, autoload_with=engine)

[13]: pltest_table
[13]: Table('pltest', MetaData(), Column('crimeclassification', VARCHAR(length=255), table=pltest), Column('seoul', VARCHAR(length=255), table=pltest), Column('busan', VARCHAR(length=255), table=pltest), Column('daegu', VARCHAR(length=255), table=pltest), Column('incheon', VARCHAR(length=255), table=pltest), Column('gwangju', VARCHAR(length=255), table=pltest), Column('daejeon', VARCHAR(length=255), table=pltest), Column('ulsan', VARCHAR(length=255), table=pltest), Column('sejong', VARCHAR(length=255), table=pltest), Column('gyeonggido', VARCHAR(length=255), table=pltest), Column('gangwondo', VARCHAR(length=255), table=pltest), Column('chungcheongdo', VARCHAR(length=255), table=pltest), Column('Jeollado', VARCHAR(length=255), table=pltest), Column('gyeongsangdo', VARCHAR(length=255), table=pltest), Column('jeju', VARCHAR(length=255), table=pltest), Column('othercities', VARCHAR(length=255), table=pltest), Column('nonurbanareas', VARCHAR(length=255), table=pltest), schema=None)

```

이렇게 테이블을 가져왔다.