



1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ

# 経済に関するオルタナティブ・データを考慮した金融マーケット予測手法の開発

Development of Financial Market Forecasting Method  
Considering Alternative Data on Economy

Itaru Aso

Graduate School of Information Engineering, Toyama Prefectural University  
t855001@st.pu-toyama.ac.jp

Toyama Prefectural Univ.



# 1.1 取り組み

## 取り組み内容

1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ

- 1 MySQL のインストール
- 2 MySQL の基本操作
- 3 Python で MySQL を操作
- 4 Python を用いて csv ファイルを MySQL にインポート



## 2.1 MySQL

1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ

### MySQL

世界中で利用されているリレーショナルデータベース

### MySQL の特徴

- ① オープンソースなため無料
- ② 大量のデータを高速に取り扱うことができる
- ③ 便利なツールが充実している



## 2.1 MySQL を活用する理由

### 理由

Twitter からツイートと日付のデータの容量が大きい  
集計や抽出するだけでも時間がかかるてしまう  
そこで、データベースを用いて取得したデータを管理する必要がある

1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ



## 3.1 基本操作

1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ

```
|mysql> select User,Host from mysql.user;
+-----+-----+
| User | Host |
+-----+-----+
| twitter | % |
| user | % |
| mysql.infoschema | localhost |
| mysql.session | localhost |
| mysql.sys | localhost |
| root | localhost |
| test_user | localhost |
+-----+-----+
7 rows in set (0.03 sec)
```

Figure: 1 ターミナルでの操作

ターミナル上で操作も可能である  
最終的にはデータベースからデータを取り出し、そのまま分析を行いたいため Python を用いて操作を行う

```
import mysql.connector

db=mysql.connector.connect(host="localhost", user="trump_tw", password="trump")

cursor=db.cursor()

cursor.execute("USE trump_tweet")
db.commit()
```



## 3.2 問題点

1. 本研究について
2. MySQL
3. Python を用いた MySQL
4. クラスタリングにおける学習データ

### 問題点

取得した Twitter のタイムスタンプは、MySQL の DATETIME 型のフォーマットにあっていない

MySQL のフォーマットに合わせてデータベースに挿入する必要がある

```
date = 'Sun Jul 29 20:12:15 +0000 2018'  
cursor.execute("SELECT str_to_date(%s, '%a %b %d %T %f %Y')" % date)  
#cursor.execute("SELECT str_to_date('Sun Jul 29 20:12:15 +0000 2018', '%a %b %d %T %f %Y')")  
rows = cursor.fetchall()
```

Figure: 1 Python での操作