

LT: データサイエンス老人会

Lightning Talk: Old Data Science Memories

30th May 2025, みんなのPython勉強会 #114

Yuta Kanzawa @yutakanzawa

Start
Python
Club
#stapy

Senior Data Scientist at Zurich Insurance Company Limited, Japan Branch



神沢雄大 Yuta Kanzawa

- データサイエンティスト@チューリッヒ保険会社



- 日本支店

- Twitter: [@yutakanzawa](https://twitter.com/yutakanzawa)

- 好きなもの：オペラとワイン

- ワーグナー
 - ブルゴーニュ (WSET Lv 3→?)

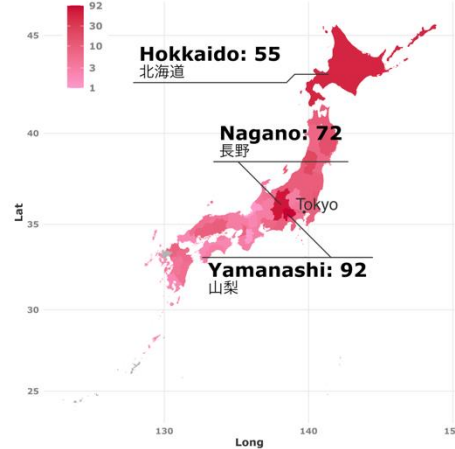
- 使用可能言語：7

- 人間：日本語、英語、ドイツ語
 - コンピューター：R, Python, SAS, SQL



ポートフォリオ

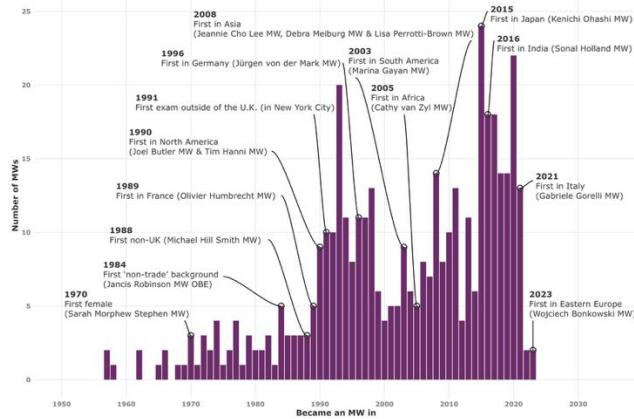
Number of Wineries in Japan in 2022, by Prefecture



Data: National Tax Agency Japan via https://www.nta.go.jp/taxes/saisei/shihei-gakkyo/saisei_onogakuryu/05/pdf/05.pdf
©Graphic: Yuta Kanzawa

415 Active Masters of Wine by Year of Qualification

As of May 2023, 500 people have gained the title since the inaugural exam in May 1953. NB: 85 deceased or resigned MWs are not counted here.



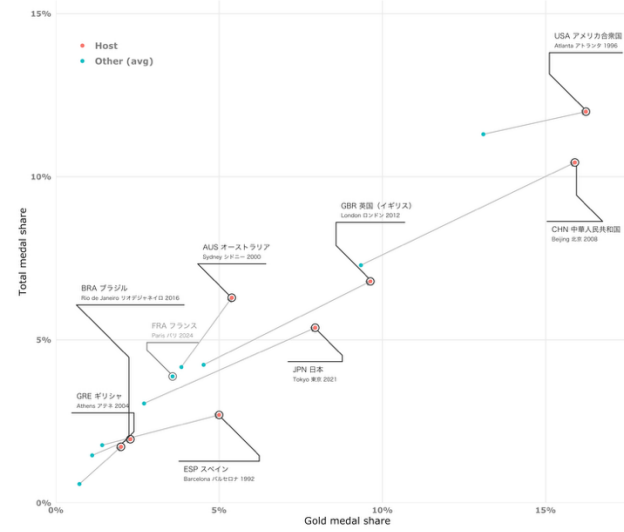
Data: The Institute of Masters of Wine via <https://www.mastersofwine.org> - Graphic: Yuta Kanzawa

Number of Qualified JSA Sommelier Excellence and Equivalents* by Year and Gender, 2013-2020



Source: Japan Sommelier Association <https://www.sommelier.jp/exam/pdf/qualifiedholders.pdf>
*Sommelier Excellence (2019-2020), Senior Sommelier (2013-2018), Senior Wine Adviser (2013-2015)

Medal shares of Olympic host countries in the past 30 years



Data: International Olympic Committee via <https://olympics.com> & <https://www.wikipedia.org> - Graphic: Yuta Kanzawa

Number of Active MWs by Region Based in
70% of active MWs are based in Europe (mostly Western Europe). NB: Some MWs are multi-based.



Data: The Institute of Masters of Wine via <https://www.mastersofwine.org> - Graphic: Yuta Kanzawa

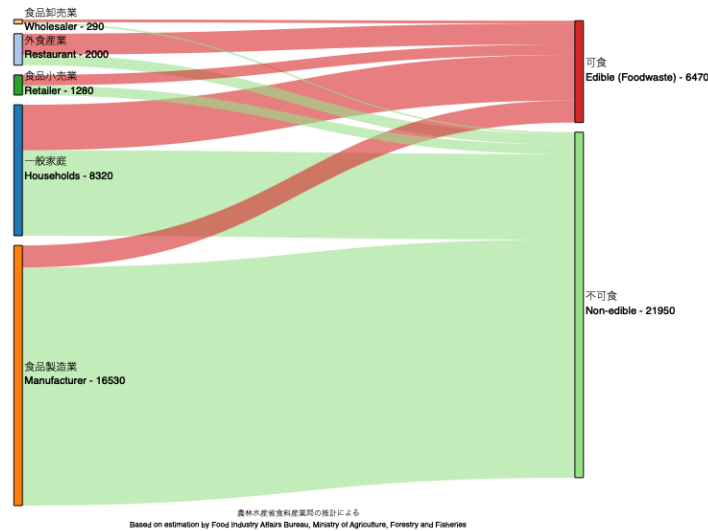
Daily maximum temperature in Tokyo, 1875-2021



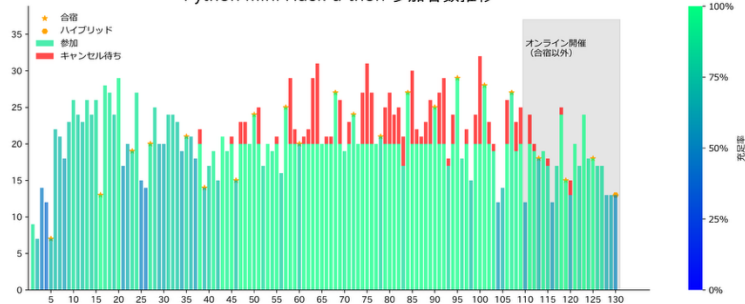
Data: Japan Meteorological Agency via <https://www.jma.go.jp> - Graphic: Yuta Kanzawa (Inspired by Cédric Scherer)

ポートフォリオ（参考までにR以外も）

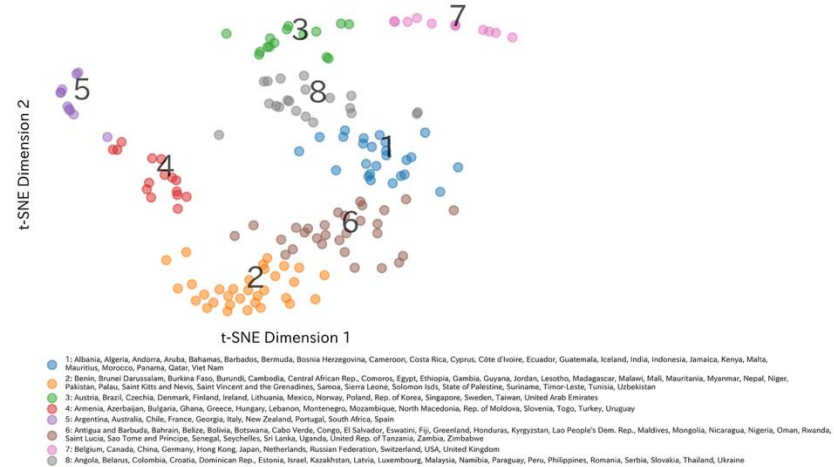
日本の食品廃棄物等の発生量（平成27年度推計） Estimated Food Disposals in Japan (FY2015)



Python mini Hack-a-thon 参加者数推移



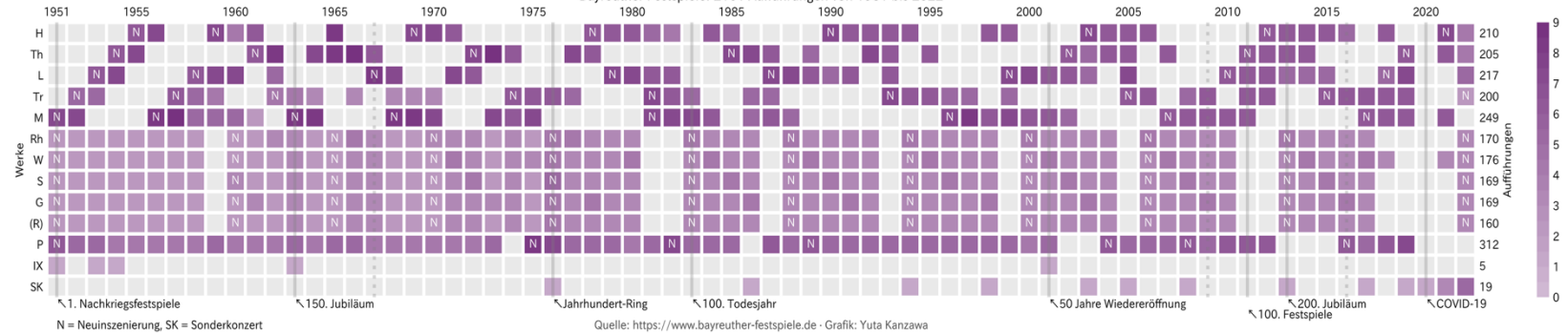
Clustering of Countries and Regions by Wine Trade Values & Production/Consumption Volumes in 2017 using t-SNE and K-Means



Sources: UN Comtrade (<https://comtrade.un.org>), FAOSTAT (<http://www.fao.org/faostat/>)



Bayreuther Festspiele: 2101 Aufführungen von 1951 bis 2022



10周年おめでとうございます！！



個人的な10年間

初参加！

2015年7月2日

業務のためのPython勉強会#2
『日々の仕事にPythonをどう活用できるのか？を考える集まりです。』



勉強会#65



1010000
1111001 ▶!!!

勉強会#89

キャリアづくり



トーク『わたしとデータサイエンス
とあなた：いままでとこれから』

2023年1月19日

みんなのPython勉強会#89

『データサイエンティストのキャリアづくり』

トーク『データサイエンスのための
リーダブルコードのススメ』

2021年1月13日

みんなのPython勉強会#65

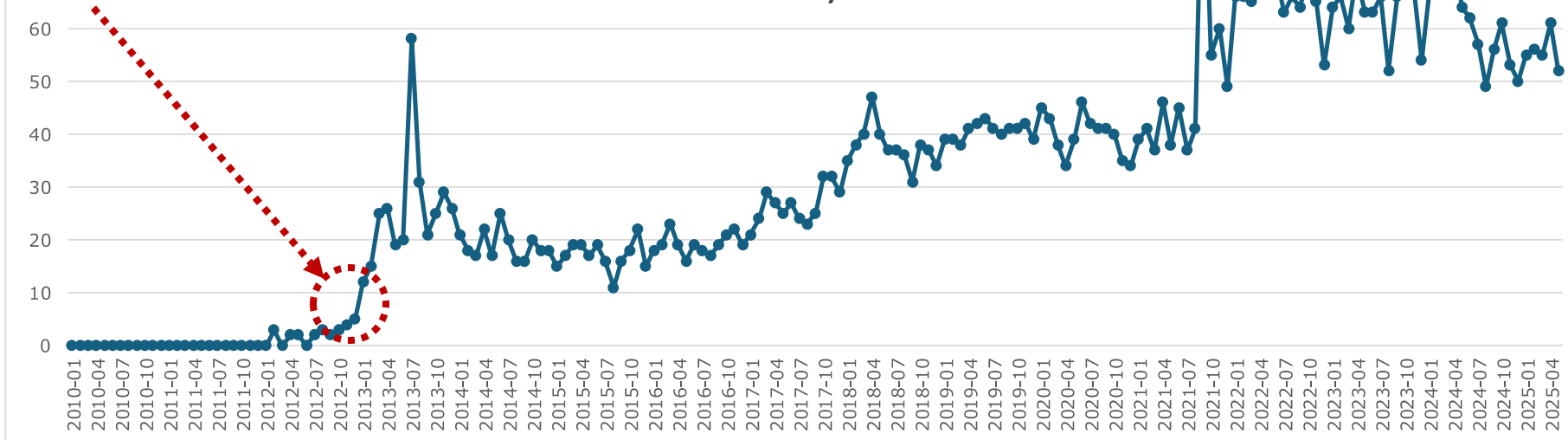
『分析・コード・共創』

「データサイエンティスト」



Googleトレンド「データサイエンティスト」（日本）

'Data Scientist: The Sexiest Job of the 21st Century'



期間中の最大値を100としたグラフ

- * 正確に訳すと「魅力的な」、「素敵な」（邦題）。
- * <https://store.hbr.org/product/harvard-business-review-october-2012/br1210>
- * <https://dhbr.diamond.jp/ud/backnumber/50ebdd876a8d1e3c6a000001>

@yutakanzawa



2010年

- 『Pythonスタートブック』初版（辻）



Googleなどの**Webサービスで人気**のスクリプト言語「Python」は、現在の言語に共通する、美しく模範的なプログラミングの学習に最適な言語です。本書は、プログラミングの知識がゼロの方でも理解できる、いちばん簡単なPythonの学習書です。プログラミングの基本の基本から、データ型を駆使した本格プログラムまで、豊富なイラストとサンプルを使って、Pythonならではのプログラミングのノウハウを楽しく解説します。

2013年

- ビッグデータ
- 深層学習（ディープラーニング）
- 『統計学が最強の学問である』（西内）
- 『データサイエンティスト養成読本』（佐藤、原田、下田、大成、奥野、中川、橋本、里、和田、早川、倉橋）



2015年

・「業務のためのPython勉強会」爆誕！

5月
27

業務のためのPython勉強会

日々の仕事にPythonをどう活用できるのか？を考える集まりです。



| | | |
|------|--------------|---------------|
| 募集内容 | 参加する 無料 | 先着順 50/20人 |
| | 発表したい！ 無料 | 先着順 0/2人 |



プログラミングでは無いお仕事をされている方々が、日頃のお仕事をPythonでどう効率化できるか？また、Pythonを使ったデータ解析がどのように業務の役に立つのか？を中心に考えていきます。現時点では、全3回を予定。（2回以降の日程は未定です）

18:30～オリエンテーション（辻、阿久津）10min

勉強会の目的、方向性を伝え、参加者で意識を共有する。スケジュール、運営スタッフの紹介。

Talk1：Pythonでプログラムを書こう！（辻）20min

Pythonまわりの便利なツールなどを紹介して、非プログラマーをPythonに誘う。

Talk2：私のPython学習奮闘記（仮）（阿久津）20min

私の独学歴を紹介。オープンラーニング（MIT6.00.1xなど）、教科書の紹介。Python学習の悩みを初学者の視点で伝える。

Talk3：データサイエンスとデータエンジニア（草薨）20min

- ・ 自己紹介
- ・ 業界の話
- ・ IT業界のトレンド
- ・ なぜ今データサイエンスか
- ・ データサイエンティストってどういう職業？
- ・ データエンジニアってどういう職業？
- ・ 技術の話
- ・ データサイエンティストに求められるスキル
- ・ データエンジニアに求められるスキル
- ・ 大規模なデータを扱う時に重要なこと
- ・ Hadoopって？
- ・ Sparkって？
- ・ 分析と機械学習
- ・ ビジネスの話
- ・ ビジネスに分析を生かしている企業
- ・ よくある悩み
- ・ 分析をビジネスに活用するために重要なこと：組織とプロセス

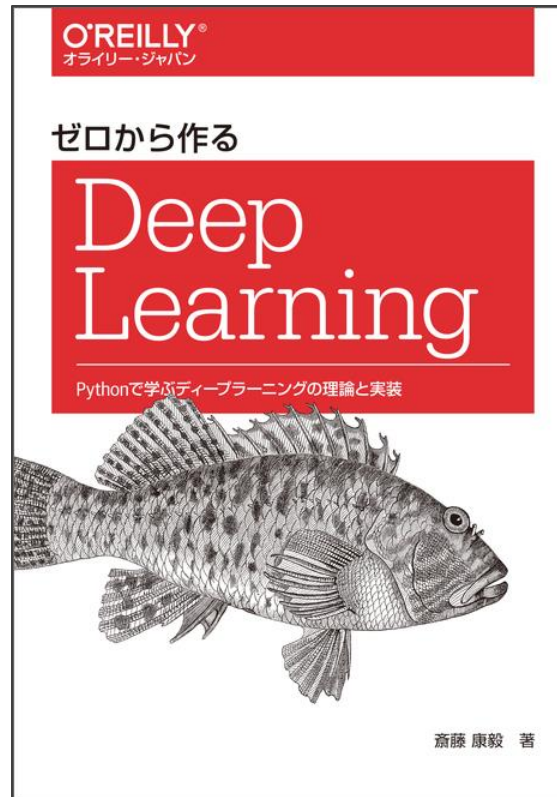
少し休憩

19:45頃～参加者の皆さんの自己紹介と、今後この回で検討する課題を決める

PCなどをご持参ください。無線LAN環境はこちらで用意する予定です。

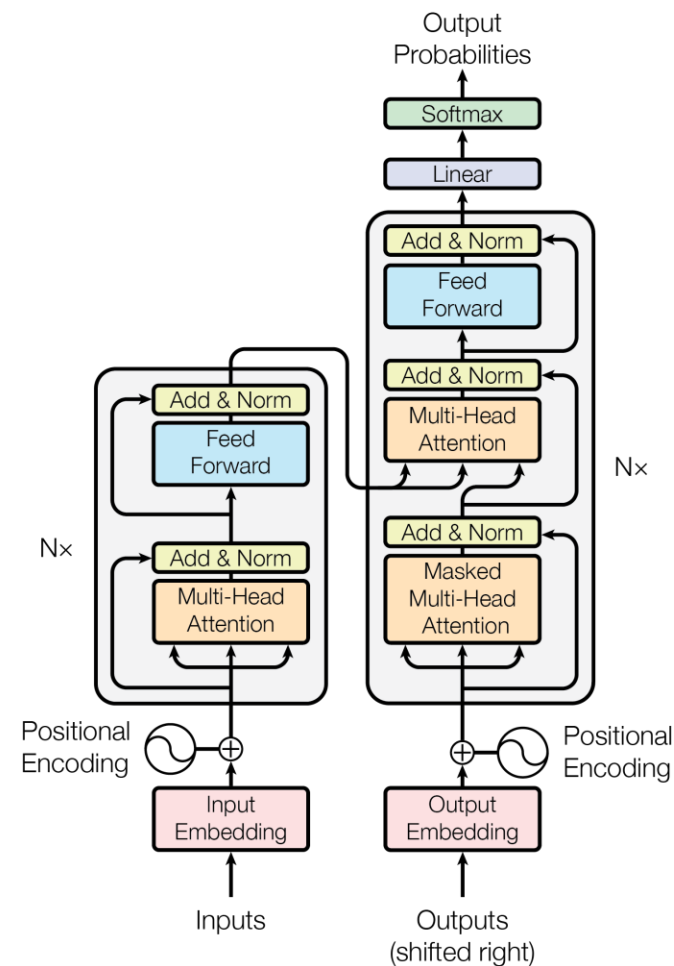
2016年

- 『ゼロから作るDeep Learning』（斎藤）



2017年

- *Attention Is All You Need*
 - NIPS 2017
 - Googleなどの研究者チーム
 - Transformer, Attention

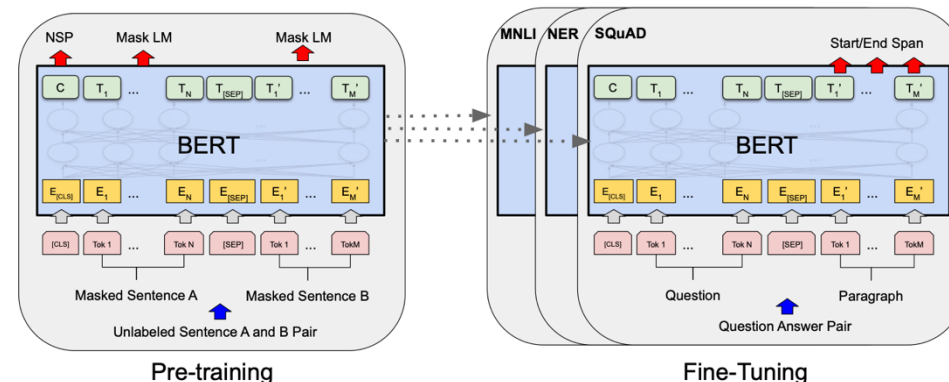


<https://doi.org/10.48550/arXiv.1706.03762>

2018年

- BERT
 - Masked Language Model
- kaggle社内ランク
 - DeNA
- GDPR施行
 - General Data Protection Regulation
 - 制定は2016年
 - Yahoo!ニュース

kaggle™



<https://doi.org/10.48550/arXiv.1810.04805>



<https://www.atlassian.com/trust/compliance/resources/gdpr>

2019年

- 内定辞退予測
 - 個人情報目的外利用
 - 法的、倫理的問題
 - Twitter上で中の人らしきアカウントが擁護して炎上という場外乱闘も。
- Chainer開発終了
 - PFN
- Anaconda一部有償化



2020年

- MLOps
- 因果推論
- Python 2 EOL
 - 1月1日
- Optuna v1.0
 - PFN
- pandas v1.0
 - 2008年の開発開始から10年以上！

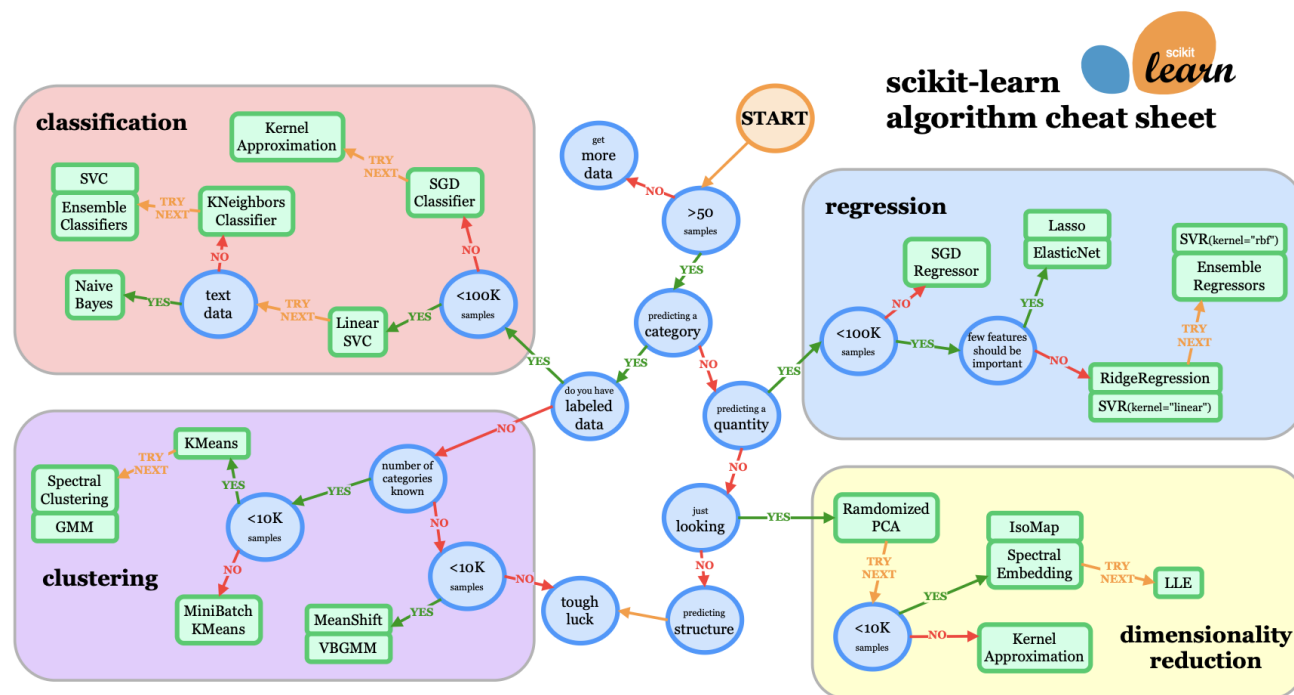


<https://x.com/pyconil/status/1212267601700802560>



2021年

- scikit-learn v1.0
 - 2007年の開発開始から10年以上！
 - 2015年時点ではv0.16
 - v1.0の直前はv0.24



https://scikit-learn.org/stable/machine_learning_map.html

2022年

みんなのPython勉強会の主催者について教えてください。

- ChatGPT (LLM)
- GPT-3自体は2020年



「みんなのPython勉強会」は、**Start Python Club (スタート・パイソン・クラブ)**が主催する、Pythonを中心とした技術共有・学習のためのコミュニティ活動です。このコミュニティは、2015年5月に東京で始まり、月に1回のペースで勉強会を開催しています。参加者は、初心者から上級者まで、さまざまなバックグラウンドを持つ人々が集まり、Pythonに関する知識や経験を共有しています。

主催者と運営メンバー

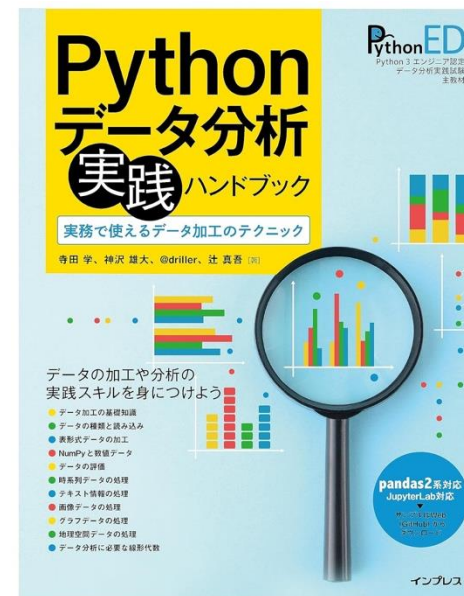
Start Python Clubの主要メンバーには、以下のような方々がいます：

- 辻 真吾 (Shingo Tsuji) : コミュニティの立ち上げメンバーであり、代表的な登壇者です。彼は、Pythonを活用したデータ解析や教育活動に従事しており、「ゼロからはじめるデータサイエンス入門～R・Python一挙両得～」などの著書もあります。 [connpass](#)
- あべんべん (Takeshi Akutsu) : コミュニティの活動報告やイベントのまとめを行っているメンバーで、勉強会の運営にも携わっています。

また、**PyQ (パイキュー)**のメンバーも登壇者として参加することがあります。例えば、鈴木たかのりさんは、PyQの運営会社であるビープラウドのメンバーであり、過去の勉強会で「Pythonで日本語処理入門 ～フリガナプログラムを作ろう～」と題した講演を行っています。 [blog.pyq.jp](#)

2023年

- pandas v2.0
- 『Pythonデータ分析 実践ハンドブック』（寺田、神沢、@driller、辻）



2024年

- 『イシューからはじめよ』改訂版（安宅）
 - 初版は2010年！



Enjoy!