

Rubyでデータ解析

@genya0407

自己紹介

- 三軒家 佑將(さんげんや ゆうすけ)
 - @genya0407 (Twitter / Github)
- 京都大学工学部 電気電子工学科 3回生
- 今年の1月までPythonista
- 今年の2月からRubyに鞍替え
 - Railsはいいぞ



普段の活動

- Webサイト/Webアプリ作成
 - 趣味/アルバイト
 - 大半はこっち
- 実験のグラフをRuby + Jupyterで描く
- SciRuby関連のPR, Issue, Document, Example ...



OSS

- daru_plotting_plotly
 - https://github.com/genya0407/daru_plotting_plotly
 - Rubyのデータフレームとグラフィブラリを連携させるライブラリ
 - きれいなグラフが書ける



今日、何を喋るか

- 正直、PyDataで喋れるようなことはない
 - Python→詳しくない
 - データサイエンス→詳しくない
- Rubyのデータサイエンス関連のこと
 - Ruby好き
 - SciRubyに(一応)Join
 - これもそんなに詳しいわけではないが...



「Rubyでデータ解析」の現状

- SciRuby Machine Learning Current Status and Future (by @mrkn)
 - <http://rubykaigi.org/2016/presentations/mrkn.html>
 - Rubyの機械学習についての発表
- 引用:
 - データサイエンスの一連のプロセスの中で
 - 「Rubyができることは、**基本的には何もない**」
 - 「Pythonは**全部**できる」
 - 「なのでみんなPythonを使う」



何もない？

- Jupyterカーネル
- データフレーム
 - **daru** (<https://github.com/v0dro/daru>)
 - 実用できるレベルにはある
 - まだまだバグ有り
- 統計ライブラリ
 - **statsample** (<https://github.com/clbustos/statsample>)
 - ドキュメント不足
 - 開発止まっている？
- グラフライブラリ
 - 乱立状態
 - rplotly, nyaplot, gruff, etc...
 - とはいえ使えるものもいくつかある

現状のRubyにできること

- 例：
 - データフレームにデータ格納＋グラフ描画
 - <https://goo.gl/00xifR>
 - 太陽電池の特性を測定したときのNotebook
 - ロジスティック回帰
 - <https://goo.gl/rl3Xjb>
 - あんまり理解してないので割愛

SciRuby

- Rubyでの科学計算やデータ可視化を促進する団体
 - 前述のライブラリもSciRubyの一部
 - Ruby自体のコミッタも何人かJoin
- とはいえ人手不足感
 - 使う人が少なすぎてStack Overflowに質問がない
 - Ruby歴半年ちょいの素人(自分)でもJoinできる
 - データサイエンスの人がJoinしたら、とても歓迎されると思われる



まとめ

- Rubyができることも(多少は)ある
 - データフレーム/グラフ/初等統計
 - 機械学習はまだまだ
- 興味あり → Join SciRuby!
 - HP:
 - <http://sciruby.com/>
 - <https://sciruby-jp.github.io/>
 - slack:
 - <https://sciruby-slack.herokuapp.com/>

ご清聴ありがとうございました

