

統計データの可視化

matplotlibとGoogle Chartの比較

目次

- 背景
- 目的
- 比較した可視化ツール
- まとめ

背景

- 統計データから世界情勢を理解したい
- 数字のデータではわかりにくい
- わかりやすい図表に可視化したい
- 適切な可視化の方法がわからない

目的

- いくつかのデータ可視化ツールの特徴を比較

使用する統計データ

- 国別マラリアによる乳幼児の死亡率(The Database,WHOより)
- 国別マラリアによる死者数(同上)
- 国別GDP(World Development Indicatorsより)

比較した可視化ツール

- Google ChartのGeoChart
- matplotlibの折れ線グラフ
- matplotlibの棒グラフ

Google Chartとは

- Googleが提供するグラフ描画ツール
- Javascriptで動くWebアプリケーション
- データ、フォーマットの指定が容易
- インタラクティブに動作

GeoChartとは

- Google Chartの機能の一つ
- 階級区分図 Choropleth mapを
描画できる
- 階級区分図…地図上の地域を
階級によって色分けした地図

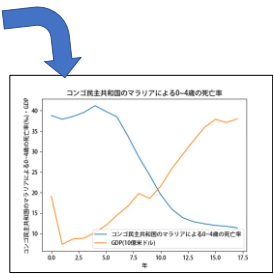


Matplotlibとは

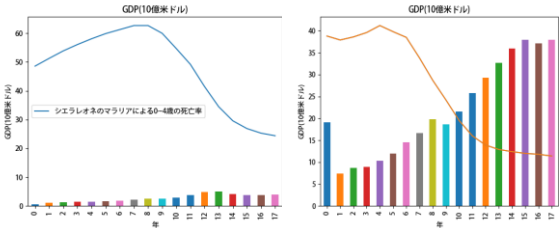
- Pythonと科学計算用ライブラリ Numpyのためのグラフ描画
ライブラリ
- 主に2次元のプロットをサポート
- 多くのグラフ形式を提供

```
cd = pd.DataFrame(Y
[[2000, 38.8, 19.1],[2001, 37.9, 7.4],[2002, 38.6,
8.7],[2003, 39.6, 8.9],Y
[2004, 41.2, 10.3],[2005, 39.8, 12.0],[2006, 38.5,
14.5],[2007, 33.8, 16.7],Y
[2008, 28.7, 19.8],[2009, 24.2, 18.6],[2010, 19.5,
21.6],[2011, 16.0, 25.8],Y
[2012, 13.9, 29.3],[2013, 12.9, 32.7],[2014, 12.4,
35.9],[2015, 12.0, 37.9],Y
[2016, 11.8, 37.1],[2017,11.4,38.0]],Y
columns = ['年', 'コンゴ民主共和国のマラリアによる0-4歳の
死亡率', 'GDP(10億米ドル)'])

cd['コンゴ民主共和国のマラリアによる0-4歳の死亡率
'].plot()
plt.title('コンゴ民主共和国のマラリアによる0-4歳の死亡率')
plt.xlabel('年')
plt.ylabel('死亡率(%)')
plt.savefig('CDMDeathRate001.png')
```



折れ線グラフ・棒グラフ

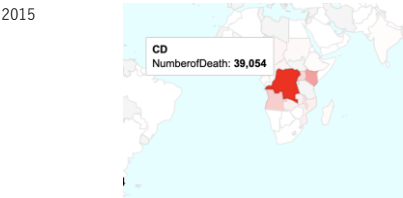
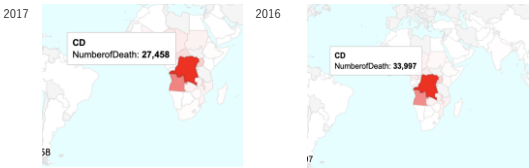


具体例

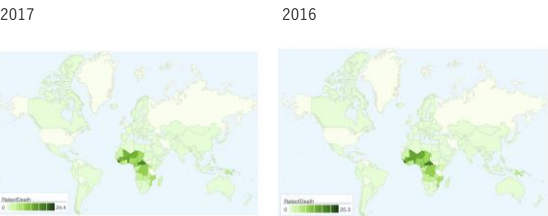
3年間のマラリアによる死亡数の推移



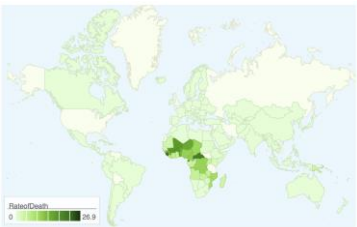
2015



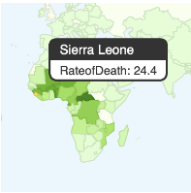
マラリアによる死亡率の推移



2015



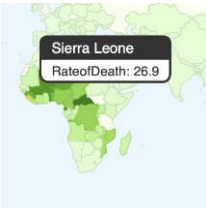
2017



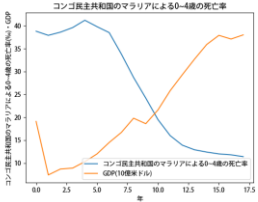
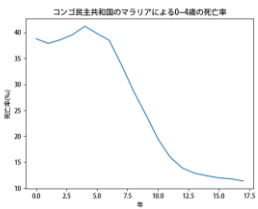
2016



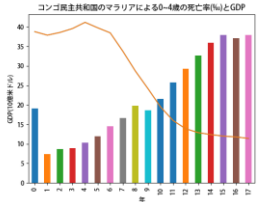
2015



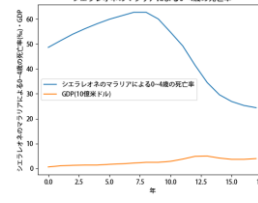
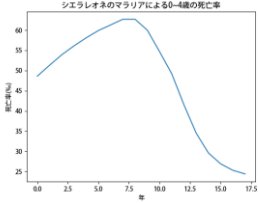
折れ線グラフで見る死亡率の推移



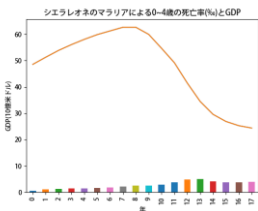
折れ線グラフと棒グラフを用いた比較



折れ線グラフで見る死亡率の推移



折れ線グラフと棒グラフを用いた比較



まとめ

折れ線グラフ・棒グラフ

- 時系列に沿った変化が見やすい
- 2つの要素を比較することができる

GeoChart(階級区分図)

- 地域特性が見やすい
- 時間的推移が見にくい