

人を動かすデータビジュアライゼーション データビジュアライゼーションを学問にするには？

GeoDev Meetup 2016-12-16

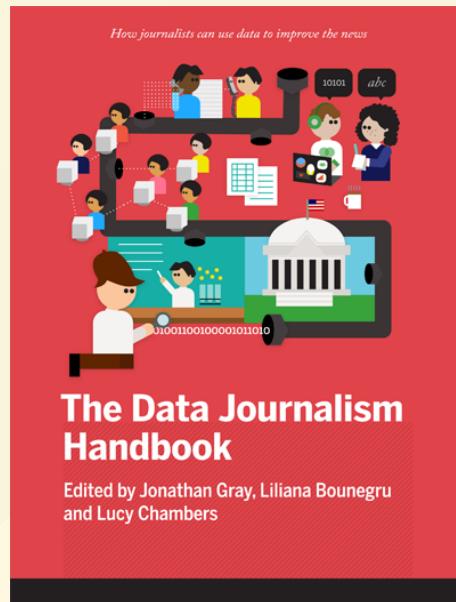


五十嵐康伸 [E2D3.org、株式会社インテリジェンス]

データをどうやって見せるか？

「The Data Journalism Handbook」日本語翻訳版

リリース記念イベント 2016-11-27



<http://datajournalismjp.github.io/handbook/>

講演内容

1. データ・ジャーナリズムとはなにか（立園 理彦）
2. いかにしてデータを手に入れるか（亀田 勇宙）
3. データ・ジャーナリズムで人々は情報をどうに消費するか（桐井 智弘）
4. データをどうやって見せるか（五十嵐 康伸）

講演内容のスライド

<https://goo.gl/xTXHlk>

知能プログラム特論

名古屋工業大学 M1

2016-12-20



自然言語処理, Linked Data／セマンティックWeb等の先端研究について調査し, 最新研究及び技術の動向を把握習得する。

講義内容を決めるために大事にしたいこと

Data Viz（データ・ビジュアライゼーション）が教育
対象、研究対象になりうるよう

大学院生（Data Vizを初めて聞く人）の人生にとっ
て、いつか役に立つように

世界（広がり）と歴史（深さ）を"体系的"に

Data Vizを学問に

起きて自由な時間の50%はE2D3

E2D3は何を目指すべきか？
E2D3をどう発展させるべきか？



研究対象としてData Viz

工学：便利な道具 ([ヒト、モノ、力ネ]、情報、自然を制御する方法、理論) を作る

理学：自然の法則を発見する

研究対象としてData Viz

工学：便利な道具 ([ヒト、モノ、力ネ]、情報、自然を制御する方法、理論) を作る

New Comp + New Dataによって

1. 新しい表現 (Data Viz) が可能に
2. 過去の表現 (Data Viz) をより手軽に、より早く可能に

→ 世界（広がり）と歴史（深さ）を"体系的"に

世界（広がり）と歴史（深さ）の体系

	地図	Tree・Netowrk	統計グラフ・チャート	ピクトグラム・写真
古代				
中世	<ul style="list-style-type: none"> ・John Snow ロンドンのコレラ地図 ・André-Michel Guerry: 犯罪レベルを示すフランスのコロプレス(塗り分け)地図 ・TED&本 スティーブ・ジョンソン: 感染地図 		<ul style="list-style-type: none"> ・William Playfair: 棒グラフ、線グラフ ・Florence Nightingale: 鳥頭グラフ ・本 統計学者としてのナイチンゲール 	
現代	<ul style="list-style-type: none"> ・本 杉浦康平: 時間のヒダ、空間のシワ 	<ul style="list-style-type: none"> ・TED マニュエル・リマ: 人類の知識を表す視覚的表現の歴史 ・本 ビジュアル・コンプレキシティ ・本 THE BOOK OF TREES 	<ul style="list-style-type: none"> ・TED デビッド・マキャンドレス: データビジュアライゼーションの美 ・本 鈴木、鈴村: データビジュアライゼーションのデザインパターン20 	
ツール	ArcGIS・QGIS	Cytoscope	Tableau・PowerBI・Plotly	E2D3・Infogram

データ関連技術の出力先は？

1. データ計測 [機械+人間]
2. データ解析、シミュレーション [機械+人間]
3. データ可視化 [必ず人間]

Data Vizは何の問題を解決するために作るのか？

大問題：国、自治体、共同体、会社の中の人 が

- 間違った方に動いている
- 正しい方向に動いているか怪しい
- 動かない（無関心、信用・信頼・懐疑、透明性）
by ジャーナリスト、研究者、アナリスト、営業・マーケ、市民

中問題：データに基づいて、人を動かす

Data Viz is a method to stimulate a human brain

小問題：情報過不足調整、主体探索、新規発見、美

6. データを届ける

6.4 データ・ジャーナリズムの馬車馬としてのビジュアライゼーション

6.5 ビジュアライゼーションの力でストーリーを伝える

6.6 異なる図は、異なる結果を意味する
Brian Suda, [optional.is](#) ***

6.7 データ・ビジュアライゼーション・DIY：便利なツール

in Data Journalism Handbook

[6.6 前半] DVの歴史

今となっては紙は静的で平坦で、ちょっと劣ったメディアだと思ってしまうけれど、実のところはもう何百年もの間にわたって、紙に書いて、印刷をすることで、信じられないほどの豊かな知識とデータをなんとかその紙に示すことを行なは成し遂げてきたわけである。いま流行りのインタラクティブ・チャートやDV、そしてインフォグラフィクスも、もとは我々が学んできた多くのベスト・プラクティスのもとに成り立っている。

[6.6 前半] 3人の偉人

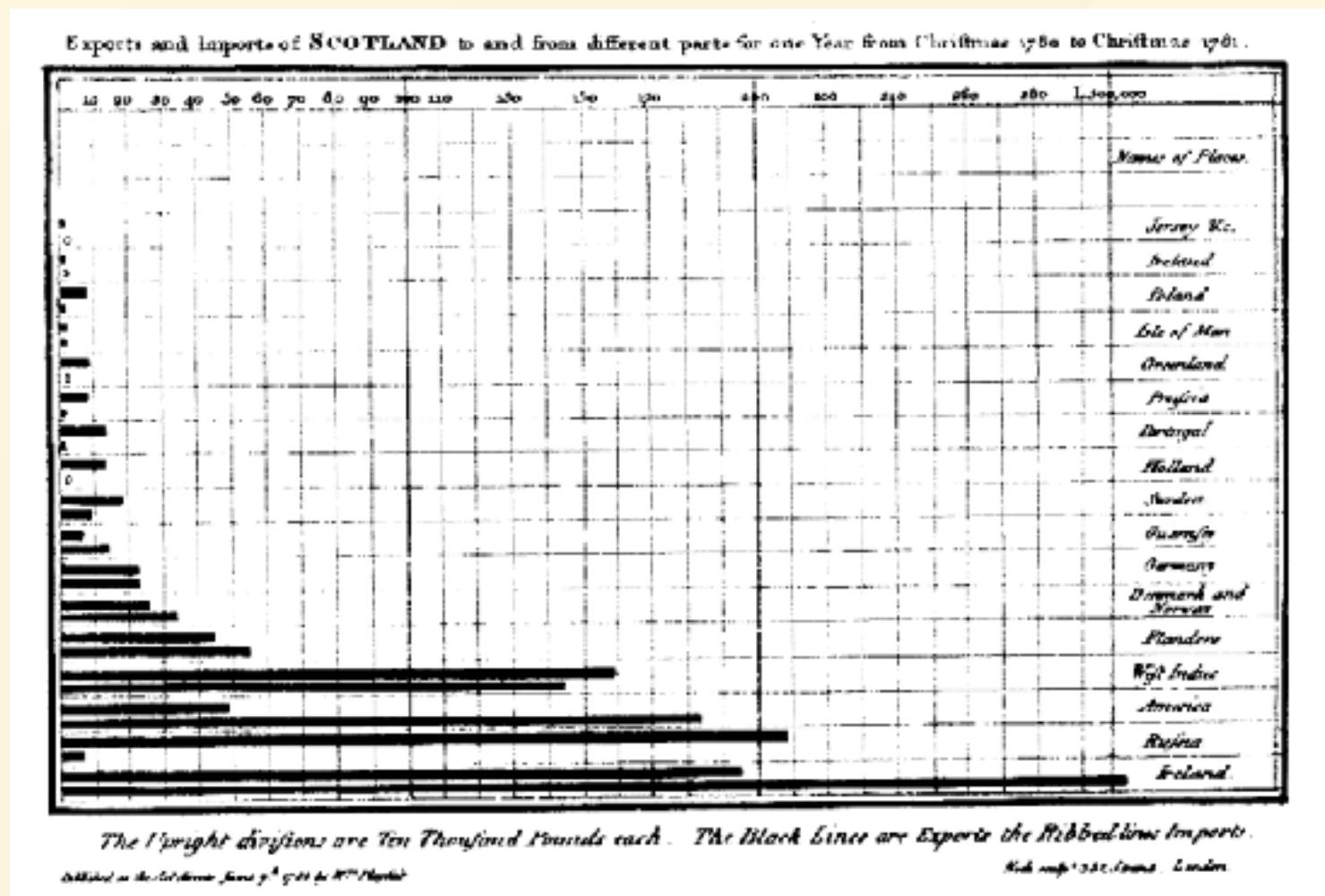
1. William Playfair (1759-1823, Sco)
2. John Snow (1813-1858, En)
3. André-Michel Guerry (1802-1866, Fr)

William Playfair (1759-1823, Sco)

多言語に通じたスコットランド人技師

William Playfairは、今日でも同じように使われている図やグラフのほとんどを、たった一人で世界に送り出した人物である。1786年の彼の本、「Commercial and Political Atlas（商業と政治の地図）」のなかで彼は、棒グラフという全く新しく、視覚的な手法を用いて、スコットランドの輸出入量を明快に示したのである。

William Playfair (1759-1823, Sco)



初期の頃の棒グラフ、Export and Import of Scotland

William Playfair (1759-1823, Sco)

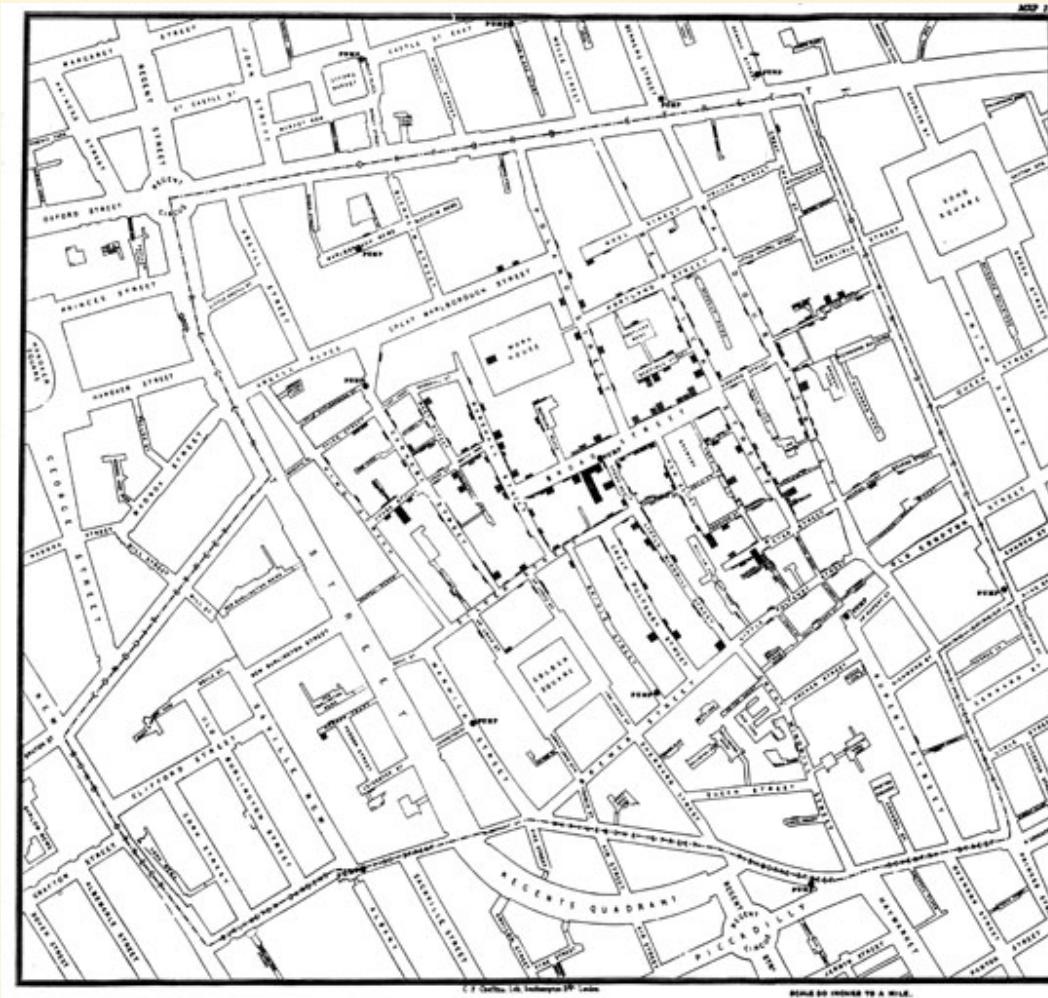
William Playfairは、1801年の本、「Statistical Breviary（統計的な聖務日課書）」の中で、世にも恐ろしい円グラフを世に広めた。これらの新しい形の図とグラフは、商業的なニーズから生まれたものであったが、時間が経つにつれて、別の者が現れて、人の命を救うために使うようになった。

John Snow (1813-1858, En)

イギリンドの医師、疫学の創設者

1854年、John Snowは今日では有名な「ロンドンのコレラ地図」を、発症が報告された地域に小さな黒い棒をつけて作成した。徐々に時間が経つに連れ、明らかに発症密度が高い地域がわかるようになり、問題の抑制措置が取ることができた。

John Snow (1813-1858, En)



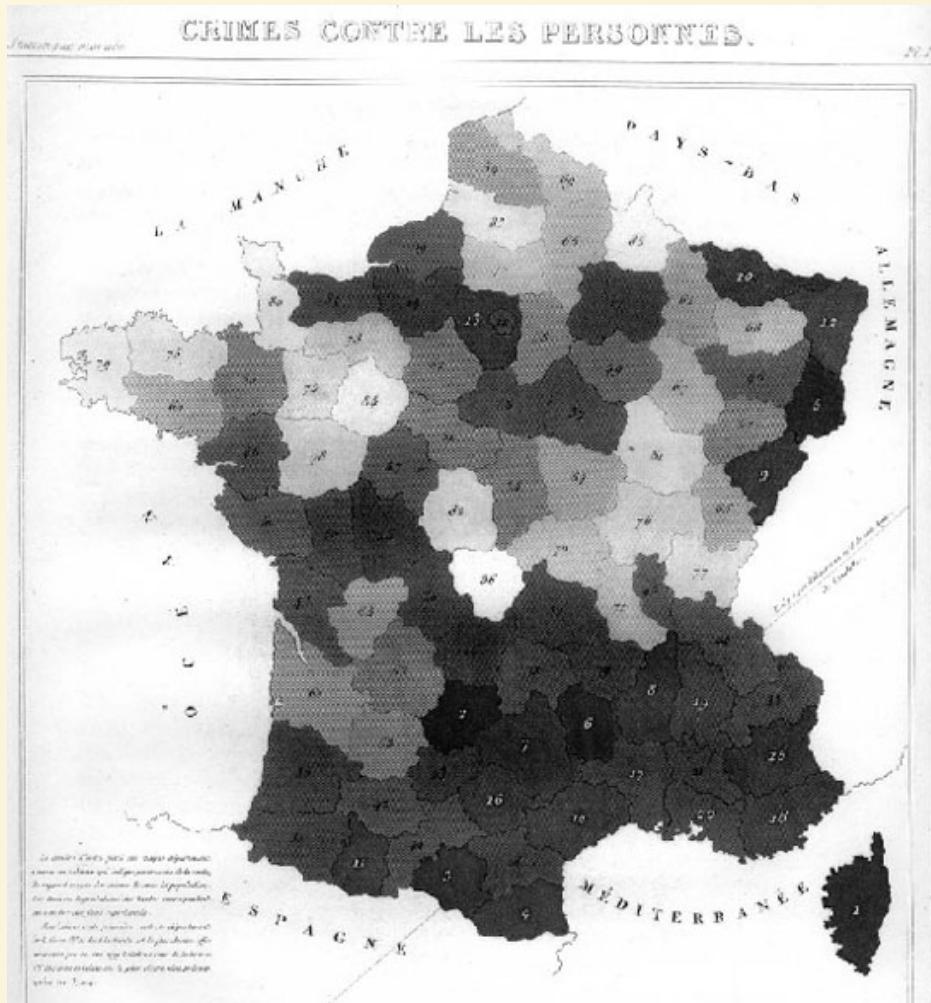
ロンドンのコレラ地図

André-Michel Guerry (1802-1866, Fr)

フランスの弁護士とアマチュアの統計学者

André-Michel Guerryは、地図に何らかの変数に応じて、地図の各々の地域を違う色で塗り分けるというアイデアを最初に世に出した人物である。彼は1829年に、犯罪レベルに応じてフランスの地図の各地区を影付けをして、初めてのコロプレス(塗り分け)地図を作成した。

André-Michel Guerry (1802-1866, Fr)



犯罪レベルを示すフランスの
コロプレス(塗り分け)地図

André-Michel Guerry (1802-1866, Fr)

今日、これらの地図は、例えば選挙区ごとに誰が誰に投票したかや、富の分布、その他の多くの地理的な要因に結びつく変数を説明するときに使われている。

6. データを届ける

6.4 データ・ジャーナリズムの馬車馬としてのビジュアライゼーション

6.5 ビジュアライゼーションの力でストーリーを伝える

Geoff McGhee, Stanford University ***

6.6 異なる図は、異なる結果を意味する

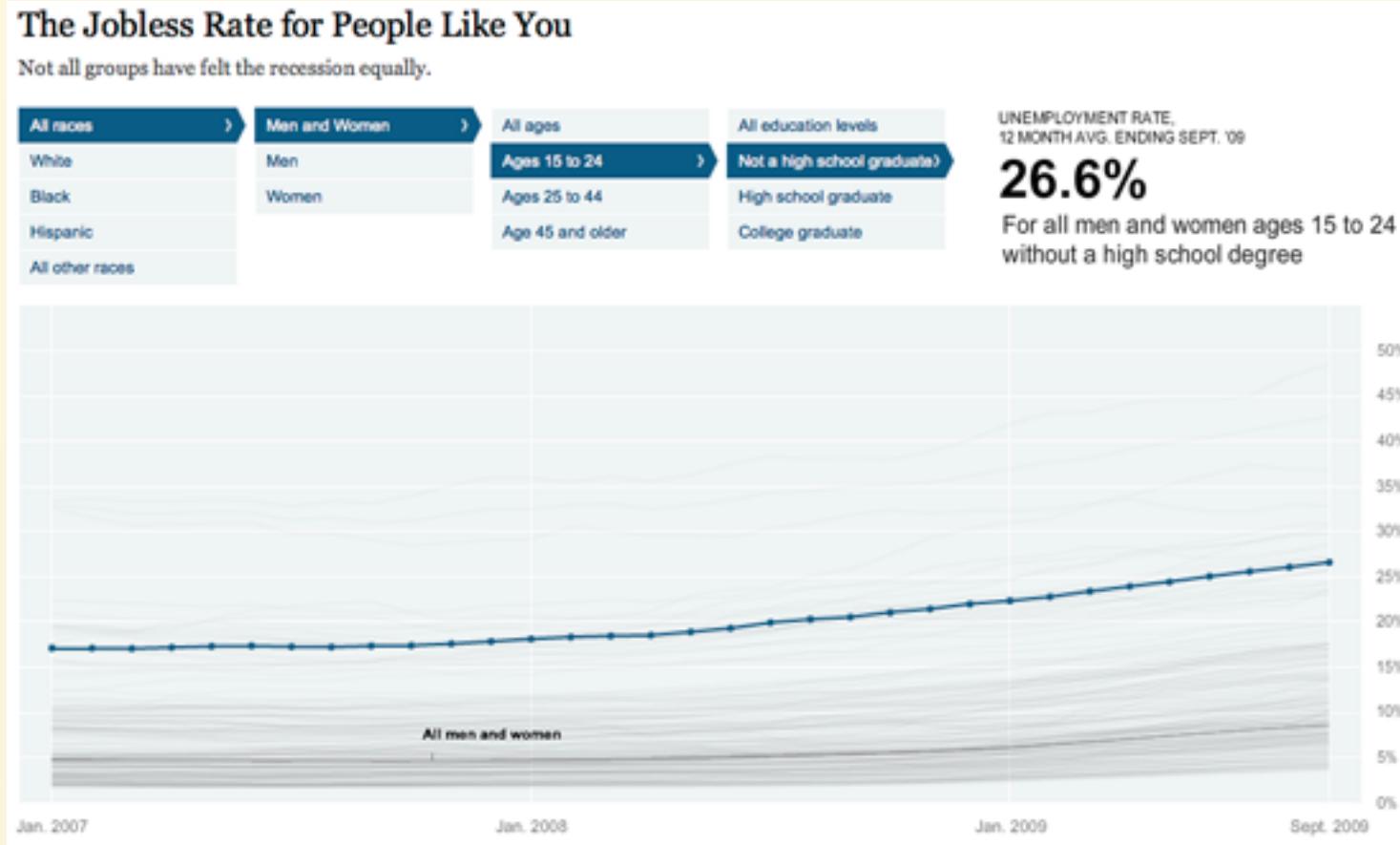
6.7 データ・ビジュアライゼーション・DIY：便利なツール

in Data Journalism Handbook

[6.5] 読者へ物語りを伝えるときにData Vizが役立つかもしれないいくつかの例

1. 新しい方法で馴染みのものを見ること
2. 時間と共に変化することを示す
3. 値の比較
4. 繋がりと流れを見せること
5. データでデザインする
6. 階層を示すこと
7. 大きなデータセットをブラウンジングする
8. 代替結果を想定する事
9. データの可視化を使用しないとき

[6.5.1] 新しい方法で馴染みのものを見ること

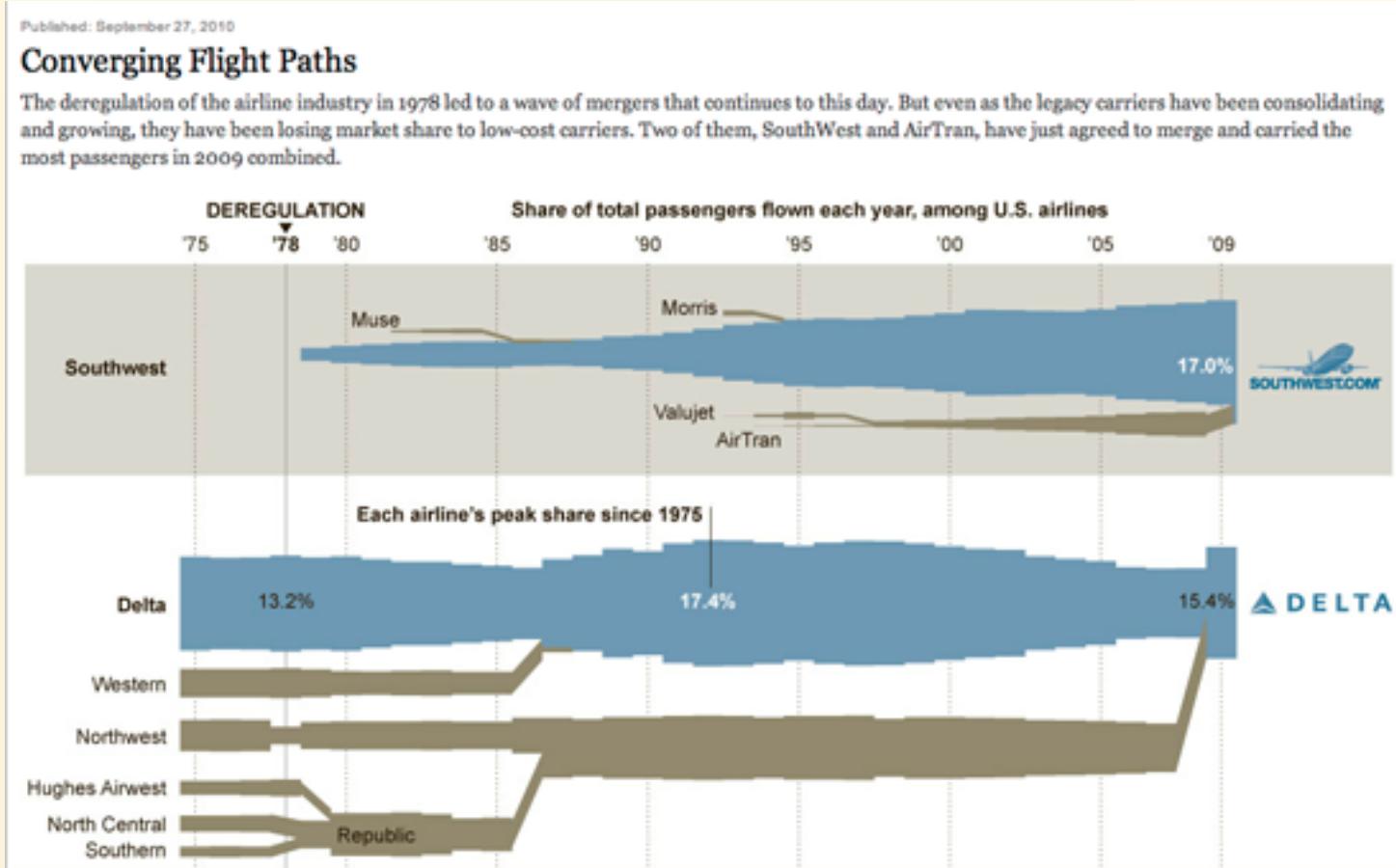


あなたのような人々のための失業率
(NYT) <https://goo.gl/2tEwcT>

[6.5.1] 新しい方法で馴染みのものを見ること

米国の失業率が9%近くでさまよっていると、ユーザーは様々な人口統計や学歴フィルターによって、米国人口を選別することが出来た。結局のところ、その内訳は学士号より上の学位を持つ中年女性から高校を卒業していない若い黒人男性の全体の半分近くで4%未満を占められていて、この格差は真新しくなかった。それを見終わったあとでさえ、良いDVは頭の中に入り、事実、傾向またプロセスの持続的なメンタルモデルが残る。

[6.5.2] 時間と共に変化することを示す

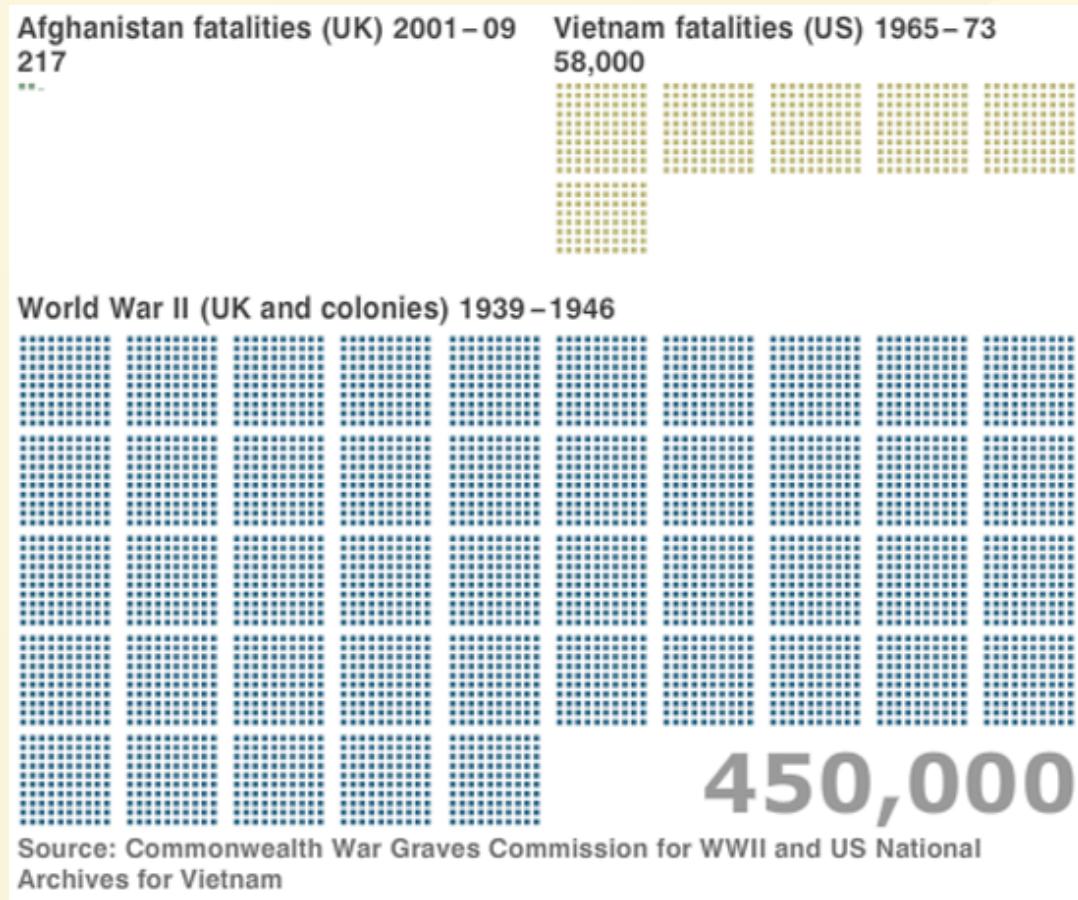


飛行経路の収束 (NYT) <https://goo.gl/DzwM>
deregulation:規制撤廃

[6.5.2] 時間と共に変化することを示す

時系列の変化を示している他の興味深いグラフは、業界再編の数十年の間についての米国的主要航空会社の市場占有率を描画したものがある。Carter政権が旅客飛行を規制緩和した後、NYTが描いたこのグラフのように、たくさんの国債融資が小さな地方航空会社から全国航空会社を作成しました。ほとんど全てのカジュアルの読者が図の水平線「X」軸を時間を表すと見るとするならば、時には、全てのVizが時系列での変化を表すべきだと考えることは容易である。

[6.5.3] 値の比較

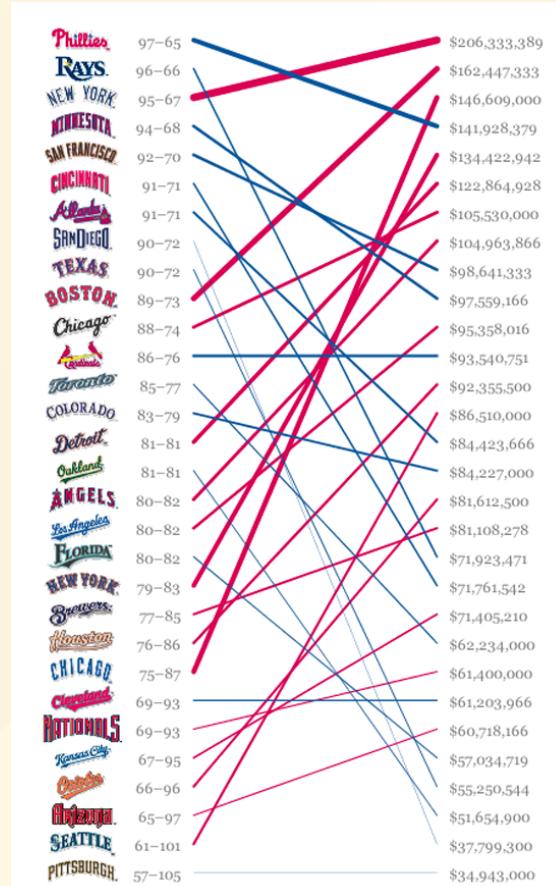


戦争の人的コストの計算 (BBC) <https://goo.gl/3pIYfy>
fatality: 死亡者（数）

[6.5.3] 値の比較

BBCが実施した犠牲者のデータベースとアニメ化されたスライドショーのように、イランとアフガニスタンの争い（2001-2009）と、ベトナムで殺された何千人の命や（1965-1973）、第二次世界大戦で無くなった何百万もの命（1939-1946）と比較することによって、DVは読者が二つやそれ以上の値の比較のするのを助ける分野においても輝く。

[6.5.4] 繋がりと流れを見せること



年棒対パフォーマンス(Ben Fry) <https://goo.gl/hqZdi5>
左：日付との成績、右：年俸合計

[6.5.4] 繋がりと流れを見せること

二つに分けられた値を比較するために、メジャーリーグベースボールチーム relative to their payrolls のパフォーマンスが計算されている Ben Fry の図を見る。左側の列では、チームは日付ごとの成績によってランク付けされていて、一方右側の列は年俸の合計である。赤色（標準以下）もしくは青色（標準以上）で描かれた線は二つの値つなぎ、駄目になった年俸の高い選手に後悔しているチームオーナーの便利な感覚を提供している。

[6.5.5] データでデザインする

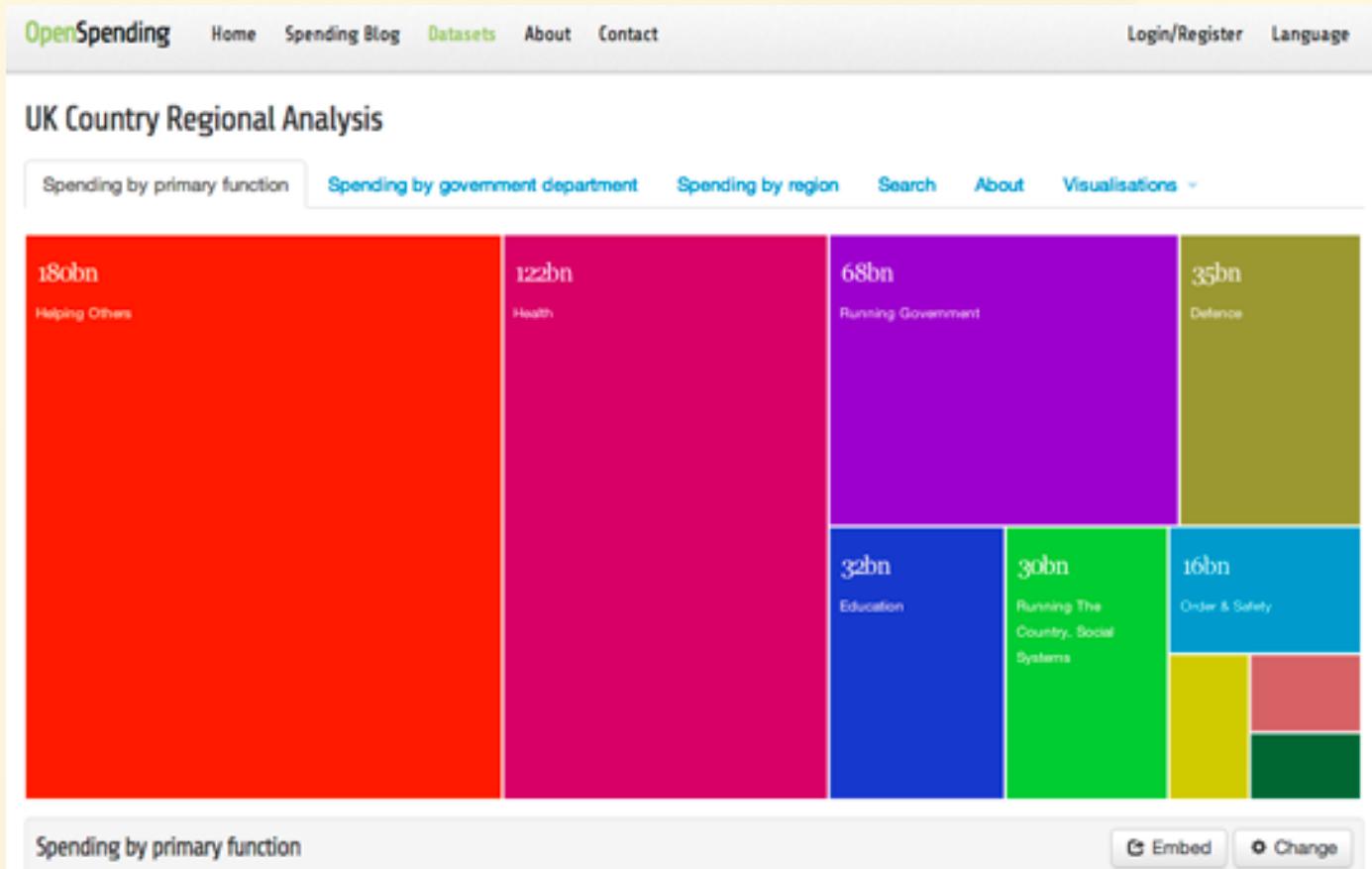
繋がりをグラフにするの方法と同様に、流れの図もいつも通り太さ・色によって情報を線の繋がりに組み込んで符号化する。例えば、線の太さはある国から他国へ渡る信用度反映している

It's All Connected: An Overview of the Euro Crisis
(NYT)

<https://goo.gl/xxqT1>

参考：西田 善彦さん、直感とビジュアライズ
<https://goo.gl/ODSI0A>

[6.5.6] 階層を示すこと



OpenSpending.org (Open Knowledge Foundation)
<https://openspending.org/>

[6.5.6] 階層を示すこと

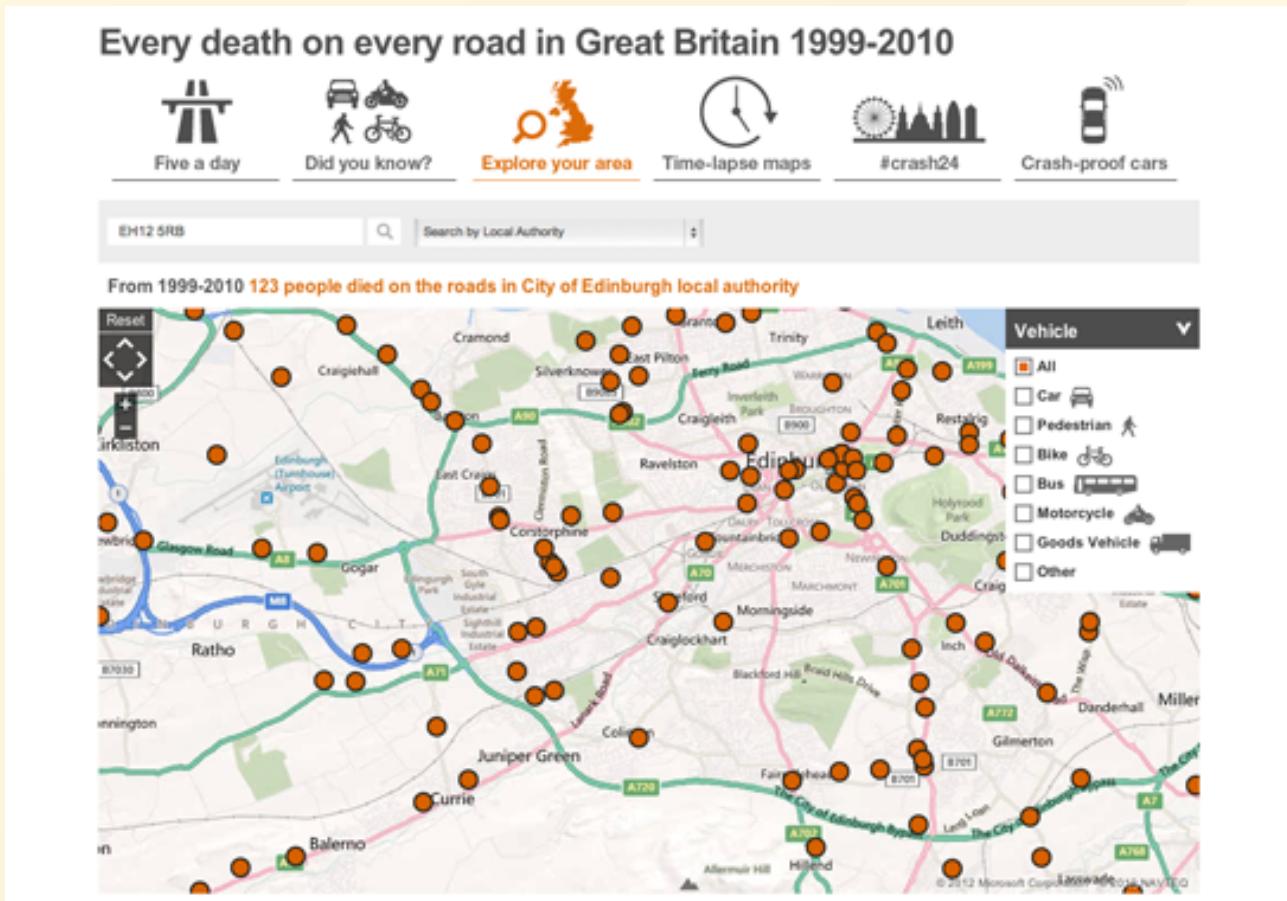
研究者Ben Shneidermanは、互いの内部で同心円上に入れ子になっている複数の箱を含む "ツリーマップ"と呼ばれる新しいVizを開発した（1991、樹形図を有限空間に埋め込む）。

<http://www.cs.umd.edu/hcil/treemap-history/>

参考：THE BOOK OF TREES—系統樹大全:知の世界を可視化するインフォグラフィックス

<https://goo.gl/SS6IWw>

[6.5.7] 大きなデータセットをブラウンジングする



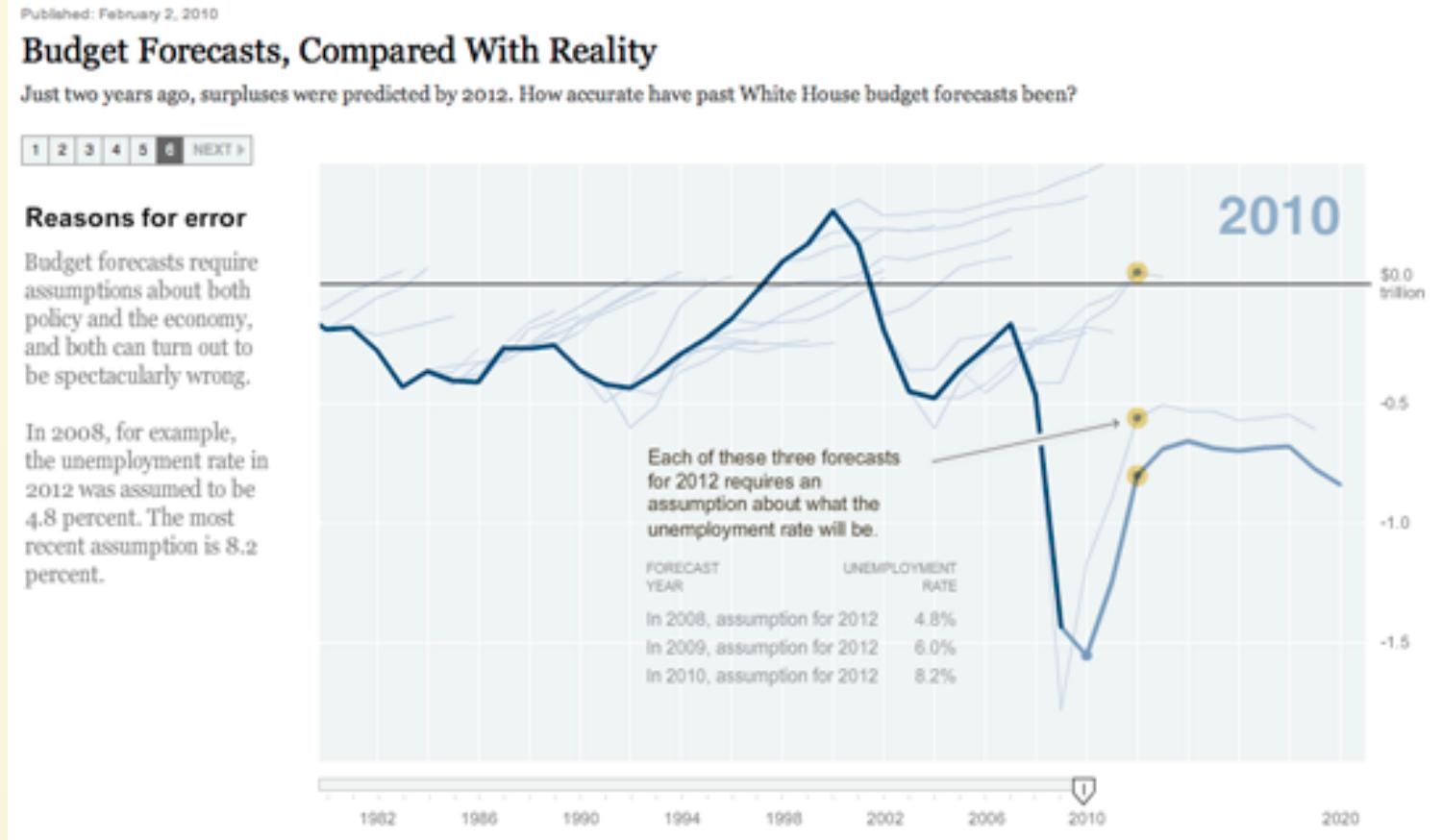
1999年から2010年イギリスの路上における
すべての死 (BBC) <https://goo.gl/qjba03>
local authority : 地方自治体

[6.5.7] 大きなデータセットをブラウンジングする

DVはよく知られている情報を話す時や新しいものを示す時とても効果的であるが、人々が検索したい真新しい情報をあなたが受け取った時に、何が起こるだろうか？

データベースのフロントエンドとしてVizの同じような賞賛に値する用途は、BBCの交通死亡の調査やWikiLeaksにあるイラクやアフガニスタンの戦争記録の開示のような大きなダンプデータを素早くインデックスするという多くの試みがある。

[6.5.8] 代替結果を想定する事



現実に対する予算見通し (NYT) <https://goo.gl/Atym>
surplus : 黒字、deficit : 赤字

[6.5.8] 代替結果を想定する事

NYTにおいて、何年にも渡る 悲惨なほど楽観的な米国の赤字予測のAmanda Coxの「ヤマアラシ図」は起こらなかったことより、面白くないことが起きたことがどれほどあるか示している。戦争と減税の十年の後に財政赤字が急増していることを示しているCoxの熱線は将来の期待が起きることがどれほど非現実的かを示している。

謝辞

- 本日お越しの皆様
- 企画、運営してくださったESRIジャパンの皆様

に御礼申し上げます

Enjoy Data Viz !

この内容なら、いつでもどこでもトーク可能
御社、御コミュニティーへお気軽に呼んでください！



五十嵐康伸 [E2D3.org、株式会社インテリジェンス]
FB: yasunobu.igarashi、Twitter: B_T_Budo