

# サンキーダイアグラムのススメ

**7<sup>th</sup> June 2020, PyData.Fukuoka Meetup #8**  
**Yuta Kanzawa @yutakanzawa**



**Data Science Senior Analyst at Janssen Pharmaceutical K.K., Tokyo**  
**A Family Company of Johnson & Johnson**



# I am...

- 神沢雄大 Yuta Kanzawa (twitter: [@yutakanzawa](https://twitter.com/yutakanzawa))
- Data scientist at Janssen Japan, Tokyo
  - A pharmaceutical company of J&J
- Opera & wine lover
  - Wagner
  - Bourgogne
- 7 languages
  - Human: Japanese, English, German
  - Computer: R, Python, SAS, SQL

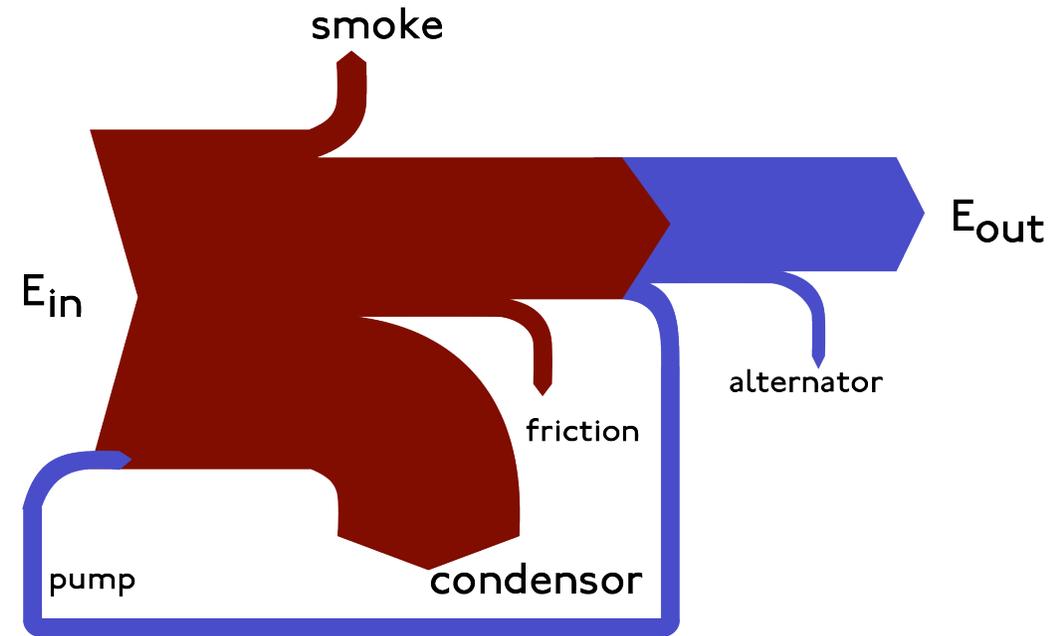


# アジェンダ

- 今日話すこと
  - サンキーダイアグラム
- 今日話さないこと
  - コード
  - R

# サンキーダイアグラムとは

- フローチャートに似た有向グラフ
  - ノード間の関係性と順序 = 流れの方向
  - 矢印の幅 = 流量
  - 量を伴う、時間・空間・状態の遷移に最適。

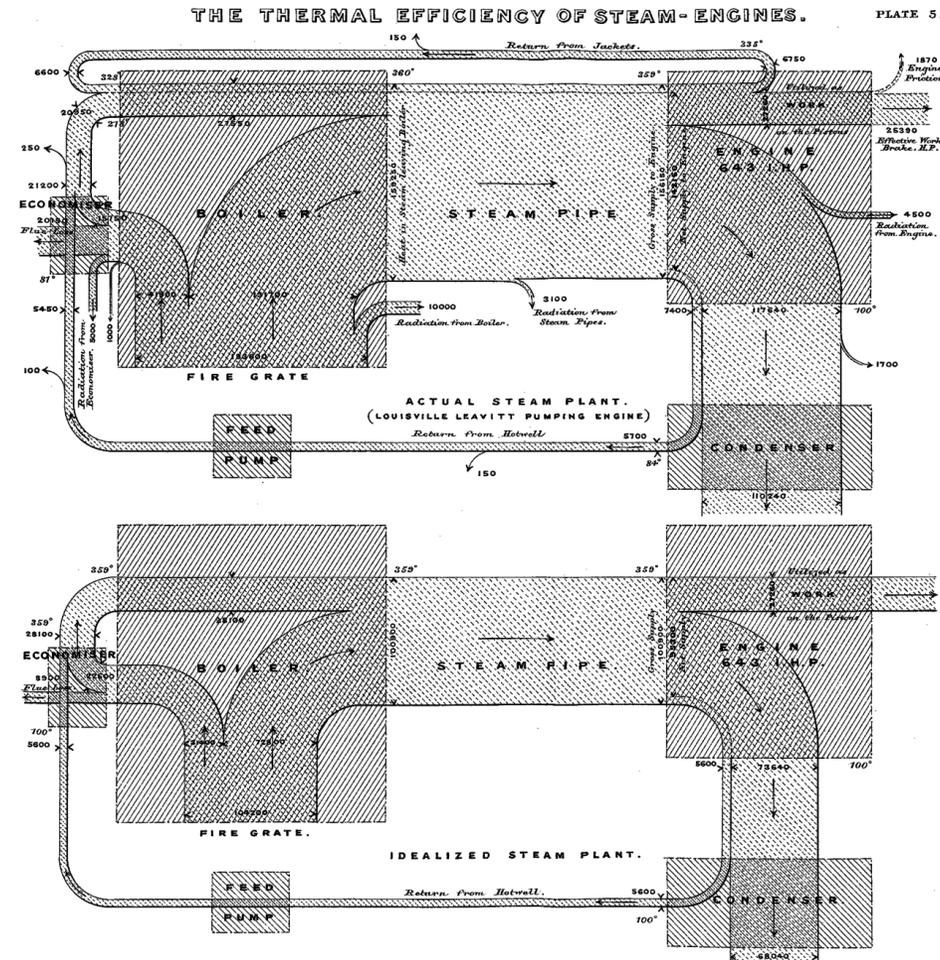


- 19世紀からある！\*1
  - 例1：蒸気機関の熱効率（サンキー、1898年）
    - 「サンキー」はこれを描いた人の名前
  - 例2：ナポレオンのロシア遠征の兵力推移（ミナール、1869年）
    - 実はミナールが先。\*2
  - サンキーもミナールもエンジニア。

\*1 [https://en.wikipedia.org/wiki/Sankey\\_diagram](https://en.wikipedia.org/wiki/Sankey_diagram)

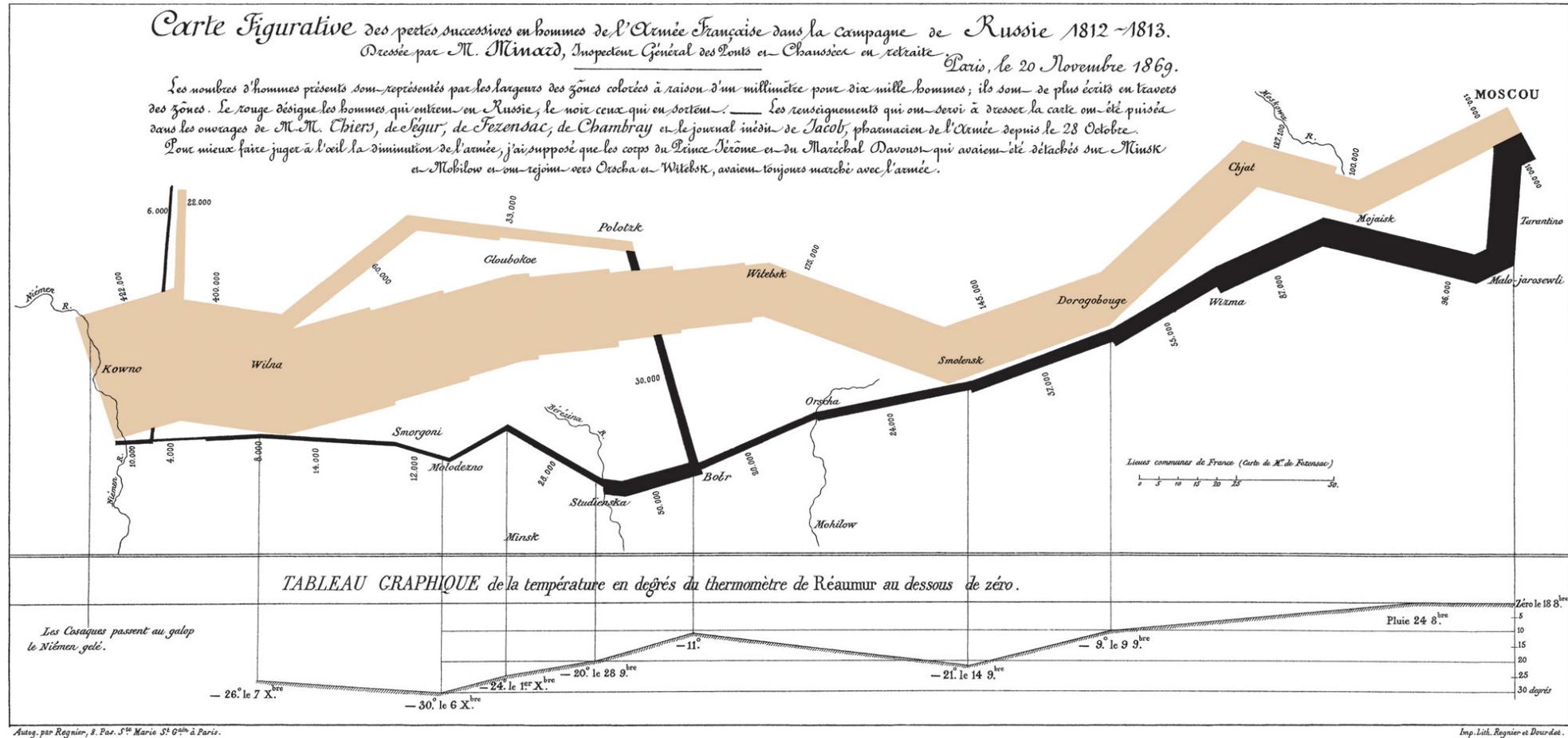
\*2 [https://commons.wikimedia.org/wiki/Charles\\_Joseph\\_Minard](https://commons.wikimedia.org/wiki/Charles_Joseph_Minard)

# 例1：蒸気機関の熱効率（サンキー、1898年）



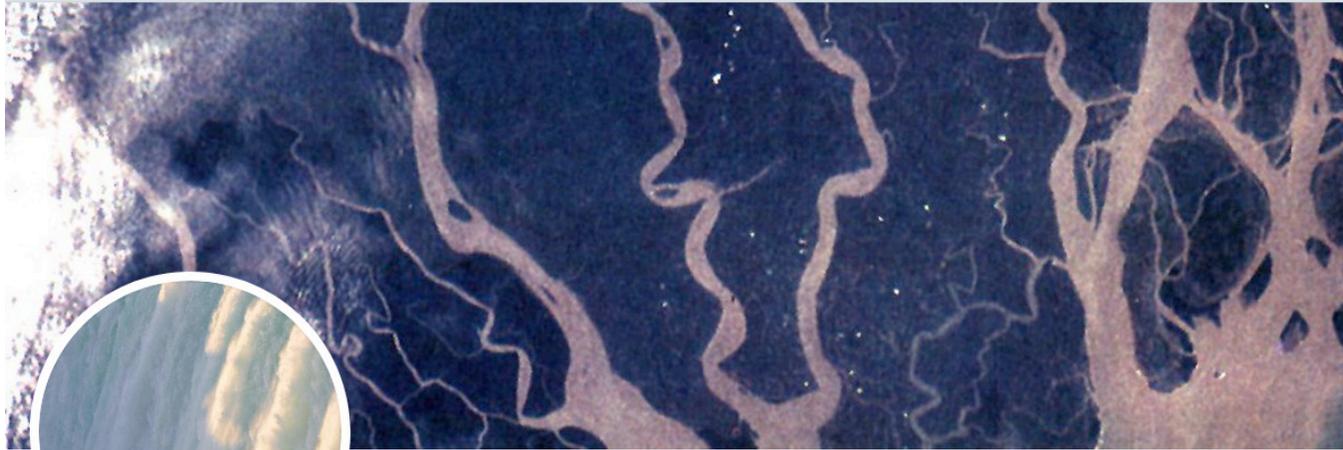
\* [https://en.wikipedia.org/wiki/Sankey\\_diagram](https://en.wikipedia.org/wiki/Sankey_diagram)

# 例2：ナポレオンのロシア遠征の兵力推移（ミナル、1869年）



\* [https://en.wikipedia.org/wiki/Sankey\\_diagram](https://en.wikipedia.org/wiki/Sankey_diagram)

# 番外編：サンキーダイアグラムファンのためのTwitterアカウント



Follow

**Sankey Diagram Lovers**

@SankeyLovers

#sankey #sankeydiagrams #alluvialdiagram #minard #dataviz #infoviz #mobile #design #flow #datavisualization #visualization #mapviz #opendata #networks #viz #ml

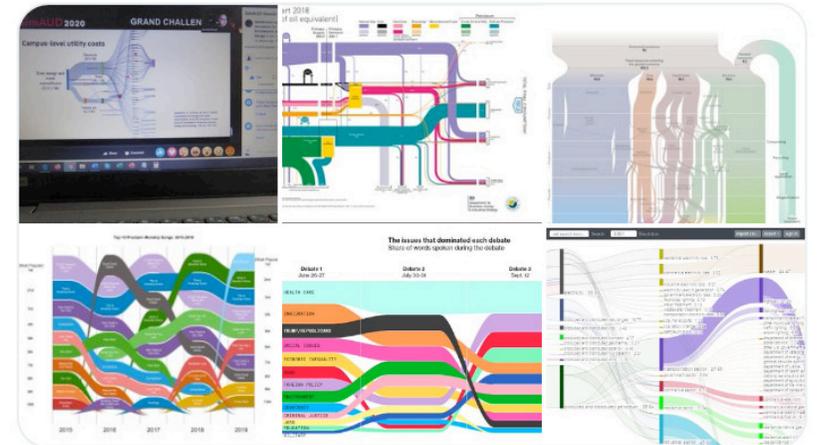
📍 Lisbon, Portugal 🔗 [behance.net/dph22](https://www.behance.net/dph22) 📅 Joined November 2017

2,997 Following 682 Followers

**New to Twitter?**

Sign up now to get your own personalized timeline!

Sign up



**You might like**

Diogo Pereira Henri

\* <https://twitter.com/sankeylovers>

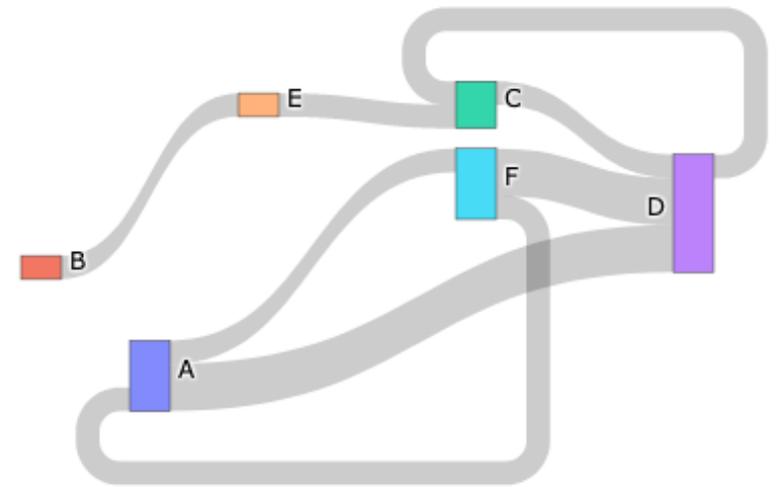
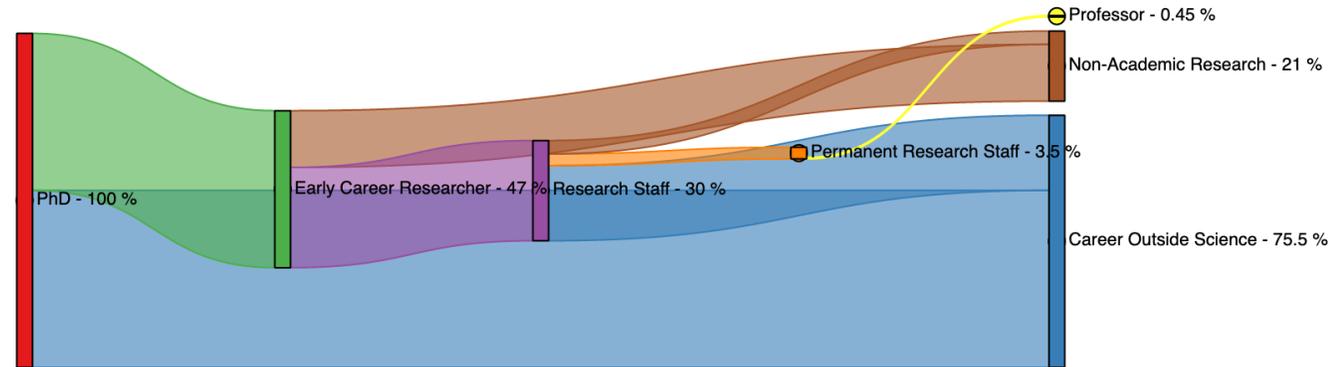
# Pythonでの実装

- HoloViews\*1

- holoviews.Sankey
  - バックエンド : BokehとMatplotlib
  - Plotlyは未実装?
- シンプル
- 非循環データのみ。

- Plotly\*2

- plotly.graph\_objects.Sankey
- 細かい設定が可能。
- 循環データもOK。
- どちらも矢印はないらしい (左→右の流れ) 。



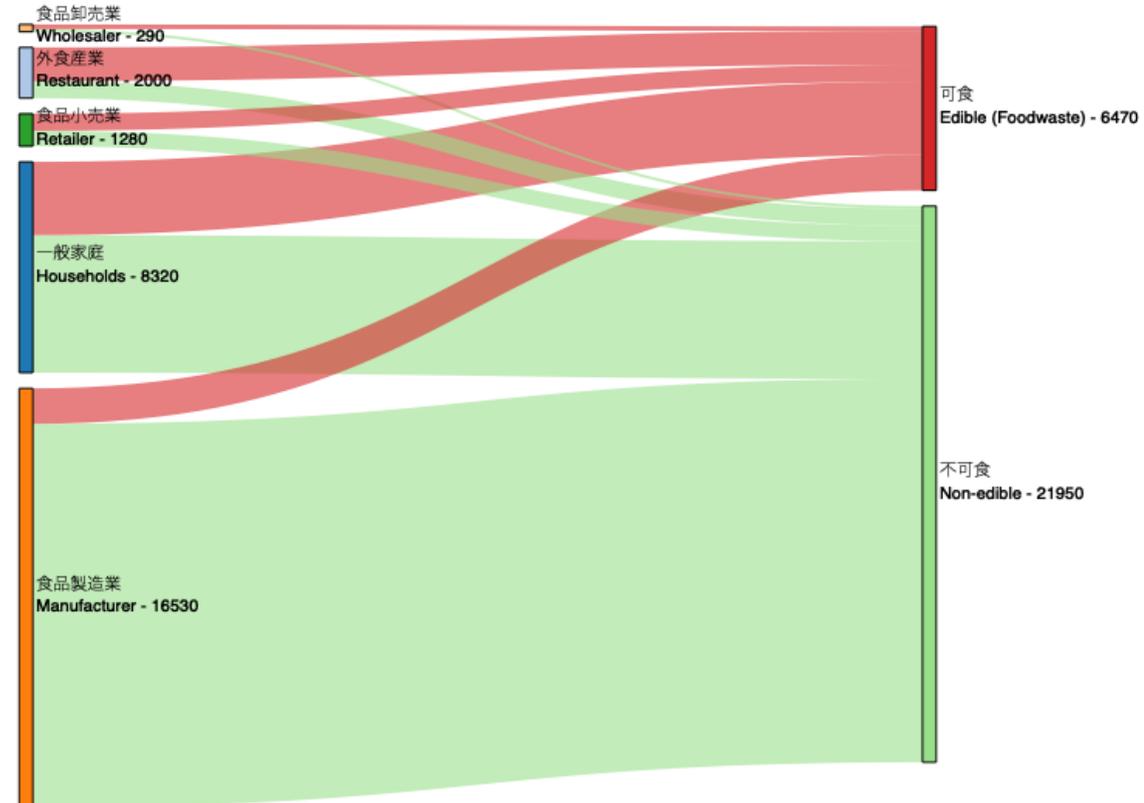
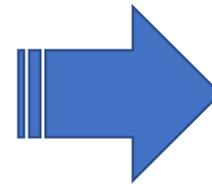
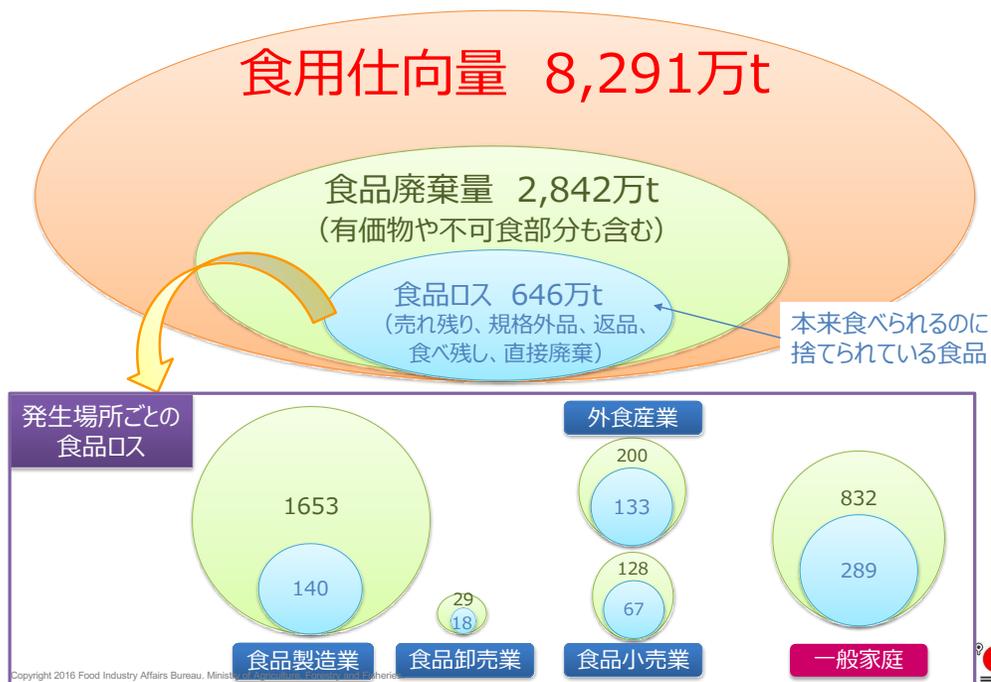
\*1 <https://holoviews.org/reference/elements/bokeh/Sankey.html>

\*2 <https://plotly.com/python/sankey-diagram/>

# 使ってみた1：日本の食品廃棄物の発生量（2015年度）

## ○ 食品廃棄物等の発生量（平成27年度推計）

農林水産省  
食料産業局



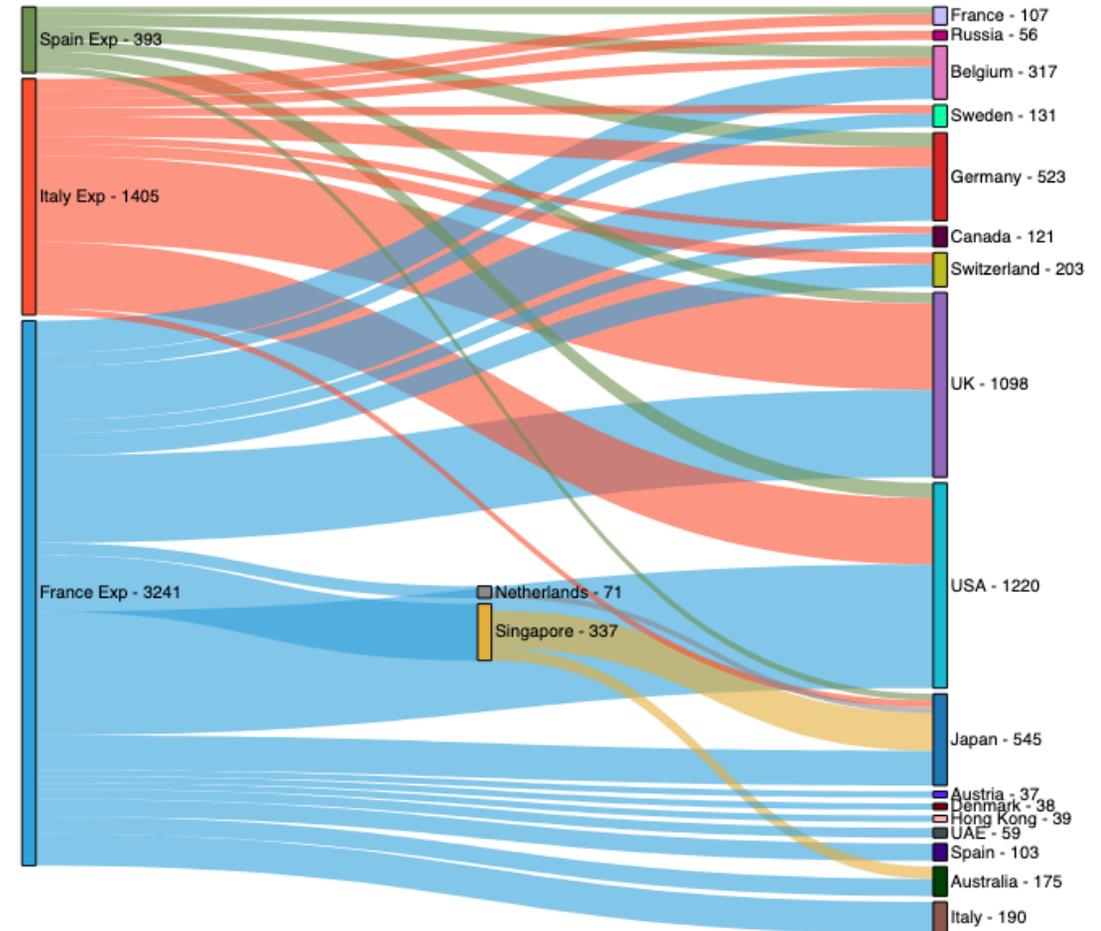
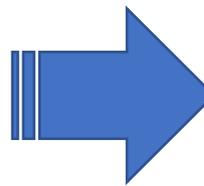
\* <https://www.vizforsocialgood.com/join-a-project/2019/4/1/jpfoodbank/>  
 \* <https://www.maff.go.jp/kinki/syouhi/mn/iken/attach/pdf/30nendo-8.pdf>

# 使ってみた2：スパークリングワインの主要取引ルート（2018年）

**aawe**  
wine economics

**World's Top Sparkling Wine Trade Routes in 2018**  
value in million US\$; Source: comtrade

1 France → USA	736	13 France → Spain	103	25 France → Utd Arab Emi	59
2 France → UK	520	14 France → Australia	103	26 Italy → Russia	56
3 Italy → UK	513	15 Spain → USA	90	27 Italy → Belgium	54
4 Italy → USA	394	16 Spain → Germany	85	28 Italy → Sweden	50
5 France → Singapore	337	17 France → Sweden	81	29 Spain → France	44
6 France → Germany	317	18 France → Canada	78	30 Italy → Canada	43
7 Singapore → Japan	225	19 Singapore → Australia	72	31 Italy → Japan	40
8 France → Japan	207	20 France → Netherlands	71	32 Spain → Japan	39
9 France → Belgium	193	21 Italy → Switzerland	71	33 France → CHN Hong Kong	39
10 France → Italy	190	22 Spain → Belgium	70	34 France → Denmark	38
11 France → Switzerland	132	23 Spain → UK	65	35 France → Austria	37
12 Italy → Germany	121	24 Italy → France	63	36 Netherlands → Japan	34



\* <https://twitter.com/wineecon/status/1260214505965060102>

# Next steps

- HoloViewsにノードの順序と位置の指定を実装したい。
  - 今はできない。
    - ノードに紐づく値の大きさをソート？
  - Plotlyなら位置を指定できるが、個人的にはHoloViewsの方が好み。
- 仕事のレポートでも使ってみたい。
  - 目的に応じて。
  - 量を伴う、時間・空間・状態の遷移

**Enjoy!**