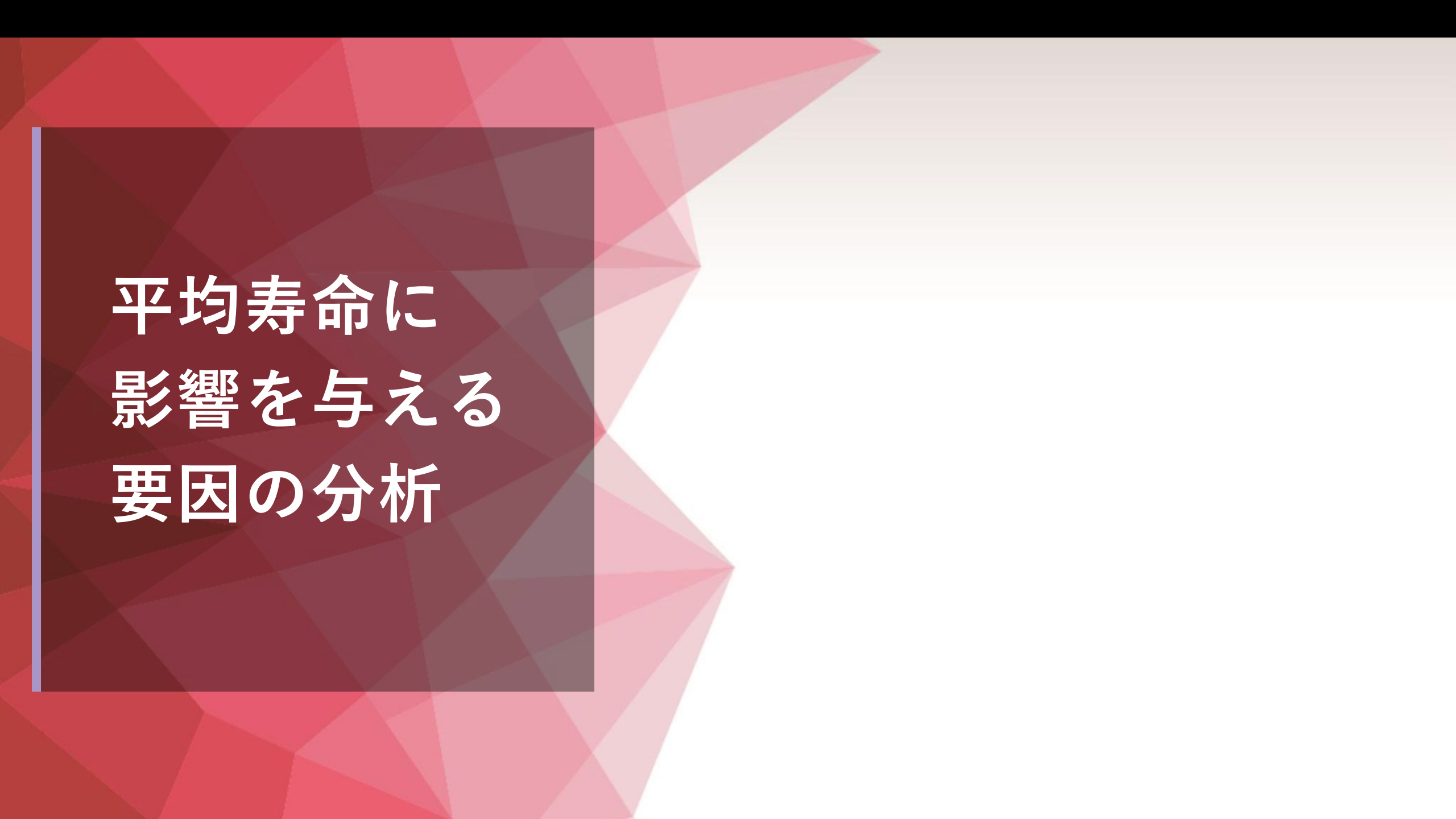




平均寿命に 影響を与える 要因の分析

1.概要

- ウェブ上で、様々な国の寿命に関するデータを見つけた。
このデータを分析することで、何が寿命に影響をもたらしているのかを知るのが目的である。
また、各項目間の相関を詳しく見ていくことで、GDPなどの寿命と相関があることが知られている要素以外にも寿命と相関がある項目を見つけることを期待している。

2. データ詳細

- Kaggle(<https://www.kaggle.com>)の、 Life Expectancy (WHO)(<https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who>)というデータを使用。
- WHOと国際連合のウェブサイトから集められた、2000年から2015年までの193カ国のデータ。
- 予防接種・死亡率・経済・社会・健康要因などの22項目×2938行。
- 22項目は以下の通り。

Country (国名), Year (年), Status (開発状況), Life expectancy (平均寿命), Adult Mortality (成人死亡率), infant deaths (乳児死亡数), Alcohol (アルコール消費量), percentage expenditure (健康支出割合), Hepatitis B (B型肝炎予防接種率), Measles (はしかの患者数), BMI (平均BMI), under-five deaths (5歳未満の死亡数), Polio (ポリオ予防接種率), Total expenditure (総医療支出割合), Diphtheria (ジフテリア予防接種率), HIV/AIDS (HIV/AIDSの死亡率), GDP (国内総生産), Population (人口), thinness 1-19 years (1～19歳の低体重率), thinness 5-9 years (5～9歳の低体重率), Income composition of resources (所得構成指数), Schooling (平均就学年数)

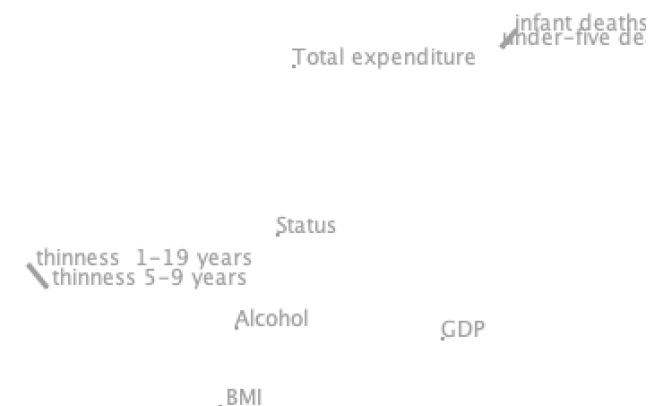
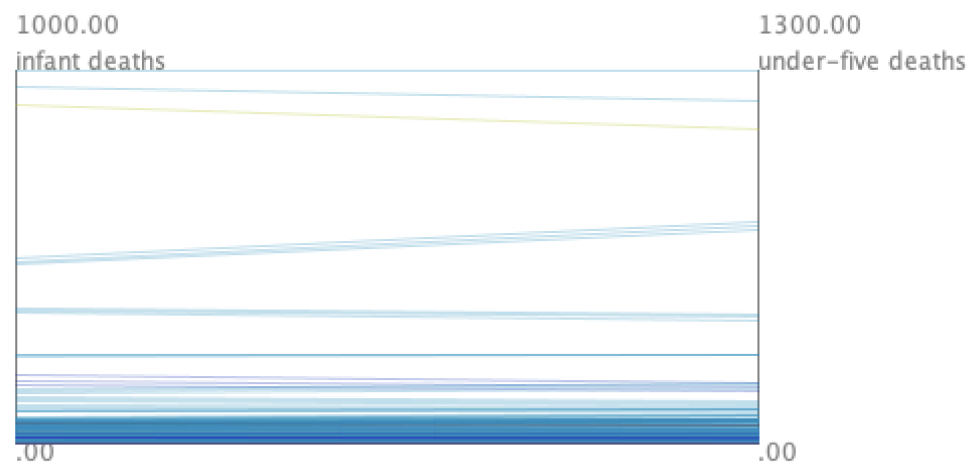
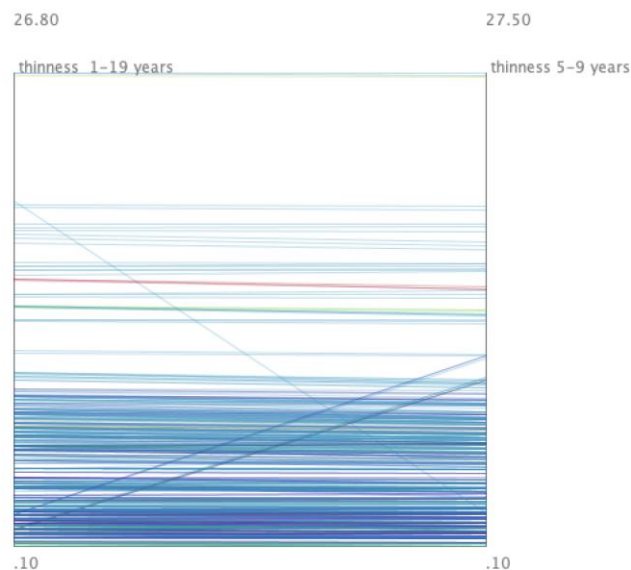
2. データ詳細

- 施した加工

- ・ 今回の目的上必要ないと考えたため、22項目のうち国名と年を削除。
- ・ Status列の値Developing, Developedをそれぞれ0,1に置換。
- ・ 空白はそのセルが属する列の平均値で埋める。
- ・ 2938行から561行に削減(各国の行を最大3つまでとした)。

4.分析結果

- 右上のバーを一番下から上げていく。最初に現れたのは以下のグラフであった。
- **分かること**
 - 1-19歳と5-9歳の低体重率には正の相関がある
 - 乳児死亡数と5歳未満の死亡数には正の相関がある
- 簡単に予測が出来そうな結果で、相関図でこの二つが最初に繋がったのは納得がいく。



4.分析結果

- 次に、相関図ではPolioとDiphtheriaが繋がり、以下のグラフが現れた。

- 分かること**

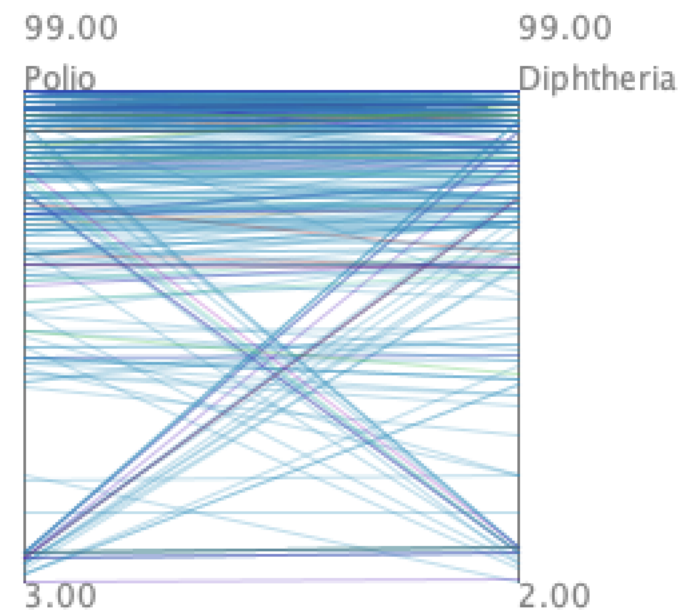
ポリオワクチン接種率と麻疹ワクチン接種率には正の相関があるが、一部負の相関も見受けられる。

- ある国において、特定のワクチンの接種率が高ければ他のワクチンの接種率も高いと考えるのは自然だと感じるが、一部に負の相関が見られた点が気になった。

元のデータを確認すると、発展途上国にこのような傾向があったことから医療資源が少ない中でどちらかの流行を抑えるために集中的に接種が進んだ可能性があるのではないかと考えた。

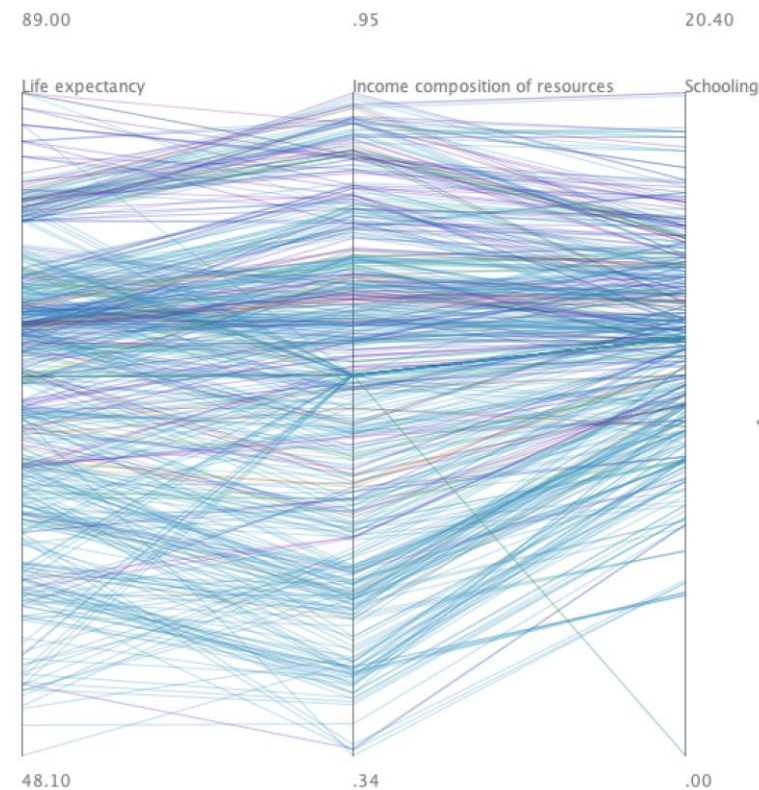
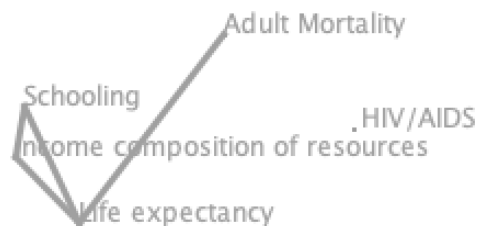
一例としてアゼルバイジャンの2006年のデータではDiphtheriaが高くPolioが低かった。

調べてみるとアゼルバイジャンは麻疹罹患率が高い国らしく、その年に麻疹ワクチンの一斉接種が行われた可能性が考えられる。



4.分析結果

- さらにバーを上げていくと、相関図でLife ExpectancyとIncome composition of resourcesが繋がり右下のグラフが現れた。
- **分かること**
 - ・ 上の二つの項目には正の相関がある
- 所得・教育・生活水準が良いほど寿命が伸びるということが分かる。さらにバーを上げると、Life Expectancyは以下の相関図のようにSchoolingにも繋がり、教育は寿命に大きく関係していそうだということが分かる。



4.分析結果

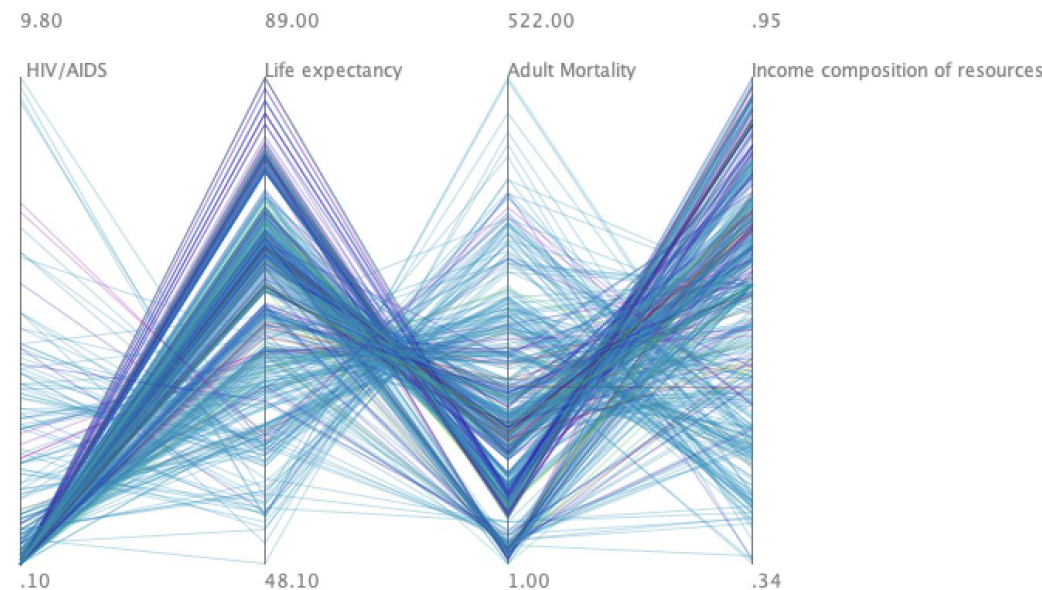
- 次に相関図上でLife Expectancyと繋がったのは、HIV/AIDSであった。また、以下のようなグラフが得られた。
- **分かること**
 - ・ 上の二つの項目には負の相関がある

- HIV/AIDSによる死亡率が低いほど寿命が長いということが分かる。

これは容易に理解できる。

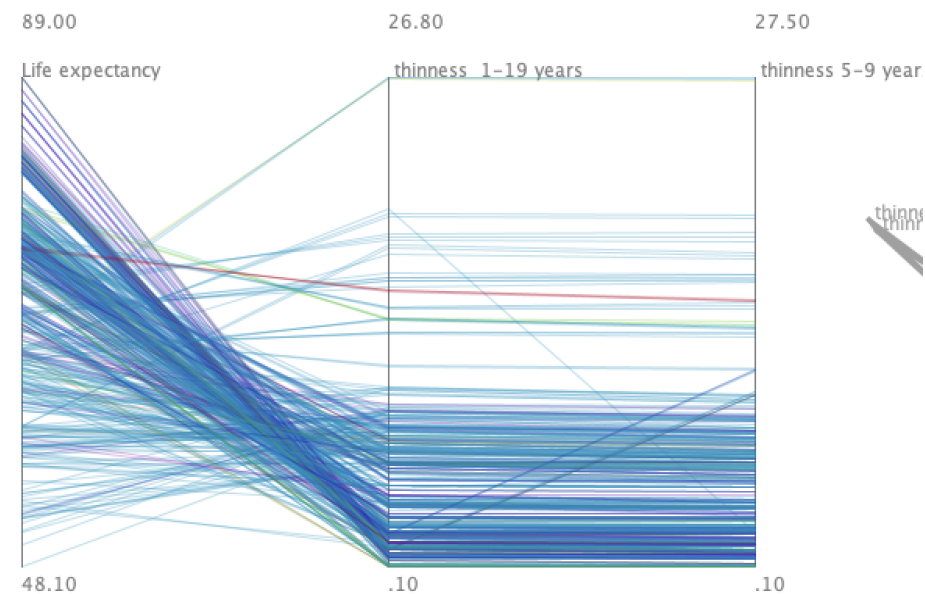
WHOによると2016年の低所得国の死因第4位はHIV/AIDSであり
発展途上国ではHIV/AIDSにより寿命が短くなっている可能性がある。

(<https://www.wakayama-med.jrc.or.jp/webmagazine/detail.php?seq=149>より)



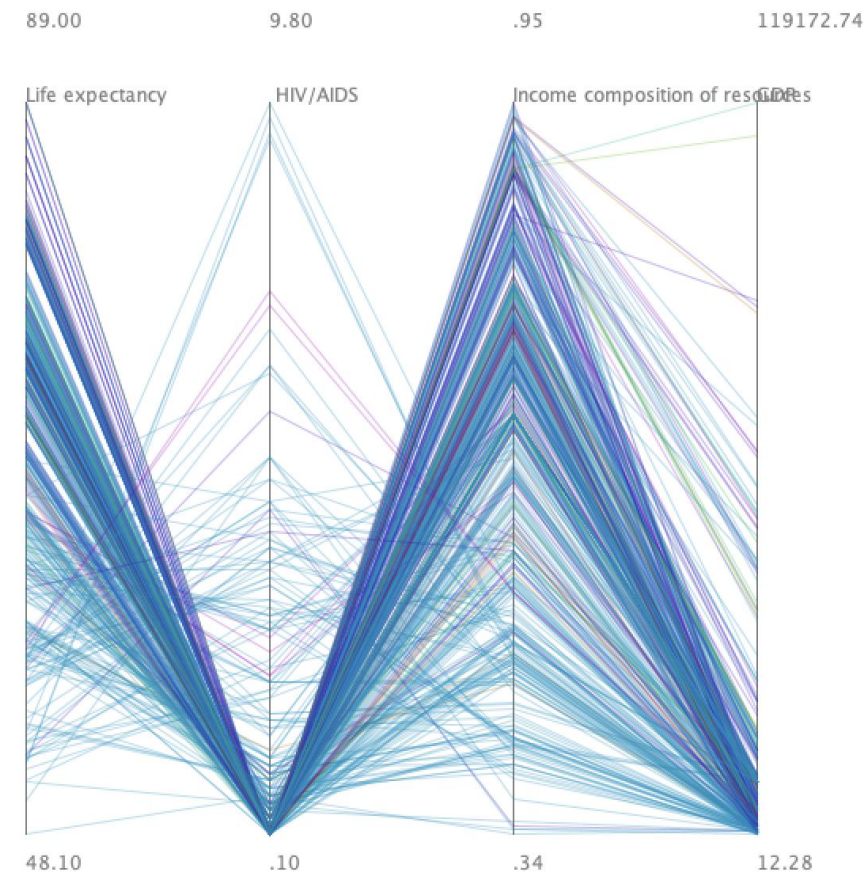
4.分析結果

- 次に相関図上でLife Expectancyと繋がったのは、thinness 1-19 yearsであった。また、以下のようなグラフが得られた。
- **分かること**
 - ・ 上の二つの項目には概ね負の相関がある
- 低体重率が低ければ栄養不足の人が少ないと予測できるため、寿命が伸びることは容易に理解できる。



4.分析結果

- さらにバーを上げていくと、以下のようなグラフが現れた。
- **分かること**
 - ・ HIV/AIDSとIncome composition of resourceには負の相関があり、Income composition of resourceとGDPには正の相関がある。
- 先ほどの「2016年の低所得国の死因第4位はHIV/AIDS」という情報と上の結果から、GDPとHIV/AIDSにも負の相関がある可能性が高いと考えられる。



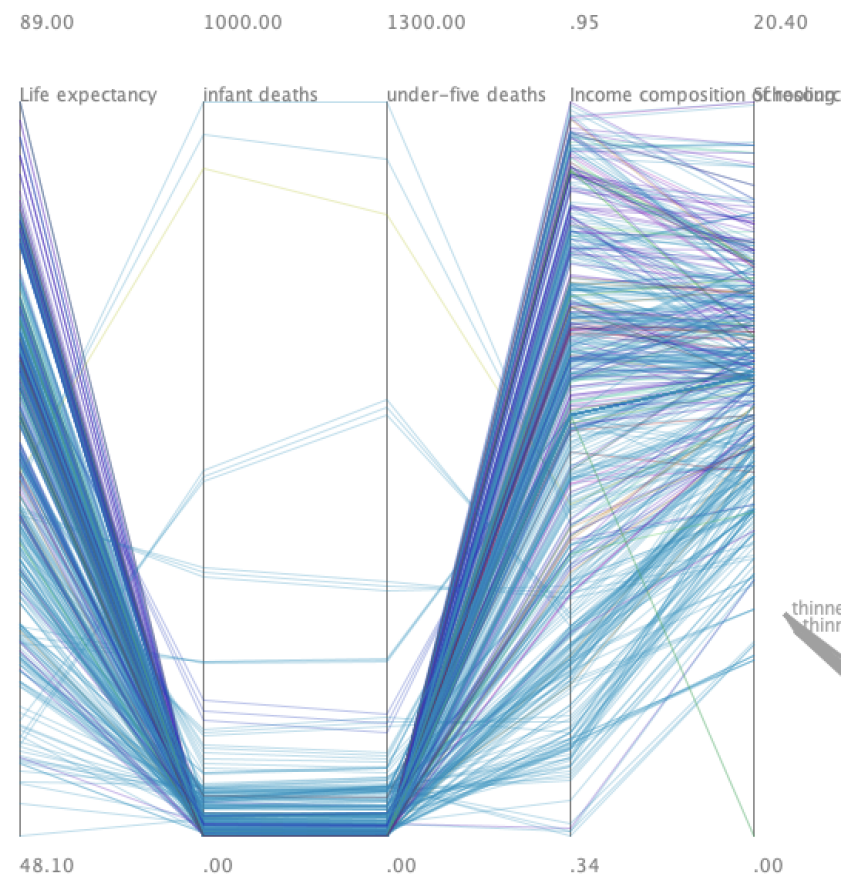
4.分析結果

- さらにバーを上げていくと、以下のようなグラフが現れ、相関図ではLife expectancyとInfant deathsが繋がった。

- **分かること**

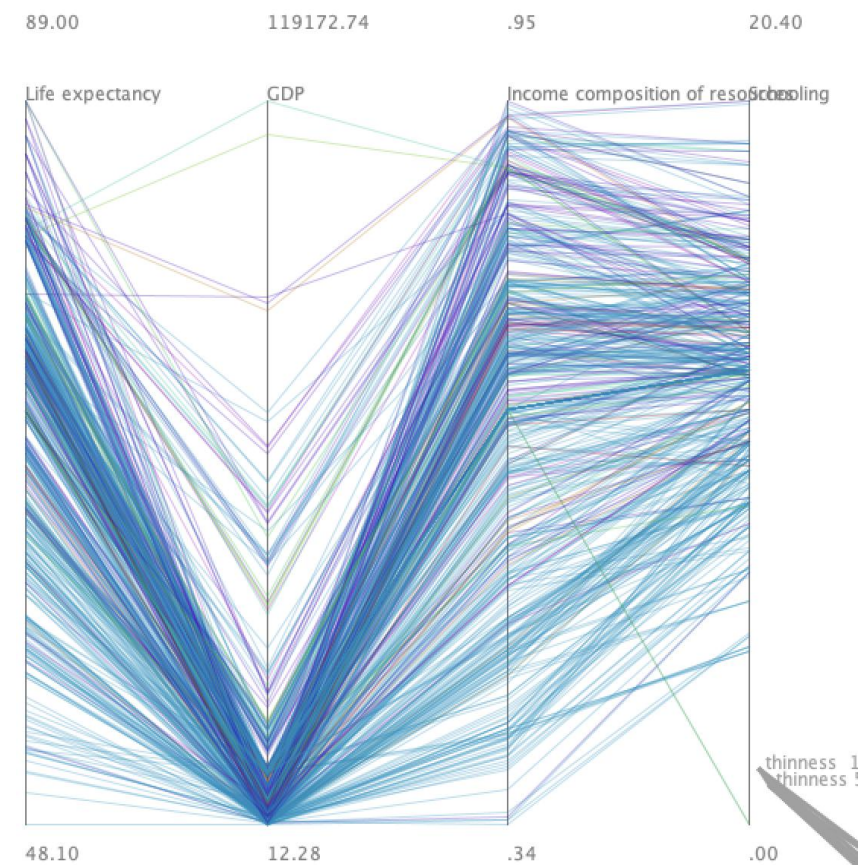
- Life expectancyとInfant deathsには正の相関がある
- 乳児死亡率が低いということから、以下のような乳児死亡の原因を排除できる程に医療体制が整っているということが予測できるので、寿命が伸びるのも容易に理解できる。
 1. 出産時の合併症
 2. 肺炎などの感染症
 3. 下痢
 4. マラリア

(https://www.worldvision.jp/children/education_04.html#d0e9d87eb78fa54e47cd213ca7606442より)



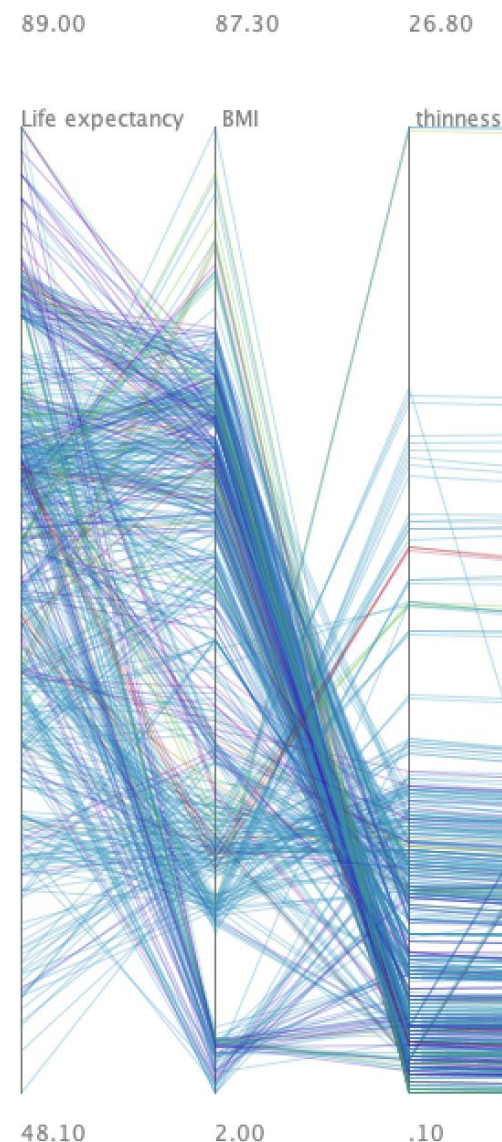
4.分析結果

- さらにバーを上げていくと、Life expectancyとGDPが繋がった。
- 分かること
 - ・ Life expectancyとGDPには正の相関がある
- GDPが低い国では医療・教育・栄養などが整っていないことが多いため寿命が短いことは容易に理解できる。



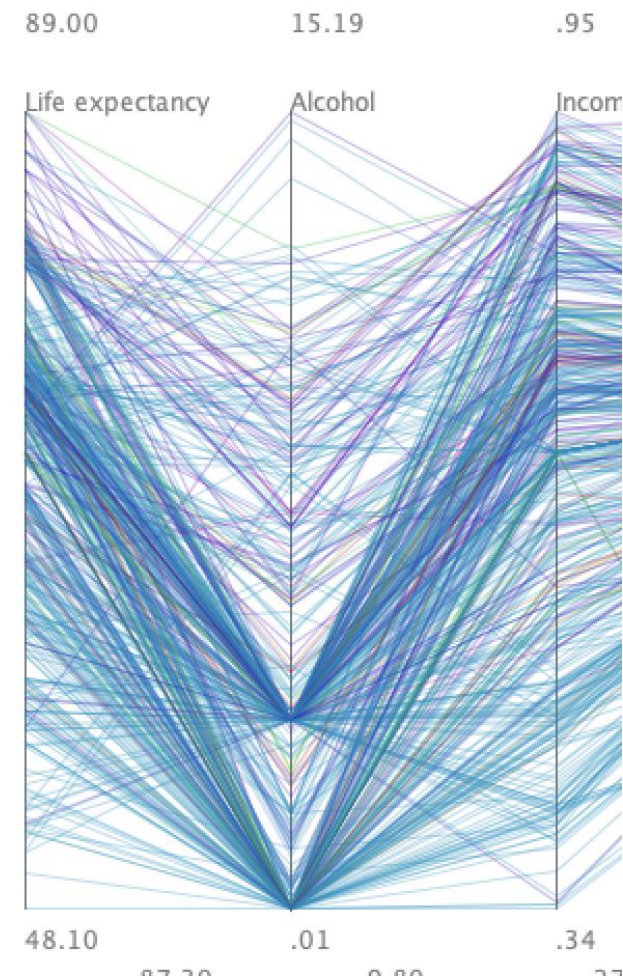
4.分析結果

- さらにバーを上げていくと、Life expectancyとBMIが繋がった。
- **分かること**
 - ・ 概ね正の相関があるように見える
 - ・ BMIが真ん中あたりを超えると、寿命が長い国が多くなる
 - ・ BMIが極端に高い国は寿命が短い
- 概ね正の相関があるのは、BMIが低いと栄養失調などの問題により寿命が短くなる可能性が高まることから理解できる。BMIが極端に高い国に関しては、生活習慣病などが寿命を短くしている可能性があると考えられる。



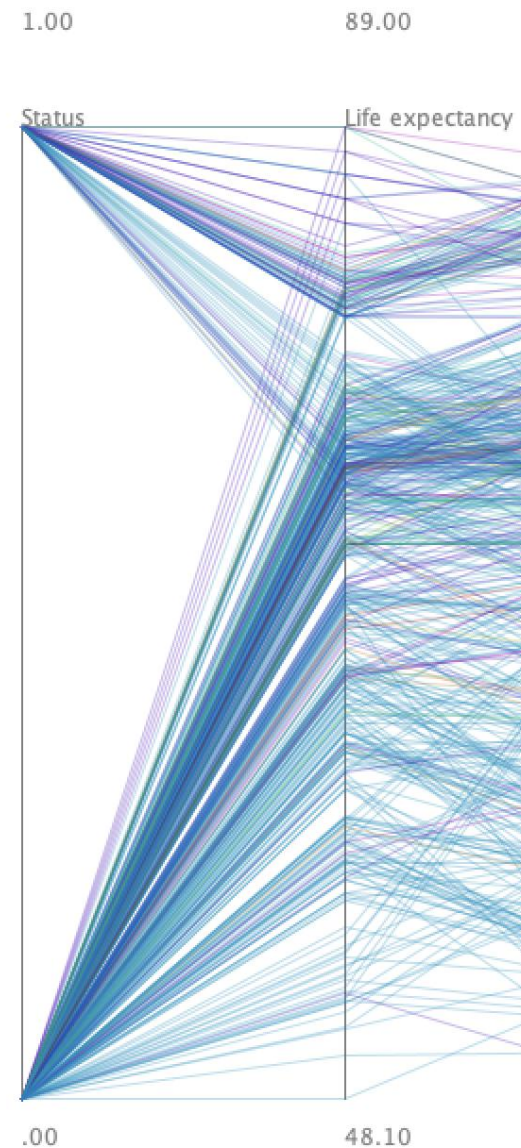
4.分析結果

- さらにバーを上げていくと、Life expectancyとAlcoholが繋がった。
- **分かること**
 - ・ Life expectancyとAlcoholには概ね正の相関がある
 - ・ しかしAlcoholが極端に大きい国の寿命は比較的短い
- Alcoholの大きさは生活水準と正の相関があると考えられるため、寿命と正の相関があることは理解できる。
元データを確認したところ、Alcoholが極端に大きい国はリトアニアであった。
調べてみると、リトアニアは一人当たりのアルコール消費量が15リットルと、世界平均の6.4リットルと比べても多く、過剰なアルコール摂取によって寿命が短くなっている可能性はあるかもしれない。
(<https://www.nikkei.com/article/DGXMZO35685370T20C18A9000000/>)



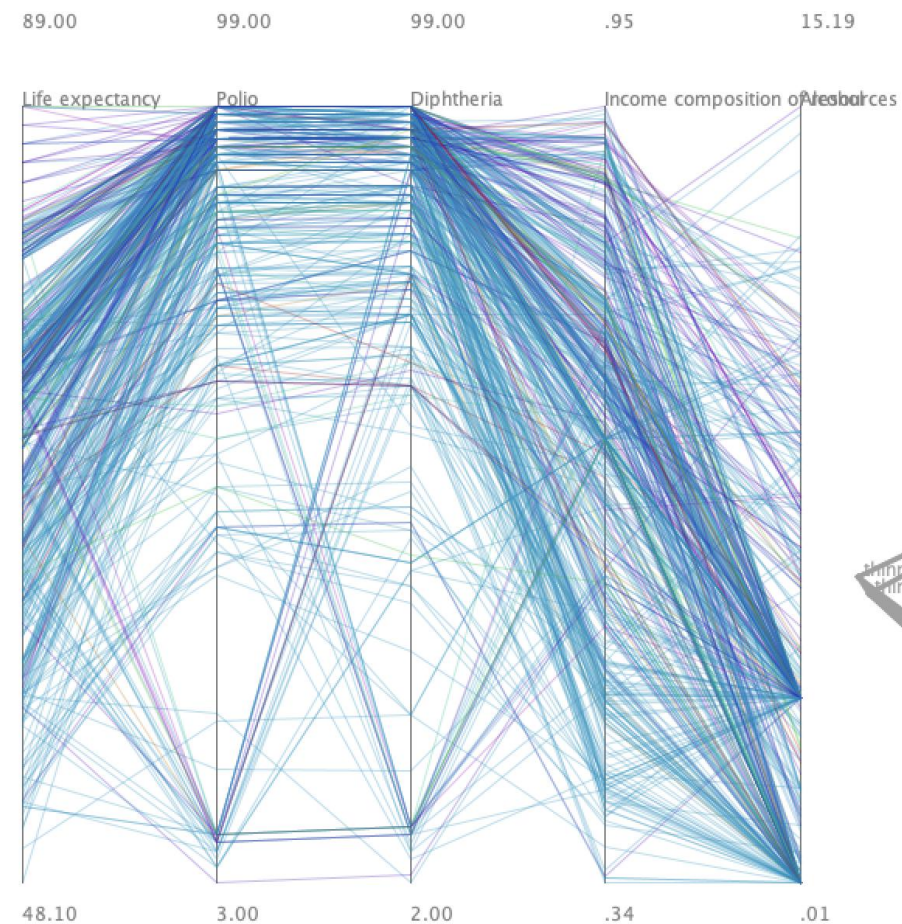
4.分析結果

- さらにバーを上げていくと、Life expectancyとStatusが繋がった。
- **分かること**
 - ・ 先進国は寿命が長く、発展途上国は寿命が短いという傾向がある
- これも、生活や教育、医療の水準の高さが影響していると考えられる



4.分析結果

- さらにバーを上げていくと、Life expectancyとPolioが繋がった。
- **分かること**
 - ・ Life expectancyとPolioには正の相関がある
- Polioはポリオワクチン接種率なので、高い方がポリオによる死亡率は低くなるし、医療水準も高いと考えられるため正の相関があることは納得がいく。



5.分析を終えて

- 寿命の長さとその他の項目について、もっと意外な関連性が見付かればと思っていたが、大体は容易に想像できるようなものだった。
- しかし、教育は思った以上に早い段階で寿命に影響を与える要素として現れたと感じた。それよりも身体的な健康の方が寿命を伸ばすことに貢献するのではないかと思っていた。