Global Well-Being Status Analysis

UML Group Project Presentation
Rui He, Becky Lau, Oliver (Yifeng) Tang

Collective well-being as a policy goal

Policymakers are increasingly focused on improving the well-being of its citizens.

“Gross National Happiness is distinguishable from Gross Domestic Product by valuing collective happiness as the goal of governance.”

Research questions

Given policymakers' interest in collective well-being

How do geographies around the world relate to each other in terms of well-being?

How do the patterns relate to what we already know about cross-cultural differences in well-being?

Given the difficulty in defining well-being and how multi-faceted well-being is

How do different kinds of well-being relate to one another?

Gallup World Poll - a high dimensional feature space
How do we reduce the complexity of this feature space?

- City or Area is a Perfect Place
- Experience Anger yesterday
- Experience Sadness yesterday
- Experience Stress yesterday
- Experience Worry yesterday
- Experienced Enjoyment yesterday
- Felt Active and productive
- Friends/Family give you positive energy
- Global Community well-being index
- Global financial well-being index
- Global Physical well-being index
- Global purpose well-being index
- Global Social Well-being index
- Global Well-being index
- Have enough money
- Learn or do something interesting
- Learn something
- Life evaluation index
- Life in five years
- Life today
- Like what you do each day
- Negative experience index
- Physical health near perfect positive health near perfect
- Positive experience index
- Recognition for improving city or area
- Smile or laugh
- Someone encourages your health
- Treated with respect
- Worried about money

Data Structure

Data from 139 geographies * 29 features

Each geography has one datapoint, aggregated from nationally representative samples

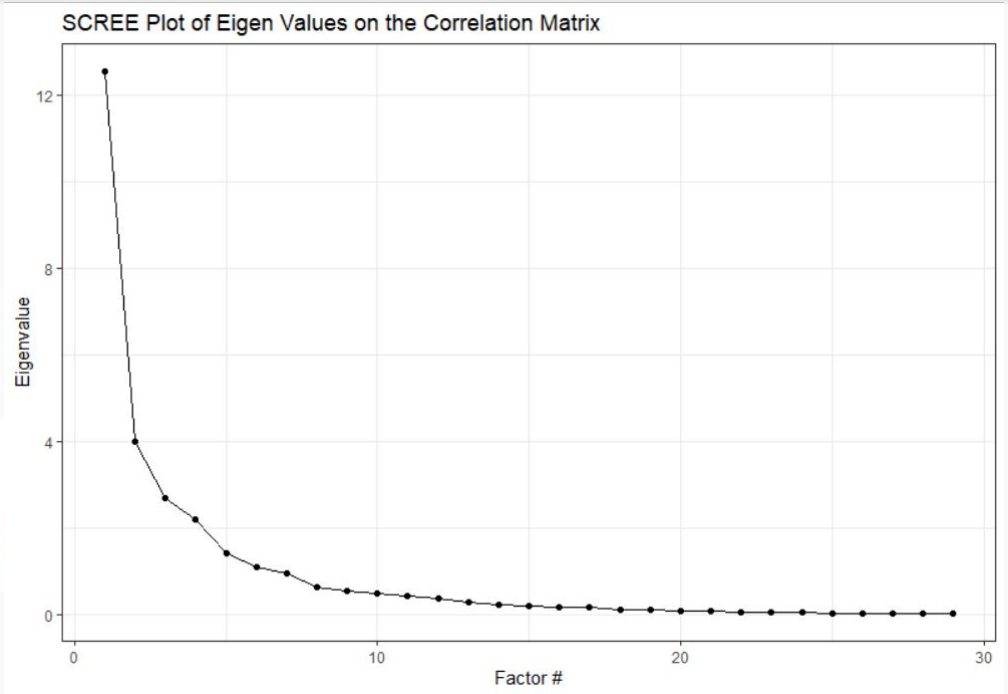
Transformed different variable types to make them comparable

- Binary Yes/No -> Percentage Yes's
- Numeric value -> No transformation
- Likert scale indicating agreement -> Average response
- Percentage Thriving, Struggling, or Suffering -> Percentage Thriving

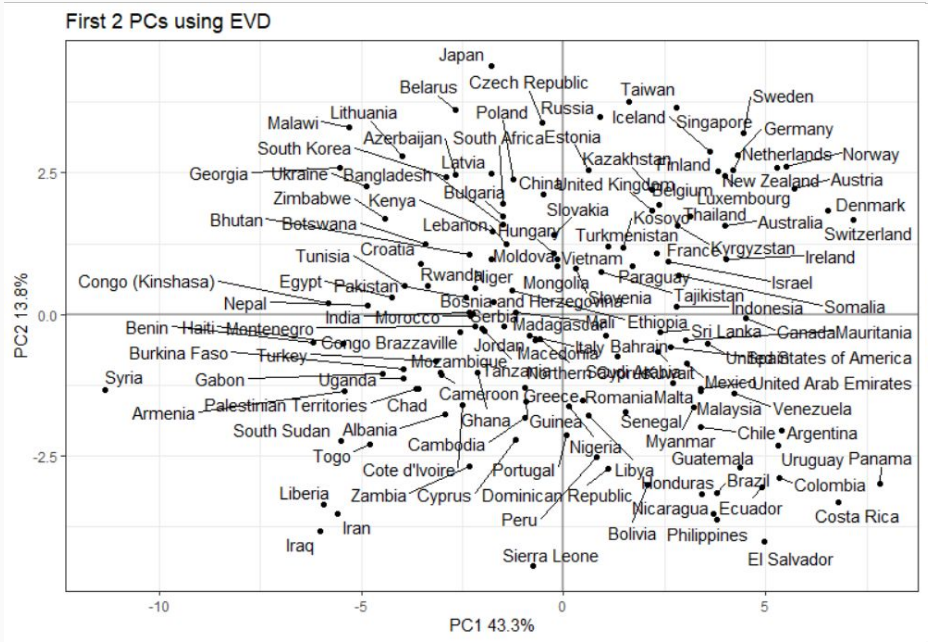
Principal Component Analysis

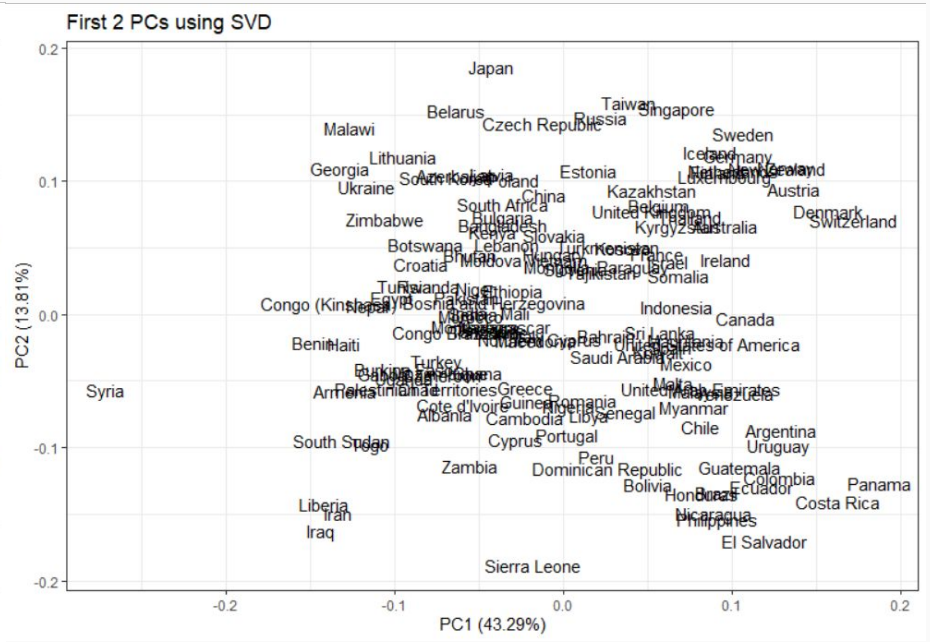
- High dimension
- High correlation
- **First 2 factors → 57%**
- First 4 factors → 74%

```
> summary(pca.out)
Importance of components:
                          PC1    PC2     PC3     PC4
Standard deviation     3.5432 2.0016 1.64018 1.48580
Proportion of Variance 0.4329 0.1381 0.09277 0.07612
Cumulative Proportion  0.4329 0.5711 0.66382 0.73995
```

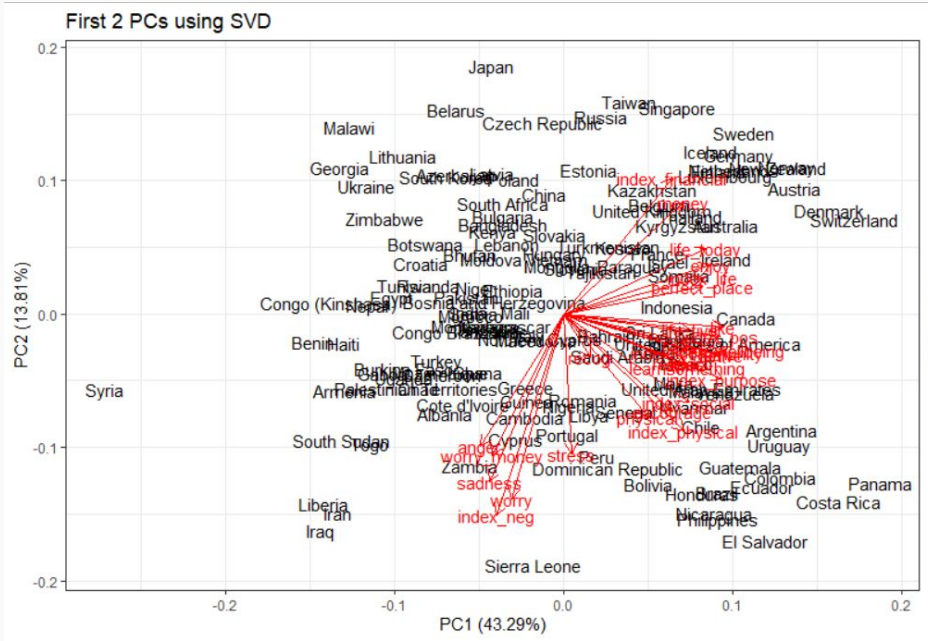


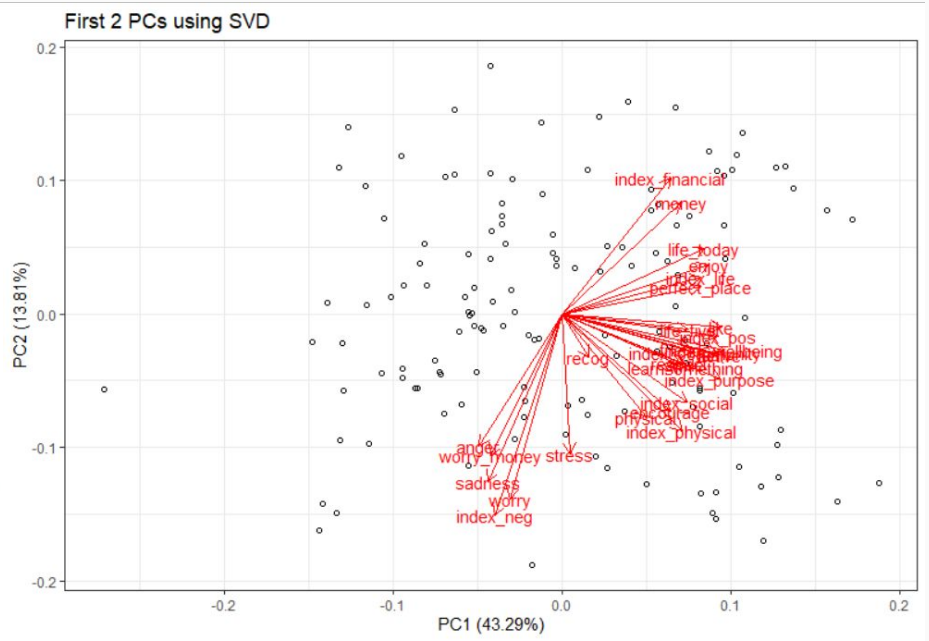
Principal Component Analysis

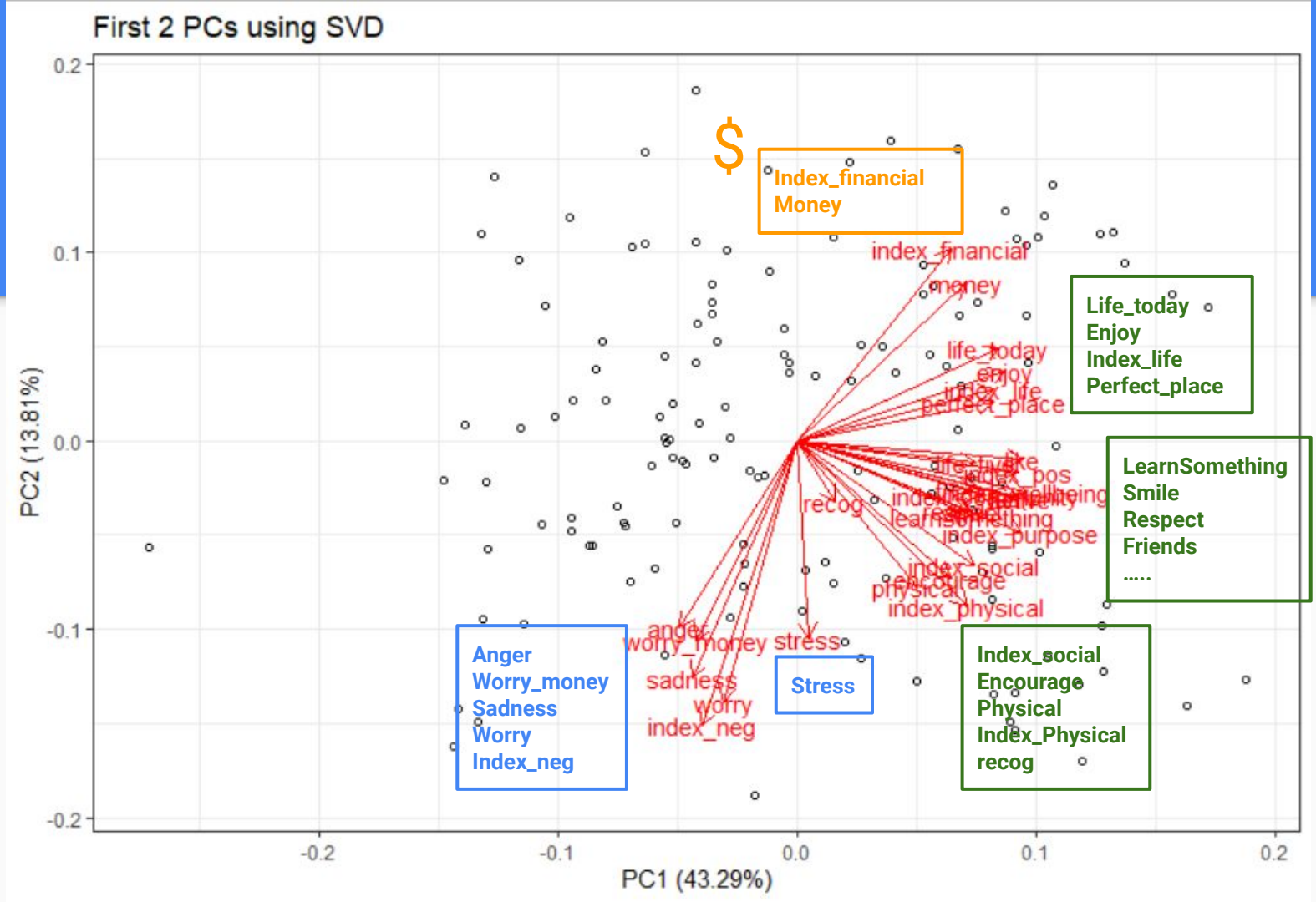




Principal Component Analysis



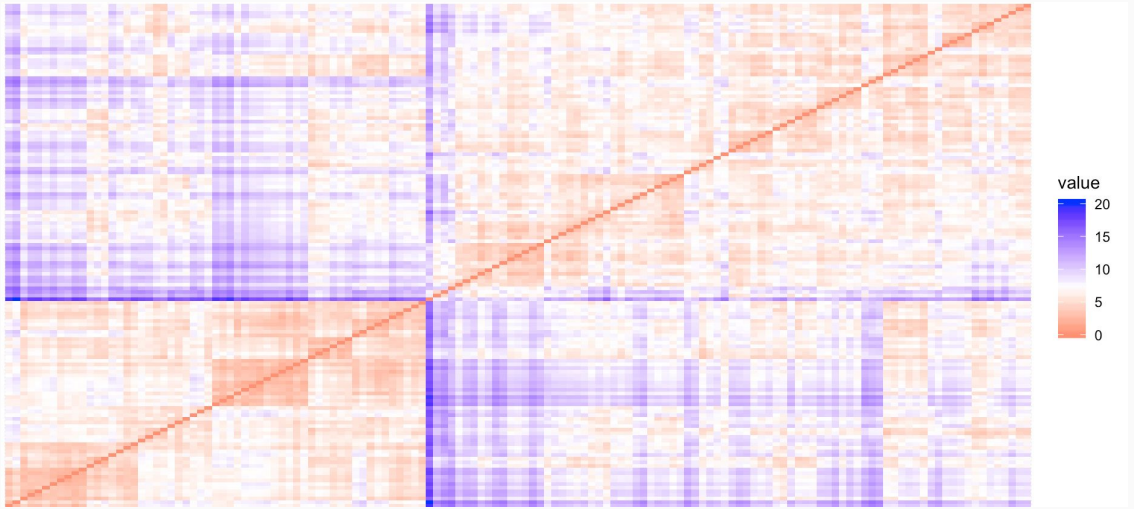



First 2 PCs using SVD
PC2 (13.81%)
PC1 (43.29%)
0.2
0.1
0.0
-0.1
-0.2
-0.2
-0.1
0.0
0.1
0.2
$
Index_financial
Money
index_financial
money
life_today
enjoy
index_life
perfect_place
Life_today
Enjoy
Index_life
Perfect_place
life_like
index_pos
recog
learnsomething
index_purpose
LearnSomething
Smile
Respect
Friends
.....
index_social
encourage
physical
index_physical
Index_social
Encourage
Physical
Index_Physical
recog
anger
worry_money
stress
sadness
worry
index_neg
Anger
Worry_money
Sadness
Worry
Index_neg
Stress

K-Mean clustering

Clusterability Diagnosis

- ODI
- Hopkins stats: 0.29



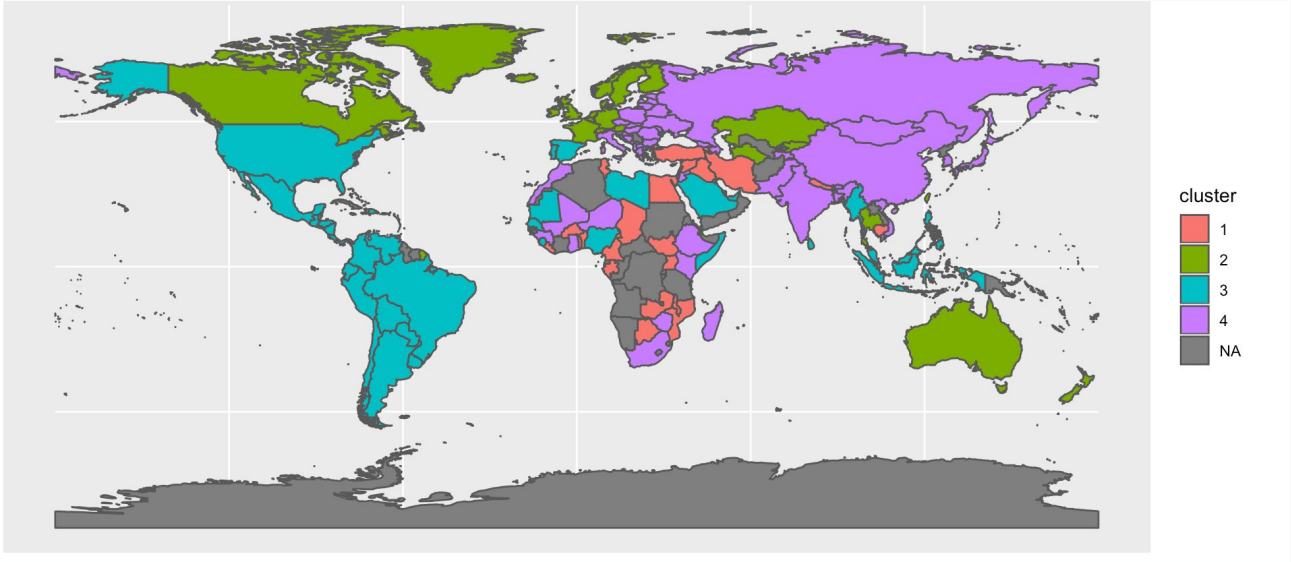
K = 4

- Internal validation

Results

Geospatial patterns:

1. Middle East + Africa (25 geographies)
2. Europe + Australia + Canada (24)
3. South American + US South Asia (37)
4. Asia + East Europe + Africa (53)



Results

1. Middle East + Africa (25)
2. Europe + Australia (24)
3. South American +US (37)
4. Asia, East Europe (53)

Cluster Means:

1. High negative, low positive
2. Low negative, high positive
3. High negative, high positive
4. Low negative, low positive

```
Cluster means:
        anger    sadness     stress      worry
1  1.07294270  1.2357999  0.3245685  0.9499872
2 -0.82407227 -0.7328706 -0.2449325 -0.8519976
3  0.07586473  0.1697124  0.3867693  0.3618213
4 -0.18590242 -0.3695370 -0.3121943 -0.3148892
      active     friends       money       learn
1 -1.3636515 -1.32785659 -0.8027719 -0.9074661
2  0.5743103  0.25440854  1.3905342  0.8050132
3  0.9352474  0.64812919  0.1350527  0.9136423
4 -0.2697418  0.05867604 -0.3452919 -0.5743100
```

Discussion

- Geospatial patterns observed
- Reflect the physical/social/financial/political environments of the regions
- Explanation
  - Emotion expressiveness?
  - Major sociopolitical factors?
- Limitations
  - aggregate data
  - potential sampling bias
  - cultural bias of questions