CE 263 final project

Happiness model:
Finding features correlated with happiness worldwide

harry durbin

"The very purpose of our life is to seek happiness."
— Dalai Lama XIV

Infectiously spread like a disease?
(spatial localization)

Dependent on external events?
(might have spatial localization)

Completely Independent
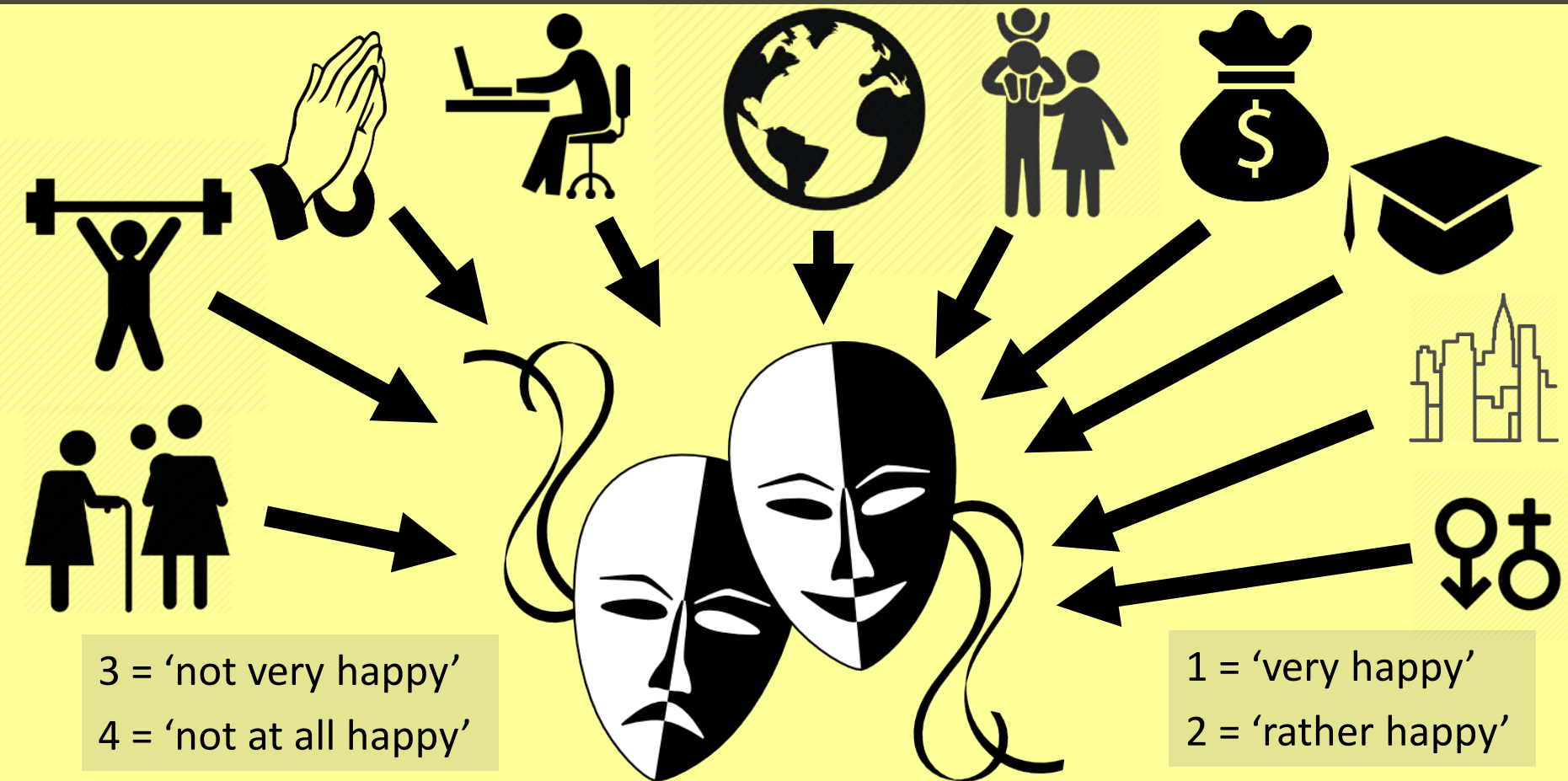(no spatial localization)

3 = 'not very happy'

4 = 'not at all happy'

1 = 'very happy'

2 = 'rather happy'

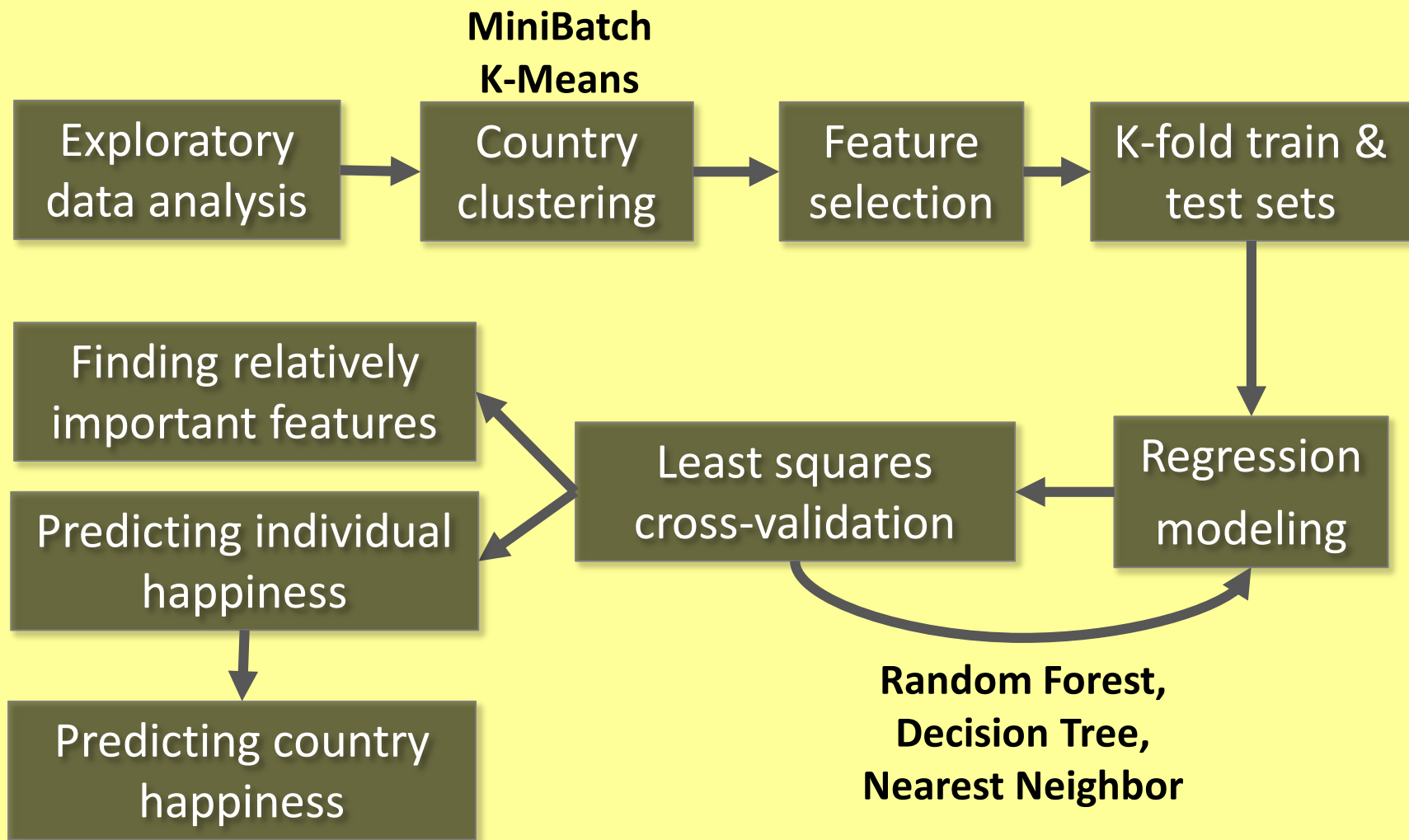**Two goals**:

1. Determine which features correlate most with happiness.

2. Create a model to try to predict happiness levels.

Approach

MiniBatch K-Means

Exploratory data analysis → Country clustering → Feature selection → K-fold train & test sets

Finding relatively important features

Predicting individual happiness ↔ Least squares cross-validation ← Regression modeling

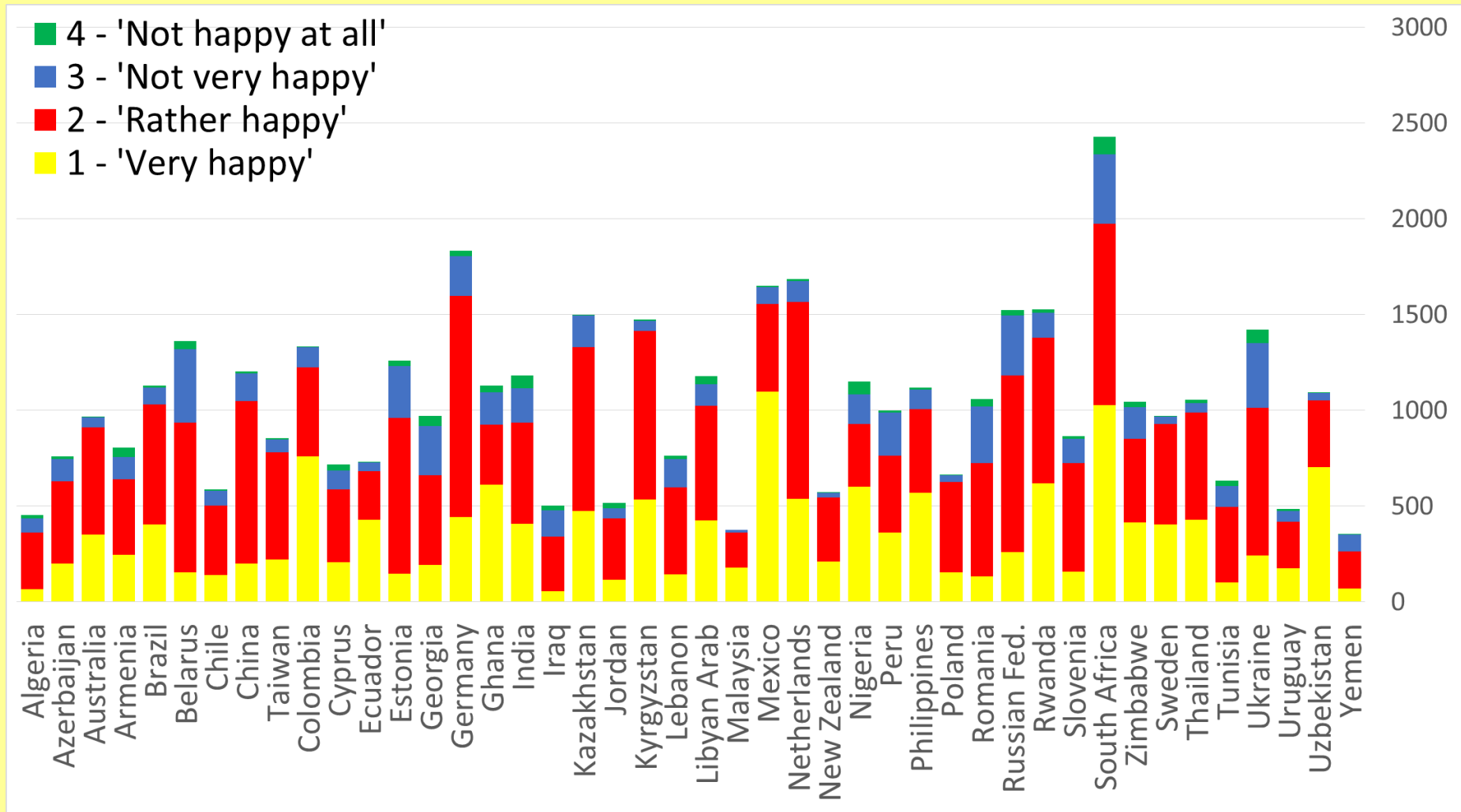Predicting country happiness

Random Forest, Decision Tree, Nearest Neighbor

Training data: 45k surveys

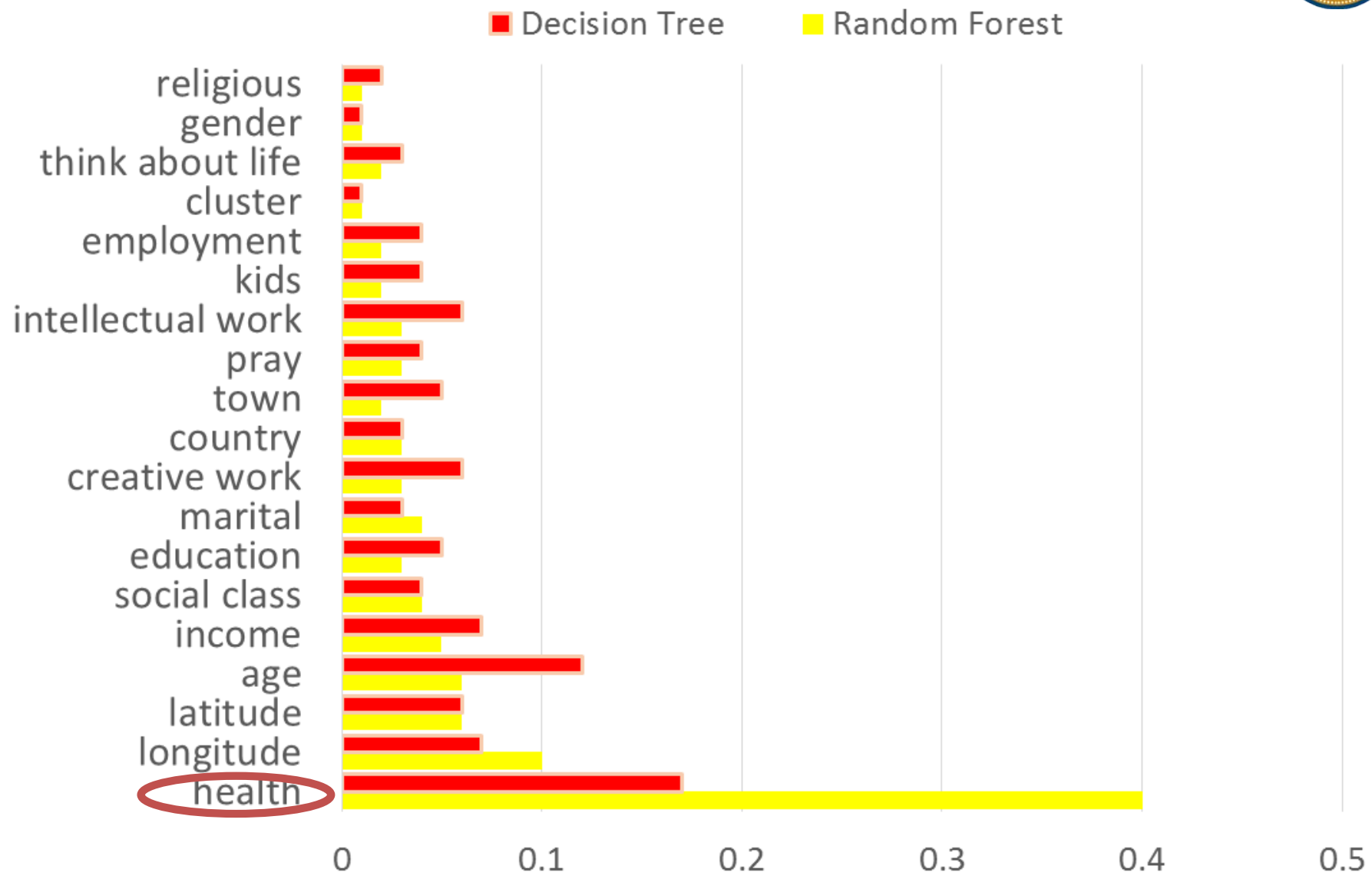**Tuning parameters**: feature selection, # trees (RF), depth of trees, # neighbors (NN)

Most people across the world say they are happy

See ya, I'm moving to **Uzbekistan!**

health has the strongest correlation with happiness

There is a mind and body connection

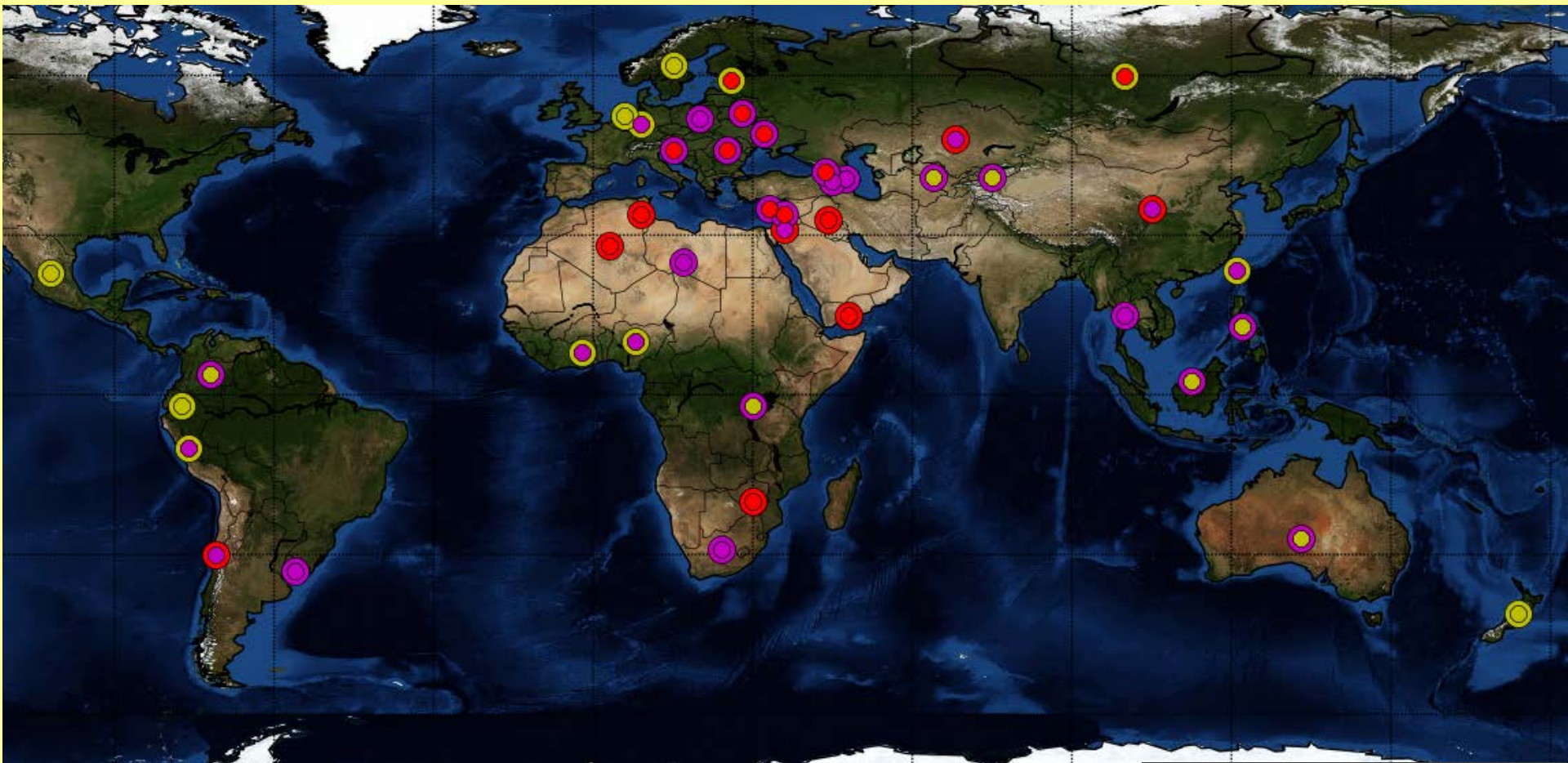# Average happiness predictions are so-so accurate

Country	Actual	Predicted
Uzbekistan	1.39	1.80
Mexico	1.40	1.75
Ecuador	1.48	1.73
Colombia	1.52	1.76
Malaysia	1.56	1.75
Philippines	1.60	1.70
Sweden	1.63	1.64
Ghana	1.67	1.76
Thailand	1.67	1.83
Kyrgyzstan	1.68	1.76
New Zealand	1.68	1.78
Australia	1.69	2.07
Rwanda	1.70	1.66
Nigeria	1.72	1.55
Brazil	1.74	1.98
Netherlands	1.76	1.89
Uruguay	1.79	1.63
Kazakhstan	1.80	1.84
South Africa	1.80	2.07
Libyan Arab	1.81	2.01
Zimbabwe	1.82	1.87
Poland	1.82	1.77

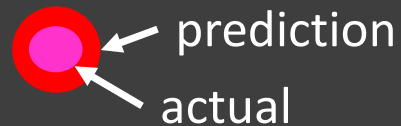| Country	Actual	Predicted
Taiwan	1.84	1.87
Peru	1.88	1.38
Germany	1.90	1.91
Chile	1.92	2.23
India	1.92	2.24
Azerbaijan	1.92	2.03
Cyprus	1.94	1.90
Armenia	1.97	2.07
China	1.97	2.08
Jordan	1.98	1.90
Slovenia	2.00	1.98
Lebanon	2.05	2.02
Yemen	2.06	1.61
Russian Fed.	2.07	1.77
Algeria	2.09	2.12
Tunisia	2.10	1.73
Estonia	2.14	1.52
Ukraine	2.17	1.78
Georgia	2.18	1.88
Romania	2.23	1.76
Belarus	2.23	1.93
Iraq	2.26	2.15

1 = 'very happy'
2 = 'rather happy'
3 = 'not very happy
4 = 'not at all happy'

● < 1.75   ● < 2.00   ● < 2.50   prediction
actual

"happiness depends upon ourselves."
- Aristotle