

A young boy with light hair and a blue baseball cap is laughing heartily with his mouth wide open. He is wearing a plaid short-sleeved shirt and light-colored pants. He is sitting on a weathered wooden bench. The background is a bright, slightly overexposed outdoor scene with greenery and sunlight.

MODULE 5 PROJECT

**UNDERSTANDING
HAPPINESS**

2017

1	Norway
2	Denmark
3	Iceland
4	Switzerland
5	Finland
6	Netherlands
7	Canada
8	New Zealand
9	Australia
10	Sweden
11	Israel
12	Costa Rica
13	Austria
14	United States
15	Ireland
16	Germany
17	Belgium
18	Luxembourg
19	United Kingdom
20	Chile

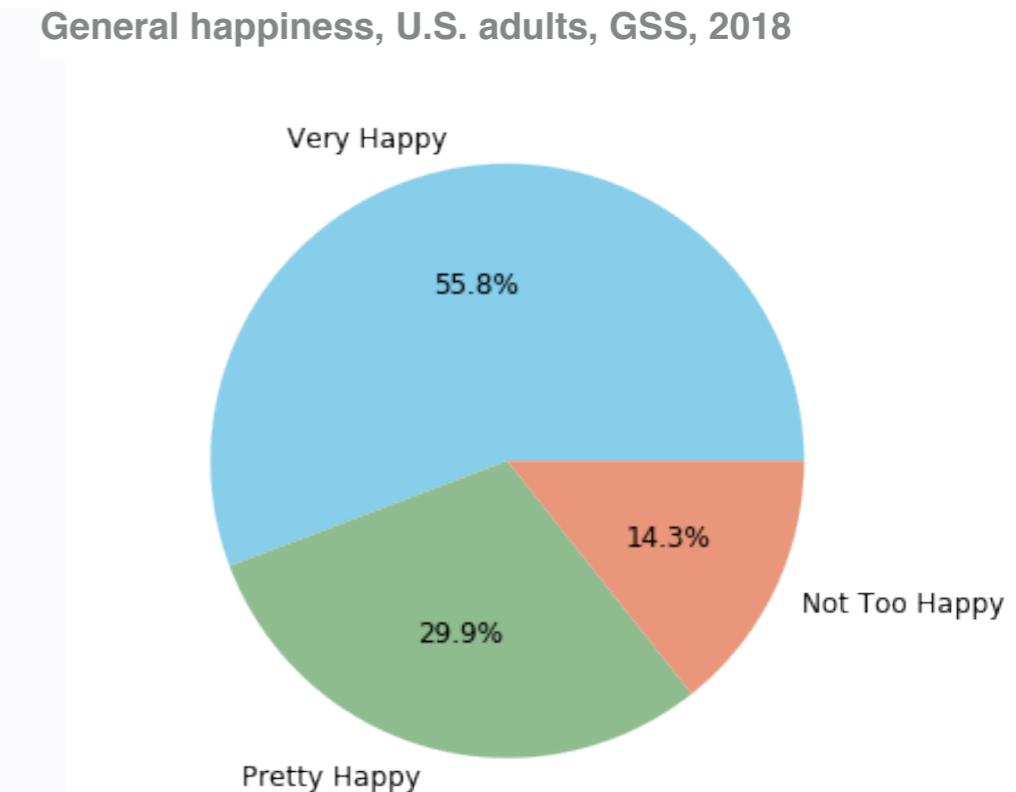
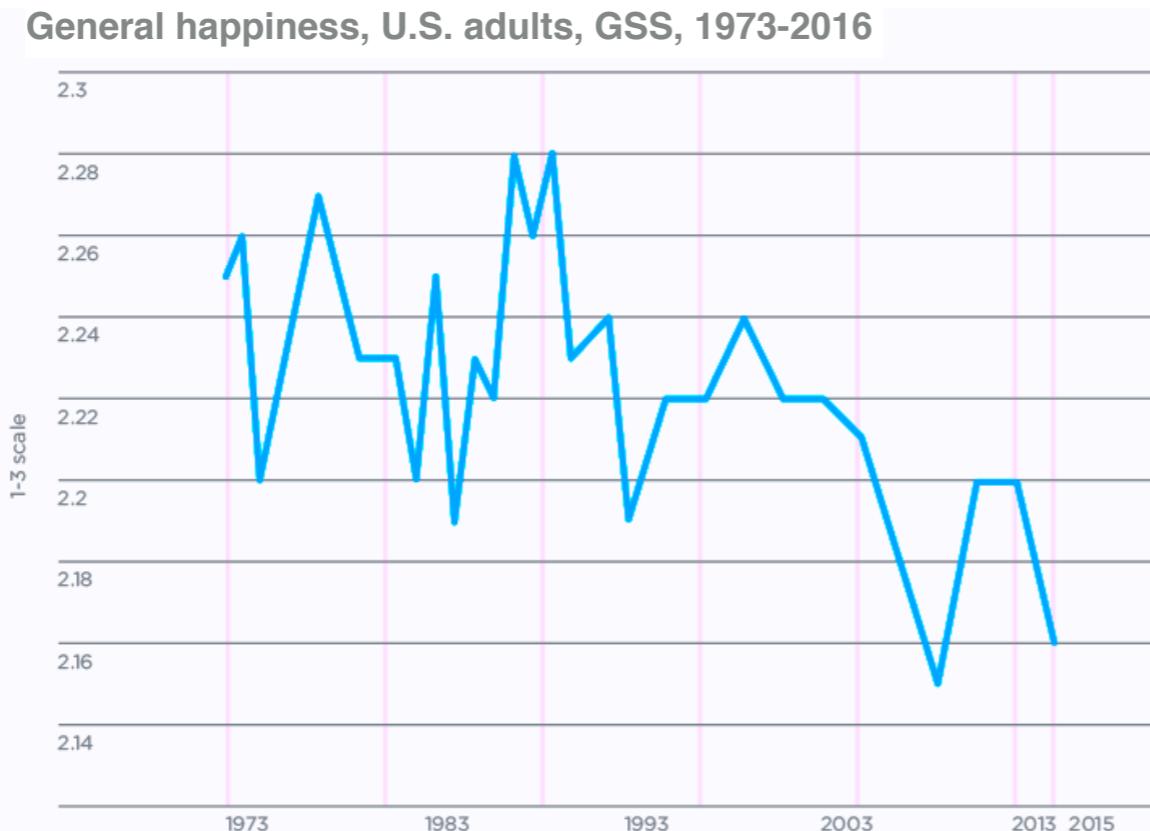
2018

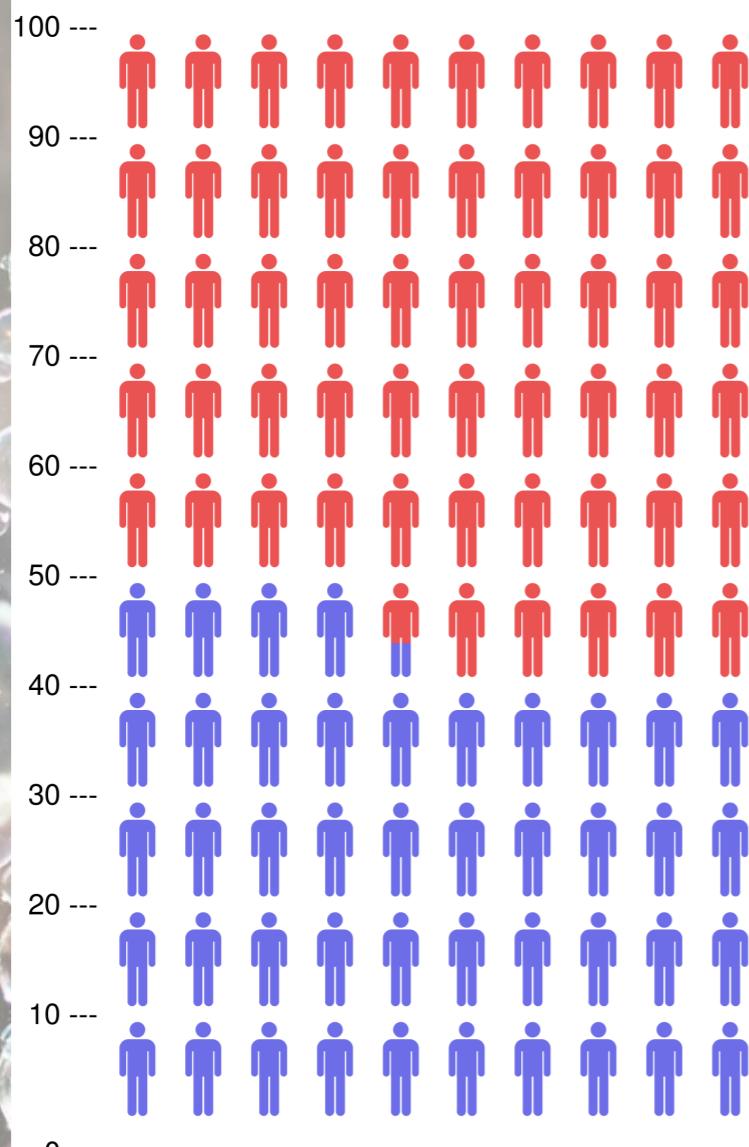
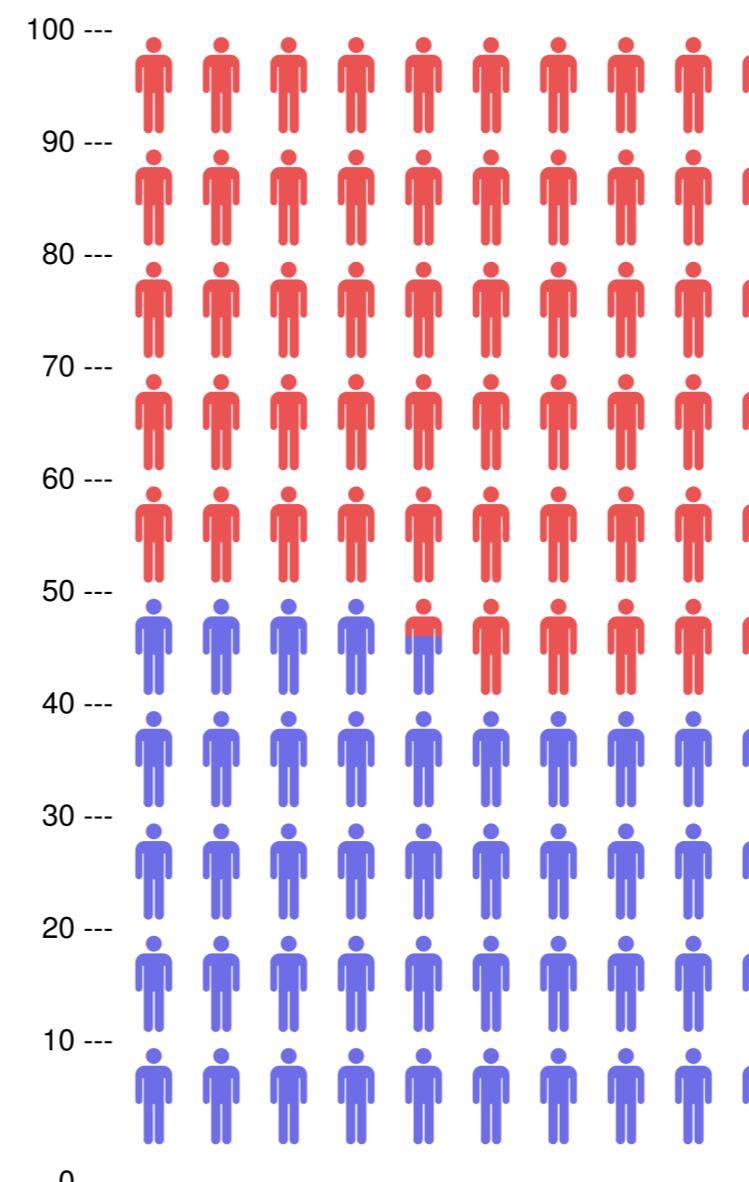
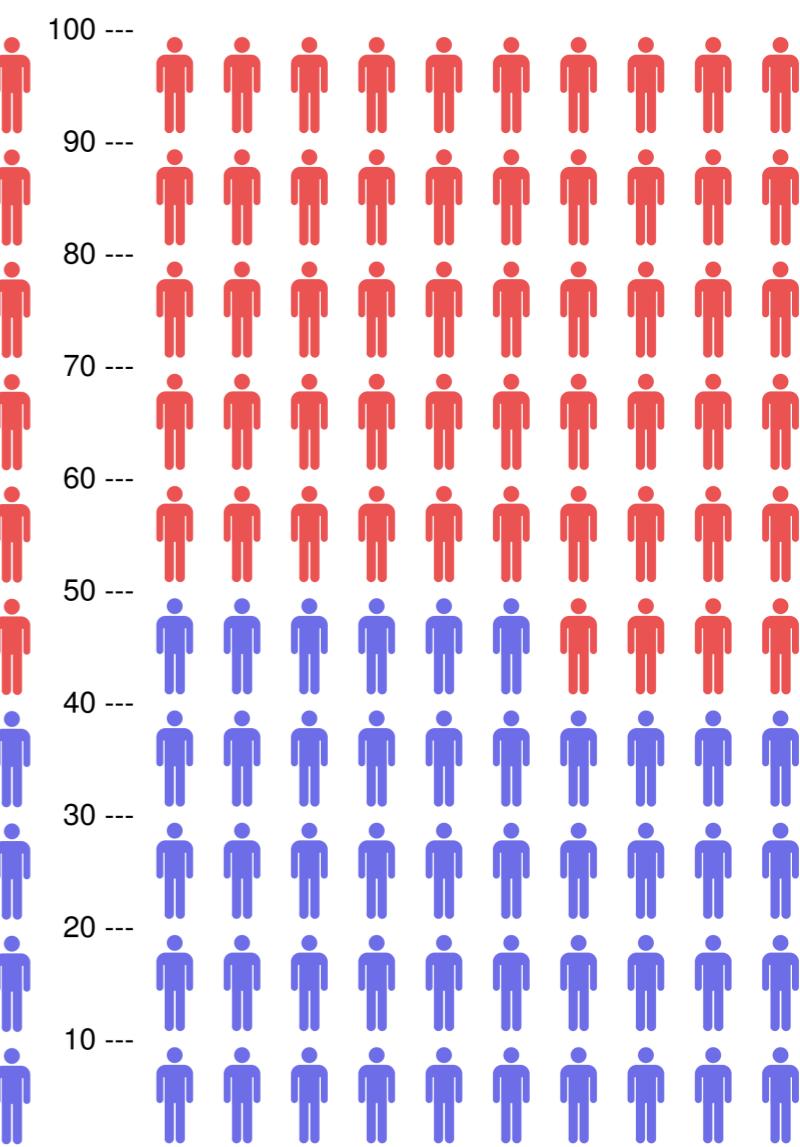
1	Finland
2	Norway
3	Denmark
4	Iceland
5	Switzerland
6	Netherlands
7	Canada
8	New Zealand
9	Sweden
10	Australia
11	United Kingdom
12	Austria
13	Costa Rica
14	Ireland
15	Germany
16	Belgium
17	Luxembourg
18	United States
19	Israel
20	United Arab Emirates

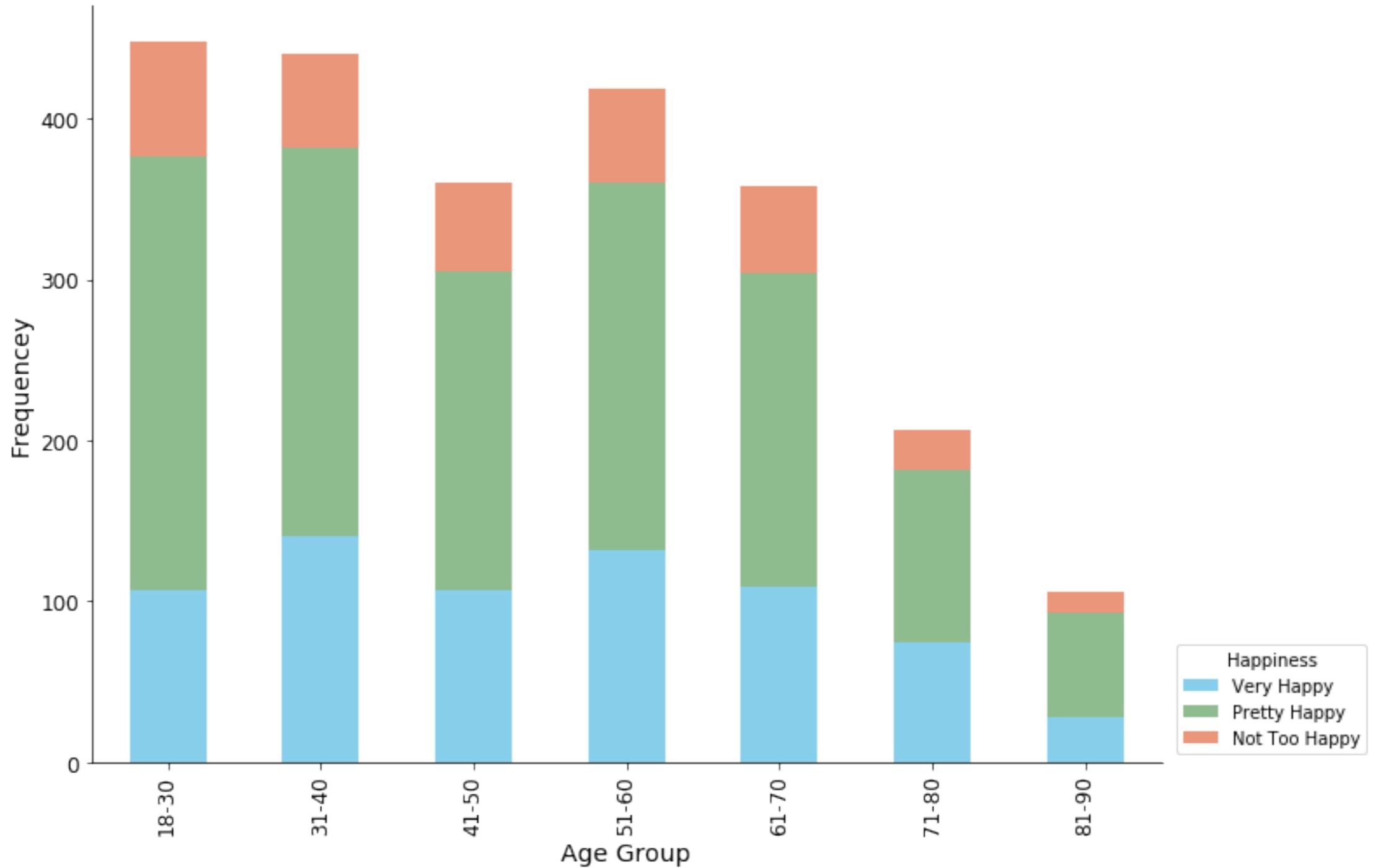
2019

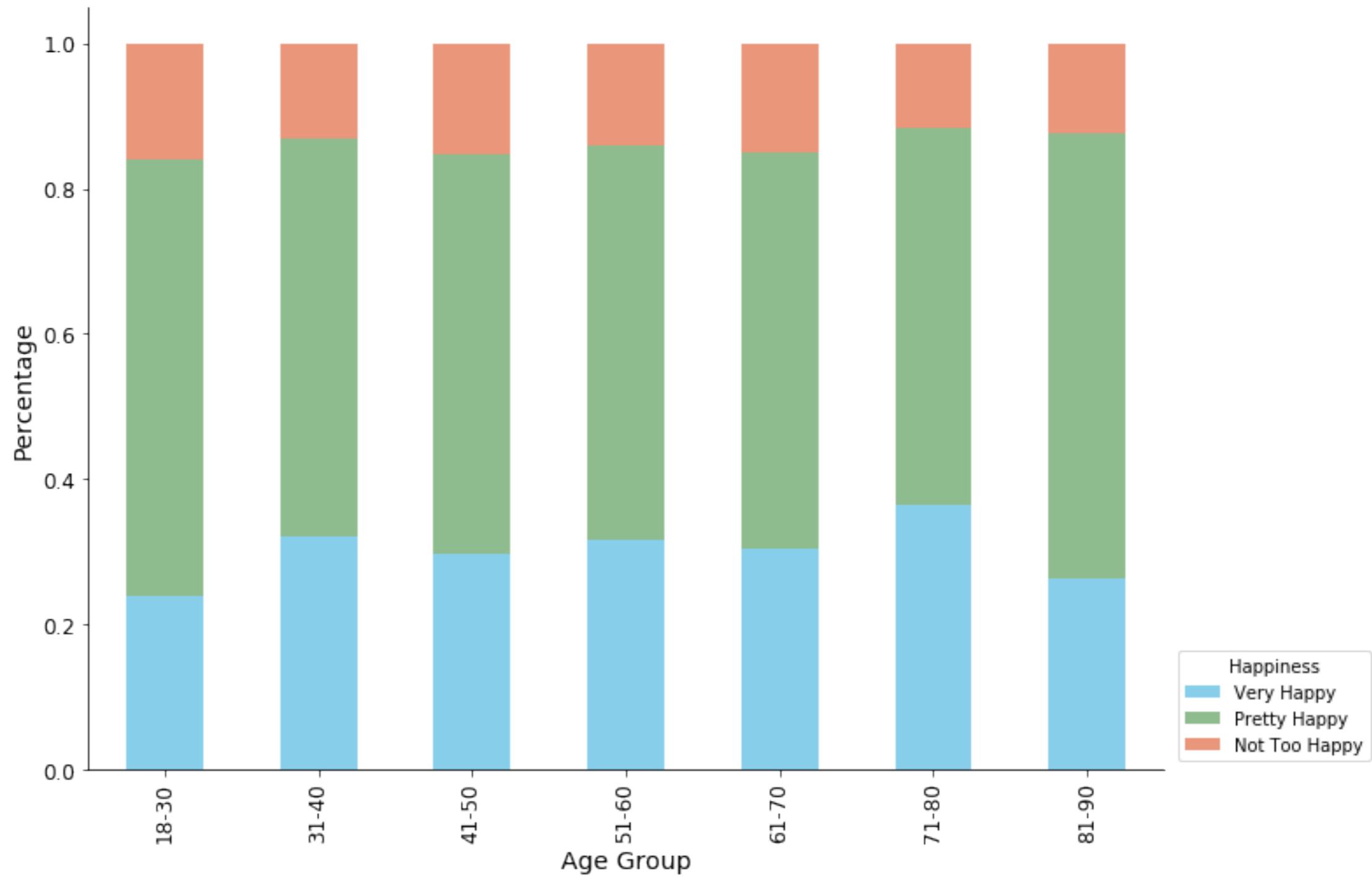
1	Finland
2	Denmark
3	Norway
4	Iceland
5	Netherlands
6	Switzerland
7	Sweden
8	New Zealand
9	Canada
10	Austria
11	Australia
12	Costa Rica
13	Israel
14	Luxembourg
15	United Kingdom
16	Ireland
17	Germany
18	Belgium
19	United States
20	Czech Republic

- ▶ General Social Survey (GSS)
- ▶ Target population : adults (18+) living in households in the United States.
- ▶ 2018 data has 2,348 observations



Very Happy**Pretty Happy****Not Too Happy**

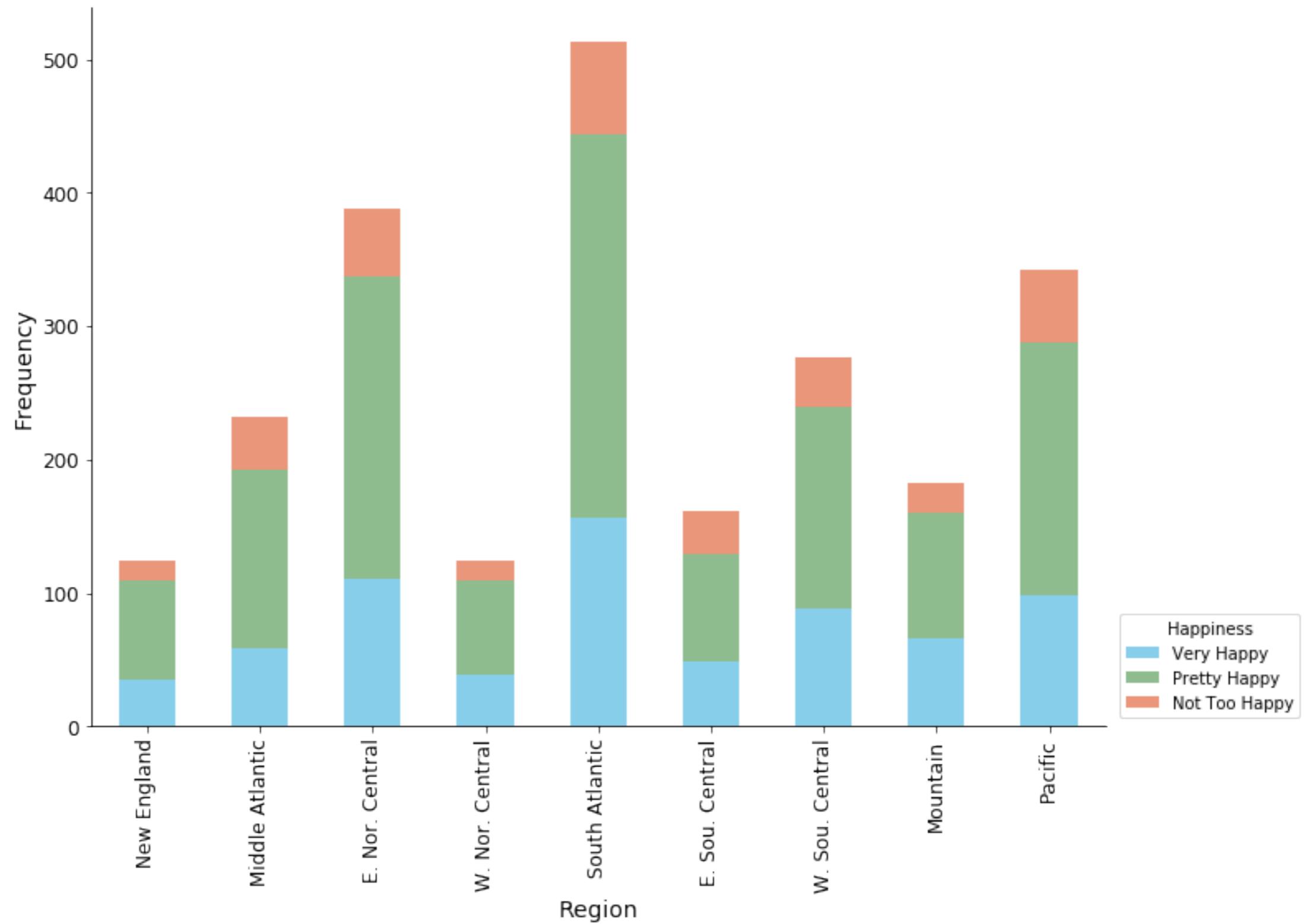






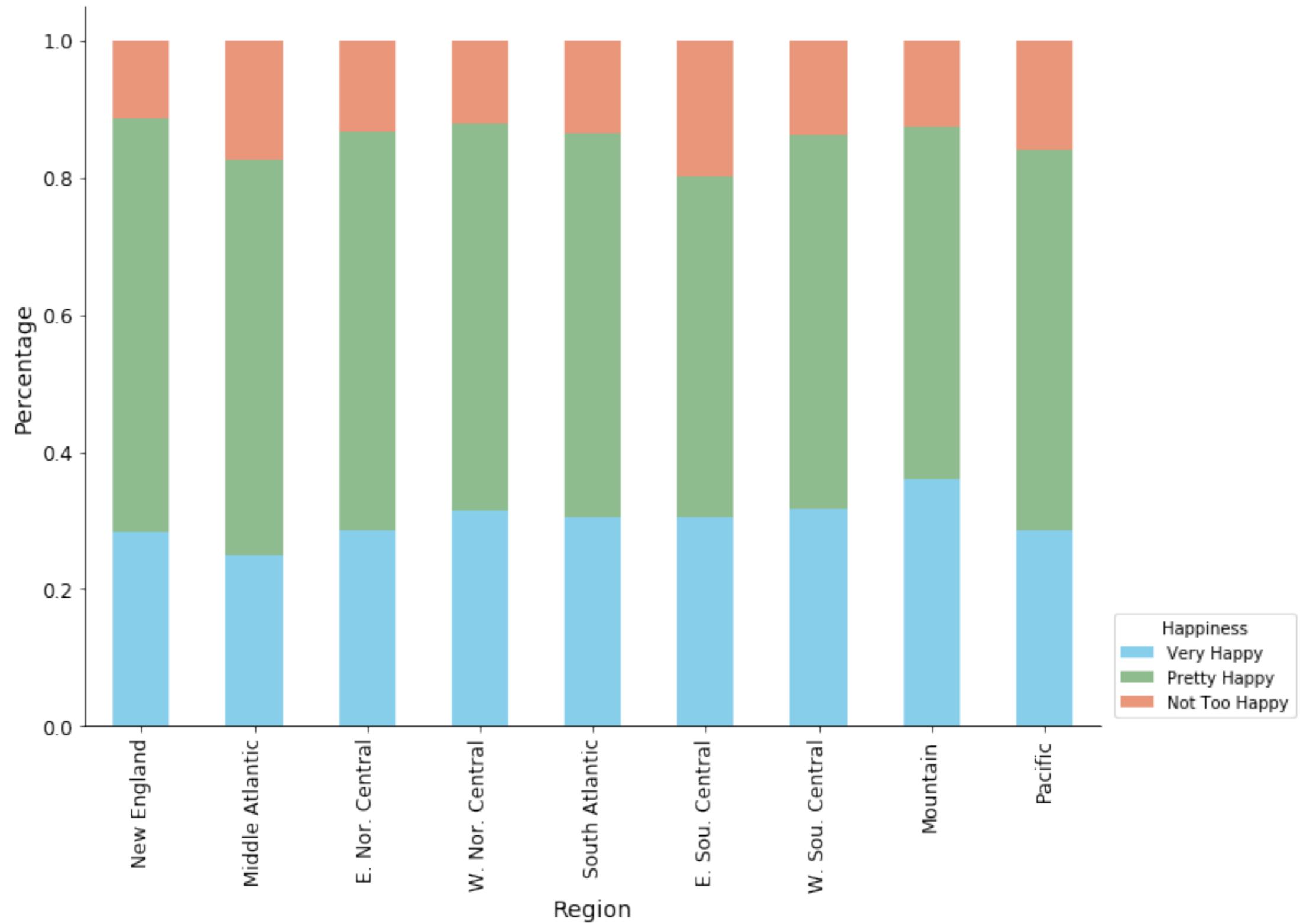
HAPPINESS BY REGION

7



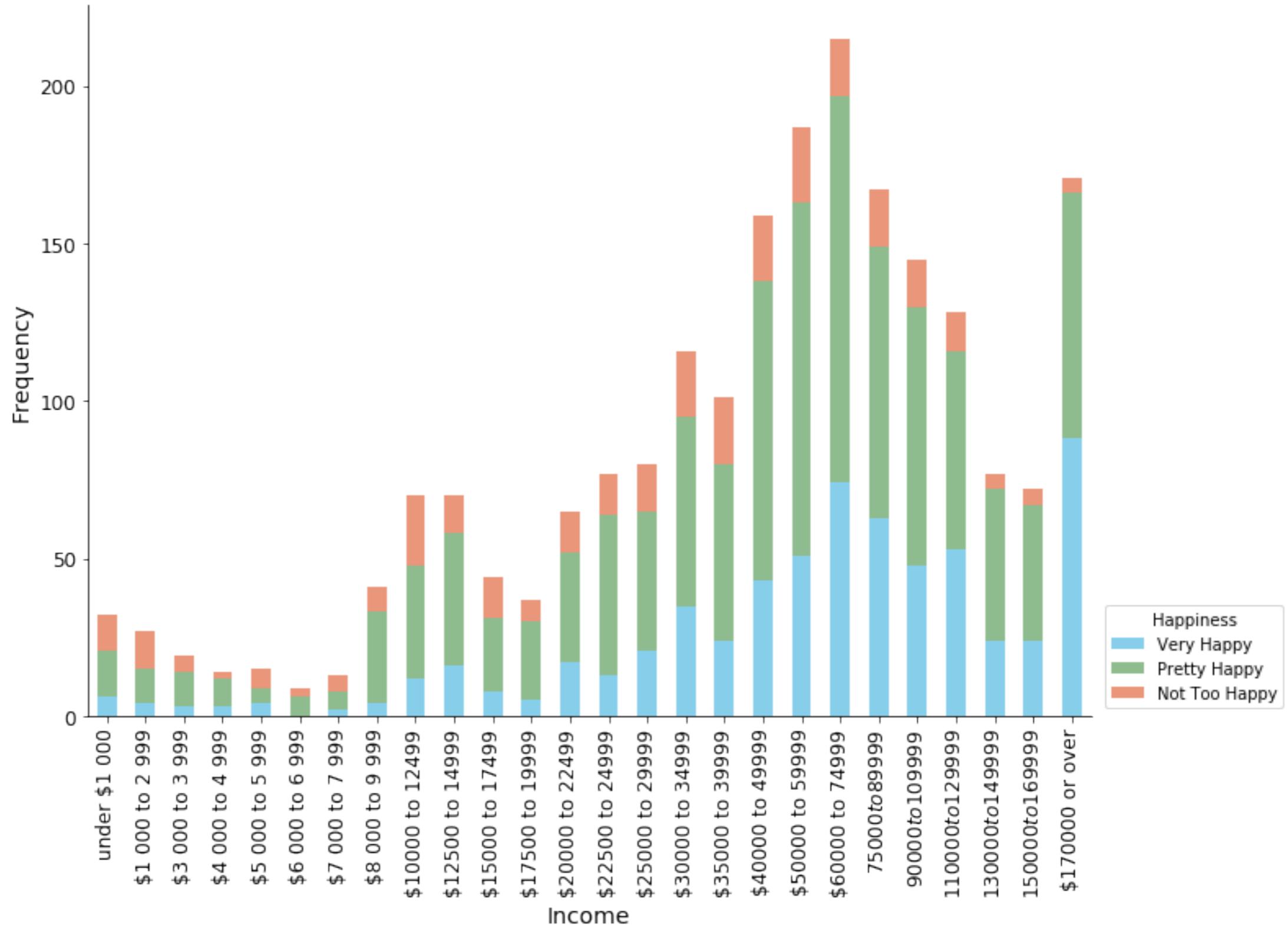
HAPPINESS BY REGION

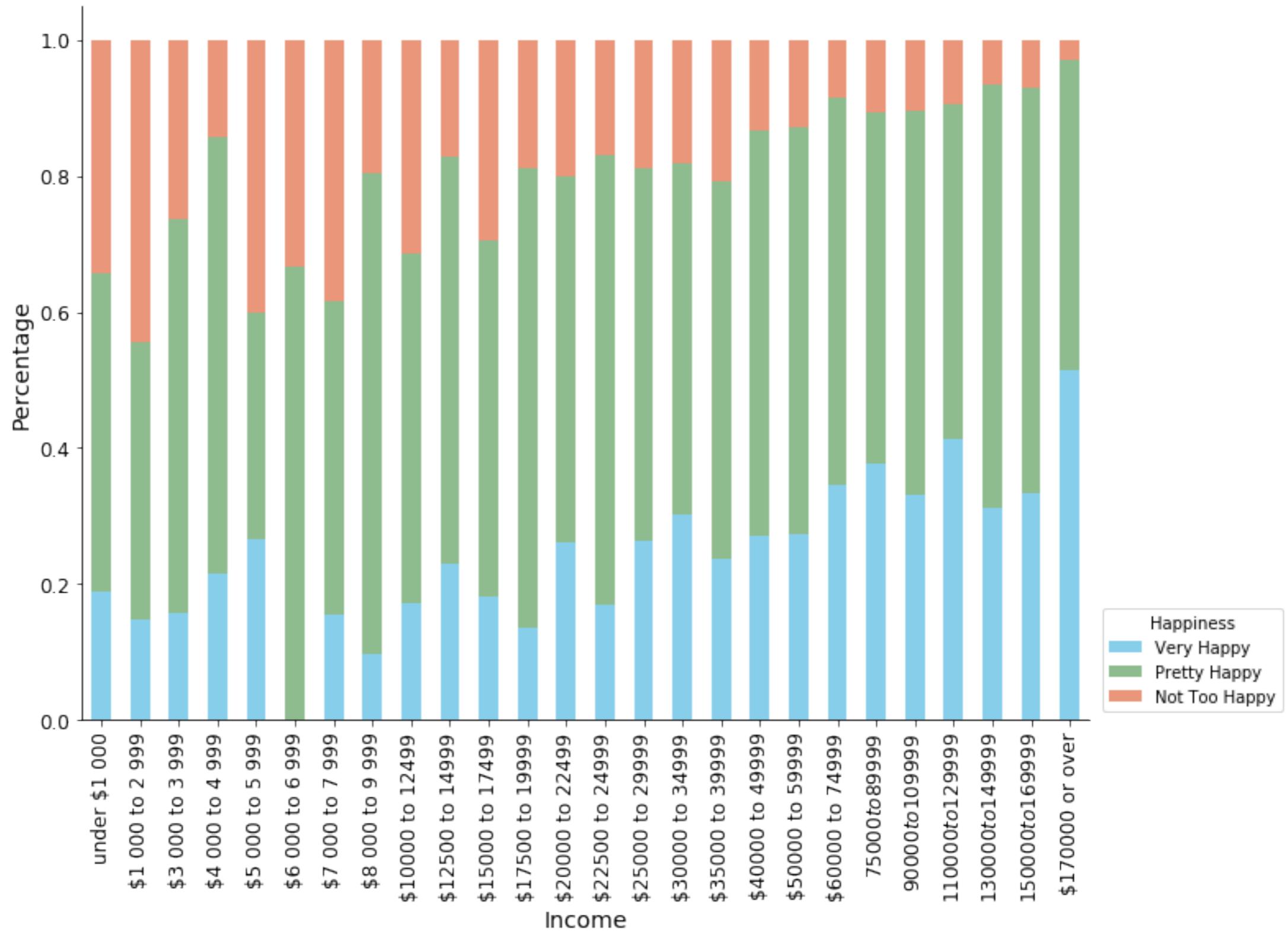
8



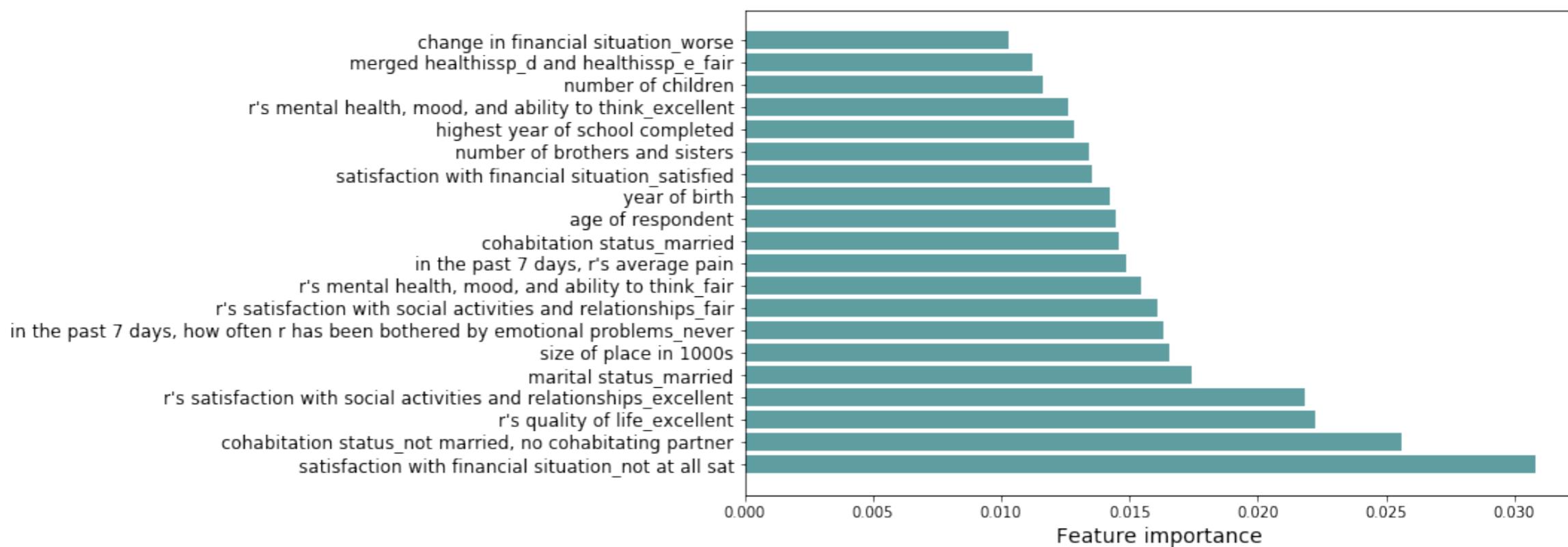
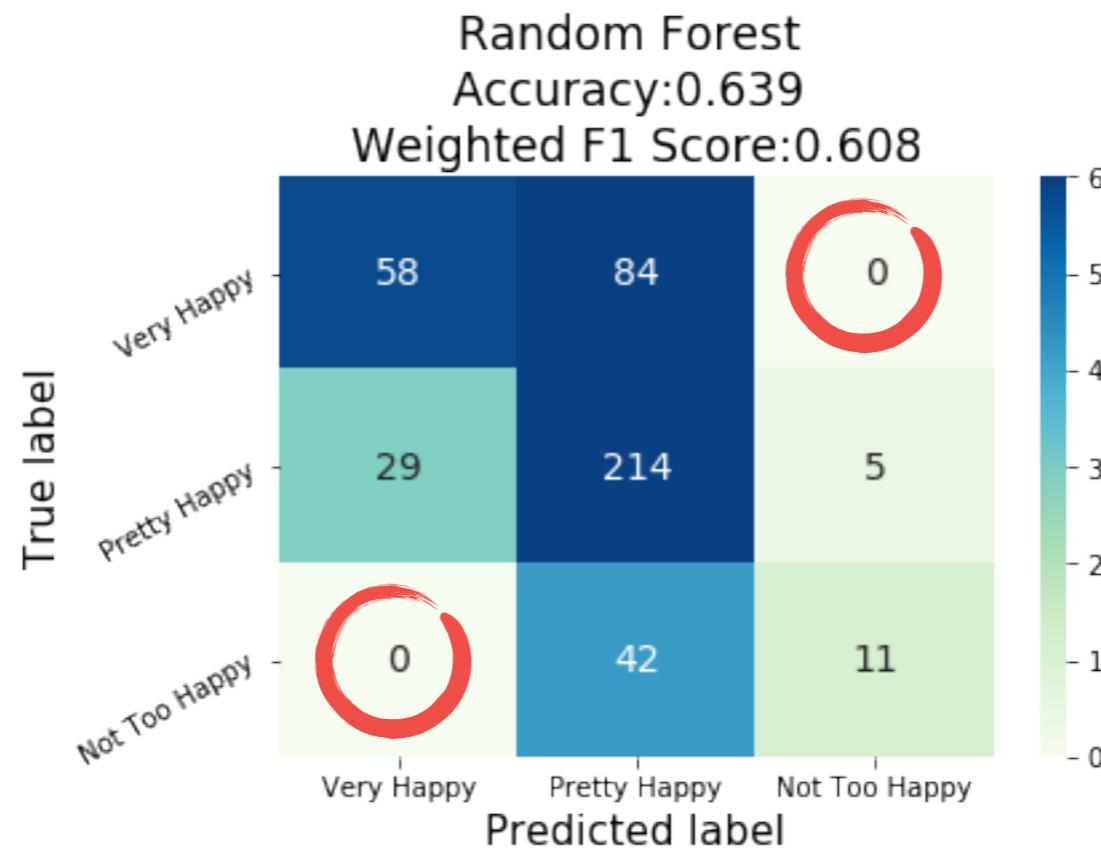
HAPPINESS BY INCOME

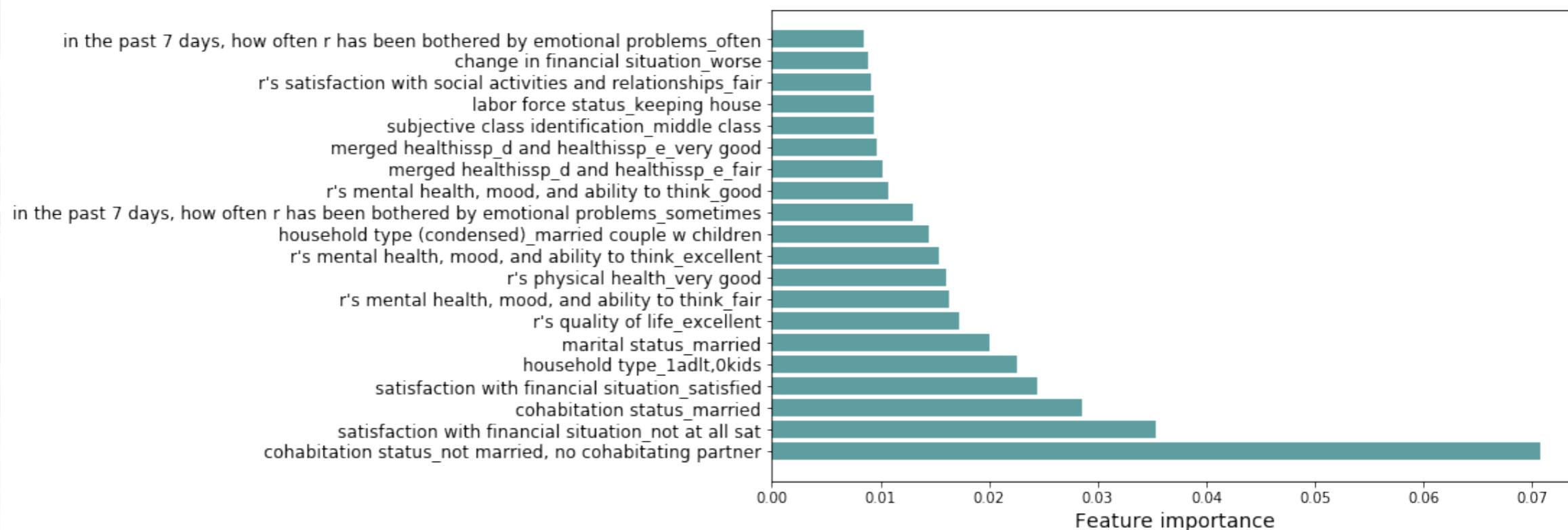
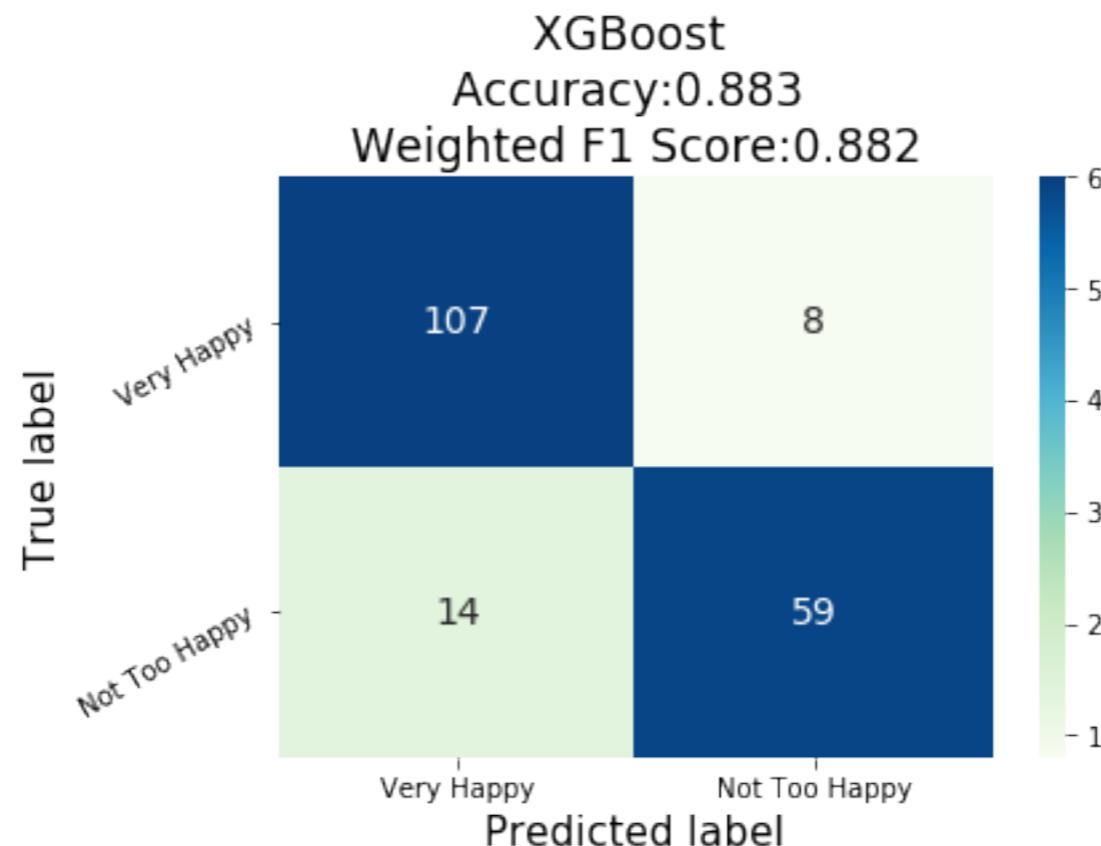
9





- 
- A vertical column of colorful bubbles of various sizes and colors, ranging from clear to iridescent blue and green, set against a dark background.
- ▶ Missing observation threshold : 1% of total
 - ▶ Feature selection
 - ▶ Highest correlation to happiness variable
 - ▶ Additional features
 - ▶ Manual review of features
 - ▶ Vanilla models on (1) all features and (2) minority classes
 - ▶ Oversampling for class imbalance problem
 - ▶ Run vanilla models to choose optimal algorithm
 - ▶ Refine the model





- ▶ Model performance is better when focused on two extreme categories
- ▶ Important Features
 - ▶ Financial satisfaction
 - ▶ Mental and physical health
 - ▶ Cohabitation situation
- ▶ Next steps
 - ▶ Better feature selection
 - ▶ Account for missing values
 - ▶ Collect more data to train the model

THANK YOU FOR YOUR TIME

