Global Happiness

By Christina, Elaine, Fred, and John

Motivation and Summary

- Core Message and Hypothesis
  - There are many factors to take into account when it comes to happiness. On a global scale, how does this measure up between countries?
  - Hypothesis: Countries with a higher happiness rank are more likely to show higher scores in various factors of happiness (GDP per capita, social support, life expectancy, freedom, and generosity), meanwhile creating a larger ecological footprint.
- Questions - why did we ask?
  - Do happier countries make more money?
  - Is there a relationship between happiness factors (i.e., social support, life expectancy, freedom, and generosity) and happiness rank ?
  - Are “happier” countries more eco-friendly?
- Were we able to answer these questions?
  - The data suggested moderate correlations between some of the variables as shown later on.

Questions and Data

-The World Happiness Report and the 2016 Global Ecological Footprint were data sources found on kaggle.com.

-The World Happiness Report provided self-reported happiness scores from their citizens via the Gallup World Poll. Data regarding a country's GDP per capita, their social support, life expectancy, freedom and generosity were assigned an indexed value based on real world data.

-Data from each year from 2015 through 2019 was provided for close to 160 countries.

-The 2016 Global Ecological Footprint provided a dataset for 2016 that allowed us to combine it's more in-depth global footprint data with the Happiness Report to further examine if there were correlations with a country's ecological standing.

-More questions - where does the United States rank with the rest of the world? Regions?

Data Cleanup and Exploration

-Happiness data involved combining and cleaning 5 CSVs

-Countries data involved renaming countries to match Happiness data

-Countries data was merged with with Happiness data

-Obtained top 10 and bottom 10 countries by for each year

-Compared United States to other countries

-Grouped Happiness data by region to look for other trends

```
# Clean the 2017 data and insert the year
happiness_2017_df = pd.read_csv(happiness_2017_file)
happiness_2017_df.insert(0, 'Year', 2017)
happiness_2017_df['Region'] = happiness_2017_df['Country'].map(happiness_2016_df.set_index('Country')['Region'])
region = happiness_2017_df['Region']
happiness_2017_df.drop(labels=['Region'], axis=1,inplace = True)
happiness_2017_df.insert(2, 'Region', region)
del happiness_2017_df['Whisker.high']
del happiness_2017_df['Whisker.low']
del happiness_2017_df['Dystopia.Residual']
happiness_2017_df.rename(columns={'Happiness.Rank': 'Happiness Rank',
                                  'Happiness.Score': 'Happiness Score',
                                  'Economy..GDP.per.Capita.': 'Economy (GDP per Capita)',
                                  'Family': 'Social Support',
                                  'Health..Life.Expectancy.': 'Health (Life Expectancy)',
                                  'Trust..Government.Corruption.': 'Trust (Government Corruption)'}, inplace=True)
trust = happiness_2017_df['Trust (Government Corruption)']
happiness_2017_df.drop(labels=['Trust (Government Corruption)'], axis=1,inplace = True)
happiness_2017_df.insert(9, 'Trust (Government Corruption)', trust)
happiness_2017_df.head()
```

	Year	Country	Region	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Social Support	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity
0	2017	Norway	Western Europe	1	7.537	1.616463	1.533524	0.796667	0.635423	0.315964	0.362012
1	2017	Denmark	Western Europe	2	7.522	1.482383	1.551122	0.792566	0.626007	0.400770	0.355280
2	2017	Iceland	Western Europe	3	7.504	1.480633	1.610574	0.833552	0.627163	0.153527	0.475540
3	2017	Switzerland	Western Europe	4	7.494	1.564980	1.516912	0.858131	0.620071	0.367007	0.290549
4	2017	Finland	Western Europe	5	7.469	1.443572	1.540247	0.809158	0.617951	0.382612	0.245483

```
# Combine all the year dataframes into a single dataframe
happiness_df_list = [happiness_2015_df, happiness_2016_df, happiness_2017_df, happiness_2018_df, happiness_2019_df]
happiness_df = pd.concat(happiness_df_list)
happiness_df

# Export combined dataframe to a csv
happiness_df.to_csv('Resources/happiness_years.csv', encoding='utf-8', index=False)
```

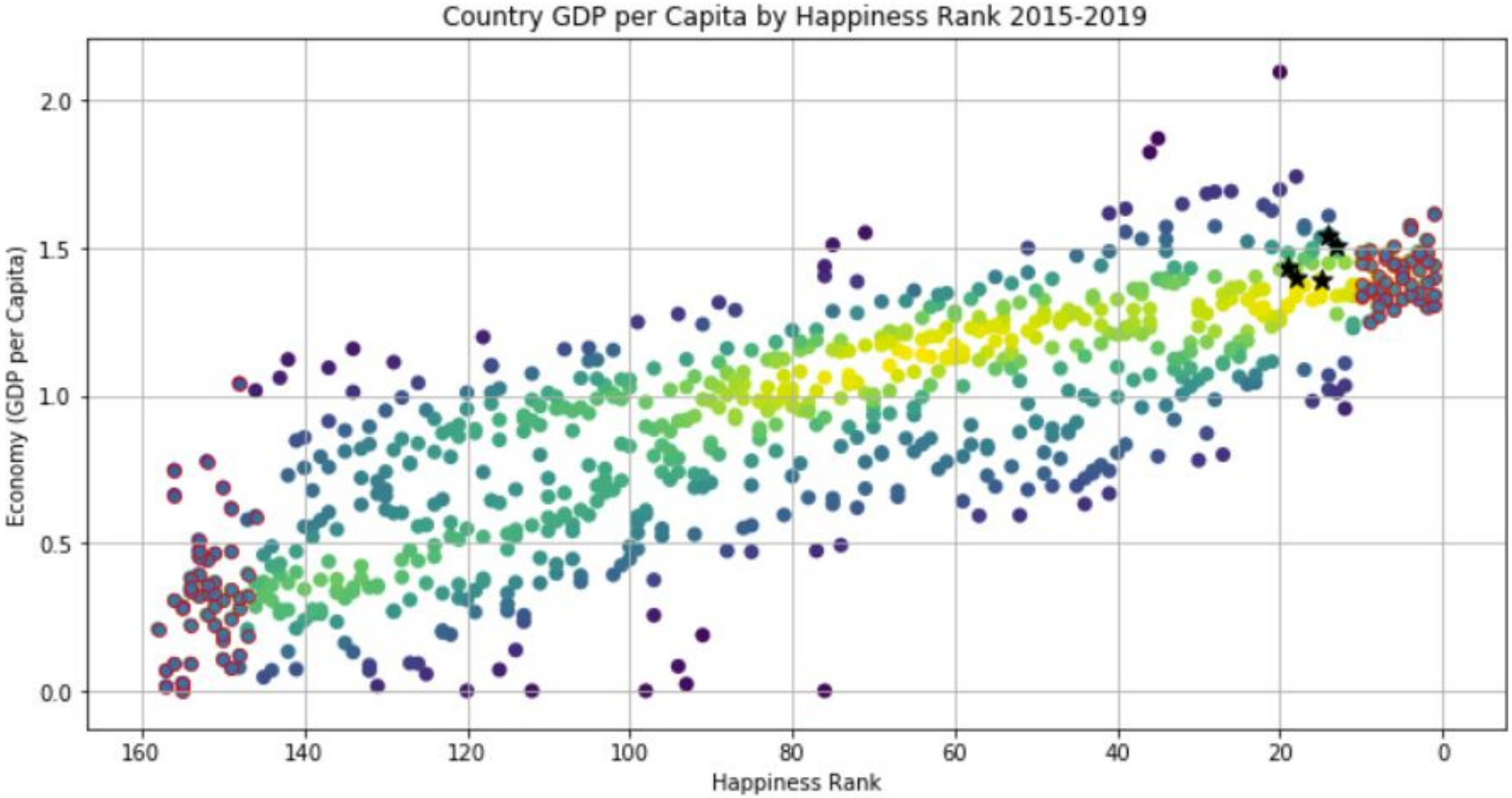
```
# Create dataframes for top 10 ranked countries
top_happiness_2015 = happiness_df.loc[happiness_df['Year'] == 2015, :].head(10)
top_happiness_2015
top_happiness_2016 = happiness_df.loc[happiness_df['Year'] == 2016, :].head(10)
top_happiness_2016
top_happiness_2017 = happiness_df.loc[happiness_df['Year'] == 2017, :].head(10)
top_happiness_2017
top_happiness_2018 = happiness_df.loc[happiness_df['Year'] == 2018, :].head(10)
top_happiness_2018
top_happiness_2019 = happiness_df.loc[happiness_df['Year'] == 2019, :].head(10)
top_happiness_2019

top_10 = [top_happiness_2015, top_happiness_2016, top_happiness_2017, top_happiness_2018, top_happiness_2019]

top_10_df = pd.concat(top_10)
top_10_df

top_10_df.to_csv('Resources/happiness_top_10.csv', encoding='utf-8', index=False)
```

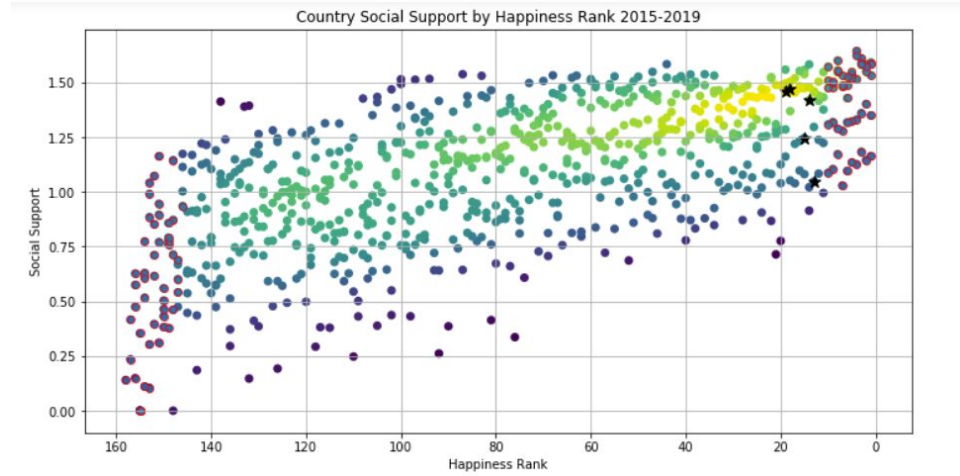
```
#Change names of Countries to match the Country names in the happiness data
replacements = {'Country': {
r'Libyan Arab Jamahiriya':'Libya',
r'Macedonia TFYR':'Macedonia',
r'Russian Federation':'Russia',
r'Syrian Arab Republic':'Syria',
r'Tanzania, United Republic of':'Tanzania',
r'United States of America':'United States',
r'Venezuela, Bolivarian Republic of':'Venezuela'}}

footprint_data1.replace(replacements, regex=True, inplace=True)

footprint_data1.head()
```

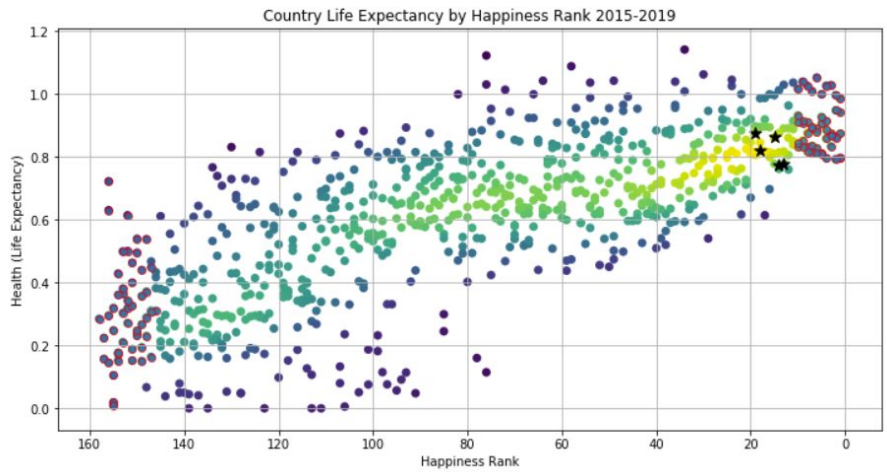
	Country	Region	Population (millions)	HDI	GDP per Capita	Cropland Footprint	Grazing Footprint	Forest Footprint	Carbon Footprint	Fish Footprint	...	Cropland	Grazing Land	Forest Land	Fishing Water	Urba Lan
0	Afghanistan	Middle East/Central Asia	29.82	0.46	$614.66	0.30	0.20	0.08	0.18	0.00	...	0.24	0.20	0.02	0.00	0.0
1	Albania	Northern/Eastern Europe	3.16	0.73	$4,534.37	0.78	0.22	0.25	0.87	0.02	...	0.55	0.21	0.29	0.07	0.0
2	Algeria	Africa	38.48	0.73	$5,430.57	0.60	0.16	0.17	1.14	0.01	...	0.24	0.27	0.03	0.01	0.0
3	Angola	Africa	20.82	0.52	$4,665.91	0.33	0.15	0.12	0.20	0.09	...	0.20	1.42	0.64	0.26	0.0
4	Antigua and Barbuda	Latin America	0.09	0.78	$13,205.10	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	Na

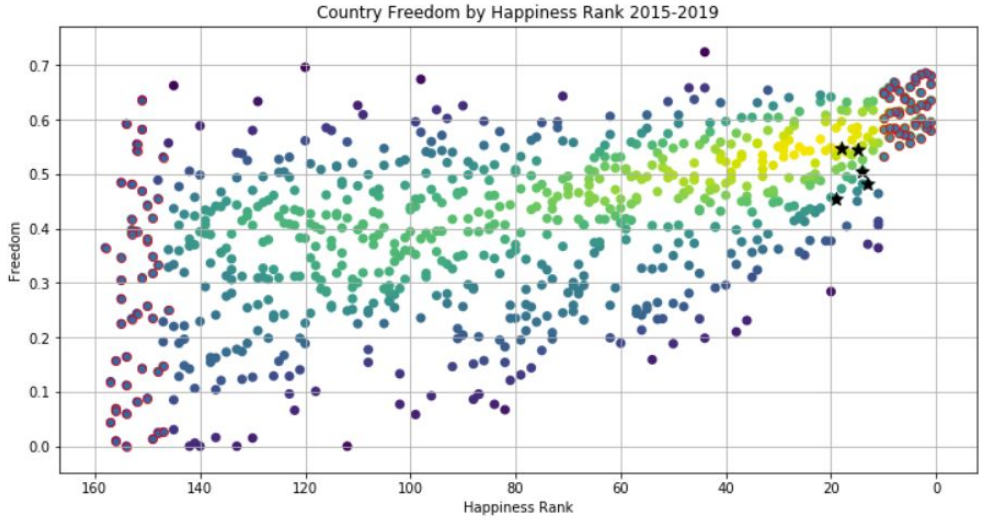
Data Analysis

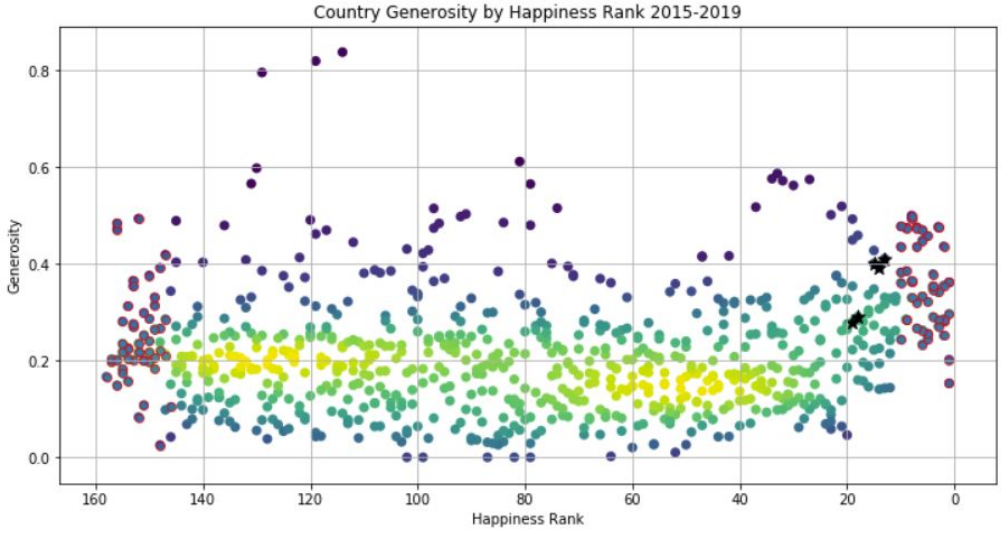
- Steps taken to analyze the data and answer each question you asked in your proposal
  - With the cleaned data, we ran correlation tests
  - We made charts and visuals for the data that showed stronger correlations, or we found surprising
  - United States data points are marked as stars (★) on the charts.

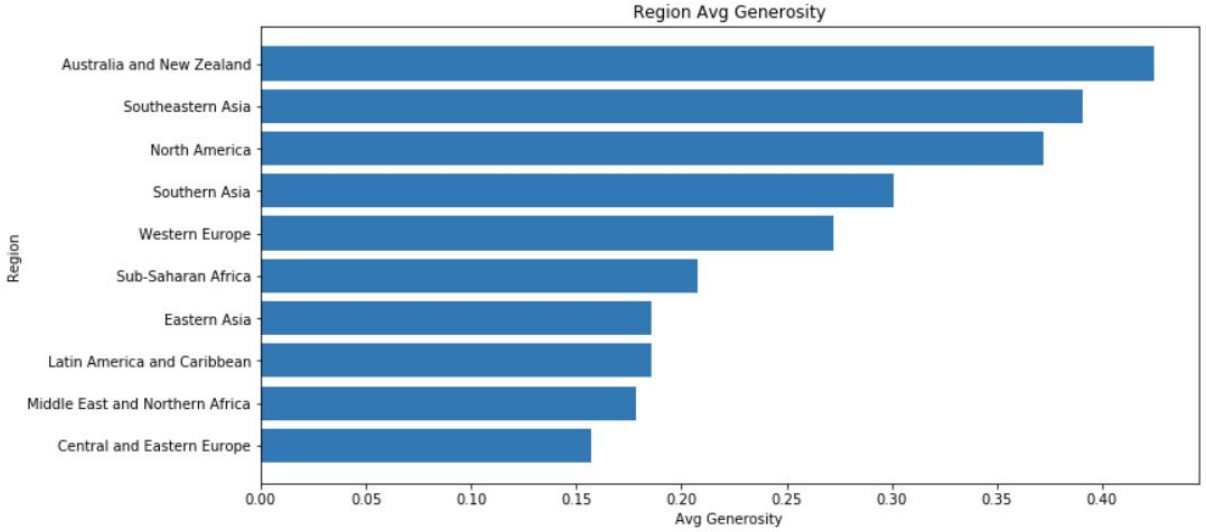

Country GDP per Capita by Happiness Rank 2015-2019
Economy (GDP per Capita)
2.0
1.5
1.0
0.5
0.0
160
140
120
100
80
60
40
20
0
Happiness Rank

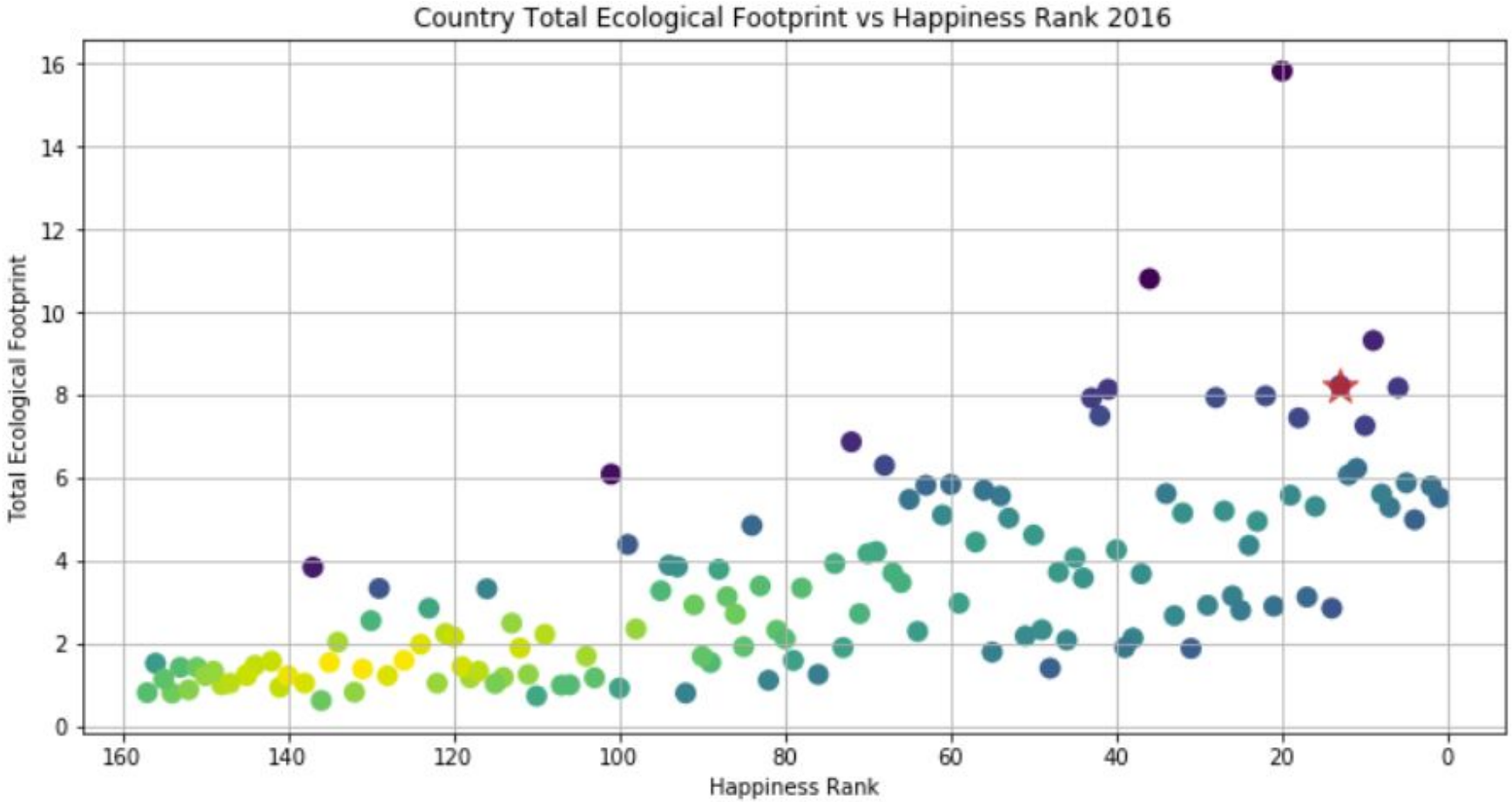

Country Social Support by Happiness Rank 2015-2019
Social Support
Happiness Rank

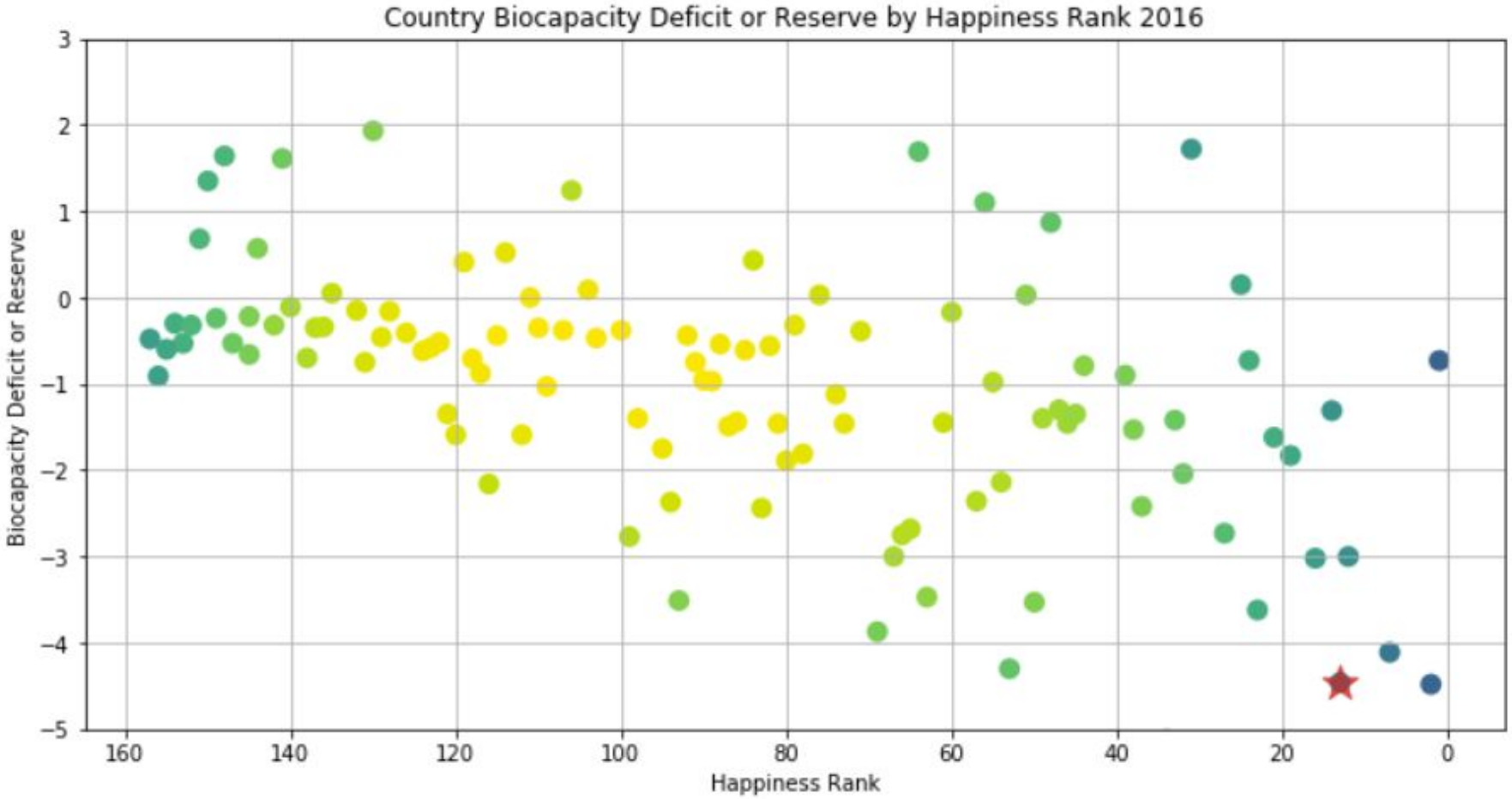

Country Life Expectancy by Happiness Rank 2015-2019
Health (Life Expectancy)
Happiness Rank


Country Freedom by Happiness Rank 2015-2019
Freedom
Happiness Rank


Country Generosity by Happiness Rank 2015-2019
Generosity
0.0
0.2
0.4
0.6
0.8
160
140
120
100
80
60
40
20
0
Happiness Rank


Region Avg Generosity
Region
Australia and New Zealand
Southeastern Asia
North America
Southern Asia
Western Europe
Sub-Saharan Africa
Eastern Asia
Latin America and Caribbean
Middle East and Northern Africa
Central and Eastern Europe
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
Avg Generosity


Country Total Ecological Footprint vs Happiness Rank 2016
Total Ecological Footprint
Happiness Rank
0
2
4
6
8
10
12
14
16
160
140
120
100
80
60
40
20
0


Country Biocapacity Deficit or Reserve by Happiness Rank 2016
Biocapacity Deficit or Reserve
3
2
1
0
−1
−2
−3
−4
−5
160
140
120
100
80
60
40
20
0
Happiness Rank

Discussion

- Countries with a higher happiness ranking show a strong positive correlation to having a higher GDP per capita
- The data suggests a moderate, positive correlation between happiness factors (i.e., social support, life expectancy, freedom, and generosity) and happiness rank.
- There is a moderate, positive correlation between happiness rank and total ecological footprint.
    - We were surprised to find that there was not a strong correlation between happiness rank and biocapacity deficit/reserve.

Post Mortem

- Difficulties:
  - The data cleanup process took longer than expected
  - Focusing our scope on specific questions based on the data we had
  - Finding the methods and definitions for the data

- Additional Questions:
  - Are there correlations between the different factors we analyzed?
  - Are the country rankings similar between the different factors?
  - Are there different correlations when looking at the outliers?
  - How do the factors change for countries over time?

Questions

-open-floor Q & A with Audience!!!!

-And GO GIT!!!! WE LOVE IT!!!!