**Slide 1 – Beth**

Happy Saturday to you all!

After tossing around lots of different ideas, our team selected to present a topic that had a bit more positivity (we hope) – World Happiness.

**Slide 2 – Beth**

Background

Upon researching more about the topic we selected, the World Happiness Report was first published in 2012. Information on our files were compiled from data gathered from the Gallup World Poll, where approximately 1,000 people from each country (average of 156 countries) were polled using the Cantril ladder. The Cantril ladder is where 0 is the lowest rank and 10 being the highest. As you can see by the graphic to the right, this will help explain the overall ranking for happiness.

**Slide 3 - Beth**

The Data

There are 6 components of data used to determine overall happiness.

* GDP – Gross Domestic Product, economic snapshot of the country
* Social Support – This includes all types of personal relationships (family, friends, even mental health support)
* Healthy life expectancy – What is the overall outlook on diet, exercise and availability to doctors?
* Freedom to make life choices – such as religion, political environment, and personal choice (marriage, family planning, etc.)
* Generosity – what citizens do to support others in their community (charity, volunteerism)
* Perceptions of Corruption – what is the misuse of public power for private benefit

This slide shows how each of these components are weighted in determining the overall happiness scores based upon the data pulled from 2015 through 2019. It was not surprising to see that GDP and social support were key components to measuring happiness.

**Slide 4 - Zack**

Initial Analysis

We opted to pull data from 2015 through 2019. There were two reasons why data from 2020 was not used: 1) the data may be too new and there could be a high potential for bad data, and 2) don’t we all just want to forget 2020 anyway.

Once we completed pulling and merging our datasets, we conducted some activities to assure that our data was clean.

* We checked for null (NaN) values, where we found none.
* We checked the data types to assure these were as expected.
* Identified and renamed instances where country name appeared to be the same, such as Trinidad and Tabago vs Trinidad & Tabago. 5 re-namings were complete.
* Removed instances where a country did not appear on all files.

**Slide 5, 6, 7 - Darrius**

What did we learn?

Speak to the different scatter plots.

* GDP vs Healthy life expectancy – this was a different plot, where we wanted to see if there was any correlation between GDP and life expectancy, and as you see there is a positive correlation.
* ­Slide 6 – we completed some additional scatter plots and measured the correlation coefficient to check each of the components measuring overall happiness. For each of the four plots shown here, you will see a positive correlation to overall happiness.
  + GDP to Happiness
  + Social support to happiness
  + Health to happiness
  + Freedom to happiness
  + Slide 7 – for the plots for generosity and corruption, we saw some interesting results.
* Generosity to happiness
* Corruption to happiness

**Slide 8 to 12 - Matt**

We then took each year’s top and bottom 10 ranked countries as stacked bar charts to see how they compared.

Talk about your thoughts / observations

**Slide 13 – Jared**

Happiness over time – Top 10

We then wanted to see the top 10 happiest countries over time to see how they scored/ranked over time.

Talk about your thoughts / observations

**Slide 14 - Jared**

Overall Correlation map

Finally, we wanted to put it all together, where this map shows how each component of happiness correlates with one another. It may take a couple minutes to really see how this tells the story.

Talk about your thoughts / observations

**Slide 15 - Darrius**

Conclusion

After cleaning the data, we identified that a lot of the top ranked countries were

European. We were able to generate scatter plots to identify the correlation between

subcategories and the overall ranking/score of each country. Using the scatter plots, we were able

to identify the positive correlations between GDP, life expectancy, and freedom as it pertains to

the overall happiness score. GDP has a positive correlation with life expectancy. Increases in

GDP coincided with higher life expectancy. Social support proved to be one of the biggest

contributors to overall happiness as seen in the pie chart identifying social support as 31.9% of

the overall score. The data sets reflect that the world is getting happier over time. Although we

were unable to pull rankings for 2020, we were able to identify a trend of increasing happiness

over a five year time span.

If there is time…. This project really had us flexing the muscle memory of what we learned through this course to date. We realize with any dataset, there are going to be some limitations and we all need to find ways to solve for them.

What additional information would be interesting…. How about climate? How or does this impact happiness?

Are there any questions?