Work in Progress

Ahalya K, Ian W, Ankita B

About the Data

- We used data from the UN and Gapminder
  - Data pulled from 2022 and 2023
- We are currently comparing HDI with happiness score for 2023
- We will compare HDI indicators with happiness later
  - Ex. GNI per capita, years of schooling, life expectancy

Data Wrangling - HDI Data

```
library(readxl)
library(tidyverse)

HDI_raw <- read_excel("C:/Users/ahaly/OneDrive - The Pennsylvania State University/Desktop/HDR25_Statistical_Annex_HDI_Table.xlsx")

HDI_clean <- HDI_raw %>%
  select(-c(1, 4:15)) %>%
  slice(-c(1:7, 82, 133, 177, 204:278)) %>%
  rename("Country" = `Table 1. Human Development Index and its components`,
         "HDI" = `...3`)
```

Data Wrangling - Happiness Data

```
library(tidyverse) #loading in tidyverse
happiRaw <- read.csv("hapiscore_whr.csv") # read in csv file
str(happiRaw) # checking structure of the data
View(happiRaw) # Viewing the data

# cleaning the data
happiClean <- happiRaw %>%
  select(-c(geo, X2005:X2022)) %>% # removing unneeded columns
  rename(
    Happiness_Score_2023 = X2023 #rename to a more appropriate name
  )

View(happiClean) # view the clean data
```

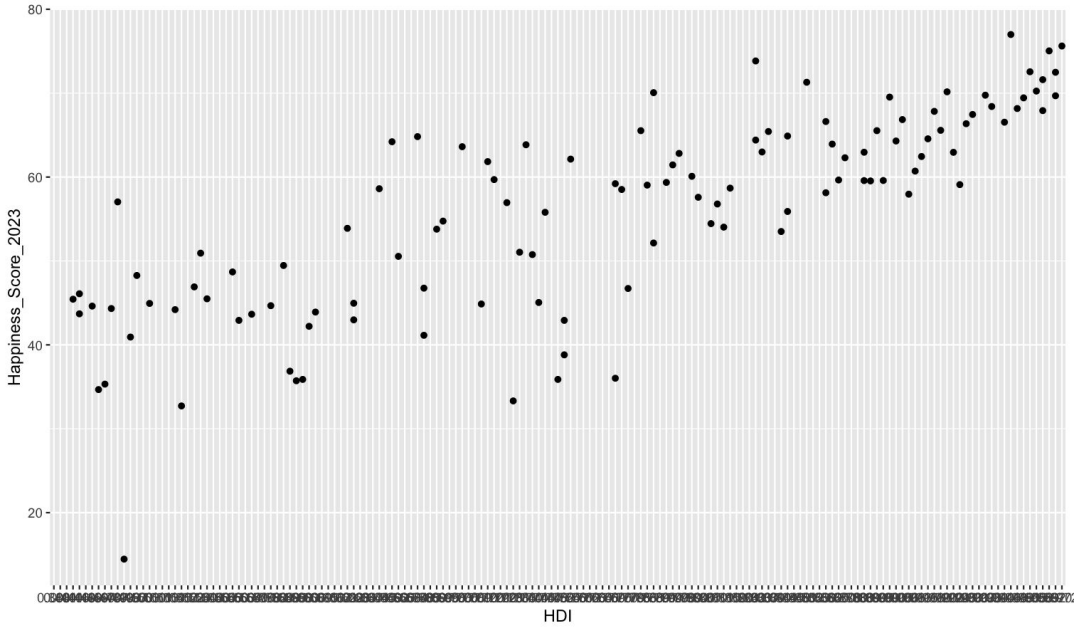
Data visualization (currently)

- scatterplot

What to add - looking to work on in the future:

- Adding a best fit line
- Fixing the x-axis scale
- Adding Title and labels
- Making the color palette color blind friendly
- Creating visualizations for 2022 data, as well as for the other attributes we will be researching

Currently we can see, there is a moderate positive correlation between a country's HDI index and its' Happiness score.



Challenges

- Finding other data sets to see a relationship with HDI
- Figuring out what type of visualizations are best to make (or maybe adding a table as well)

Any questions?