3. DataVisualization:

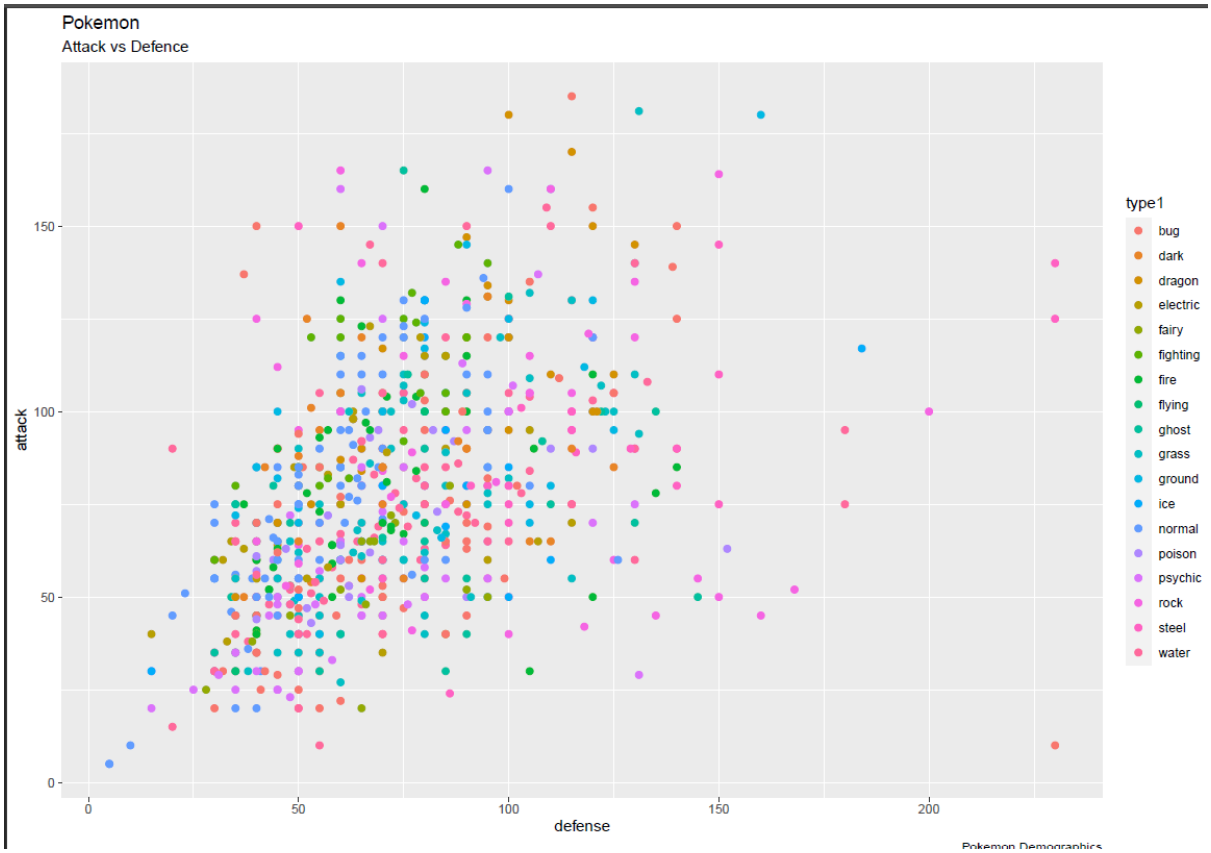
As I approach the end of a very long and intense couple of weeks, the goal very nearly in sight, I have already begun to critique my approach to the tasks. Maybe this is premature since I haven't finished yet, but I can already admit that I spent (and wasted) way too much time trying to find a dataset to work with, and even more time trying to decide what kind of graph would best portray the data. Hours were spent trying things out and then binning the results.

It's not clear but it seems you may have wanted to see all those abortive attempts to produce something of value. I regret my petulance in trashing the vast majority of what I did, if for no other reason than to illustrate the amount of time I've put into these tasks.

Not a time to stress the material I didn't keep and focus on what I did.

Practice:

I made a scatterplot of the attack vs defence capabilities of all (currently) discovered Pokemon for my kids. There were too many to name, so it's divided by colour into the various classes. They were mildly amused though not enlightened so I gave up on the dataset there and then.



(full-sized pdf versions can be found at the end of the document)

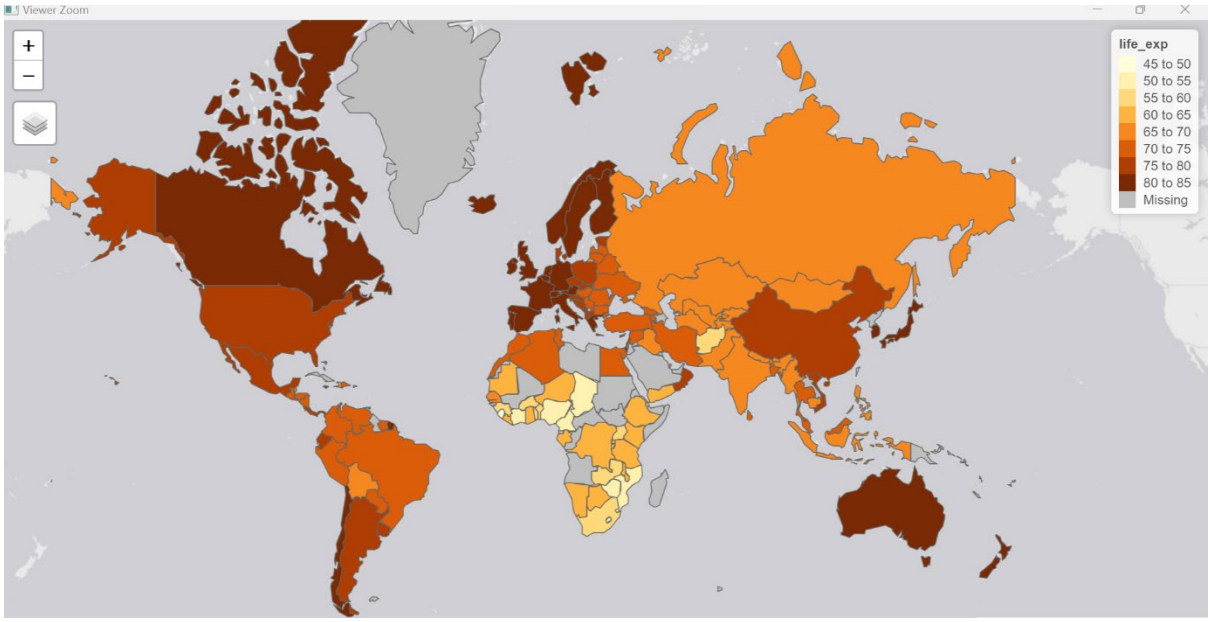
World Maps:

I eventually settled on a 'World' database which contained coordinates so I could plot on a world map. I wasted far too much time playing with all the various functions available, and I hope to develop this interest more when I have free time (somewhere around 2025 by my reckoning).

```r
library(tmap)
data("World")

tmap_mode("view")

  tm_shape(World)+
  tm_polygons("life_exp")
```



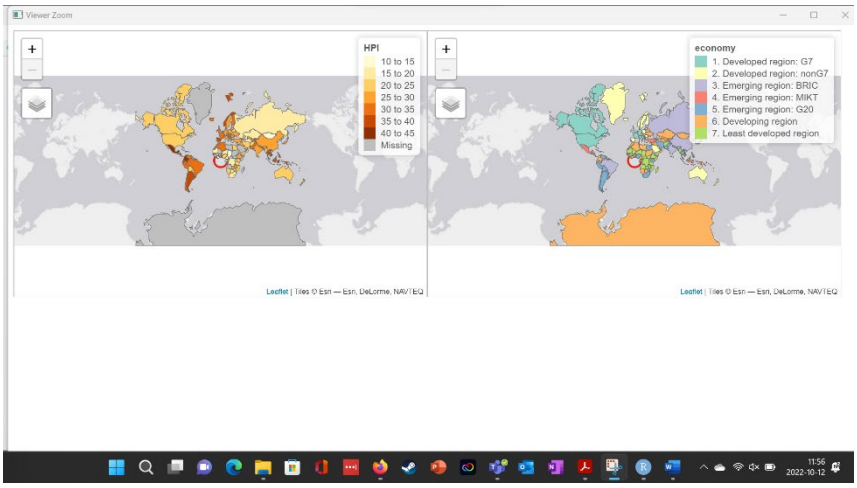
With the inbuilt view mode and 'polygon' function one can instantly create informative and interactive maps. I created a map showing life expectancy around the world which made for some sober viewing.

There were many fantastic functions, and the ones that stood out were the interactive ones, mainly as they were so easy to produce and are more pedagogic and memorable by their nature.

```r
library(tmap)
data("World")

tmap_mode("view")

tm_shape(World) +
  tm_polygons(c("HPI", "economy")) +
  tm_facets(sync = TRUE, ncol = 2)
```



By using the 'tm_facets' function one can create maps that sit side by side and by 'syncing' them it's possible to access more information on each scan.

Scatterplot:

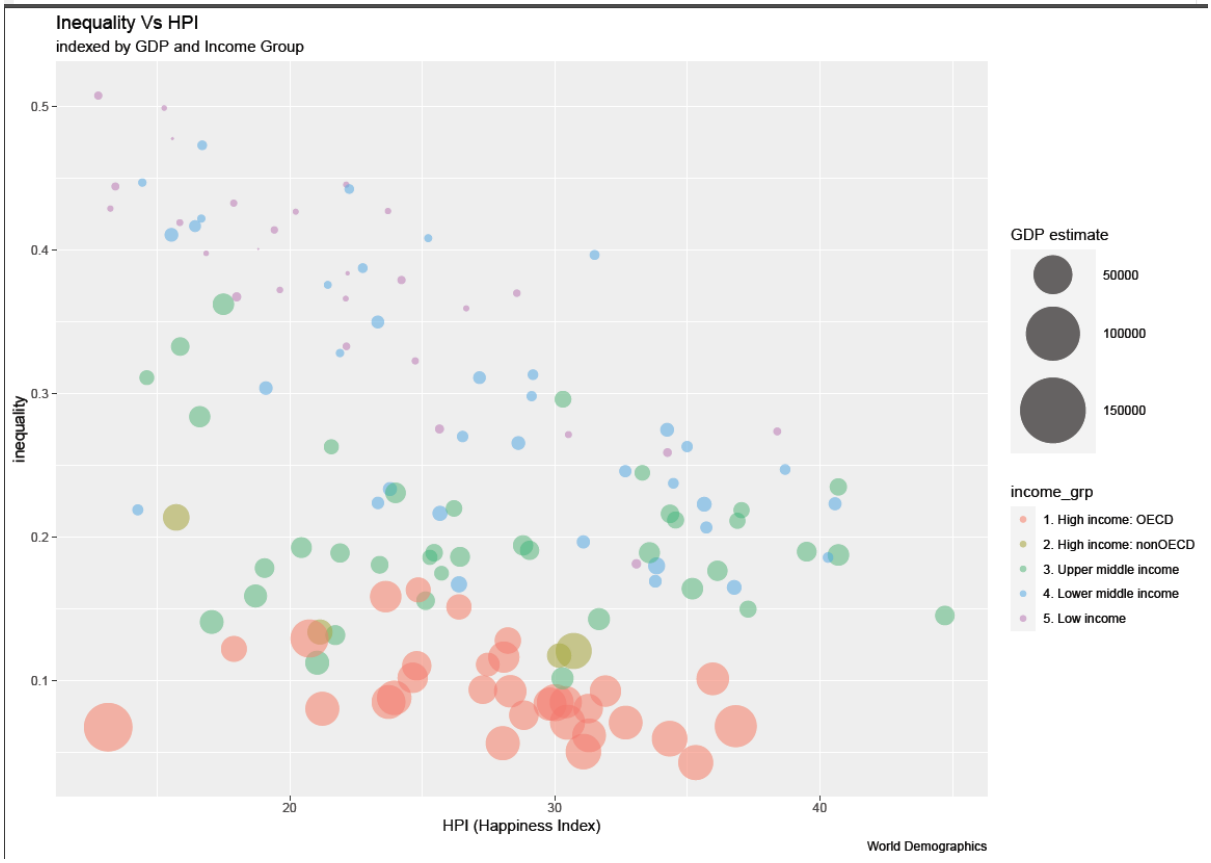
I wanted to create a scatterplot next in order to compare as many different factors as possible. In the limited time that I've been studying graphing with R, this has been the most powerful visual use of data that I've come across.

I chose to compare levels of inequality in countries to the HPI (happiness index) to see if racial or sexual discrimination was clearly reflected in the 'happiness' of each country's citizens. I added a legend for the income groups represented and the GDP of the countries as well.

```
gg2 <- ggplot(World, aes(HPI, inequality, col= income_grp, size= gdp_cap_est))+

  geom_point(alpha=0.5)+
  scale_size(range = c(.1, 24), name="GDP estimate")+

labs(title="Inequality Vs HPI", subtitle="indexed by GDP and Income Group",
     x="HPI (Happiness Index)",
    caption="World Demographics")
```



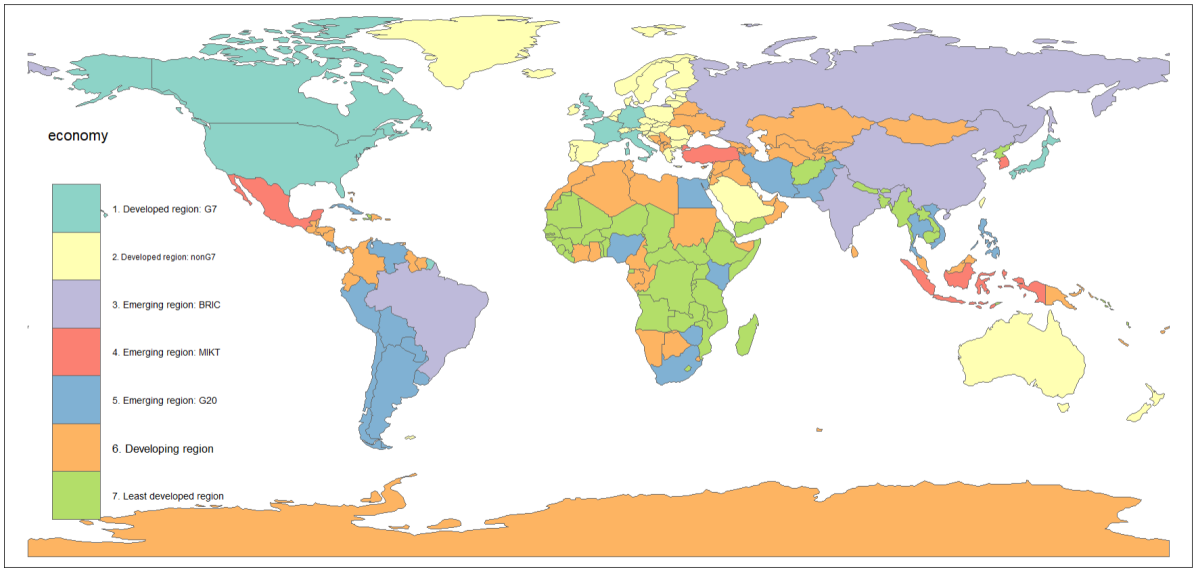
I really loved this graph. It clearly shows how without a stable economy and a relatively high standard of living it is hard for any country to achieve viable and visible equality (disregarding for now any cultural factors).

It is equally interested to note how spread out the HP index was among high earning countries. It raises many questions (for instance who is that country in the bottom left with lots of money, little inequality and with a miniscule amount of happiness to show for it. Tragic!

And the number of middle to high income countries that report very high indexes lends more credence to the cliché that 'money can't buy you everything'.

Boxplot:

Lastly, I created a boxplot to show the happiness of the different economic groups in the world. The world bank decides which group a country belongs to. Here is a map:




```
tmap_mode("plot")

tm_shape(World)+
  tm_polygons("economy")
```

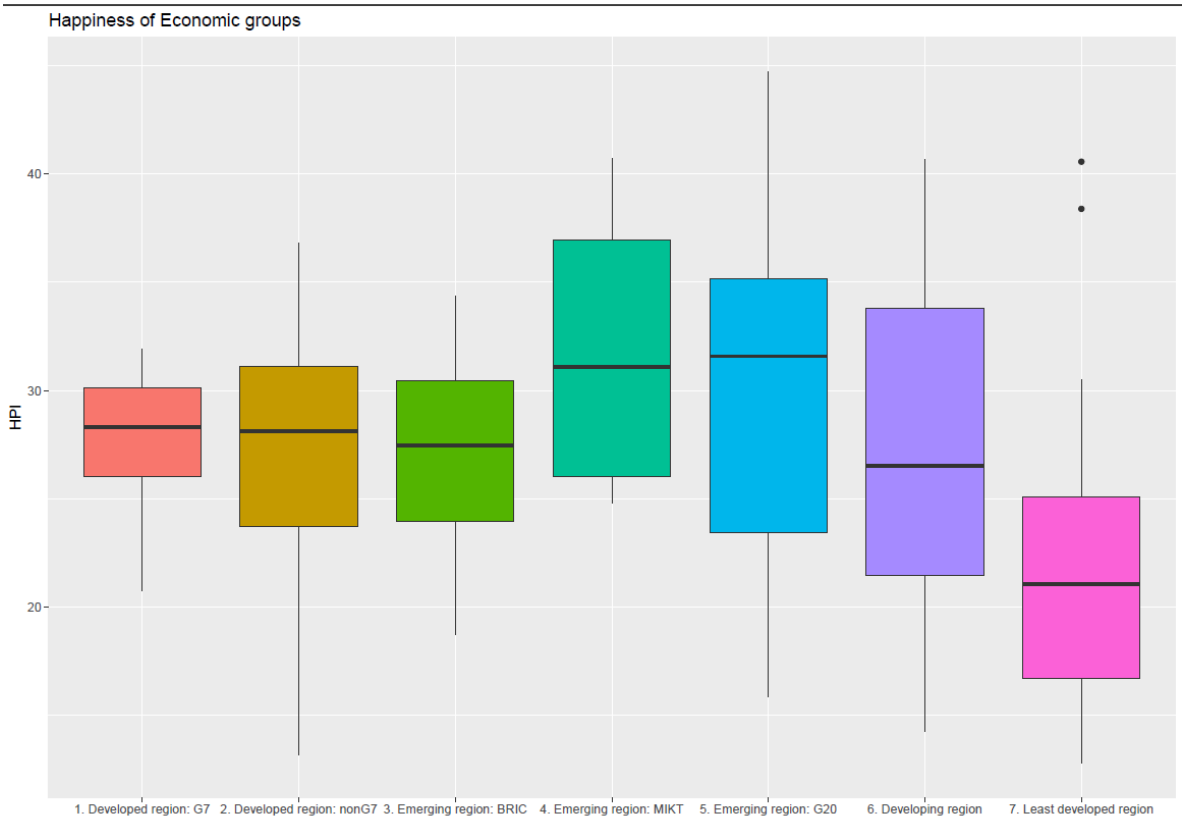
Again I used the tmap package to plot the map.

I created a boxplot with a focus on the economic groups, I wanted it to be as simple as possible.

```
## create plot using economy as the x and happiness as the y, the fill
gg3 <- ggplot(World, aes(x=economy, y=HPI, fill=economy)) +

## call the boxplot through the geom function
  geom_boxplot()+
  labs(title="Happiness of Economic groups")+
xlab("Economic Group") +
  ## there is no legend so that the whole plot could fit on the page
  theme(legend.position="none") +
  xlab("")
```

Here is how the boxplot of happiness in these regions came out:



Not surprising that those with little money have little to be happy about, though it is surprising to see how middling and uninspiring the happiness measurement of the G7 countries was. There is something optimistic and heartening about seeing how high happiness is among emerging countries. If there is anything to conclude from this chart it is that how we measure happiness is fraught with difficulty, contradiction, and cultural relevance.