# Pitch-Text Praxisseminar Data Analytics for Business and Psychology

**Methods and Analysis**

To analyse our questions on the correlation between stringency measures on happiness ratings per year across countries worldwide, we needed to gather our data from different resources. We accessed the happiness data from the open-source platforms Kaggle, which was made available by a research group that collected this data for the World-Happiness reports from 2019-2022. Happiness ratings are the composite score of several indicators of happiness, such as freedom-to-make choices, health expectancy and others. The data on Covid was drawn from OurWorldInData.org, of which we used unique cases/deaths per million to account for population differences. The stringency data was provided by the Oxford University project called OxCGRT. They calculated a Government Stringency Index as the composite measure of nine of the response metrics, including among others school and workplace closures, restrictions on internal movement and stay-at-home requirements.

Having gathered the relevant data, we created a total data frame, combining the three resources by country name and year. As happiness was presented as a single measure for one year, we computed one single variable per year for the other variables at question. These were total death rates per million and the average stringency measures for each country per year. We obtained data for 115 countries.

Our primary outcome were happiness ratings for each country and our primary predictor was the stringency measure. To guide our understanding of the data, we conducted a regression analysis to see whether there were was a relationship between these variables. To address the questions about the influence of death rates and regional differences on the correlation between happiness and stringency, we then looked at the impact of yearly total covid deaths and regional differences on this relationship. We analysed this relationship for each year between 2020 and 2022, thus creating a separate graph for each year.

To conclude our analysis, our final visual output consisted of three sets of graphs, one for each year, for the relationship between happiness and stringency, colouring for regional differences and using size differences for each point in the graph to show differences Covid-Death cases. Our final short presentation graph will show all variables combined, faceted by year and regions.