Description of the data set for the manuscript "**A machine learning approach to quantifying the specificity of color-emotion associations and their cultural differences"**

Each row represents the ratings of the intensity of 20 emotions associated with the color term given by variable "color", by participant identified by variable "subject"

**Variables:**

*lang*: language in which the survey was presented = mother tongue = country of origin. CN = China, DE = Germany, GB = United Kingdom, GR = Greece,

*subject*: code for the participant

*color*: color term

*admiration* to *shame*: ratings of emotion intensity. 0 = no association between the color term and a given emotion. 5 = strongest intensity of an associated emotion.