## COSC522

M8.10 Project Proposal Overview

## Group 5

Atul Gupta, Britton Almy, Mary Proctor, Armando Rangel

# Predicting Happiness Score: The Role of Socio-Cultural and Political Factors in Regional Well-being

## Overview:

This project focused on developing a robust regression-based machine learning model to accurately predict the happiness scores of various global regions. The team utilized the **World Happiness Report** dataset available on Kaggle and conducted detailed **feature extraction and selection** through **correlation analysis**, eliminating variables that were either irrelevant or highly correlated with others.

The selected regression algorithms included **Support Vector Regression (SVR)** and **Deep Learning Scalar Regression**. Both models were trained and fine-tuned to optimize predictive accuracy. Model performance was evaluated using a **holdout test set**, with metrics such as **Mean Absolute Error (MAE)**, **Mean Squared Error (MSE)**, and **R-squared** serving as evaluation criteria.

Each team member also tested assigned subsets of the dataset independently, applying both regression models. The individual results were later compiled and compared to assess consistency across subsets and identify any discrepancies or unique insights.

## Goals:

1. Achieve an R-squared value close to 1, demonstrating strong predictive accuracy.
2. Provide actionable insights into the socio-cultural and political factors most significantly contributing to regional happiness.
3. Identify limitations within our models and outline potential improvements for future research.

Project Timeline (4 Weeks)

### Week 1:

* Acquire and preprocess dataset.
* Conduct initial exploratory data analysis (EDA).
* Define and assign subsets of the data for individual testing by each team member.

### Week 2:

* Perform detailed correlation analysis for feature selection.
* Begin model development, training initial SVR and Deep Learning models.
* Start individual testing with assigned subsets.

### Week 3:

* Fine-tune model parameters and conduct validation checks.
* Individual members complete testing and record findings.

### Week 4:

* Consolidate and analyze individual testing results.
* Evaluate model performance using predefined metrics (MAE, MSE, R-squared).
* Compile final results and insights into a comprehensive report.

## Expected Outcomes:

The project was expected to yield **highly accurate predictive models** for regional happiness scores, provide **valuable insights for policymakers**, and contribute to the broader understanding of global happiness and well-being through data-driven analysis.