The economics of Happiness

LIlian Sao de Rivera #4
Data Mining Class

"Can economic indicators like GDP, GINI, and Governance Indexes predict the score of happiness of a country?"

Index of Happiness

- Based on a poll made by Gallup
- Canadian Institute for Advanced Research
  - First Report 2006
- Measures the level of happiness base on GDP, Family relations, Freedom , Generosity, Perception of Corruption, Health.
-

"Happiness is increasingly considered the proper measure of social progress and the goal of public policy. "

-Word Happiness Report - 2017.

Why this project?

Savings

Investment

Ratiring

Wellness

Project Structure

- Understanding the information
- Preprocessing
- Machine Learning Algorithms
- Implementation
- Conclusions and Results

Datasets 1/3

Happiness Score

Years	2017
No. Observations	156 countries
No. of Features	12
Type of file	CSV file

GDP

Years	2006 thru 2017
No. Observations	264 Countries
No. of Features	63
Type of File	CSV

Datasets 2/3

Governance Indicators

Years of Information	1996 thru 2018
No. Observations	216 Countries
No. of Features	7
Type of File	Excel File

GINI

Years	2006 thru 2017
No. Observations	264 Countries
No. of Features	63
Type of File	CSV

Datasets 3/3

Country ISO Codes
(Apha-3)

No. Observations	264 Countries
No. of Features	9
Type of File	CSV

Datasets Challenges

ISO : first char has an odd value

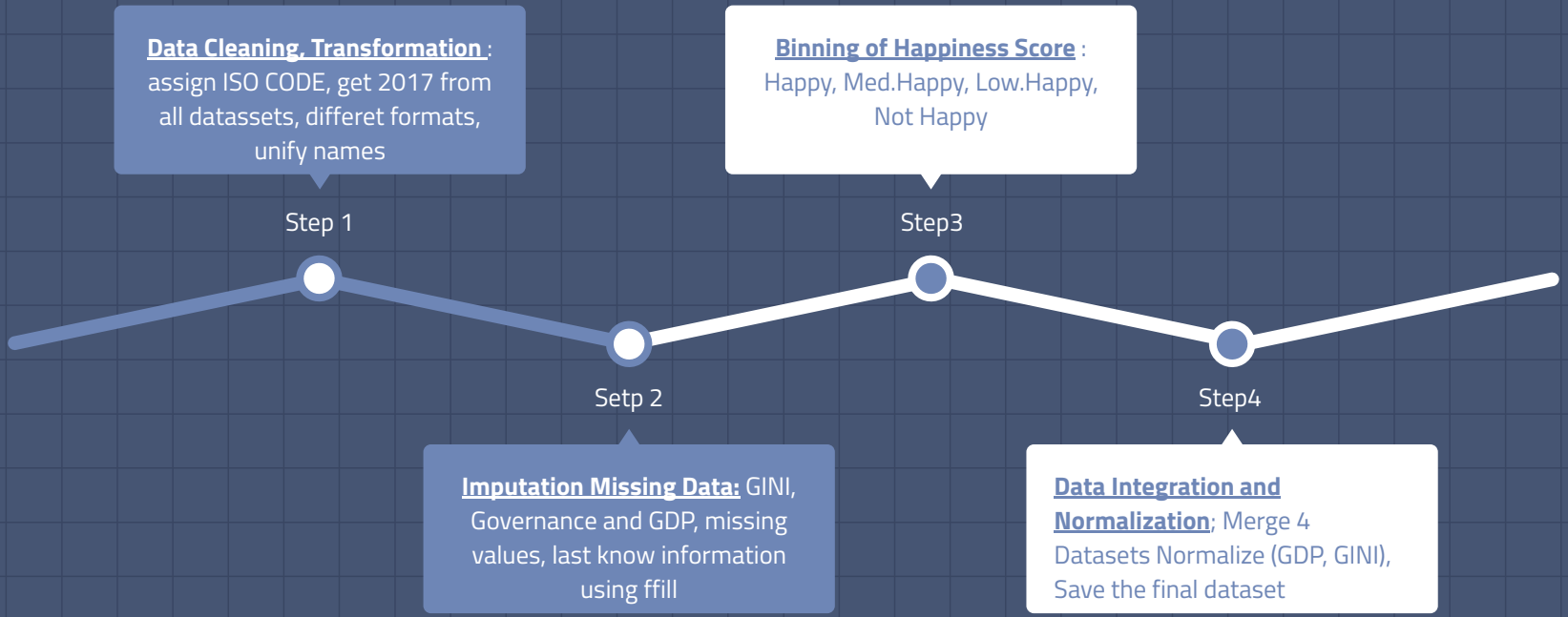
GINI: for each year, not all countries have information. The indicator ranges from 26 to 65, estimate by year is stack horizontally

GDP: for each year, not all countries have information. The amount of the indicator is in dollars, estimate by year is stack horizontally

Governance : One excel sheet per indicator with several information columns. Estimates by year are stacked horizontally

Happiness: The only feature to join with other indicators is the name of the country

Datasets- Preprocessing



Final Dataset (140 observations)

Country	Country_code	HappinessScale	
HappinessScore	GDP	GINI	
ViolenceandAcccountability	PoliticalStability	GovEffectiveness	
RegulatoryQuality	RuleofLaW	ControlofCorruption	

ML Algorithms

Decision Tree Classifier

The target is a four class feature.

Expected to predict happiness

Less than 25 features

Continuos values

Gridi function used is Entropy

Random Forest Classifier

Expected to validate or simplify the DT Classifier as a comparative approach

Delivers importance of features, importat to further analysis

Base on trees, os enhances the DT results

Implementation : PyQt5

Purpose

Create a desktop application with interactive response to manipulate the parameters and the analysis of information.

Design

EDA Analysis

- Graphs

Dashboards

- Decision Tree
- Random Forest

Results

Easy to change parameters

Easy to add or drop features

Immediate response without changes in the code

Technical Summary

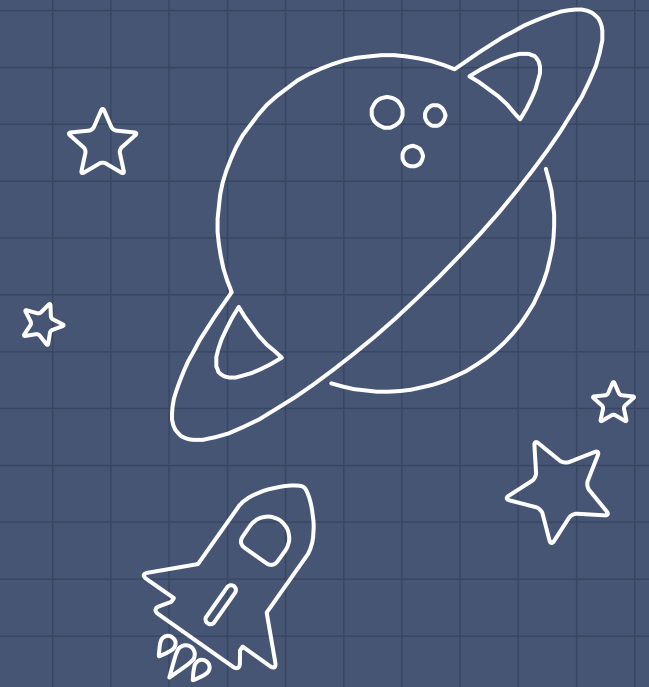
- The work was divided in two parts
  - Preprocessing of Datasets
    - Data was unified as one clean and useful dataset.
  - Desktop application
    - The dataset is loaded in the main() function and the application uses this information.

DESKTOP APPLICATION

PyQt5 , Example of Dashboard

DEMO APP

The economics of happiness



Conclusions-2

- The simpler the better : Random Forest seems the best classifier for this experiment

- PyQt5 application was useful to create the iterative Dashboards.

Conclusions-1

- Random Forest shows better results with the same features and sample size than Decision Tree classifier.
- GINI and Political Stability showed a low contribution to the model
- Decision Tree results can be enhanced by dropping GINI, PoliticalStability.

Recommendations

-The EDA analysis showed high correlation a regression, thus a regression model could be implemented

-There seems missing information in the results

Other features could be included

Health factors , Education Factors

-More graphics are needed it in the EDA analysis

-The analysis was made over 1 year, the inclusio of more years would be

THANKS!

Any questions?

You can find me at

- lsaodr@gwu.edu

CREDITS

- Presentation template by SlidesCarnival
- World Bank Database
- World Happiness Report