ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Analysis of Countries Happiness

ONE LOVE. ONE FUTURE.

Agenda

- Introduction
- Data Collection and Pre-processing
- Exploratory Data Analysis and Visualization
- Conclusion

HUST

hust.edu.vn fb.com/dhbkhn

Introduction

Introduction

- Rapid development of global economy: global Gross Domestic Product (GDP) has more than tripled since 1990 [1].
- Economic growth does not always guarantee well-being: one in five adults in the US had received mental health treatment in 2019 [2].
- Alternative measurement: happiness score by World Happiness Report.
- Which socioeconomic factors contribute to happiness?

→ Analyzing the **correlation** of **countries happiness scores** with **socioeconomic factors**.

Methodology

- **Data Collection and Preprocessing:** gather data from multiple sources and transform to high quality single database.
- **Exploratory Data Analysis:** summarize and understand data.
- **Data Visualization:** understand and communicate insights from data through visual representation.

HUST

hust.edu.vn fb.com/dhbkhn

Data Collection and Pre-processing

Countries Happiness Score

- Response variable.
- Annual happiness report provided by World Happiness Report [3]-[5].
- Scale from 0 to 10, with 10 is the highest score.
- Averaged of 2017-2019 period.

Countries Socioeconomic Factors

- Availability challenge.
- Fourteen explanatory variables.
- Public data from The World Bank, United Nations Development Programme (UNDP), World Health Organization (WHO), United Nations Office on Drugs and Crime (UNOCD).

List of Socioeconomic Factors

- Population density (people per sq. km of land area) [7].
- GDP per capita (current US$) [8].
- PM2.5 air pollution, mean annual exposure ($\mu g/m^3$) [9].
- Forest area (% of land area) [10].
- Consumer price index (2010 = 100) [11].
- Mean years of schooling [12].
- Life expectancy at birth (years) [13].
- Gender Development Index (GDI) [14].
- CO_2 emissions per capita (metric tons) [15].
- Unemployment rate (% of total labor force) [16].
- Healthy life expectancy at birth (HALE) [17].
- Current health expenditure (% of GDP) [18].
- Age-standardized suicide rate (per 100,000 population) [19].
- Prison population (per 100,000 population) [20].

Data Pre-processing

- Multi-source data integration by three-letter country code ISO 3166-1 alpha-3 [6].
- Handling missing data:
  - Country with more than 10 missing values: remove.
  - Otherwise: replace with mean value of the region.

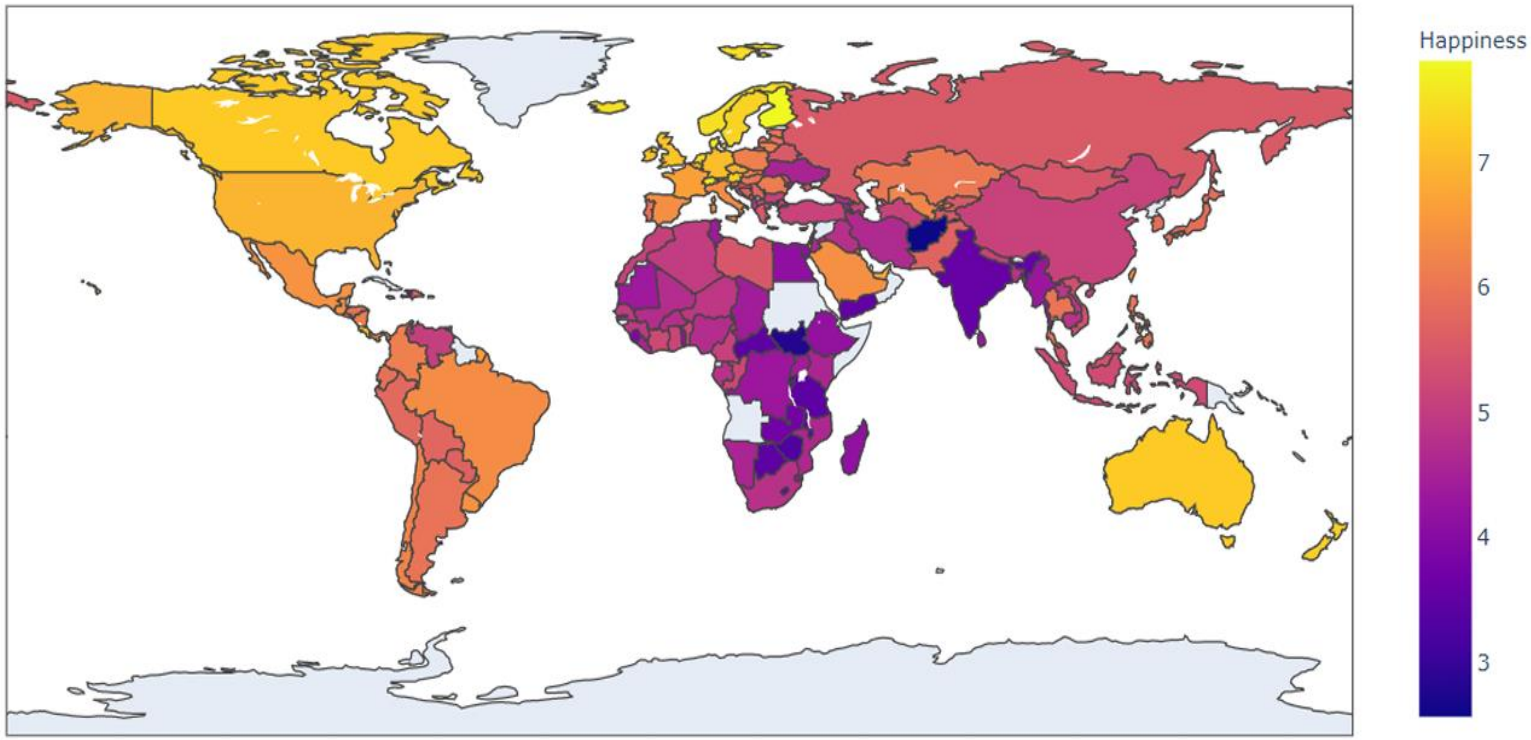
HUST

hust.edu.vn fb.com/dhbkhn

Exploratory Data Analysis and Visualization

How Happy Are People Globally?

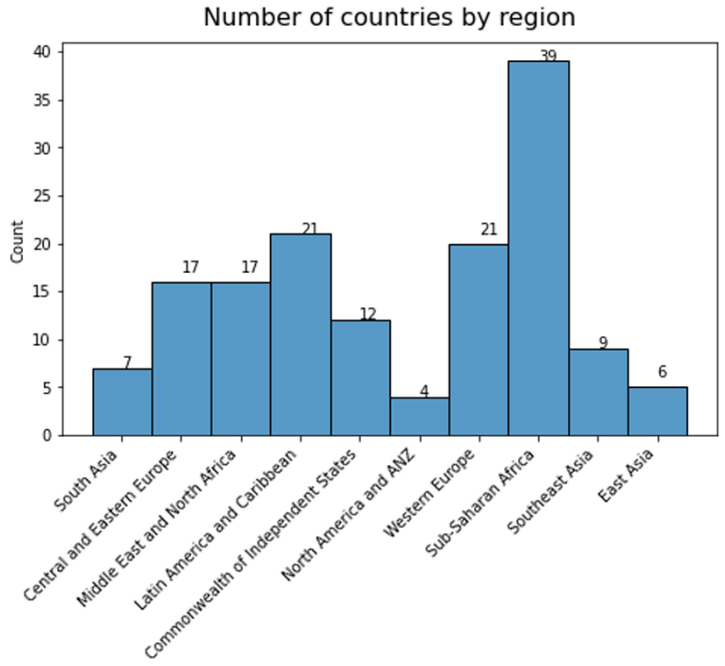
- Happier regions: Europe and North America.
- Less happy region: Africa.



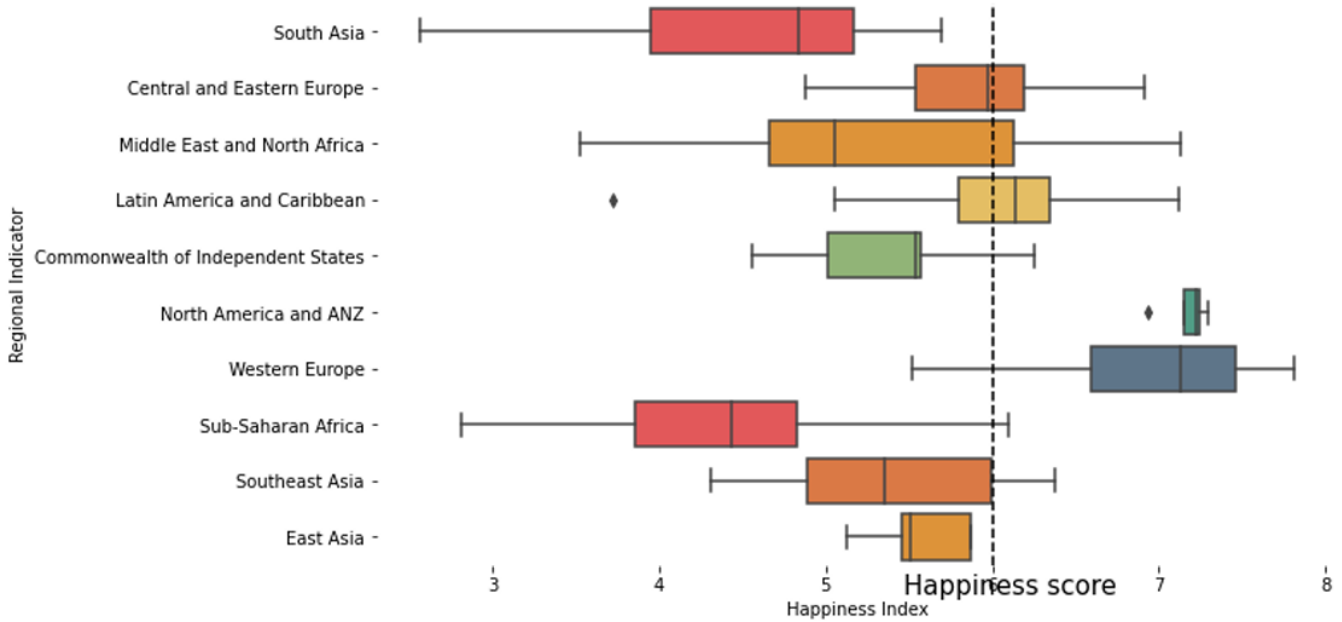
Global map of countries with happiness score.

Happiness Score By Regions

- Happy regions: Western Europe, North America and ANZ.
- Least happy region: Sub-Sahara Africa.
- Central and Eastern Europe: average happiness score.



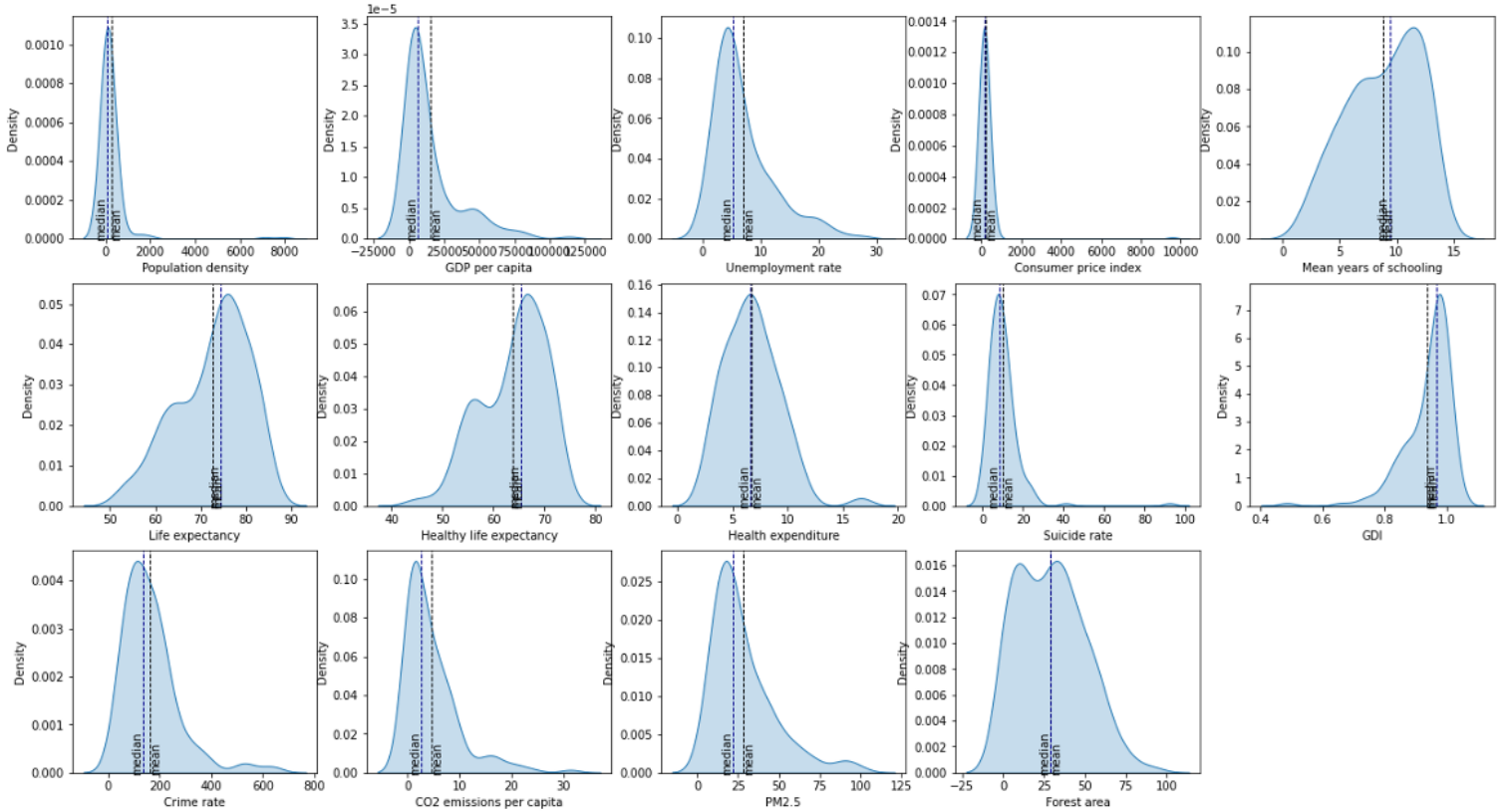
Bar chart for number of countries by region



Box plot of happiness score by regions. Dash line indicated average value.

Socioeconomic Variables Statistics

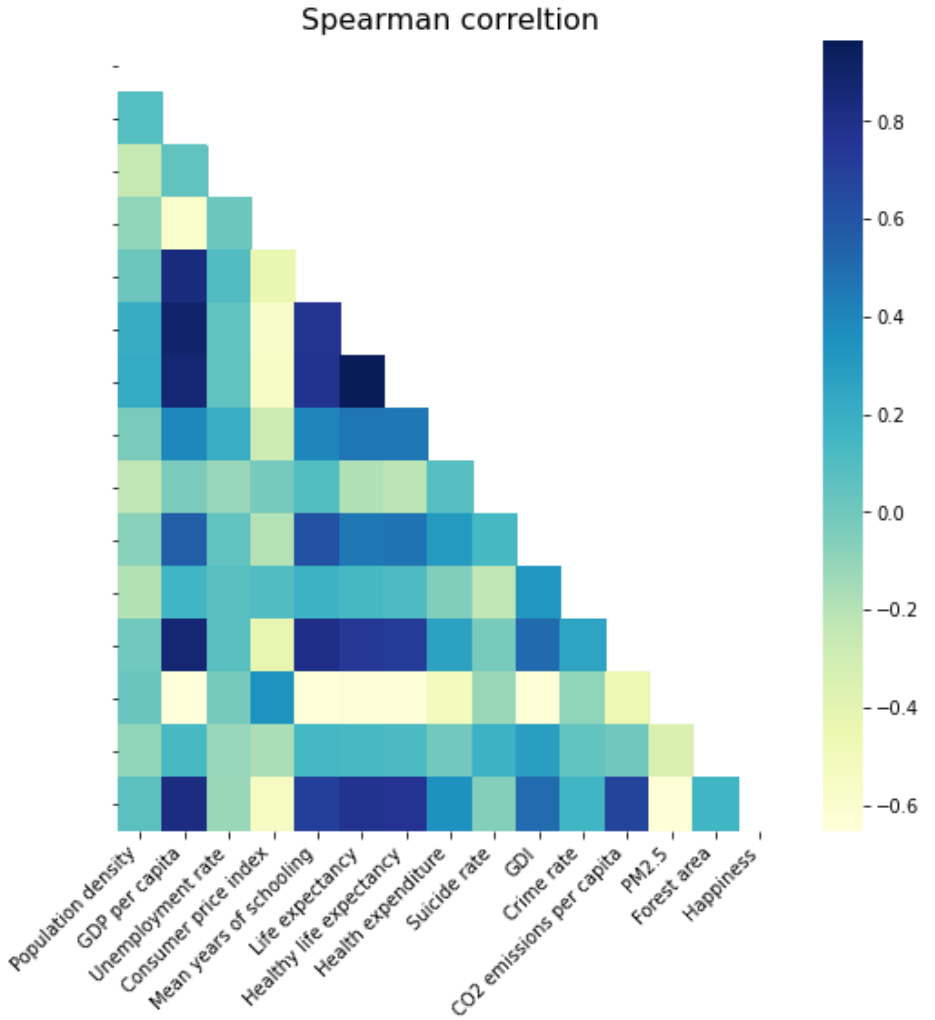
- Density plot of socioeconomic variables.
- Approximately follow normal distribution.
- Variables with outliers: population density, consumer price index, suicide rate.



Density plot of socioeconomic variables.

Correlation of Happiness Score with Socioeconomic Factors

- Happiness score is:
  - **Positively correlated** with: GDP per capita, Mean years of schooling, Life expectancy, Healthy life expectancy, CO2 emissions per capita, and GDI.
  - **Negatively correlated** with: Consumer price index and PM 2.5.
  - **Uncorrelated** with: population density, unemployment rate, health expenditure, suicide rate, crime rate, and forest area.

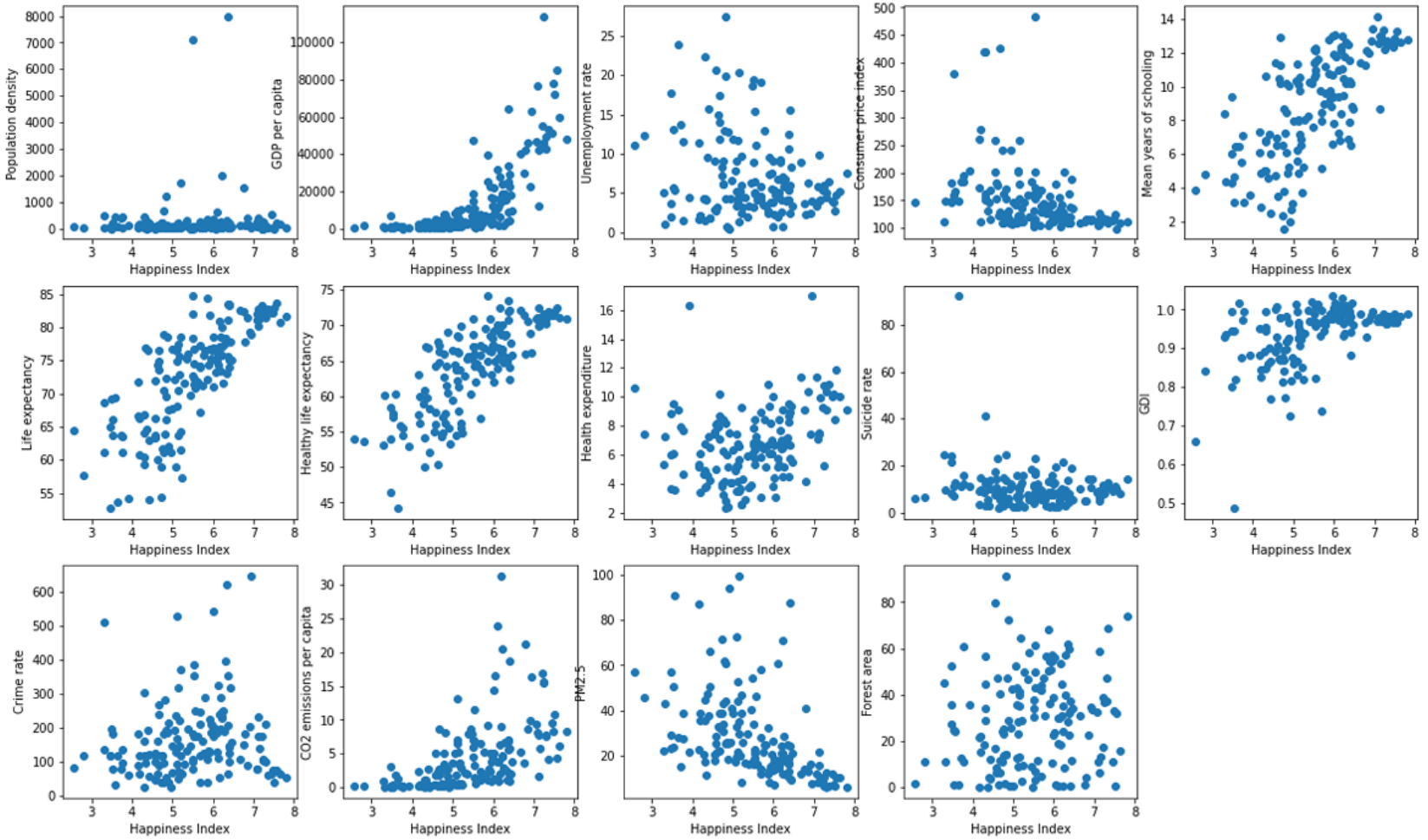


Heatmap visualization of spearman correlation between variables.

⚠ Correlation does not imply causation!

Scatter Plots of Happiness Score with Socioeconomic Factors

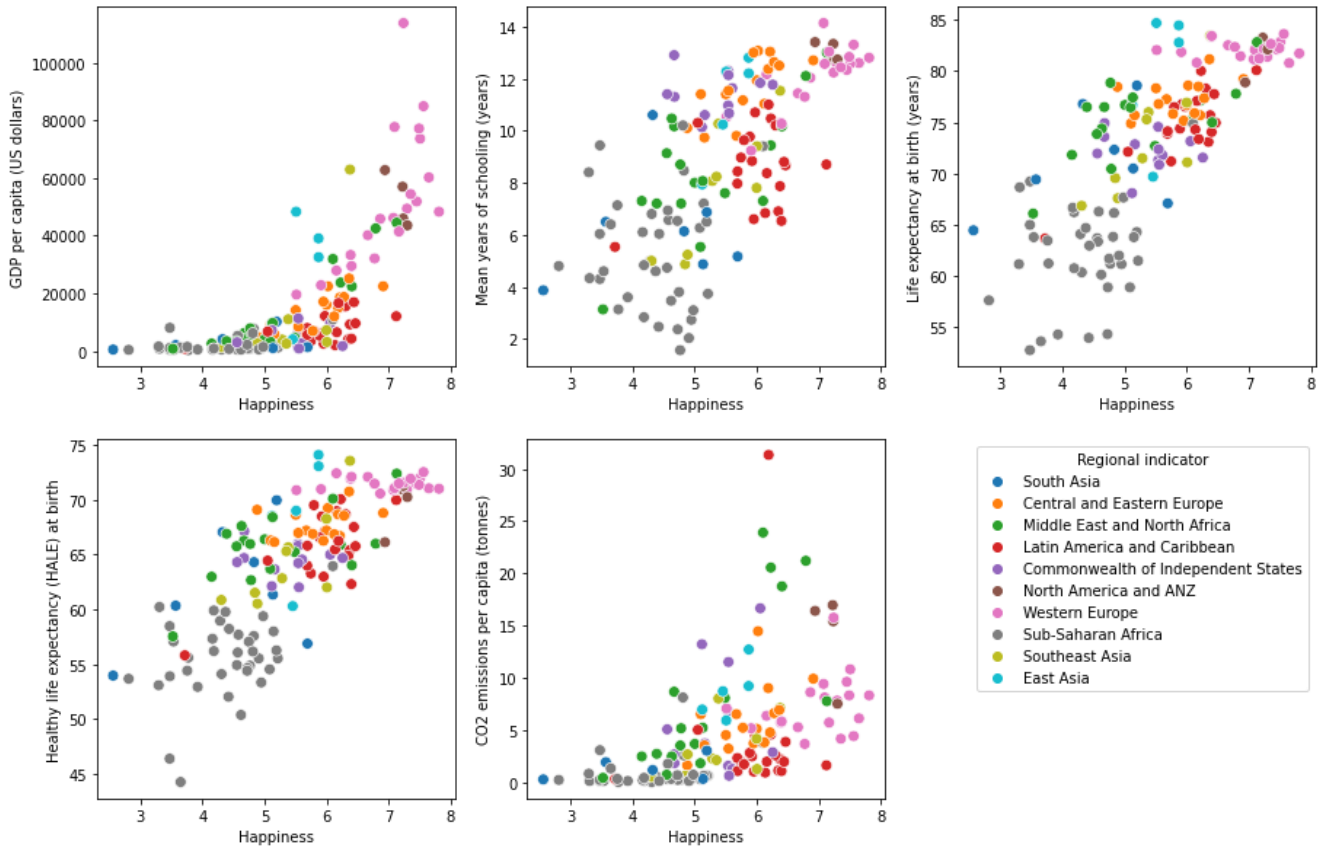
- More details of happiness score aforementioned correlations:
  - **Linear relationship**, e.g. life expectancy and mean years of schooling.
  - **Exponential relationship**, e.g. GDP per capita.



Scatter plots of happiness score with socioeconomic factors.

Scatter Plots of Happiness Score with Socioeconomic Factors

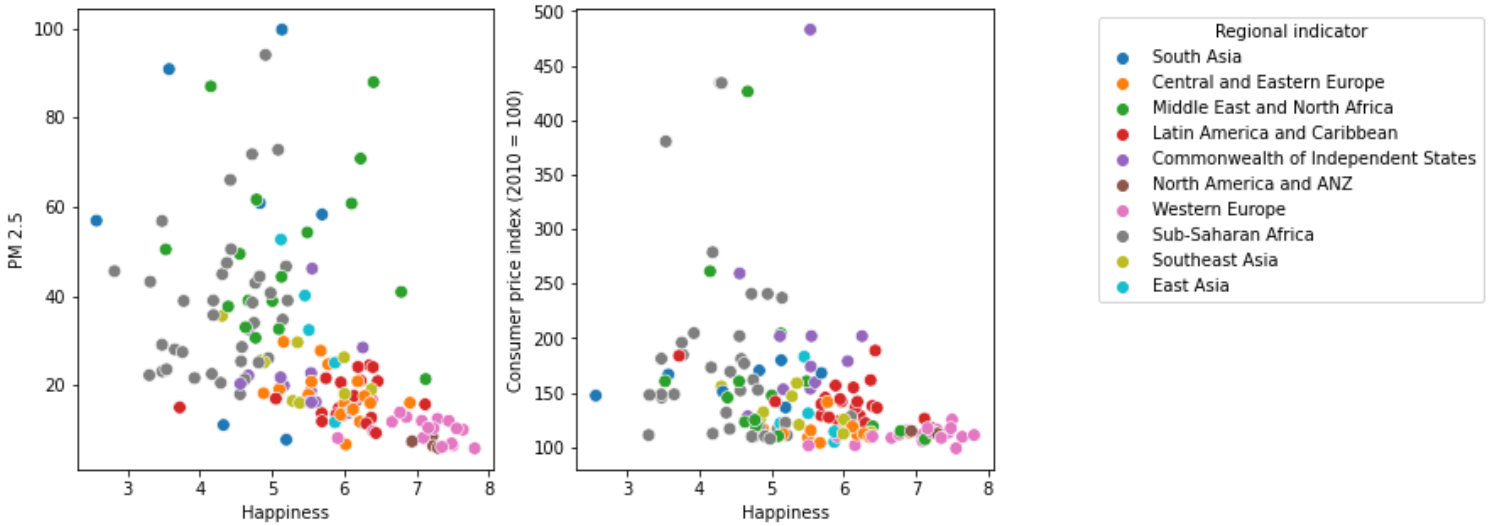
- Scatter plots of correlated variables with happiness score, grouped by region.
- Countries in the same region forms a cluster, e.g. Western Europe.



Scatter plots of happiness score with positively correlated variables by region.

Scatter Plots of Happiness Score with Socioeconomic Factors

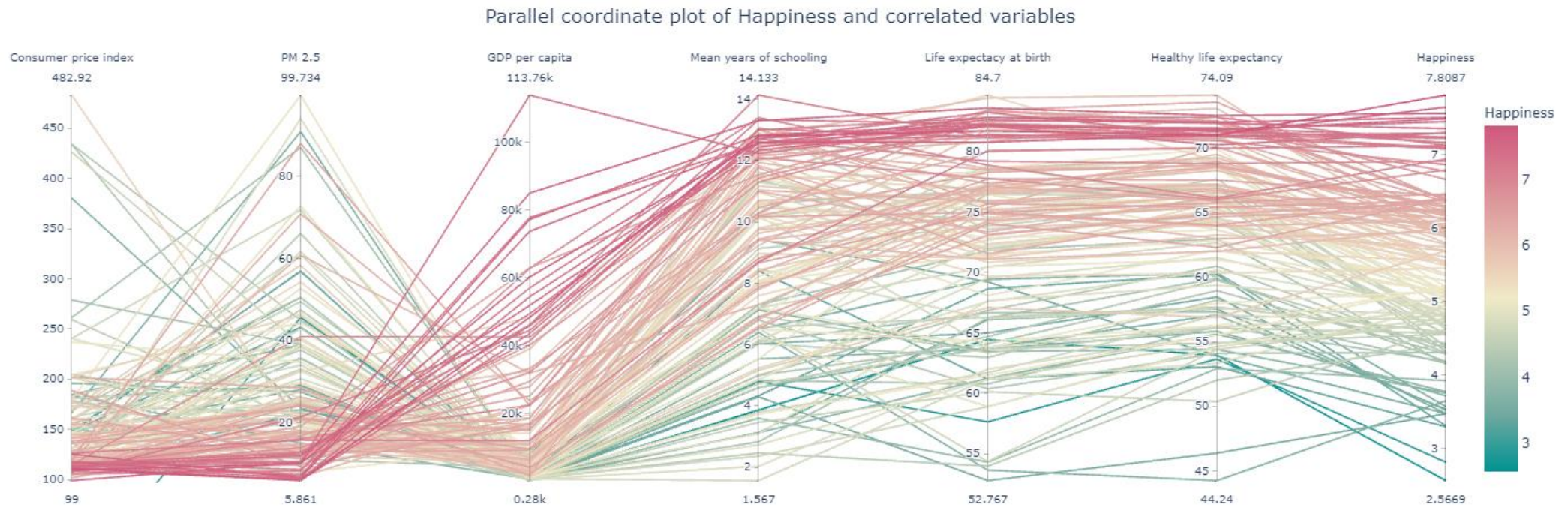
- Scatter plots of correlated variables with happiness score, grouped by region.
- Countries in the same region forms a cluster, e.g. Western Europe.



Scatter plots of happiness score with negatively correlated variables by region.

Characteristics of Happy Country

- A happy country has:
  - Low living inflation (low consumer price index) and good air quality (low PM 2.5).
  - Above-averaged GDP per capita.
  - Good education (high mean years of schooling) and health care system (high life expectancy at birth and healthy life expectancy).



Parallel coordinate plot of happiness score and correlated variables.

HUST

hust.edu.vn fb.com/dhbkhn

Conclusion

Conclusion

- Analysis of correlation of countries happiness score with socioeconomic factors.
- Factors correlated with happiness score:
  - Positive correlation: GDP per capita, Mean years of schooling, Life expectancy, Healthy life expectancy, CO2 emissions per capita, and GDI.
  - Negative correlation: Consumer price index and PM 2.5.

References

[1] World Bank. (n.d.). GDP (current US$) [Data set]. The World Bank. Retrieved [Jan 3, 2022], from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

[2] Terlizzi, E. P., & Zablotsky, B. (2020). Mental health treatment among adults: United States, 2019.

[3] Helliwell, J. F., Layard, R., & Sachs, J. (Eds.). (2017). World Happiness Report 2017. Sustainable Development Solutions Network. https://worldhappiness.report/ed/2017/

[4] Helliwell, J. F., Layard, R., & Sachs, J. (Eds.). (2018). World Happiness Report 2018. Sustainable Development Solutions Network. https://worldhappiness.report/ed/2018/

[5] Helliwell, J. F., Layard, R., Sachs, J., & De Neve, J.-E. (Eds.). (2019). World Happiness Report 2019. Sustainable Development Solutions Network. https://worldhappiness.report/ed/2019/

[6] International Organization for Standardization (ISO). (2006). Iso 3166-1: codes for the representation of names of countries and their subdivisions–Part 1: country codes.

[7] World Bank. (n.d.). Population density (people per sq. km of land area). Retrieved [Dec 3, 2021], from https://data.worldbank.org/indicator/EN.POP.DNST

[8] World Bank. (n.d.). GDP per capita (current US$). Retrieved [Dec 3, 2021], from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

[9] World Bank. (n.d.). PM2.5 air pollution, mean annual exposure (μg/m^3). Retrieved [Dec 3, 2021], from https://data.worldbank.org/indicator/EN.ATM.PM25.MC.M3

References

[10] World Bank. (n.d.). Forest area (% of land area). Retrieved [Dec 3, 2021], from https://data.worldbank.org/indicator/AG.LND.FRST.ZS

[11] World Bank. (n.d.). Consumer price index (2010 = 100). Retrieved [Dec 3, 2021], from https://data.worldbank.org/indicator/FP.CPI.TOTL

[12] United Nations Development Programme. (n.d.). Mean years of schooling. Retrieved [Dec 3, 2021], from http://hdr.undp.org/en/indicators/103006

[13] United Nations Development Programme. (n.d.). Life expectancy at birth (years). Retrieved [Dec 3, 2021], from http://hdr.undp.org/en/indicators/69206

[14] United Nations Development Programme. (n.d.). Gender Development Index (GDI). Retrieved [Dec 3, 2021], from http://hdr.undp.org/en/indicators/137906

[15] United Nations Development Programme. (n.d.). CO_2 emissions per capita (metric tons). Retrieved [Dec 3, 2021], from http://hdr.undp.org/en/indicators/195606

[16] United Nations Development Programme. (n.d.). Unemployment rate (% of total labor force). Retrieved [Dec 3, 2021], from http://hdr.undp.org/en/indicators/140606

[17] World Health Organization. (n.d.). Healthy life expectancy at birth (HALE). Retrieved [Dec 3, 2021], from https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-ghe-hale-healthy-life-expectancy-at-birth

References

[18] World Health Organization. (n.d.). Current health expenditure (% of GDP). Retrieved [Dec 3, 2021], from https://www.who.int/data/gho/data/indicators/indicator-details/GHO/current-health-expenditure-(che)-as-percentage-of-gross-domestic-product-(gdp)-(-)

[19] World Health Organization. (n.d.). Age-standardized suicide rate (per 100,000 population). Retrieved [Dec 3, 2021], from https://www.who.int/data/gho/data/themes/mental-health/suicide-rates

[20] United Nations Office on Drugs and Crime. (n.d.). Prison population (per 100,000 population). Retrieved [Dec 3, 2021], from https://dataunodc.un.org/content/prison-population-national-vs-regional-estimates

HUST

hust.edu.vn fb.com/dhbkhn

THANK YOU !