Guided Practice Project Overview

Estimated time: 10 minutes

In this practice project, you will use a real-world data set and practice generative AI to generate Python codes that can perform data preparation, analysis, visualization and dashboarding.

Project Scenario

You have been employed as a data analyst by a Healthcare consultancy firm which has been conducting a survey on the state of global happiness annually. The World Happiness Report offers valuable insights into factors influencing happiness across countries. The firm wants you to produce a report to find out whether there are demographic, regional, and/or economic characteristics that lead to a better life.

About the data set

The World Happiness Report is a landmark survey of the state of global happiness. The reports review the state of happiness in the world today and show how the new science of happiness explains personal and national variations in happiness. This is a public dataset available on the Kaggle website as World Happiness Report under the CC0: Public Domain license.

For this guided practice project, you will work on the **year 2016** data, which has been slightly modified for the purpose of this guided practice project.

Hold down the Ctrl key while you click to download it from here: 2016 data

Attributes of this dataset have been explained below.

Variable	Description
Country	Name of the country
Region	Region the country belongs to
Happiness Rank	Rank of the country based on the Happiness Score
Happiness Score	A metric measured in 2016 by asking the sampled people the question: "How would you rate your happiness?"
Lower Confidence Interval	Lower confidence interval of the Happiness Score
Upper Confidence Interval	Upper confidence interval of the Happiness Score
Economy (GDP per Capita)	The extent to which GDP contributes to the calculation of the Happiness Score
Family	The extent to which family contributes to the calculation of the Happiness Score
Health (Life Expectancy)	The extent to which life expectancy contributes to the calculation of the Happiness Score
Freedom	The extent to which freedom contributes to the calculation of the Happiness Score
Trust (Government Corruption)	The extent to which trust contributes to the calculation of the Happiness Score
Generosity	The extent to which generosity contributes to the calculation of the Happiness Score
Dystopia Residual	Dystopia is an imaginary country that has the world's least-happy people. The residuals, or unexplained components, differ for each country, reflecting the extent to which the six variables either over- or under-explain average 2014-2016 life evaluations. These residuals have an average value of approximately zero over the whole set of countries.

Tasks in the project

The project tasks are data preparation, analysis, visualization, and dashboarding. Based on the data set, you must write prompts to generate the Python codes for performing specific tasks. You can access a JupyterLite-based testing environment to test the generated codes using the Generative AI classroom prompts.

The tasks assigned to you are as follows:

1. There might be a few missing values in the dataset. Data cleaning will be a part of the assignment.
2. You have to perform exploratory data analysis to draw keen insights on the data:
    - Identify the GDP per capita and Healthy Life Expectancy of the top 10 countries.and represent it as a bar chart.
    - Find the correlation between the Economy (GDP per Capita), Family, Health (Life Expectancy), Freedom, Trust (Government Corruption), Generosity, and Happiness Score.
    - Create a scatter plot to identify the effect of GDP per Capita on Happiness Score in various Regions.
    - Create a pie chart to present Happiness Score by region.
    - Create a map to display GDP per capita of countries and include Healthy Life Expectancy as a tooltip.
3. Create a dashboard with at least four of the above visualizations.
4. Generate the narrative to present the dashboard.

Author(s)

Dr. Pooja