**Proyecto: Análisis del World Happiness Report**

**Objetivo:**  
Explorarán y analizarán el World Happiness Report para identificar patrones y tendencias en la felicidad global, y cómo diferentes factores contribuyen a la felicidad en distintos países. A través de este proyecto, los estudiantes desarrollarán habilidades en el análisis de datos y la visualización para entender mejor las métricas de bienestar y su impacto.

**Parte 1: Preprocesamiento de Datos**

1. **Importación y Exploración Inicial:**
   * Carga el dataset en un entorno de programación como Jupyter Notebook utilizando pandas.
   * Examina las primeras filas del conjunto de datos para familiarizarte con su estructura y contenido.
2. **Estadísticas Básicas:**
   * Usa pandas.describe() para calcular estadísticas descriptivas y entender la distribución de cada variable.
   * Revisa la distribución de las variables categóricas y numéricas con pandas.value\_counts().
3. **Revisión de Valores Faltantes:**
   * Identifica y maneja los valores faltantes utilizando técnicas de imputación o eliminación, según sea apropiado.
   * Visualiza los valores faltantes con un mapa de calor usando seaborn para identificar patrones.
4. **Conversión de Datos Categóricos:**
   * Si es necesario, convierte las variables categóricas a numéricas utilizando pd.get\_dummies() o funciones de mapeo.
5. **Detección de Outliers:**
   * Identifica y maneja outliers en las variables numéricas utilizando boxplots y análisis estadístico.
6. **Renombrado de Columnas:**
   * Asegúrate de que los nombres de las columnas sean claros y descriptivos para facilitar el análisis.

**Parte 2: Análisis Exploratorio de Datos (EDA)**

1. **Distribución de la Felicidad:**
   * Crea un histograma para visualizar la distribución de la "Happiness Score" y analiza las tendencias generales.
2. **Análisis Univariado:**
   * **GDP per Capita:** Analiza la distribución del PIB per cápita y su relación con la felicidad.
   * **Life Expectancy:** Visualiza la distribución de la esperanza de vida y su impacto en la felicidad.
   * **Generosity y Corruption:** Crea gráficos de barras para explorar la distribución de generosidad y corrupción.
3. **Análisis Bivariado:**
   * **Happiness Score vs. GDP per Capita:** Crea un gráfico de dispersión para analizar cómo el PIB per cápita influye en el puntaje de felicidad.
   * **Happiness Score vs. Life Expectancy:** Usa gráficos de dispersión o de burbujas para visualizar la relación entre la felicidad y la esperanza de vida.
   * **Social Support vs. Happiness Score:** Analiza cómo el apoyo social afecta la felicidad utilizando gráficos de dispersión.
4. **Análisis Comparativo:**
   * **Comparación entre Países:** Crea gráficos de barras para comparar la felicidad en diferentes países o regiones.
   * **Comparación Temporal:** Analiza cómo han cambiado los puntajes de felicidad de un año a otro si los datos están disponibles para múltiples años.
5. **Exploración de Factores Contributivos:**
   * Analiza cómo los diferentes factores (como Generosity, Trust Government Corruption, etc.) contribuyen al puntaje de felicidad.
   * Crea gráficos de barras apiladas o diagramas de áreas para visualizar las contribuciones de cada factor.

**Parte 3: Visualización Avanzada**

1. **Visualizaciones Interactivas:**
   * Utiliza plotly o bokeh para crear visualizaciones interactivas que permitan una exploración más detallada de los datos.
2. **Dashboard de Visualización:**
   * Crea un dashboard utilizando herramientas como Dash o Streamlit para presentar visualizaciones clave de manera accesible y atractiva.
3. **Exploración Adicional:**
   * Anima a los estudiantes a identificar y explorar visualizaciones adicionales que puedan aportar información relevante sobre la felicidad global.

**Evaluación del Proyecto**

* **Calidad del Preprocesamiento:**
  + Evalúa la capacidad de los estudiantes para limpiar y preparar los datos para el análisis.
* **Profundidad del Análisis Exploratorio:**
  + Valora la habilidad para identificar patrones y tendencias significativas en los datos.
* **Efectividad de las Visualizaciones:**
  + Mide la claridad y efectividad de las visualizaciones para comunicar los hallazgos.
* **Diseño del Dashboard:**
  + Evalúa el diseño y la funcionalidad del dashboard en términos de usabilidad y presentación clara de la información.

**Conclusión**

Este proyecto proporciona a una oportunidad para explorar cómo los diferentes factores contribuyen a la felicidad en los países de todo el mundo. Al enfocarse en el análisis de datos y la visualización, desarrollarán habilidades críticas en el análisis de bienestar, útiles para carreras en ciencias sociales, políticas públicas y análisis de datos. Además, les ayudará a entender cómo las métricas de felicidad pueden informar las decisiones y mejorar el bienestar social.

**Caso 5**

**About Dataset**

**Context**

The World Happiness Report is a landmark survey of the state of global happiness. The first report was published in 2012, the second in 2013, the third in 2015, and the fourth in the 2016 Update. The World Happiness 2017, which ranks 155 countries by their happiness levels, was released at the United Nations at an event celebrating International Day of Happiness on March 20th. The report continues to gain global recognition as governments, organizations and civil society increasingly use happiness indicators to inform their policy-making decisions. Leading experts across fields – economics, psychology, survey analysis, national statistics, health, public policy and more – describe how measurements of well-being can be used effectively to assess the progress of nations. 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.

**Content**

The happiness scores and rankings use data from the Gallup World Poll. The scores are based on answers to the main life evaluation question asked in the poll. This question, known as the Cantril ladder, asks respondents to think of a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale. The scores are from nationally representative samples for the years 2013-2016 and use the Gallup weights to make the estimates representative. The columns following the happiness score estimate the extent to which each of six factors – economic production, social support, life expectancy, freedom, absence of corruption, and generosity – contribute to making life evaluations higher in each country than they are in Dystopia, a hypothetical country that has values equal to the world’s lowest national averages for each of the six factors. They have no impact on the total score reported for each country, but they do explain why some countries rank higher than others.

**Inspiration**

What countries or regions rank the highest in overall happiness and each of the six factors contributing to happiness? How did country ranks or scores change between the 2015 and 2016 as well as the 2016 and 2017 reports? Did any country experience a significant increase or decrease in happiness?

**What is Dystopia?**

Dystopia is an imaginary country that has the world’s least-happy people. The purpose in establishing Dystopia is to have a benchmark against which all countries can be favorably compared (no country performs more poorly than Dystopia) in terms of each of the six key variables, thus allowing each sub-bar to be of positive width. The lowest scores observed for the six key variables, therefore, characterize Dystopia. Since life would be very unpleasant in a country with the world’s lowest incomes, lowest life expectancy, lowest generosity, most corruption, least freedom and least social support, it is referred to as “Dystopia,” in contrast to Utopia.

**What do the columns succeeding the Happiness Score(like Family, Generosity, etc.) describe?**

The following columns: GDP per Capita, Family, Life Expectancy, Freedom, Generosity, Trust Government Corruption describe the extent to which these factors contribute in evaluating the happiness in each country.  
The Dystopia Residual metric actually is the Dystopia Happiness Score(1.85) + the Residual value or the unexplained value for each country as stated in the previous answer.

If you add all these factors up, you get the happiness score so it might be un-reliable to model them to predict Happiness Scores.