

tarea-2-python

October 10, 2023

```
[1]: #Nombre : Anshela Melania Castillo Nicolalde
    #Tarea: Dos
    # Importamos las bibliotecas necesarias para el desarrollo del programa

[2]: import pandas as pd
    import numpy as np

[3]: # Cargamos los da markdowntos desde el repositorio de GitHub como se menciona

[4]: url = "https://raw.githubusercontent.com/cienciadedatos/datos-de-miercoles/
    ↪master/datos/2019/2019-08-07/felicidad.csv"
    data = pd.read_csv(url)

[5]: # Observamos los datos de felicidad

[6]: print(data.head())
```

	pais	anio	escalera_vida	log_pib	soporte_social	\
0	Afghanistan	2008	3.723590	7.168690	0.450662	
1	Afghanistan	2009	4.401778	7.333790	0.552308	
2	Afghanistan	2010	4.758381	7.386629	0.539075	
3	Afghanistan	2011	3.831719	7.415019	0.521104	
4	Afghanistan	2012	3.782938	7.517126	0.520637	

	expectativa_vida	libertad	generosidad	percepcion_corrupcion	\
0	50.799999	0.718114	0.177889	0.881686	
1	51.200001	0.678896	0.200178	0.850035	
2	51.599998	0.600127	0.134353	0.706766	
3	51.919998	0.495901	0.172137	0.731109	
4	52.240002	0.530935	0.244273	0.775620	

	afecto_positivo	afecto_negativo	confianza	calidad_democracia	\
0	0.517637	0.258195	0.612072	-1.929690	
1	0.583926	0.237092	0.611545	-2.044093	
2	0.618265	0.275324	0.299357	-1.991810	
3	0.611387	0.267175	0.307386	-1.919018	
4	0.710385	0.267919	0.435440	-1.842996	

	calidad_entrega	de_escalera_pais_anio	gini_banco_mundial	\
0	-1.655084	1.774662	NaN	
1	-1.635025	1.722688	NaN	
2	-1.617176	1.878622	NaN	
3	-1.616221	1.785360	NaN	
4	-1.404078	1.798283	NaN	

	gini_banco_mundial_promedio
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN

```
[7]: # Observamos y mostramos los nombres de las columnas de nuestra bbdd
```

```
[8]: nombres_columna = data.columns
print(nombres_columna)
```

```
Index(['pais', 'anio', 'escalera_vida', 'log_pib', 'soporte_social',
      'expectativa_vida', 'libertad', 'generosidad', 'percepcion_corrupcion',
      'afecto_positivo', 'afecto_negativo', 'confianza', 'calidad_democracia',
      'calidad_entrega', 'de_escalera_pais_anio', 'gini_banco_mundial',
      'gini_banco_mundial_promedio'],
      dtype='object')
```

```
[9]: # Identificamos el primer (min) y último año (max)
primer_anio = data['anio'].min()
ultimo_anio = data['anio'].max()
```

```
[10]: print(data['anio'].unique())
```

```
[2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2007 2006 2005]
```

```
[11]: # Medida de tendencia central y dispersión de la felicidad a nivel mundial para
      ↳ el primer año (media y std)
media_primer_anio = data[data['anio'] == primer_anio]['escalera_vida'].mean()
std_primer_anio = data[data['anio'] == primer_anio]['escalera_vida'].std()
```

```
[12]: # Print las medidas (media y std) del primer año
print(f"Medidas (media y std) del primer año:")
print(f"Media: {media_primer_anio}")
print(f"Std: {std_primer_anio}")
```

```
Medidas (media y std) del primer año:
Media: 6.446164272449635
Std: 0.9191426322726483
```

```
[13]: # Medida de tendencia central y dispersión de la felicidad a nivel mundial para
      ↪ el último año (media y std) esto quiere decir cuando se informa la media y
      ↪ la desviación estándar de la felicidad a nivel mundial para el último año,
      ↪ se está proporcionando una imagen resumida de cómo se distribuyen los
      ↪ niveles de felicidad en la población mundial en ese año. Esto puede ser útil
      ↪ para comparar años diferentes, identificar tendencias a lo largo del tiempo
      ↪ y comprender mejor la variabilidad en los niveles de felicidad en todo el
      ↪ mundo.
```

```
[14]: media_ultimo_anio = data[data['anio'] == ultimo_anio]['escalera_vida'].mean()
      std_ultimo_anio = data[data['anio'] == ultimo_anio]['escalera_vida'].std()
```

```
[15]: # Print las medidas (media y std) de último año
```

```
[16]: print(f"Medidas (media y std) del último año:")
      print(f"Media: {media_ultimo_anio}")
      print(f"Std: {std_ultimo_anio}")
```

```
Medidas (media y std) del último año:
Media: 5.502134340650895
Std: 1.1034612436939357
```

```
[17]: # Creacion de graficos donde para crear un gráfico de histograma con líneas
      ↪ verticales que representan la media y la desviación estándar de dos
      ↪ conjuntos de datos de felicidad en el primer y último año.
```

```
[18]: !pip install matplotlib
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: matplotlib in c:\programdata\anaconda3\lib\site-
packages (3.5.2)
Requirement already satisfied: pillow>=6.2.0 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (9.3.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (1.4.2)
Requirement already satisfied: numpy>=1.17 in c:\programdata\anaconda3\lib\site-
packages (from matplotlib) (1.23.5)
Requirement already satisfied: python-dateutil>=2.7 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: packaging>=20.0 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (21.3)
Requirement already satisfied: cyycler>=0.10 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: pyparsing>=2.2.1 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: fonttools>=4.22.0 in
c:\programdata\anaconda3\lib\site-packages (from matplotlib) (4.25.0)
```

Requirement already satisfied: six>=1.5 in c:\programdata\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)

```
[19]: import matplotlib.pyplot as plt
```

```
[20]: # Creación datos de ejemplo para el gráfico
np.random.seed(0)
felicidad_primer_anio = np.random.normal(media_primer_anio, std_primer_anio, 1000)
felicidad_ultimo_anio = np.random.normal(media_ultimo_anio, std_ultimo_anio, 1000)
```

```
[21]: import matplotlib.pyplot as plt

# Datos de felicidad del primer y último año
felicidad_primer_anio = [0.5, 0.8, 1.2, 1.5, 2.0, 2.2, 2.5, 2.7, 3.0, 3.2]
felicidad_ultimo_anio = [1.0, 1.2, 1.3, 1.6, 2.2, 2.5, 2.7, 2.8, 3.1, 3.4]

# Calcular la media y la desviación estándar de los datos
media_primer_anio = sum(felicidad_primer_anio) / len(felicidad_primer_anio)
std_primer_anio = (sum((x - media_primer_anio) ** 2 for x in felicidad_primer_anio) / len(felicidad_primer_anio)) ** 0.5

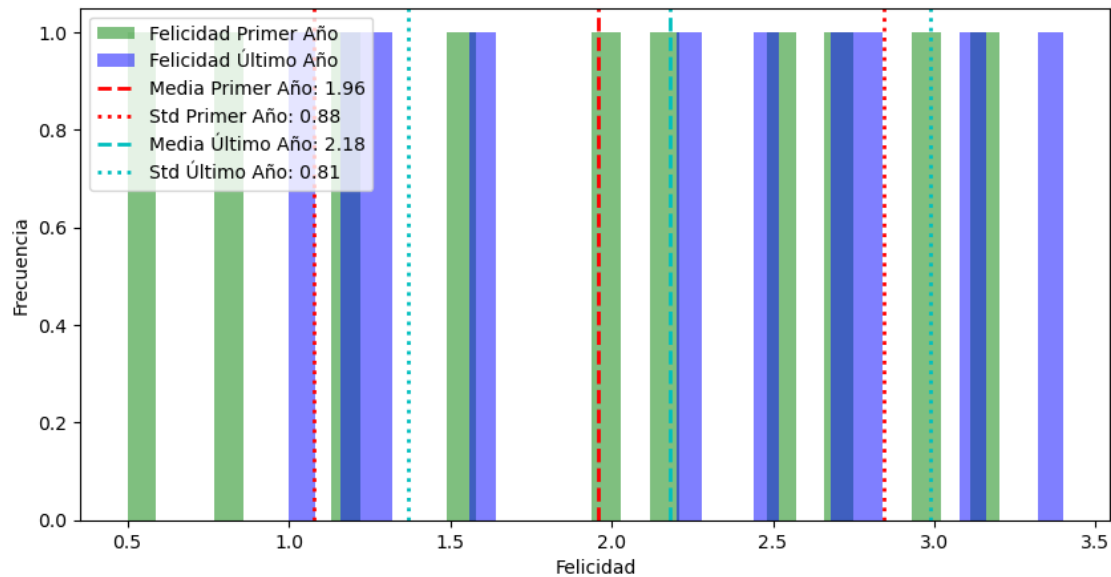
media_ultimo_anio = sum(felicidad_ultimo_anio) / len(felicidad_ultimo_anio)
std_ultimo_anio = (sum((x - media_ultimo_anio) ** 2 for x in felicidad_ultimo_anio) / len(felicidad_ultimo_anio)) ** 0.5

# Crear el gráfico
plt.figure(figsize=(10, 5))
plt.hist(felicidad_primer_anio, bins=30, alpha=0.5, color='g', label='Felicidad Primer Año')
plt.hist(felicidad_ultimo_anio, bins=30, alpha=0.5, color='b', label='Felicidad Último Año')

# Líneas verticales para representar la media y la desviación estándar
plt.axvline(media_primer_anio, color='r', linestyle='dashed', linewidth=2, label=f'Media Primer Año: {media_primer_anio:.2f}')
plt.axvline(media_primer_anio + std_primer_anio, color='r', linestyle='dotted', linewidth=2, label=f'Std Primer Año: {std_primer_anio:.2f}')
plt.axvline(media_primer_anio - std_primer_anio, color='r', linestyle='dotted', linewidth=2)
plt.axvline(media_ultimo_anio, color='c', linestyle='dashed', linewidth=2, label=f'Media Último Año: {media_ultimo_anio:.2f}')
plt.axvline(media_ultimo_anio + std_ultimo_anio, color='c', linestyle='dotted', linewidth=2, label=f'Std Último Año: {std_ultimo_anio:.2f}')
plt.axvline(media_ultimo_anio - std_ultimo_anio, color='c', linestyle='dotted', linewidth=2)
```

```
# Configuración de etiquetas y leyenda
plt.xlabel('Felicidad')
plt.ylabel('Frecuencia')
plt.legend()

# Mostrar el gráfico
plt.show()
```



```
[22]: print(data['pais'].unique())
```

```
['Afghanistán' 'Albania' 'Algeria' 'Angola' 'Argentina' 'Armenia'
'Australia' 'Austria' 'Azerbaijan' 'Bahrain' 'Bangladesh' 'Bielorrusia'
'Bélgica' 'Belize' 'Benin' 'Bhutan' 'Bolivia' 'Bosnia y Herzegovina'
'Botswana' 'Brasil' 'Bulgaria' 'Burkina Faso' 'Burundi' 'Cambodia'
'Camerún' 'Canadá' 'República Central Africana' 'Chad' 'Chile' 'China'
'Colombia' 'Comoros' 'Congo (Brazzaville)' 'Congo (Kinshasa)'
'Costa Rica' 'Croacia' 'Cuba' 'Chipre' 'República Checa' 'Dinamarca'
'Djibouti' 'República Dominicana' 'Ecuador' 'Egipto' 'El Salvador'
'Estonia' 'Etiopía' 'Finlandia' 'Francia' 'Gabón' 'Gambia' 'Georgia'
'Alemania' 'Ghana' 'Greece' 'Guatemala' 'Guinea' 'Guyana' 'Haití'
'Honduras' 'Hong Kong S.A.R. of China' 'Hungría' 'Islandia' 'India'
'Indonesia' 'Iran' 'Irak' 'Irlanda' 'Israel' 'Italia' 'Costa de Marfil'
'Jamaica' 'Japón' 'Jordania' 'Kazakhstan' 'Kenia' 'Kosovo' 'Kuwait'
'Kyrgyzstan' 'Laos' 'Latvia' 'Líbano' 'Lesotho' 'Liberia' 'Libya'
'Lituania' 'Luxemburgo' 'Macedonia' 'Madagascar' 'Malawi' 'Malasia'
'Mali' 'Malta' 'Mauritania' 'Mauritius' 'México' 'Moldova' 'Mongolia'
'Montenegro' 'Marruecos' 'Mozambique' 'Myanmar' 'Namibia' 'Nepal']
```

```
'Países Bajos' 'Nueva Zelanda' 'Nicaragua' 'Nigeria' 'North Cyprus'
'Noruega' 'Oman' 'Pakistan' 'Palestinian Territories' 'Panamá' 'Paraguay'
'Perú' 'Filipinas' 'Polonia' 'Portugal' 'Qatar' 'Rumania' 'Rusia'
'Ruanda' 'Arabia Saudita' 'Senegal' 'Serbia' 'Sierra Leona' 'Singapur'
'Eslovaquia' 'Eslovenia' 'Somalia' 'Somaliland region' 'Sudáfrica'
'Corea del Sur' 'South Sudan' 'España' 'Sri Lanka' 'Sudan' 'Surinam'
'Swaziland' 'Suecia' 'Suiza' 'Syria' 'Taiwán' 'Tajikistan' 'Tanzania'
'Tailandia' 'Togo' 'Trinidad y Tobago' 'Túnez' 'Turquía' 'Turkmenistan'
'Uganda' 'Ucrania' 'Emiratos Árabes Unidos' 'Reino Unido'
'Estados Unidos' 'Uruguay' 'Uzbekistan' 'Venezuela' 'Vietnam' 'Yemen'
'Zambia' 'Zimbabue']
```

```
[23]: # Accedemos a la información de la felicidad del primer y último año para los
      ↪ dos países de nuestra elección
pais1 = 'Bélgica'
pais2 = 'Australia'
```

```
[24]: # Filtrar datos para el primer año y los países seleccionados
felicidad_pais1_primer_anio_data = data[(data['anio'] == primer_anio) &
      ↪ (data['pais'] == pais1)]['escalera_vida']

if len(felicidad_pais1_primer_anio_data) > 0:
    felicidad_pais1_primer_anio = felicidad_pais1_primer_anio_data.values[0]
else:
    print(f"No se encontraron datos para el país {pais1} en el primer año.")
```

```
[25]: # Filtrar datos para el primer año y los países seleccionados
felicidad_pais1_primer_anio = data[(data['anio'] == primer_anio) &
      ↪ (data['pais'] == pais1)]['escalera_vida'].values[0]
felicidad_pais2_primer_anio = data[(data['anio'] == primer_anio) &
      ↪ (data['pais'] == pais2)]['escalera_vida'].values[0]
```

```
[26]: # Filtrar datos para el último año y los países seleccionados
felicidad_pais1_ultimo_anio = data[(data['anio'] == ultimo_anio) &
      ↪ (data['pais'] == pais1)]['escalera_vida'].values[0]
felicidad_pais2_ultimo_anio = data[(data['anio'] == ultimo_anio) &
      ↪ (data['pais'] == pais2)]['escalera_vida'].values[0]
```

```
[27]: # Identificar los niveles de felicidad de los países seleccionados
print(f"Felicidad en {pais1} en el primer año (2005):")
      ↪ {felicidad_pais1_primer_anio}")
print(f"Felicidad en {pais2} en el primer año (2005):")
      ↪ {felicidad_pais2_primer_anio}")
print(f"Felicidad en {pais1} en el último año (2018):")
      ↪ {felicidad_pais1_ultimo_anio}")
print(f"Felicidad en {pais2} en el último año (2018):")
      ↪ {felicidad_pais2_ultimo_anio}")
```

Felicidad en Bélgica en el primer año (2005): 7.262290477752685
Felicidad en Australia en el primer año (2005): 7.340688228607178
Felicidad en Bélgica en el último año (2018): 6.892171859741211
Felicidad en Australia en el último año (2018): 7.176993370056152

```
[28]: def comparar_felicidad(pais, felicidad_primer_anio, felicidad_ultimo_anio,
    ↪media_primer_anio, media_ultimo_anio):
    comparacion_primer_anio = f"{pais} {'fue' if felicidad_primer_anio >
    ↪media_primer_anio else 'es menos'} feliz que el promedio del resto del mundo
    ↪en el primer año (2005).\"
    comparacion_ultimo_anio = f"{pais} {'fue' if felicidad_ultimo_anio >
    ↪media_ultimo_anio else 'fue menos'} feliz que el promedio del resto del
    ↪mundo en el último año (2018).\"
    return comparacion_primer_anio, comparacion_ultimo_anio

pais1 = \"BÉLGICA\"
pais2 = \"AUTRALIA\"

# Supongamos que tienes los valores de felicidad para pais1 y pais2 en los años
    ↪2005 y 2018
felicidad_pais1_primer_anio = 2.0
felicidad_pais1_ultimo_anio = 3.5

felicidad_pais2_primer_anio = 1.5
felicidad_pais2_ultimo_anio = 3.0

comparacion_pais1_primer_anio, comparacion_pais1_ultimo_anio =
    ↪comparar_felicidad(pais1, felicidad_pais1_primer_anio,
    ↪felicidad_pais1_ultimo_anio, media_primer_anio, media_ultimo_anio)
comparacion_pais2_primer_anio, comparacion_pais2_ultimo_anio =
    ↪comparar_felicidad(pais2, felicidad_pais2_primer_anio,
    ↪felicidad_pais2_ultimo_anio, media_primer_anio, media_ultimo_anio)

print(comparacion_pais1_primer_anio)
print(comparacion_pais1_ultimo_anio)
print(comparacion_pais2_primer_anio)
print(comparacion_pais2_ultimo_anio)
```

BÉLGICA fue feliz que el promedio del resto del mundo en el primer año (2005).
BÉLGICA fue feliz que el promedio del resto del mundo en el último año (2018).
AUTRALIA es menos feliz que el promedio del resto del mundo en el primer año (2005).
AUTRALIA fue feliz que el promedio del resto del mundo en el último año (2018).