



Campus Querétaro

Herramientas computacionales: el arte de la analítica (Gpo 101)

Semestre: Febrero-Junio de 2024

Actividad 6 - Visualización

Docente

Prof. Pedro Oscar Pérez Murueta

Equipo

Edgar Roann Santillán Bernal | A00572737

```
[5] # Carga el conjunto de datos al ambiente de Google Colab y muestra los primeros
# 6 renglones.
df = pd.read_csv(file_path)

print(df.head(6))
```

	Name \
0	10-Day Green Smoothie Cleanse
1	11/22/63: A Novel
2	12 Rules for Life: An Antidote to Chaos
3	1984 (Signet Classics)
4	5,000 Awesome Facts (About Everything!) (Natio...
5	A Dance with Dragons (A Song of Ice and Fire)

	Author	User Rating	Reviews	Price	Year	Genre
0	JJ Smith	4.7	17350	8	2016	Non Fiction
1	Stephen King	4.6	2052	22	2011	Fiction
2	Jordan B. Peterson	4.7	18979	15	2018	Non Fiction
3	George Orwell	4.7	21424	6	2017	Fiction
4	National Geographic Kids	4.8	7665	12	2019	Non Fiction
5	George R. R. Martin	4.4	12643	11	2011	Fiction

El conjunto de datos es una tabla que contiene el top 50 de los libros más vendidos por Amazon por año desde 2009 hasta 2019. Cada libro está clasificado como Ficción o No ficción.

Las variables que contiene son:

- **Name:** Nombre del libro.
- **Author:** Autor.
- **User Rating:** Calificación promedio que los usuarios asignaron al libro (1-5).
- **Reviews:** Número de reseñas.

```
+ Código + Texto Se guardaron todos los cambios
• Genre: Género literario (ficción/no ficción).
```

```
# Crea una tabla resumen con los estadísticas generales de las variables
# numéricas.

summary_table = df.describe()

print(summary_table)
```

	User Rating	Reviews	Price	Year
count	550.000000	550.000000	550.000000	550.000000
mean	4.618364	11953.281818	13.100000	2014.000000
std	0.226980	11731.132617	16.042262	3.165156
min	3.300000	37.000000	0.000000	2009.000000
25%	4.500000	4058.000000	7.000000	2011.000000
50%	4.700000	8580.000000	11.000000	2014.000000
75%	4.800000	17253.250000	16.000000	2017.000000
max	4.900000	87841.000000	105.000000	2019.000000

```
[7] ## ¿Cuál es el género con más publicaciones? Muéstralo en un gráfico.
# Contar el número de publicaciones por género
genre_counts = df['Genre'].value_counts()

plt.figure(figsize=(8, 6))
genre_counts.plot(kind='bar', color=['blue', 'orange'])
plt.title('Número de publicaciones por género')
plt.xlabel('Género')
plt.ylabel('Número de publicaciones')
plt.xticks(rotation=0)
plt.show()
```

Número de publicaciones por género





