import pandas as pd

import requests

from bs4 import BeautifulSoup

import matplotlib.pyplot as plt

import seaborn as sns

# Lists to store extracted data

Book\_name = []

Prices = []

Review = []

# Loop through multiple pages

for i in range(2, 5):

url = f"https://books.toscrape.com/catalogue/category/books/fiction\_10/page-{i}.html"

response = requests.get(url)

soup = BeautifulSoup(response.text, "lxml")

# Find all book containers

books = soup.find\_all("article", class\_="product\_pod")

for book in books:

# Extract title

title = book.h3.a["title"] if book.h3.a else "Unknown"

# Extract price

price\_tag = book.find("p", class\_="price\_color")

price = price\_tag.get\_text(strip=True) if price\_tag else "£0.00"

# Extract rating

rating\_tag = book.find("p", class\_="star-rating")

rating = rating\_tag["class"][1] if rating\_tag else "Zero"

# Append data

Book\_name.append(title)

Prices.append(price)

Review.append(rating)

# Store data in a Pandas DataFrame

df = pd.DataFrame({"Book Title": Book\_name, "Price": Prices, "Rating": Review})

# Convert rating text to numerical values

rating\_map = {"Zero": 0, "One": 1, "Two": 2, "Three": 3, "Four": 4, "Five": 5}

df["Rating"] = df["Rating"].map(rating\_map)

# Convert prices to float (remove £ symbol)

df["Price"] = df["Price"].str.replace(r"[^\d.]", "", regex=True).astype(float)

# Set Seaborn style

sns.set\_style("whitegrid")

#Highest Rated Books

max\_rating = df["Rating"].max() # Get highest rating

highest\_rated\_books = df[df["Rating"] == max\_rating] # Filter books with highest rating

plt.figure(figsize=(10, 5))

sns.barplot(y=highest\_rated\_books["Book Title"], x=highest\_rated\_books["Rating"], palette="Blues\_d")

plt.title(f"Highest Rated Books (Rating: {max\_rating})", fontsize=14)

plt.xlabel("Star Rating")

plt.ylabel("Book Title")

plt.show()

#Prices vs Rating

plt.figure(figsize=(8, 5))

sns.scatterplot(x=df["Rating"], y=df["Price"], hue=df["Rating"], palette="coolwarm", s=100)

plt.title("Book Prices vs Ratings", fontsize=14)

plt.xlabel("Star Rating")

plt.ylabel("Price (£)")

plt.show()

#No. of Books per Rating

plt.figure(figsize=(8, 5))

sns.countplot(x=df["Rating"], palette="viridis")

plt.title("Number of Books per Star Rating", fontsize=14)

plt.xlabel("Star Rating")

plt.ylabel("Count of Books")

plt.show()