

# Pandas Assignment: Series, DataFrames & GroupBy

## Assignment Overview

This assignment covers:

- Basic Pandas operations (Series, DataFrames, filtering, indexing, nulls)
- GroupBy operations (aggregation, multiple groupings)
- Exploratory Data Analysis (EDA) on a real-world dataset

## Dataset to Use

Dataset Name: Netflix Movies and TV Shows Dataset

Download From: <https://www.kaggle.com/datasets/shivamb/netflix-shows>

File to Use: netflix\_titles.csv

## Part 1: Basic Pandas Operations

Q1. Load and explore the dataset:

- Import necessary libraries
- Load the dataset using Pandas
- Display first and last 5 rows
- Print the shape, column names, and info summary of the dataset

Q2. Data Cleaning:

- Check for missing values and count them per column
- Drop all rows where the title or type is missing
- Convert the `date\_added` column to datetime format

Q3. Filtering and indexing:

- Show only movies released after 2015
- Filter all shows from the country India
- Select only Movie type entries with duration over 90 minutes

# Pandas Assignment: Series, DataFrames & GroupBy

Q4. Column operations:

- Create a new column called `release\_decade` based on the release\_year
- Display unique values in type, rating, and country

## Part 2: GroupBy Operations

Q5. Aggregations:

- Group by type and count the number of titles in each category
- Group by country and find the top 5 countries with the most titles
- Group by rating and find average release year for each rating type

Q6. Multiple aggregations:

- Group by release\_year and show:
  - Total number of titles
  - Count of Movies and TV Shows separately

Q7. Advanced:

- Create a new column `is\_recent` True if `release\_year` > 2015
- Group by country and count how many recent shows are there per country

## Part 3: Analysis Questions

Answer briefly:

1. What is the most common rating for TV Shows?
2. Which country has the most content on Netflix?
3. Are most Netflix titles older or recent (after 2015)?
4. Whats the average release year for Netflix Movies?

## Submission Guidelines

Submit your `.ipynb`` file via Google Classroom

## **Pandas Assignment: Series, DataFrames & GroupBy**

Keep your code clean and well-commented

Include a 2-3 line conclusion at the end summarizing your findings