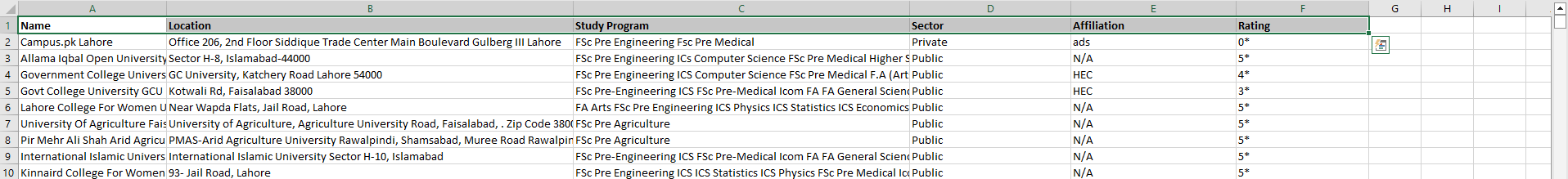
**Intermediate Colleges of Pakistan EDA**

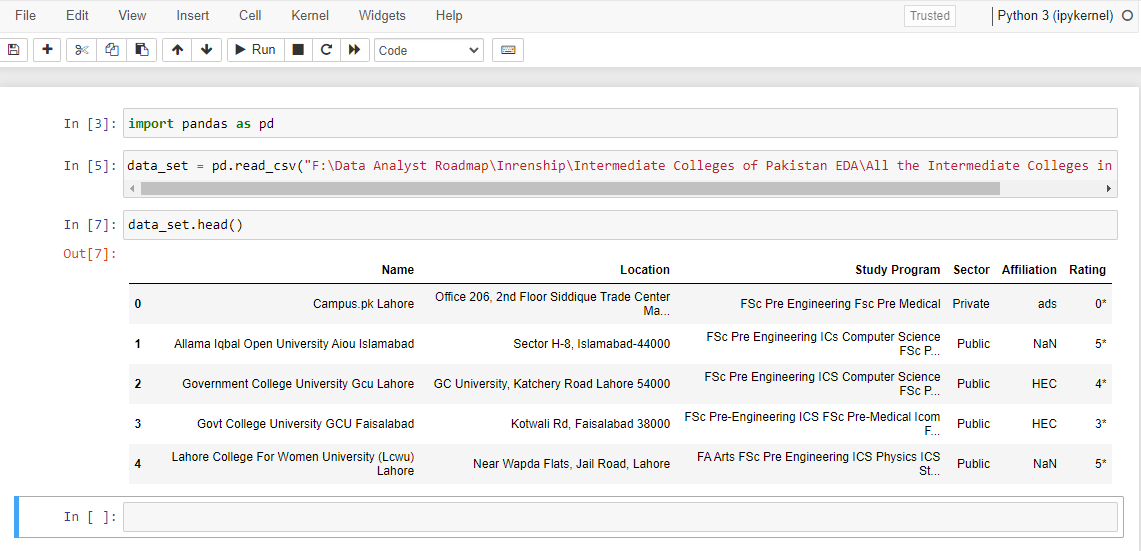
Intermediate Colleges of Pakistan EDA project entails analyzing educational data from across Pakistan to identify performance trends and regional disparities among intermediate colleges. Interns will use EDA to dissect factors such as academic results, faculty quality, and resources, aiming to highlight areas needing improvement. The analysis seeks to inform targeted interventions to bridge educational gaps between different regions, ultimately enhancing the national education standard.

**Data Collection**

First Data Collection is important to apply EDA. I will use “All the Intermediate Colleges in Pakistan” dataset to complete this project.



After collecting the data, import complete dataset to apply EDA process in ‘**Jupyter Notebook’.**

****

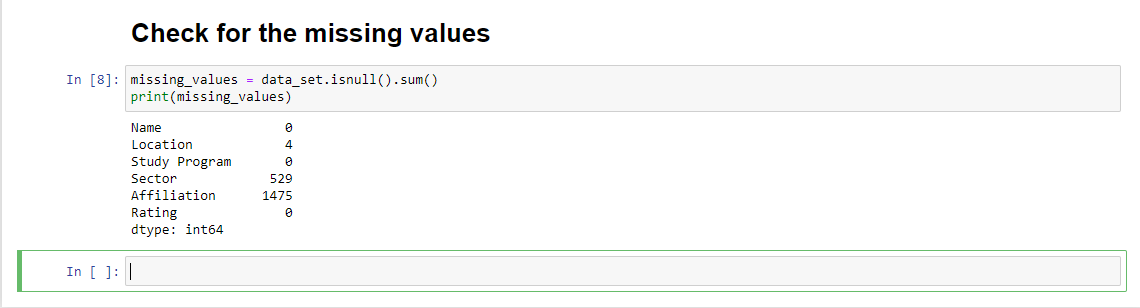
Using Pandas library import the dataset. Using “**.head()”** function show starting 5 values of each column (Name, Location, Study Program, Sector, Affiliation, & Rating).

**Data Cleaning:**

 Handle Missing Values

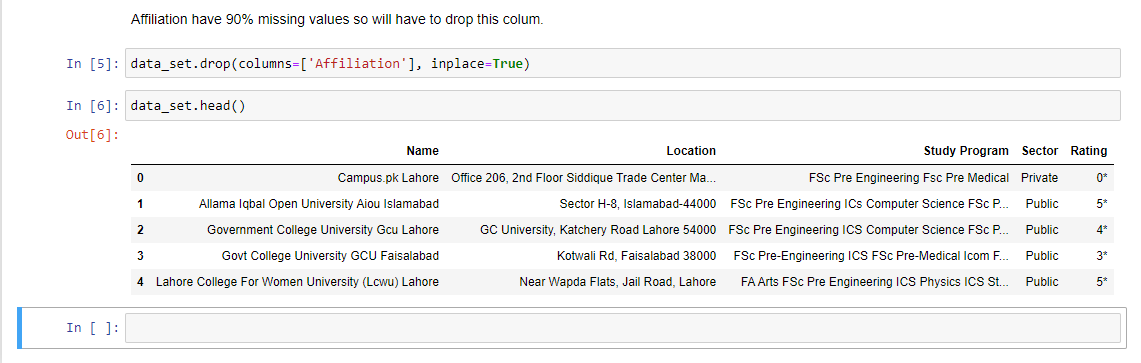
 Correct Inconsistencies

 Verify Data Types

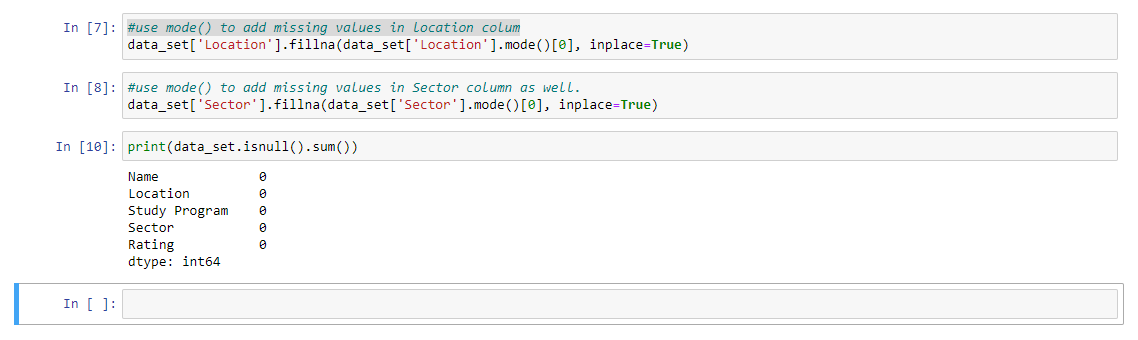


**Removing or Filling Missing Data**

The **Affiliation** column has missing values in **90.49%** of the data, it might not provide meaningful information. You can consider **removing this column**, as filling in such a large number of missing values may lead to unreliable results.

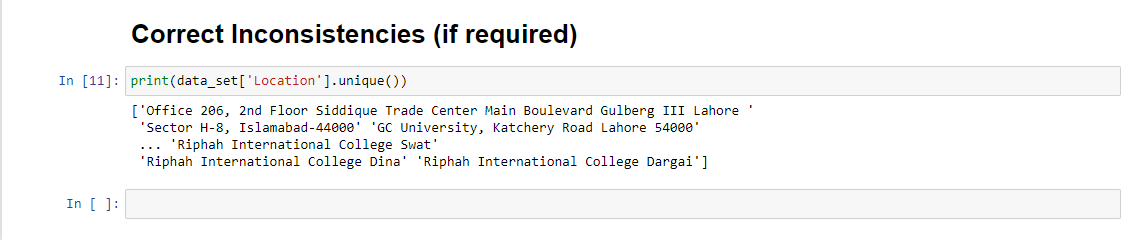


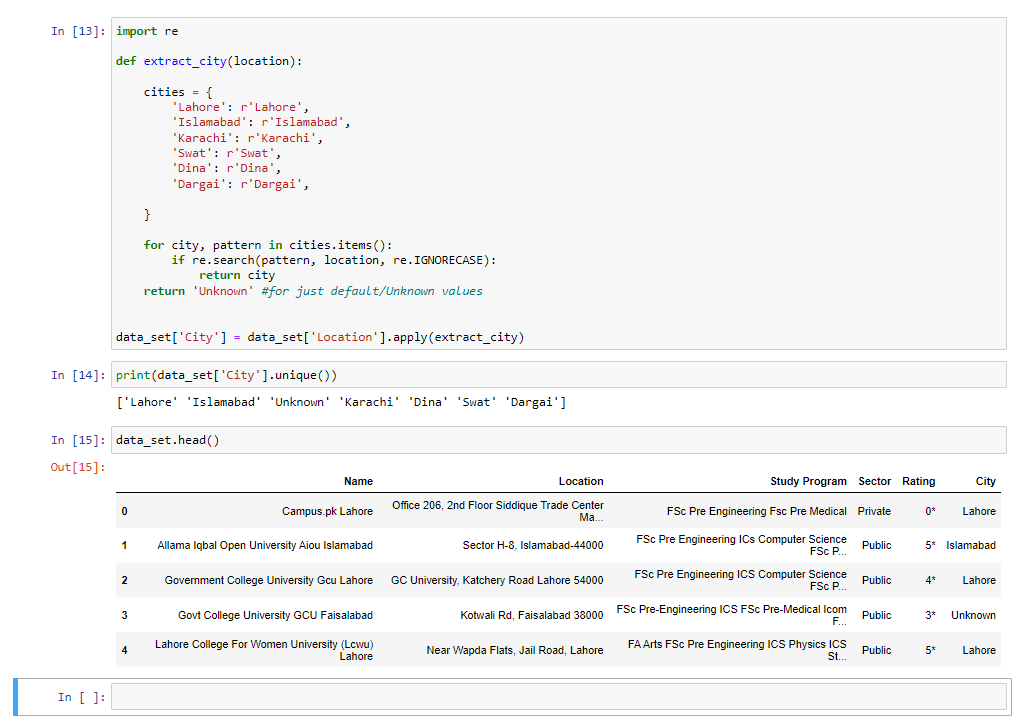
In ‘location’ column, the number of missing values is very small (0.25%) so fill the missing values using the most frequent value (mode), Also apply this method on ‘sector’ column.



Now, the dataset is clean. Zero missing value in the dataset.

In **Inconsistencies** step we have a problem in location column, because **Location** column contains detailed addresses, which could make it difficult to analyze regional trends or group data effectively. That’s why we will extract city in location column and add one more column in our dataset name (city).

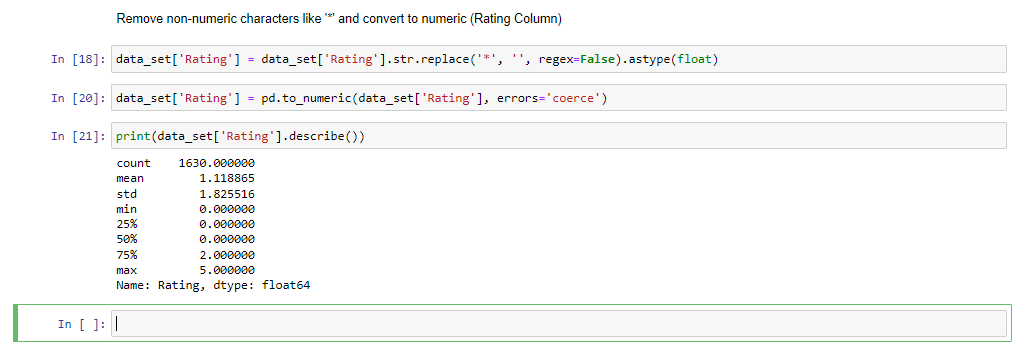


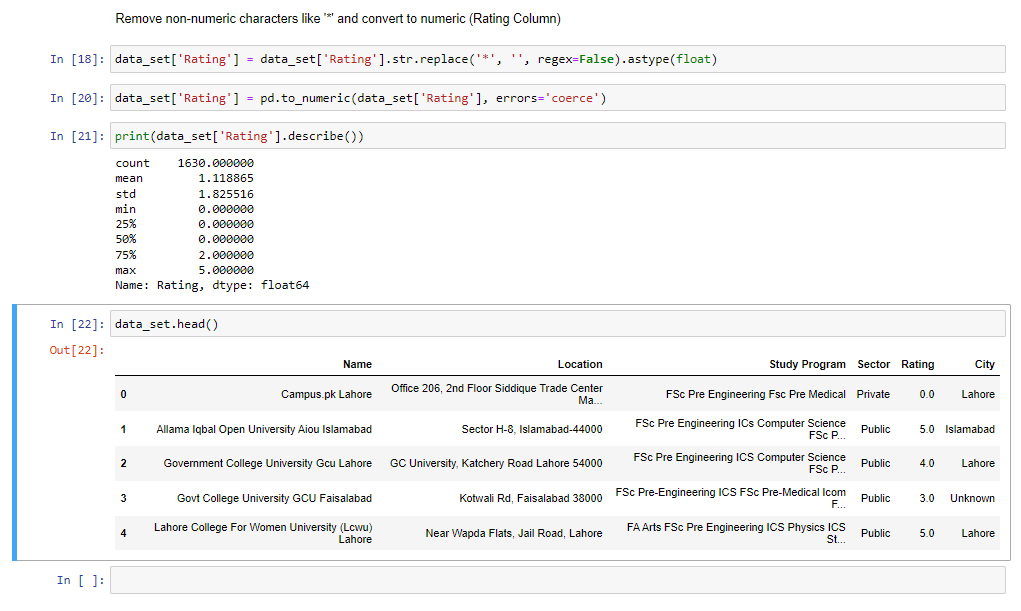


City column added in the dataset as you can see in SS.

After that we need to verify **Data Types** for all the columns.

In Rating column have non-numeric data type so it will remove it or change it numeric data type.



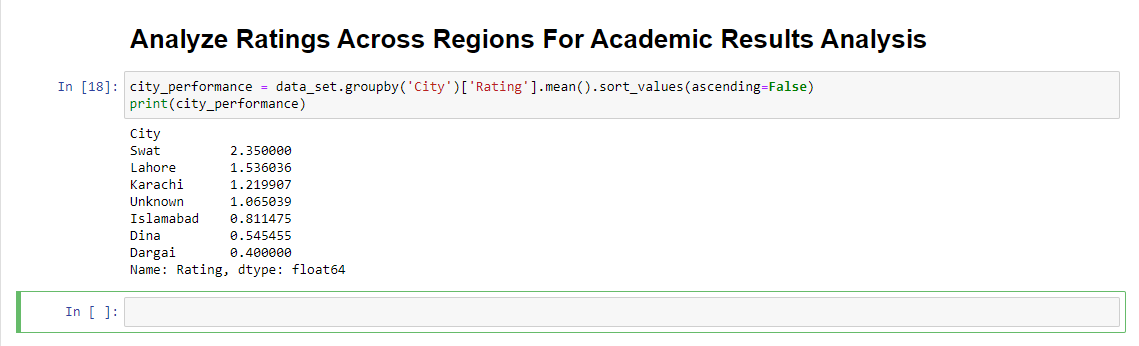


Now the dataset is completely clean, we can apply different analysis steps to find meaning dull insights and provide recommendations to do better Education System in Our Colleges.

**In-depth Exploratory Data Analysis (EDA)**

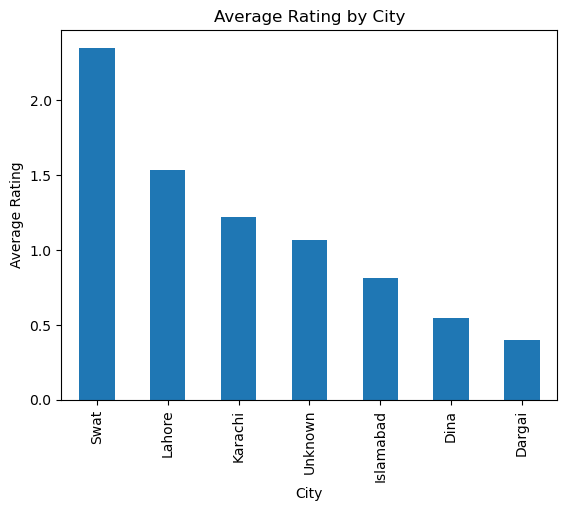
**Academic Results Analysis**

Analyzing ratings across regions using group by method.



After that, Visualize the Distribution:

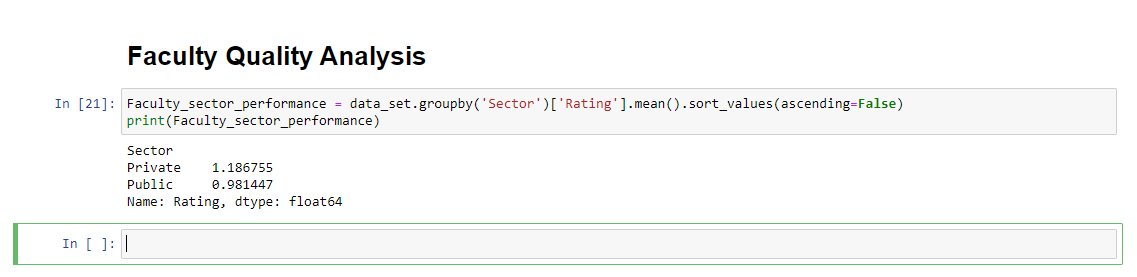
Using Matplotlib library



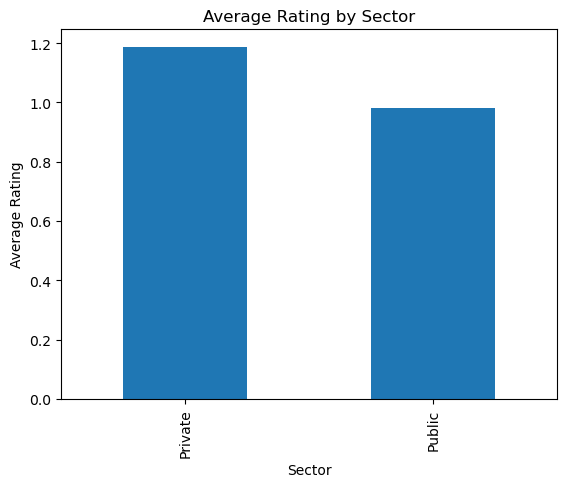
The analysis shows significant regional disparities in ratings. Swat has the highest average rating above 2.0, indicating relatively better performance. In contrast, Dargai has the lowest average rating at 0.4, suggesting potential areas for improvement.

**Faculty Quality Analysis**

In the given dataset, we don't have direct data on faculty quality. Use Rating and Sector column to find the insights.



Visualize the performance:



The analysis of ratings by sector reveals a slight difference between private and public institutions. Private sector institutions have an average rating of approximately 1.2, while public sector institutions have a slightly lower average rating of 1.0. This suggests that private institutions may have better faculty quality or other factors contributing to higher ratings, though the difference is modest.

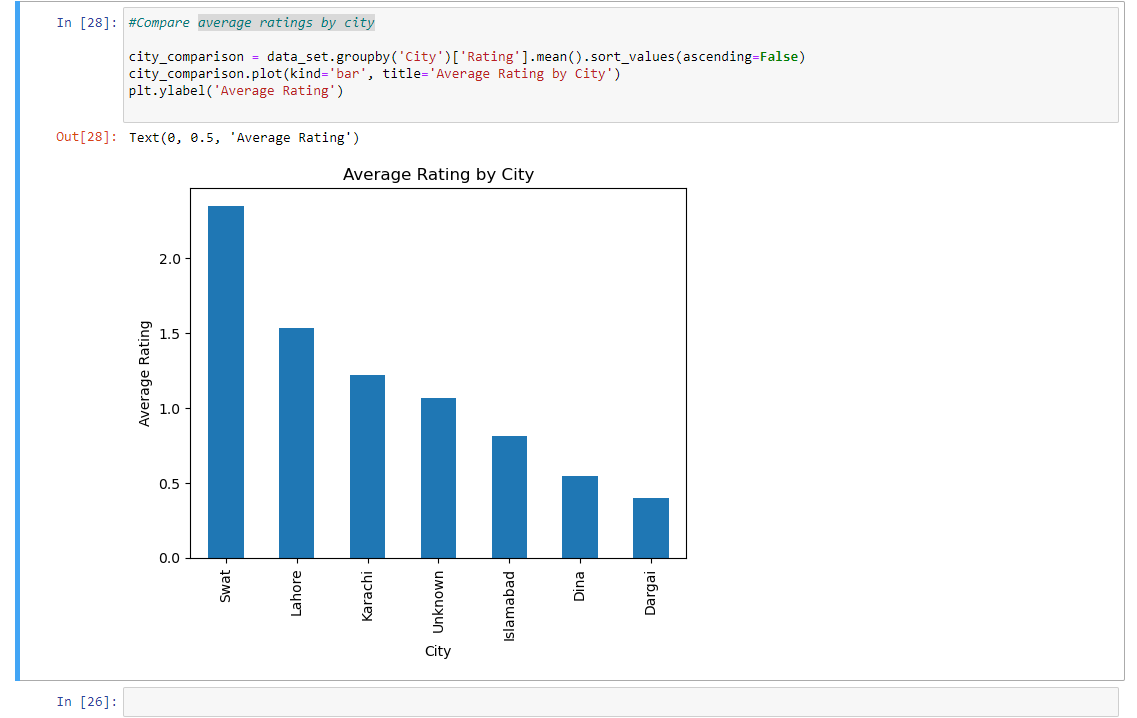
**Resource Allocation**

While direct data on resources was not available, the analysis of sector-based ratings provides some insights. Private sector institutions have a slightly higher average rating **(1.2)** compared to public sector institutions **(1.0).** This could suggest that private institutions might have better resources or more efficient resource allocation, which may contribute to their slightly higher performance.

**Regional Disparity Analysis**

I will compare and highlight disparities across different regions (cities) in terms of performance (Rating).

**Compare Average Ratings by City**

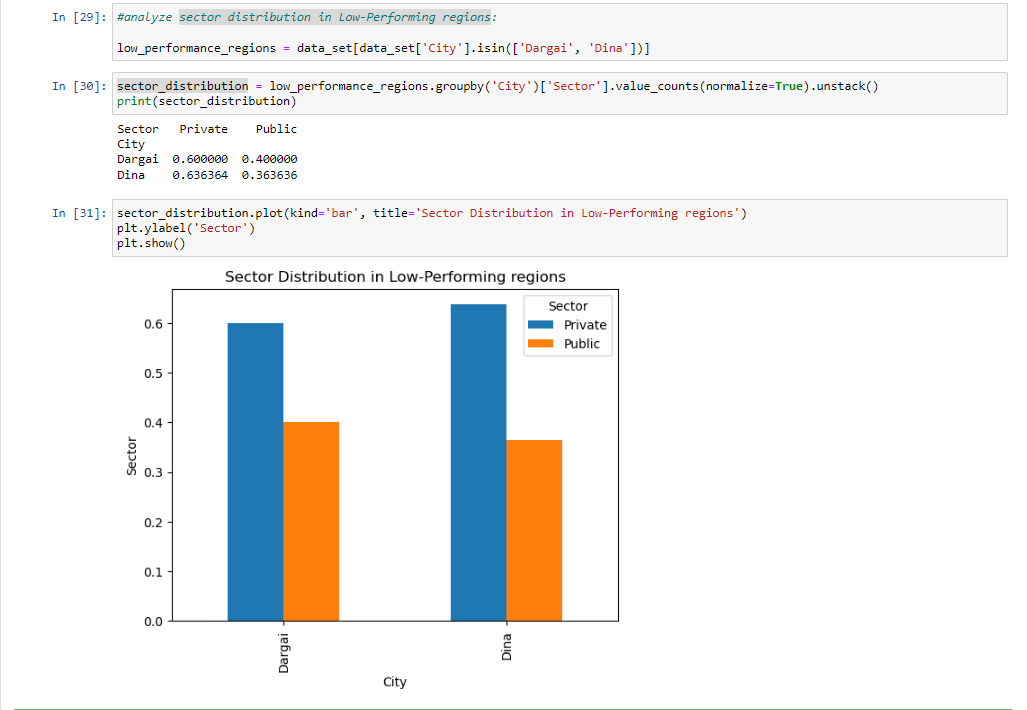


The analysis shows significant disparities in ratings across regions. Swat leads with a rating of 2.0 above, while Dargai is at the bottom with a rating of 0.5. Major cities like Lahore (1.5) and Karachi (1.3) perform moderately, but Islamabad (0.9) and smaller towns like Dina (0.6) and Dargai (0.5) lag behind. These findings suggest a need for targeted interventions in lower-performing regions.

**Identify Gaps**

Analyze Sector Distribution in Low-Performing Regions

Focus on cities with the lowest ratings, such as Dargai and Dina, and also check the distribution of sectors (Public vs. Private).



The analysis shows that Dargai and Dina, which have the lowest ratings, are primarily served by private sector institutions 60% in Dargai and 63.6% in Dina.

Despite the dominance of private institutions, these regions still perform poorly, indicating that factors like resource availability or the quality of education might be affecting their performance.

**The Regional Analysis** revealed significant disparities in educational performance across Pakistan. Swat emerged as the top-performing region, while Dargai and Dina were among the lowest, despite a high concentration of private sector institutions. These findings suggest that simply being in the private sector doesn't guarantee higher performance.

**Synthesis and Recommendations**

Consolidate all the insights gathered from EDA process into a clear and comprehensive narrative about the state of intermediate education in Pakistan.

**Integrated Findings:**

* The state of intermediate education in Pakistan is complicated and varies a lot depending on the region and sector. For example, Swat is doing really well with high educational ratings, while areas like Dargai and Dina lag far behind, even though they have many private schools. This suggests that factors like the quality and availability of educational resources play a bigger role in educational success than whether a school is public or private.
* Lahore (1.5) and Karachi (1.3), both medium-sized regions, perform slightly better than Swat but still don't rank among the top cities. The small advantage private schools have over public ones in cities like Lahore and Karachi highlights how resources and management affect academic success.
* There are also noticeable differences between regions, with poorer-performing areas likely facing challenges like fewer resources, outdated facilities, or less qualified teachers, which need more attention.

**Recommendations:**

Here are some recommendations that you will have to understand for better result.

1. **Resource Allocation for Low-Performing Regions (Like: Dargai & Dina):** Increase funding for educational resources such as libraries, labs, and digital tools. This will help bridge the gap with better-performing regions like Swat.
2. **Faculty Development:** Introduce teacher training and development programs in low-performing regions, focusing on modern teaching methods and improving educational outcomes.
3. **Replication of High-Performing Practices (Like: Swat):** Investigate the strategies used in Swat and replicate them in other regions. This may include improved resource management, teacher training, or administrative practices.
4. **Public Sector Improvement:** Advocate for increased funding and better resource management for public sector institutions, particularly in underperforming areas.
5. **Private Sector Quality:** Ensure that private institutions in low-performing areas have access to adequate resources and are held to high educational standards.