

## AP EAMCET SQL Project

### Project Overview

This project aims to analyze the education system in Andhra Pradesh, India, using SQL to explore various aspects of the data. The analysis focuses on understanding the distribution of colleges and branches, fee structures, convener seat allocations, affiliations, and regional variations in educational institutions. The data is stored in a database named [Apeducation] with two key tables: [colleges] and [Branch].

### Database Structure

- **[Branch]:** Contains information about different academic branches (e.g., engineering, medicine) offered by colleges.
  - **Bid:** Unique identifier for each branch.
  - **Bcode:** Branch code.
  - **Bname:** Branch name.
- **[colleges]:** Contains information about colleges across Andhra Pradesh.
  - **Institute\_Code:** Unique code for each college.
  - **Institute\_Name:** Name of the college.
  - **Place:** Location of the college.
  - **District\_Name:** District where the college is located.
  - **Region:** Geographical region of the college.
  - **College\_Type:** Type of college (e.g., government, private).
  - **Minority:** Whether the college is a minority institution.
  - **Co\_Educ:** Whether the college is co-educational.
  - **Affiliated\_To:** University or body the college is affiliated to.
  - **Branch\_Code:** Foreign key referring to Bcode from the [Branch] table.
  - **Fee:** Fee structure for the branch.
  - **Convener\_Seats:** Number of seats available through the convener quota.

### Objectives

This project seeks to address the following questions through SQL queries:

1. **College Distribution:** Understand how colleges are distributed across different districts and regions of Andhra Pradesh.
2. **Branch Analysis:** Identify which academic branches are offered by the most colleges and in which regions.
3. **Fee and Seat Allocation:** Explore fee structures and seat availability across various branches and colleges.
4. **Affiliation Trends:** Analyze the affiliations of colleges and their distribution by university.
5. **Regional Insights:** Understand the differences between urban and rural areas in terms of college distribution.
6. **Time Trends (if applicable):** Analyze the growth trends of colleges and branches over time, provided that historical data is available.
7. **Minority and Co-educational Institutions:** Identify minority institutions and co-educational trends across regions.

### SQL Queries

This project leverages SQL queries to perform data analysis. Some key queries include:

- **College and Branch Analysis:**

- How many colleges are there in each district?
- Which branches are offered by the most number of colleges?
- **Fee and Seat Analysis:**
  - What is the average fee for each branch across all colleges?
  - Which colleges have the highest and lowest fee structures?
- **Affiliation and Educational Quality:**
  - What are the top universities with the most affiliated colleges?
- **Regional Trends:**
  - What are the differences in number of colleges and branches across regions?

For detailed SQL queries, refer to the queries.sql file.

## **Data Quality and Integrity**

During the project, we also check for data quality issues such as missing or inconsistent data, including:

- Branch codes that are missing or not properly linked.
- Incomplete or outdated fee structures.

SQL queries are also used to clean and validate the data to ensure meaningful analysis.

## **How to Use**

### **Prerequisites**

- **SQL Server or Any Compatible Database Engine**
- Basic knowledge of SQL

### **Step 1: Clone the Repository**

```
git clone https://github.com/yourusername/ap-education-sql-analysis.git
```

```
cd ap-education-sql-analysis
```

### **Step 2: Import the Data**

1. Import the [Apeducation] database structure and data into your SQL environment.
2. Make sure the [colleges] and [Branch] tables are correctly loaded.

### **Step 3: Run Queries**

You can explore the predefined SQL queries by running them in your SQL environment. These queries are located in the queries.sql file.

### **Step 4: Customize**

Feel free to modify the queries or add additional analysis based on your specific research needs or data insights.

## **Project Output**

The output of this project will include:

- A comprehensive analysis of the college and branch distribution in Andhra Pradesh.
- Insightful data on fee structures and seat allocations.
- Regional and affiliation-based trends.
- Identification of potential data quality issues.

### **Future Enhancements**

- **Time-Based Trends:** If more historical data becomes available, future analysis could focus on how the education system has evolved over time.
- **Student Performance Data:** Incorporating student enrollment and placement data could help analyze educational outcomes by branch or college.
- **Visualization:** Add data visualizations to better represent the regional and branch-wise trends.

### **Contributing**

If you would like to contribute to this project, please fork the repository and submit a pull request with your updates. Contributions could include new queries, optimizations, or additional insights.