



# EDUVERA

## Educational Analytics Platform

Empowering Schools, Teachers, and Parents with Data-Driven Insights

---

Instructor: Eng. Ahmed samir



EDUVERA



وزارة الاتصالات  
وتقنيهedia المعلومات



## TEAM

Manar Sayed Mohamed  
Bassel Mohamed ALI  
Sama Ahmed Mohamed  
Youssef Ahmed Ibrahim  
Mostafa Mohamed Abdelaziz

# MAIN CHAPTERS

**Chapter 1: Introduction to the Dataset**

**Chapter 2: Data Preparation & Modeling**

**Chapter 3: Exploratory Data Analysis**

**Chapter 4 : Data Visualization using Tableau**

**Chapter 5 : Recommendations & Deliverables**

1

# Introduction to the Dataset

## Project Focus:

Understanding and preparing educational data from 7 sheets.

## Why It Matters:

Enables schools, teachers, and parents to use clear, accurate insights for decision-making.

Cloud

# PROJECT OVERVIEW

- Eduvera is a platform designed to help schools, teachers, and parents work smarter together by turning school data into meaningful insights.
- Our project has two main parts: building the Eduvera website with performance tracking and collaboration tools, and developing a data analysis system that transforms raw data into actionable insights for better decision making and improved student outcomes.

# SYSTEM OVERVIEW

## Main Entities

### Schools

Stores school master data  
PK: SchoolID  
Parent table for Students & Employees

### Students

Stores student personal info  
Links to School & Enrollment  
PK: StudentID  
FK: SchoolID, EnrollmentCode

### Enrollment

Defines grade + subjects per enrollment code  
PK: EnrollmentCode  
Used to assign students to academic tracks

### StudentPerformance

Tracks academic performance per student  
,Includes: Midterm, Final  
,Assignments, Attendance  
Total Score  
,FKs: StudentID  
EnrollmentCode

## **Employee**

**Stores staff data  
(teachers, assistants)  
(office staff)**  
**PK: EmployeeID**  
**FK: SchoolID**

## **Staff Performance**

**Tracks employee performance metrics**  
**Includes: Tasks, Teaching Load, Evaluations**  
**FK: EmployeeID, SchoolID**

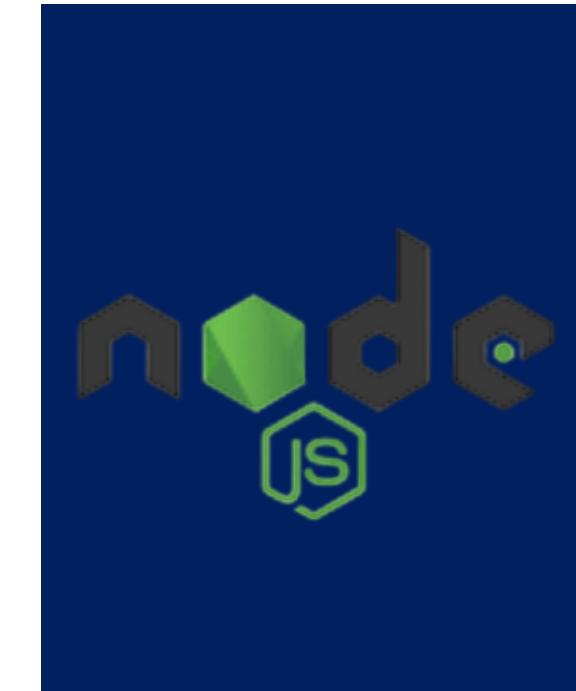
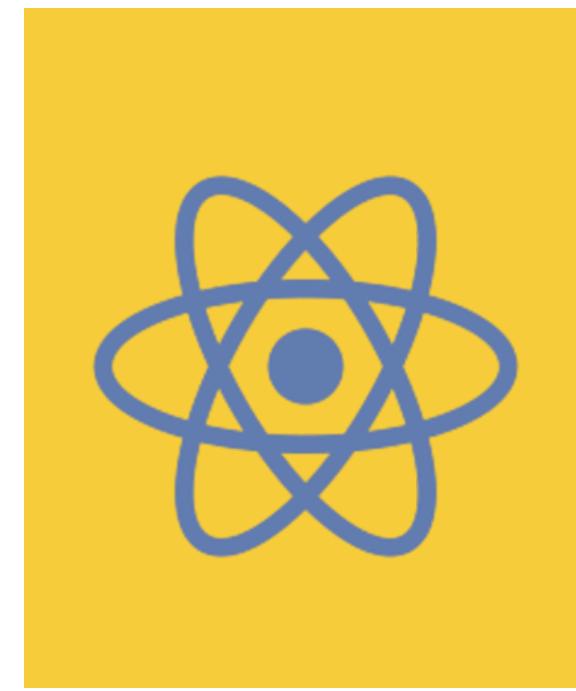
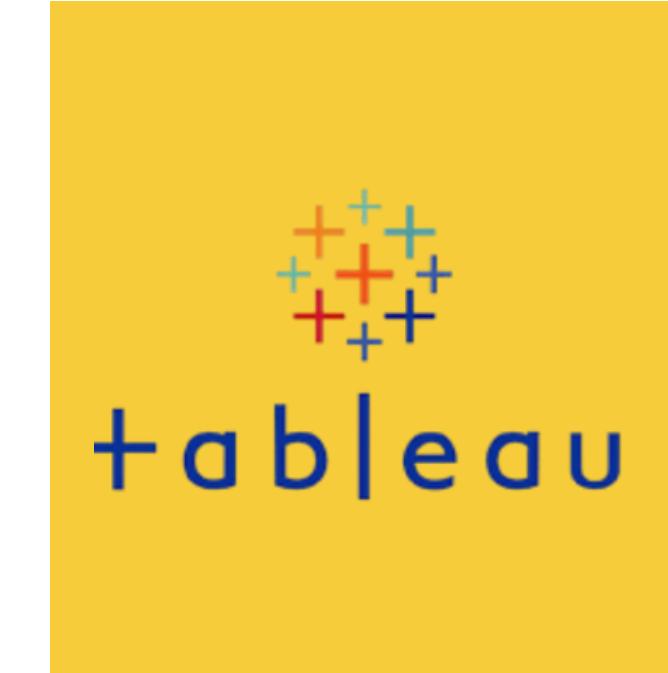
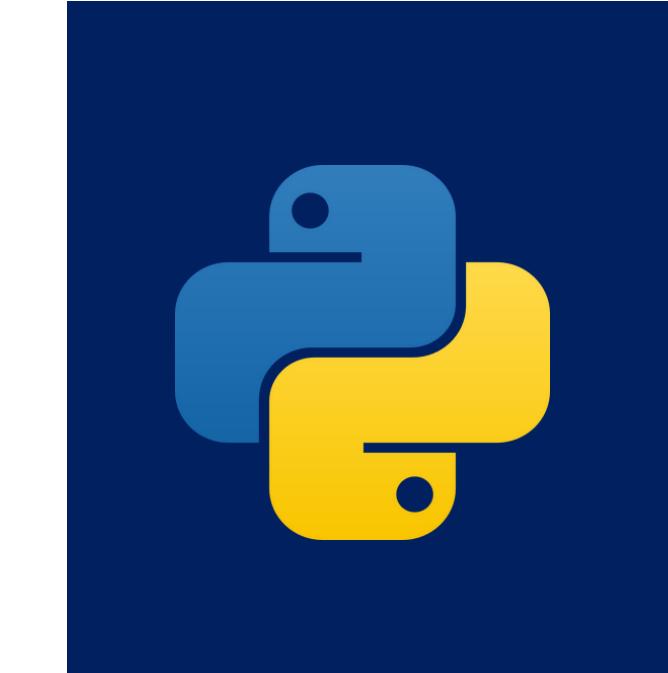
## **Student Employee**

**Maps students to teachers and assistants**  
**FKs: StudentID, EnrollmentCode, TeacherID, AssistantID**  
**Includes subject taught to each student**

# Dataset Structure Summary

The dataset represents a comprehensive educational system with fifteen schools across four governorates, serving 6,250 high school students and 692 staff members (278 teachers and 414 teaching assistants). It provides detailed academic and performance data, along with HR information, forming an integrated platform to support analysis and efficient decision making.

# Tools used





2

## Data Preparation & Modeling

### Why?

We used Data Cleaning and Wrangling to prepare the data for analysis. This step helps us remove errors and missing values, organize the data in a clear structure, and standardize formats so that comparisons are accurate. By doing this, we make sure our analysis and visualizations are reliable and meaningful.

# Data Preparation & Modeling

## Data Inspection:

- Used `info()` and `describe()` to understand structure, data types, and missing values.

## Handling Missing Values:

- Imputed missing numerical values using suitable statistics (median or mean as appropriate).
- Filled missing categorical values with placeholders like "Unknown" or "Not Provided".

## Removing Duplicates & Unnecessary Columns:

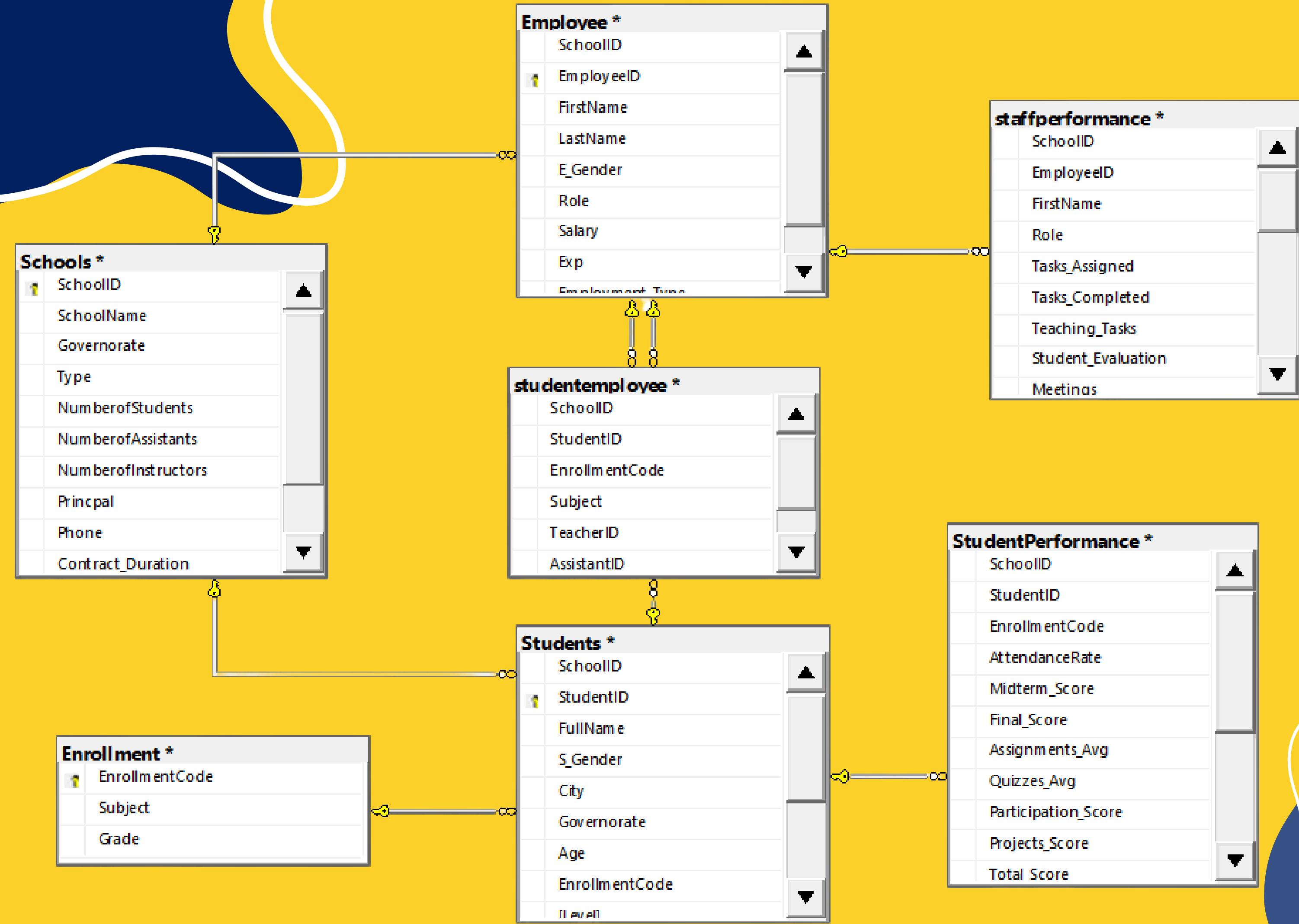
- Used Python `(drop_duplicates())` to remove redundant rows.
- Dropped columns that didn't contribute to analysis (e.g., redundant IDs, extra text fields).

## Data Formatting:

- Converted date columns into datetime format.
- Standardized capitalization and formats for names and IDs.

## Merging Sheets & Validation :

- Combined the 7 sheets (Schools, Employees, Students, Enrollment, etc.) into one structured dataset for analysis.
- Checked dataset consistency and saved cleaned CSVs for analysis and dashboard creation.





# Python Automation for Data Cleaning

## Importing & Loading Data

```
import pandas as pd
import numpy as np

df = pd.read_excel(r"C:\Users\OMEN\Desktop\Elearningsystem.xlsx")
```

## Checking Duplicates and Nulls

- Duplicates: Identify rows that appear more than once to ensure data uniqueness.
- Nulls: Detect missing values to maintain data completeness and accuracy.

```
duplicates = df[df.duplicated()]
missing_rows = df[df.isnull().any()]

print(duplicates)
print(missing_rows)
```

## Detect Outliers

- Identify unusually high or low values
- Ensures data consistency and prevents skewed analysis
- Supports accurate HR and performance insights

```
Q1 = df['Salary'].quantile(0.25)
Q3 = df['Salary'].quantile(0.75)
IQR = Q3 - Q1

lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

outliers = df[(df['Salary'] < lower_bound) | (df['Salary'] > upper_bound)]
```

## Rename Key Columns

- Standardize column names for consistency across datasets
- Remove spaces, fix capitalization, and make names meaningful

```
df.rename(columns={
    'Student ID': 'StudentID', 'School ID': 'SchoolID',
    'Full Name': 'FullName', 'Address': 'City',
    'Enrollment ID': 'EnrollmentCode', 'Track': 'Level',
}, inplace=True)
```

# Data Inspection and Missing Value Handling

- Inspecting Dataset Structure
- Handling Missing Values

```
print("Shape:", df.shape)
print(df.head())
print(df.dtypes)
print(df.isnull().sum())
df['SchoolID'] = df['SchoolID'].fillna('Not Specified')
df['FirstName'] = df['FirstName'].fillna('Not Specified')
df['Attendance_Days'] = df['Attendance_Days'].fillna(df['Attendance_Days'].mean()).round(0).astype(int)
df['Overtime_Hours'] = df['Overtime_Hours'].fillna(df['Overtime_Hours'].mean()).round(0).astype(int)
```

# Performing Key Calculations

- Gather all relevant data
- Apply the correct calculation methods
- Assign proper weights to different components

```
df['Performance_Score'] = (
    0.5 * df['Task_Completion_Rate'] + 0.3 * df['Attendance_Rate']
    + 0.1 * df['Behavior_Score'] + 0.1 * work_hours_component).round(2)
```

## Data Type Conversion in Key Columns

- Convert numeric columns and handle invalid entries.
- Transform date columns to proper datetime format.
- Standardize text columns for consistency and readability.

```
students_df['Age'] = pd.to_numeric(students_df['Age'], errors='coerce')
numeric_cols = ['Midterm_Score', 'Final_Score', 'Assignments_Avg', 'Quizzes_Avg', 'Projects_Score']

for col in numeric_cols:
    if col in student_perf_df.columns:
        student_perf_df[col] = pd.to_numeric(student_perf_df[col], errors='coerce')

schools_df["Foundation_Date"] = pd.to_datetime(schools_df.get("Foundation_Date"), errors="coerce")

cols_to_str = ['FullName', 'City', 'Governorate', 'Grade', 'Status', 'Type', 'SpecialNeeds']
for col in cols_to_str:
    if col in students_df.columns:
        students_df[col] = students_df[col].astype(str).str.title()
```

## Export Cleaned Sheets

```
df.to_excel(r"C:\Users\OMEN\Desktop\Elearningcleaned.xlsx", index=False)
```



## Chapter 3: Exploratory Data Analysis Supported by SQL Queries

- We used SQL for data analysis to extract meaningful insights from our dataset. It helps us:
- Access and retrieve data efficiently from multiple tables.
- Perform calculations like averages, counts, and summaries to evaluate student and staff performance.
- Identify trends and patterns across schools, students, and staff.
- Answer key analytical questions to support decision-making.
- Ensure accurate, structured, and actionable insights for stakeholders.

# Governorate Analysis

How many schools are in each governorate ?

```
SELECT Governorate, COUNT(SchoolID) AS Number_of_Schools  
FROM Schools GROUP BY Governorate ORDER BY Number_of_Schools DESC;
```

Results Messages

Governorate	Number_of_Schools
Alexandria	5
Cairo	5
Giza	3
Qaliobia	2

How are students distributed across governates?

```
SELECT Governorate,  
COUNT(*) AS Number_of_Students,  
ROUND((COUNT(*) * 100.0 / (SELECT COUNT(*) FROM Students)), 2) AS Percentage  
FROM Students GROUP BY Governorate  
ORDER BY Number_of_Students DESC;
```

Results Messages

Governorate	Number_of_Students	Percentage
Cairo	2020	32.32000000000000
Alexandria	1990	31.84000000000000
Giza	1280	20.48000000000000
Qaliobia	960	15.36000000000000

# School Analysis

What are the top 3 schools ranked by student performance percentage?

```
SELECT TOP 3
    s.SchoolName,
    CAST(AVG(sp.Performance_Score)
        AS DECIMAL(5,2))AS Avg_Staff_Performance
FROM Schools s JOIN StaffPerformance sp
    ON s.SchoolID = sp.SchoolID
GROUP BY s.SchoolName
ORDER BY Avg_Staff_Performance DESC;
```

SchoolName	Avg_Staff_Performance
Al Manhal	85.94
El Nile	85.10
El Tahrir	84.72

```
SELECT TOP 3 sc.SchoolName,
    ROUND(AVG(sp.Total_Score), 2)
    AS Avg_Student_Performance
FROM Students st JOIN StudentPerformance sp
    ON st.StudentID = sp.StudentID
JOIN Schools sc
    ON st.SchoolID = sc.SchoolID
GROUP BY sc.SchoolName
ORDER BY Avg_Student_Performance DESC;
```

SchoolName	Avg_Student_Performance
Sakara	85.96
El Nile	85.13
Al Manhal	84.54

What are the top 3 schools by average Employee performance score?

# Staff Analysis

What's the average experience for instructors vs TAs, broken down by gender?

```
SELECT
    Role, E_Gender,
    AVG(Exp) AS Average_Experience
FROM Employee
GROUP BY Role, E_Gender;
```

Results Messages

Role	E_Gender	Average_Experience
Instructor	F	14
Teaching Assistant	F	4
Instructor	M	16
Teaching Assistant	M	4

What's the average salary for instructors vs TAs, broken down by gender?

```
SELECT
    Role, E_Gender,
    AVG(Salary) AS Avg_Salary
FROM Employee
GROUP BY Role, E_Gender;
```

Results Messages

Role	E_Gender	Avg_Salary
Instructor	F	18709
Teaching Assistant	F	8605
Instructor	M	18287
Teaching Assistant	M	9499

# Student Analysis

How are male/female students distributed across Grades?

```
SELECT  
    Class, S_Gender,  
    COUNT(*) AS StudentCount  
FROM Students  
GROUP BY Class, S_Gender  
ORDER BY Class, S_Gender;
```

Class	S_Gender	StudentCount
Grade 10	F	618
Grade 10	M	919
Grade 11	F	636
Grade 11	M	1014
Grade 12	F	1284
Grade 12	M	1779

How many students fall into each status category?

```
SELECT  
    Status,  
    COUNT(*) AS StudentCount  
FROM Students  
GROUP BY Status
```

Status	StudentCount
Pending	424
Active	5826

# 4 | Data visualization using tableau

We used Tableau for data visualization to transform our dataset into clear and meaningful insights. It helps us:

- Present complex data in simple, interactive visual formats.
- Identify trends and patterns more quickly through charts and dashboards.
- Compare student performance, attendance, and engagement across schools.
- Highlight key insights that support better decision-making.
- Communicate our findings effectively to teachers, staff, and stakeholders.



SABRA



Staff



Student



# of GOVS

4

# of Schools

15

# of Students

6.250

# of Instructors

278

# of Assistants

414

## Governorate Analysis Dashboard

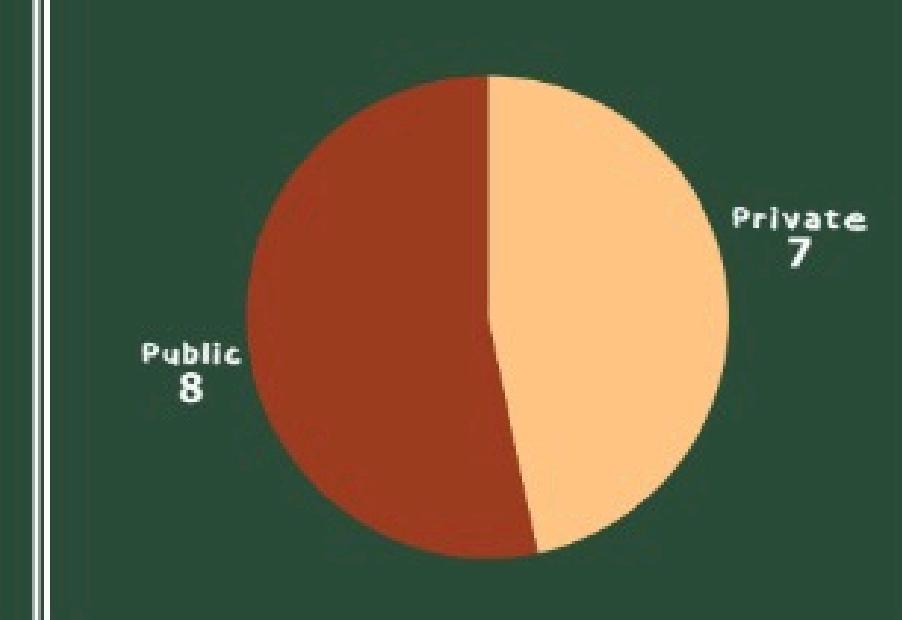
### Student Distribution



### Schools Distribution



### School Type Distribution



### Top 3 Schools



### Govs perf by Staff



### Govs perf by students



SABRA



Staff



Student



# of students

6,250

# of staff

692

# of Instructors

278

# of Assistants

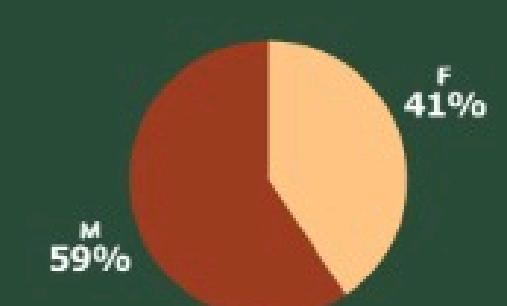
414

## School Analysis Dashboard

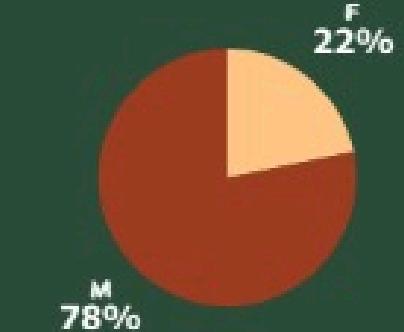
Type

- (All)
- Private
- Public

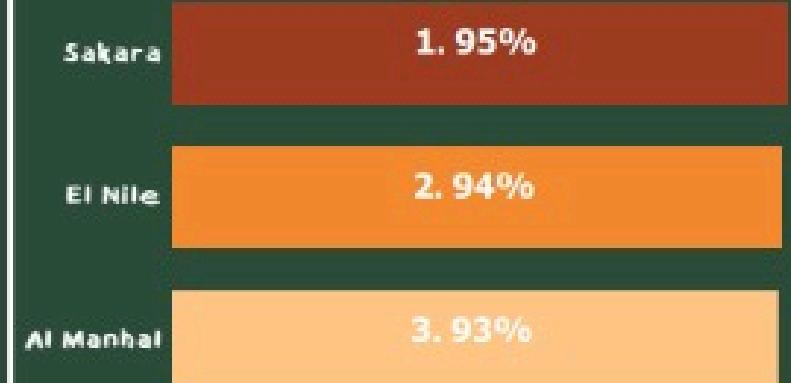
### Student Gender Distribution



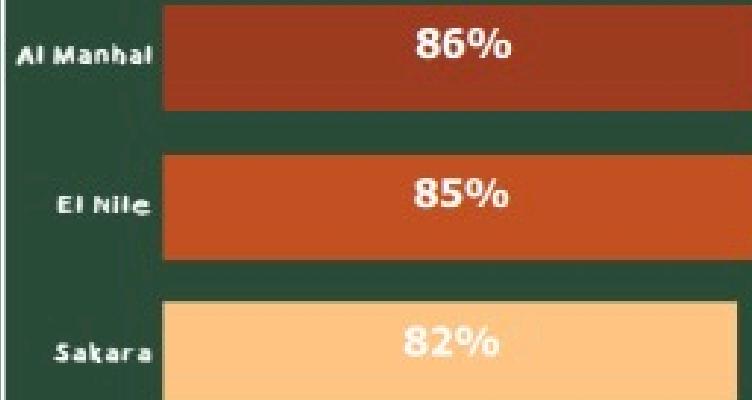
### Staff Gender Distribution



### Top 3 Schools by students perf



### Top 3 Schools by staff perf



### Staff by Role - All



### Avg Staff Exp



SABDA



Staff



Student

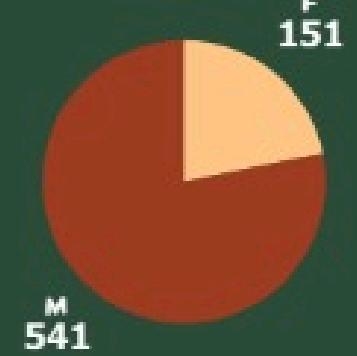


**Employment Type**  
 (All)  
 Full Time  
 Part time

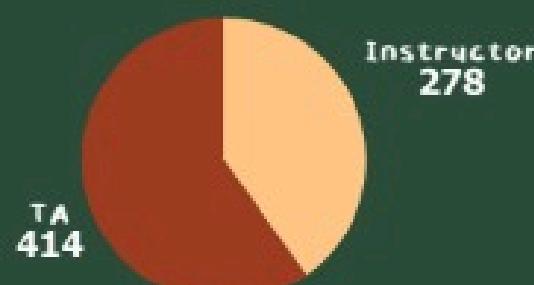
**School Type**  
 (All)  
 Private  
 Public

## Staff Analysis Dashboard

**GENDER Distribution**



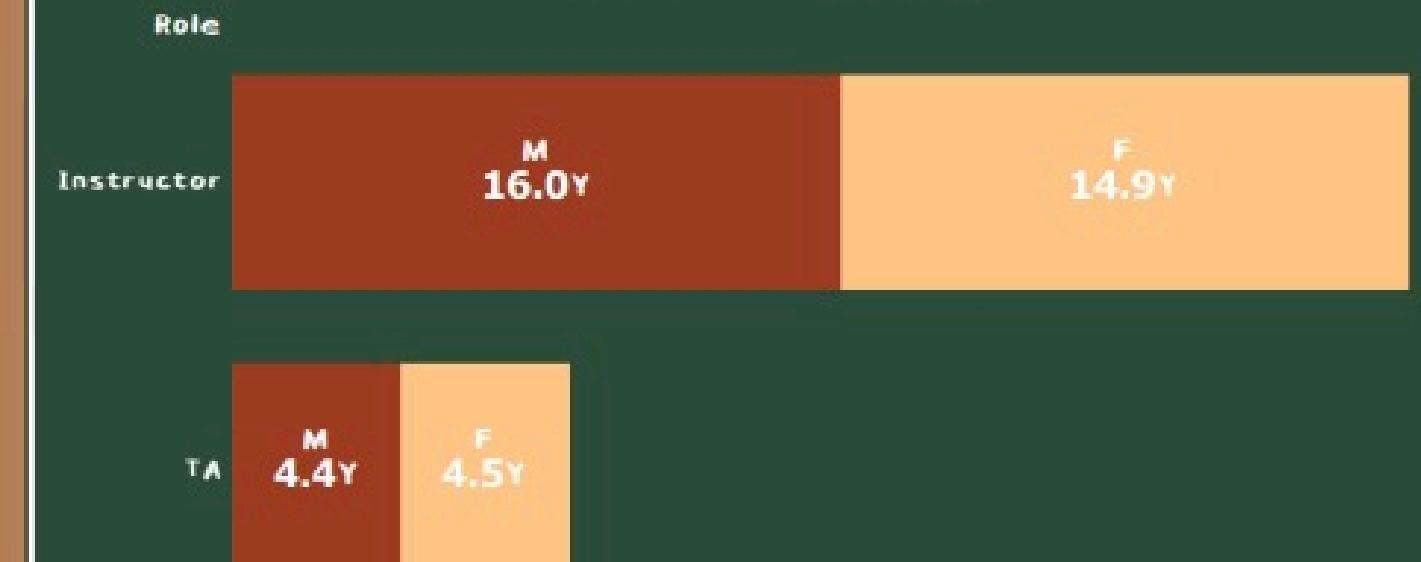
**ROLE Distribution**



**SALARY Distribution per ROLE**



**YEARS OF EXPERIENCE**



**ROLE & GENDER DISTRIBUTION**



**SALARY VS EXPERIENCE**



SABRA



Staff



Student



# of Students

6.250

# of active students

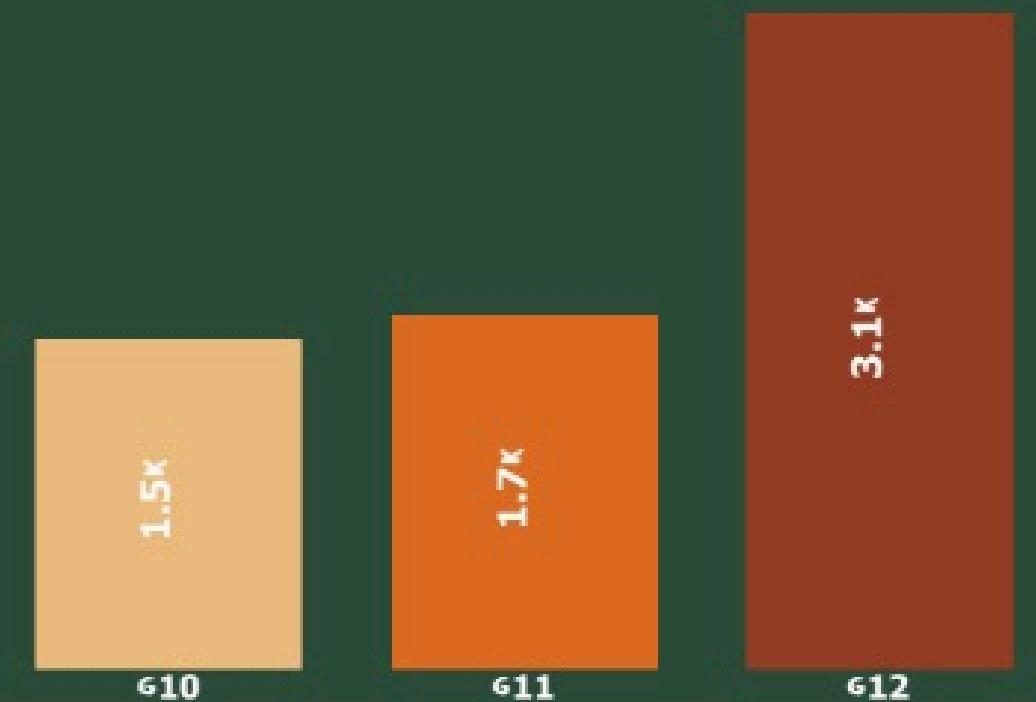
6.160

## Student Analysis Dashboard

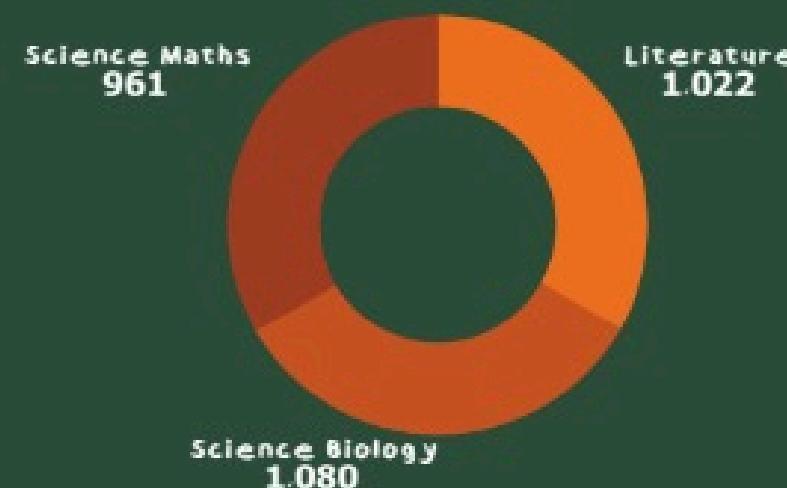
### Governorate

- (All)
- Alexandria
- Cairo
- Giza
- Qaliobia

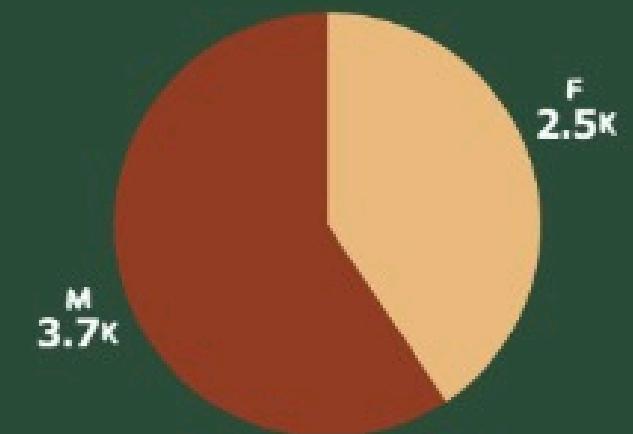
### LEVEL DISTRIBUTION



### Grade 12 CLASS DISTRIBUTION



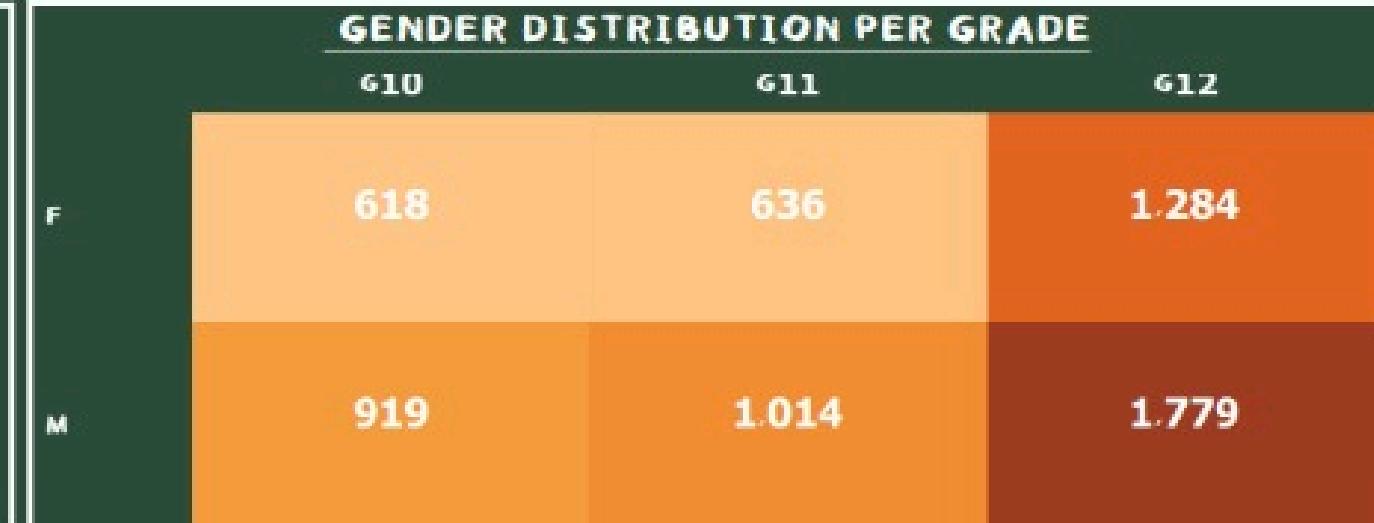
### GENDER DISTRIBUTION



### STUDENT STATUS



### GENDER DISTRIBUTION PER GRADE



SABRA



Staff



Student



#of Instructors

278

Instructor Tasks

10.266

Avg. Instructor Performance

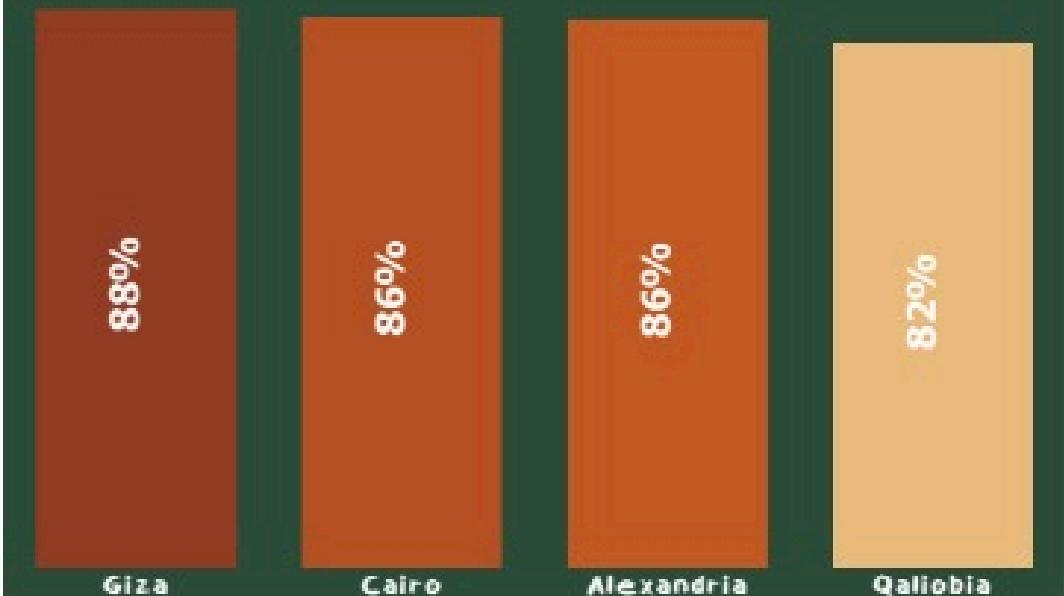
86%

## Instructor Performance Dashboard

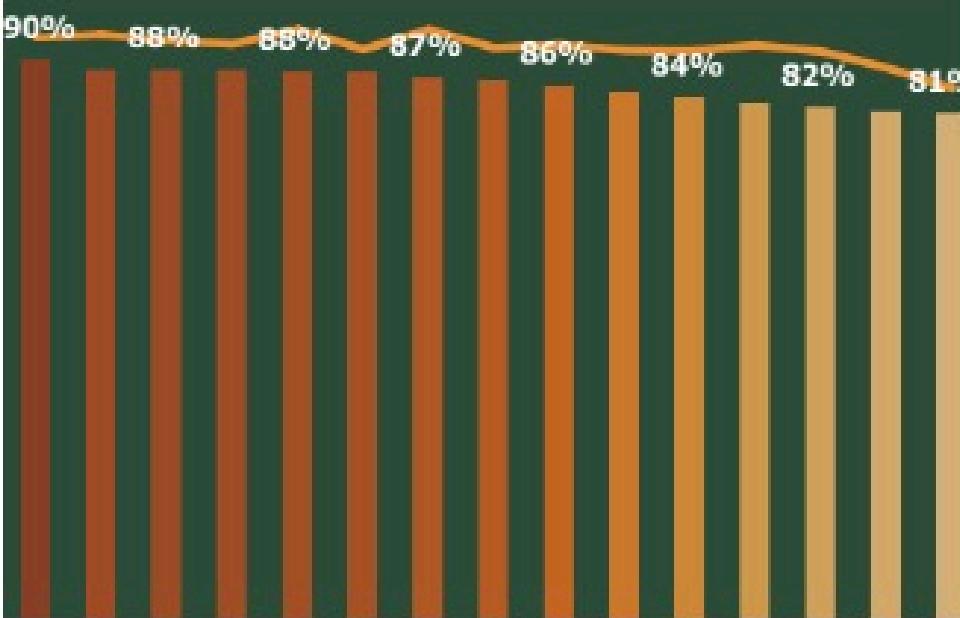
Governorate

- (All)
- Alexandria
- Cairo
- Giza
- Qaliobia

### Top Performing Instructor Per Gov



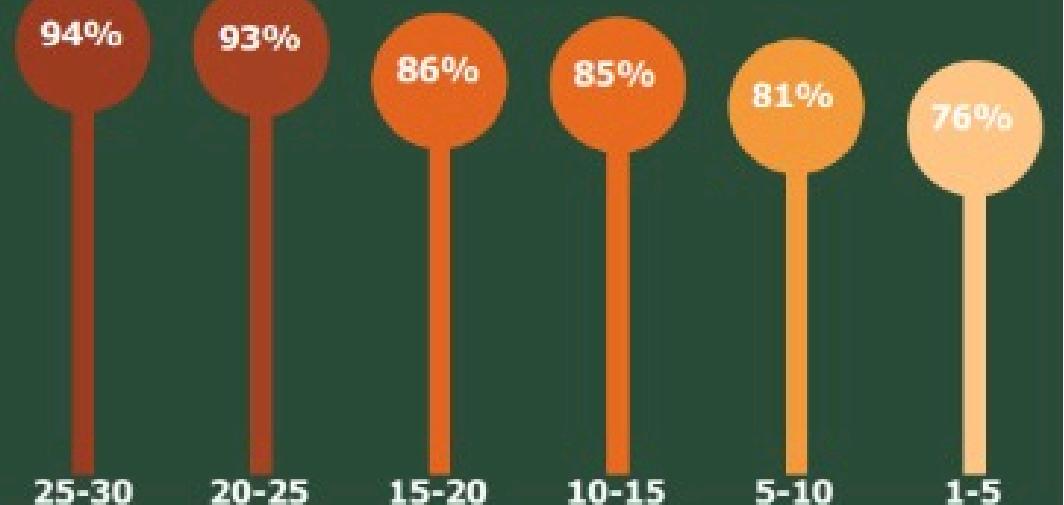
### Instructor Attendance Vs Performance



### Top performing Instructor per school



### Experience Years vs Performance



### Does Number of Tasks Affect Performance



SAKURA



# of Assistant

414

Assistants Tasks

5.658

Avg. Assistant

Performance

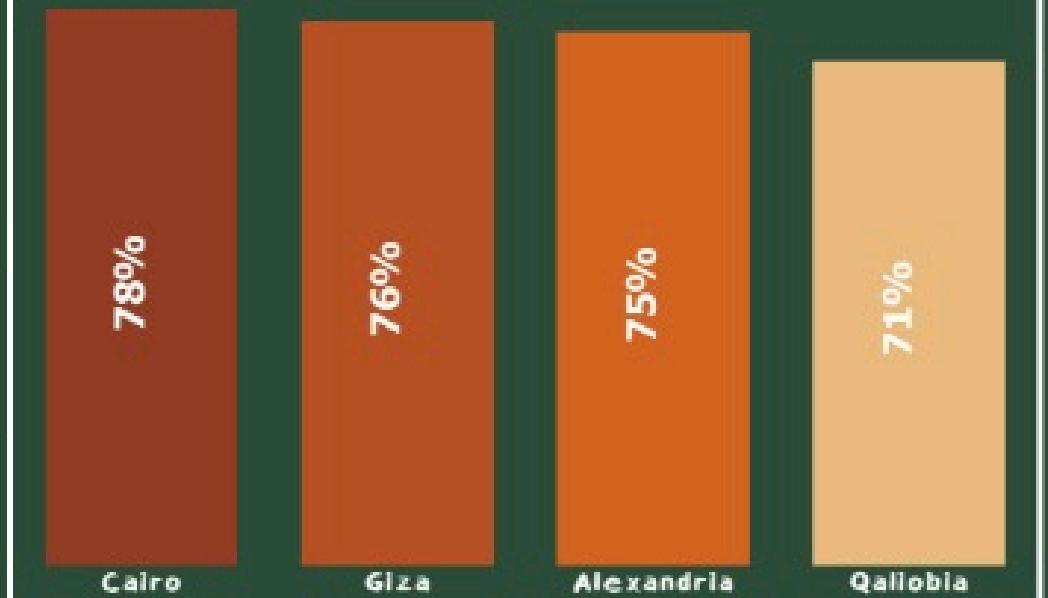
75%

## Assistant Performance Dashboard

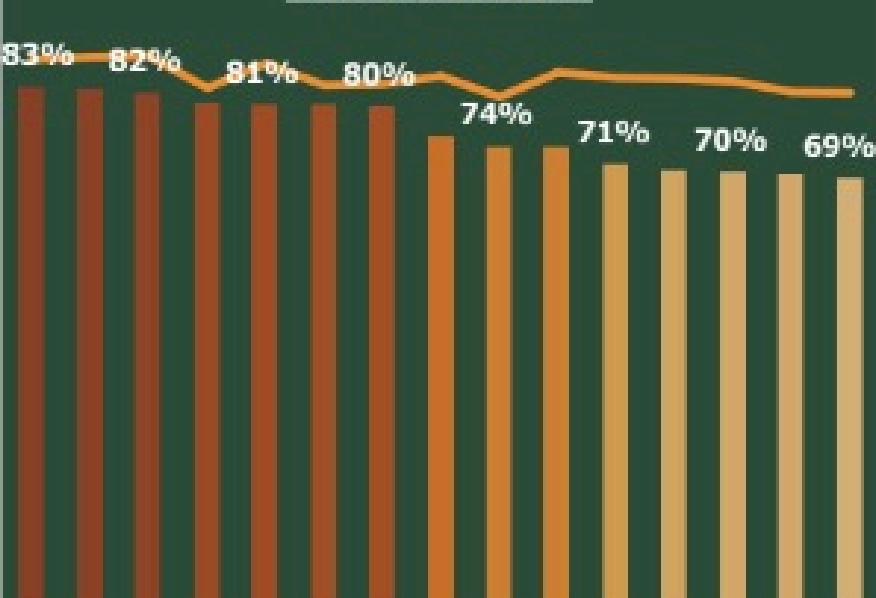
Governorate

- (All)
- Alexandria
- Cairo
- Giza
- Qaliobia

### Top Performing Assistant Per Gov



### Assistant Attendance Vs Performance



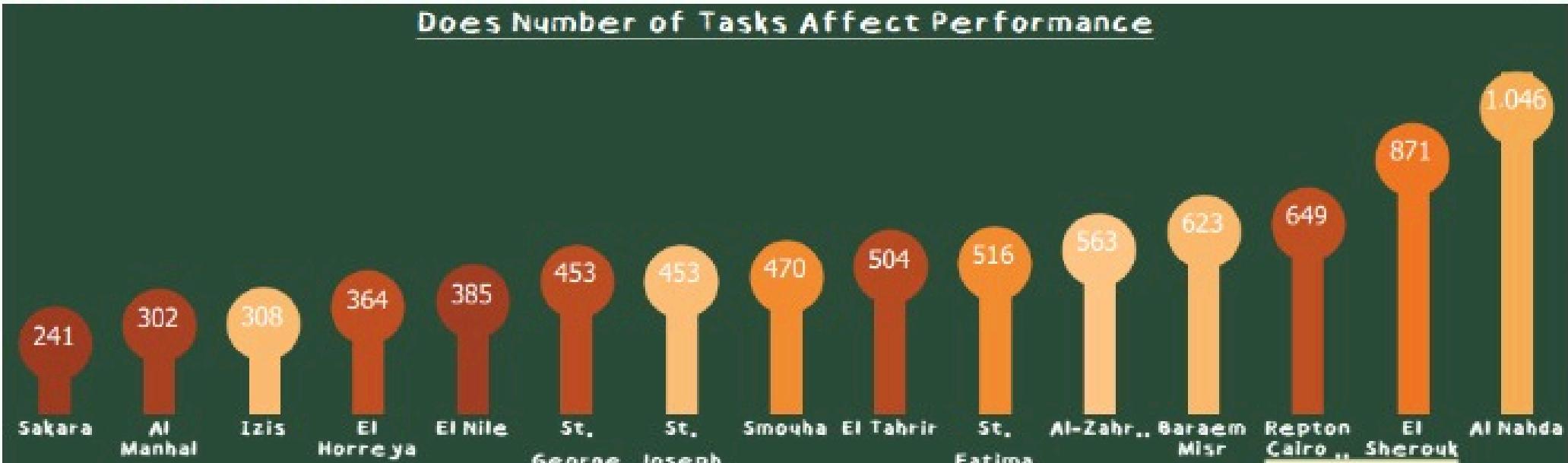
### Performance of Assistants per School



### Experience Years vs Performance



### Does Number of Tasks Affect Performance



SABOORA



Staff



Student



### Avg. Attendance Rate

81%

### Avg. Study Hours

18H

### # of Students

6.250

### Avg. Student Performance

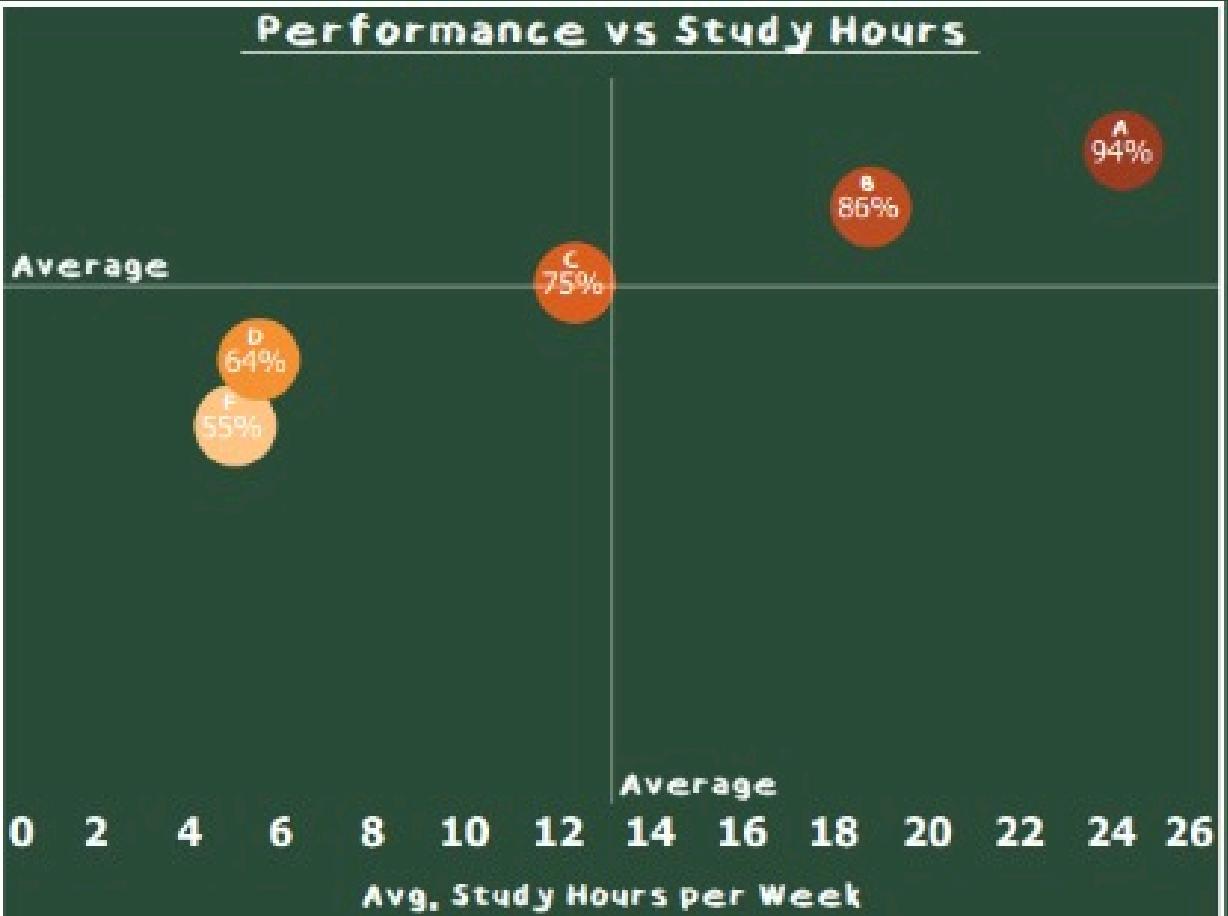
83%

### Performance vs Major



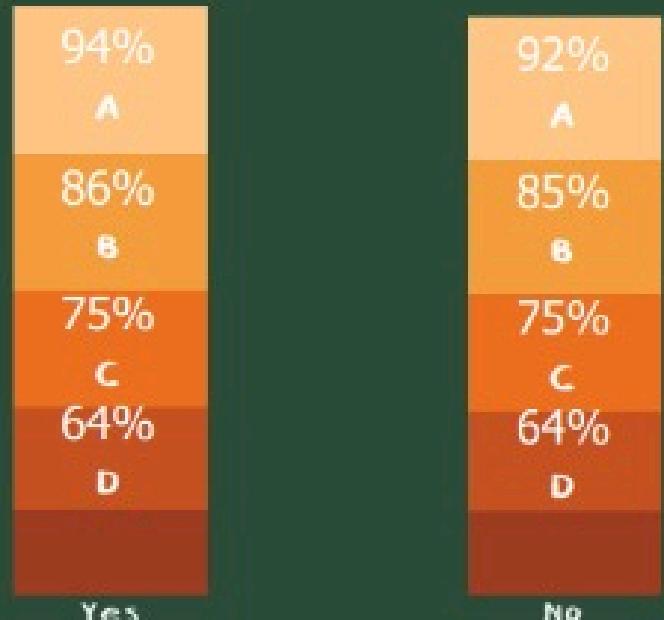
## Student Performance Dashboard

### Performance vs Study Hours

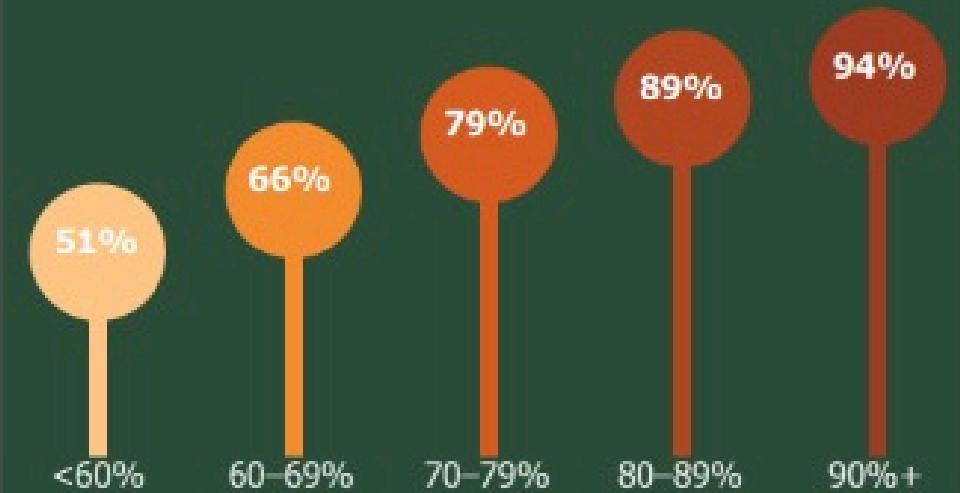


### Performance vs Internet Access

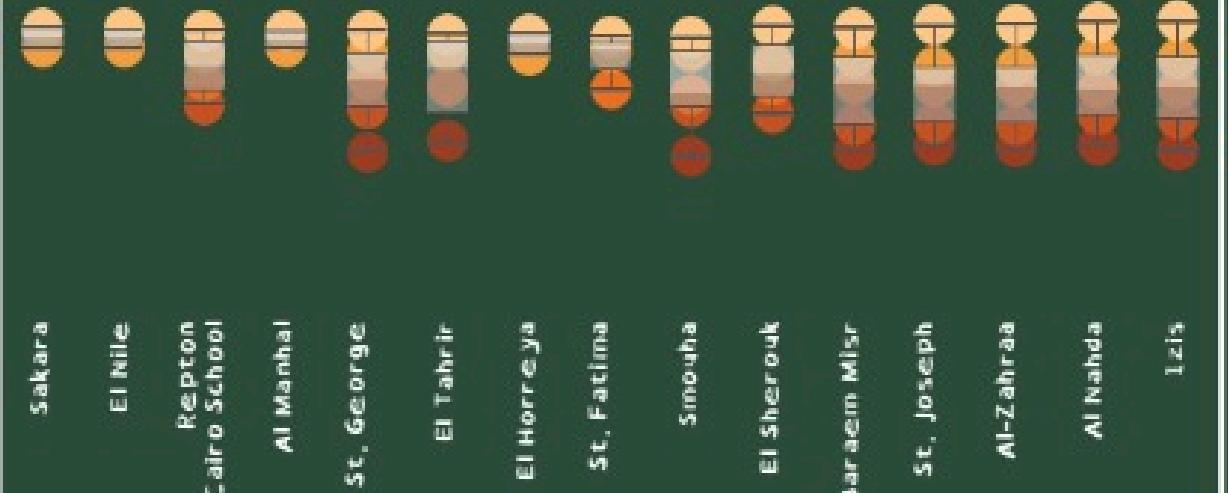
#### Internet Access at Home



### Performance vs Attendance Rate



### Distribution of Performance by School



### Performance vs Parent Education Level



# **Chapter 5:** **Recommendations** **& Deliverables**



# Actionable Insights for Enhancing Government Performance

Key Elements Driving Governorate Performance

## Strengthen Low-Performing Governorates

- Focus intervention programs on governorates with lowest staff (Qalubeya-75%) and student (Qalubeya-73%) performance
- Deploy experienced educators and additional resources to raise performance above 80% threshold

## Bridge Public-Private Education Gap

- Increase private school participation beyond current levels to provide diverse educational options
- Offer incentives for private sector investment in underserved governorates

## Replicate Effective Practices from High-Performing Schools

- Identify and analyze the practices used by the top-performing schools within the current dataset to understand the drivers behind their strong outcomes.
- Apply these proven practices across other schools in the system, especially those showing lower performance levels, to elevate overall results.

## Expand Capacity for Schools Showing High Student Concentrations

- Some schools in the current system exhibit significantly high student counts, indicating potential capacity pressure within the sampled institutions.
- Consider expanding infrastructure or redistributing enrollment within these schools to maintain optimal learning conditions and manageable class sizes.

# Actionable Insights for Enhancing School Performance

Key Elements Driving School Performance

Extreme Staff Distribution Inequality Across Schools

- Conduct comprehensive staffing needs assessment based on student enrollment per school
- Redistribute staff to ensure equitable student-to-teacher ratios across all schools

Standardize Staff-to-Student Ratios System-Wide

- Establish target ratios: 1 instructor per 22 students, 1 assistant per 15 students
- Audit current distribution and reallocate resources to meet standards in all schools
- Schools exceeding capacity should receive priority hiring or temporary staff support

Leverage Experience Gap

- Establish school-to-school mentorship partnerships pairing experienced schools with newer ones
- Create rotation/exchange programs where veteran staff spend time training at schools with less experience
- Provide intensive professional development for schools averaging under 8 years experience

Critical Shortage of Instructors Relative to Assistants

- Shift hiring priority to instructors - target 350 instructors (25% increase) over next 18 months
- Review if assistants are being forced to perform instructor duties without proper qualifications

# Operational Recommendations for Staff Improvement

Key Elements Driving Staff Performance

## Address Critical Gender Imbalance

- Only 30% of staff are female (141 out of 682 total staff)
- Launch targeted female recruitment initiative to reach minimum 40% representation within 3 years
- Focus particularly on instructor roles where gender gap is most pronounced

## Review Compensation Equity

- Instructors earn \$18.3k while TAs earn \$9.5k (nearly 2x difference)
- Conduct salary benchmarking to ensure competitive rates that retain talent across both roles
- Introduce performance bonuses tied to student outcomes and staff development

## Leverage Experienced Instructors for System-Wide Impact

- Instructors average 16+ years experience but may lack knowledge-sharing mechanisms
- Establish peer learning networks where senior instructors train newer staff
- Document best practices from veteran educators to standardize quality across all schools

## Develop Career Pathways for Teaching Assistants

- TAs have only 4.47 years average experience vs instructors' 16.07 years, indicating high turnover or lack of progression
- Create structured promotion program allowing qualified TAs to advance to instructor positions
- Provide mentorship pairing experienced instructors with TAs for skill development

# Performance Insights: Instructors & Assistants

Key Factors Driving Instructor & Assistant Performance



Average Performance improvement



Replicate Success Models from Top Schools



Accelerate Development



Excessive Task Load Destroys Performance

- Prioritize tasks and redistribute workload to balance team capacity
- Implement time management strategies and set clear performance expectations

- Establish a mentorship system pairing less experienced team members with seasoned veterans
- Reduce time to peak performance from 5 years to 2 years through competency-based training programs

- Organize cross-school workshops where top performers share best practices
- Pilot proven strategies from high-performing schools in bottom performers

- Introduce performance-based incentives for teams showing improvement
- Implement monthly monitoring with immediate support/intervention when performance drops below 70%

# Academic Improvement Recommendations for Students

- Key Factors Influencing Student Performance

Major Significantly Impacts Performance



- Investigate why Science-Biology performance lags behind - is it curriculum difficulty, teaching quality, or resource gaps?
- Provide additional support, tutoring, and teaching resources specifically for Science programs

Attendance is THE Critical Success Factor



- Implement aggressive attendance improvement initiative with target of 85%+ attendance for all students
- Identify and address barriers preventing regular attendance (transportation, health, family issues, engagement)
- Create early warning system to intervene when students miss 2+ consecutive days

Study Hours Show Diminishing Returns - Quality Over Quantity

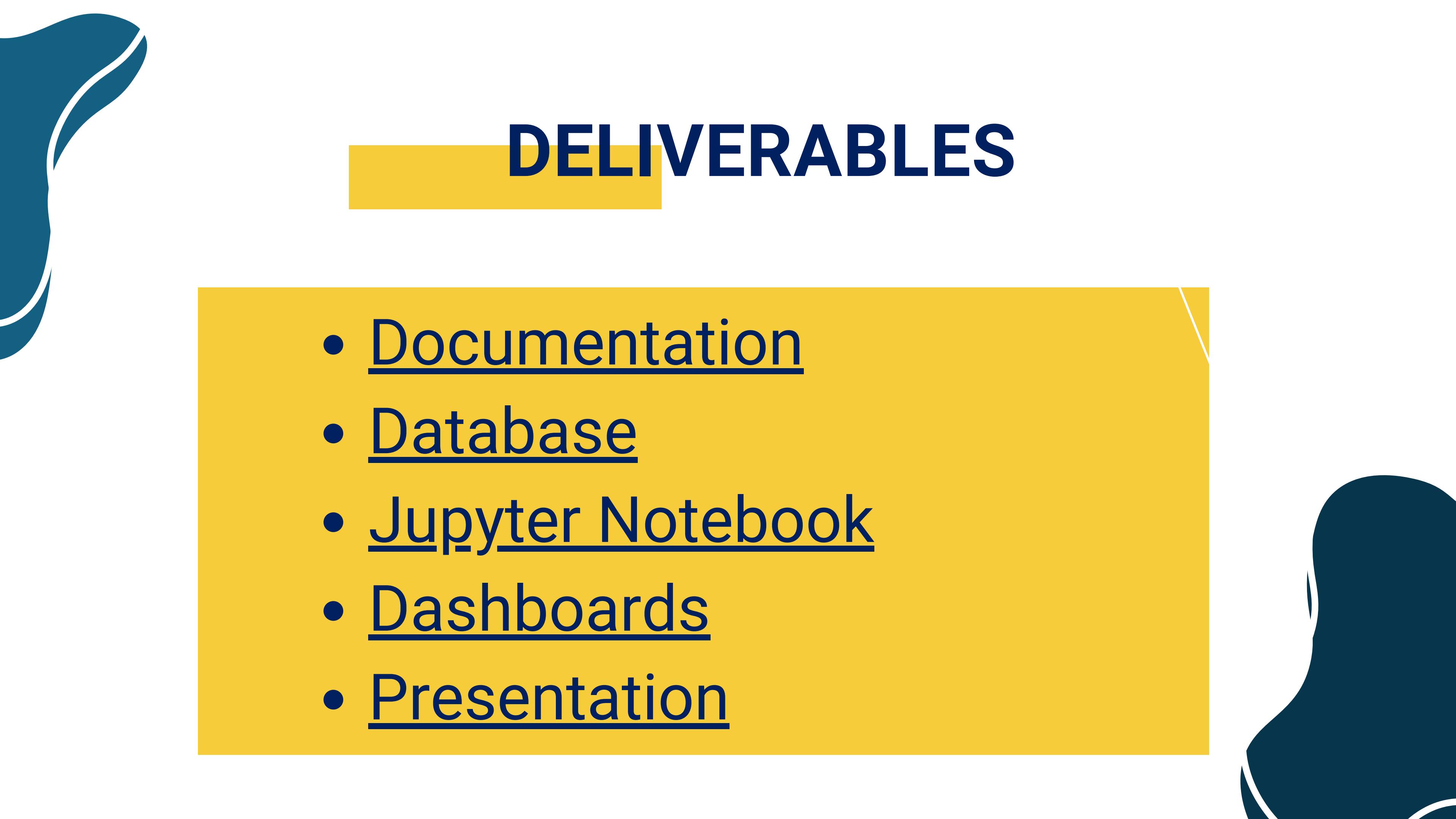


- Educate students and parents that excessive study hours (20+ hours) cause burnout and reduce effectiveness
- Promote optimal study range of 14-20 hours per week with focus on effective study techniques
- Train students in efficient learning strategies rather than encouraging marathon study sessions

Digital Divide Creates Performance Gap



- Launch digital access initiative to provide internet connectivity or alternative resources to students without home access
- Create after-school computer lab programs with extended hours for students lacking home internet
- Develop offline learning materials and resources for students with connectivity challenges



# DELIVERABLES

- Documentation
- Database
- Jupyter Notebook
- Dashboards
- Presentation

A yellow sunburst graphic consisting of eight curved, radiating shapes of varying lengths.

**THANK YOU**