**Team BAS**

**Members:  
• Tim Nguyen  
• Abdo Mohamed  
• Amaka Okalla  
• Mustafa Alghuralbawl**

**Presentation Outline**

**1. Introduction  
Presenter: Tim Nguyen**

* **Objective: Investigate the 4-year graduation rates at Texas public universities between 2020-2022.**
* **Approach: Utilize data analysis techniques to sort, analyze, and visualize graduation rates.**
* **Expected Outcome: Identify institutions with the highest and lowest graduation rates, calculate the average graduation rate, and present the findings visually.**
* **Team Member Contributions evenly divided:** 
  + **Tim Nguyen: Creation of SoW, researched and gathered presentation ideas, produced main code, managed repository, and debugged code. Created README.md file.**
  + **Abdo Mohamed: Enhanced code to display graphs with improved visuals. Handled code debugging and final presentation preparation.**
  + **Amaka Okalla: Modified the SoW, formatted project timelines, added team signatures, and implemented error handling for database and file operations.**
  + **Mustafa Alghuralbawl: Added comments to each line of code for clarity and easy debugging. Will explain the code during the presentation recording.**

**2. Scope of Work (SoW)  
Presenter: Amaka Okalla**

* **Project objectives and goals**
* **Task distribution among team members**
* **Timeline and deadlines**

**3. Environment Setup  
Presenter: Abdo Mohamed**

* **Packages to install on PyCharm interpreter:** 
  + **pandas**
  + **numpy**
  + **matplotlib**
  + **sqlalchemy**
  + **mysql-connector-python**

**4. Data Collection and Preprocessing  
Presenter: All Members**

* **Task: Import libraries, handle datasets, connect to the database, read CSV files, and store data in SQL tables.**
* **Responsibility: All members (Each member will briefly explain their role in this section)**

**5. Data Analysis and Calculation  
Presenter: All Members**

* **Task: Calculate average graduation rates, identify top and bottom institutions, and sort all institutions.**
* **Responsibility: All members (Each member will briefly explain their role in this section)**

**6. Visualization  
Presenters: Tim Nguyen and Abdo Mohamed**

* **Task: Create plots for the highest, top 3, bottom 3, and all institutions with graduation rates.** 
  + **Tim Nguyen will explain the visualization approach, while Abdo Mohamed will walk through the visual improvements.**

**7. Database Management  
Presenter: All Members**

* **Task: Manage database connections and queries.**
* **Responsibility: All members (Each member will briefly explain their role in this section)**

**8. Code Walkthrough  
Presenter: Mustafa Alghuralbawl**

* **Task: Explain the code structure and key sections:** 
  + **Data Import**
  + **Database Connection**
  + **Data Analysis Functions**
  + **Visualization/Plotting**

**9. Database Connection  
Presenter: Amaka Okalla**

* **Task: Show successful connection to MySQL database, table creation, and data insertion process.**

**10. Data Analysis and Visualization  
Presenters: Tim Nguyen and Abdo Mohamed**

* **Demonstrate the following functions:** 
  + **Top Institution**
  + **Top 3 Institutions**
  + **Bottom 3 Institutions**
  + **All Institutions Sorted by Graduation Rates**
* **Present visualizations:** 
  + **Highest Graduation Rate Institution**
  + **Top 3 Institutions**
  + **Bottom 3 Institutions**
  + **All Institutions and Average Graduation Rate**
* **Tim Nguyen will demonstrate the analysis results, while Abdo Mohamed will present the visualizations.**

**11. Code Explanation  
Presenter: Mustafa Alghuralbawl**

* **Libraries Used:** 
  + **pandas: For data handling**
  + **numpy: For numerical operations**
  + **matplotlib: For data visualization**
  + **sqlalchemy: For database connection**
* **Database Connection Process:** 
  + **Set up credentials**
  + **Handle connection errors**
* **Data Handling:** 
  + **Read CSV file**
  + **Store data in SQL table**
* **Data Analysis Functions:** 
  + **Calculate average graduation rates**
  + **Identify top and bottom institutions**
* **Plotting Functions:** 
  + **Bar plot for top institution**
  + **Bar plot for top 3 institutions**
  + **Bar plot for bottom 3 institutions**
  + **Bar plot for all institutions and average**

**12. Results and Findings  
Presenter: Tim Nguyen**

* **Top performing institutions**
* **Bottom performing institutions**
* **Average graduation rate across institutions**

**13. Conclusion  
Presenter: Abdo Mohamed**

* **Key insights**
* **Challenges faced**
* **Future improvements**

**14. Q&A Session  
Presenter: All Members**

* **Open the floor for questions**
* **Address audience queries**