**Технічне завдання: Аналіз даних шкільних учнів**

**Мета проекту:**

Проведення всебічного аналізу та візуалізації даних про шкільних учнів, щоб зрозуміти, як різні фактори впливають на їх академічні успіхи.

**Інструменти:**

* Python для обробки та аналізу даних

**Опис датасету:**

1. **Student Information:**
   * StudentID: Унікальний ідентифікатор, присвоєний кожному учню (від 1001 до 3392).
2. **Demographic Details:**
   * Age: Вік учнів, що варіюється від 15 до 18 років.
   * Gender: Стать учнів, де 0 означає чоловічу стать, а 1 означає жіночу стать.
   * Ethnicity: Етнічна приналежність учнів, закодована наступним чином:
     + 0: Кавказька раса
     + 1: Афроамериканці
     + 2: Азіати
     + 3: Інші
   * ParentalEducation: Рівень освіти батьків, закодований наступним чином:
     + 0: Відсутня
     + 1: Середня школа
     + 2: Деякий коледж
     + 3: Бакалавр
     + 4: Вища
3. **Study Habits:**
   * StudyTimeWeekly: Тижневий час, витрачений на навчання, у годинах (від 0 до 20).
   * Absences: Кількість пропусків протягом навчального року (від 0 до 30).
   * Tutoring: Наявність додаткових занять, де 0 означає відсутність, а 1 означає наявність додаткових занять.
4. **Parental Involvement:**
   * ParentalSupport: Рівень підтримки з боку батьків, закодований наступним чином:
     + 0: Відсутня
     + 1: Низька
     + 2: Помірна
     + 3: Висока
     + 4: Дуже висока
5. **Extracurricular Activities:**
   * Extracurricular: Участь у позашкільних заходах, де 0 означає відсутність участі, а 1 означає участь.
   * Sports: Участь у спортивних секціях, де 0 означає відсутність участі, а 1 означає участь.
   * Music: Участь у музичних гуртках, де 0 означає відсутність участі, а 1 означає участь.
   * Volunteering: Участь у волонтерській діяльності, де 0 означає відсутність участі, а 1 означає участь.
6. **Academic Performance:**
   * GPA: Середній бал (Grade Point Average) за шкалою від 2.0 до 4.0, що залежить від навчальних звичок, участі батьків та позашкільної діяльності.
7. **Target Variable: Grade Class:**
   * GradeClass: Класифікація оцінок учнів на основі їх середнього бала (GPA):
     + 0: 'A' (GPA >= 3.5)
     + 1: 'B' (3.0 <= GPA < 3.5)
     + 2: 'C' (2.5 <= GPA < 3.0)
     + 3: 'D' (2.0 <= GPA < 2.5)
     + 4: 'F' (GPA < 2.0)

**Завдання:**

1. **Завантаження та підготовка даних:**
   * Завантажити датасет.
   * Перевірити дані на наявність відсутніх значень та аномалій.
   * Привести дані до зручного для аналізу формату.
2. **Описова статистика:**
   * Розрахувати основні статистичні показники для всіх змінних (середнє значення, медіана, стандартне відхилення тощо).
   * Визначити розподіл учнів за віком, статтю, етнічністю та рівнем освіти батьків.
3. **Аналіз навчальних звичок:**
   * Вивчити кореляцію між навчальним часом на тиждень та кількістю пропусків.
   * Аналіз впливу додаткових занять на академічні успіхи.
4. **Аналіз участі батьків:**
   * Дослідити, як рівень підтримки батьків впливає на академічні результати учнів.
5. **Аналіз позашкільної діяльності:**
   * Вивчити вплив участі у позашкільних заходах (спортивні секції, музичні гуртки, волонтерство) на GPA.
6. **Аналіз академічних результатів:**
   * Проаналізувати розподіл GPA серед учнів.
   * Визначити основні фактори, що впливають на класифікацію оцінок (GradeClass).
7. **Висновки та рекомендації:**
   * Оформити основні висновки з аналізу даних.
   * Надати рекомендації щодо покращення навчальних успіхів учнів на основі отриманих даних.

**Результати:**

Проект повинен включати:

* Код на Python для аналізу даних.
* Документ з висновками та рекомендаціями.