### **1. Data Understanding**

Dataset ini memuat informasi tentang individu yang telah menyelesaikan gelar sarjana (*Bachelor’s degree*) dan sedang mempertimbangkan untuk melanjutkan ke program MBA. Data mencakup faktor demografis, latar belakang akademik, pengalaman kerja, faktor keuangan, dan aspirasi karir. Dataset ini bertujuan untuk menganalisis pola dan faktor utama yang mempengaruhi keputusan MBA.

#### **1.1 Informasi Dataset**

* **Sumber**: Kaggle
* **Jumlah Baris**: 10000
* **Jumlah Kolom**: 20
* **Target Variable**: Decided to Pursue MBA? (*Yes/No*)

#### **1.2 Deskripsi Variabel**

1. **Person ID**: Identifikasi unik untuk individu.
2. **Age**: Usia individu pada saat membuat keputusan MBA.
3. **Gender**: Jenis kelamin individu (Male, Female, Other).
4. **Undergraduate Major**: Jurusan saat sarjana (Engineering, Business, Arts, Science, dll.).
5. **Undergraduate GPA**: Indeks prestasi kumulatif sarjana (skala 0–4).
6. **Years of Work Experience**: Pengalaman kerja (tahun) sebelum keputusan MBA.
7. **Current Job Title**: Jabatan saat ini (Analyst, Manager, Consultant, dll.).
8. **Annual Salary (Before MBA)**: Gaji tahunan dalam USD sebelum mengambil MBA.
9. **Has Management Experience**: Apakah memiliki pengalaman manajerial (Yes/No).
10. **GRE/GMAT Score**: Skor ujian masuk MBA (GRE atau GMAT).
11. **Undergrad University Ranking**: Peringkat institusi sarjana.
12. **Entrepreneurial Interest**: Minat pada kewirausahaan (skala 1–10).
13. **Networking Importance**: Pentingnya jaringan dalam keputusan (skala 1–10).
14. **MBA Funding Source**: Sumber pendanaan MBA (Self-funded, Loan, Scholarship, Employer).
15. **Desired Post-MBA Role**: Aspirasi pekerjaan pasca-MBA (Consultant, Executive, Startup Founder, dll.).
16. **Expected Post-MBA Salary**: Ekspektasi gaji setelah menyelesaikan MBA.
17. **Location Preference (Post-MBA)**: Preferensi lokasi pasca-MBA (Domestic, International).
18. **Reason for MBA**: Alasan utama mengambil MBA (Career Growth, Skill Enhancement, Entrepreneurship, dll.).
19. **Online vs. On-Campus MBA**: Preferensi format pembelajaran MBA.
20. **Decided to Pursue MBA?**: Keputusan individu untuk melanjutkan MBA (Yes/No).

#### **1.3 Tujuan Dataset**

Dataset ini dirancang untuk:

1. Mengidentifikasi pola keputusan MBA berdasarkan faktor individu dan profesional.
2. Menganalisis pengaruh pengalaman kerja, gaji, dan minat pada kewirausahaan terhadap keputusan MBA.
3. Memahami hubungan antara aspirasi karir pasca-MBA dengan faktor akademik dan keuangan.

### **2. Penentuan Goals**

#### **2.1 Tujuan Bisnis**

Memahami faktor-faktor utama yang mempengaruhi keputusan individu untuk melanjutkan ke program MBA. Hasil analisis dapat digunakan oleh lembaga pendidikan, organisasi bisnis, atau konsultan karir untuk menyusun strategi yang lebih baik dalam menarik calon mahasiswa MBA.

#### **2.2 Tujuan Analisis**

1. **Prediksi Keputusan MBA**: Mengembangkan model prediktif untuk mengklasifikasikan apakah seseorang akan memutuskan untuk mengambil MBA berdasarkan faktor demografis, akademik, dan profesional.
2. **Analisis Faktor Pengaruh**: Mengidentifikasi variabel utama yang mempengaruhi keputusan, seperti pengalaman kerja, skor GRE/GMAT, atau alasan mengambil MBA.
3. **Segmentasi Individu**: Membuat segmentasi berdasarkan preferensi pembelajaran, pendanaan, dan aspirasi karir untuk menemukan pola yang signifikan.
4. **Rekomendasi Data-Driven**: Memberikan wawasan berbasis data bagi calon mahasiswa untuk membantu mereka membuat keputusan yang lebih baik.

#### **2.3 Langkah-Langkah Proyek**

1. Melakukan *exploratory data analysis (EDA)* untuk memahami distribusi data dan pola hubungan antar variabel.
2. Membersihkan data untuk menangani nilai hilang, outlier, dan inkonsistensi.
3. Membuat visualisasi data untuk menyajikan wawasan utama secara intuitif.
4. Melatih model prediktif (contoh: Logistic Regression, Decision Tree, Random Forest) untuk memprediksi keputusan MBA.
5. Mengevaluasi performa model menggunakan metrik seperti *accuracy*, *F1-score*, dan *AUC-ROC*.
6. Menyusun laporan akhir yang mencakup wawasan analisis, model prediktif, dan rekomendasi.