PREDICTIVE STATISTICAL MODELING OF COLLEGE ADMISSION SUCCESS RATES

1. Client Context

An educational consultancy based in the U.S. approached us to create a predictive model to assess the probability of college admission based on applicant profiles. The client worked with high school seniors applying to top 100 U.S. colleges and aimed to provide data-backed admission likelihoods for each student profile to guide application strategy.

2. Problem Statement

The consultancy needed a statistical model that could:

- Predict admission likelihood based on quantifiable factors such as GPA, SAT scores, extracurriculars, intended major, and geographic location.
- Provide a report that explains which variables contribute most significantly to admission decisions.
- Present predictions in a form understandable to both counselors and students.

3. Tools and Techniques Used

- **Software:** SPSS and Excel
- **Methods:**
    - Logistic regression analysis for binary outcome prediction (admitted vs. not admitted)
    - Correlation matrix and multicollinearity check using Variance Inflation Factor (VIF)
    - ROC curve analysis and AUC value to assess model performance
    - Chi-square tests to assess categorical variable associations (e.g., intended major vs. admission)
    - Descriptive statistics for profiling successful applicants
- **Reporting:** Professional report in APA style, including graphical dashboards for field counselors

4. Data Overview

- **Sample Size:** 2,500 historical applicant records from 2020–2023
- **Key Variables:**
    - Admission_Status (binary: 1 = admitted, 0 = rejected)
    - GPA (continuous, 0–4 scale)
    - SAT_Score (continuous, 400–1600)
    - Extracurricular_Score (ordinal, 1–5)
    - Intended_Major (categorical)
    - Geographic_Region (categorical, 5 levels)

All data was anonymized. Missing values in SAT scores (3.4%) were imputed using median scores within geographic regions.

5. Key Findings

- **Model Accuracy:** Logistic regression yielded an accuracy of **83.6%** with an AUC score of **0.89**, indicating high predictive strength.
- **Significant Predictors:**
    - GPA and SAT scores were the strongest predictors ($p < 0.001$)
    - Students applying for STEM majors had 18% lower odds of admission when controlling for GPA and scores, suggesting higher competition in these fields
    - Extracurricular involvement showed a moderate positive effect ($p = 0.03$), especially for non-STEM applicants
- **Insights Visualization:**
    - Predicted probabilities were visualized in Excel as a color-coded admissions dashboard for counselors
    - Coefficient interpretations were presented with odds ratios for practical application

6. Deliverables

- A 16-page report detailing:

    - Summary of variable selection and model testing
    - Interpretation of coefficients with real-world meaning
    - Visual decision trees created from model outputs
    - Model limitations and assumptions for non-statistical stakeholders
- Customized Excel tool for:
    - Inputting new student profiles
    - Automatically displaying admission probability with narrative recommendations

7. Client Impact

The consultancy integrated the model into their internal application advising software, enabling counselors to run real-time predictions during client meetings. As a result, students received more realistic guidance on school selection, and application success rate improved by 12% year-over-year. The consultancy used the model as a differentiator in their marketing pitch for college admission packages.