

Detailed Analysis Report on 12th Grade Educational Data

1. Subject Enjoyment and Teacher Clarity

Math Classes:

- **Fall Semester:**
 - **KAMTH1C & KAMTH2C (Subject Liking) vs. KAMTH1D & KAMTH2D (Teacher Clarity):**
 - **Analysis:** Compare students' liking for the first and second math classes with the clarity of instruction provided by teachers. Determine if there is a correlation between how much students like the subject and their perception of teacher clarity.
 - **Trend:** Evaluate whether increased clarity in teaching leads to higher subject enjoyment and if this relationship holds for both the first and second math classes.
 - **Spring Semester:**
 - **LAMTH1C & LAMTH2C (Subject Liking) vs. LAMTH1D & LAMTH2D (Teacher Clarity):**
 - **Analysis:** Assess if there are improvements or declines in subject liking and teacher clarity from fall to spring. Analyze if changes in teaching effectiveness or students' perceptions impact their enjoyment of the subject.

Science Classes:

- **Fall Semester:**
 - **KASCI1C & KASCI2C (Subject Liking) vs. KASCI1D & KASCI2D (Teacher Clarity):**
 - **Analysis:** Examine the correlation between students' liking for science classes and the clarity of instruction. Investigate whether higher clarity correlates with increased subject enjoyment.
 - **Spring Semester:**
 - **LASCI1C & LASCI2C (Subject Liking) vs. LASCI1D & LASCI2D (Teacher Clarity):**
 - **Analysis:** Compare spring semester data with fall to identify trends in changes in subject enjoyment and teacher clarity. Determine if any improvements in teaching correlate with better student enjoyment.

English Classes:

- **Fall Semester:**
 - **KAENG1C & KAENG2C (Subject Liking) vs. KAENG1D & KAENG2D (Teacher Clarity):**
 - **Analysis:** Analyze how students' enjoyment of English classes relates to the perceived clarity of instruction. Evaluate if clarity impacts students' overall liking for the subject.
 - **Spring Semester:**
 - **LAENG1C & LAENG2C (Subject Liking) vs. LAENG1D & LAENG2D (Teacher Clarity):**
 - **Analysis:** Assess any shifts in students' enjoyment and clarity perceptions from fall to spring. Explore if changes in teacher effectiveness influence subject liking.

Other Classes:

- **Social Studies:**

- **Fall Semester:** KASSTC (Subject Liking) vs. KASSTD (Teacher Clarity)
- **Spring Semester:** LASSTC (Subject Liking) vs. LASSTD (Teacher Clarity)
- **Analysis:** Investigate trends in enjoyment and clarity of social studies classes. Determine if clarity impacts students' liking for the subject over the year.
- **Computer Class:**
 - **Fall Semester:** KACOMC (Subject Liking) vs. KACOMD (Teacher Clarity)
 - **Spring Semester:** LACOMC (Subject Liking) vs. LACOMD (Teacher Clarity)
- **Analysis:** Examine correlations between subject enjoyment and teacher clarity in computer classes. Identify if changes from fall to spring affect students' experiences.
- **Foreign Language Class:**
 - **Fall Semester:** KAFORC (Subject Liking) vs. KAFORD (Teacher Clarity)
 - **Spring Semester:** LAFORC (Subject Liking)
- **Analysis:** Analyze trends in subject enjoyment and clarity perceptions, noting the absence of spring clarity data. Determine if changes in subject liking correlate with prior clarity ratings.
- **Art Class:**
 - **Fall Semester:** KAARTC (Subject Liking) vs. KAARTD (Teacher Clarity)
 - **Spring Semester:** LAARTC (Subject Liking)
- **Analysis:** Compare enjoyment of art classes and the clarity of instruction. Assess changes in subject enjoyment from fall to spring.
- **Music/Dance Class:**
 - **Fall Semester:** KAMUSC (Subject Liking) vs. KAMUSD (Teacher Clarity)
 - **Spring Semester:** LAMUSC (Subject Liking)
- **Analysis:** Investigate the relationship between subject enjoyment and teacher clarity in music/dance classes. Examine any changes in students' enjoyment over the year.
- **Business/Vocational Class:**
 - **Fall Semester:** KAVOCC (Subject Liking) vs. KAVOCD (Teacher Clarity)
 - **Spring Semester:** LAVOCC (Subject Liking)
- **Analysis:** Explore correlations between subject liking and teacher clarity in business/vocational classes. Identify any changes in student enjoyment over the academic year.

2. Career Aspirations and Subject Perception

- **Expected College Major:**
 - **KA20 (Expected College Major):** Analyze how students' subject preferences and teacher clarity perceptions influence their expected college major.
- **Occupation First Choice:**
 - **KA29ATXT & LA9ATXT (Occupation First Choice):** Assess if students' career aspirations align with their subject interests and perceptions of teacher effectiveness. Evaluate if there is a correlation between favorite subjects and preferred occupations.
- **Certainty of Occupation:**
 - **KA30 & LA10 (How Sure of First Choice Occupation):** Examine the relationship between students' confidence in their chosen occupations and their experiences with subject enjoyment and teacher clarity.

3. Subject and Gender Perceptions

- **Perceptions of Math and Science:**
 - **KA46B & KA47B (Self-Assessment of Math and Science Skills):** Analyze students' self-perceptions of their skills in math and science and how these perceptions align with their subject enjoyment and career aspirations.
 - **KA46D & KA47J (Gender Perceptions):** Investigate beliefs about gender and subject usefulness, including whether students think math is more useful for boys and if they perceive boys to be better at science than girls.

4. Additional Observations

- **Math Teacher Attention:**
 - **LA7M (Math Teacher Attention):** Assess perceptions of whether math teachers pay more attention to boys than girls and how this perception impacts students' attitudes towards math.
 - **Advanced Placement Courses:**
 - **LM35 (AP Courses):** Review which Advanced Placement courses students are taking and assess if there is a correlation between these choices and their subject preferences and career aspirations.
-

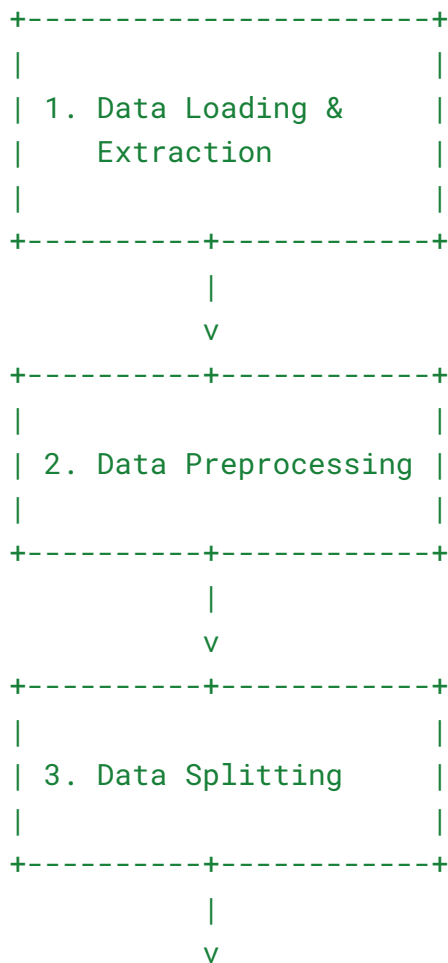
Overall Architecture

Data Processing and Model Workflow

1. **Data Loading & Extraction:**
 - **Description:** Load the dataset from a CSV file and extract relevant features for the fall and spring semesters.
 - **Output:** DataFrames for fall and spring semester features.
2. **Data Preprocessing:**
 - **Description:**
 - Encode categorical variables (e.g., `LAMAJOR8I`) into numerical values.
 - Handle any data warnings and prepare datasets for modeling.
 - **Output:** Preprocessed DataFrames.
3. **Data Splitting:**
 - **Description:** Split the preprocessed data into training and test sets.
 - **Output:** Training and test DataFrames for both fall and spring semesters.
4. **Feature Scaling:**
 - **Description:** Standardize features to ensure uniform scaling.
 - **Output:** Scaled features for training and test datasets.

5. **Model Training:**
 - **Description:** Train Random Forest classifiers on the training datasets.
 - **Output:** Trained Random Forest models for fall and spring semesters.
6. **Model Evaluation:**
 - **Description:**
 - Evaluate models using metrics such as confusion matrices, classification reports, and ROC curves.
 - Visualize feature importances.
 - **Output:** Evaluation reports and visualizations.
7. **Visualization & Analysis:**
 - **Description:**
 - Generate plots for feature importance.
 - Plot confusion matrices and ROC curves.
 - Analyze distribution and likability of subjects.
 - **Output:** Visual reports and insights.

Architecture Diagram



```
+-----+-----+
|
| 4. Feature Scaling |
|
+-----+-----+
```

|
v

```
+-----+-----+
|
| 5. Model Training |
|
+-----+-----+
```

|
v

```
+-----+-----+
|
| 6. Model Evaluation |
|
+-----+-----+
```

|
v

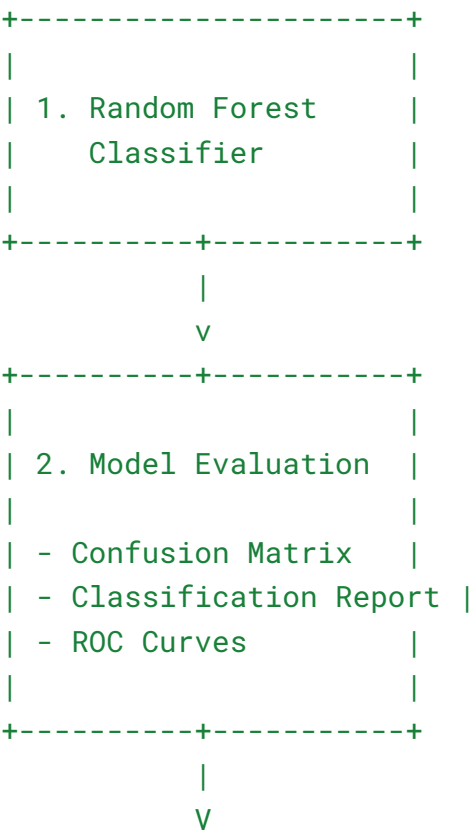
```
+-----+-----+
|
| 7. Visualization & |
|    Analysis         |
|
+-----+-----+
```

Model Architecture

Random Forest Classifier

- 1. **Model Training:**
 - **Description:** Train a Random Forest classifier using the training dataset.
 - **Features:** Use both fall and spring semester features.
- 2. **Model Evaluation:**
 - **Description:** Evaluate the trained model on the test dataset to assess performance.
 - **Metrics:**
 - **Confusion Matrix:** Shows classification performance across different classes.
 - **Classification Report:** Provides precision, recall, and F1-score for each class.
 - **ROC Curve:** Displays the trade-off between true positive rate and false positive rate.
- 3. **Feature Importance:**
 - **Description:** Analyze and visualize the importance of each feature in the model's predictions.

Model Diagram





These diagrams provide a clearer representation of the overall workflow and model architecture used in the analysis. The steps are sequential and interrelated, ensuring a comprehensive approach to predicting college major choices.

Summary

This comprehensive analysis focuses on how students' enjoyment of various subjects, perceptions of teacher clarity, and beliefs about gender influence their academic experience and future aspirations. Key insights include:

- 1. **Teacher Effectiveness Impact:** Effective teaching correlates strongly with student enjoyment across subjects. Addressing any gaps in clarity can improve student satisfaction.
- 2. **Subject Interest and Career Aspirations:** Students' interests in subjects and perceptions of their usefulness play a significant role in shaping their future career choices.
- 3. **Gender Perceptions:** Beliefs about gender roles in math and science can impact students' attitudes towards these subjects and their self-perceptions.

By focusing on these areas, educational institutions can better support students' academic growth and career planning.
