

TrainToTech: A Career Transition Story

Sanhita Saxena

20th February, 2024

1. Abstract

The lack of skills and employability among students after completing their undergraduate degrees leads them to pursue higher education or government job examinations. With a rising population and intense competition, students not only face unemployment but also end up having gap years of more than two years on their resumes, making it difficult for them to secure jobs. Many students opt for career transitions and take up courses to enter the lucrative tech job market, but often struggle with choosing the right courses and end up wasting time and money on overpriced or irrelevant programs. This issue can be addressed by leveraging machine learning models: they can help evaluate students' interests and learning patterns, providing tailored course recommendations to upskill effectively and secure jobs in fields they will enjoy and excel in.

2. Problem Statement

India's education system stands at a crossroads: while globally renowned leaders emerge from its elite institutions, a vast majority of graduates face grim prospects with limited skills and high unemployment rates. A staggering 42.3% of young graduates under 25 remain jobless, as per a report by Azim Premji University. This paradox intensifies due to the existence of substandard private colleges, devoid of qualified faculty, practical training, and job placements. Students, caught in this cycle, chase entrance exams for perceived better degrees and government jobs, often ending up with the dreaded "gap years" due to fierce competition. This reality leads to a growing chorus questioning the value of higher education, pushing them to seek alternative paths.

3. Customer Need Assessment

The target customers for this business idea are students who have had gap years in their resumes and want to transition into the tech field to upskill and secure jobs. These students often face challenges in determining the most suitable role for them and selecting appropriate courses to enhance their skills. Transitioning to a new field is

unfamiliar territory, and these students are usually on a budget and in urgent need of employment.

- Employability
- Industry-relevant skills
- Affordable upskilling courses
- Custom-tailored job role suggestions
- Career Guidance
- Best transitioning fields suitable
- Upskilling in a short amount of time
- Practical application-based courses

4. Target Specification

The proposed system aims to assist students in their career transitions by providing a comprehensive skills test that evaluates their cognitive abilities, personality type, learning style, and personal preferences. Based on the test result, the system will recommend the most suitable tech job roles for the students. Additionally, the system will offer affordable online courses tailored to their specific learning styles. The target specification of the system is to empower students with personalized guidance and resources to upskill effectively and secure jobs in the tech industry.

5. External Search

India's education system faces a stark reality: over 80% of engineering graduates lack crucial skills for employment. This "unemployability crisis" stems from a two-pronged issue: student mindset and outdated curriculum. Many students falsely believe a college degree guarantees success, neglecting necessary skill development. Further, universities prioritize attendance and rote learning over practical skills and industry exposure, leaving graduates with irrelevant knowledge. The curriculum itself fails to keep pace with

technological advancements, leaving graduates ill-equipped for high-demand fields like AI and data science.

In this vacuum, online platforms are emerging as potential solutions. These platforms offer skills assessments and upskilling courses specifically tailored to market needs, promising both speed and affordability compared to traditional education. They focus on the very skills lacking in graduates: communication, problem-solving, and adaptability. Additionally, they address the outdated curriculum by emphasizing new-age technologies in their courses.

This shift towards online platforms is not just a convenience; it is a necessity. The global job market demands dynamic individuals with relevant skills, not just degrees. Organizations prioritize performance over formal qualifications, recognizing the value of diverse talent. The rise of online platforms signifies a crucial acknowledgement: students need upskilling and reskilling to remain relevant in a rapidly evolving world. While universities must reform their curriculum and prioritize practical skills, online platforms offer a promising solution to bridge the skill gap and improve employability, paving the way for a future where talent, not outdated degrees, defines success.

6. Benchmarking Alternate Products

While we have identified competitors like MyPlan, Career Launcher, and Guild Education, a closer look reveals a unique opportunity. None of them specifically cater to gap year students seeking tech upskilling, a significant and under-served population.

Though, some offer related services, their features are limited and lack the comprehensive approach we propose. Our platform excels by:

- Focusing on a specific niche: Gap year students seeking tech careers, addressing a large, unmet need.
- Offering unique features: Holistic skills assessment, personalized tech field recommendations, and affordable upskilling course curation.
- Filling the gaps: Where competitors fall short, we provide a complete solution.

Our platform tackles a specific pain point with features tailored to a specific niche, potentially reaching a large, unemployed population currently ignored by existing platforms.

Feature	MyPlan	Career Launcher	Guild Education	TrainToTech
Target Audience	Recent graduates	Recent graduates	Established professionals	Gap year students (India)
Skills Assessment	General	General	Specific to job roles	Five-tier, cognitive, personality, learning styles, knowledge-based, personal interest
Career Recommendations	Broad fields	Specific jobs	Specific learning pathways	Personalized tech field & job roles
Course Recommendations	No	Limited options	Curated learning paths	Partnered platforms & affordable options
Monetization	Paid test + premium features	Subscription & commission	Subscription & training fees	Test fee + commission, referrals & premium features
Machine Learning	Basic recommendations	Basic job matching	Advanced skills & career pathing	Complex personalization & skill-gap analysis
Strengths	Broad audience, diverse features	Job-specific focus, connections	Established platform, learning pathways	Niche Indian market, detailed insights
Weaknesses	Generic	Limited	High cost,	New competitor,

Feature	MyPlan	Career Launcher	Guild Education	TrainToTech
	recommendations, no course options	features, basic ML	professional focus	requires user trust

7. Applicable Regulations

- Maintaining transparency and fairness
- Data security and privacy regulations
- Setting up grievance redressal mechanisms and dispute resolution
- Accessibility for all
- Consumer protection regulations
- Data transparency

8. Applicable Constraints

- Limited data and data biasness: gathering enough initial data for accurate models in the beginning can pose a challenge, one of the contributing factors for the same can be data privacy concerns.
- Interpretability: the high granularity of the skills test is going to require complex models which will be difficult to interpret and have to be mindfully constructed.
- Marketing and customer acquisition: can face initial challenges in gaining trust and the students are already in a state of urgency and might prefer more experience platforms over one that is relatively new.
- Competition: the competitors are few but have been in the game for a while, so distinguishing the product while trying to work with limited resources will need careful planning and consideration.
- Monetization strategy: careful adjustments have to be made to make the product affordable for students while also maintaining a steady revenue generation.
- Technical infrastructure and expertise requirement: will require experts and resources to maintain a robust platform while keeping the data security in check.

9. Business Opportunity and Monetization

- India, with 65% of its population under the age of 35 has 42.4% of its graduate population unemployed. The key factors contributing to this are the outdated-college curriculums and lack of industry-ready skills. To make better professional opportunities for themselves most of these students opt to appear for entrance exams for either higher education or government jobs. However, due to the increasing competition, often these students end up with no success and unaccounted gap years in their resumes. These students have untapped potential which if given the right path can join the workforce and contribute effectively to the economy of the country. With the booming tech sector in India, these students transition into fields with little to no knowledge and end up paying hefty fees and wasting valuable time on useless courses. Our platform is going to tend to the needs of this specific niche of students and help them identify the specific tech fields best suitable for them and provide them with course recommendations based on the same.
- There are multiple avenues of monetization throughout this project, as mentioned below:
 - i. Direct user payment:
 - Skills test fee: An affordable fee that will be charged for the comprehensive skills test (that will provide the best suitable job roles, strengths, weaknesses, and learning style for that specific user, along with the scope and growth in that role).
 - Subscription model: This will provide effective, recommendations and study plans with estimated timelines based on the user's specific learning style.
 - ii. Premium features (added later after assessing the growth of the platform): This will include personalized career coaching and mentorship for another additional fee.
 - iii. Course commissions:
 - Affiliate partnerships: Collaborating with promising and affordable platforms and earning commissions on recommended courses/boot camps completed by the users.

- Curated course bundles: providing custom-tailored course packages at a premium price.
- iv. Partnerships:
 - Course providers: Partner with educational institutions, and technical education companies to make this assessment test as part of their programs with a referral fee.
 - Insights and analytics: Aggregating anonymized data (with user consent) and providing insights about skills trends and talent pool.

10. Concept Generation

The idea of making a career guidance and transition app is heavily influenced by my personal experience. This app aims to empower gap-year students with an affordable, user-friendly career guidance platform. Through a comprehensive skills test and AI-powered recommendations, it helps them bridge the gap between their skills and the job market, boosting personal confidence, employability, and ultimately, the national economy.

11. Final Product Prototype

The final product is a career transition guidance app which helps graduates (with gap years) who wish to transition into the field of technology. The user is required to appear for a comprehensive skills test which will test the user on criteria like their cognitive ability, learning style, personality type, their base-level knowledge about the technology field and their interests. The results from this test will be used to suggest suitable tech job roles that the user is highly likely to excel at along with career growth. Moreover, based on the suggested job roles, the user will also be recommended affordable online courses that the user can enrol in to upskill thus providing the user a one-stop solution for all their career transition needs.

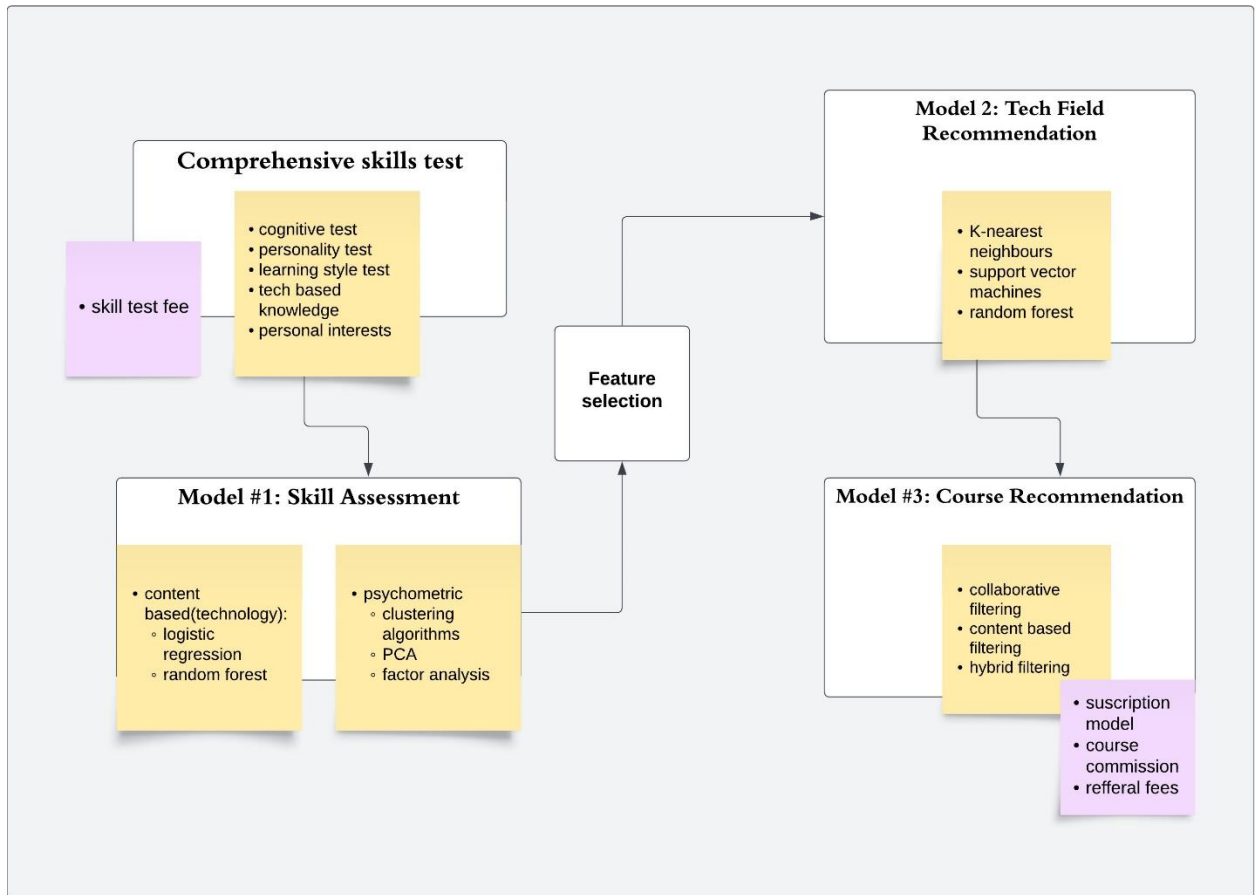


Figure 1: Schematic diagram of the final product.

12. Minimum Viable Product Features:

- Skills Assessment:
 - Questionnaire to assess cognitive capacity (numerical, verbal, logical reasoning) using validated question formats.
 - Interactive coding challenges of varying difficulty levels to evaluate technical aptitude.
 - Personality assessment based on established frameworks like the Big Five (simplified version).
 - Questions to gauge interests in specific technology areas.
 - Questionnaires or interactive tasks to identify preferred learning styles (visual, kinesthetic, auditory).
- Recommendation Engine:

- #1 Recommender system model using synthetic data or expert knowledge to map assessment results to suitable tech fields based on required skills and user preferences.
- #2 Recommender system model to suggest affordable online courses relevant to the recommended field (partnered and affordable courses).
- User Interface:
 - A user-friendly website or mobile app (both for a wider reach, taking resources availability into consideration) for taking the assessment.
 - Presents the results in a visually appealing and understandable format.
 - Allow users to explore recommended tech fields and course options.
 - Suggests user custom-tailored study suggestions and tips aligning with their learning styles.
 - Integrated payment gateway to charge for the assessment (introductory discounts).
- Additional Features:
 - Content: Provides informative articles and resources related to tech upskilling and career paths.
 - Community: Offer a forum or chat functionality for users to connect and learn from each other.
 - Feedback: Implements a feedback mechanism to gather user input and improve the assessment and recommendations.

13. Code implementation (MVP dataset)

Implemented code can be found at this link: https://github.com/sanhiitaa/feynnlabs-internship-2024/blob/main/minimum_viable_product_dataset_EDA.ipynb

This section depicts code implementation for performing Exploratory Data Analysis (EDA) on the output dataset generated by the project's three models. The code provides insights into the dataset's characteristics and relationships, aiding in understanding its

potential applications and implications, thus forming the foundational basis for developing the minimum viable product.

Importing necessary libraries

▼ Note

This is a synthetic dataset used to depict what the final output generated by the 3 models would look like. It is meant to serve as a foundational component in the development of the minimum viable product.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

+ Code + Text

▼ Loading the dataset

```
[3] df=pd.read_csv('/content/syntheticdata.csv')
df.columns
```

Index(['Student_ID', 'Name', 'Age', 'Educational_Background', 'Reason_for_Gap_Year', 'Desired_Tech_Field', 'Cognitive_Capacity_Score', 'Personality_Traits', 'Interests', 'Learning_Styles', 'Recommended_Tech_Field', 'Probability_Score', 'Course_Name', 'Provider', 'Course_Cost_Rs', 'Test_Fee_Rs', 'Commission_Rs'], dtype='object')

▼ Exploratory Data Analysis

[] # first few rows of the dataset
df.head(10)

	Student_ID	Name	Age	Educational_Background	Reason_for_Gap_Year	Desired_Tech_Field	Cognitive_Capacity_Score	Personality_Traits	Interests	Learning_Styles	Recommended_Tech_Field
0	1	Aryan	23	Bachelor's in Engineering	Competitive Exam Preparation	Data Science	85	Analytical	Data Analysis	Visual	Machine Learning
1	2	Priya	22	Bachelor's in Mathematics	Family Commitments	Artificial Intelligence	78	Creative	Machine Learning	Auditory	Web Development
2	3	Rohan	24	Bachelor's in Computer Science	Health Issues	Web Development	92	Organized	Front-end Development	Kinesthetic	Cybersecurity
3	4	Kavita	21	Bachelor's in Physics	Personal Interest	Cybersecurity	80	Reserved	Ethical Hacking	Visual	Digital Marketing
4	5	Rahul	25	Bachelor's in Business Administration	Family Business Responsibilities	Digital Marketing	75	Outgoing	Social Media Marketing	Auditory	Machine Learning

Statistical

```
[ ] # statistics
df.describe()
```

	Student_ID	Age	Cognitive_Capacity_Score	Probability_Score	Course_Cost_Rs	Test_Fee_Rs	Commission_Rs
count	70.000000	70.000000	70.000000	70.000000	69.000000	69.0	69.000000
mean	35.500000	22.942857	85.900000	0.770286	3024.637681	500.0	151.231884
std	20.351085	1.261305	3.257767	0.030973	379.010242	0.0	18.950512
min	1.000000	21.000000	75.000000	0.650000	1800.000000	500.0	90.000000
25%	18.250000	22.000000	84.250000	0.750000	2800.000000	500.0	140.000000
50%	35.500000	23.000000	86.000000	0.770000	3000.000000	500.0	150.000000
75%	52.750000	24.000000	88.000000	0.790000	3200.000000	500.0	160.000000
max	70.000000	25.000000	92.000000	0.820000	3800.000000	500.0	190.000000

```
[ ] # rows in the dataset
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 70 entries, 0 to 69
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Student_ID            70 non-null    int64
1   Name                  70 non-null    object
2   Age                   70 non-null    int64
3   Educational_Background 70 non-null    object
4   Reason_for_Gap_Year    70 non-null    object
5   Desired_Tech_Field     70 non-null    object
6   Cognitive_Capacity_Score 70 non-null    int64
7   Personality_Traits     70 non-null    object
8   Interests             70 non-null    object
9   Learning_Styles       70 non-null    object
10  Recommended_Tech_Field  70 non-null    object
11  Probability_Score      70 non-null    float64
12  Course_Name           69 non-null    object
13  Provider              69 non-null    object
14  Course_Cost_Rs        69 non-null    float64
15  Test_Fee_Rs          69 non-null    float64
16  Commission_Rs         69 non-null    float64
dtypes: float64(4), int64(3), object(10)
memory usage: 9.4+ KB
```

```
[ ] # checking for null values
print(df.isnull().sum(), '\n')
print('\n','***20)
print("total null values in the dataset: ", df.isnull().sum().sum())

# checking for duplicated values
print('\n','***20)
print("\ntotal duplicated values in the dataset: ", df.duplicated().sum())
```

```
Student_ID      0
Name            0
Age            0
Educational_Background  0
Reason_for_Gap_Year    0
Desired_Tech_Field     0
Cognitive_Capacity_Score  0
Personality_Traits     0
Interests           0
Learning_Styles       0
Recommended_Tech_Field  0
Probability_Score      0
Course_Name         1
Provider           1
Course_Cost_Rs      1
Test_Fee_Rs        1
Commission_Rs       1
dtype: int64

*****
total null values in the dataset:  5
*****

total duplicated values in the dataset:  0
```

```
[ ] # unique values in categorical columns
for column in df.select_dtypes(include=['object']):
    print(f"{column}: {df[column].nunique()} \n")
```

```
Name: 58
Educational_Background: 8
Reason_for_Gap_Year: 5
Desired_Tech_Field: 45
Personality_Traits: 7
Interests: 64
Learning_Styles: 3
Recommended_Tech_Field: 45
Course_Name: 54
Provider: 3
```

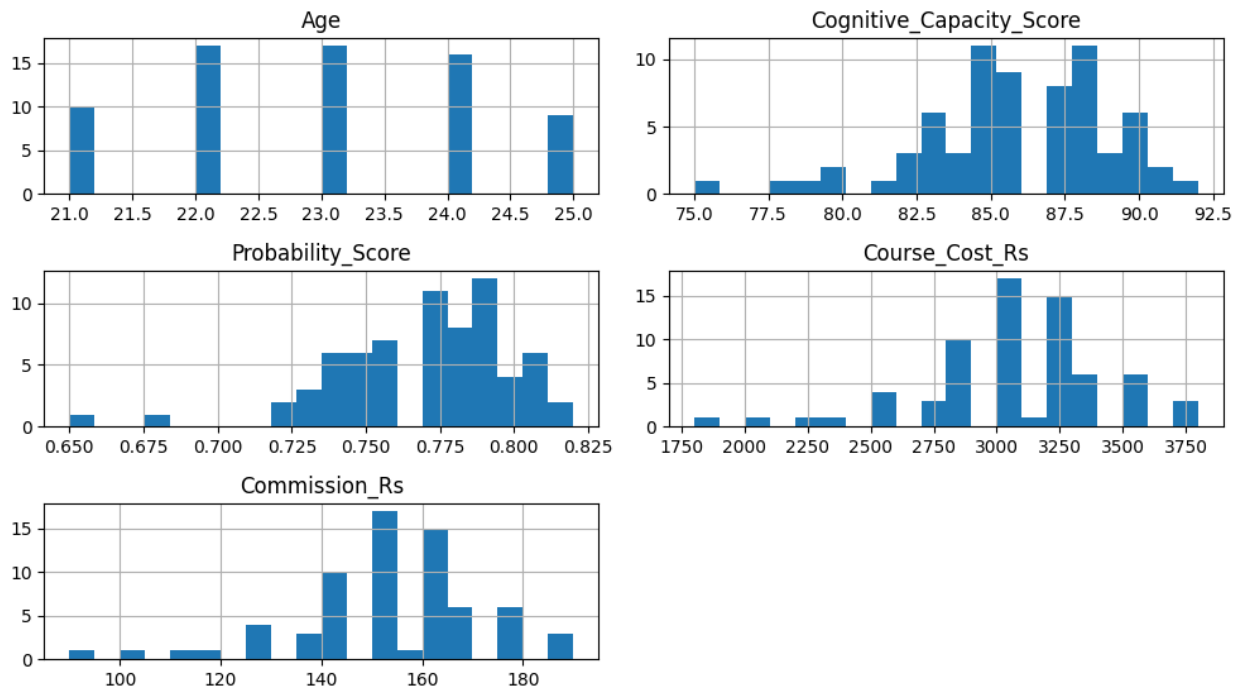
Visualization

```
[ ] # numerical columns
num=[column for column in df.columns if df.dtypes[column]!='object']
num.pop(0); #excluding student id column
# excluding the test fee column as it is a constant
num.pop(4)
```

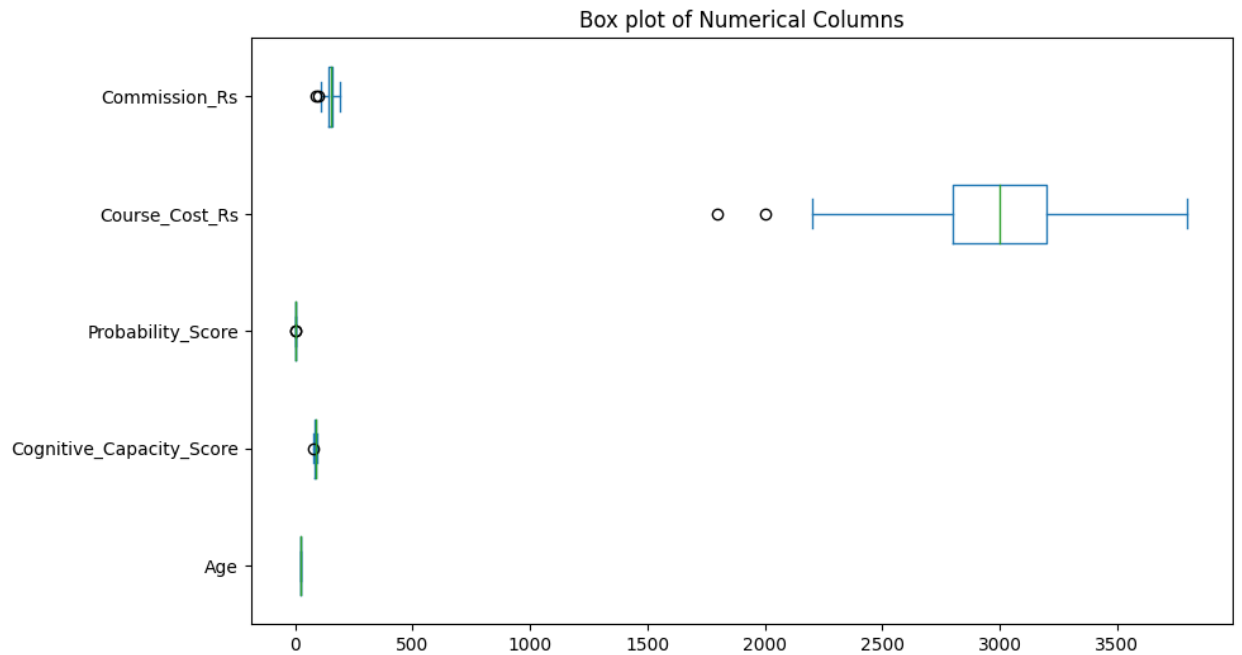
```
'Test_Fee_Rs'
```

```
[ ] # plotting histograms for numerical columns to check distributions
df[num].hist(bins=20, figsize=(10, 6))
plt.suptitle('Histogram of numerical columns');
plt.tight_layout()
```

Histogram of numerical columns

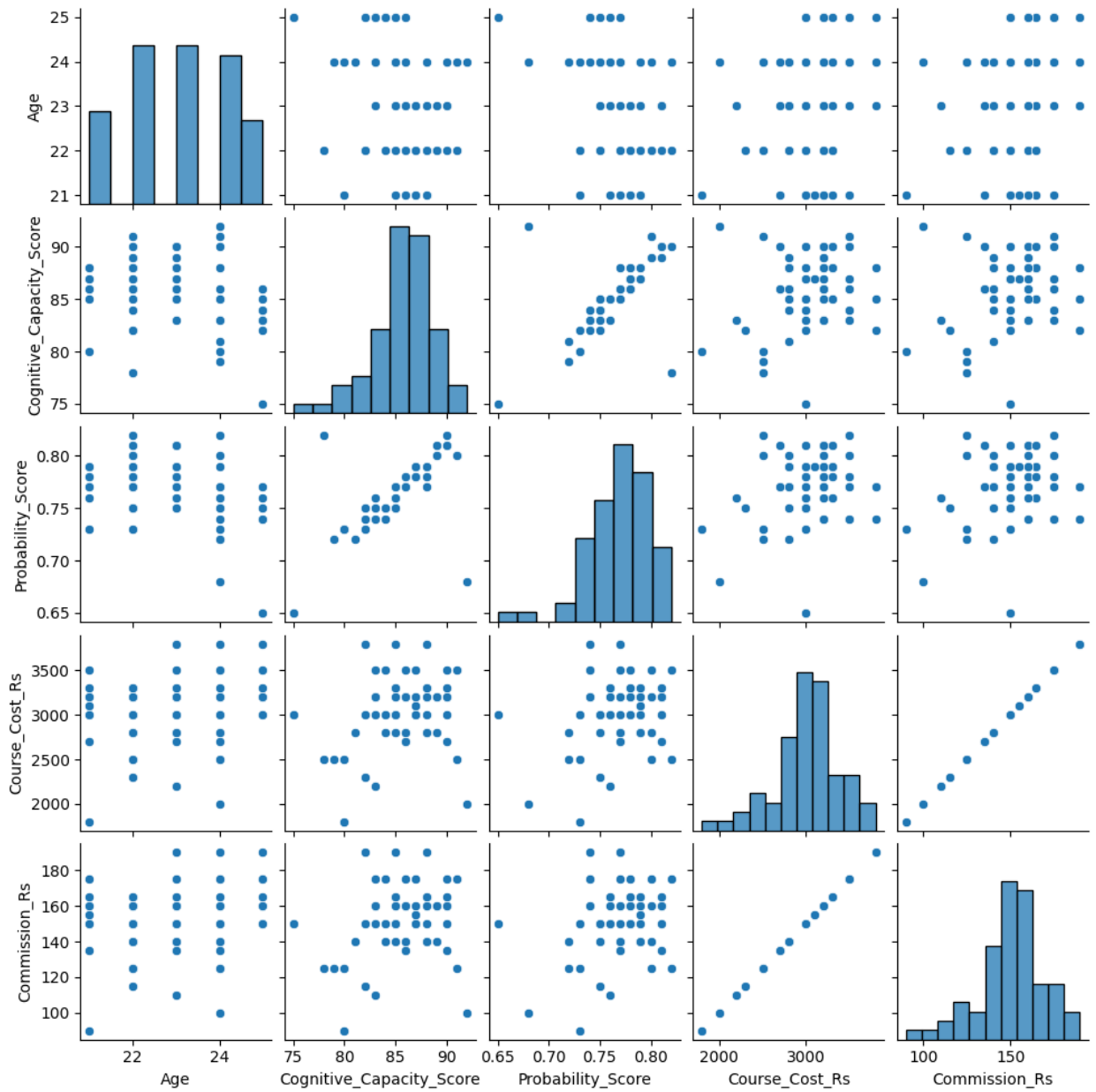


```
[ ] # plotting boxplot to identify outliers
df[num].plot(kind='box', vert=False, figsize=(10, 6))
plt.title('Box plot of Numerical columns')
plt.show()
```



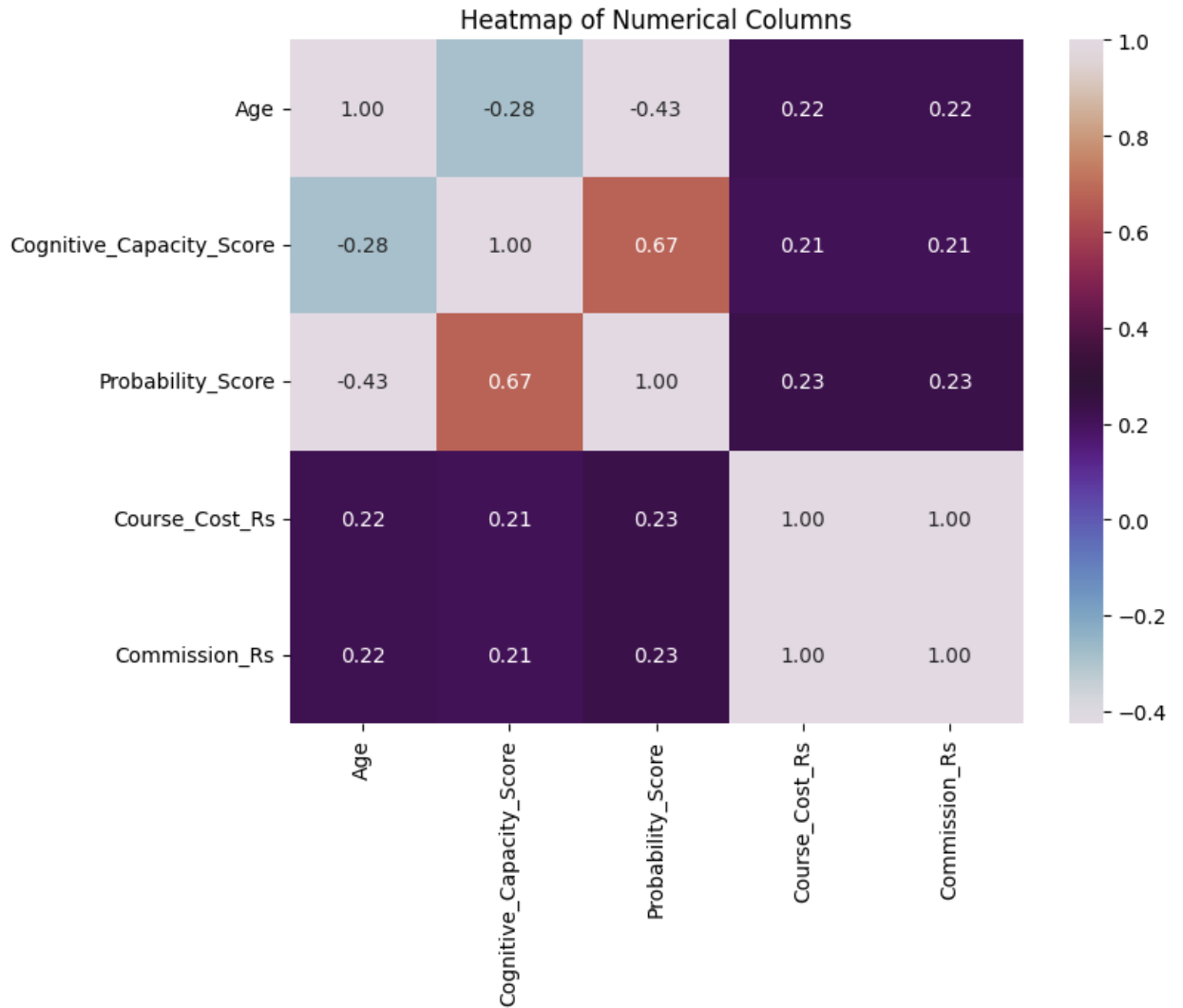
```
[ ] # plotting pairplot to determine relationships between different numerical columns
plt.figure(figsize=(6, 6))
pairplot = sns.pairplot(df[num], height= 2, aspect=1)
plt.suptitle('Pairplot of Numerical Columns', y=1.02)
plt.show()
```

Pairplot of Numerical Columns



```
# calculating the correlation matrix
correlation_matrix = df[num].corr()

# plotting the heatmap to understand correlation between columns
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap="twilight", fmt=".2f")
plt.title('Heatmap of Numerical Columns')
plt.show()
```



14. Conclusion

The app aims to offer a data-driven solution for gap-year students who seek to transition into the technology field. By combining a holistic skills test, and recommendations driven by powerful machine learning models, it aims to empower students to identify skills they can excel at by upskilling and joining the workforce with confidence, thus achieving their career goals in the dynamic and ever-booming tech industry.