

ChatGPT

Prompting


Send a message.



A prompt guides the model to generate useful output

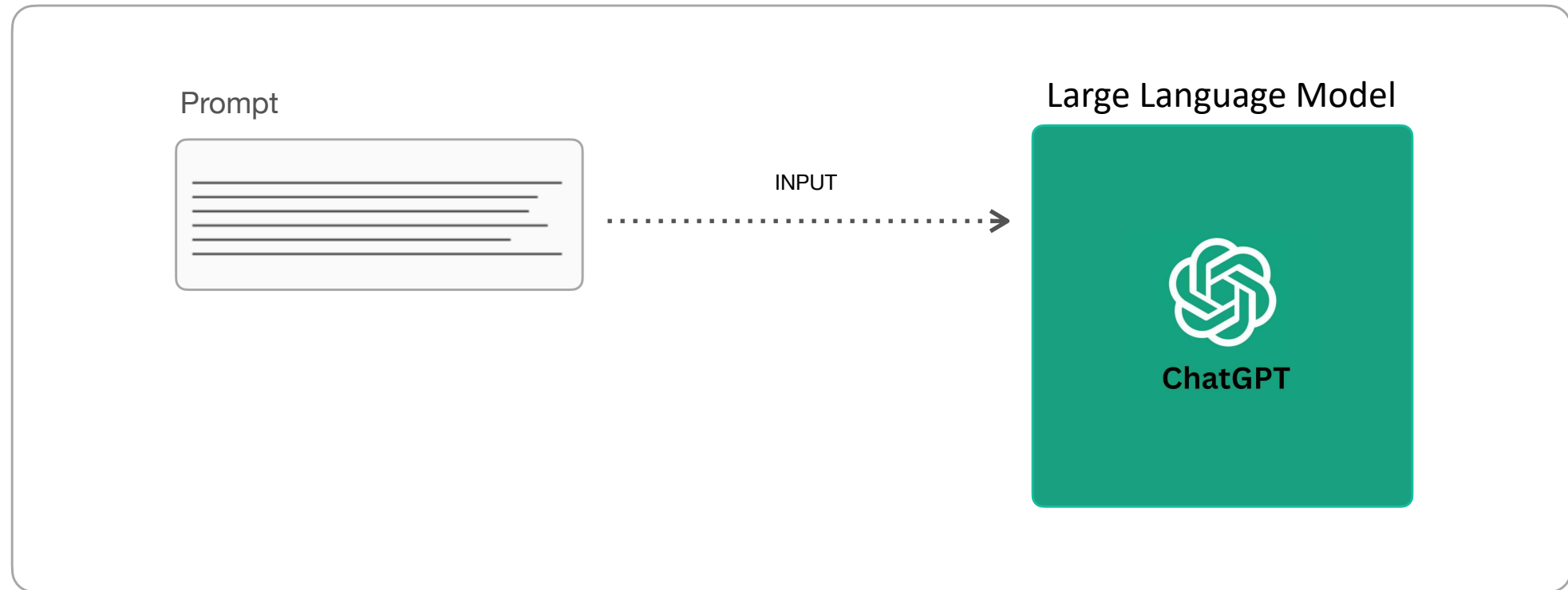
A prompt guides the model to generate useful output

Prompt

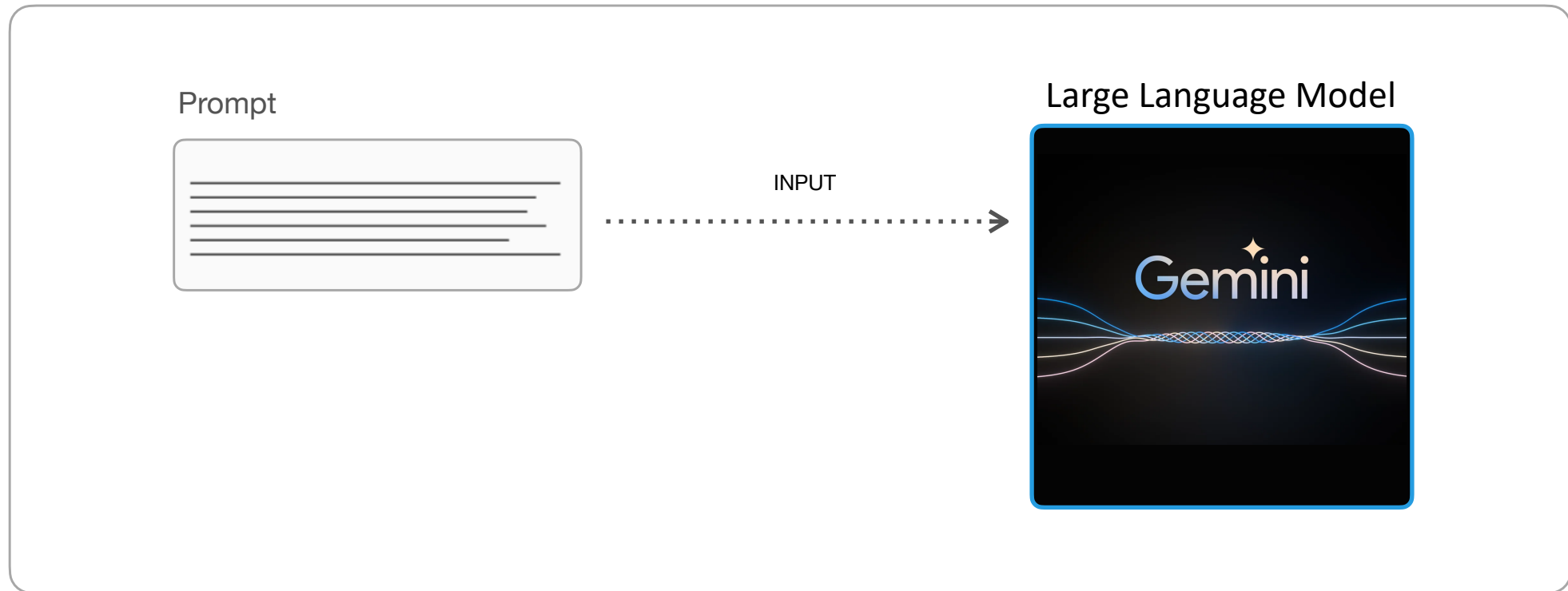


A light gray rounded rectangle representing a text input field. Inside the rectangle, there are six horizontal lines of varying lengths, stacked vertically, indicating a multi-line text input area for a prompt.

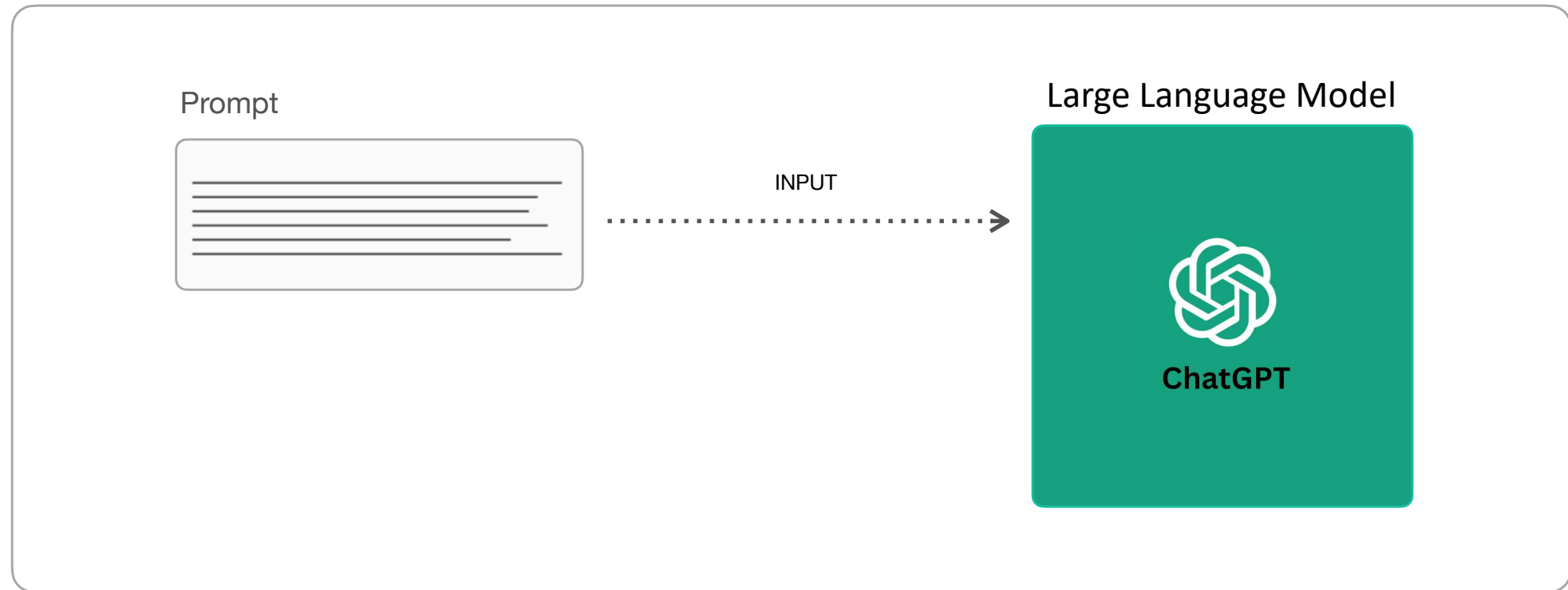
A prompt guides the model to generate useful output



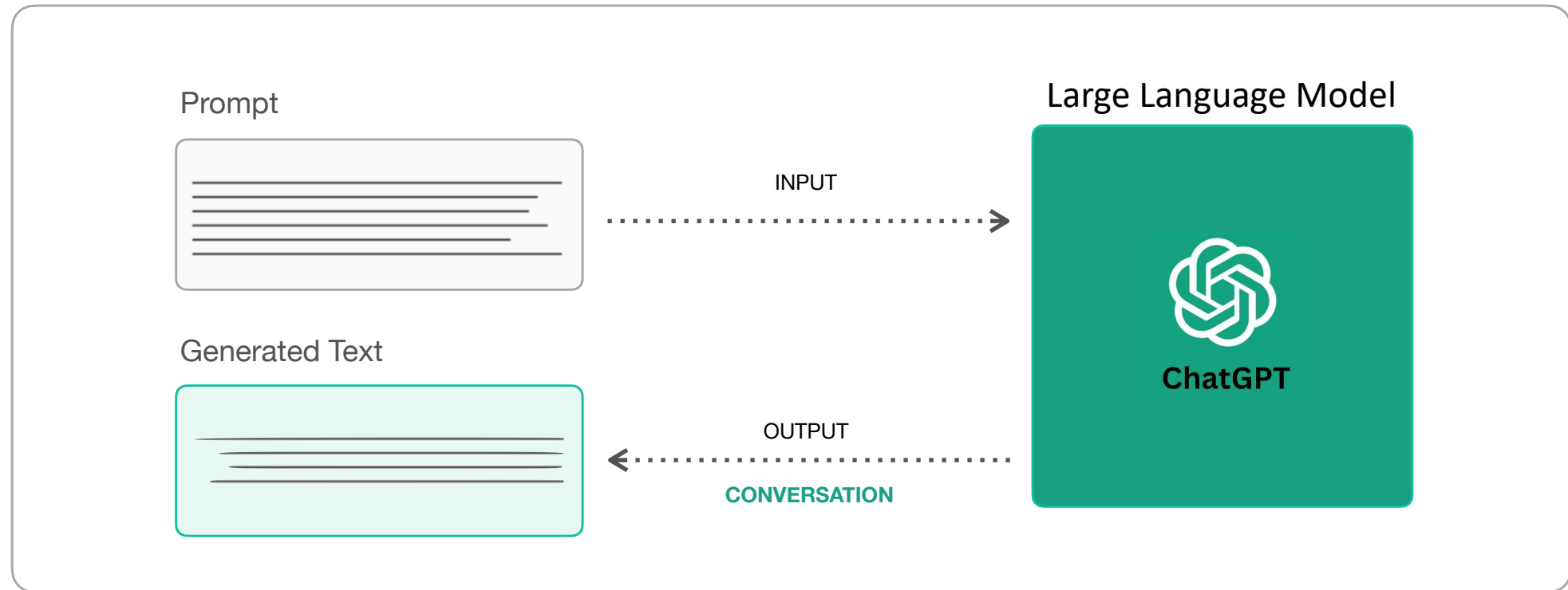
A prompt guides the model to generate useful output



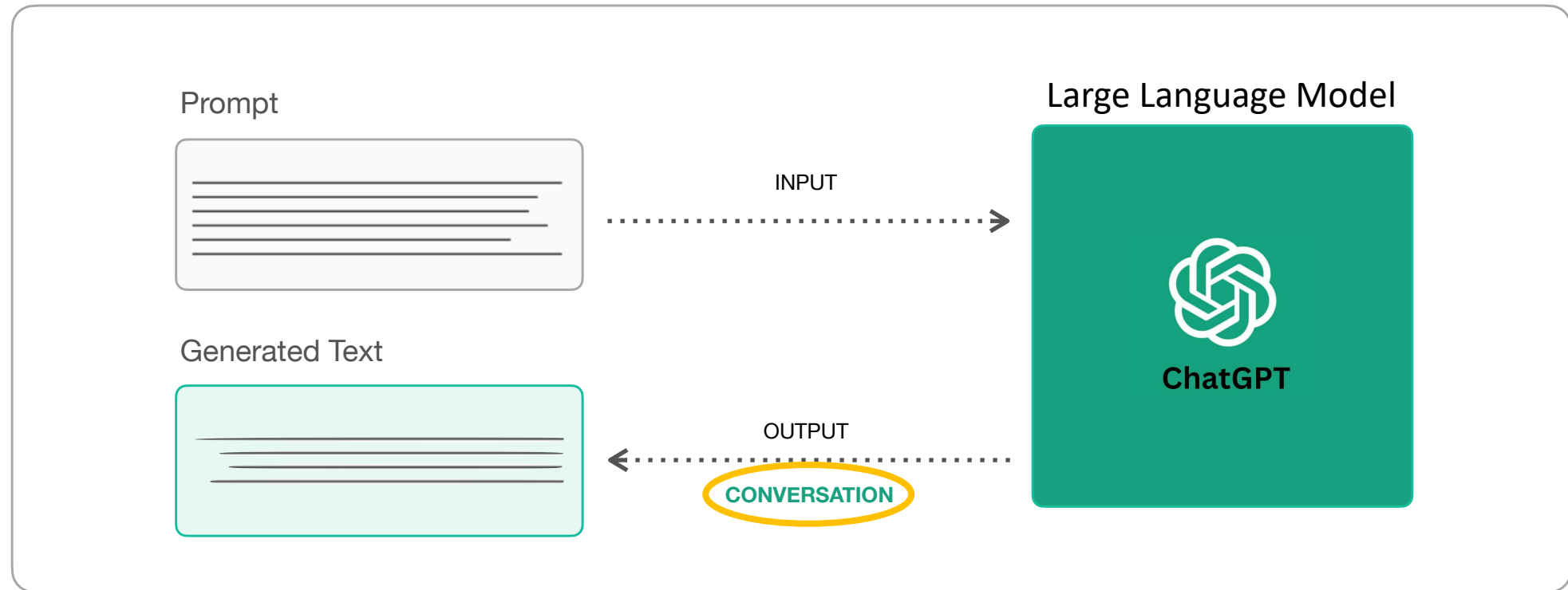
A prompt guides the model to generate useful output



A prompt guides the model to generate useful output



A prompt guides the model to generate useful output



Coding

Coding



Write code in Python that will drop all of the missing values in a data set.

Coding



Write code in Python that will drop all of the missing values in a data set.

python

 Copy code

```
import pandas as pd

# Sample dataset with missing values
data = {
    'A': [1, 2, None, 4],
    'B': [None, 2, 3, 4],
    'C': [1, None, 3, 4]
}

# Create a DataFrame
df = pd.DataFrame(data)

# Drop rows with any missing values
df_cleaned = df.dropna()

# If you want to drop columns with any missing values instead, use:
# df_cleaned = df.dropna(axis=1)
```

Coding



Write code in Python that will drop all of the missing values in a data set.

Coding

P Write code in Python that will drop all of the missing values in a data set.

P Explain the following code:

Coding

P Write code in Python that will drop all of the missing values in a data set.

P Explain the following code:

P Translate the following code from Python to R.

Coding

P Write code in Python that will drop all of the missing values in a data set.

P Explain the following code:

P Translate the following code from Python to R.

P What is the error in the following code?

Data Analysis



grad_admit_.csv
Spreadsheet



Display the first few rows of the dataset.



grad_admit_.csv

Spreadsheet



Display the first few rows of the dataset.



ChatGPT

Here are the first few rows of the dataset displayed as a table:

GRE Score	TOEFL Score	University Rating	SOP	LOR	GPA	Research	Admitted
337	118	4	4.5	4.5	3.89	1	1
324	107	4	4.0	4.5	3.58	1	1
316	104	3	3.0	3.5	3.23	1	0
322	110	3	3.5	2.5	3.50	1	1
314	103	2	2.0	3.0	3.31	0	0



Describe the columns in the dataset.

P Describe the columns in the dataset.



ChatGPT

The dataset consists of the following columns:

1. **GRE Score:** This column represents the Graduate Record Examination score of the applicant.
2. **TOEFL Score:** It indicates the Test of English as a Foreign Language score of the applicant.
3. **University Rating:** This is a rating of the university the applicant is applying to, likely on a scale.
4. **SOP:** Statement of Purpose strength, possibly rated on a scale.
5. **LOR:** Letter of Recommendation strength, possibly rated on a scale.
6. **GPA:** Undergraduate Grade Point Average of the applicant.
7. **Research:** This is a binary indicator (1 or 0) representing whether the applicant has research experience.
8. **Admitted:** This is a binary indicator (1 or 0) representing whether the applicant was admitted. [\[>-\]](#)



How many students with a GRE less than the median and a GPA greater than the median were admitted?



How many students with a GRE less than the median and a GPA greater than the median were admitted?



ChatGPT

There are **15 students** with a GRE Score less than the median and a GPA greater than the median who were admitted. `[>-]`



How many students with a GRE less than the median and a GPA greater than the median were admitted?



ChatGPT

There are **15 students** with a GRE Score less than the median and a GPA greater than the median who were admitted. [[↗](#)]



How many students with a GPA less than the median and a GRE greater than the median were admitted?



How many students with a GRE less than the median and a GPA greater than the median were admitted?



ChatGPT

There are **15 students** with a GRE Score less than the median and a GPA greater than the median who were admitted. [↗]



How many students with a GPA less than the median and a GRE greater than the median were admitted?



ChatGPT

There are **11 students** with a GPA less than the median and a GRE Score greater than the median who were admitted. [↗]



Can you do some creative feature engineering to create a new feature?



Can you do some creative feature engineering to create a new feature?



ChatGPT

Certainly! Here are a few ideas for creative feature engineering for this dataset:



Can you do some creative feature engineering to create a new feature?



ChatGPT

Certainly! Here are a few ideas for creative feature engineering for this dataset:

1. **Composite Score:** You could create a "Composite Score" that combines GRE, TOEFL, GPA, SOP, and LOR into a single metric. This might be a weighted sum where more important features contribute more to the score.
2. **Research Excellence:** Instead of having a binary variable for research, create a "Research Excellence" feature that not only considers whether they have research experience but also considers the university's rating (assuming higher-rated universities provide better research opportunities).
3. **SOP*LOR Interaction:** Sometimes the interaction between how well a student writes their statement of purpose (SOP) and the strength of their letters of recommendation (LOR) can be more telling than the two features individually. You could create an interaction feature by multiplying the two.



Can you do some creative feature engineering to create a new feature?



ChatGPT

Let's implement the "Composite Score" as an example:

- **Composite Score** = $0.3 \times \text{Normalized GRE} +$
 $0.2 \times \text{Normalized TOEFL} +$
 $0.3 \times \text{GPA} +$
 $0.1 \times \text{SOP} +$
 $0.1 \times \text{LOR}$

Here, weights are assumed for demonstration, and normally you would decide these based on domain knowledge or iterative model testing. We'll also normalize GRE and TOEFL scores to be between 0 and 1 for this calculation.

P Can you do some creative feature engineering to create a new feature?



ChatGPT

Here are the first few rows of the updated dataset with the newly engineered "Composite Score" feature:

GRE Score	TOEFL Score	University Rating	SOP	LOR	GPA	Research	Admitted	Normalized GRE	Normalized TOEFL	Composite Score
337	118	4	4.5	4.5	3.89	1	1	0.94	0.928571	2.534714
324	107	4	4.0	4.5	3.58	1	1	0.68	0.535714	2.235143
316	104	3	3.0	3.5	3.23	1	0	0.52	0.428571	1.860714
322	110	3	3.5	2.5	3.50	1	1	0.64	0.642857	1.970571
314	103	2	2.0	3.0	3.31	0	0	0.48	0.392857	1.715571

GRE Score	TOEFL Score	University Rating	SOP	LOR	GPA	Research	Admitted	Normalized GRE	Normalized TOEFL	Composite Score
337	118	4	4.5	4.5	3.89	1	1	0.94	0.928571	2.534714
324	107	4	4.0	4.5	3.58	1	1	0.68	0.535714	2.235143
316	104	3	3.0	3.5	3.23	1	0	0.52	0.428571	1.860714
322	110	3	3.5	2.5	3.50	1	1	0.64	0.642857	1.970571
314	103	2	2.0	3.0	3.31	0	0	0.48	0.392857	1.715571



Drop GRE Score, TOEFL Score, SOP, LOR and GPA. Place the column Admitted at the end.



Drop GRE Score, TOEFL Score, SOP, LOR and GPA. Place the column Admitted at the end.



ChatGPT

Here is the modified DataFrame with "Admitted" placed at the end:

University Rating	Research	Composite Score	Admitted
4	1	2.534714	1
4	1	2.235143	1
3	1	1.860714	0
3	1	1.970571	1
2	0	1.715571	0

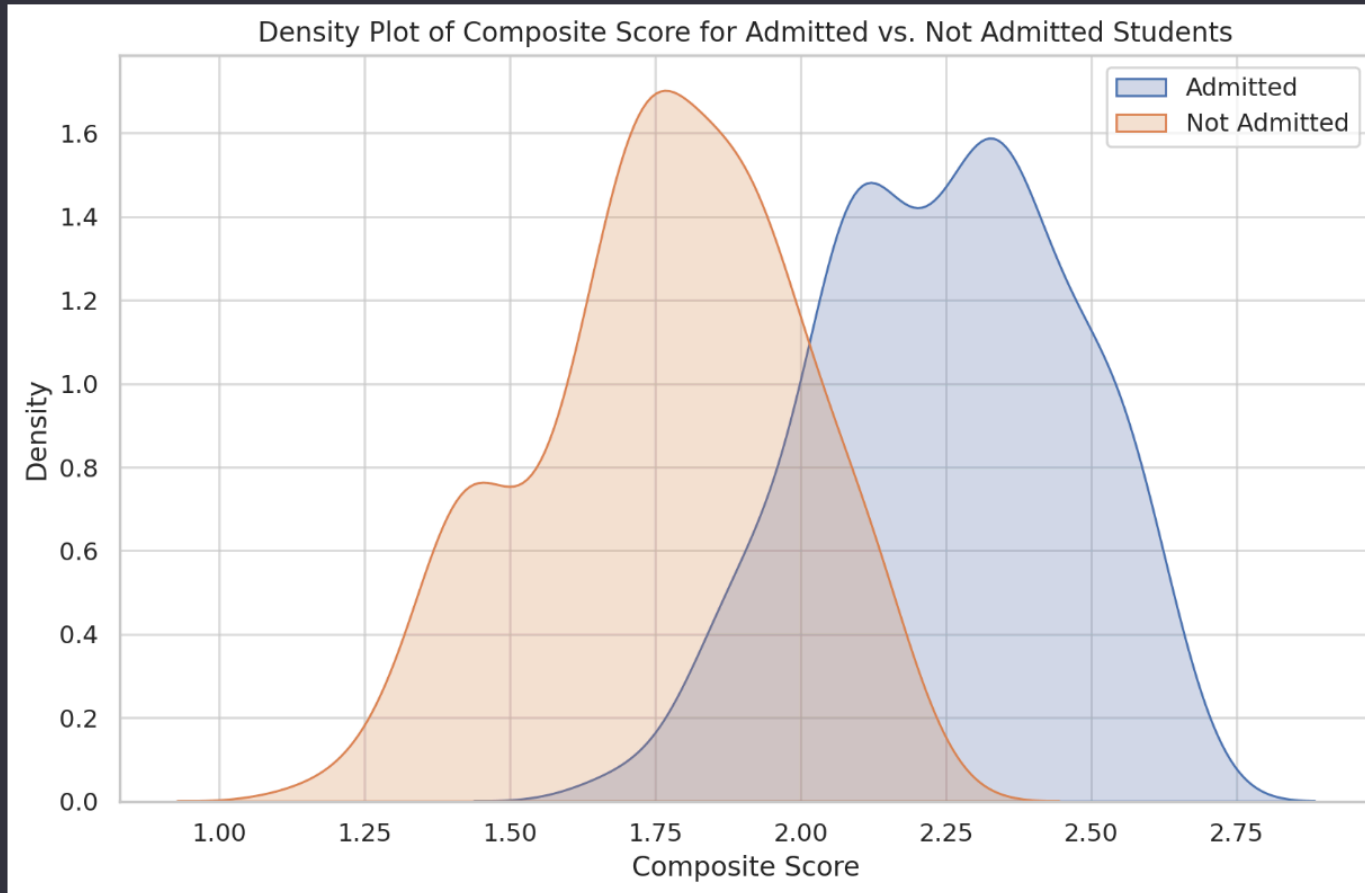
University Rating	Research	Composite Score	Admitted
4	1	2.534714	1
4	1	2.235143	1
3	1	1.860714	0
3	1	1.970571	1
2	0	1.715571	0



Create a plot comparing the Composite Score of those who were admitted to those who weren't admitted.



ChatGPT



Here is the density plot of the Composite Score for students who were admitted versus those who were not. The plot shows the distribution of Composite Scores for both groups, indicating how they differ and overlap. Generally, you might observe trends like higher composite scores being more common among admitted students. [>-]



Provide me with a new CSV file that includes the feature transformations that you made.



Provide me with a new CSV file that includes the feature transformations that you made.



ChatGPT

The modified DataFrame, including the new “Composite Score” feature, has been saved to a CSV file. You can download it using the following link:

[Download the CSV file with the Composite Score \[↗-\]](#)