Phase 4:

In an AI-driven exploration and prediction project, you typically follow a structured workflow that includes Exploratory Data Analysis (EDA), Feature Engineering, and Predictive Modeling. Here's how you can perform each step using Python:

1.Exploratory Data Analysis (EDA):

EDA helps you understand your data, identify patterns, and gather insights. You can use libraries like Pandas, Matplotlib, and Seaborn.

```python
import pandas as pd

import numpy as np

import matplotlib.pyplotas plt

import seaborn as sns

data = pd.read_csv('your_dataset.csv')

print(data.head())

print(data.info())

print(data.describe())

sns.histplot(data['feature1'], kde=True)

sns.boxplot(x='category', y='feature2', data=data)

correlation_matrix = data.corr()

sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm")

sns.countplot(x='category', data=data)
```

2.Feature Engineering:

Feature engineering involves creating new features or transforming existing ones to improve model performance.

```python
data['new_feature'] = data['feature1'] * data['feature2']

data['log_feature1'] = np.log(data['feature1'])

data['categorical_feature'] = pd.get_dummies(data['category'])
```

3.Predictive Modeling:

In this step, you build and evaluate machine learning models to make predictions.

```python
from sklearn.model_selection import train_test_split

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import accuracy_score, classification_report

X = data.drop('target', axis=1)

y = data['target']

model = RandomForestClassifier()

model.fit(X_train, y_train)

y_pred = model.predict(X_test)

accuracy =    accuracy_score(y_test, y_pred)

print(f"Accuracy: {accuracy:.2f}")

print(classification_report(y_test, y_pred))
```