

```

# Q1: What is a parameter?
"""
A parameter is a configuration variable that is internal to the model and estimated from data.
Examples include weights in linear regression or decision thresholds in trees.
"""

# Q2: What is correlation?
"""
Correlation is a statistical measure that expresses the extent to which two variables are linearly related.
It ranges from -1 to +1.
"""

# Q3: What does negative correlation mean?
"""
A negative correlation means that as one variable increases, the other decreases.
Example: Hours of entertainment vs exam performance.
"""

# Q4: Define Machine Learning. What are the main components in Machine Learning?
"""
Machine Learning is the field of study that gives computers the ability to learn from data.
Key components: Data, Features, Model, Loss Function, Optimizer.
"""

# Q5: How does loss value help in determining whether the model is good or not?
"""
The loss value quantifies the error between predicted and true values.
Lower loss → better model performance.
"""

# Q6: What are continuous and categorical variables?
"""
Continuous: Numeric values with infinite possibilities (e.g., height, salary).
Categorical: Non-numeric labels (e.g., gender, city).
"""

# Q7: How do we handle categorical variables in Machine Learning? What are the common techniques?
"""
We convert categories to numbers using techniques like:
- Label Encoding
- One-Hot Encoding
- Ordinal Encoding
"""

# Q8: What do you mean by training and testing a dataset?
"""
Training set is used to train the model.
Testing set is used to evaluate the model's generalization on unseen data.
"""

# Q9: What is sklearn.preprocessing?
"""
A module in scikit-learn that provides functions for preprocessing data:
- Scaling
- Encoding
- Normalization
"""

# Q10: What is a Test set?
"""
The test set is a portion of the dataset kept aside to evaluate the model after training.
It simulates real-world data.
"""

# Q11: How do we split data for model fitting (training and testing) in Python?
from sklearn.model_selection import train_test_split

# Example:
# X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Q12: How do you approach a Machine Learning problem?
"""
1. Understand the problem
2. Collect and clean data
3. Perform EDA
4. Feature Engineering
5. Model training
6. Evaluation
7. Tuning
"""

```

```

# Q13: Why do we have to perform EDA before fitting a model to the data?
"""
EDA helps us understand data patterns, spot missing values, detect outliers,
and choose the right features and transformations.
"""

# Q14: What is correlation?
"""
Correlation measures how strongly two variables are related linearly.
"""

# Q15: What does negative correlation mean?
"""
It means when one variable increases, the other tends to decrease.
"""

# Q16: How can you find correlation between variables in Python?
import pandas as pd

df = pd.DataFrame({'a': [1, 2, 3], 'b': [3, 2, 1]})
correlation = df.corr()
print("Correlation Matrix:\n", correlation)

# Q17: What is causation? Explain difference between correlation and causation with an example.
"""
Causation: One variable directly affects another.
Correlation: Variables move together, not necessarily causally.
Example: Ice cream sales and drowning are correlated, but not causal.
"""

# Q18: What is an Optimizer? What are different types of optimizers? Explain each with an example.
"""
An optimizer updates model parameters to minimize loss.
Types:
- SGD: Simple gradient step
- Adam: Adaptive gradients with momentum
- RMSProp: Scaled learning rates
"""

# Q19: What is sklearn.linear_model ?
"""
A module that provides linear models like:
- LinearRegression
- LogisticRegression
- Ridge, Lasso, etc.
"""

# Q20: What does model.fit() do? What arguments must be given?
"""
Trains the model using the training data.
Arguments: model.fit(X_train, y_train)
"""

# Q21: What does model.predict() do? What arguments must be given?
"""
Predicts target values using the trained model.
Arguments: model.predict(X_test)
"""

# Q22: What are continuous and categorical variables?
"""
Continuous = Numeric
Categorical = Labels or categories
"""

# Q23: What is feature scaling? How does it help in Machine Learning?
"""
Scaling transforms features to a common scale.
Helps models converge faster and avoid dominance by large-range features.
"""

# Q24: How do we perform scaling in Python?
from sklearn.preprocessing import StandardScaler

# Example:
# scaler = StandardScaler()
# X_scaled = scaler.fit_transform(X)

# Q25: What is sklearn.preprocessing?
"""
Module for data transformations like:
- Scaling

```

```
- Normalization  
- Encoding  
Used before feeding data into models.  
"""
```

Double-click (or enter) to edit