Addressing missing data technique summary

1. Dropping missing data using Pandas:

```
import pandas as pd
import numpy as np

# Drop rows with missing data
df.dropna(axis=0, inplace=True)
# Drop columns with missing data
df.dropna(axis=1, inplace=True)

# Drop specific observations or variables
df.drop([1, 2], axis=0, inplace=True) # Drop rows 1 and 2
df.drop('VariableName', axis=1, inplace=True) # Drop column 'VariableName'
```

2. Imputing missing data using Pandas and Numpy:

```
import pandas as pd
import numpy as np

# Forward-fill or backward-fill missing values
df.fillna(method='ffill', inplace=True) # Forward-fill
df.fillna(method='bfill', inplace=True) # Backward-fill

# Mean, median, or mode imputation
df.fillna(df.mean(), inplace=True) # Mean imputation
df.fillna(df.median(), inplace=True) # Median imputation
df.fillna(df.mode().iloc[0], inplace=True) # Mode imputation

# Interpolation
df.interpolate(inplace=True) # Default method is linear interpolatio
# Replace missing values in a specific column with a specific value
df['ColumnName'].fillna(0, inplace=True) # Replace missing values with 0
```

3. Predicting missing data using Scikit-learn:

```
import pandas as pd
import numpy as np

# Split data into complete and incomplete sets
complete_data = df.dropna()
incomplete_data = df[df.isnull().any(axis=1)]

# Select features and target variable
X_complete = complete_data.drop('TargetVariable', axis=1)
y_complete = complete_data['TargetVariable']
X_incomplete = incomplete_data.drop('TargetVariable', axis=1)

# Train machine learning model
model = RandomForestRegressor()
model.fit(X_complete, y_complete)

# Predict missing values
y_pred = model.predict(X_incomplete)

# Fill in missing values in original data set
df.loc[df['TargetVariable'].isnull(), 'TargetVariable'] = y_pred
```

4. Multiple imputation using Scikit-learn:

```
import pandas as pd
import numpy as np

# Create multiple imputed data sets
imputer = IterativeImputer(random_state=0)
imputed_data = imputer.fit_transform(df)

# Analyze each imputed data set separately
for i in range(imputed_data.shape[0]):
    data_set = imputed_data[i, :]
    # Perform analysis on data_set

#Handling missing data in time series using Pandas:
# Fill in missing values using linear interpolation
df['ColumnName'].interpolate(method='linear', inplace=True)
# Fill in missing values using spline interpolation
df['ColumnName'].interpolate(method='spline', order=3, inplac
```