



## Pandas Advance Quiz – 2

8 out of 8 correct

1. Given a Pandas DataFrame df with a column 'date', how can you create a new column 'month' that contains the month value of each date?

- ☐ `df['month'] = df['date'].month`
- ☐ `df['month'] = df['date'].apply(lambda x: x.month)`
- ☐ `df['month'] = pd.to_datetime(df['date']).month`
- ☒ `df['month'] = df['date'].dt.month`

**Explanation:** The dt accessor is used to access the datetime components of a column in a Pandas DataFrame. Here, we can use the dt.month attribute to extract the month value of each date and create a new column.

2. How can you group a Pandas DataFrame df by the 'category' column and calculate the mean of the 'value' column for each group?

- ☐ `df.groupby('category').mean('value')`
- ☒ `df.groupby('category')['value'].mean()`
- ☐ `df.mean('category', 'value')`



- ☐ `df.mean(groupby='category', column='value')`

**Explanation:** The `groupby()` method is used to group a DataFrame by one or more columns, and the `mean()` method can be used to calculate the mean value of a specific column for each group. Here, we can use the syntax `df.groupby('category')['value'].mean()` to group by the 'category' column and calculate the mean of the 'value' column for each group.

3. How can you create a scatter plot of two columns 'x' and 'y' in a Pandas DataFrame df?

- ☐ `df.plot(x='x', y='y', kind='scatter')`
- ☐ `df.scatter(x='x', y='y')`
- ☐ `df.plot.scatter(x='x', y='y')`
- ☒ `df.plot(kind='scatter', x='x', y='y')`

**Explanation:** The `plot()` method can be used to create various types of plots, including scatter plots, and we can specify the columns for the x and y axes using the x and y arguments. Here, we can use the syntax `df.plot(kind='scatter', x='x', y='y')` to create a scatter plot of columns 'x' and 'y'.

4. How can you read a CSV file 'data.csv' into a Pandas DataFrame and display the first 5 rows of data?

- ☐ `df = pd.read_csv('data.csv', nrows=5)`
- ☐ `df = pd.read_csv('data.csv', skiprows=5)`
- ☐ `df = pd.read_csv('data.csv').head()`
- ☒ `df = pd.read_csv('data.csv').iloc[:5]`

**Explanation:** The `read_csv()` method is used to read a CSV file into a Pandas DataFrame. We can use the `iloc[]` method to select the first 5 rows of data. Here, we can use the syntax `df = pd.read_csv('data.csv').iloc[:5]` to read the CSV file 'data.csv' and display the first 5 rows of data.

5. Which method in Pandas can be used to create a time delta object?

- ☒ `pd.to_timedelta()`
- ☐ `pd.timedelta()`
- ☐ `pd.create_timedelta()`
- ☐ `pd.time_delta()`

**Explanation:** The `to_timedelta()` method in Pandas can be used to create a time delta object from a string, number or timedelta-like input.

6. How can you convert a column in a Pandas DataFrame to categorical data?

- ☐ `df.convert_to_categorical()`
- ☐ `df.to_categorical()`
- ☒ `df.astype('category')`
- ☐ `df.categorical()`

**Explanation:** The `astype()` method in Pandas can be used to convert a column to categorical data by passing 'category' as the argument.

7. Which method in Pandas can be used to create a scatter plot?

- ☒ `df.plot(kind='scatter')`
- ☐ `df.plot(kind='line')`
- ☐ `df.plot(kind='hist')`
- ☐ `df.plot(kind='bar')`

**Explanation:** The `plot()` method in Pandas can be used to create different types of plots, including scatter plots by specifying 'scatter' as the kind argument.

8. How can you drop a column from a Pandas DataFrame?

- ☐ `df.drop_column('column_name')`
- ☐ `df.remove_column('column_name')`
- ☒ `df.drop('column_name', axis=1)`
- ☐ `df.remove('column_name', axis=1)`

**Explanation:** The `drop()` method in Pandas can be used to drop a column from a DataFrame by specifying the column name and `axis=1`. Option a and b are not valid Pandas methods and d is not a valid argument for the drop method.

Submit