

Using merge_asof() to create dataset - Corrected

The screenshot shows a web browser window with the URL `campus.datacamp.com/courses/joining-data-with-pandas/merging-ordered-and-time-series-data?ex=7`. The page displays an exercise titled "Using merge_asof() to create dataset". The instructions on the left side of the interface are as follows:

- The `merge_asof()` function can be used to create datasets where you have a table of start and stop dates, and you want to use them to create a flag in another table. You have been given `gdp`, which is a table of quarterly GDP values of the US during the 1980s. Additionally, the table `recession` has been given to you. It holds the starting date of every US recession since 1980, and the date when the recession was declared to be over. Use `merge_asof()` to merge the tables and create a status flag if a quarter was during a recession. Finally, to check your work, plot the data in a bar chart.
- The tables `gdp` and `recession` have been loaded for you.

Below the instructions, there is a "Take Hint (30 XP)" button. On the right side of the interface, there is a code editor with the following Python code:

```
1 # Merge gdp and recession on date using merge_asof()
2 gdp_recession = _____
3
4 # Create a list based on the row value of gdp_recession['econ_status']
5 is_recession = [ _____ if s == 'recession' else 'g' for s in gdp_recession['econ_status']]
6
7 # Plot a bar chart of gdp_recession
8 gdp_recession.plot(kind=_____, x=_____, color=_____, rot=90)
9 plt.show()
```

At the bottom of the code editor, there are buttons for "Run Code" and "Submit Answer". Below the code editor is a "Python Shell" area with the prompt `In [1]:`.

Question:

Using `merge_asof()`, merge `gdp` and `recession` on `date`, with `gdp` as the left table. Save to the variable `gdp_recession`. Create a list using a list comprehension and a conditional expression, named `is_recession`, where for each row if the `gdp_recession['econ_status']` value is equal to 'recession' then enter 'r' else 'g'. Using `gdp_recession`, plot a bar chart of `gdp` versus `date`, setting the `color` argument equal to `is_recession`.

Answer:

```
# Merge gdp and recession on date using merge_asof
gdp_recession = pd.merge_asof(
    gdp, recession,
    on='date'
)
```

```
# Create a list based on the row value of gdp_recession['econ_status']
is_recession = [
    'r' if s == 'recession' else 'g'
    for s in gdp_recession['econ_status']
]
```

```
# Plot a bar chart of gdp_recession
```

```
gdp_recession.plot(  
    kind='bar',  
    x='date',  
    y='gdp',  
    color=is_recession,  
    rot=90  
)  
plt.show()
```

Code Explanation:

1. `gdp_recession = pd.merge_asof(...)`:

This line merges the `gdp` and `recession` dataframes on the `'date'` column using the `'merge_asof'` function. The `'gdp'` dataframe is set as the left table, aligning rows based on the nearest previous date.

2. `is_recession = [...]`:

This line creates a list comprehension that iterates through each value in the `'econ_status'` column of the `gdp_recession` dataframe. If the value is `'recession'`, `'r'` is added to the list; otherwise, `'g'` is added.

3. `gdp_recession.plot(...)`:

This line creates a bar chart of the `'gdp'` values versus `'date'` from the `gdp_recession` dataframe. The `'color'` argument is set to the `is_recession` list, coloring the bars red (`'r'`) or green (`'g'`) based on the economic status.