

How can I concatenate strings from using groupby in Pandas?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I concatenate strings from using groupby in Pandas?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159554>

The process of concatenating strings from using the groupby function in Pandas involves grouping a DataFrame by a specific column or multiple columns and then applying the .agg() method to combine the strings within each group. This allows for the creation of a new column containing concatenated strings, providing a convenient way to consolidate and analyze data in a more organized manner. By utilizing this method, users can efficiently manipulate and extract valuable insights from their data.

Pandas: Concatenate Strings from Using GroupBy

You can use the following basic syntax to concatenate strings from using GroupBy in pandas:

```
df.groupby(, as_index=False).agg({'string_var': ' '.join})
```

This particular formula groups rows by the group_var column and then concatenates the strings in the string_var column.

The following example shows how to use this syntax in practice.

Example: How to Concatenate Strings from Using GroupBy

Suppose we have the following pandas DataFrame:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'store': ,
```

```
'quarter': ,  
'employee': })
```

```
#view DataFrame  
print(df)
```

We can use the following syntax to group the rows of the DataFrame by store and quarter and then concatenate the strings in the employee column:

```
#group by store and quarter, then concatenate  
employee strings  
df.groupby(, as_index=False).agg({'employee': ' '.join})
```

```
store quarter employee  
0 A 1 Andy Bob  
1 A 2 Chad Diane  
2 B 1 Elana Frank  
3 B 2 George Hank
```

The result is a DataFrame grouped by store and quarter with the strings in the employee column concatenated together with a space.

We could also concatenate the strings using a different

separator such as the & symbol:

#group by store and quarter, then concatenate employee strings

```
df.groupby(, as_index=False).agg({'employee': ' & '.join})
```

store quarter employee

0 A 1 Andy & Bob

1 A 2 Chad & Diane

2 B 1 Elana & Frank

3 B 2 George & Hank

Notice that the strings in the employee column are now separated by the & symbol.

Note: You can find the complete documentation for the GroupBy operation in pandas .

Additional Resources

The following tutorials explain how to perform other common operations in pandas: