

How do I select rows in Pandas that do not start with a specific string?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I select rows in Pandas that do not start with a specific string?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=155064>

Pandas is a popular data analysis library in Python that allows users to manipulate and analyze data efficiently. One common task in data analysis is selecting rows of data based on certain criteria. In order to select rows in Pandas that do not start with a specific string, users can use the "str.startswith()" method to check if a string starts with a given prefix. By using the "~" symbol, users can invert the boolean values and select rows that do not meet the specified criteria. This method is particularly useful for filtering out unwanted data and extracting only the relevant information for analysis.

Pandas: Select Rows that Do Not Start with String

You can use the following basic syntax to select rows that do not start with a specific string in a pandas DataFrame:

```
df
```

This particular formula selects all rows in the DataFrame where the column called my_column does not start with the string this or the string that.

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

Example: Select Rows that Do Not Start with String in Pandas

Suppose we have the following pandas DataFrame that contains information about sales for various stores:

```
import pandas as pd
```

```
#create DataFrame
```

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

```
#view DataFrame
```

```
print(df)
```

```
store sales
```

```
0 Upper East 150
```

```
1 Upper West 224
```

```
2 Lower East 250
```

```
3 West 198
```

```
4 CTR 177
```

We can use the following syntax to select all rows in the DataFrame that do not start with the strings 'Upper' or 'Lower' in the store column:

```
#select all rows where store does not start with 'Upper'  
or 'Lower'
```

```
df
```

```
store sales
```

```
3 West 198
```

```
4 CTR 177
```

Notice that the only rows returned are the ones where the store column does not start with 'Upper' or 'Lower.'

If you'd like, you can also define the tuple of strings outside of the startswith() function:

```
#define tuple of strings
```

```
some_strings = ('Upper', 'Lower')
```

```
#select all rows where store does not start with strings  
in tuple
```

```
df
```

```
store sales
```

```
3 West 198
```

```
4 CTR 177
```

This produces the same result as the previous method.

Note: You can find the complete documentation for the startswith function in pandas .

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