

How can I fix a KeyError in Pandas? (With an example)

Authored by
stats writer

May 5, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I fix a KeyError in Pandas? (With an example)*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142859>

KeyError in Pandas is a common error that occurs when trying to access a dictionary key that does not exist. This can be fixed by identifying the missing key and either creating it or modifying the code to avoid accessing it. For example, if a dataframe has columns titled "Name" and "Age", attempting to access the column "Gender" would result in a KeyError. To fix this, the code can be modified to only access the existing columns or the "Gender" column can be added to the dataframe. By properly handling the KeyError, the code can be made more efficient and accurate in working with the dataframe.

Fix KeyError in Pandas (With Example)

One error you may encounter when using pandas is:

KeyError: 'column_name'

This error occurs when you attempt to access some column in a pandas DataFrame that does not exist.

Typically this error occurs when you simply misspell a column names or include an accidental space before or after the column name.

The following example shows how to fix this error in practice.

How to Reproduce the Error

Suppose we create the following pandas DataFrame:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'points': ,  
'assists': ,  
'rebounds': })
```

```
#view DataFrame
```

```
df
```

```
points assists rebounds
```

```
0 25 5 11
```

```
1 12 7 8
```

```
2 15 7 10
```

```
3 14 9 6
```

```
4 19 12 6
```

```
5 23 9 5
```

```
6 25 9 9
```

```
7 29 4 12
```

Then suppose we attempt to print the values in a column called 'point':

```
#attempt to print values in 'point' column
```

```
print(df)
```

KeyError Traceback (most recent call last)

```
/srv/conda/envs/notebook/lib/python3.7/site-  
packages/pandas/core/indexes/base.py in get_loc(self,  
key, method, tolerance)
```

3360 try:

-> 3361 return self._engine.get_loc(casted_key)

3362 except KeyError as err:

```
/srv/conda/envs/notebook/lib/python3.7/site-  
packages/pandas/_libs/index.pyx in  
pandas._libs.index.IndexEngine.get_loc()
```

```
/srv/conda/envs/notebook/lib/python3.7/site-  
packages/pandas/_libs/index.pyx in  
pandas._libs.index.IndexEngine.get_loc()
```

```
pandas/_libs/hashtable_class_helper.pxi in  
pandas._libs.hashtable.PyObjectHashTable.get_item()
```

```
pandas/_libs/hashtable_class_helper.pxi in  
pandas._libs.hashtable.PyObjectHashTable.get_item()
```

KeyError: 'point'

Since there is no 'point' column in our DataFrame, we receive a KeyError.

How to Fix the Error

The way to fix this error is to simply make sure we spell the column name correctly.

If we're unsure of all of the column names in the DataFrame, we can use the following syntax to print each column name:

```
#display all column names of DataFrame  
print(df.columns.tolist())
```

We can see that there is a column called 'points', so we can fix our error by spelling the column name correctly:

```
#print values in 'points' column  
print(df)
```

0 25

1 12

2 15

3 14

4 19

5 23

6 25

7 29

Name: points, dtype: int64

The following tutorials explain how to fix other common errors in Python:

ARABPSYCHOLOGY.COM