

4.7.3 Get the Number of Schools

Great work on getting the total number of students! It's a good thing you didn't drop those incorrect names, or you would still be cleaning data and behind in your work. Now it's time to get the total number of schools in the school district.

REWIND

Remember we found the total number of schools by using the `count()` method on the `school_data_df` DataFrame. The output was 15.

There are two ways we can get the number of schools. One is to use the `school_data_df` DataFrame and assign a variable to one of the columns in the DataFrame, like this:

```
# Calculate the total number of schools.  
school_count = school_data_df["school_name"].count()  
school_count
```

After running the code, the output is 15:

```
# Calculate the total number of schools.  
school_count = school_data_df["school_name"].count()  
school_count  
  
15
```

We can't use the `count()` method on the `school_data_complete_df["school_name"]` column because this would give us a value of 39,170. If we want to use `school_data_complete_df`, we first need to get the unique items in the `["school_name"]` column by using the `unique()` method. This method will return a "ndarray", or n-dimensional array of all the unique values of that column.

Add the following code to a new cell and run the cell. The output is shown below.

```
# Calculate the total number of schools  
school_count_2 = school_data_complete_df["school_name"].unique()  
school_count_2
```

```
array(['Huang High School', 'Figueroa High School', 'Shelton High School',  
      'Hernandez High School', 'Griffin High School',  
      'Wilson High School', 'Cabrera High School', 'Bailey High School',  
      'Holden High School', 'Pena High School', 'Wright High School',  
      'Rodriguez High School', 'Johnson High School', 'Ford High School',  
      'Thomas High School'], dtype=object)
```

The output of using the `unique()` method on a Series is a ndarray of all the high schools—and you know how to get the number of items in an array!



NOTE

For more information, see the [Pandas documentation on the unique\(\) method](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.unique.html) [_](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.unique.html)(<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.unique.html>)[_](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.unique.html).

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.