**Name -Monika Gariya**

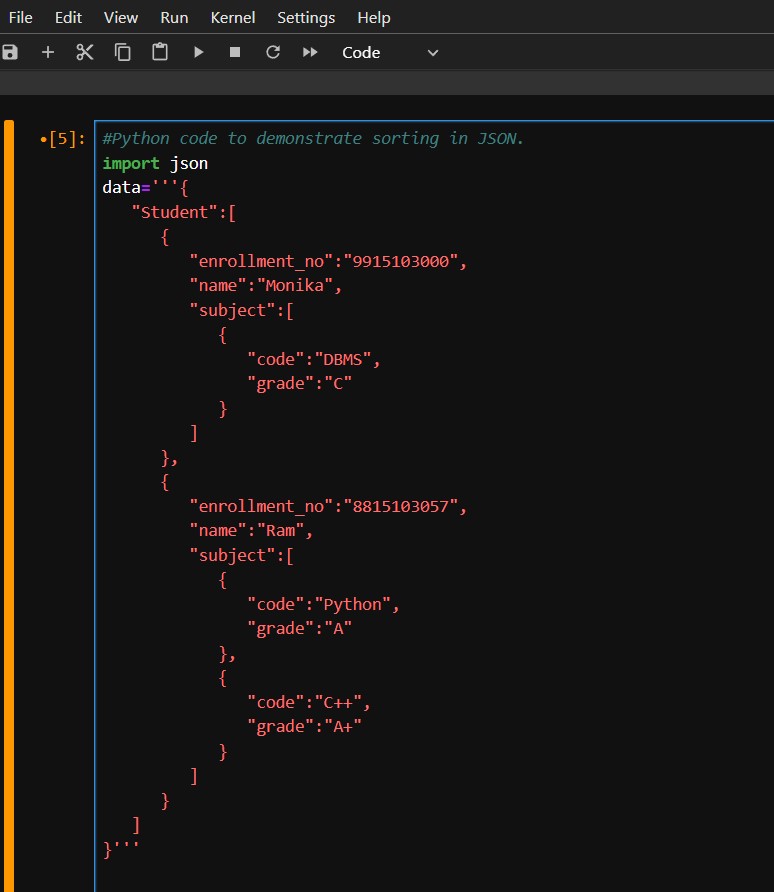
**Email-** [**monikagariya2023@gmail.com**](mailto:monikagariya2023@gmail.com)

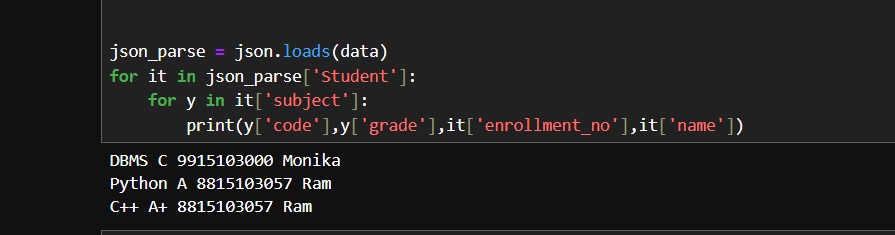
**Data Engineering Batch 1**

**Date – 01-02-2024**

**Topic** – **Sorting in JSON, tuple****, Enriching data using NumPy & Pandas, Different Types of Joins in Pandas**

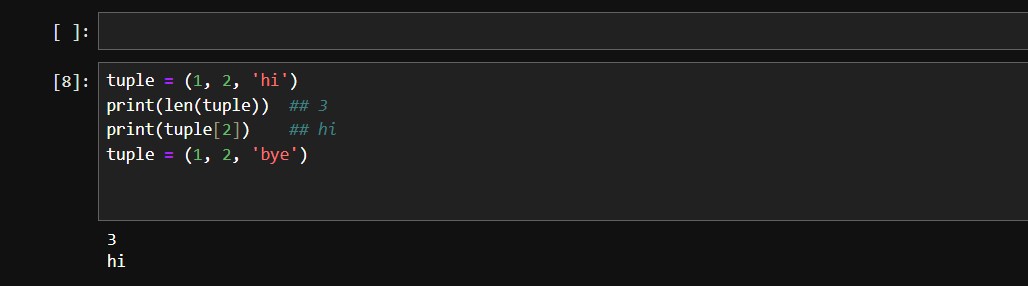
**1)Python code to demonstrate sorting in JSON**.





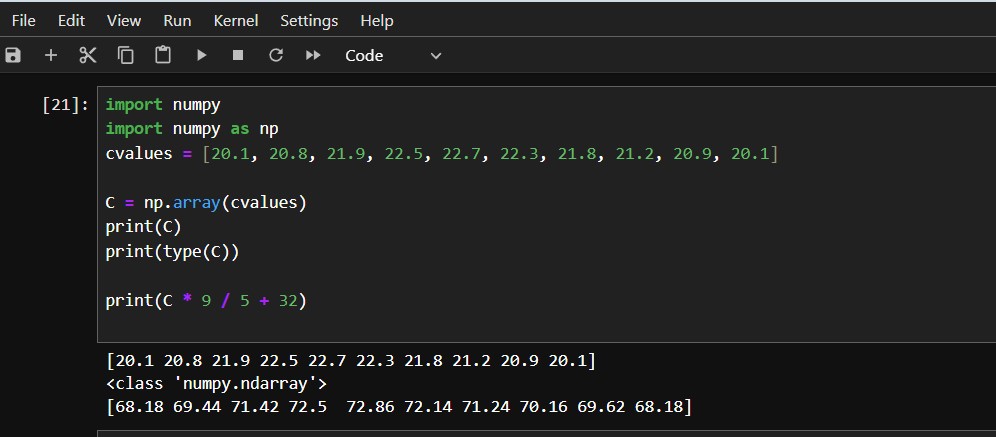
**2)Tuple**

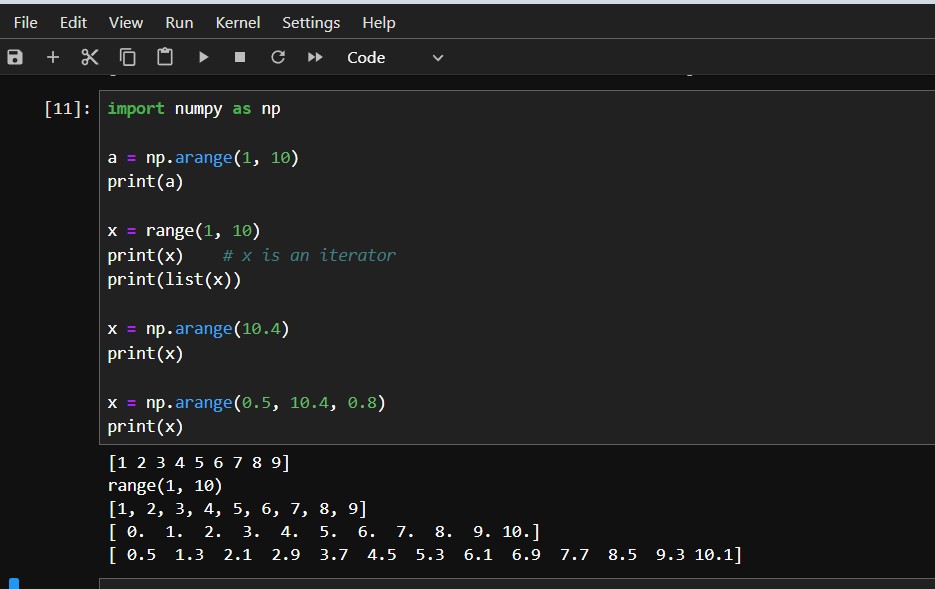
Tuples are used to store multiple items in a single variable. Tuple is one of 4 built-in data types in Python used to store collections of data, the other 3 are [List](https://www.w3schools.com/python/python_lists.asp), [Set](https://www.w3schools.com/python/python_sets.asp), and [Dictionary](https://www.w3schools.com/python/python_dictionaries.asp), all with different qualities and usage. A tuple is a collection which is ordered and unchangeable. Tuples are written with round brackets.

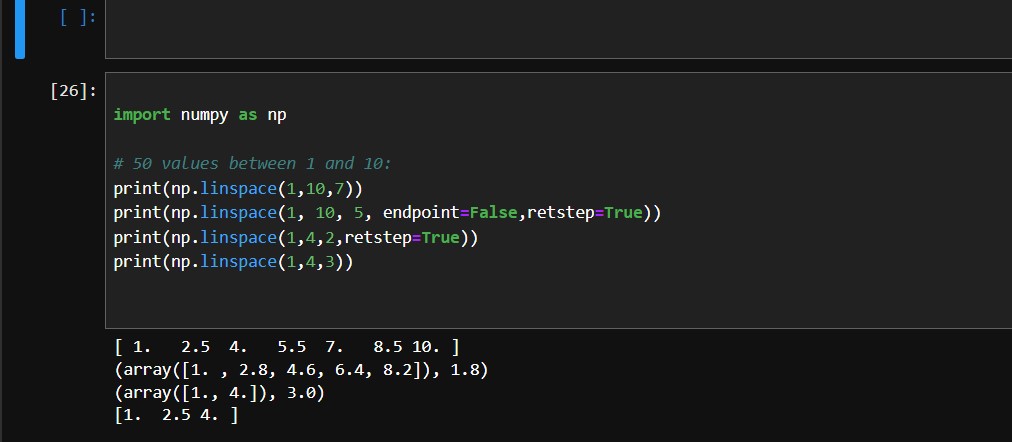


**3) Enriching data using NumPy & Pandas.**

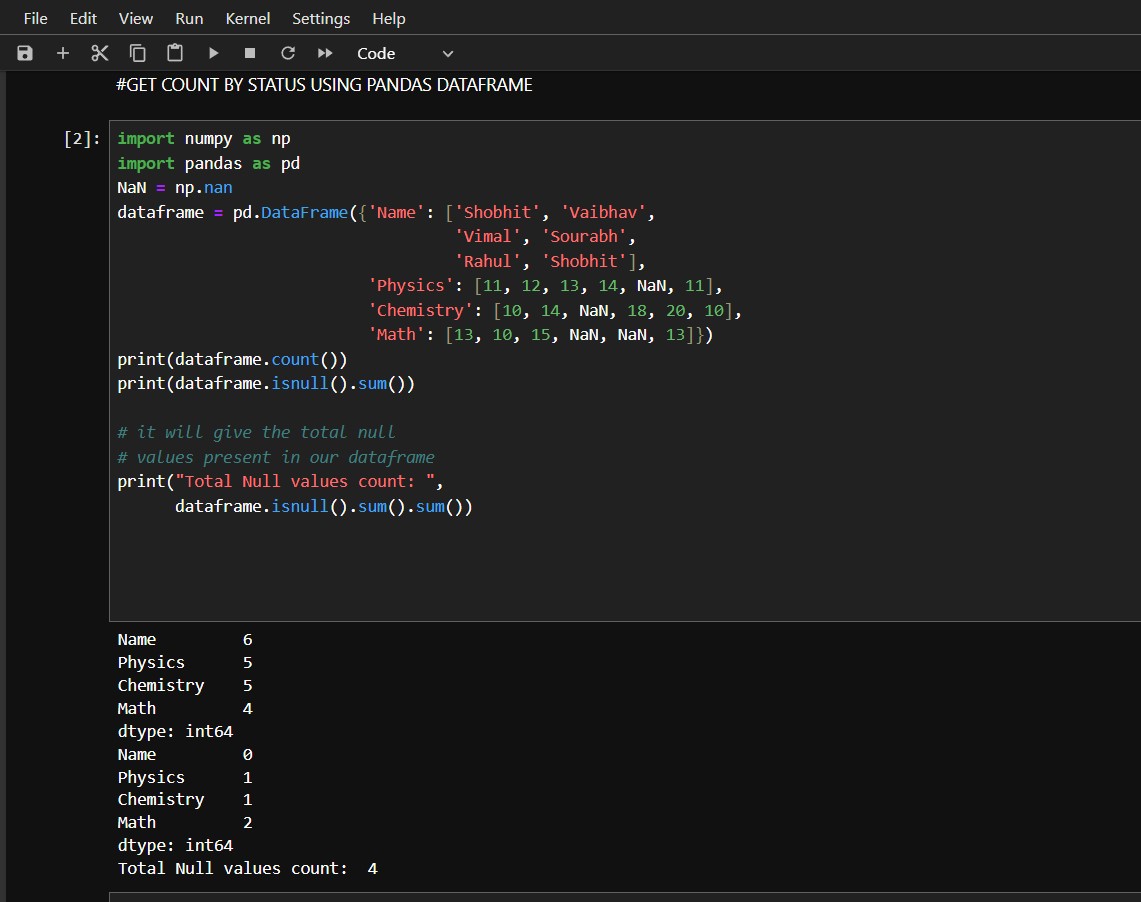
**1)NumPy** - NumPy stands for Numerical Python. NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, Fourier transform, and matrices. It is an open-source project and you can use it freely.



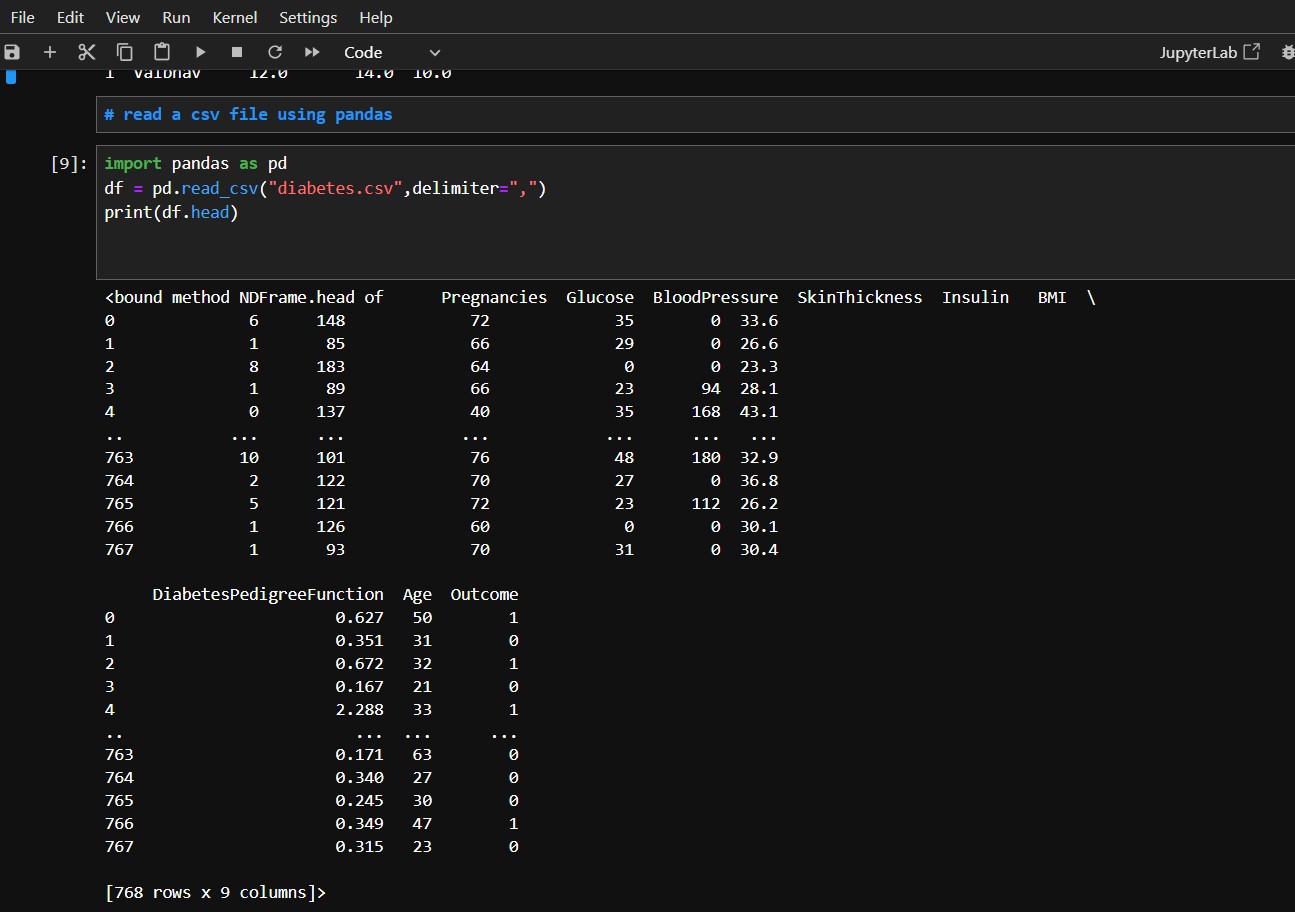




GET COUNT STATUS BY PANDAS DATAFRAME



READ A CSV FILE USING PANDAS



**Different Types of Joins in Pandas**

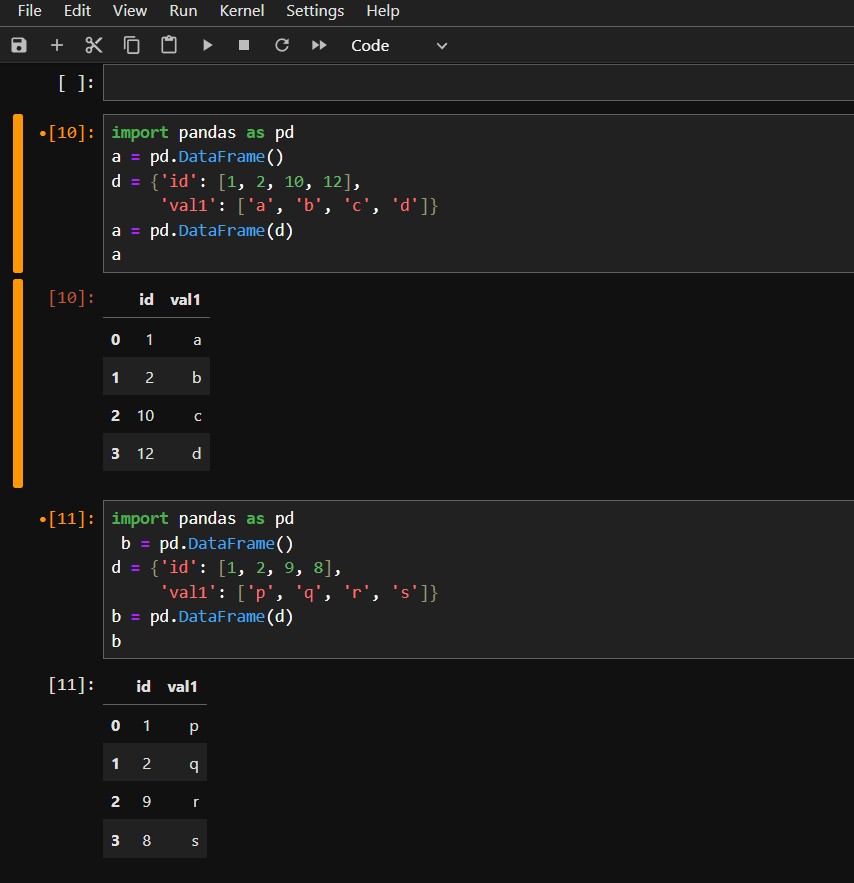
There are five types of Joins in [Pandas](https://www.geeksforgeeks.org/pandas-tutorial/).

* Inner Join
* Left Outer Join
* Right Outer Join
* Full Outer Join or simply Outer Join
* Index Join

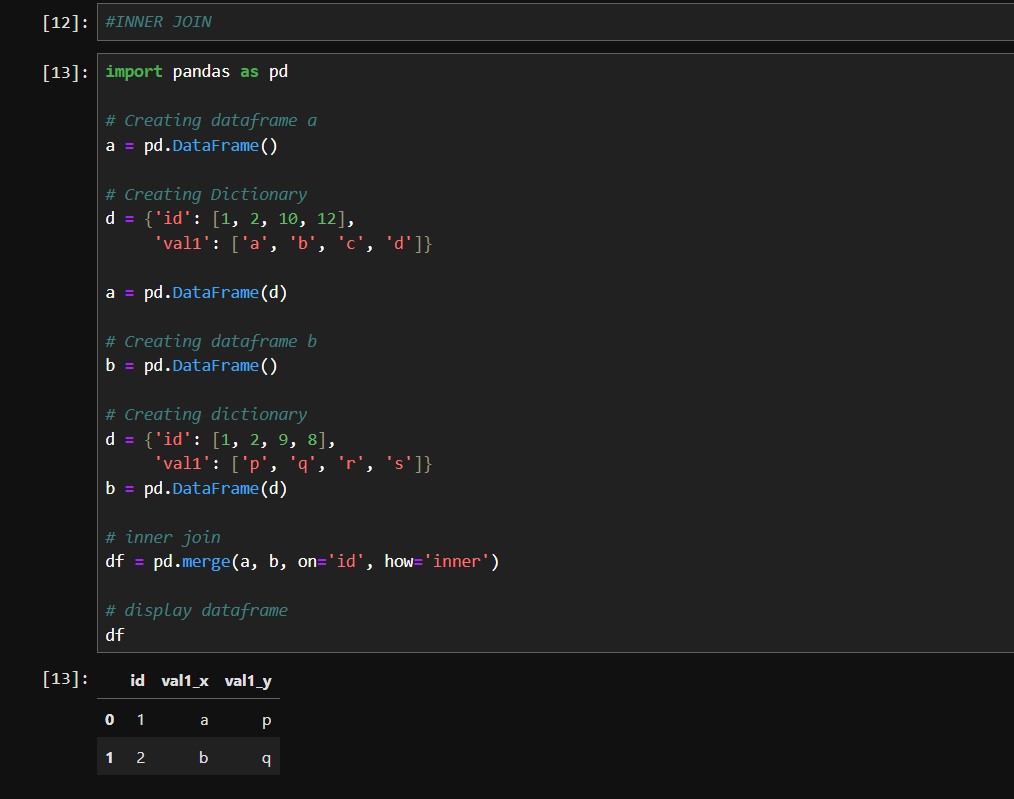
**1)Pandas Inner Join**

Inner join is the most common type of join you’ll be working with. It returns a Data frame with only those rows that have common characteristics. This is similar to the intersection of two sets.

Suppose we have two tables a, b:

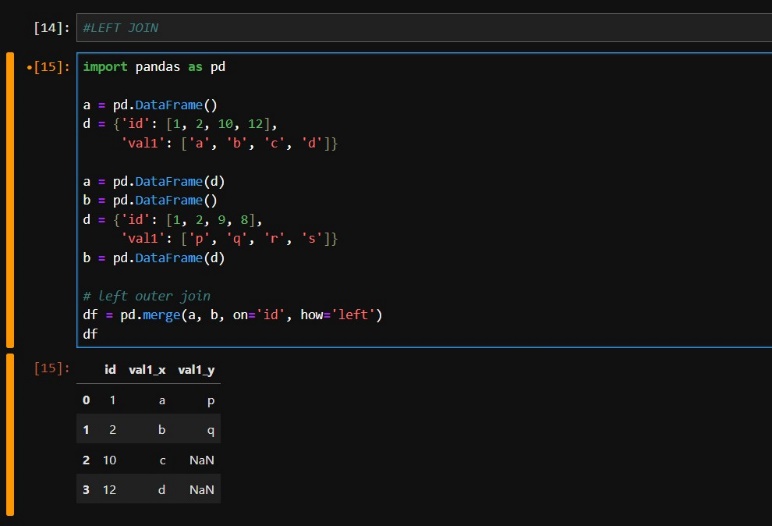


Now Applying Inner join in a and b table we got the result:



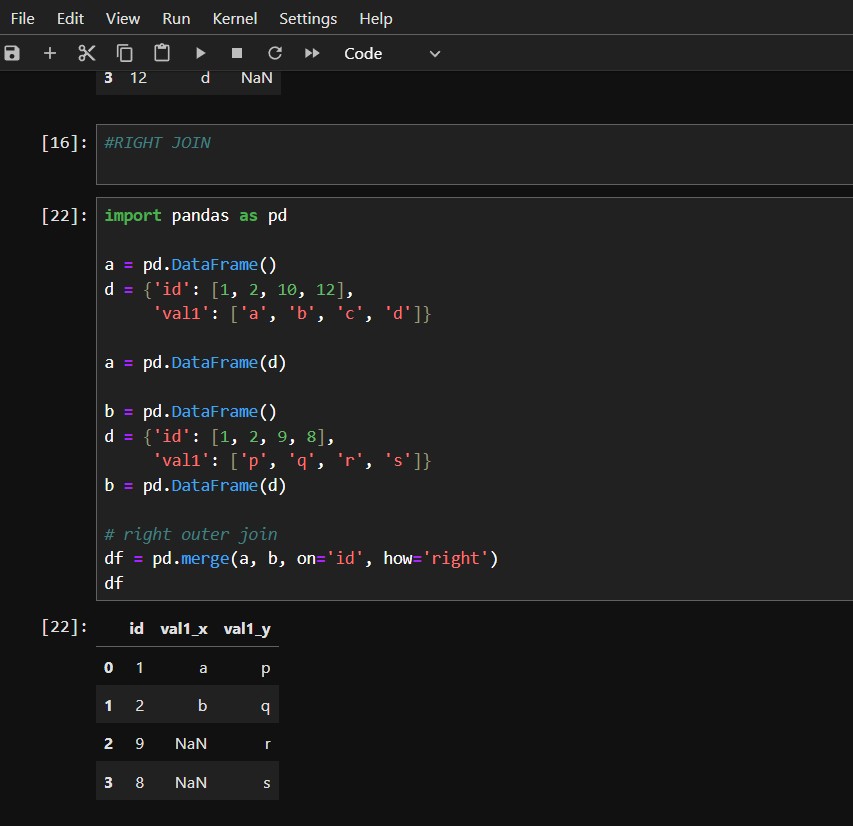
**2)Pandas Left Join**

With a left outer join, all the records from the first Dataframe will be displayed, irrespective of whether the keys in the first Dataframe can be found in the second Dataframe. Whereas, for the second Dataframe, only the records with the keys in the second Dataframe that can be found in the first Dataframe will be displayed.



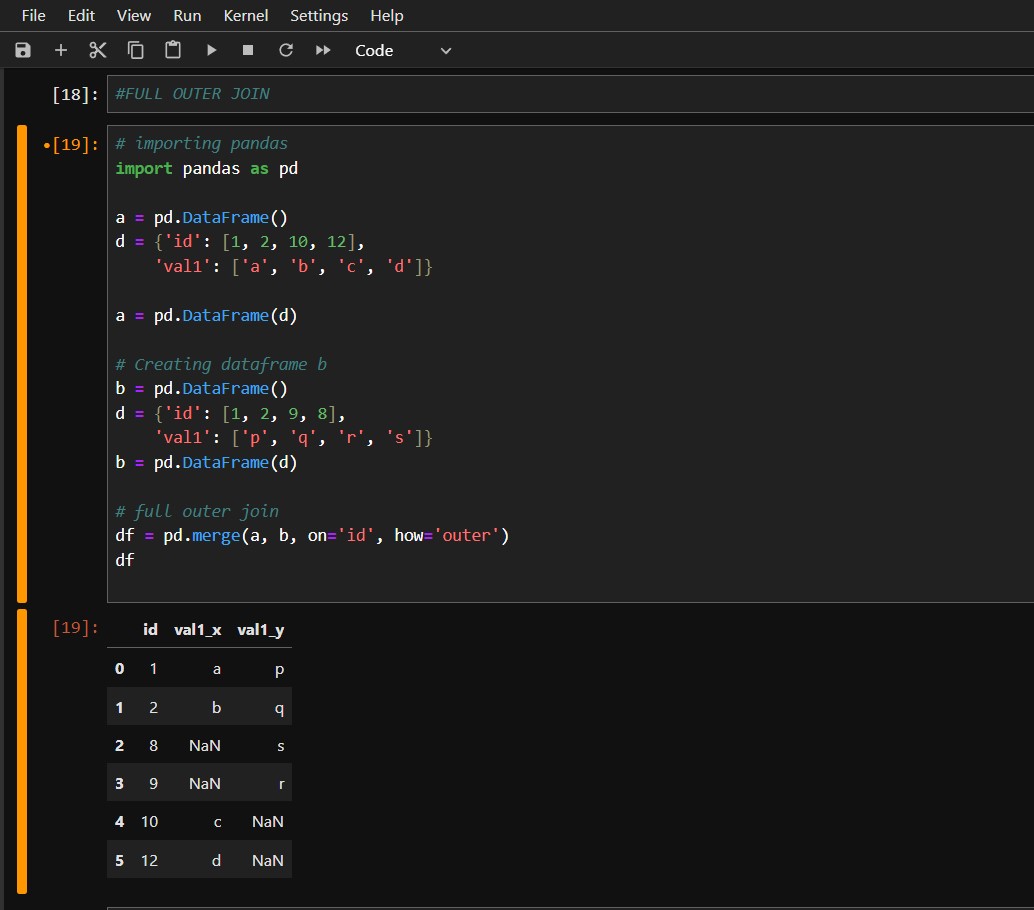
**Pandas Right Outer Join**

For a right join, all the records from the second Dataframe will be displayed. However, only the records with the keys in the first Dataframe that can be found in the second Dataframe will be displayed.



**Pandas Full Outer Join**

A full outer join returns all the rows from the left Dataframe, and all the rows from the right Dataframe, and matches up rows where possible, with NaN elsewhere. But if the Dataframe is complete, then we get the same output.



**Pandas Index Join**

To merge the Dataframe on indices pass left\_index and right\_index arguments as True i.e. both the Dataframes are merged on an index using default Inner Join.

