**Coding Assessment Spark**

**Name : Jeevan Sai Badana**

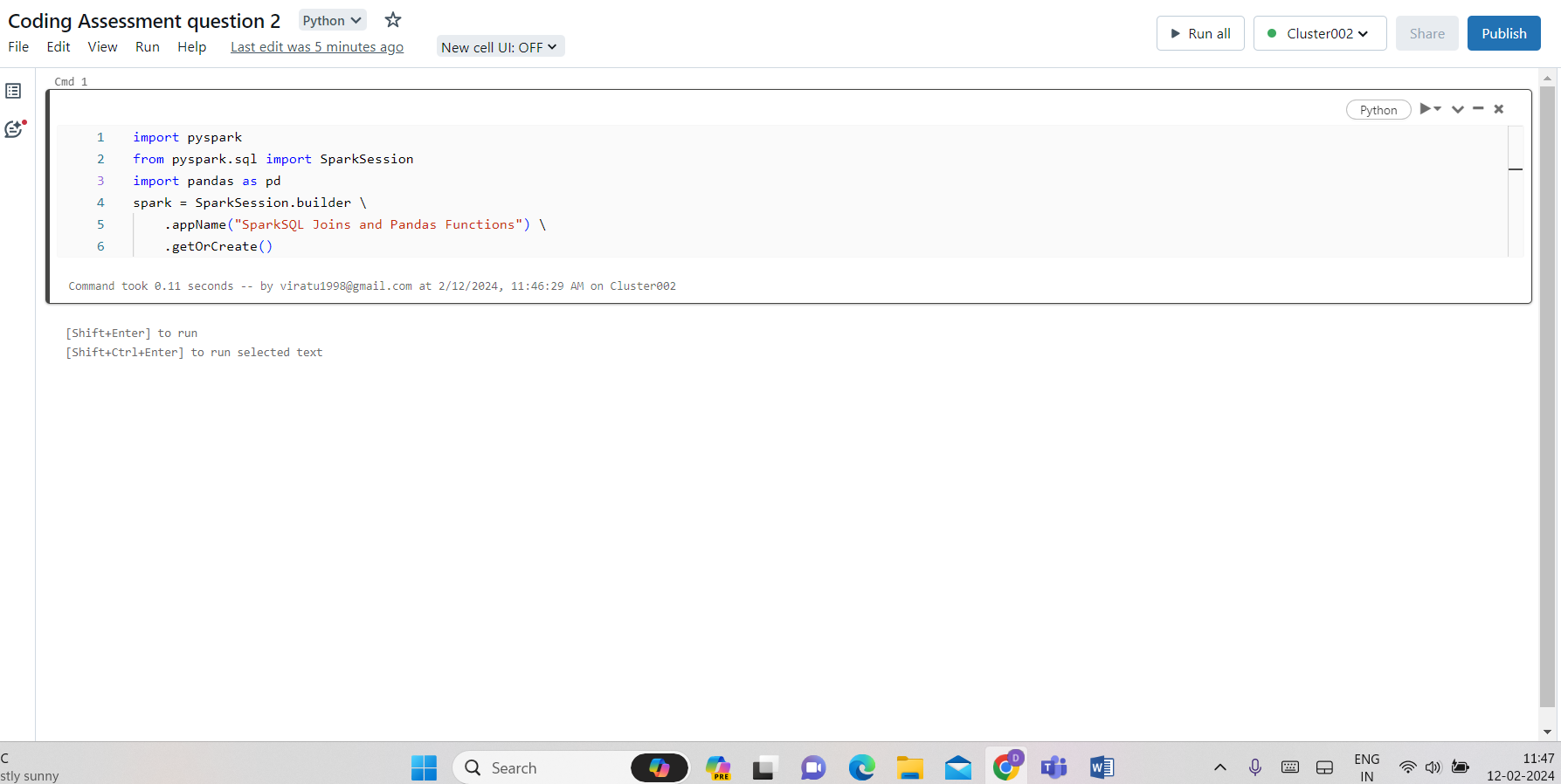
**Mail:** [jeevansai100@gmail.com](mailto:jeevansai100@gmail.com)

DATE: 12-02-2024

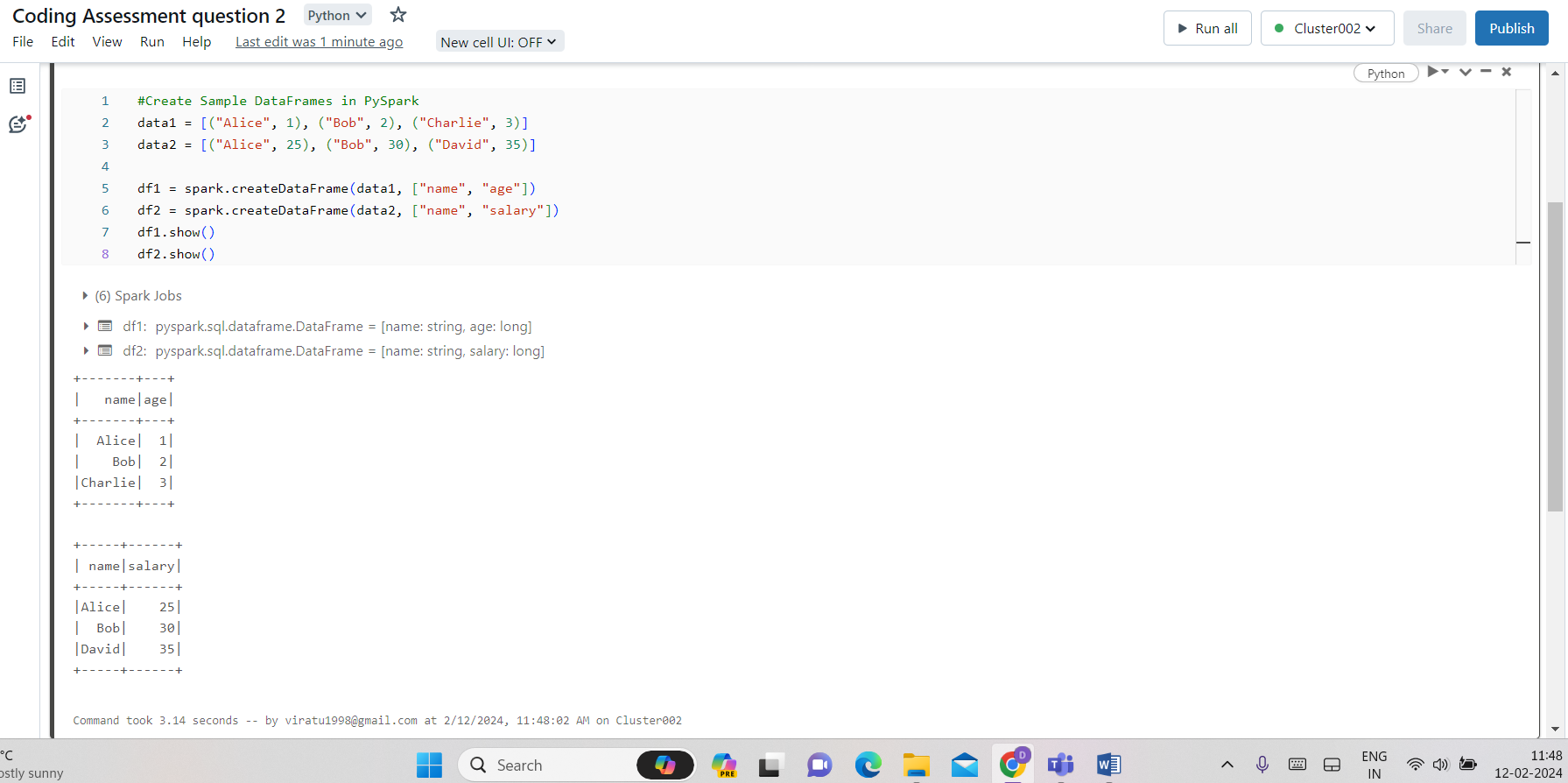
Question number 2

**Execute Pyspark -sparksql joins & Applying Functions in a Pandas DataFrame**

First we import pyspark and spark sessions as below

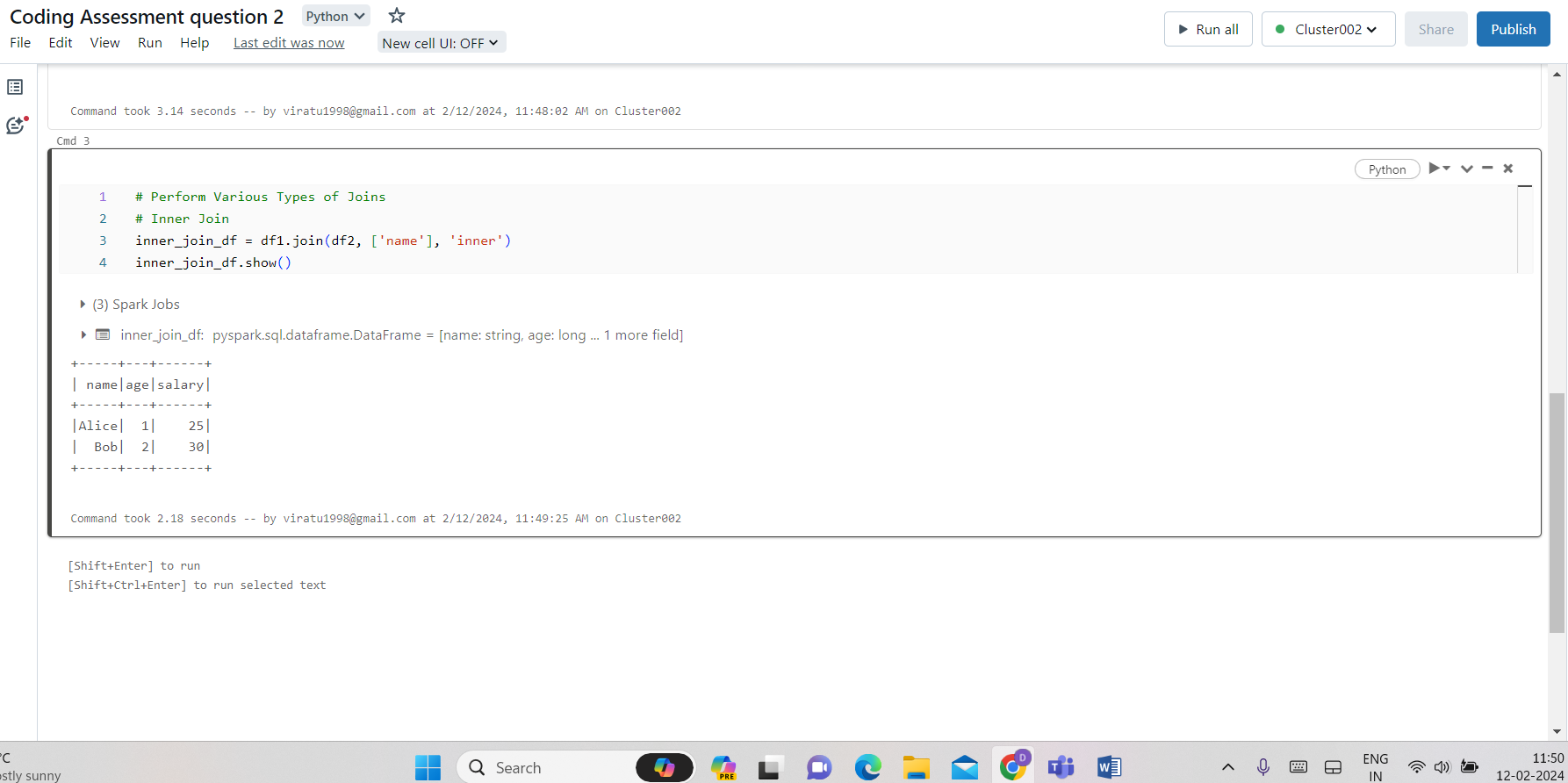


Now we create two data frames

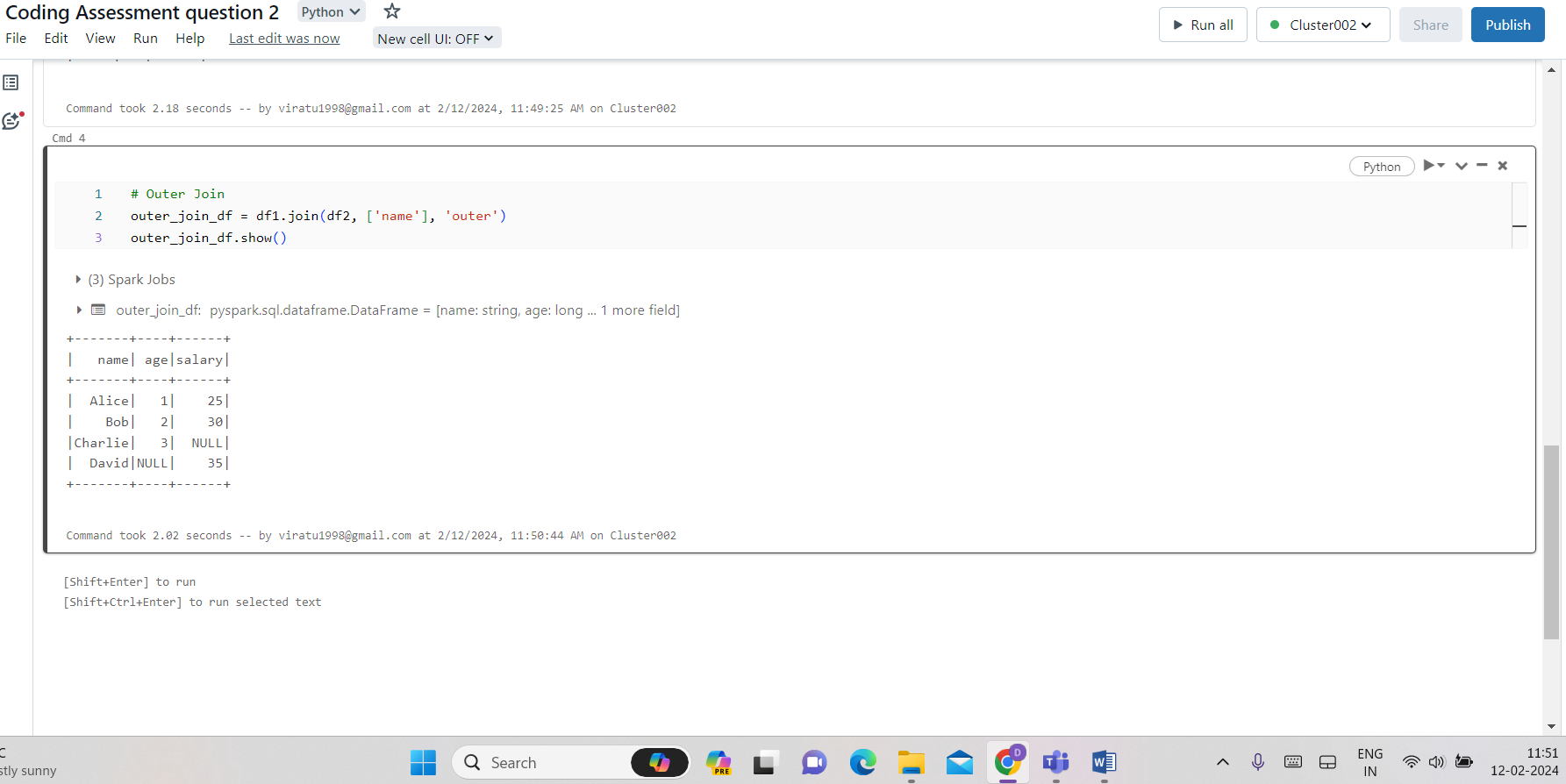


**Perfoming various joins**

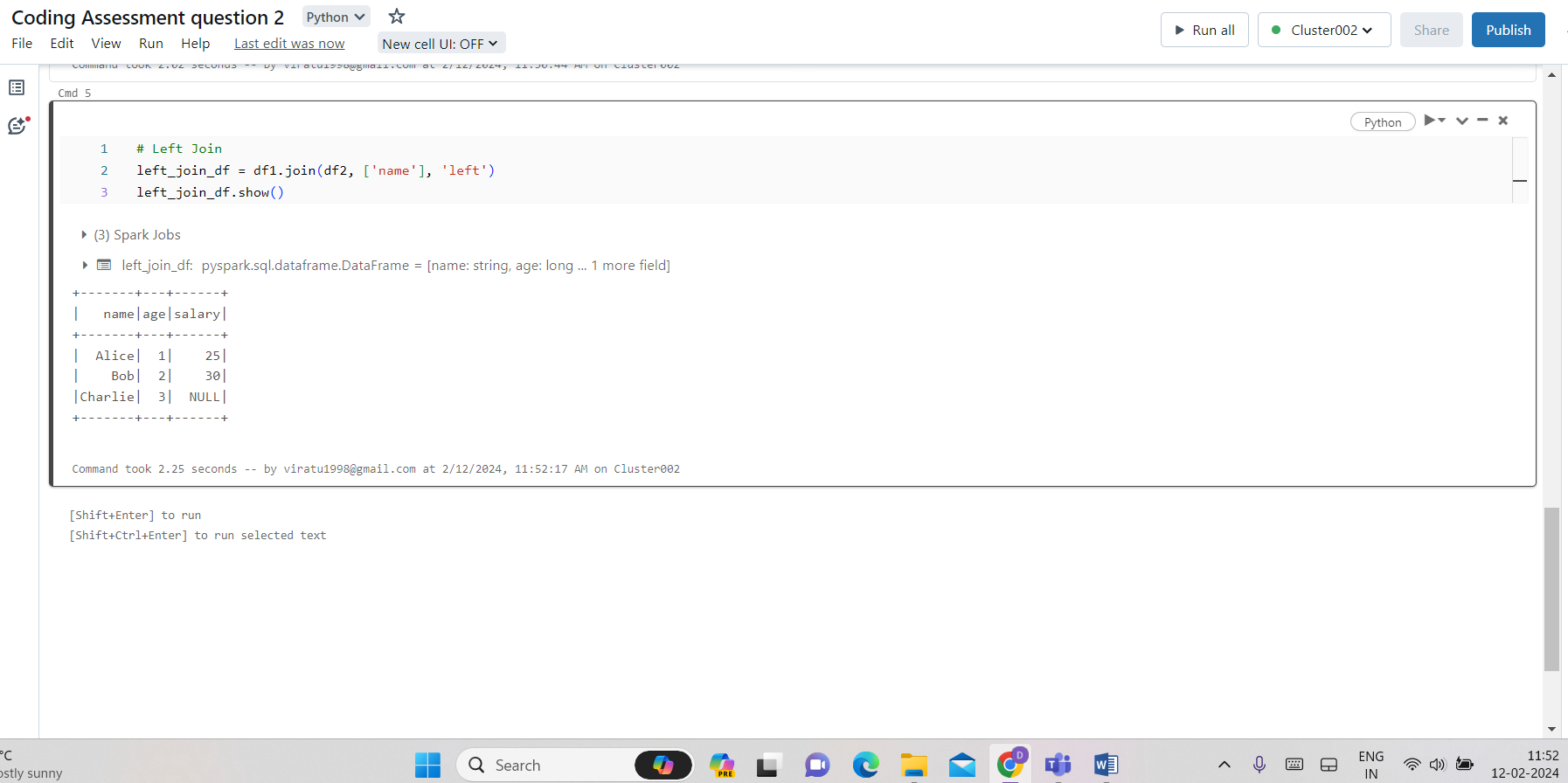
Inner join



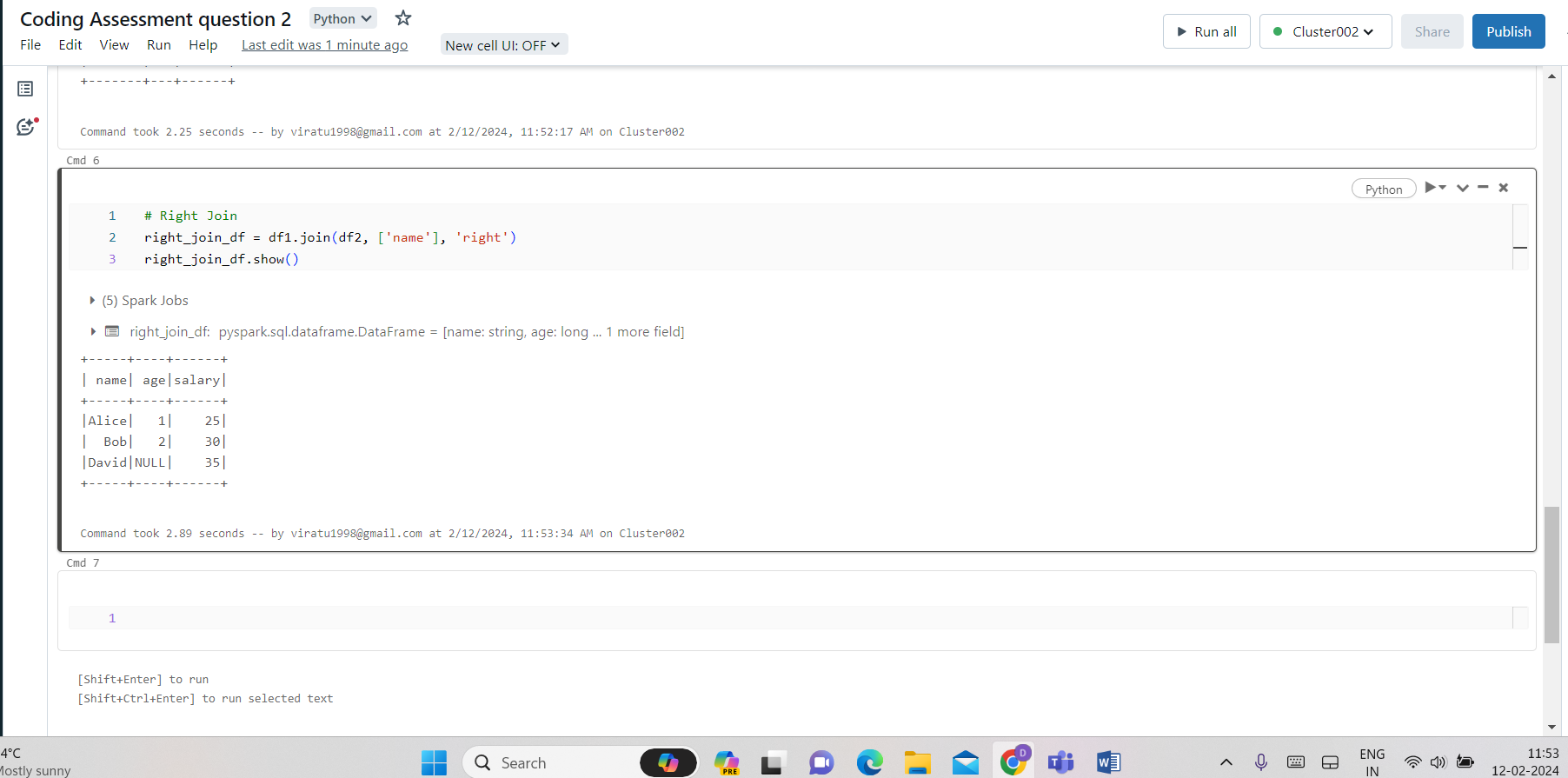
Outer join:



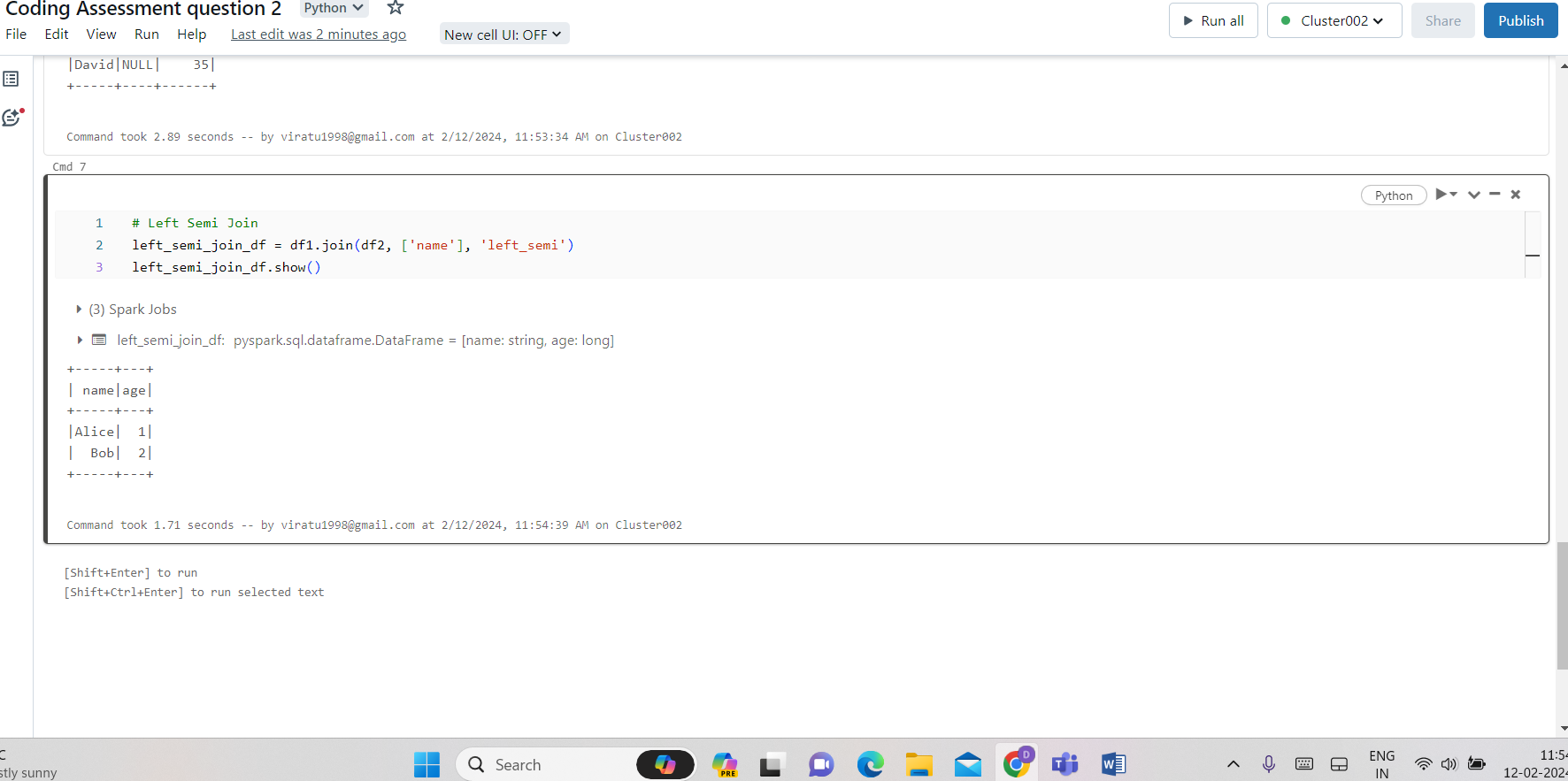
Left join:



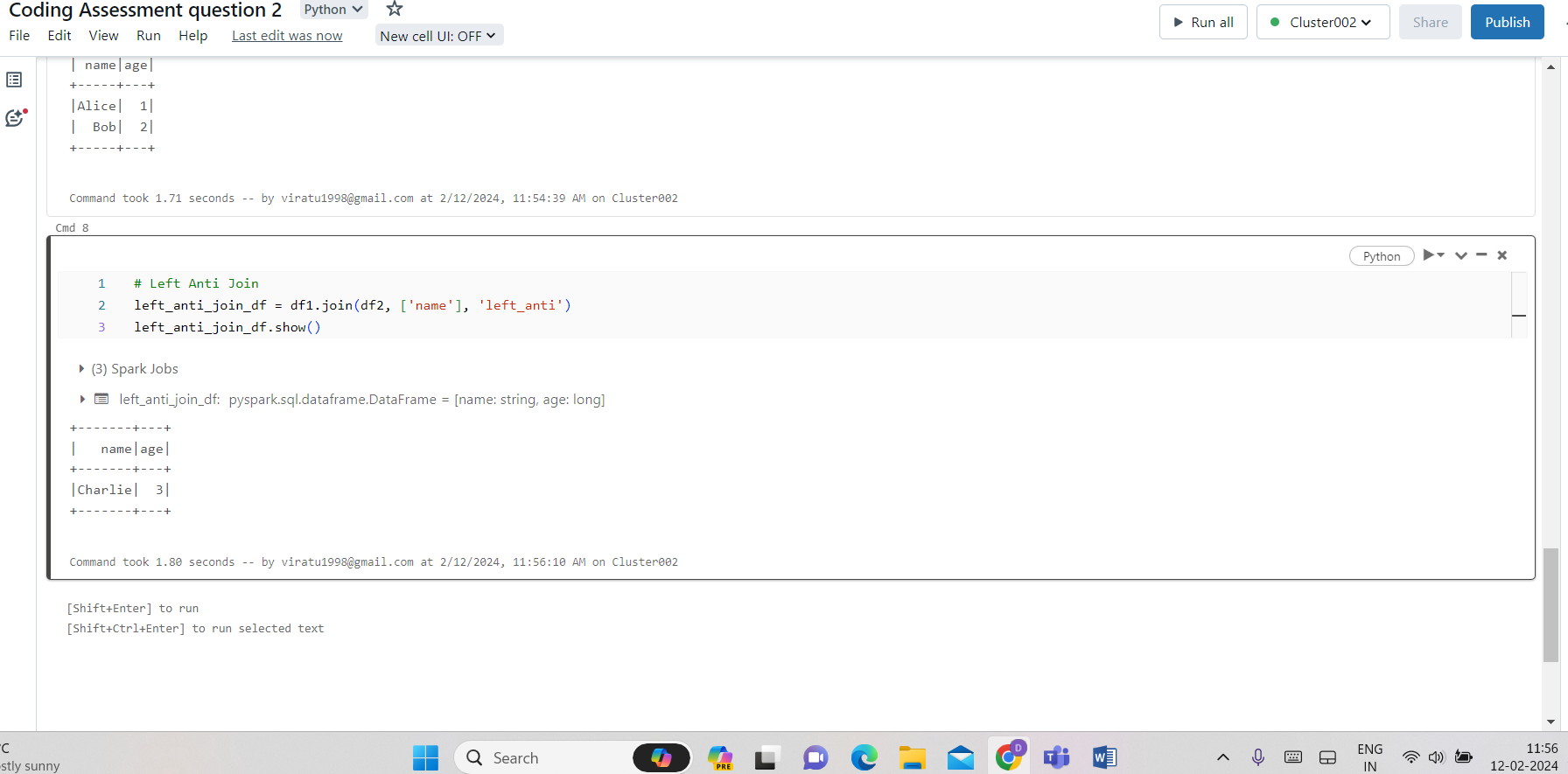
Right join:



Left semi join:



Left anti join:



Convert PySpark DataFrame to Pandas DataFrame and Apply Function:

