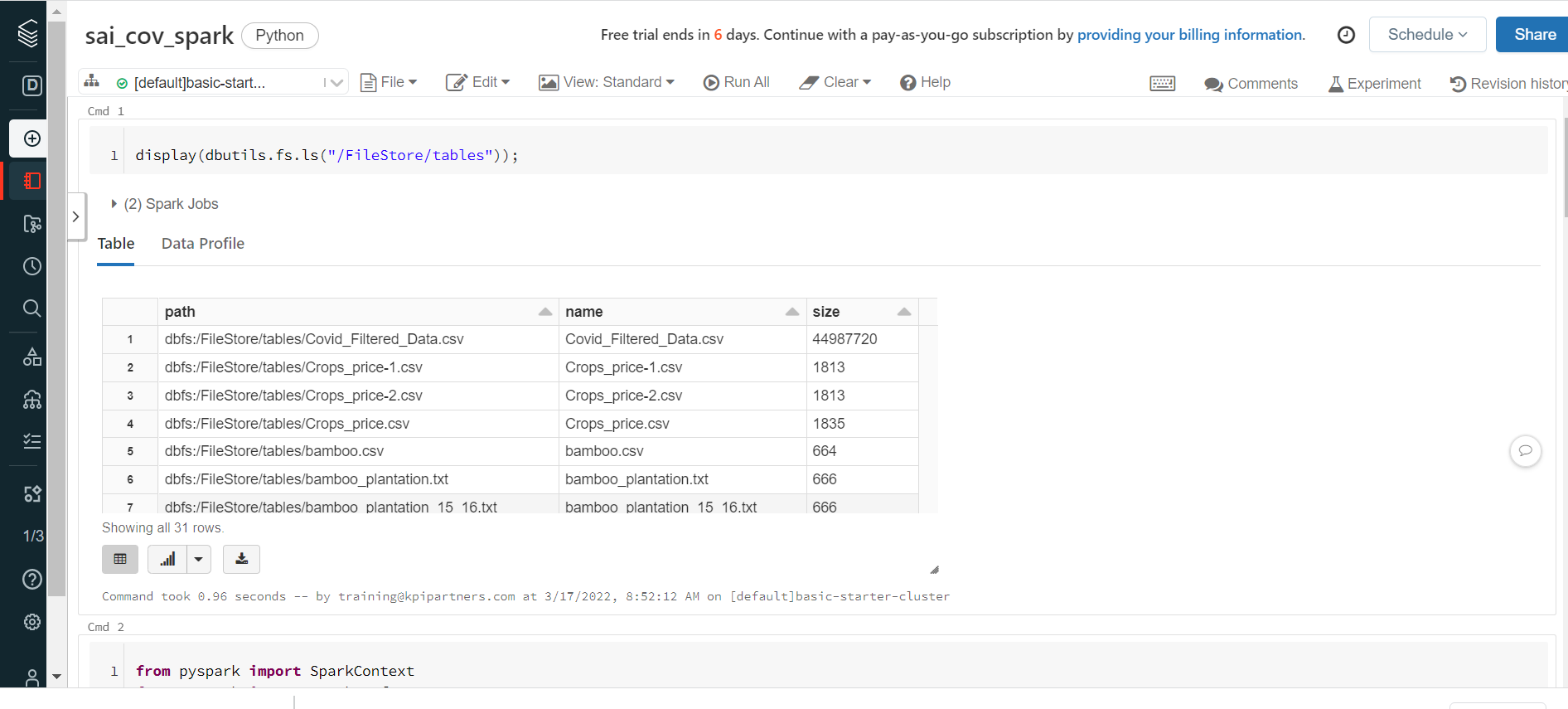
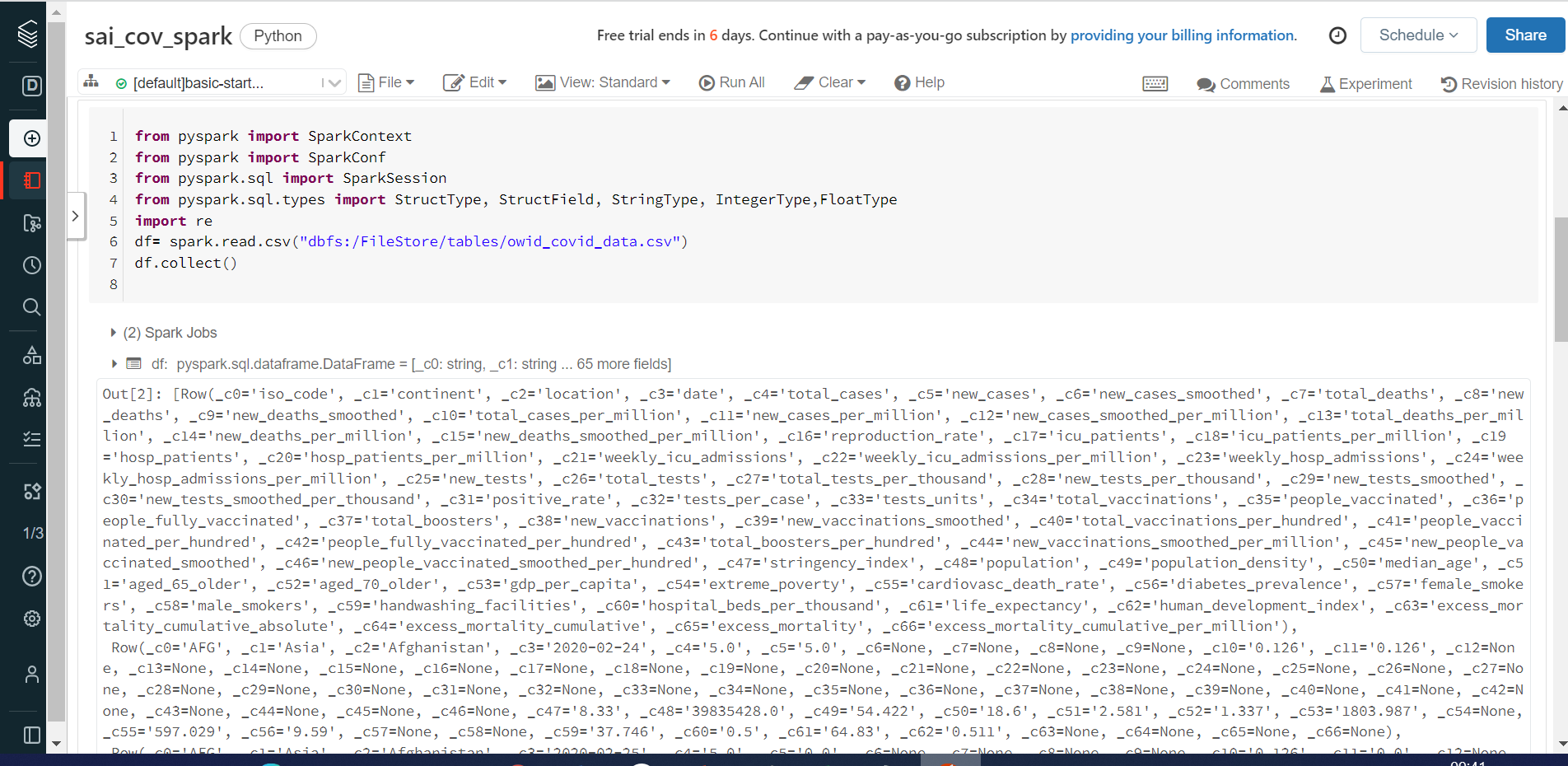
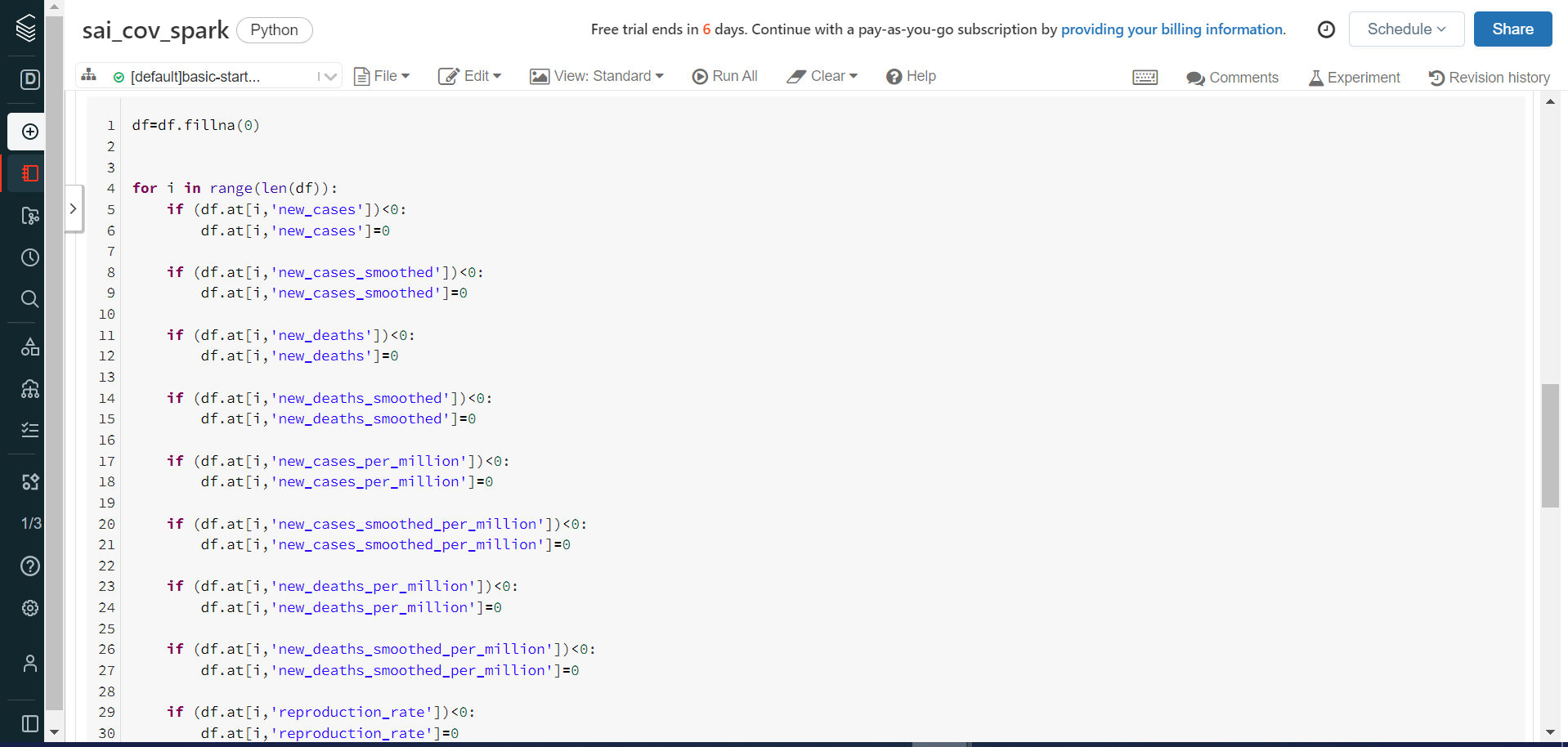
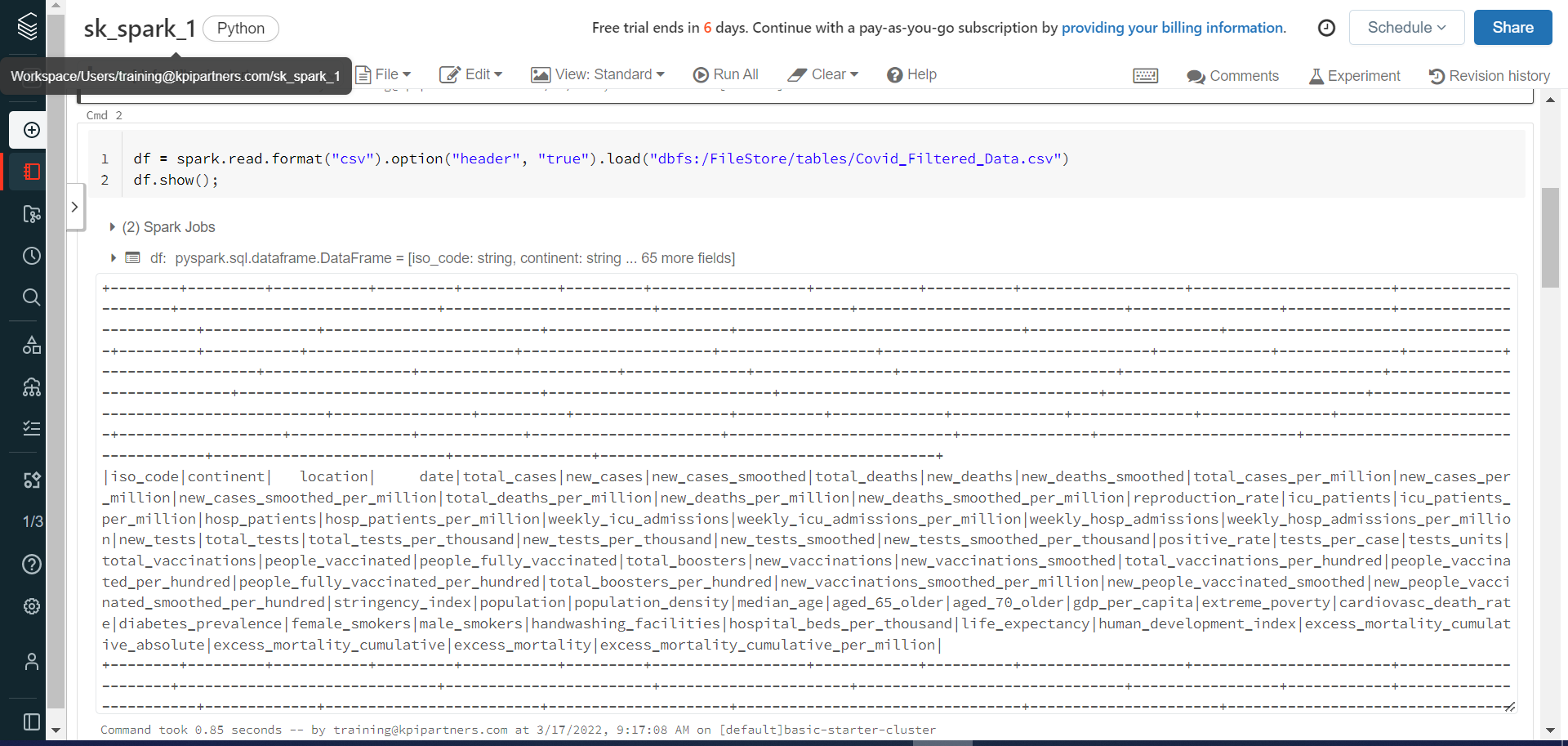
**COVID USECASE USING SPARK,SCALA AND SQL**

****



1. Replace all the NEGATIVE and NULL valued records with 0



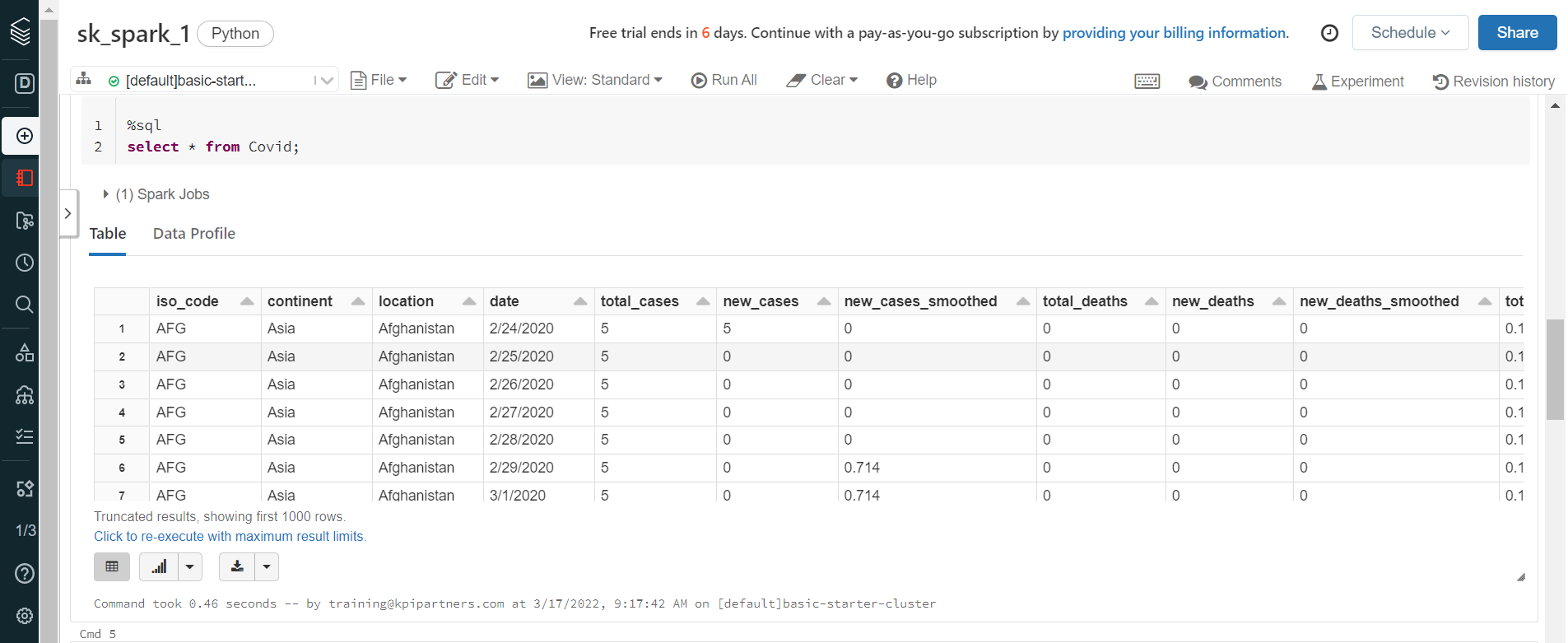
­­­

2.Remove the records where the continent column having blanks.

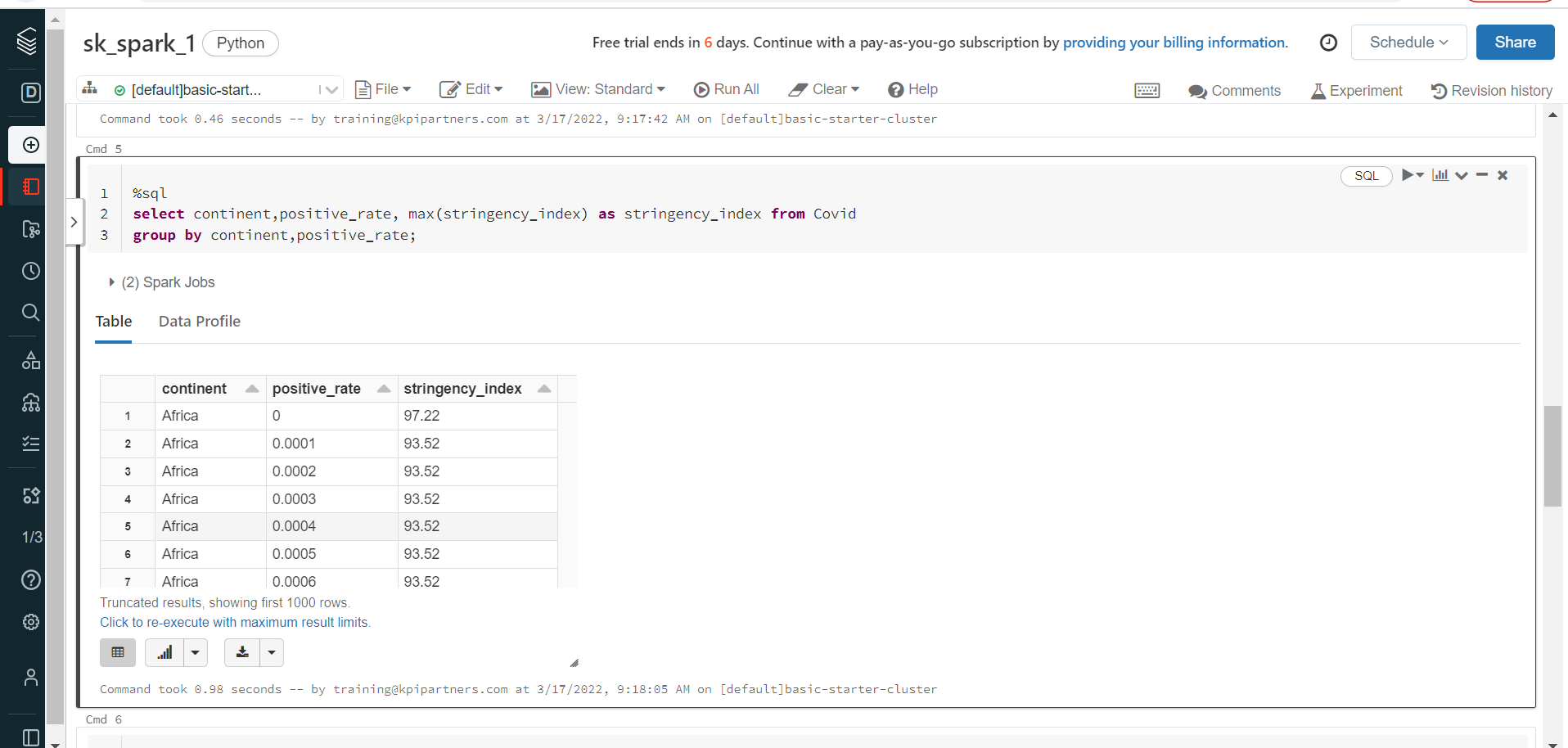
for i in range(len(df)):

if (df.at[i,'continent'])==0 :

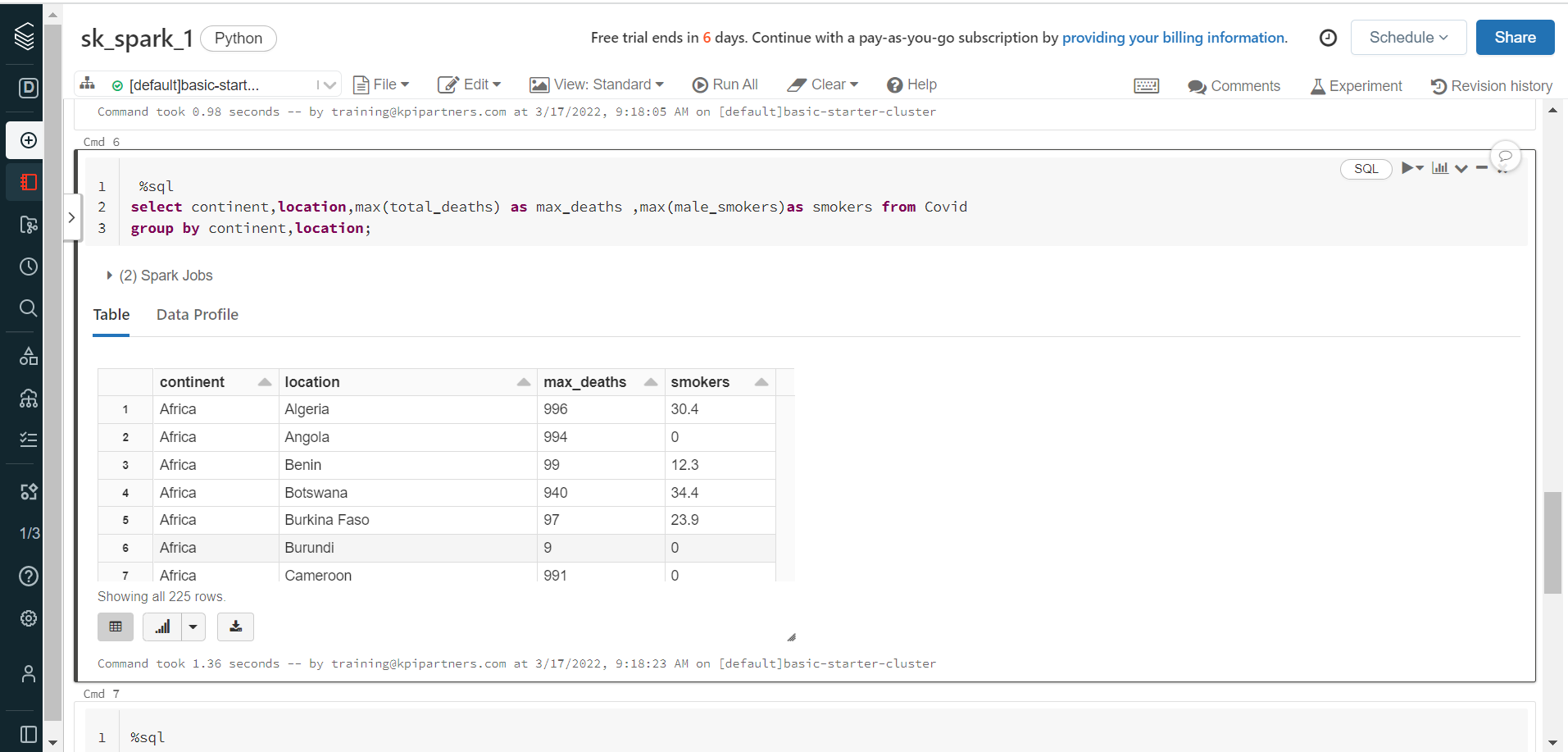
df.drop([i], axis=0, inplace=True)



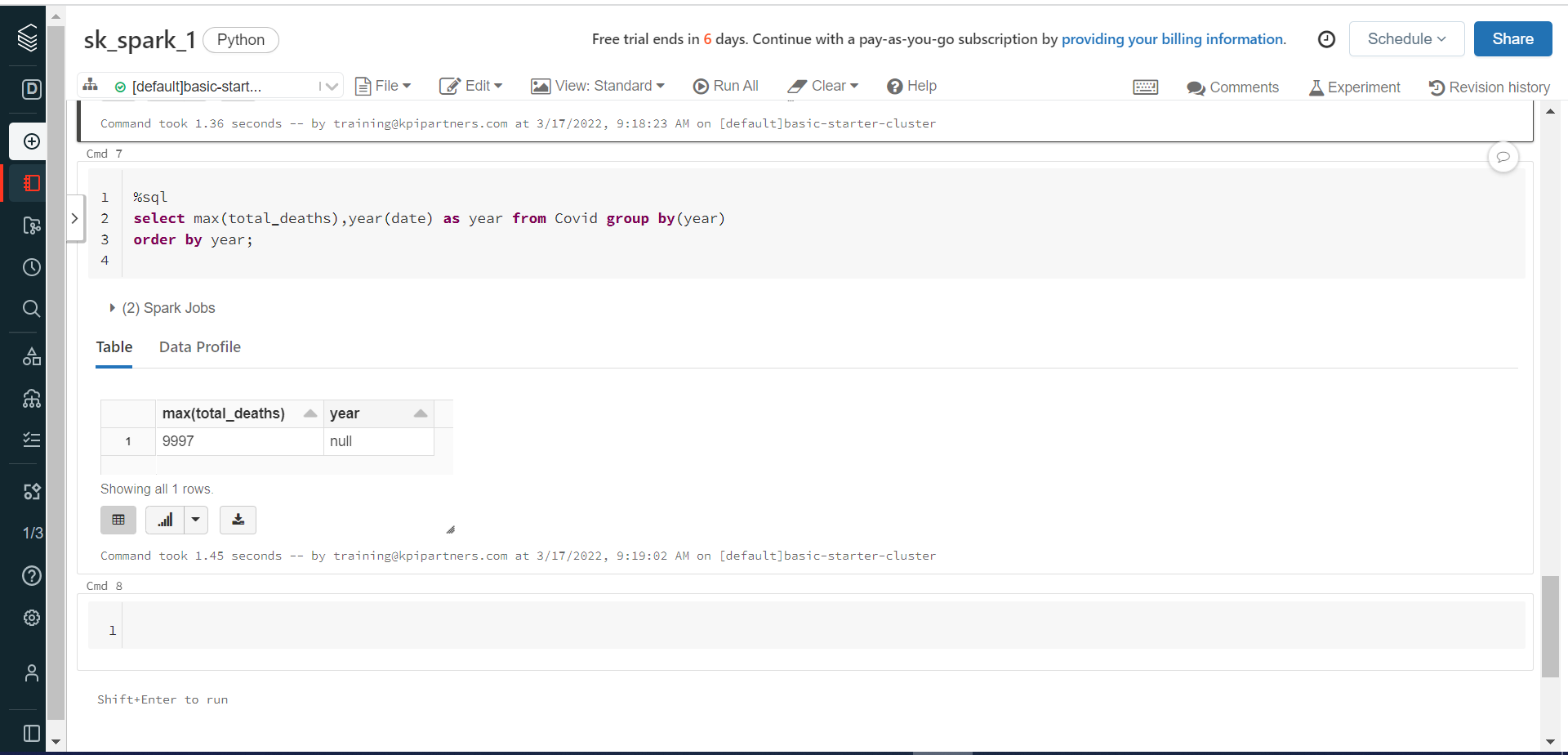
3. Display the continent,positivity rate where the strigency index is maximum



4. Which continent and location having max of deaths and male\_smokers

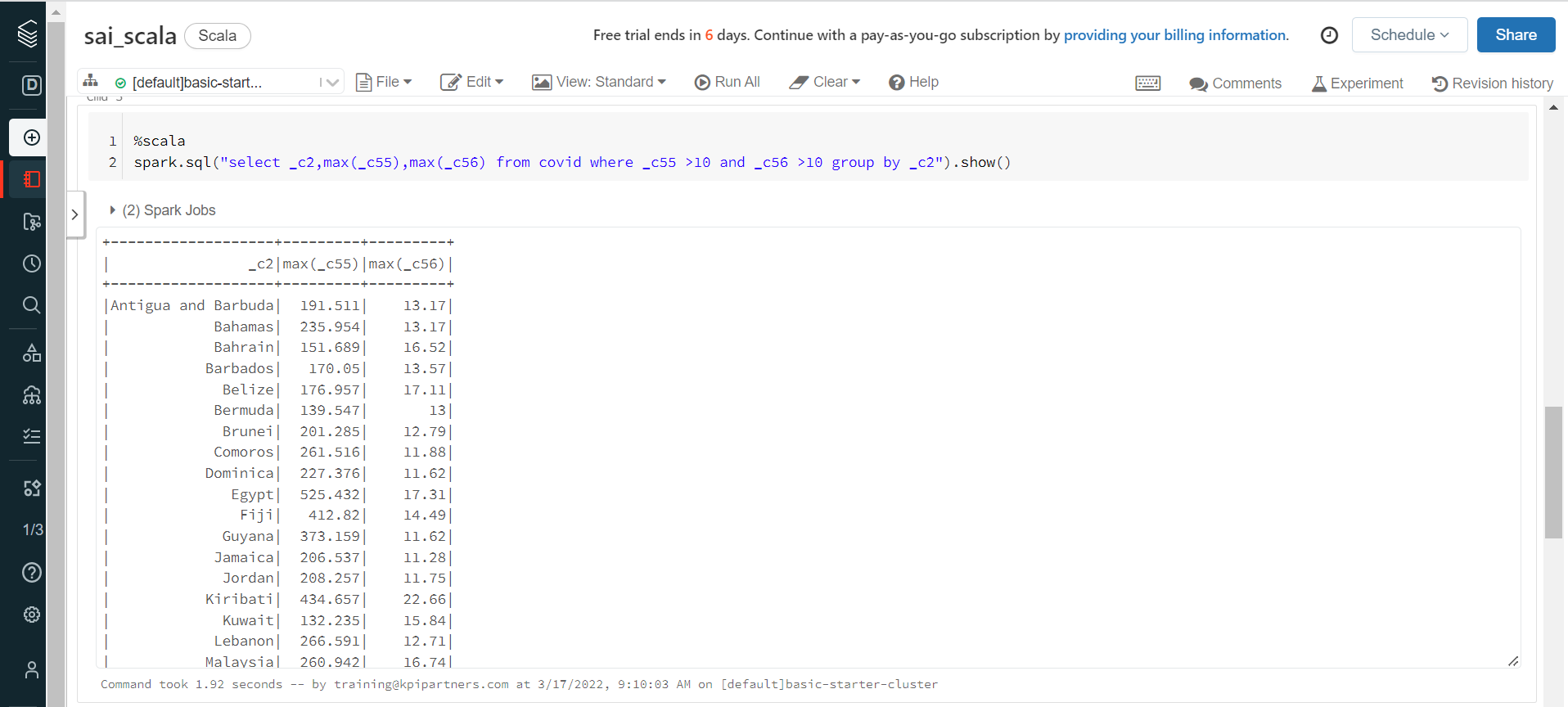


5. Which year having the highest number of deaths



6. Which location having lowest recovery rate for each continent

7. Countries with max cardiovasc\_death\_rate and diabetes\_prevalence greater than 10



8. Which location has highest number of vaccinated people in each continent.

