1. Write program to load the given dataset using pandas and create a column name “change” and the values of this column should be the difference between columns “Open” and “close”.

Format of the given dataset is:

Date,Open,High,Low,Close,Volume,Adj Close

7/14/2014,202,216,199.15,212.05,3014200,212.05

2. Calculate the “moving average” using Adj Close column.

[Hint: use rolling\_mean] and display last 10 records.

3. Now plot “Adj Close” and “mavg”. So that the plot should appear like below.



1. Search for the **Join** module, and drag it to the canvas below both the **Diabetic Data** dataset and the **Admission Mapping** dataset. Then connect the output ports from the **Diabetic Data** dataset and the **Admission Mapping** dataset to the **Dataset1** and **Dataset2** input ports of the **Join** module respectively.
2. Select the **Join** module, and in the **Properties** pane, set the following properties:
   1. **Join key columns for L**: Launch the column selector and select **admission\_type\_id**
   2. **Join key columns for R**: Launch the column selector and select **admission\_type\_id**
   3. **Join type**: Left Outer Join
   4. **Keep right key column**: *Unselected*

**Note:** The key columns in this example have the same name, but this is not a requirement to join columns from two datasets. By using a left outer join, the **Join** module will retain any rows in the left dataset (the diabetic data) that do not have a matching **admission\_type\_id** column in the right dataset (the admissions mapping data).