**ind\_actividad\_cliente**

Replace them with 0 or replace them with a percentage of the 0’s and 1’s.

* Calculation: #0 / sum(0,1) and the multiply this by the nulls which will give us how many 0’s should replace the nulls

**segmento**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

**tiprel\_1mes**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

**fecha\_alta**

Find the middle date, given the range of dates.

**Sexo**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

**ind\_nuevo**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

**ind\_empleado**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

Note that there is a ‘S’, which need to considered to be a null value

Code: data\_df.loc[data\_df['ind\_empleado'].isnull()] = 'N'

**Renta**

We will either fill the null with mean/average/median/mode. Need to do some research

**Indext**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**

**Nomprov**

Deal with in a similar fashion as we did for **ind\_actividad\_cliente**