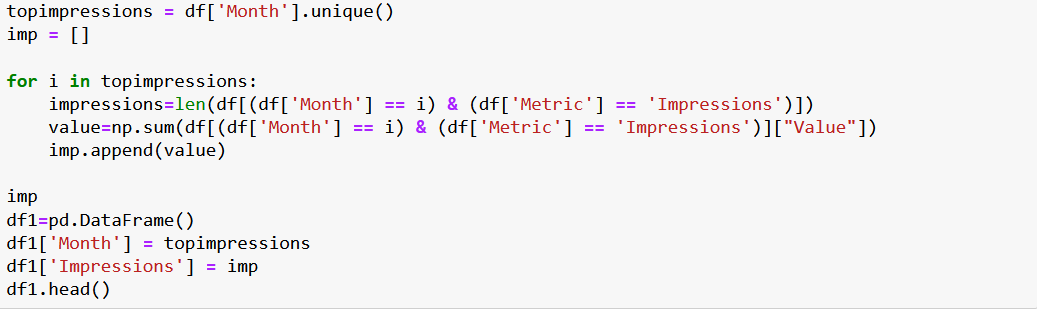
**Answers**

(In this file, I will just be attaching the code and the outputs of the questions asked, you could surely understand the methodology, which I used to solve these in the PY files.)

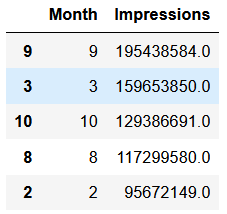
Campaign data

1. **Find out the top 5 months in terms of impressions**

After separating the month from the date we can easily find the top 5 month in terms of impression, The below is the codes

****

We get the results in unsorted formats, after sorting the values in descending format we get the below output.

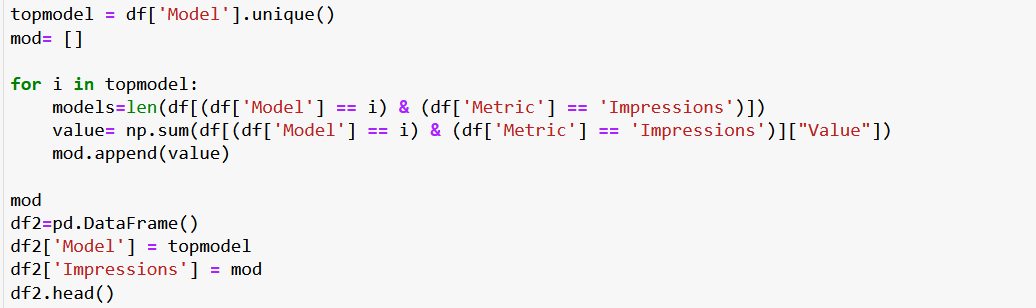
****

Here we can see that the top 5 months in terms of impressions are:

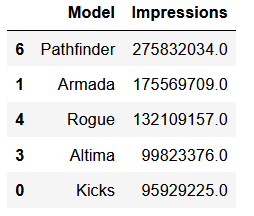
* 9
* 3
* 10
* 8
* 2

1. **Which model has maximum impressions?**

This is an very simple solution, Please go through the code below

****

We get the results in unsorted formats, after sorting the values in descending format we get the below output.

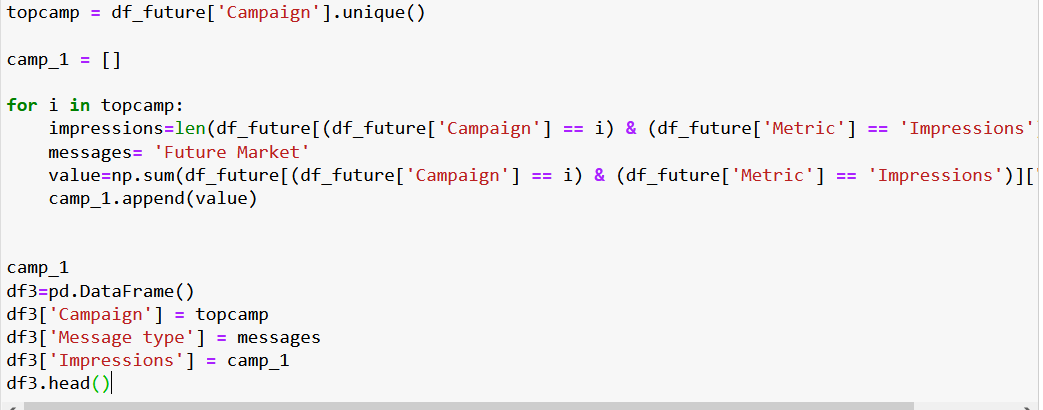


As you can see the model which has the maximum impressions is “Pathfinder”.

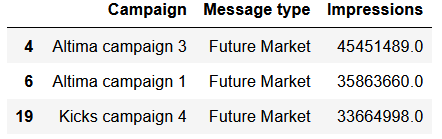
1. **Find out the top 3 campaigns in terms of impressions for every message type**

As we have two types of message types, “Future Market” and “Near Market” we can find the solutions separately.

**For “Future Market”**



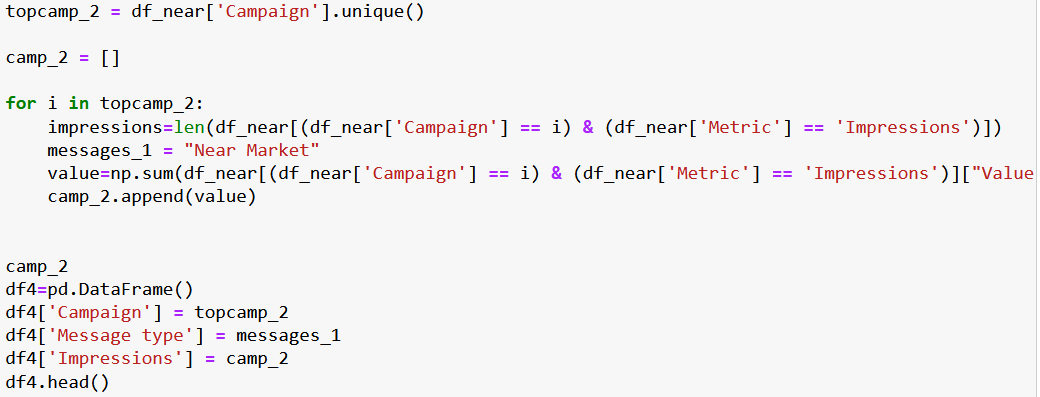
Hence the output will be



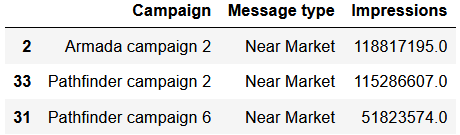
As you can see the top three campaigns in terms of impressions for future market are:

* Altima campaign 3
* Altima campaign 1
* Altima campaign 4

**For “Near Market”**



Hence the output will be

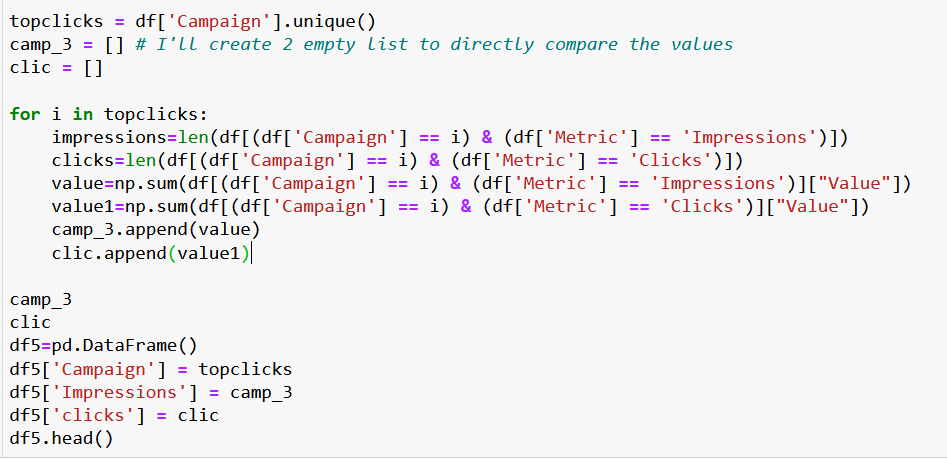
****

As you can see the top three campaigns in terms of impressions for near market are:

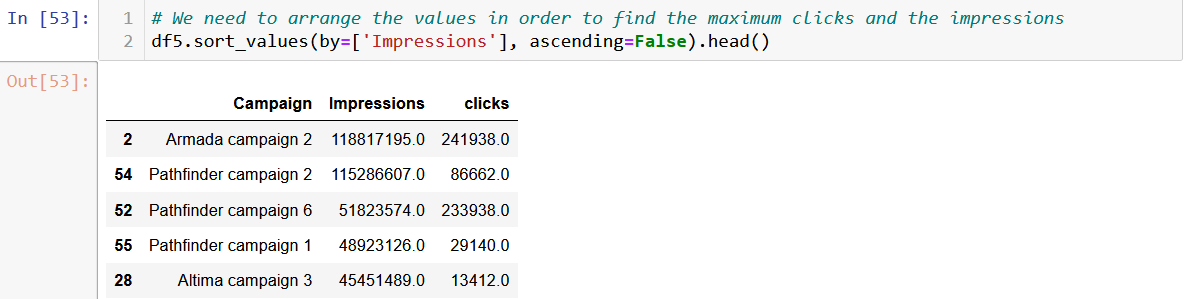
* Altima campaign 2
* Pathfinder campaign 2
* Pathfinder campaign 6

1. **Does the campaign with maximum impressions also have maximum clicks?**

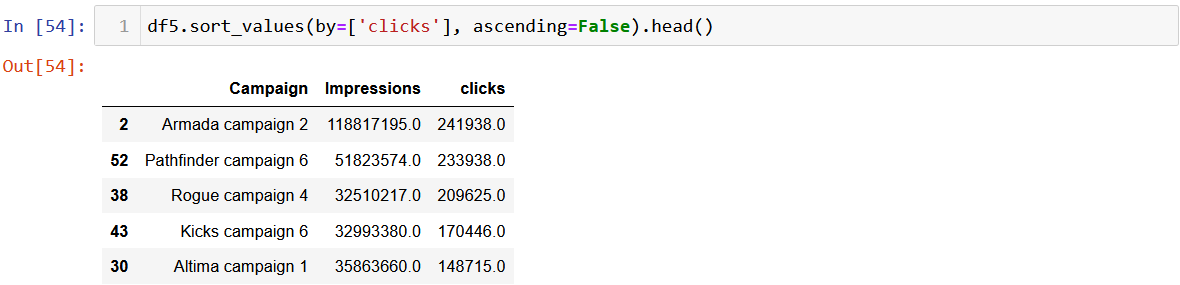
Yes, the maximum impressions also have maximum clicks, here we have the same code but I created two empty lists and appended the values of impressions and clicks, Later I sorted the list according to impressions first and then clicks and got the same results.

****

Then sorted the values according to impressions first



Again, re-sorted it in terms of clicks

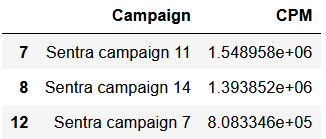


1. **Find out the top 3 campaigns in terms of CPM (CPM stands for cost per mile which means cost per 1000 impressions. CPM = total media cost/total impressions \* 1000)**

Code:



The output is:



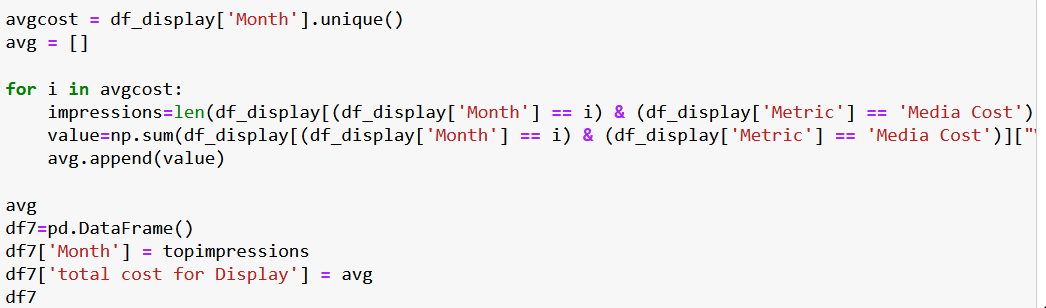
Therefore, you can see that the top three campaign in terms of CPM are:

* Sentra campaign 11
* Sentra campaign 14
* Sentra campaign 7

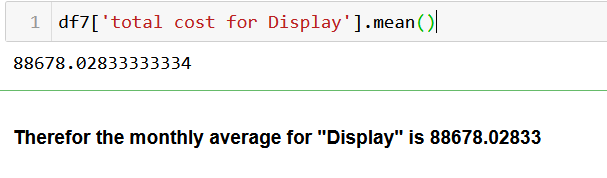
1. **What is the monthly average media cost for different media types?**

Here we can calculate the media cost separately

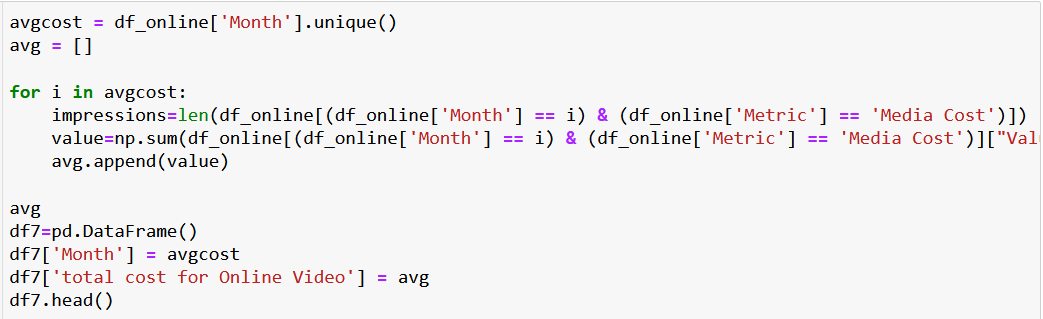
For “Online Video”



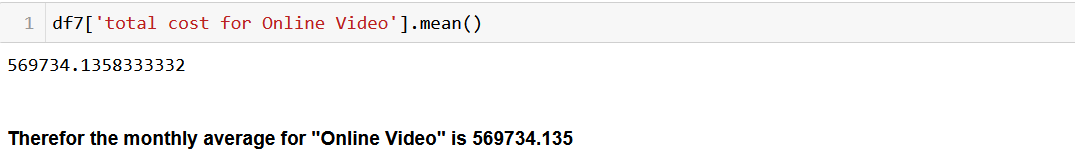
We get output:



For “Display”



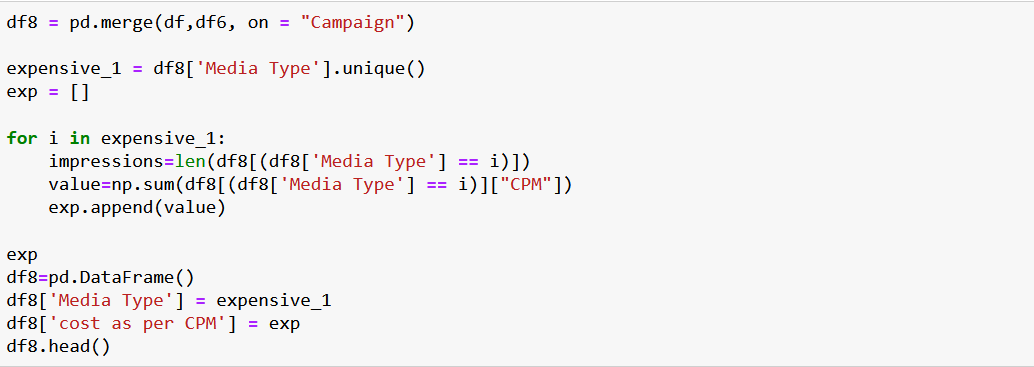
The output is



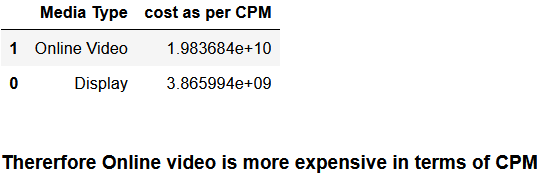
(Note: Here I have founded the average for every month even though some of months didn’t have any value, If we take the avg of the month which only consist the media cost of particular type then the value will vary)

1. **Which media type is more expensive in terms of CPM?**

As we had already calculated the CPM we will just need to add media type and calculate the CPM.

****

The output will be unsorted we can sort it and get the most expensive media type in terms of CPM



1. **When would you use a Pie chart, a Bar chart and a Line Chart? Use the given dataset and provide an example for each type of chart.**

* When you want to find the percentage of different categories within each column for categorical data we use Pie chart.
* But when you want to compare each category against each other within a column for a categorical dataset we use a bar chart.
* Line chart is effective when the column values need to be analyzed based on time (month, quarters, year)

(Note: I have just plotted the bar graph in this model as this data is very big for mu system to operate, It was crashing the system I have plotted many graphs for the please have a look)

