

After opening the file and turning it into a dataframe, I extracted the month from the transaction date using the 'split' function, used the if and elif functions to obtain the month names, and inserted a series of the month names equivalents in the main dataframe (df). I made a list of the months from January to June so I made use of the month column of the dataframe and dropped the duplicates. To start the first table, I separated the values under transaction items, turned the separated values into a dataframe, and dropped the unnecessary column since I only needed the numbers. To add the numbers, I turned the type of the values in the column from string to integer. I used the 'groupby' function to get the sum quantity for each product every month and I was finally able to generate the first table by turning my function into a dictionary and turning the dictionary into a dataframe. To start the second table, I got the items and corresponding values from the main data frame and dropped the repeating products. Next, I used the 'str.contains.' function to get the products that were by themselves and had only one quantity. Next, I made a data frame with just the product and cost. Following this, I made a function to get the price of a specific item. To get the total sales per month of each product, I made another function which multiplied the quantity sold per product and the price function I made, and turned this into a data frame. Originally, the rows and columns of the dataframe were switched so I used the 'transpose' function. For the third table, I first made a function that returned the transaction data of customers per month and another which gave me only the names of monthly customers and turned them both into dictionaries to easily get the values per month. For the repeaters function, I figured that January should have 0 repeaters since it is the first month. Next, I changed the month from name to number by getting the index to get the past month by subtracting the index by 1. I then created a series of current and past buyers and got the intersection of both using the 'merge' function, and used the 'len' function to get the count of names. For the inactive function, I did the same with January. For the past months, I used the first dictionary I made for the third table and the colon(:) to get the month from the start up until the "current" month. To get the dataframe of the past months, I used the 'isin' function and dropped the repeating names. Again, I derived the intersection of the current and past customers and subtracted this from the past customers. For the engaged function, I wanted to get the list of customers for each of the months prior to the "current" month. After searching for many hours, I came across the while loop. After numerous tries and going back and forth from the internet, I was finally able to use it in the function. After making all the tables, I wanted to have a few charts and graphs. I searched up how to do this and found that I can make graphs using just pandas.