In order to start my analysis, I went through the paper to see the deliverables needed. The first thing that came to mind was the last note in the context part of the brief file wherein it was mentioned that each row is sometimes populated with multiple semicolon separated items. Hence, the first thing that I did was to clean the data and explode the content of one row into multiple rows. To work with the data, I downloaded the JSON file and then converted it into a Pandas DataFrame with the same logic that I used to track the International Space Station. Once the DataFrame was created, I split the transaction items column with the delimiter (";") to get individual items per purchase then exploded it into different rows. I once again used the split function to split the transaction item column with the delimiter (",") to separate the brand name, the product name, and the quantity. The quantity was then converted into the int data type so that it could work better for that require integer inputs. I could not find a way to insert the unique product names into a for loop to get the individual price so the indexing for this part was done manually. I then created a new column that contains the price per product multiplied by the quantity sold in order to get the total price value of the rows. Also, the transaction date was a string so I converted it into a datetime data type and extracted the month and year for the groupby function. The groupby was then applied to the DataFrame and is grouped by transaction date and the product name, which sums up the quantity sold and the product value. For the loyalty table, I tried using the CohortIndex in order to see the number of repeaters per month but it was referencing the first month of purchase per unique username instead of the previous month. Hence the table generated is off. I used the diagonals of the table generated to add all instances wherein customers in that month also purchased in the previous month. However, this uses the first purchase as reference. Hence, I added an additional analysis that will show the number of months that each user has made at least one purchase to try and make up for the lack of the others. Other supporting analysis can be seen in the notebook and the powerpoint.