1. Import churn dataset into the PBI desktop using get data and selecting python
    a. Write code to import dataset using method read_csv
    b. Select transform data
2. This opens the Query Editor and gives you a lot of options to perform cleaning, reshaping, and transformation of data.
    a. convert the **customer_nw_category** variable into a text field as these represent the Customer Net Worth Category and it should not be used as a continuous variable.
    b. select the column, go to **Data Type,** and change the data type to Text. Power Query records this step under the **Applied Steps** section. It is a good practice to rename this step, for easy recall. We will rename it to "nw_cat Text".
    c. Similarly, we will transform the **churn** column into a logical variable, representing True for 1 (churned) and False for 0 (not churned) and rename the step to "churn – True/False".
3. After you have completed the transformation step, click on Close & Apply (on the top left corner) to apply these transformations to the data.
4. we will demonstrate how to create a correlation matrix heatmap using Python's correlation function. This heatmap will be displayed on the Report section in Power BI.
5. Head over to the **Report** section in Power BI and click on Python visual denoted by **Py** symbol under the **Visualizations** section. At the left, you will notice an empty Python visual appearing and a Python script Editor popping up at the bottom. In other words, Power BI gives you the option of creating visualizations with scripts.
6. You will notice that currently, the **Values** field is empty.
7. To illustrate the correlation heatmap, we will get all the continuous variables into the **Values** field, namely, age, all average monthly balance columns, current, and previous month balance and current and previous month transaction columns, a number of dependents, and vintage (the time of association). This is an important step. Otherwise, Power BI wouldn't recognize these variables to be part of the visualization.
8. As we get the variables into the **Values** field, the Python script is automatically populated with the following codes:

```
import matplotlib.pyplot as plt
import seaborn as sns

# create the correlation matrix on the dataset
corr = dataset.corr()
# create a heatmap of the correlation matrix
sns.heatmap(corr, cmap="YlGnBu")
# show plot
plt.show()
```

9. Finally, after running this script using the Run Script button, it produces a correlation matrix heatmap.
10. we apply a filter of churn = True or False using the blue boxes to observe the heatmap for the two groups of customers separately.