{

"cells": [

{

"cell\_type": "code",

"execution\_count": 19,

"id": "01578712",

"metadata": {},

"outputs": [],

"source": [

"# Import Libraries\n",

"import pandas as pd\n",

"import numpy as np\n",

"import matplotlib.pyplot as plt\n",

"import scipy.cluster.hierarchy as sch\n",

"from sklearn.cluster import AgglomerativeClustering\n",

"from sklearn.preprocessing import normalize"

]

},

{

"cell\_type": "code",

"execution\_count": 20,

"id": "f055e4e6",

"metadata": {},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

" .dataframe tbody tr th:only-of-type {\n",

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

"<table border=\"1\" class=\"dataframe\">\n",

" <thead>\n",

" <tr style=\"text-align: right;\">\n",

" <th></th>\n",

" <th>ID#</th>\n",

" <th>Balance</th>\n",

" <th>Qual\_miles</th>\n",

" <th>cc1\_miles</th>\n",

" <th>cc2\_miles</th>\n",

" <th>cc3\_miles</th>\n",

" <th>Bonus\_miles</th>\n",

" <th>Bonus\_trans</th>\n",

" <th>Flight\_miles\_12mo</th>\n",

" <th>Flight\_trans\_12</th>\n",

" <th>Days\_since\_enroll</th>\n",

" <th>Award?</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

" <th>0</th>\n",

" <td>1</td>\n",

" <td>28143</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>174</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7000</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>1</th>\n",

" <td>2</td>\n",

" <td>19244</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>215</td>\n",

" <td>2</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6968</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>2</th>\n",

" <td>3</td>\n",

" <td>41354</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>4123</td>\n",

" <td>4</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7034</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3</th>\n",

" <td>4</td>\n",

" <td>14776</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6952</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>4</th>\n",

" <td>5</td>\n",

" <td>97752</td>\n",

" <td>0</td>\n",

" <td>4</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>43300</td>\n",

" <td>26</td>\n",

" <td>2077</td>\n",

" <td>4</td>\n",

" <td>6935</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>...</th>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3994</th>\n",

" <td>4017</td>\n",

" <td>18476</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>8525</td>\n",

" <td>4</td>\n",

" <td>200</td>\n",

" <td>1</td>\n",

" <td>1403</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3995</th>\n",

" <td>4018</td>\n",

" <td>64385</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>981</td>\n",

" <td>5</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1395</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3996</th>\n",

" <td>4019</td>\n",

" <td>73597</td>\n",

" <td>0</td>\n",

" <td>3</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>25447</td>\n",

" <td>8</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1402</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3997</th>\n",

" <td>4020</td>\n",

" <td>54899</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>1401</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3998</th>\n",

" <td>4021</td>\n",

" <td>3016</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1398</td>\n",

" <td>0</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>3999 rows × 12 columns</p>\n",

"</div>"

],

"text/plain": [

" ID# Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles \\\n",

"0 1 28143 0 1 1 1 174 \n",

"1 2 19244 0 1 1 1 215 \n",

"2 3 41354 0 1 1 1 4123 \n",

"3 4 14776 0 1 1 1 500 \n",

"4 5 97752 0 4 1 1 43300 \n",

"... ... ... ... ... ... ... ... \n",

"3994 4017 18476 0 1 1 1 8525 \n",

"3995 4018 64385 0 1 1 1 981 \n",

"3996 4019 73597 0 3 1 1 25447 \n",

"3997 4020 54899 0 1 1 1 500 \n",

"3998 4021 3016 0 1 1 1 0 \n",

"\n",

" Bonus\_trans Flight\_miles\_12mo Flight\_trans\_12 Days\_since\_enroll \\\n",

"0 1 0 0 7000 \n",

"1 2 0 0 6968 \n",

"2 4 0 0 7034 \n",

"3 1 0 0 6952 \n",

"4 26 2077 4 6935 \n",

"... ... ... ... ... \n",

"3994 4 200 1 1403 \n",

"3995 5 0 0 1395 \n",

"3996 8 0 0 1402 \n",

"3997 1 500 1 1401 \n",

"3998 0 0 0 1398 \n",

"\n",

" Award? \n",

"0 0 \n",

"1 0 \n",

"2 0 \n",

"3 0 \n",

"4 1 \n",

"... ... \n",

"3994 1 \n",

"3995 1 \n",

"3996 1 \n",

"3997 0 \n",

"3998 0 \n",

"\n",

"[3999 rows x 12 columns]"

]

},

"execution\_count": 20,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"# Import Dataset\n",

"airline=pd.read\_csv('EastWestAirlines.csv')\n",

"airline"

]

},

{

"cell\_type": "code",

"execution\_count": 21,

"id": "50c31753",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

"<class 'pandas.core.frame.DataFrame'>\n",

"RangeIndex: 3999 entries, 0 to 3998\n",

"Data columns (total 12 columns):\n",

" # Column Non-Null Count Dtype\n",

"--- ------ -------------- -----\n",

" 0 ID# 3999 non-null int64\n",

" 1 Balance 3999 non-null int64\n",

" 2 Qual\_miles 3999 non-null int64\n",

" 3 cc1\_miles 3999 non-null int64\n",

" 4 cc2\_miles 3999 non-null int64\n",

" 5 cc3\_miles 3999 non-null int64\n",

" 6 Bonus\_miles 3999 non-null int64\n",

" 7 Bonus\_trans 3999 non-null int64\n",

" 8 Flight\_miles\_12mo 3999 non-null int64\n",

" 9 Flight\_trans\_12 3999 non-null int64\n",

" 10 Days\_since\_enroll 3999 non-null int64\n",

" 11 Award? 3999 non-null int64\n",

"dtypes: int64(12)\n",

"memory usage: 375.0 KB\n"

]

}

],

"source": [

"airline.info()"

]

},

{

"cell\_type": "code",

"execution\_count": 22,

"id": "53f46c3c",

"metadata": {},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

" .dataframe tbody tr th:only-of-type {\n",

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

"<table border=\"1\" class=\"dataframe\">\n",

" <thead>\n",

" <tr style=\"text-align: right;\">\n",

" <th></th>\n",

" <th>Balance</th>\n",

" <th>Qual\_miles</th>\n",

" <th>cc1\_miles</th>\n",

" <th>cc2\_miles</th>\n",

" <th>cc3\_miles</th>\n",

" <th>Bonus\_miles</th>\n",

" <th>Bonus\_trans</th>\n",

" <th>Flight\_miles\_12mo</th>\n",

" <th>Flight\_trans\_12</th>\n",

" <th>Days\_since\_enroll</th>\n",

" <th>Award?</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

" <th>0</th>\n",

" <td>28143</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>174</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7000</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>1</th>\n",

" <td>19244</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>215</td>\n",

" <td>2</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6968</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>2</th>\n",

" <td>41354</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>4123</td>\n",

" <td>4</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7034</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3</th>\n",

" <td>14776</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6952</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>4</th>\n",

" <td>97752</td>\n",

" <td>0</td>\n",

" <td>4</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>43300</td>\n",

" <td>26</td>\n",

" <td>2077</td>\n",

" <td>4</td>\n",

" <td>6935</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>...</th>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3994</th>\n",

" <td>18476</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>8525</td>\n",

" <td>4</td>\n",

" <td>200</td>\n",

" <td>1</td>\n",

" <td>1403</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3995</th>\n",

" <td>64385</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>981</td>\n",

" <td>5</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1395</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3996</th>\n",

" <td>73597</td>\n",

" <td>0</td>\n",

" <td>3</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>25447</td>\n",

" <td>8</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1402</td>\n",

" <td>1</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3997</th>\n",

" <td>54899</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>1401</td>\n",

" <td>0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3998</th>\n",

" <td>3016</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1398</td>\n",

" <td>0</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>3999 rows × 11 columns</p>\n",

"</div>"

],

"text/plain": [

" Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles \\\n",

"0 28143 0 1 1 1 174 \n",

"1 19244 0 1 1 1 215 \n",

"2 41354 0 1 1 1 4123 \n",

"3 14776 0 1 1 1 500 \n",

"4 97752 0 4 1 1 43300 \n",

"... ... ... ... ... ... ... \n",

"3994 18476 0 1 1 1 8525 \n",

"3995 64385 0 1 1 1 981 \n",

"3996 73597 0 3 1 1 25447 \n",

"3997 54899 0 1 1 1 500 \n",

"3998 3016 0 1 1 1 0 \n",

"\n",

" Bonus\_trans Flight\_miles\_12mo Flight\_trans\_12 Days\_since\_enroll \\\n",

"0 1 0 0 7000 \n",

"1 2 0 0 6968 \n",

"2 4 0 0 7034 \n",

"3 1 0 0 6952 \n",

"4 26 2077 4 6935 \n",

"... ... ... ... ... \n",

"3994 4 200 1 1403 \n",

"3995 5 0 0 1395 \n",

"3996 8 0 0 1402 \n",

"3997 1 500 1 1401 \n",

"3998 0 0 0 1398 \n",

"\n",

" Award? \n",

"0 0 \n",

"1 0 \n",

"2 0 \n",

"3 0 \n",

"4 1 \n",

"... ... \n",

"3994 1 \n",

"3995 1 \n",

"3996 1 \n",

"3997 0 \n",

"3998 0 \n",

"\n",

"[3999 rows x 11 columns]"

]

},

"execution\_count": 22,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"airline2=airline.drop(['ID#'],axis=1)\n",

"airline2"

]

},

{

"cell\_type": "code",

"execution\_count": 23,

"id": "802f6b2b",

"metadata": {},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

" .dataframe tbody tr th:only-of-type {\n",

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

"<table border=\"1\" class=\"dataframe\">\n",

" <thead>\n",

" <tr style=\"text-align: right;\">\n",

" <th></th>\n",

" <th>Balance</th>\n",

" <th>Qual\_miles</th>\n",

" <th>cc1\_miles</th>\n",

" <th>cc2\_miles</th>\n",

" <th>cc3\_miles</th>\n",

" <th>Bonus\_miles</th>\n",

" <th>Bonus\_trans</th>\n",

" <th>Flight\_miles\_12mo</th>\n",

" <th>Flight\_trans\_12</th>\n",

" <th>Days\_since\_enroll</th>\n",

" <th>Award?</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

" <th>0</th>\n",

" <td>0.970414</td>\n",

" <td>0.0</td>\n",

" <td>0.000034</td>\n",

" <td>0.000034</td>\n",

" <td>0.000034</td>\n",

" <td>0.006000</td>\n",

" <td>0.000034</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.241371</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" <tr>\n",

" <th>1</th>\n",

" <td>0.940209</td>\n",

" <td>0.0</td>\n",

" <td>0.000049</td>\n",

" <td>0.000049</td>\n",

" <td>0.000049</td>\n",

" <td>0.010504</td>\n",

" <td>0.000098</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.340437</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" <tr>\n",

" <th>2</th>\n",

" <td>0.981113</td>\n",

" <td>0.0</td>\n",

" <td>0.000024</td>\n",

" <td>0.000024</td>\n",

" <td>0.000024</td>\n",

" <td>0.097817</td>\n",

" <td>0.000095</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.166880</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3</th>\n",

" <td>0.904428</td>\n",

" <td>0.0</td>\n",

" <td>0.000061</td>\n",

" <td>0.000061</td>\n",

" <td>0.000061</td>\n",

" <td>0.030605</td>\n",

" <td>0.000061</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.425527</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" <tr>\n",

" <th>4</th>\n",

" <td>0.912226</td>\n",

" <td>0.0</td>\n",

" <td>0.000037</td>\n",

" <td>0.000009</td>\n",

" <td>0.000009</td>\n",

" <td>0.404078</td>\n",

" <td>0.000243</td>\n",

" <td>0.019383</td>\n",

" <td>0.000037</td>\n",

" <td>0.064718</td>\n",

" <td>0.000009</td>\n",

" </tr>\n",

" <tr>\n",

" <th>...</th>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3994</th>\n",

" <td>0.905810</td>\n",

" <td>0.0</td>\n",

" <td>0.000049</td>\n",

" <td>0.000049</td>\n",

" <td>0.000049</td>\n",

" <td>0.417949</td>\n",

" <td>0.000196</td>\n",

" <td>0.009805</td>\n",

" <td>0.000049</td>\n",

" <td>0.068784</td>\n",

" <td>0.000049</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3995</th>\n",

" <td>0.999649</td>\n",

" <td>0.0</td>\n",

" <td>0.000016</td>\n",

" <td>0.000016</td>\n",

" <td>0.000016</td>\n",

" <td>0.015231</td>\n",

" <td>0.000078</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.021659</td>\n",

" <td>0.000016</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3996</th>\n",

" <td>0.944948</td>\n",

" <td>0.0</td>\n",

" <td>0.000039</td>\n",

" <td>0.000013</td>\n",

" <td>0.000013</td>\n",

" <td>0.326726</td>\n",

" <td>0.000103</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.018001</td>\n",

" <td>0.000013</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3997</th>\n",

" <td>0.999592</td>\n",

" <td>0.0</td>\n",

" <td>0.000018</td>\n",

" <td>0.000018</td>\n",

" <td>0.000018</td>\n",

" <td>0.009104</td>\n",

" <td>0.000018</td>\n",

" <td>0.009104</td>\n",

" <td>0.000018</td>\n",

" <td>0.025509</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3998</th>\n",

" <td>0.907271</td>\n",

" <td>0.0</td>\n",

" <td>0.000301</td>\n",

" <td>0.000301</td>\n",

" <td>0.000301</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.000000</td>\n",

" <td>0.420546</td>\n",

" <td>0.000000</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>3999 rows × 11 columns</p>\n",

"</div>"

],

"text/plain": [

" Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles \\\n",

"0 0.970414 0.0 0.000034 0.000034 0.000034 0.006000 \n",

"1 0.940209 0.0 0.000049 0.000049 0.000049 0.010504 \n",

"2 0.981113 0.0 0.000024 0.000024 0.000024 0.097817 \n",

"3 0.904428 0.0 0.000061 0.000061 0.000061 0.030605 \n",

"4 0.912226 0.0 0.000037 0.000009 0.000009 0.404078 \n",

"... ... ... ... ... ... ... \n",

"3994 0.905810 0.0 0.000049 0.000049 0.000049 0.417949 \n",

"3995 0.999649 0.0 0.000016 0.000016 0.000016 0.015231 \n",

"3996 0.944948 0.0 0.000039 0.000013 0.000013 0.326726 \n",

"3997 0.999592 0.0 0.000018 0.000018 0.000018 0.009104 \n",

"3998 0.907271 0.0 0.000301 0.000301 0.000301 0.000000 \n",

"\n",

" Bonus\_trans Flight\_miles\_12mo Flight\_trans\_12 Days\_since\_enroll \\\n",

"0 0.000034 0.000000 0.000000 0.241371 \n",

"1 0.000098 0.000000 0.000000 0.340437 \n",

"2 0.000095 0.000000 0.000000 0.166880 \n",

"3 0.000061 0.000000 0.000000 0.425527 \n",

"4 0.000243 0.019383 0.000037 0.064718 \n",

"... ... ... ... ... \n",

"3994 0.000196 0.009805 0.000049 0.068784 \n",

"3995 0.000078 0.000000 0.000000 0.021659 \n",

"3996 0.000103 0.000000 0.000000 0.018001 \n",

"3997 0.000018 0.009104 0.000018 0.025509 \n",

"3998 0.000000 0.000000 0.000000 0.420546 \n",

"\n",

" Award? \n",

"0 0.000000 \n",

"1 0.000000 \n",

"2 0.000000 \n",

"3 0.000000 \n",

"4 0.000009 \n",

"... ... \n",

"3994 0.000049 \n",

"3995 0.000016 \n",

"3996 0.000013 \n",

"3997 0.000000 \n",

"3998 0.000000 \n",

"\n",

"[3999 rows x 11 columns]"

]

},

"execution\_count": 23,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"# Normalize heterogenous numerical data \n",

"airline2\_norm=pd.DataFrame(normalize(airline2),columns=airline2.columns)\n",

"airline2\_norm"

]

},

{

"cell\_type": "code",

"execution\_count": 24,

"id": "d7fcfe4a",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 720x504 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"# Create Dendrograms\n",

"plt.figure(figsize=(10, 7)) \n",

"dendograms=sch.dendrogram(sch.linkage(airline2\_norm,'complete'))"

]

},

{

"cell\_type": "code",

"execution\_count": 25,

"id": "016a7a31",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"AgglomerativeClustering(n\_clusters=5)"

]

},

"execution\_count": 25,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"# Create Clusters (y)\n",

"hclusters=AgglomerativeClustering(n\_clusters=5,affinity='euclidean',linkage='ward')\n",

"hclusters"

]

},

{

"cell\_type": "code",

"execution\_count": 26,

"id": "34c01a50",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"2 1547\n",

"4 1191\n",

"3 579\n",

"1 453\n",

"0 229\n",

"Name: clustersid, dtype: int64"

]

},

"execution\_count": 26,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y=pd.DataFrame(hclusters.fit\_predict(airline2\_norm),columns=['clustersid'])\n",

"y['clustersid'].value\_counts()"

]

},

{

"cell\_type": "code",

"execution\_count": 27,

"id": "d69d6243",

"metadata": {},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

" .dataframe tbody tr th:only-of-type {\n",

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

"<table border=\"1\" class=\"dataframe\">\n",

" <thead>\n",

" <tr style=\"text-align: right;\">\n",

" <th></th>\n",

" <th>Balance</th>\n",

" <th>Qual\_miles</th>\n",

" <th>cc1\_miles</th>\n",

" <th>cc2\_miles</th>\n",

" <th>cc3\_miles</th>\n",

" <th>Bonus\_miles</th>\n",

" <th>Bonus\_trans</th>\n",

" <th>Flight\_miles\_12mo</th>\n",

" <th>Flight\_trans\_12</th>\n",

" <th>Days\_since\_enroll</th>\n",

" <th>Award?</th>\n",

" <th>clustersid</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

" <th>0</th>\n",

" <td>28143</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>174</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7000</td>\n",

" <td>0</td>\n",

" <td>4</td>\n",

" </tr>\n",

" <tr>\n",

" <th>1</th>\n",

" <td>19244</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>215</td>\n",

" <td>2</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6968</td>\n",

" <td>0</td>\n",

" <td>2</td>\n",

" </tr>\n",

" <tr>\n",

" <th>2</th>\n",

" <td>41354</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>4123</td>\n",

" <td>4</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>7034</td>\n",

" <td>0</td>\n",

" <td>2</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3</th>\n",

" <td>14776</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>6952</td>\n",

" <td>0</td>\n",

" <td>2</td>\n",

" </tr>\n",

" <tr>\n",

" <th>4</th>\n",

" <td>97752</td>\n",

" <td>0</td>\n",

" <td>4</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>43300</td>\n",

" <td>26</td>\n",

" <td>2077</td>\n",

" <td>4</td>\n",

" <td>6935</td>\n",

" <td>1</td>\n",

" <td>3</td>\n",

" </tr>\n",

" <tr>\n",

" <th>...</th>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" <td>...</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3994</th>\n",

" <td>18476</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>8525</td>\n",

" <td>4</td>\n",

" <td>200</td>\n",

" <td>1</td>\n",

" <td>1403</td>\n",

" <td>1</td>\n",

" <td>3</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3995</th>\n",

" <td>64385</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>981</td>\n",

" <td>5</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1395</td>\n",

" <td>1</td>\n",

" <td>4</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3996</th>\n",

" <td>73597</td>\n",

" <td>0</td>\n",

" <td>3</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>25447</td>\n",

" <td>8</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1402</td>\n",

" <td>1</td>\n",

" <td>2</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3997</th>\n",

" <td>54899</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>500</td>\n",

" <td>1</td>\n",

" <td>1401</td>\n",

" <td>0</td>\n",

" <td>4</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3998</th>\n",

" <td>3016</td>\n",

" <td>0</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>1</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>0</td>\n",

" <td>1398</td>\n",

" <td>0</td>\n",

" <td>2</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>3999 rows × 12 columns</p>\n",

"</div>"

],

"text/plain": [

" Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles Bonus\_miles \\\n",

"0 28143 0 1 1 1 174 \n",

"1 19244 0 1 1 1 215 \n",

"2 41354 0 1 1 1 4123 \n",

"3 14776 0 1 1 1 500 \n",

"4 97752 0 4 1 1 43300 \n",

"... ... ... ... ... ... ... \n",

"3994 18476 0 1 1 1 8525 \n",

"3995 64385 0 1 1 1 981 \n",

"3996 73597 0 3 1 1 25447 \n",

"3997 54899 0 1 1 1 500 \n",

"3998 3016 0 1 1 1 0 \n",

"\n",

" Bonus\_trans Flight\_miles\_12mo Flight\_trans\_12 Days\_since\_enroll \\\n",

"0 1 0 0 7000 \n",

"1 2 0 0 6968 \n",

"2 4 0 0 7034 \n",

"3 1 0 0 6952 \n",

"4 26 2077 4 6935 \n",

"... ... ... ... ... \n",

"3994 4 200 1 1403 \n",

"3995 5 0 0 1395 \n",

"3996 8 0 0 1402 \n",

"3997 1 500 1 1401 \n",

"3998 0 0 0 1398 \n",

"\n",

" Award? clustersid \n",

"0 0 4 \n",

"1 0 2 \n",

"2 0 2 \n",

"3 0 2 \n",

"4 1 3 \n",

"... ... ... \n",

"3994 1 3 \n",

"3995 1 4 \n",

"3996 1 2 \n",

"3997 0 4 \n",

"3998 0 2 \n",

"\n",

"[3999 rows x 12 columns]"

]

},

"execution\_count": 27,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"# Adding clusters to dataset\n",

"airline2['clustersid']=hclusters.labels\_\n",

"airline2"

]

},

{

"cell\_type": "code",

"execution\_count": 28,

"id": "f32048ad",

"metadata": {},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

" .dataframe tbody tr th:only-of-type {\n",

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead tr th {\n",

" text-align: left;\n",

" }\n",

"</style>\n",

"<table border=\"1\" class=\"dataframe\">\n",

" <thead>\n",

" <tr>\n",

" <th></th>\n",

" <th>clustersid</th>\n",

" <th>Balance</th>\n",

" <th>Qual\_miles</th>\n",

" <th>cc1\_miles</th>\n",

" <th>cc2\_miles</th>\n",

" <th>cc3\_miles</th>\n",

" <th>Bonus\_miles</th>\n",

" <th>Bonus\_trans</th>\n",

" <th>Flight\_miles\_12mo</th>\n",

" <th>Flight\_trans\_12</th>\n",

" <th>Days\_since\_enroll</th>\n",

" <th>Award?</th>\n",

" </tr>\n",

" <tr>\n",

" <th></th>\n",

" <th></th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" <th>mean</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

" <th>0</th>\n",

" <td>0</td>\n",

" <td>5524.222707</td>\n",

" <td>8.755459</td>\n",

" <td>1.000000</td>\n",

" <td>1.000000</td>\n",

" <td>1.000000</td>\n",

" <td>584.532751</td>\n",

" <td>2.401747</td>\n",

" <td>66.982533</td>\n",

" <td>0.209607</td>\n",

" <td>4875.301310</td>\n",

" <td>0.131004</td>\n",

" </tr>\n",

" <tr>\n",

" <th>1</th>\n",

" <td>1</td>\n",

" <td>31066.514349</td>\n",

" <td>111.415011</td>\n",

" <td>3.200883</td>\n",

" <td>1.026490</td>\n",

" <td>1.070640</td>\n",

" <td>40266.935982</td>\n",

" <td>17.289183</td>\n",

" <td>626.754967</td>\n",

" <td>1.812362</td>\n",

" <td>4205.624724</td>\n",

" <td>0.858720</td>\n",

" </tr>\n",

" <tr>\n",

" <th>2</th>\n",

" <td>2</td>\n",

" <td>81201.080802</td>\n",

" <td>136.521008</td>\n",

" <td>2.115061</td>\n",

" <td>1.013575</td>\n",

" <td>1.000646</td>\n",

" <td>16350.149968</td>\n",

" <td>13.574014</td>\n",

" <td>488.550743</td>\n",

" <td>1.340659</td>\n",

" <td>4285.891403</td>\n",

" <td>0.312217</td>\n",

" </tr>\n",

" <tr>\n",

" <th>3</th>\n",

" <td>3</td>\n",

" <td>69569.894646</td>\n",

" <td>97.257340</td>\n",

" <td>3.326425</td>\n",

" <td>1.032815</td>\n",

" <td>1.022453</td>\n",

" <td>35743.675302</td>\n",

" <td>17.784111</td>\n",

" <td>406.804836</td>\n",

" <td>1.274611</td>\n",

" <td>4090.832470</td>\n",

" <td>0.525043</td>\n",

" </tr>\n",

" <tr>\n",

" <th>4</th>\n",

" <td>4</td>\n",

" <td>94957.590260</td>\n",

" <td>215.220823</td>\n",

" <td>1.141058</td>\n",

" <td>1.005038</td>\n",

" <td>1.002519</td>\n",

" <td>3524.928631</td>\n",

" <td>5.640638</td>\n",

" <td>461.104954</td>\n",

" <td>1.521411</td>\n",

" <td>3736.071369</td>\n",

" <td>0.230898</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" clustersid Balance Qual\_miles cc1\_miles cc2\_miles cc3\_miles \\\n",

" mean mean mean mean mean \n",

"0 0 5524.222707 8.755459 1.000000 1.000000 1.000000 \n",

"1 1 31066.514349 111.415011 3.200883 1.026490 1.070640 \n",

"2 2 81201.080802 136.521008 2.115061 1.013575 1.000646 \n",

"3 3 69569.894646 97.257340 3.326425 1.032815 1.022453 \n",

"4 4 94957.590260 215.220823 1.141058 1.005038 1.002519 \n",

"\n",

" Bonus\_miles Bonus\_trans Flight\_miles\_12mo Flight\_trans\_12 \\\n",

" mean mean mean mean \n",

"0 584.532751 2.401747 66.982533 0.209607 \n",

"1 40266.935982 17.289183 626.754967 1.812362 \n",

"2 16350.149968 13.574014 488.550743 1.340659 \n",

"3 35743.675302 17.784111 406.804836 1.274611 \n",

"4 3524.928631 5.640638 461.104954 1.521411 \n",

"\n",

" Days\_since\_enroll Award? \n",

" mean mean \n",

"0 4875.301310 0.131004 \n",

"1 4205.624724 0.858720 \n",

"2 4285.891403 0.312217 \n",

"3 4090.832470 0.525043 \n",

"4 3736.071369 0.230898 "

]

},

"execution\_count": 28,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"airline2.groupby('clustersid').agg(['mean']).reset\_index()"

]

},

{

"cell\_type": "code",

"execution\_count": 29,

"id": "84c6b90c",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"<matplotlib.collections.PathCollection at 0x298d8cc9490>"

]

},

"execution\_count": 29,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 720x504 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"# Plot Clusters\n",

"plt.figure(figsize=(10, 7)) \n",

"plt.scatter(airline2['clustersid'],airline2['Balance'], c=hclusters.labels\_) "

]

}

],

"metadata": {

"kernelspec": {

"display\_name": "Python 3 (ipykernel)",

"language": "python",

"name": "python3"

},

"language\_info": {

"codemirror\_mode": {

"name": "ipython",

"version": 3

},

"file\_extension": ".py",

"mimetype": "text/x-python",

"name": "python",

"nbconvert\_exporter": "python",

"pygments\_lexer": "ipython3",

"version": "3.9.7"

}

},

"nbformat": 4,

"nbformat\_minor": 5

}