{

"cells": [

{

"cell\_type": "markdown",

"id": "e70dabbc",

"metadata": {},

"source": [

"### FY22-Analytics-Course-MD\n"

]

},

{

"cell\_type": "markdown",

"id": "9880ee47",

"metadata": {},

"source": [

"### Milestones"

]

},

{

"cell\_type": "markdown",

"id": "85c8b618",

"metadata": {},

"source": [

"Github Link : https://github.com/Dream43210/UCDPA\_Mary-Drea-Project-for-Data-Analysis-Course"

]

},

{

"cell\_type": "markdown",

"id": "5416a55a",

"metadata": {},

"source": [

"#### 1. Data\n",

"o The project should use a real-world dataset and include a reference to the\n",

"source in the report.\n",

"\n",

"1.0 The data used here is a combination of 2 flat file. Both Files added in \n",

" Github :https://github.com/Dream43210/UCDPA\_Mary-Drea-Project-for-Data-Analysis-Course\n",

"1.1 Data from Salesforce CRM ( work data with all confidential information masked) \n",

" File Name : Dataset for Project 2 22.08.csv\n",

"1.2: https://www.worldometers.info/world-population/population-by-country/ \n",

" File Name : Country Data Source Worldometer.\n",

"Population by Country. Countries in the world by population (2022)\n",

"Data based on the latest United Nations Population Division estimation\n",

"Source: Worldometer (www.Worldometers.info)\n",

"Elaboration of data by United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2019 Revision. (Medium-fertility variant).\n"

]

},

{

"cell\_type": "markdown",

"id": "898d5ddf",

"metadata": {},

"source": [

"\*Note : Adding imports that will be used throughout the project.Additions may be required at later satges. "

]

},

{

"cell\_type": "code",

"execution\_count": 19,

"id": "8eed4a6d",

"metadata": {},

"outputs": [],

"source": [

"import pandas as pd"

]

},

{

"cell\_type": "code",

"execution\_count": 20,

"id": "3e349266",

"metadata": {},

"outputs": [],

"source": [

"import matplotlib.pyplot as plt"

]

},

{

"cell\_type": "code",

"execution\_count": 21,

"id": "e00d7db8",

"metadata": {},

"outputs": [],

"source": [

"import numpy as np"

]

},

{

"cell\_type": "code",

"execution\_count": 22,

"id": "0359302f",

"metadata": {},

"outputs": [],

"source": [

"import seaborn as sns"

]

},

{

"cell\_type": "markdown",

"id": "9095628a",

"metadata": {},

"source": [

"#### 2. Importing\n",

"o Import data from a flat file (.csv, .xls, xlsx, .txt, etc.)\n",

"o Retrieve data using online SQL, APIs, or web scraping. \n",

"\n",

"-For the purposes of this Project, I have used 2 CSV files relevant to Customer Planning, as this is the focus in this analysis. "

]

},

{

"cell\_type": "code",

"execution\_count": 26,

"id": "c7d19216",

"metadata": {},

"outputs": [],

"source": [

"df1 = pd.read\_csv (r\"C:\\Users\\dream\\OneDrive - VMware, Inc\\VMwareCorp\\Desktop\\Dataset for Project 2 22.08.csv\")"

]

},

{

"cell\_type": "code",

"execution\_count": 27,

"id": "83bb5e3a",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

" Customer Business Plan: Plan # Customer Account: Account Name \\\n",

"0 CP-0000037961 ABC123 \n",

"1 CP-0000037955 ABC300 \n",

"2 CP-0000038896 ABC432 \n",

"3 CP-0000037959 ABC472 \n",

"4 CP-0000026118 ABC555 \n",

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

"1790 CP-0000025913 ABC1325 \n",

"1791 CP-0000038983 ABC1566 \n",

"1792 CP-0000038937 ABC1595 \n",

"1793 CP-0000026154 ABC1824 \n",

"1794 CP-0000038460 ABC208 \n",

"\n",

" Customer Account: Zip/Postal Code \\\n",

"0 100000 \n",

"1 84 \n",

"2 700000 \n",

"3 10000 \n",

"4 0 \n",

"... ... \n",

"1790 C1057AAH \n",

"1791 1123 \n",

"1792 1009 \n",

"1793 16000 \n",

"1794 1010 \n",

"\n",

" Customer Account: Address 1 \\\n",

"0 Sá»‘ nhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 22nd Floor, Keangnam Landmark 72 Tower E6,, \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St \n",

"3 18A - Cong Hoa Street, Ward 12, Tan Binh District \n",

"4 2nd floor, 2B1 Dam Trau, Bach Dang Ward, \n",

"... ... \n",

"1790 Parana 771 \n",

"1791 Av Santa Fe 1845 5B \n",

"1792 Carlos Pellegrini 1163, PB \n",

"1793 85 Lotissement E'nasr \n",

"1794 Rr. Murat Toptani, Gjergji Center, \n",

"\n",

" Customer Account: D&B Address line 1 \\\n",

"0 Sá»‘ NhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 E6 Pham Hung Street \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St., Wa... \n",

"3 15 Duy Tan Dich Vong Hau \n",

"4 7 A2 Dam Trau, Bach Dang Ward Hai Ba Trung \n",

"... ... \n",

"1790 NaN \n",

"1791 Av Santa Fe 1845 5b \n",

"1792 Carlos Pellegrini 1163, Pb \n",

"1793 NaN \n",

"1794 Rruga Murat Toptani \n",

"\n",

" Customer Business Plan: Created Date Created Quarter \\\n",

"0 11/03/2022 Q1 \n",

"1 11/03/2022 Q1 \n",

"2 31/03/2022 Q1 \n",

"3 11/03/2022 Q1 \n",

"4 20/04/2021 Q2 \n",

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

"1790 31/03/2021 Q1 \n",

"1791 05/04/2022 Q2 \n",

"1792 04/04/2022 Q2 \n",

"1793 30/04/2021 Q2 \n",

"1794 17/03/2022 Q1 \n",

"\n",

" Customer Business Plan: Last Modified Date Year Plan Status \\\n",

"0 13/04/2022 2023 Submitted to Customer \n",

"1 13/04/2022 2023 Approved \n",

"2 21/08/2022 2023 Approved \n",

"3 29/04/2022 2023 Approved \n",

"4 02/03/2022 2022 Approved \n",

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

"1790 13/04/2021 2022 Draft \n",

"1791 30/06/2022 2023 Approved \n",

"1792 21/08/2022 2023 Approved \n",

"1793 09/03/2022 2022 Draft \n",

"1794 13/06/2022 2023 Approved \n",

"\n",

" Was the plan originally Approved? Customer Account: Is Direct Reseller \\\n",

"0 0 0 \n",

"1 1 0 \n",

"2 1 0 \n",

"3 1 0 \n",

"4 1 0 \n",

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

"1790 0 0 \n",

"1791 1 0 \n",

"1792 1 0 \n",

"1793 0 0 \n",

"1794 1 0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"0 0 \n",

"1 0 \n",

"2 0 \n",

"3 0 \n",

"4 0 \n",

"... ... \n",

"1790 0 \n",

"1791 0 \n",

"1792 0 \n",

"1793 0 \n",

"1794 0 \n",

"\n",

" Customer Account: Is Distributor Account \\\n",

"0 0 \n",

"1 0 \n",

"2 1 \n",

"3 0 \n",

"4 0 \n",

"... ... \n",

"1790 0 \n",

"1791 0 \n",

"1792 1 \n",

"1793 0 \n",

"1794 0 \n",

"\n",

" Customer Account: Is Aggregator Customer Account: Is Services Only \\\n",

"0 0 0 \n",

"1 0 0 \n",

"2 0 0 \n",

"3 0 0 \n",

"4 0 0 \n",

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

"1790 0 0 \n",

"1791 0 0 \n",

"1792 1 0 \n",

"1793 0 0 \n",

"1794 0 0 \n",

"\n",

" Customer Account: Is Cloud Customer Account: Is Resell \\\n",

"0 0 1 \n",

"1 0 1 \n",

"2 0 0 \n",

"3 0 1 \n",

"4 0 1 \n",

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

"1790 0 1 \n",

"1791 0 1 \n",

"1792 0 0 \n",

"1793 0 1 \n",

"1794 0 1 \n",

"\n",

" Customer Account: Is Subscriptions Eligible \\\n",

"0 1 \n",

"1 1 \n",

"2 1 \n",

"3 0 \n",

"4 0 \n",

"... ... \n",

"1790 1 \n",

"1791 1 \n",

"1792 0 \n",

"1793 0 \n",

"1794 1 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"0 0 0 \n",

"1 0 0 \n",

"2 0 1 \n",

"3 0 1 \n",

"4 0 0 \n",

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

"1790 1 1 \n",

"1791 1 1 \n",

"1792 0 0 \n",

"1793 1 1 \n",

"1794 1 1 \n",

"\n",

" Define Customerâ€™s marketing capability \\\n",

"0 3 Customer has integrated marketing function; ... \n",

"1 4 Customer has integrated marketing function; ... \n",

"2 4 Customer has integrated marketing function; ... \n",

"3 3 Customer has integrated marketing function; ... \n",

"4 3 Customer has integrated marketing function; ... \n",

"... ... \n",

"1790 5 Customer runs business outcome marketing cam... \n",

"1791 1 Customer has limited marketing capability \n",

"1792 5 Customer runs business outcome marketing cam... \n",

"1793 NaN \n",

"1794 NaN \n",

"\n",

" Define Customerâ€™s pre-sales scalability \\\n",

"0 3 Customer pre-Â­sales could scale up and there... \n",

"1 4 Customer pre-Â­sales is currently scaling up \n",

"2 3 Customer pre-Â­sales could scale up and there... \n",

"3 3 Customer pre-Â­sales could scale up and there... \n",

"4 4 Customer pre-Â­sales is currently scaling up \n",

"... ... \n",

"1790 3 Customer pre-Â­sales could scale up and there... \n",

"1791 5 Customer pre-Â­sales is already very scalable \n",

"1792 5 Customer pre-Â­sales is already very scalable \n",

"1793 NaN \n",

"1794 NaN \n",

"\n",

" Customer Account: Managed Customer Managed Customer Plan End Date \\\n",

"0 Yes 1 03/02/2023 \n",

"1 No 0 03/02/2023 \n",

"2 Yes 1 03/02/2023 \n",

"3 No 0 03/02/2023 \n",

"4 No 0 28/01/2022 \n",

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

"1790 No 0 28/01/2022 \n",

"1791 No 0 03/02/2023 \n",

"1792 No 0 03/02/2023 \n",

"1793 No 0 28/01/2022 \n",

"1794 No 0 03/02/2023 \n",

"\n",

" Executive Sponspor Present Present Quarter Region Geo \\\n",

"0 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"1 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"2 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"3 0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"4 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

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

"1790 0 Q1,Q2,Q3,Q4 LATAM AMER \n",

"1791 1 Q1,Q2,Q3,Q4 LATAM AMER \n",

"1792 1 Q1,Q2,Q3,Q4 LATAM AMER \n",

"1793 0 Q1,Q2,Q3,Q4 SEMEA EMEA \n",

"1794 0 Q1,Q2,Q3,Q4 CEMEA EMEA \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \\\n",

"0 0 40 1 \n",

"1 0 2510 1 \n",

"2 0 91 0 \n",

"3 0 80 1 \n",

"4 0 150 0 \n",

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

"1790 0 48 0 \n",

"1791 0 11 0 \n",

"1792 0 33 1 \n",

"1793 0 0 0 \n",

"1794 0 0 0 \n",

"\n",

" KPI Compliant Failure Reason Place Country \n",

"0 NaN VIETNAM \n",

"1 NaN VIETNAM \n",

"2 Goals and Objectives- Min. 1 Goal Required VIETNAM \n",

"3 NaN VIETNAM \n",

"4 Goals and Objectives- Min. 1 Goal Required; Pl... VIETNAM \n",

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

"1790 NaN ARGENTINA \n",

"1791 NaN ARGENTINA \n",

"1792 NaN ARGENTINA \n",

"1793 NaN ALGERIA \n",

"1794 Planner- Incomplete Planned Total FY >1 ALBANIA \n",

"\n",

"[1795 rows x 35 columns]\n"

]

}

],

"source": [

"print (df1)"

]

},

{

"cell\_type": "markdown",

"id": "1239c9fc",

"metadata": {},

"source": [

"I want to be able to see all the Columns Heading in a easy to view option. Updating the max column display settings"

]

},

{

"cell\_type": "code",

"execution\_count": 24,

"id": "b98db52d",

"metadata": {},

"outputs": [],

"source": [

"pd.set\_option('display.max\_columns', None)"

]

},

{

"cell\_type": "code",

"execution\_count": 28,

"id": "162fb28f",

"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>Customer Business Plan: Plan #</th>\n",

" <th>Customer Account: Account Name</th>\n",

" <th>Customer Account: Zip/Postal Code</th>\n",

" <th>Customer Account: Address 1</th>\n",

" <th>Customer Account: D&amp;B Address line 1</th>\n",

" <th>Customer Business Plan: Created Date</th>\n",

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

" <th>Customer Business Plan: Last Modified Date</th>\n",

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

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

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

" <th>Define Customerâ€™s marketing capability</th>\n",

" <th>Define Customerâ€™s pre-sales scalability</th>\n",

" <th>Customer Account: Managed Customer</th>\n",

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

" <th>Plan End Date</th>\n",

" <th>Executive Sponspor Present Present</th>\n",

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

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

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

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

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

" <th>KPI Compliant Failure Reason</th>\n",

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>CP-0000037961</td>\n",

" <td>ABC123</td>\n",

" <td>100000</td>\n",

" <td>Sá»‘ nhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T...</td>\n",

" <td>Sá»‘ NhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T...</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>13/04/2022</td>\n",

" <td>2023</td>\n",

" <td>Submitted to Customer</td>\n",

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

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

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

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

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

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

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

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

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

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

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

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>40</td>\n",

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000037955</td>\n",

" <td>ABC300</td>\n",

" <td>84</td>\n",

" <td>22nd Floor, Keangnam Landmark 72 Tower E6,,</td>\n",

" <td>E6 Pham Hung Street</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>13/04/2022</td>\n",

" <td>2023</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>4 Customer has integrated marketing function; ...</td>\n",

" <td>4 Customer pre-Â­sales is currently scaling up</td>\n",

" <td>No</td>\n",

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>2510</td>\n",

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000038896</td>\n",

" <td>ABC432</td>\n",

" <td>700000</td>\n",

" <td>Tuoi Tre Tower, No. 60A, Hoang Van Thu St</td>\n",

" <td>Tuoi Tre Tower, No. 60A, Hoang Van Thu St., Wa...</td>\n",

" <td>31/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>21/08/2022</td>\n",

" <td>2023</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>4 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

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

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>91</td>\n",

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

" <td>Goals and Objectives- Min. 1 Goal Required</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000037959</td>\n",

" <td>ABC472</td>\n",

" <td>10000</td>\n",

" <td>18A - Cong Hoa Street, Ward 12, Tan Binh District</td>\n",

" <td>15 Duy Tan Dich Vong Hau</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>29/04/2022</td>\n",

" <td>2023</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

" <td>No</td>\n",

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>80</td>\n",

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000026118</td>\n",

" <td>ABC555</td>\n",

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

" <td>2nd floor, 2B1 Dam Trau, Bach Dang Ward,</td>\n",

" <td>7 A2 Dam Trau, Bach Dang Ward Hai Ba Trung</td>\n",

" <td>20/04/2021</td>\n",

" <td>Q2</td>\n",

" <td>02/03/2022</td>\n",

" <td>2022</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>4 Customer pre-Â­sales is currently scaling up</td>\n",

" <td>No</td>\n",

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

" <td>28/01/2022</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>150</td>\n",

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

" <td>Goals and Objectives- Min. 1 Goal Required; Pl...</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Customer Business Plan: Plan # Customer Account: Account Name \\\n",

"0 CP-0000037961 ABC123 \n",

"1 CP-0000037955 ABC300 \n",

"2 CP-0000038896 ABC432 \n",

"3 CP-0000037959 ABC472 \n",

"4 CP-0000026118 ABC555 \n",

"\n",

" Customer Account: Zip/Postal Code \\\n",

"0 100000 \n",

"1 84 \n",

"2 700000 \n",

"3 10000 \n",

"4 0 \n",

"\n",

" Customer Account: Address 1 \\\n",

"0 Sá»‘ nhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 22nd Floor, Keangnam Landmark 72 Tower E6,, \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St \n",

"3 18A - Cong Hoa Street, Ward 12, Tan Binh District \n",

"4 2nd floor, 2B1 Dam Trau, Bach Dang Ward, \n",

"\n",

" Customer Account: D&B Address line 1 \\\n",

"0 Sá»‘ NhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 E6 Pham Hung Street \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St., Wa... \n",

"3 15 Duy Tan Dich Vong Hau \n",

"4 7 A2 Dam Trau, Bach Dang Ward Hai Ba Trung \n",

"\n",

" Customer Business Plan: Created Date Created Quarter \\\n",

"0 11/03/2022 Q1 \n",

"1 11/03/2022 Q1 \n",

"2 31/03/2022 Q1 \n",

"3 11/03/2022 Q1 \n",

"4 20/04/2021 Q2 \n",

"\n",

" Customer Business Plan: Last Modified Date Year Plan Status \\\n",

"0 13/04/2022 2023 Submitted to Customer \n",

"1 13/04/2022 2023 Approved \n",

"2 21/08/2022 2023 Approved \n",

"3 29/04/2022 2023 Approved \n",

"4 02/03/2022 2022 Approved \n",

"\n",

" Was the plan originally Approved? Customer Account: Is Direct Reseller \\\n",

"0 0 0 \n",

"1 1 0 \n",

"2 1 0 \n",

"3 1 0 \n",

"4 1 0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"0 0 \n",

"1 0 \n",

"2 0 \n",

"3 0 \n",

"4 0 \n",

"\n",

" Customer Account: Is Distributor Account Customer Account: Is Aggregator \\\n",

"0 0 0 \n",

"1 0 0 \n",

"2 1 0 \n",

"3 0 0 \n",

"4 0 0 \n",

"\n",

" Customer Account: Is Services Only Customer Account: Is Cloud \\\n",

"0 0 0 \n",

"1 0 0 \n",

"2 0 0 \n",

"3 0 0 \n",

"4 0 0 \n",

"\n",

" Customer Account: Is Resell Customer Account: Is Subscriptions Eligible \\\n",

"0 1 1 \n",

"1 1 1 \n",

"2 0 1 \n",

"3 1 0 \n",

"4 1 0 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"0 0 0 \n",

"1 0 0 \n",

"2 0 1 \n",

"3 0 1 \n",

"4 0 0 \n",

"\n",

" Define Customerâ€™s marketing capability \\\n",

"0 3 Customer has integrated marketing function; ... \n",

"1 4 Customer has integrated marketing function; ... \n",

"2 4 Customer has integrated marketing function; ... \n",

"3 3 Customer has integrated marketing function; ... \n",

"4 3 Customer has integrated marketing function; ... \n",

"\n",

" Define Customerâ€™s pre-sales scalability \\\n",

"0 3 Customer pre-Â­sales could scale up and there... \n",

"1 4 Customer pre-Â­sales is currently scaling up \n",

"2 3 Customer pre-Â­sales could scale up and there... \n",

"3 3 Customer pre-Â­sales could scale up and there... \n",

"4 4 Customer pre-Â­sales is currently scaling up \n",

"\n",

" Customer Account: Managed Customer Managed Customer Plan End Date \\\n",

"0 Yes 1 03/02/2023 \n",

"1 No 0 03/02/2023 \n",

"2 Yes 1 03/02/2023 \n",

"3 No 0 03/02/2023 \n",

"4 No 0 28/01/2022 \n",

"\n",

" Executive Sponspor Present Present Quarter Region Geo \\\n",

"0 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"1 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"2 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"3 0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"4 1 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \\\n",

"0 0 40 1 \n",

"1 0 2510 1 \n",

"2 0 91 0 \n",

"3 0 80 1 \n",

"4 0 150 0 \n",

"\n",

" KPI Compliant Failure Reason Place Country \n",

"0 NaN VIETNAM \n",

"1 NaN VIETNAM \n",

"2 Goals and Objectives- Min. 1 Goal Required VIETNAM \n",

"3 NaN VIETNAM \n",

"4 Goals and Objectives- Min. 1 Goal Required; Pl... VIETNAM "

]

},

"execution\_count": 28,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.head()"

]

},

{

"cell\_type": "code",

"execution\_count": 29,

"id": "b736a425",

"metadata": {},

"outputs": [],

"source": [

"df2 = pd.read\_csv (r\"C:\\Users\\dream\\OneDrive - VMware, Inc\\VMwareCorp\\Desktop\\Country data from Worldometer in UTF8.csv\")"

]

},

{

"cell\_type": "code",

"execution\_count": 30,

"id": "5011d1dc",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

" Place Country Population Yearly Change Net Change \\\n",

"0 Albania 2877797 -0.0011 -3120 \n",

"1 Algeria 43851044 0.0185 797990 \n",

"2 Argentina 45195774 0.0093 415097 \n",

"3 Australia 25499884 0.0118 296686 \n",

"4 Austria 9006398 0.0057 51296 \n",

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

"92 United Arab Emirates 9890402 0.0123 119873 \n",

"93 United Kingdom 67886011 0.0053 355839 \n",

"94 United States 331002651 0.0059 1937734 \n",

"95 Uruguay 3473730 0.0035 11996 \n",

"96 Vietnam 97338579 0.0091 876473 \n",

"\n",

" Density (P/KmÂ²) Land Area (KmÂ²) Migrants (net) Fert. Rate Med. Age \\\n",

"0 105 27400 -14000 1.6 36 \n",

"1 18 2381740 -10000 3.1 29 \n",

"2 17 2736690 4800 2.3 32 \n",

"3 3 7682300 158246 1.8 38 \n",

"4 109 82409 65000 1.5 43 \n",

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

"92 118 83600 40000 1.4 33 \n",

"93 281 241930 260650 1.8 40 \n",

"94 36 9147420 954806 1.8 38 \n",

"95 20 175020 -3000 2.0 36 \n",

"96 314 310070 -80000 2.1 32 \n",

"\n",

" Urban Pop % World Share \n",

"0 0.63 0.0004 \n",

"1 0.73 0.0056 \n",

"2 0.93 0.0058 \n",

"3 0.86 0.0033 \n",

"4 0.57 0.0012 \n",

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

"92 0.86 0.0013 \n",

"93 0.83 0.0087 \n",

"94 0.83 0.0425 \n",

"95 0.96 0.0004 \n",

"96 0.38 0.0125 \n",

"\n",

"[97 rows x 11 columns]\n"

]

}

],

"source": [

"print (df2)"

]

},

{

"cell\_type": "code",

"execution\_count": 31,

"id": "57283fef",

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

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

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

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

" <th>Density (P/KmÂ²)</th>\n",

" <th>Land Area (KmÂ²)</th>\n",

" <th>Migrants (net)</th>\n",

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

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

" <th>Urban Pop %</th>\n",

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

" <td>2877797</td>\n",

" <td>-0.0011</td>\n",

" <td>-3120</td>\n",

" <td>105</td>\n",

" <td>27400</td>\n",

" <td>-14000</td>\n",

" <td>1.6</td>\n",

" <td>36</td>\n",

" <td>0.63</td>\n",

" <td>0.0004</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" <td>43851044</td>\n",

" <td>0.0185</td>\n",

" <td>797990</td>\n",

" <td>18</td>\n",

" <td>2381740</td>\n",

" <td>-10000</td>\n",

" <td>3.1</td>\n",

" <td>29</td>\n",

" <td>0.73</td>\n",

" <td>0.0056</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" <td>45195774</td>\n",

" <td>0.0093</td>\n",

" <td>415097</td>\n",

" <td>17</td>\n",

" <td>2736690</td>\n",

" <td>4800</td>\n",

" <td>2.3</td>\n",

" <td>32</td>\n",

" <td>0.93</td>\n",

" <td>0.0058</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" <td>25499884</td>\n",

" <td>0.0118</td>\n",

" <td>296686</td>\n",

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

" <td>7682300</td>\n",

" <td>158246</td>\n",

" <td>1.8</td>\n",

" <td>38</td>\n",

" <td>0.86</td>\n",

" <td>0.0033</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" <td>9006398</td>\n",

" <td>0.0057</td>\n",

" <td>51296</td>\n",

" <td>109</td>\n",

" <td>82409</td>\n",

" <td>65000</td>\n",

" <td>1.5</td>\n",

" <td>43</td>\n",

" <td>0.57</td>\n",

" <td>0.0012</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Place Country Population Yearly Change Net Change Density (P/KmÂ²) \\\n",

"0 Albania 2877797 -0.0011 -3120 105 \n",

"1 Algeria 43851044 0.0185 797990 18 \n",

"2 Argentina 45195774 0.0093 415097 17 \n",

"3 Australia 25499884 0.0118 296686 3 \n",

"4 Austria 9006398 0.0057 51296 109 \n",

"\n",

" Land Area (KmÂ²) Migrants (net) Fert. Rate Med. Age Urban Pop % \\\n",

"0 27400 -14000 1.6 36 0.63 \n",

"1 2381740 -10000 3.1 29 0.73 \n",

"2 2736690 4800 2.3 32 0.93 \n",

"3 7682300 158246 1.8 38 0.86 \n",

"4 82409 65000 1.5 43 0.57 \n",

"\n",

" World Share \n",

"0 0.0004 \n",

"1 0.0056 \n",

"2 0.0058 \n",

"3 0.0033 \n",

"4 0.0012 "

]

},

"execution\_count": 31,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df2.head()"

]

},

{

"cell\_type": "markdown",

"id": "5478a0c9",

"metadata": {},

"source": [

"Using the below functions to review the Datatypes. And to review where there may be data issues. Missing values in row 4,21,22.\n",

"\n",

"4: D&B Address info. Leaving as is. This shows where there may be opportunities for data hygiene improvements in Customer mapping operations.\n",

"\n",

"21.Define Customerâ€™s marketing capability. Leaving as is. This is not a mandatory field in CRM (Salesforce). In this scenario, Null references are valuable. This Null reference will be interrupted as Limited Marketing Capability. Data will be shared with Marketing Operations, to follow up and define measures to handle new visibility.\n",

"\n",

"22.This field will be made mandatory in CRM (Salesforce).\n"

]

},

{

"cell\_type": "code",

"execution\_count": 32,

"id": "75fc47be",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

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

"RangeIndex: 1795 entries, 0 to 1794\n",

"Data columns (total 35 columns):\n",

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

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

" 0 Customer Business Plan: Plan # 1795 non-null object\n",

" 1 Customer Account: Account Name 1795 non-null object\n",

" 2 Customer Account: Zip/Postal Code 1795 non-null object\n",

" 3 Customer Account: Address 1 1795 non-null object\n",

" 4 Customer Account: D&B Address line 1 1477 non-null object\n",

" 5 Customer Business Plan: Created Date 1795 non-null object\n",

" 6 Created Quarter 1795 non-null object\n",

" 7 Customer Business Plan: Last Modified Date 1795 non-null object\n",

" 8 Year 1795 non-null int64 \n",

" 9 Plan Status 1795 non-null object\n",

" 10 Was the plan originally Approved? 1795 non-null int64 \n",

" 11 Customer Account: Is Direct Reseller 1795 non-null int64 \n",

" 12 Customer Account: Is Direct Service Provider 1795 non-null int64 \n",

" 13 Customer Account: Is Distributor Account 1795 non-null int64 \n",

" 14 Customer Account: Is Aggregator 1795 non-null int64 \n",

" 15 Customer Account: Is Services Only 1795 non-null int64 \n",

" 16 Customer Account: Is Cloud 1795 non-null int64 \n",

" 17 Customer Account: Is Resell 1795 non-null int64 \n",

" 18 Customer Account: Is Subscriptions Eligible 1795 non-null int64 \n",

" 19 Customer Account: Cloud Eligible 1795 non-null int64 \n",

" 20 Customer Account: Cloud Activated 1795 non-null int64 \n",

" 21 Define Customerâ€™s marketing capability 1380 non-null object\n",

" 22 Define Customerâ€™s pre-sales scalability 1372 non-null object\n",

" 23 Customer Account: Managed Customer 1795 non-null object\n",

" 24 Managed Customer 1795 non-null int64 \n",

" 25 Plan End Date 1795 non-null object\n",

" 26 Executive Sponspor Present Present 1795 non-null int64 \n",

" 27 Quarter 1795 non-null object\n",

" 28 Region 1795 non-null object\n",

" 29 Geo 1795 non-null object\n",

" 30 Skip Customer Approval 1795 non-null int64 \n",

" 31 Total # of Employees 1795 non-null int64 \n",

" 32 KPI Compliant 1795 non-null int64 \n",

" 33 KPI Compliant Failure Reason 499 non-null object\n",

" 34 Place Country 1795 non-null object\n",

"dtypes: int64(17), object(18)\n",

"memory usage: 490.9+ KB\n"

]

}

],

"source": [

"df1.info()"

]

},

{

"cell\_type": "code",

"execution\_count": 33,

"id": "fc19a365",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

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

"RangeIndex: 97 entries, 0 to 96\n",

"Data columns (total 11 columns):\n",

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

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

" 0 Place Country 97 non-null object \n",

" 1 Population 97 non-null int64 \n",

" 2 Yearly Change 97 non-null float64\n",

" 3 Net Change 97 non-null int64 \n",

" 4 Density (P/KmÂ²) 97 non-null int64 \n",

" 5 Land Area (KmÂ²) 97 non-null int64 \n",

" 6 Migrants (net) 97 non-null int64 \n",

" 7 Fert. Rate 97 non-null float64\n",

" 8 Med. Age 97 non-null int64 \n",

" 9 Urban Pop % 97 non-null object \n",

" 10 World Share 97 non-null float64\n",

"dtypes: float64(3), int64(6), object(2)\n",

"memory usage: 8.5+ KB\n"

]

}

],

"source": [

"df2.info()"

]

},

{

"cell\_type": "markdown",

"id": "d37c9cf3",

"metadata": {},

"source": [

"Using Seaborn Pairplot to see an overview of all the dataset visually. A pairplot, plots a pairwise relationships in a dataset. The pairplot function creates a grid of Axes such that each variable in data will by shared in the y-axis across a single row and in the x-axis across a single column. Source https://pythonbasics.org/seaborn-pairplot/#:~:text=A%20pairplot%20plot%20a%20pairwise,axis%20across%20a%20single%20column.\n"

]

},

{

"cell\_type": "code",

"execution\_count": 121,

"id": "cbab8893",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"<seaborn.axisgrid.PairGrid at 0x25fffbae910>"

]

},

"execution\_count": 121,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 3060x3060 with 306 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.pairplot(df1)"

]

},

{

"cell\_type": "markdown",

"id": "e0e099aa",

"metadata": {},

"source": [

"Using the same method below , with a limited number of datapoints. This allows me to investigate and analyse which Quarter are the highest number of Plans created? Q1 is the main Quarter for Plan creation. This aligns to the current Planning Program guidelines. However, when taken in conjunction with Plan creation decreases, this may present an opportunity to review guidelines. Do all Customers Plan to the same Fiscal year dates? Further analysis will be carried out on Planning over a 2- or 3-year period. Is Plan creation increasing or decreasing? In this scenario , decreasing. As customer attrition is an element of any transition to SaaS business models, this is an expected result. One of the new program metrics informed by this will be to increase Customer retention. New customer development is an existing company OKR. ( Objective & Key Result)"

]

},

{

"cell\_type": "code",

"execution\_count": 79,

"id": "02d46d85",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 720x720 with 20 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 4284x4284 with 306 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"g = sns.pairplot(df1, vars=[\"Created Quarter\", \"Year\",\"Geo\",\"Was the plan originally Approved?\" ])\n",

"g.fig.suptitle(\"Plan Overview\", y=1.05)\n",

"plt.rcParams[\"figure.figsize\"] = [8.5, 5.50]\n",

"g = sns.pairplot(df1, height=3.5)\n",

"plt.rcParams[\"figure.autolayout\"] = True\n"

]

},

{

"cell\_type": "code",

"execution\_count": 43,

"id": "077cddc3",

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

" <td>20</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>56</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>19</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

" <th>UNITED ARAB EMIRATES</th>\n",

" <td>25</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>87</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>181</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

" <td>18</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>97 rows Ã— 1 columns</p>\n",

"</div>"

],

"text/plain": [

" Plan Status\n",

"Place Country \n",

"ALBANIA 1\n",

"ALGERIA 1\n",

"ARGENTINA 20\n",

"AUSTRALIA 56\n",

"AUSTRIA 19\n",

"... ...\n",

"UNITED ARAB EMIRATES 25\n",

"UNITED KINGDOM 87\n",

"UNITED STATES 181\n",

"URUGUAY 4\n",

"VIETNAM 18\n",

"\n",

"[97 rows x 1 columns]"

]

},

"execution\_count": 43,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.pivot\_table(\n",

" index='Place Country',\n",

" values='Plan Status',\n",

" aggfunc=np.count\_nonzero\n",

")"

]

},

{

"cell\_type": "code",

"execution\_count": 35,

"id": "25ee3ae4",

"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>Was the plan originally Approved?</th>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1266</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

" <th>Pending Internal Approval</th>\n",

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

" </tr>\n",

" <tr>\n",

" <th>Pending re-approval</th>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

" <th>Submitted to Customer</th>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Was the plan originally Approved?\n",

"Plan Status \n",

"Approved 1266\n",

"Draft 3\n",

"Pending Internal Approval 0\n",

"Pending re-approval 4\n",

"Rejected 0\n",

"Submitted to Customer 2"

]

},

"execution\_count": 35,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.pivot\_table(\n",

" index='Plan Status',\n",

" values='Was the plan originally Approved?',\n",

" aggfunc=np.count\_nonzero\n",

")"

]

},

{

"cell\_type": "code",

"execution\_count": 123,

"id": "6606f90e",

"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>Customer Account: Is Distributor Account</th>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

" <td>6</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

" <th>UNITED ARAB EMIRATES</th>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>13</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>10</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>97 rows Ã— 1 columns</p>\n",

"</div>"

],

"text/plain": [

" Customer Account: Is Distributor Account\n",

"Place Country \n",

"ALBANIA 0\n",

"ALGERIA 0\n",

"ARGENTINA 6\n",

"AUSTRALIA 4\n",

"AUSTRIA 4\n",

"... ...\n",

"UNITED ARAB EMIRATES 8\n",

"UNITED KINGDOM 13\n",

"UNITED STATES 10\n",

"URUGUAY 0\n",

"VIETNAM 4\n",

"\n",

"[97 rows x 1 columns]"

]

},

"execution\_count": 123,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.pivot\_table(\n",

" index='Place Country',\n",

" values='Customer Account: Is Distributor Account',\n",

" aggfunc=np.count\_nonzero\n",

")"

]

},

{

"cell\_type": "code",

"execution\_count": 163,

"id": "77020588",

"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>count</th>\n",

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

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

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

" </tr>\n",

" <tr>\n",

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

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

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1296</td>\n",

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

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

" <td>661</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>269</td>\n",

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

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

" <td>138</td>\n",

" </tr>\n",

" <tr>\n",

" <th>Pending Internal Approval</th>\n",

" <td>14</td>\n",

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

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

" <td>9</td>\n",

" </tr>\n",

" <tr>\n",

" <th>Pending re-approval</th>\n",

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

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

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

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

" </tr>\n",

" <tr>\n",

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

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

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

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

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

" </tr>\n",

" <tr>\n",

" <th>Submitted to Customer</th>\n",

" <td>211</td>\n",

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

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

" <td>76</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" count unique top freq\n",

"Plan Status \n",

"Approved 1296 3 EMEA 661\n",

"Draft 269 3 APAC 138\n",

"Pending Internal Approval 14 3 APAC 9\n",

"Pending re-approval 4 1 AMER 4\n",

"Rejected 1 1 APAC 1\n",

"Submitted to Customer 211 3 AMER 76"

]

},

"execution\_count": 163,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.groupby('Plan Status')['Geo'].describe()"

]

},

{

"cell\_type": "code",

"execution\_count": 37,

"id": "a05e5ab7",

"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>Was the plan originally Approved?</th>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

" <td>11</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>15</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>17</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

" <th>UNITED ARAB EMIRATES</th>\n",

" <td>21</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>54</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>88</td>\n",

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

" <td>13</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>97 rows Ã— 1 columns</p>\n",

"</div>"

],

"text/plain": [

" Was the plan originally Approved?\n",

"Place Country \n",

"ALBANIA 1\n",

"ALGERIA 0\n",

"ARGENTINA 11\n",

"AUSTRALIA 15\n",

"AUSTRIA 17\n",

"... ...\n",

"UNITED ARAB EMIRATES 21\n",

"UNITED KINGDOM 54\n",

"UNITED STATES 88\n",

"URUGUAY 1\n",

"VIETNAM 13\n",

"\n",

"[97 rows x 1 columns]"

]

},

"execution\_count": 37,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.pivot\_table(\n",

" index='Place Country',\n",

" values='Was the plan originally Approved?',\n",

" aggfunc=np.count\_nonzero\n",

")"

]

},

{

"cell\_type": "code",

"execution\_count": 164,

"id": "23814081",

"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>count</th>\n",

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

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

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

" <th>25%</th>\n",

" <th>50%</th>\n",

" <th>75%</th>\n",

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

" </tr>\n",

" <tr>\n",

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

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

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

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

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

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

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1296.0</td>\n",

" <td>0.976852</td>\n",

" <td>0.150432</td>\n",

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

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>269.0</td>\n",

" <td>0.011152</td>\n",

" <td>0.105210</td>\n",

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

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

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

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

" <td>1.0</td>\n",

" </tr>\n",

" <tr>\n",

" <th>Pending Internal Approval</th>\n",

" <td>14.0</td>\n",

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

" <th>Pending re-approval</th>\n",

" <td>4.0</td>\n",

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

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

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" <td>1.0</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>1.0</td>\n",

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

" <th>Submitted to Customer</th>\n",

" <td>211.0</td>\n",

" <td>0.009479</td>\n",

" <td>0.097126</td>\n",

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

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

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

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

" <td>1.0</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" count mean std min 25% 50% 75% max\n",

"Plan Status \n",

"Approved 1296.0 0.976852 0.150432 0.0 1.0 1.0 1.0 1.0\n",

"Draft 269.0 0.011152 0.105210 0.0 0.0 0.0 0.0 1.0\n",

"Pending Internal Approval 14.0 0.000000 0.000000 0.0 0.0 0.0 0.0 0.0\n",

"Pending re-approval 4.0 1.000000 0.000000 1.0 1.0 1.0 1.0 1.0\n",

"Rejected 1.0 0.000000 NaN 0.0 0.0 0.0 0.0 0.0\n",

"Submitted to Customer 211.0 0.009479 0.097126 0.0 0.0 0.0 0.0 1.0"

]

},

"execution\_count": 164,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.groupby('Plan Status')['Was the plan originally Approved?'].describe()"

]

},

{

"cell\_type": "markdown",

"id": "a4529c33",

"metadata": {},

"source": [

"#### 3. Preparation\n",

"o Create pandas DataFrame\n",

"o Sorting, indexing, grouping\n",

"o Drop duplicates, replace missing values\n",

"o Merge DataFrames"

]

},

{

"cell\_type": "markdown",

"id": "d2ca997d",

"metadata": {},

"source": [

"#### Merging Dataframes : source : Using Merge DataFrames from tutorial : :https://www.datasciencelearner.com/how-to-join-two-csv-files-in-python-using-pandas/. Using The concat() function, as losing rows of data with other options. \n"

]

},

{

"cell\_type": "code",

"execution\_count": 66,

"id": "5ea93fee",

"metadata": {},

"outputs": [],

"source": [

"import glob\n"

]

},

{

"cell\_type": "code",

"execution\_count": 67,

"id": "31aaae92",

"metadata": {},

"outputs": [],

"source": [

"import os"

]

},

{

"cell\_type": "code",

"execution\_count": 173,

"id": "c483555c",

"metadata": {},

"outputs": [],

"source": [

"df1 = pd.read\_csv(r\"C:\\Users\\dream\\OneDrive - VMware, Inc\\VMwareCorp\\Desktop\\Dataset for Project 2 22.08.csv\")\n"

]

},

{

"cell\_type": "code",

"execution\_count": 175,

"id": "a1ae8364",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

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

"RangeIndex: 1795 entries, 0 to 1794\n",

"Data columns (total 35 columns):\n",

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

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

" 0 Customer Business Plan: Plan # 1795 non-null object\n",

" 1 Customer Account: Account Name 1795 non-null object\n",

" 2 Customer Account: Zip/Postal Code 1795 non-null object\n",

" 3 Customer Account: Address 1 1795 non-null object\n",

" 4 Customer Account: D&B Address line 1 1477 non-null object\n",

" 5 Customer Business Plan: Created Date 1795 non-null object\n",

" 6 Created Quarter 1795 non-null object\n",

" 7 Customer Business Plan: Last Modified Date 1795 non-null object\n",

" 8 Year 1795 non-null int64 \n",

" 9 Plan Status 1795 non-null object\n",

" 10 Was the plan originally Approved? 1795 non-null int64 \n",

" 11 Customer Account: Is Direct Reseller 1795 non-null int64 \n",

" 12 Customer Account: Is Direct Service Provider 1795 non-null int64 \n",

" 13 Customer Account: Is Distributor Account 1795 non-null int64 \n",

" 14 Customer Account: Is Aggregator 1795 non-null int64 \n",

" 15 Customer Account: Is Services Only 1795 non-null int64 \n",

" 16 Customer Account: Is Cloud 1795 non-null int64 \n",

" 17 Customer Account: Is Resell 1795 non-null int64 \n",

" 18 Customer Account: Is Subscriptions Eligible 1795 non-null int64 \n",

" 19 Customer Account: Cloud Eligible 1795 non-null int64 \n",

" 20 Customer Account: Cloud Activated 1795 non-null int64 \n",

" 21 Define Customerâ€™s marketing capability 1380 non-null object\n",

" 22 Define Customerâ€™s pre-sales scalability 1372 non-null object\n",

" 23 Customer Account: Managed Customer 1795 non-null object\n",

" 24 Managed Customer 1795 non-null int64 \n",

" 25 Plan End Date 1795 non-null object\n",

" 26 Executive Sponspor Present Present 1795 non-null int64 \n",

" 27 Quarter 1795 non-null object\n",

" 28 Region 1795 non-null object\n",

" 29 Geo 1795 non-null object\n",

" 30 Skip Customer Approval 1795 non-null int64 \n",

" 31 Total # of Employees 1795 non-null int64 \n",

" 32 KPI Compliant 1795 non-null int64 \n",

" 33 KPI Compliant Failure Reason 499 non-null object\n",

" 34 Place Country 1795 non-null object\n",

"dtypes: int64(17), object(18)\n",

"memory usage: 490.9+ KB\n"

]

}

],

"source": [

"df1.info()"

]

},

{

"cell\_type": "code",

"execution\_count": 174,

"id": "d811c0ed",

"metadata": {},

"outputs": [],

"source": [

"df2 = pd.read\_csv (r\"C:\\Users\\dream\\OneDrive - VMware, Inc\\VMwareCorp\\Desktop\\Country data from Worldometer in UTF8.csv\")\n"

]

},

{

"cell\_type": "code",

"execution\_count": 176,

"id": "a05b6bc6",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

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

"RangeIndex: 1795 entries, 0 to 1794\n",

"Data columns (total 35 columns):\n",

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

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

" 0 Customer Business Plan: Plan # 1795 non-null object\n",

" 1 Customer Account: Account Name 1795 non-null object\n",

" 2 Customer Account: Zip/Postal Code 1795 non-null object\n",

" 3 Customer Account: Address 1 1795 non-null object\n",

" 4 Customer Account: D&B Address line 1 1477 non-null object\n",

" 5 Customer Business Plan: Created Date 1795 non-null object\n",

" 6 Created Quarter 1795 non-null object\n",

" 7 Customer Business Plan: Last Modified Date 1795 non-null object\n",

" 8 Year 1795 non-null int64 \n",

" 9 Plan Status 1795 non-null object\n",

" 10 Was the plan originally Approved? 1795 non-null int64 \n",

" 11 Customer Account: Is Direct Reseller 1795 non-null int64 \n",

" 12 Customer Account: Is Direct Service Provider 1795 non-null int64 \n",

" 13 Customer Account: Is Distributor Account 1795 non-null int64 \n",

" 14 Customer Account: Is Aggregator 1795 non-null int64 \n",

" 15 Customer Account: Is Services Only 1795 non-null int64 \n",

" 16 Customer Account: Is Cloud 1795 non-null int64 \n",

" 17 Customer Account: Is Resell 1795 non-null int64 \n",

" 18 Customer Account: Is Subscriptions Eligible 1795 non-null int64 \n",

" 19 Customer Account: Cloud Eligible 1795 non-null int64 \n",

" 20 Customer Account: Cloud Activated 1795 non-null int64 \n",

" 21 Define Customerâ€™s marketing capability 1380 non-null object\n",

" 22 Define Customerâ€™s pre-sales scalability 1372 non-null object\n",

" 23 Customer Account: Managed Customer 1795 non-null object\n",

" 24 Managed Customer 1795 non-null int64 \n",

" 25 Plan End Date 1795 non-null object\n",

" 26 Executive Sponspor Present Present 1795 non-null int64 \n",

" 27 Quarter 1795 non-null object\n",

" 28 Region 1795 non-null object\n",

" 29 Geo 1795 non-null object\n",

" 30 Skip Customer Approval 1795 non-null int64 \n",

" 31 Total # of Employees 1795 non-null int64 \n",

" 32 KPI Compliant 1795 non-null int64 \n",

" 33 KPI Compliant Failure Reason 499 non-null object\n",

" 34 Place Country 1795 non-null object\n",

"dtypes: int64(17), object(18)\n",

"memory usage: 490.9+ KB\n"

]

}

],

"source": [

"df1.info()"

]

},

{

"cell\_type": "code",

"execution\_count": 44,

"id": "a61bb4f3",

"metadata": {},

"outputs": [],

"source": [

"merged\_data = [df1, df2]\n",

"merged\_data = pd.concat(merged\_data)\n"

]

},

{

"cell\_type": "code",

"execution\_count": 210,

"id": "1a24cee1",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

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

"RangeIndex: 1795 entries, 0 to 1794\n",

"Data columns (total 35 columns):\n",

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

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

" 0 Customer Business Plan: Plan # 1795 non-null object\n",

" 1 Customer Account: Account Name 1795 non-null object\n",

" 2 Customer Account: Zip/Postal Code 1795 non-null object\n",

" 3 Customer Account: Address 1 1795 non-null object\n",

" 4 Customer Account: D&B Address line 1 1477 non-null object\n",

" 5 Customer Business Plan: Created Date 1795 non-null object\n",

" 6 Created Quarter 1795 non-null object\n",

" 7 Customer Business Plan: Last Modified Date 1795 non-null object\n",

" 8 Year 1795 non-null int64 \n",

" 9 Plan Status 1795 non-null object\n",

" 10 Was the plan originally Approved? 1795 non-null int64 \n",

" 11 Customer Account: Is Direct Reseller 1795 non-null int64 \n",

" 12 Customer Account: Is Direct Service Provider 1795 non-null int64 \n",

" 13 Customer Account: Is Distributor Account 1795 non-null int64 \n",

" 14 Customer Account: Is Aggregator 1795 non-null int64 \n",

" 15 Customer Account: Is Services Only 1795 non-null int64 \n",

" 16 Customer Account: Is Cloud 1795 non-null int64 \n",

" 17 Customer Account: Is Resell 1795 non-null int64 \n",

" 18 Customer Account: Is Subscriptions Eligible 1795 non-null int64 \n",

" 19 Customer Account: Cloud Eligible 1795 non-null int64 \n",

" 20 Customer Account: Cloud Activated 1795 non-null int64 \n",

" 21 Define Customerâ€™s marketing capability 1380 non-null object\n",

" 22 Define Customerâ€™s pre-sales scalability 1372 non-null object\n",

" 23 Customer Account: Managed Customer 1795 non-null object\n",

" 24 Managed Customer 1795 non-null int64 \n",

" 25 Plan End Date 1795 non-null object\n",

" 26 Executive Sponspor Present Present 1795 non-null int64 \n",

" 27 Quarter 1795 non-null object\n",

" 28 Region 1795 non-null object\n",

" 29 Geo 1795 non-null object\n",

" 30 Skip Customer Approval 1795 non-null int64 \n",

" 31 Total # of Employees 1795 non-null int64 \n",

" 32 KPI Compliant 1795 non-null int64 \n",

" 33 KPI Compliant Failure Reason 499 non-null object\n",

" 34 Place Country 1795 non-null object\n",

"dtypes: int64(17), object(18)\n",

"memory usage: 490.9+ KB\n"

]

}

],

"source": [

"df1.info()"

]

},

{

"cell\_type": "code",

"execution\_count": 45,

"id": "526d9a2e",

"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>Customer Business Plan: Plan #</th>\n",

" <th>Customer Account: Account Name</th>\n",

" <th>Customer Account: Zip/Postal Code</th>\n",

" <th>Customer Account: Address 1</th>\n",

" <th>Customer Account: D&amp;B Address line 1</th>\n",

" <th>Customer Business Plan: Created Date</th>\n",

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

" <th>Customer Business Plan: Last Modified Date</th>\n",

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

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

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

" <th>Define Customerâ€™s marketing capability</th>\n",

" <th>Define Customerâ€™s pre-sales scalability</th>\n",

" <th>Customer Account: Managed Customer</th>\n",

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

" <th>Plan End Date</th>\n",

" <th>Executive Sponspor Present Present</th>\n",

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

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

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

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

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

" <th>KPI Compliant Failure Reason</th>\n",

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

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

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

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

" <th>Density (P/KmÂ²)</th>\n",

" <th>Land Area (KmÂ²)</th>\n",

" <th>Migrants (net)</th>\n",

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

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

" <th>Urban Pop %</th>\n",

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>CP-0000037961</td>\n",

" <td>ABC123</td>\n",

" <td>100000</td>\n",

" <td>Sá»‘ nhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T...</td>\n",

" <td>Sá»‘ NhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T...</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>13/04/2022</td>\n",

" <td>2023.0</td>\n",

" <td>Submitted to Customer</td>\n",

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

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

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

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

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

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

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

" <td>1.0</td>\n",

" <td>1.0</td>\n",

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

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

" <td>1.0</td>\n",

" <td>03/02/2023</td>\n",

" <td>1.0</td>\n",

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>40.0</td>\n",

" <td>1.0</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000037955</td>\n",

" <td>ABC300</td>\n",

" <td>84</td>\n",

" <td>22nd Floor, Keangnam Landmark 72 Tower E6,,</td>\n",

" <td>E6 Pham Hung Street</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>13/04/2022</td>\n",

" <td>2023.0</td>\n",

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

" <td>1.0</td>\n",

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

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

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

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

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

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

" <td>1.0</td>\n",

" <td>1.0</td>\n",

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

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

" <td>4 Customer has integrated marketing function; ...</td>\n",

" <td>4 Customer pre-Â­sales is currently scaling up</td>\n",

" <td>No</td>\n",

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

" <td>03/02/2023</td>\n",

" <td>1.0</td>\n",

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>2510.0</td>\n",

" <td>1.0</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000038896</td>\n",

" <td>ABC432</td>\n",

" <td>700000</td>\n",

" <td>Tuoi Tre Tower, No. 60A, Hoang Van Thu St</td>\n",

" <td>Tuoi Tre Tower, No. 60A, Hoang Van Thu St., Wa...</td>\n",

" <td>31/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>21/08/2022</td>\n",

" <td>2023.0</td>\n",

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

" <td>1.0</td>\n",

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

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

" <td>1.0</td>\n",

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

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

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

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

" <td>1.0</td>\n",

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

" <td>1.0</td>\n",

" <td>4 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

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

" <td>1.0</td>\n",

" <td>03/02/2023</td>\n",

" <td>1.0</td>\n",

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>91.0</td>\n",

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

" <td>Goals and Objectives- Min. 1 Goal Required</td>\n",

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000037959</td>\n",

" <td>ABC472</td>\n",

" <td>10000</td>\n",

" <td>18A - Cong Hoa Street, Ward 12, Tan Binh District</td>\n",

" <td>15 Duy Tan Dich Vong Hau</td>\n",

" <td>11/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>29/04/2022</td>\n",

" <td>2023.0</td>\n",

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

" <td>1.0</td>\n",

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

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

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

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

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

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

" <td>1.0</td>\n",

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

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

" <td>1.0</td>\n",

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

" <td>No</td>\n",

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>80.0</td>\n",

" <td>1.0</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000026118</td>\n",

" <td>ABC555</td>\n",

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

" <td>2nd floor, 2B1 Dam Trau, Bach Dang Ward,</td>\n",

" <td>7 A2 Dam Trau, Bach Dang Ward Hai Ba Trung</td>\n",

" <td>20/04/2021</td>\n",

" <td>Q2</td>\n",

" <td>02/03/2022</td>\n",

" <td>2022.0</td>\n",

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

" <td>1.0</td>\n",

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

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

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

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

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

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

" <td>1.0</td>\n",

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

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>4 Customer pre-Â­sales is currently scaling up</td>\n",

" <td>No</td>\n",

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

" <td>28/01/2022</td>\n",

" <td>1.0</td>\n",

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>150.0</td>\n",

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

" <td>Goals and Objectives- Min. 1 Goal Required; Pl...</td>\n",

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Customer Business Plan: Plan # Customer Account: Account Name \\\n",

"0 CP-0000037961 ABC123 \n",

"1 CP-0000037955 ABC300 \n",

"2 CP-0000038896 ABC432 \n",

"3 CP-0000037959 ABC472 \n",

"4 CP-0000026118 ABC555 \n",

"\n",

" Customer Account: Zip/Postal Code \\\n",

"0 100000 \n",

"1 84 \n",

"2 700000 \n",

"3 10000 \n",

"4 0 \n",

"\n",

" Customer Account: Address 1 \\\n",

"0 Sá»‘ nhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 22nd Floor, Keangnam Landmark 72 Tower E6,, \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St \n",

"3 18A - Cong Hoa Street, Ward 12, Tan Binh District \n",

"4 2nd floor, 2B1 Dam Trau, Bach Dang Ward, \n",

"\n",

" Customer Account: D&B Address line 1 \\\n",

"0 Sá»‘ NhÃ  17E, Tá»• 12A, PhÆ°á»ng ThÆ°á»£ng ÄÃ¬nh, Quáº­n T... \n",

"1 E6 Pham Hung Street \n",

"2 Tuoi Tre Tower, No. 60A, Hoang Van Thu St., Wa... \n",

"3 15 Duy Tan Dich Vong Hau \n",

"4 7 A2 Dam Trau, Bach Dang Ward Hai Ba Trung \n",

"\n",

" Customer Business Plan: Created Date Created Quarter \\\n",

"0 11/03/2022 Q1 \n",

"1 11/03/2022 Q1 \n",

"2 31/03/2022 Q1 \n",

"3 11/03/2022 Q1 \n",

"4 20/04/2021 Q2 \n",

"\n",

" Customer Business Plan: Last Modified Date Year Plan Status \\\n",

"0 13/04/2022 2023.0 Submitted to Customer \n",

"1 13/04/2022 2023.0 Approved \n",

"2 21/08/2022 2023.0 Approved \n",

"3 29/04/2022 2023.0 Approved \n",

"4 02/03/2022 2022.0 Approved \n",

"\n",

" Was the plan originally Approved? Customer Account: Is Direct Reseller \\\n",

"0 0.0 0.0 \n",

"1 1.0 0.0 \n",

"2 1.0 0.0 \n",

"3 1.0 0.0 \n",

"4 1.0 0.0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"0 0.0 \n",

"1 0.0 \n",

"2 0.0 \n",

"3 0.0 \n",

"4 0.0 \n",

"\n",

" Customer Account: Is Distributor Account Customer Account: Is Aggregator \\\n",

"0 0.0 0.0 \n",

"1 0.0 0.0 \n",

"2 1.0 0.0 \n",

"3 0.0 0.0 \n",

"4 0.0 0.0 \n",

"\n",

" Customer Account: Is Services Only Customer Account: Is Cloud \\\n",

"0 0.0 0.0 \n",

"1 0.0 0.0 \n",

"2 0.0 0.0 \n",

"3 0.0 0.0 \n",

"4 0.0 0.0 \n",

"\n",

" Customer Account: Is Resell Customer Account: Is Subscriptions Eligible \\\n",

"0 1.0 1.0 \n",

"1 1.0 1.0 \n",

"2 0.0 1.0 \n",

"3 1.0 0.0 \n",

"4 1.0 0.0 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"0 0.0 0.0 \n",

"1 0.0 0.0 \n",

"2 0.0 1.0 \n",

"3 0.0 1.0 \n",

"4 0.0 0.0 \n",

"\n",

" Define Customerâ€™s marketing capability \\\n",

"0 3 Customer has integrated marketing function; ... \n",

"1 4 Customer has integrated marketing function; ... \n",

"2 4 Customer has integrated marketing function; ... \n",

"3 3 Customer has integrated marketing function; ... \n",

"4 3 Customer has integrated marketing function; ... \n",

"\n",

" Define Customerâ€™s pre-sales scalability \\\n",

"0 3 Customer pre-Â­sales could scale up and there... \n",

"1 4 Customer pre-Â­sales is currently scaling up \n",

"2 3 Customer pre-Â­sales could scale up and there... \n",

"3 3 Customer pre-Â­sales could scale up and there... \n",

"4 4 Customer pre-Â­sales is currently scaling up \n",

"\n",

" Customer Account: Managed Customer Managed Customer Plan End Date \\\n",

"0 Yes 1.0 03/02/2023 \n",

"1 No 0.0 03/02/2023 \n",

"2 Yes 1.0 03/02/2023 \n",

"3 No 0.0 03/02/2023 \n",

"4 No 0.0 28/01/2022 \n",

"\n",

" Executive Sponspor Present Present Quarter Region Geo \\\n",

"0 1.0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"1 1.0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"2 1.0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"3 0.0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"4 1.0 Q1,Q2,Q3,Q4 SEAK-CHANNEL APAC \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \\\n",

"0 0.0 40.0 1.0 \n",

"1 0.0 2510.0 1.0 \n",

"2 0.0 91.0 0.0 \n",

"3 0.0 80.0 1.0 \n",

"4 0.0 150.0 0.0 \n",

"\n",

" KPI Compliant Failure Reason Place Country \\\n",

"0 NaN VIETNAM \n",

"1 NaN VIETNAM \n",

"2 Goals and Objectives- Min. 1 Goal Required VIETNAM \n",

"3 NaN VIETNAM \n",

"4 Goals and Objectives- Min. 1 Goal Required; Pl... VIETNAM \n",

"\n",

" Population Yearly Change Net Change Density (P/KmÂ²) Land Area (KmÂ²) \\\n",

"0 NaN NaN NaN NaN NaN \n",

"1 NaN NaN NaN NaN NaN \n",

"2 NaN NaN NaN NaN NaN \n",

"3 NaN NaN NaN NaN NaN \n",

"4 NaN NaN NaN NaN NaN \n",

"\n",

" Migrants (net) Fert. Rate Med. Age Urban Pop % World Share \n",

"0 NaN NaN NaN NaN NaN \n",

"1 NaN NaN NaN NaN NaN \n",

"2 NaN NaN NaN NaN NaN \n",

"3 NaN NaN NaN NaN NaN \n",

"4 NaN NaN NaN NaN NaN "

]

},

"execution\_count": 45,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data.head()"

]

},

{

"cell\_type": "markdown",

"id": "18c63a03",

"metadata": {},

"source": [

"####o Sorting, indexing, grouping\n",

"Grouping used in line 128. "

]

},

{

"cell\_type": "code",

"execution\_count": 188,

"id": "024062c9",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"0 False\n",

"1 False\n",

"2 False\n",

"3 False\n",

"4 False\n",

"5 False\n",

"6 False\n",

"7 False\n",

"8 False\n",

"9 False\n",

"10 False\n",

"11 False\n",

"12 False\n",

"13 False\n",

"14 False\n",

"15 False\n",

"16 False\n",

"17 False\n",

"18 False\n",

"19 False\n",

"20 False\n",

"21 False\n",

"22 False\n",

"23 False\n",

"24 False\n",

"25 False\n",

"26 False\n",

"27 False\n",

"28 False\n",

"29 False\n",

"30 False\n",

"31 False\n",

"32 False\n",

"33 False\n",

"34 False\n",

"35 False\n",

"36 False\n",

"37 False\n",

"38 False\n",

"39 False\n",

"40 False\n",

"41 False\n",

"42 False\n",

"43 False\n",

"44 False\n",

"45 False\n",

"46 False\n",

"47 False\n",

"48 False\n",

"dtype: bool"

]

},

"execution\_count": 188,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data.duplicated(keep=False)"

]

},

{

"cell\_type": "code",

"execution\_count": 192,

"id": "5a6fbc42",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"KPI Compliant Quarter \n",

"1 Q1,Q2,Q3,Q4 32\n",

"0 Q1,Q2,Q3,Q4 17\n",

"dtype: int64"

]

},

"execution\_count": 192,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data.groupby([\"KPI Compliant\", \"Quarter\"]).size().sort\_values(ascending=False)"

]

},

{

"cell\_type": "code",

"execution\_count": 193,

"id": "5aca3501",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"Created Quarter Year\n",

"Q1 2023 24\n",

" 2022 20\n",

"Q2 2022 4\n",

" 2023 1\n",

"dtype: int64"

]

},

"execution\_count": 193,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data.groupby([\"Created Quarter\", \"Year\"]).size().sort\_values(ascending=False)"

]

},

{

"cell\_type": "code",

"execution\_count": 146,

"id": "07cb0726",

"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>Customer Business Plan: Plan #</th>\n",

" <th>Customer Account: Account Name</th>\n",

" <th>Customer Account: Zip/Postal Code</th>\n",

" <th>Customer Account: Address 1</th>\n",

" <th>Customer Account: D&amp;B Address line 1</th>\n",

" <th>Customer Business Plan: Created Date</th>\n",

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

" <th>Customer Business Plan: Last Modified Date</th>\n",

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

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

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

" <th>Define Customerâ€™s marketing capability</th>\n",

" <th>Define Customerâ€™s pre-sales scalability</th>\n",

" <th>Customer Account: Managed Customer</th>\n",

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

" <th>Plan End Date</th>\n",

" <th>Executive Sponspor Present Present</th>\n",

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

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

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

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

" <th>KPI Compliant Failure Reason</th>\n",

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

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

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

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

" <th>Density (P/KmÂ²)</th>\n",

" <th>Land Area (KmÂ²)</th>\n",

" <th>Migrants (net)</th>\n",

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

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

" <th>Urban Pop %</th>\n",

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

" </tr>\n",

" <tr>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>CP-0000038786</td>\n",

" <td>ABC175</td>\n",

" <td>7217</td>\n",

" <td>W-1305,1306 SKV1 Center,11, Dangsan-ro 41-gil,...</td>\n",

" <td>43 Siheungdaero 70gil</td>\n",

" <td>27/03/2022</td>\n",

" <td>Q1</td>\n",

" <td>26/04/2022</td>\n",

" <td>2023</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>3 Customer has integrated marketing function; ...</td>\n",

" <td>3 Customer pre-Â­sales could scale up and there...</td>\n",

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

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

" <td>03/02/2023</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>31</td>\n",

" <td>Planner- Incomplete Planned Total FY &gt;1</td>\n",

" <td>South Korea</td>\n",

" <td>51269185</td>\n",

" <td>0.0009</td>\n",

" <td>43877</td>\n",

" <td>527</td>\n",

" <td>97230</td>\n",

" <td>11731</td>\n",

" <td>1.1</td>\n",

" <td>44</td>\n",

" <td>0.82</td>\n",

" <td>0.0066</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>CP-0000024379</td>\n",

" <td>ABC257</td>\n",

" <td>6101</td>\n",

" <td>620, Eonju-ro,</td>\n",

" <td>620 Eonju-ro, Gangnam-gu</td>\n",

" <td>01/02/2021</td>\n",

" <td>Q1</td>\n",

" <td>02/03/2022</td>\n",

" <td>2022</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

" <td>4 Customer has integrated marketing function; ...</td>\n",

" <td>4 Customer pre-Â­sales is currently scaling up</td>\n",

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

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

" <td>28/01/2022</td>\n",

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

" <td>Q1,Q2,Q3,Q4</td>\n",

" <td>SEAK-CHANNEL</td>\n",

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

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

" <td>176</td>\n",

" <td>None</td>\n",

" <td>South Korea</td>\n",

" <td>51269185</td>\n",

" <td>0.0009</td>\n",

" <td>43877</td>\n",

" <td>527</td>\n",

" <td>97230</td>\n",

" <td>11731</td>\n",

" <td>1.1</td>\n",

" <td>44</td>\n",

" <td>0.82</td>\n",

" <td>0.0066</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Customer Business Plan: Plan # Customer Account: Account Name \\\n",

"KPI Compliant \n",

"0 CP-0000038786 ABC175 \n",

"1 CP-0000024379 ABC257 \n",

"\n",

" Customer Account: Zip/Postal Code \\\n",

"KPI Compliant \n",

"0 7217 \n",

"1 6101 \n",

"\n",

" Customer Account: Address 1 \\\n",

"KPI Compliant \n",

"0 W-1305,1306 SKV1 Center,11, Dangsan-ro 41-gil,... \n",

"1 620, Eonju-ro, \n",

"\n",

" Customer Account: D&B Address line 1 \\\n",

"KPI Compliant \n",

"0 43 Siheungdaero 70gil \n",

"1 620 Eonju-ro, Gangnam-gu \n",

"\n",

" Customer Business Plan: Created Date Created Quarter \\\n",

"KPI Compliant \n",

"0 27/03/2022 Q1 \n",

"1 01/02/2021 Q1 \n",

"\n",

" Customer Business Plan: Last Modified Date Year Plan Status \\\n",

"KPI Compliant \n",

"0 26/04/2022 2023 Approved \n",

"1 02/03/2022 2022 Approved \n",

"\n",

" Was the plan originally Approved? \\\n",

"KPI Compliant \n",

"0 1 \n",

"1 1 \n",

"\n",

" Customer Account: Is Direct Reseller \\\n",

"KPI Compliant \n",

"0 0 \n",

"1 0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"KPI Compliant \n",

"0 0 \n",

"1 0 \n",

"\n",

" Customer Account: Is Distributor Account \\\n",

"KPI Compliant \n",

"0 0 \n",

"1 0 \n",

"\n",

" Customer Account: Is Aggregator \\\n",

"KPI Compliant \n",

"0 0 \n",

"1 0 \n",

"\n",

" Customer Account: Is Services Only Customer Account: Is Cloud \\\n",

"KPI Compliant \n",

"0 0 0 \n",

"1 0 0 \n",

"\n",

" Customer Account: Is Resell \\\n",

"KPI Compliant \n",

"0 1 \n",

"1 1 \n",

"\n",

" Customer Account: Is Subscriptions Eligible \\\n",

"KPI Compliant \n",

"0 1 \n",

"1 1 \n",

"\n",

" Customer Account: Cloud Eligible \\\n",

"KPI Compliant \n",

"0 1 \n",

"1 1 \n",

"\n",

" Customer Account: Cloud Activated \\\n",

"KPI Compliant \n",

"0 1 \n",

"1 1 \n",

"\n",

" Define Customerâ€™s marketing capability \\\n",

"KPI Compliant \n",

"0 3 Customer has integrated marketing function; ... \n",

"1 4 Customer has integrated marketing function; ... \n",

"\n",

" Define Customerâ€™s pre-sales scalability \\\n",

"KPI Compliant \n",

"0 3 Customer pre-Â­sales could scale up and there... \n",

"1 4 Customer pre-Â­sales is currently scaling up \n",

"\n",

" Customer Account: Managed Customer Managed Customer \\\n",

"KPI Compliant \n",

"0 Yes 1 \n",

"1 Yes 1 \n",

"\n",

" Plan End Date Executive Sponspor Present Present Quarter \\\n",

"KPI Compliant \n",

"0 03/02/2023 1 Q1,Q2,Q3,Q4 \n",

"1 28/01/2022 1 Q1,Q2,Q3,Q4 \n",

"\n",

" Region Geo Skip Customer Approval \\\n",

"KPI Compliant \n",

"0 SEAK-CHANNEL APAC 0 \n",

"1 SEAK-CHANNEL APAC 0 \n",

"\n",

" Total # of Employees KPI Compliant Failure Reason \\\n",

"KPI Compliant \n",

"0 31 Planner- Incomplete Planned Total FY >1 \n",

"1 176 None \n",

"\n",

" Place Country Population Yearly Change Net Change \\\n",

"KPI Compliant \n",

"0 South Korea 51269185 0.0009 43877 \n",

"1 South Korea 51269185 0.0009 43877 \n",

"\n",

" Density (P/KmÂ²) Land Area (KmÂ²) Migrants (net) Fert. Rate \\\n",

"KPI Compliant \n",

"0 527 97230 11731 1.1 \n",

"1 527 97230 11731 1.1 \n",

"\n",

" Med. Age Urban Pop % World Share \n",

"KPI Compliant \n",

"0 44 0.82 0.0066 \n",

"1 44 0.82 0.0066 "

]

},

"execution\_count": 146,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"gk = merged\_data.groupby('KPI Compliant')\n",

"gk.first()"

]

},

{

"cell\_type": "markdown",

"id": "902b7c33",

"metadata": {},

"source": [

"###Sorting"

]

},

{

"cell\_type": "code",

"execution\_count": 58,

"id": "405b7b57",

"metadata": {},

"outputs": [],

"source": [

"merged\_data.sort\_values(by=['Land Area (KmÂ²)'], inplace=True)"

]

},

{

"cell\_type": "markdown",

"id": "42e75512",

"metadata": {},

"source": [

"### Drop duplicates, replace missing values \n",

"Drop duplicates : keeping the last instance of each duplicate. keep{â€˜firstâ€™, â€˜lastâ€™, False}, default â€˜firstâ€™\n",

"Determines which duplicates (if any) to keep. - first : Drop duplicates except for the first occurrence. - last : Drop duplicates except for the last occurrence. - False : Drop all duplicates.\n",

"\n",

"See In 188 for Duplicate clean. Data loss on removing dups, not appropriate in this exercise. I want to see the Yearly Customer Plan, for each Customer. Duplication of data expected where Plan exists for FY22 & FY23. \n"

]

},

{

"cell\_type": "markdown",

"id": "29e5b9b9",

"metadata": {},

"source": [

"Replace a Value/ Column Name. Updating Column Year to Year Created, to allow for easier data interpretation. "

]

},

{

"cell\_type": "code",

"execution\_count": 239,

"id": "00424e2c",

"metadata": {},

"outputs": [],

"source": [

"merged\_data = merged\_data.rename(columns = {'Year':'Year Created'})"

]

},

{

"cell\_type": "markdown",

"id": "fd46823c",

"metadata": {},

"source": [

"#### 4. Analysis\n",

"o Conditional statements, looping, groupby\n",

"o Define a custom function to create reusable code\n",

"o Use NumPy functions\n",

"o Dictionary or Lists"

]

},

{

"cell\_type": "code",

"execution\_count": 235,

"id": "312224be",

"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>Year</th>\n",

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

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

" <th>Executive Sponspor Present Present</th>\n",

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

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

" <th>Population (2020)</th>\n",

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.0</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1.795000e+03</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.454039</td>\n",

" <td>0.710306</td>\n",

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

" <td>0.001114</td>\n",

" <td>0.167131</td>\n",

" <td>0.098607</td>\n",

" <td>0.013928</td>\n",

" <td>0.295265</td>\n",

" <td>0.798886</td>\n",

" <td>0.768802</td>\n",

" <td>0.550975</td>\n",

" <td>0.716435</td>\n",

" <td>0.693593</td>\n",

" <td>0.666852</td>\n",

" <td>0.004457</td>\n",

" <td>731.118663</td>\n",

" <td>0.568245</td>\n",

" <td>1.524749e+08</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.498022</td>\n",

" <td>0.453747</td>\n",

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

" <td>0.033370</td>\n",

" <td>0.373197</td>\n",

" <td>0.298217</td>\n",

" <td>0.117223</td>\n",

" <td>0.456289</td>\n",

" <td>0.400945</td>\n",

" <td>0.421716</td>\n",

" <td>0.497533</td>\n",

" <td>0.450854</td>\n",

" <td>0.461129</td>\n",

" <td>0.471470</td>\n",

" <td>0.066629</td>\n",

" <td>4676.174683</td>\n",

" <td>0.495459</td>\n",

" <td>3.102685e+08</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>0.000000e+00</td>\n",

" </tr>\n",

" <tr>\n",

" <th>25%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>1.077845e+07</td>\n",

" </tr>\n",

" <tr>\n",

" <th>50%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>54.000000</td>\n",

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

" <td>5.930869e+07</td>\n",

" </tr>\n",

" <tr>\n",

" <th>75%</th>\n",

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>193.000000</td>\n",

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

" <td>1.264765e+08</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>105049.000000</td>\n",

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

" <td>1.439324e+09</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Year Was the plan originally Approved? \\\n",

"count 1795.000000 1795.000000 \n",

"mean 2022.454039 0.710306 \n",

"std 0.498022 0.453747 \n",

"min 2022.000000 0.000000 \n",

"25% 2022.000000 0.000000 \n",

"50% 2022.000000 1.000000 \n",

"75% 2023.000000 1.000000 \n",

"max 2023.000000 1.000000 \n",

"\n",

" Customer Account: Is Direct Reseller \\\n",

"count 1795.0 \n",

"mean 0.0 \n",

"std 0.0 \n",

"min 0.0 \n",

"25% 0.0 \n",

"50% 0.0 \n",

"75% 0.0 \n",

"max 0.0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"count 1795.000000 \n",

"mean 0.001114 \n",

"std 0.033370 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Distributor Account \\\n",

"count 1795.000000 \n",

"mean 0.167131 \n",

"std 0.373197 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Aggregator Customer Account: Is Services Only \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.098607 0.013928 \n",

"std 0.298217 0.117223 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 0.000000 0.000000 \n",

"75% 0.000000 0.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Cloud Customer Account: Is Resell \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.295265 0.798886 \n",

"std 0.456289 0.400945 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 1.000000 \n",

"50% 0.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Subscriptions Eligible \\\n",

"count 1795.000000 \n",

"mean 0.768802 \n",

"std 0.421716 \n",

"min 0.000000 \n",

"25% 1.000000 \n",

"50% 1.000000 \n",

"75% 1.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.550975 0.716435 \n",

"std 0.497533 0.450854 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Managed Customer Executive Sponspor Present Present \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.693593 0.666852 \n",

"std 0.461129 0.471470 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \\\n",

"count 1795.000000 1795.000000 1795.000000 \n",

"mean 0.004457 731.118663 0.568245 \n",

"std 0.066629 4676.174683 0.495459 \n",

"min 0.000000 0.000000 0.000000 \n",

"25% 0.000000 0.000000 0.000000 \n",

"50% 0.000000 54.000000 1.000000 \n",

"75% 0.000000 193.000000 1.000000 \n",

"max 1.000000 105049.000000 1.000000 \n",

"\n",

" Population (2020) \n",

"count 1.795000e+03 \n",

"mean 1.524749e+08 \n",

"std 3.102685e+08 \n",

"min 0.000000e+00 \n",

"25% 1.077845e+07 \n",

"50% 5.930869e+07 \n",

"75% 1.264765e+08 \n",

"max 1.439324e+09 "

]

},

"execution\_count": 235,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df.describe()"

]

},

{

"cell\_type": "code",

"execution\_count": 236,

"id": "98e40575",

"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>Year</th>\n",

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

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

" <th>Executive Sponspor Present Present</th>\n",

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.0</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.454039</td>\n",

" <td>0.710306</td>\n",

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

" <td>0.001114</td>\n",

" <td>0.167131</td>\n",

" <td>0.098607</td>\n",

" <td>0.013928</td>\n",

" <td>0.295265</td>\n",

" <td>0.798886</td>\n",

" <td>0.768802</td>\n",

" <td>0.550975</td>\n",

" <td>0.716435</td>\n",

" <td>0.693593</td>\n",

" <td>0.666852</td>\n",

" <td>0.004457</td>\n",

" <td>731.118663</td>\n",

" <td>0.568245</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.498022</td>\n",

" <td>0.453747</td>\n",

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

" <td>0.033370</td>\n",

" <td>0.373197</td>\n",

" <td>0.298217</td>\n",

" <td>0.117223</td>\n",

" <td>0.456289</td>\n",

" <td>0.400945</td>\n",

" <td>0.421716</td>\n",

" <td>0.497533</td>\n",

" <td>0.450854</td>\n",

" <td>0.461129</td>\n",

" <td>0.471470</td>\n",

" <td>0.066629</td>\n",

" <td>4676.174683</td>\n",

" <td>0.495459</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

" <th>25%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" </tr>\n",

" <tr>\n",

" <th>50%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>54.000000</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>75%</th>\n",

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>193.000000</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>105049.000000</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Year Was the plan originally Approved? \\\n",

"count 1795.000000 1795.000000 \n",

"mean 2022.454039 0.710306 \n",

"std 0.498022 0.453747 \n",

"min 2022.000000 0.000000 \n",

"25% 2022.000000 0.000000 \n",

"50% 2022.000000 1.000000 \n",

"75% 2023.000000 1.000000 \n",

"max 2023.000000 1.000000 \n",

"\n",

" Customer Account: Is Direct Reseller \\\n",

"count 1795.0 \n",

"mean 0.0 \n",

"std 0.0 \n",

"min 0.0 \n",

"25% 0.0 \n",

"50% 0.0 \n",

"75% 0.0 \n",

"max 0.0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"count 1795.000000 \n",

"mean 0.001114 \n",

"std 0.033370 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Distributor Account \\\n",

"count 1795.000000 \n",

"mean 0.167131 \n",

"std 0.373197 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Aggregator Customer Account: Is Services Only \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.098607 0.013928 \n",

"std 0.298217 0.117223 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 0.000000 0.000000 \n",

"75% 0.000000 0.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Cloud Customer Account: Is Resell \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.295265 0.798886 \n",

"std 0.456289 0.400945 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 1.000000 \n",

"50% 0.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Subscriptions Eligible \\\n",

"count 1795.000000 \n",

"mean 0.768802 \n",

"std 0.421716 \n",

"min 0.000000 \n",

"25% 1.000000 \n",

"50% 1.000000 \n",

"75% 1.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.550975 0.716435 \n",

"std 0.497533 0.450854 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Managed Customer Executive Sponspor Present Present \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.693593 0.666852 \n",

"std 0.461129 0.471470 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \n",

"count 1795.000000 1795.000000 1795.000000 \n",

"mean 0.004457 731.118663 0.568245 \n",

"std 0.066629 4676.174683 0.495459 \n",

"min 0.000000 0.000000 0.000000 \n",

"25% 0.000000 0.000000 0.000000 \n",

"50% 0.000000 54.000000 1.000000 \n",

"75% 0.000000 193.000000 1.000000 \n",

"max 1.000000 105049.000000 1.000000 "

]

},

"execution\_count": 236,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df1.describe()"

]

},

{

"cell\_type": "code",

"execution\_count": 237,

"id": "4774994c",

"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>Population</th>\n",

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

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

" <th>Density (P/KmÂ²)</th>\n",

" <th>Land Area (KmÂ²)</th>\n",

" <th>Migrants (net)</th>\n",

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>5.615652e+07</td>\n",

" <td>-0.002206</td>\n",

" <td>-1.360239e+07</td>\n",

" <td>328.567010</td>\n",

" <td>1.052792e+06</td>\n",

" <td>17738.020619</td>\n",

" <td>2.140206</td>\n",

" <td>34.546392</td>\n",

" <td>0.007206</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>1.544058e+08</td>\n",

" <td>0.102795</td>\n",

" <td>1.387921e+08</td>\n",

" <td>1113.919206</td>\n",

" <td>2.506054e+06</td>\n",

" <td>154560.739621</td>\n",

" <td>0.862851</td>\n",

" <td>8.084219</td>\n",

" <td>0.019812</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.000000e+00</td>\n",

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

" <td>-1.366418e+09</td>\n",

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

" <td>3.000000e+02</td>\n",

" <td>-532687.000000</td>\n",

" <td>1.100000</td>\n",

" <td>18.000000</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>25%</th>\n",

" <td>5.792202e+06</td>\n",

" <td>0.001700</td>\n",

" <td>1.199600e+04</td>\n",

" <td>49.000000</td>\n",

" <td>6.271000e+04</td>\n",

" <td>-16000.000000</td>\n",

" <td>1.600000</td>\n",

" <td>29.000000</td>\n",

" <td>0.000700</td>\n",

" </tr>\n",

" <tr>\n",

" <th>50%</th>\n",

" <td>1.713487e+07</td>\n",

" <td>0.008700</td>\n",

" <td>8.790200e+04</td>\n",

" <td>103.000000</td>\n",

" <td>2.419300e+05</td>\n",

" <td>4800.000000</td>\n",

" <td>1.900000</td>\n",

" <td>34.000000</td>\n",

" <td>0.002200</td>\n",

" </tr>\n",

" <tr>\n",

" <th>75%</th>\n",

" <td>5.126918e+07</td>\n",

" <td>0.014100</td>\n",

" <td>4.475630e+05</td>\n",

" <td>225.000000</td>\n",

" <td>7.435320e+05</td>\n",

" <td>36527.000000</td>\n",

" <td>2.400000</td>\n",

" <td>42.000000</td>\n",

" <td>0.006600</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>1.439324e+09</td>\n",

" <td>0.029800</td>\n",

" <td>5.540090e+06</td>\n",

" <td>8358.000000</td>\n",

" <td>1.637687e+07</td>\n",

" <td>954806.000000</td>\n",

" <td>5.400000</td>\n",

" <td>48.000000</td>\n",

" <td>0.184700</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Population Yearly Change Net Change Density (P/KmÂ²) \\\n",

"count 9.700000e+01 97.000000 9.700000e+01 97.000000 \n",

"mean 5.615652e+07 -0.002206 -1.360239e+07 328.567010 \n",

"std 1.544058e+08 0.102795 1.387921e+08 1113.919206 \n",

"min 0.000000e+00 -1.000000 -1.366418e+09 0.000000 \n",

"25% 5.792202e+06 0.001700 1.199600e+04 49.000000 \n",

"50% 1.713487e+07 0.008700 8.790200e+04 103.000000 \n",

"75% 5.126918e+07 0.014100 4.475630e+05 225.000000 \n",

"max 1.439324e+09 0.029800 5.540090e+06 8358.000000 \n",

"\n",

" Land Area (KmÂ²) Migrants (net) Fert. Rate Med. Age World Share \n",

"count 9.700000e+01 97.000000 97.000000 97.000000 97.000000 \n",

"mean 1.052792e+06 17738.020619 2.140206 34.546392 0.007206 \n",

"std 2.506054e+06 154560.739621 0.862851 8.084219 0.019812 \n",

"min 3.000000e+02 -532687.000000 1.100000 18.000000 0.000000 \n",

"25% 6.271000e+04 -16000.000000 1.600000 29.000000 0.000700 \n",

"50% 2.419300e+05 4800.000000 1.900000 34.000000 0.002200 \n",

"75% 7.435320e+05 36527.000000 2.400000 42.000000 0.006600 \n",

"max 1.637687e+07 954806.000000 5.400000 48.000000 0.184700 "

]

},

"execution\_count": 237,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"df2.describe()"

]

},

{

"cell\_type": "code",

"execution\_count": 47,

"id": "48f6c9dd",

"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>Year</th>\n",

" <th>Was the plan originally Approved?</th>\n",

" <th>Customer Account: Is Direct Reseller</th>\n",

" <th>Customer Account: Is Direct Service Provider</th>\n",

" <th>Customer Account: Is Distributor Account</th>\n",

" <th>Customer Account: Is Aggregator</th>\n",

" <th>Customer Account: Is Services Only</th>\n",

" <th>Customer Account: Is Cloud</th>\n",

" <th>Customer Account: Is Resell</th>\n",

" <th>Customer Account: Is Subscriptions Eligible</th>\n",

" <th>Customer Account: Cloud Eligible</th>\n",

" <th>Customer Account: Cloud Activated</th>\n",

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

" <th>Executive Sponspor Present Present</th>\n",

" <th>Skip Customer Approval</th>\n",

" <th>Total # of Employees</th>\n",

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

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

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

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

" <th>Density (P/KmÂ²)</th>\n",

" <th>Land Area (KmÂ²)</th>\n",

" <th>Migrants (net)</th>\n",

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.0</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>1795.000000</td>\n",

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>9.700000e+01</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" <td>97.000000</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.454039</td>\n",

" <td>0.710306</td>\n",

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

" <td>0.001114</td>\n",

" <td>0.167131</td>\n",

" <td>0.098607</td>\n",

" <td>0.013928</td>\n",

" <td>0.295265</td>\n",

" <td>0.798886</td>\n",

" <td>0.768802</td>\n",

" <td>0.550975</td>\n",

" <td>0.716435</td>\n",

" <td>0.693593</td>\n",

" <td>0.666852</td>\n",

" <td>0.004457</td>\n",

" <td>731.118663</td>\n",

" <td>0.568245</td>\n",

" <td>5.615652e+07</td>\n",

" <td>-0.002206</td>\n",

" <td>-1.360239e+07</td>\n",

" <td>328.567010</td>\n",

" <td>1.052792e+06</td>\n",

" <td>17738.020619</td>\n",

" <td>2.140206</td>\n",

" <td>34.546392</td>\n",

" <td>0.007206</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.498022</td>\n",

" <td>0.453747</td>\n",

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

" <td>0.033370</td>\n",

" <td>0.373197</td>\n",

" <td>0.298217</td>\n",

" <td>0.117223</td>\n",

" <td>0.456289</td>\n",

" <td>0.400945</td>\n",

" <td>0.421716</td>\n",

" <td>0.497533</td>\n",

" <td>0.450854</td>\n",

" <td>0.461129</td>\n",

" <td>0.471470</td>\n",

" <td>0.066629</td>\n",

" <td>4676.174683</td>\n",

" <td>0.495459</td>\n",

" <td>1.544058e+08</td>\n",

" <td>0.102795</td>\n",

" <td>1.387921e+08</td>\n",

" <td>1113.919206</td>\n",

" <td>2.506054e+06</td>\n",

" <td>154560.739621</td>\n",

" <td>0.862851</td>\n",

" <td>8.084219</td>\n",

" <td>0.019812</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>0.000000e+00</td>\n",

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

" <td>-1.366418e+09</td>\n",

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

" <td>3.000000e+02</td>\n",

" <td>-532687.000000</td>\n",

" <td>1.100000</td>\n",

" <td>18.000000</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>25%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>5.792202e+06</td>\n",

" <td>0.001700</td>\n",

" <td>1.199600e+04</td>\n",

" <td>49.000000</td>\n",

" <td>6.271000e+04</td>\n",

" <td>-16000.000000</td>\n",

" <td>1.600000</td>\n",

" <td>29.000000</td>\n",

" <td>0.000700</td>\n",

" </tr>\n",

" <tr>\n",

" <th>50%</th>\n",

" <td>2022.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>54.000000</td>\n",

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

" <td>1.713487e+07</td>\n",

" <td>0.008700</td>\n",

" <td>8.790200e+04</td>\n",

" <td>103.000000</td>\n",

" <td>2.419300e+05</td>\n",

" <td>4800.000000</td>\n",

" <td>1.900000</td>\n",

" <td>34.000000</td>\n",

" <td>0.002200</td>\n",

" </tr>\n",

" <tr>\n",

" <th>75%</th>\n",

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

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

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

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

" <td>193.000000</td>\n",

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

" <td>5.126918e+07</td>\n",

" <td>0.014100</td>\n",

" <td>4.475630e+05</td>\n",

" <td>225.000000</td>\n",

" <td>7.435320e+05</td>\n",

" <td>36527.000000</td>\n",

" <td>2.400000</td>\n",

" <td>42.000000</td>\n",

" <td>0.006600</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2023.000000</td>\n",

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

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

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

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

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

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

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

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

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

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

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

" <td>1.000000</td>\n",

" <td>1.000000</td>\n",

" <td>1.000000</td>\n",

" <td>105049.000000</td>\n",

" <td>1.000000</td>\n",

" <td>1.439324e+09</td>\n",

" <td>0.029800</td>\n",

" <td>5.540090e+06</td>\n",

" <td>8358.000000</td>\n",

" <td>1.637687e+07</td>\n",

" <td>954806.000000</td>\n",

" <td>5.400000</td>\n",

" <td>48.000000</td>\n",

" <td>0.184700</td>\n",

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Year Was the plan originally Approved? \\\n",

"count 1795.000000 1795.000000 \n",

"mean 2022.454039 0.710306 \n",

"std 0.498022 0.453747 \n",

"min 2022.000000 0.000000 \n",

"25% 2022.000000 0.000000 \n",

"50% 2022.000000 1.000000 \n",

"75% 2023.000000 1.000000 \n",

"max 2023.000000 1.000000 \n",

"\n",

" Customer Account: Is Direct Reseller \\\n",

"count 1795.0 \n",

"mean 0.0 \n",

"std 0.0 \n",

"min 0.0 \n",

"25% 0.0 \n",

"50% 0.0 \n",

"75% 0.0 \n",

"max 0.0 \n",

"\n",

" Customer Account: Is Direct Service Provider \\\n",

"count 1795.000000 \n",

"mean 0.001114 \n",

"std 0.033370 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Distributor Account \\\n",

"count 1795.000000 \n",

"mean 0.167131 \n",

"std 0.373197 \n",

"min 0.000000 \n",

"25% 0.000000 \n",

"50% 0.000000 \n",

"75% 0.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Is Aggregator Customer Account: Is Services Only \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.098607 0.013928 \n",

"std 0.298217 0.117223 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 0.000000 0.000000 \n",

"75% 0.000000 0.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Cloud Customer Account: Is Resell \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.295265 0.798886 \n",

"std 0.456289 0.400945 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 1.000000 \n",

"50% 0.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Customer Account: Is Subscriptions Eligible \\\n",

"count 1795.000000 \n",

"mean 0.768802 \n",

"std 0.421716 \n",

"min 0.000000 \n",

"25% 1.000000 \n",

"50% 1.000000 \n",

"75% 1.000000 \n",

"max 1.000000 \n",

"\n",

" Customer Account: Cloud Eligible Customer Account: Cloud Activated \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.550975 0.716435 \n",

"std 0.497533 0.450854 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Managed Customer Executive Sponspor Present Present \\\n",

"count 1795.000000 1795.000000 \n",

"mean 0.693593 0.666852 \n",

"std 0.461129 0.471470 \n",

"min 0.000000 0.000000 \n",

"25% 0.000000 0.000000 \n",

"50% 1.000000 1.000000 \n",

"75% 1.000000 1.000000 \n",

"max 1.000000 1.000000 \n",

"\n",

" Skip Customer Approval Total # of Employees KPI Compliant \\\n",

"count 1795.000000 1795.000000 1795.000000 \n",

"mean 0.004457 731.118663 0.568245 \n",

"std 0.066629 4676.174683 0.495459 \n",

"min 0.000000 0.000000 0.000000 \n",

"25% 0.000000 0.000000 0.000000 \n",

"50% 0.000000 54.000000 1.000000 \n",

"75% 0.000000 193.000000 1.000000 \n",

"max 1.000000 105049.000000 1.000000 \n",

"\n",

" Population Yearly Change Net Change Density (P/KmÂ²) \\\n",

"count 9.700000e+01 97.000000 9.700000e+01 97.000000 \n",

"mean 5.615652e+07 -0.002206 -1.360239e+07 328.567010 \n",

"std 1.544058e+08 0.102795 1.387921e+08 1113.919206 \n",

"min 0.000000e+00 -1.000000 -1.366418e+09 0.000000 \n",

"25% 5.792202e+06 0.001700 1.199600e+04 49.000000 \n",

"50% 1.713487e+07 0.008700 8.790200e+04 103.000000 \n",

"75% 5.126918e+07 0.014100 4.475630e+05 225.000000 \n",

"max 1.439324e+09 0.029800 5.540090e+06 8358.000000 \n",

"\n",

" Land Area (KmÂ²) Migrants (net) Fert. Rate Med. Age World Share \n",

"count 9.700000e+01 97.000000 97.000000 97.000000 97.000000 \n",

"mean 1.052792e+06 17738.020619 2.140206 34.546392 0.007206 \n",

"std 2.506054e+06 154560.739621 0.862851 8.084219 0.019812 \n",

"min 3.000000e+02 -532687.000000 1.100000 18.000000 0.000000 \n",

"25% 6.271000e+04 -16000.000000 1.600000 29.000000 0.000700 \n",

"50% 2.419300e+05 4800.000000 1.900000 34.000000 0.002200 \n",

"75% 7.435320e+05 36527.000000 2.400000 42.000000 0.006600 \n",

"max 1.637687e+07 954806.000000 5.400000 48.000000 0.184700 "

]

},

"execution\_count": 47,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data.describe()"

]

},

{

"cell\_type": "code",

"execution\_count": 261,

"id": "7f68d49a",

"metadata": {},

"outputs": [],

"source": [

"my\_str = '1'\n",

"my\_list = ['36','29','32','38','43',\n",

"'28',\n",

"'42',\n",

"'26',\n",

"'24',\n",

"'33',\n",

"'45',\n",

"'41',\n",

"'35',\n",

"'31',\n",

"'44',\n",

"'19',\n",

"'25',\n",

"'46',\n",

"'22',\n",

"'23',\n",

"'30',\n",

"'21',\n",

"'47',\n",

"'48',\n",

"'20',\n",

"'37',\n",

"'40',\n",

"'18',\n",

"'34',\n",

"]\n",

"result1 = my\_list[int(my\_str)]"

]

},

{

"cell\_type": "code",

"execution\_count": 262,

"id": "2b6c0462",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

"29\n"

]

}

],

"source": [

"print(result1)"

]

},

{

"cell\_type": "code",

"execution\_count": 277,

"id": "6d44d089",

"metadata": {},

"outputs": [],

"source": [

"data = [36,\n",

"29,\n",

"32,\n",

"38,\n",

"43,\n",

"28,\n",

"42,\n",

"26,\n",

"24,\n",

"33,\n",

"45,\n",

"41,\n",

"35,\n",

"31,\n",

"44,\n",

"19,\n",

"25,\n",

"46,\n",

"22,\n",

"23,\n",

"30,\n",

"21,\n",

"47,\n",

"48,\n",

"20,\n",

"37,\n",

"40,\n",

"18,\n",

"34,\n",

"]\n",

"\n",

"df = pd.DataFrame(data, columns=['Numbers'])"

]

},

{

"cell\_type": "code",

"execution\_count": 278,

"id": "2e4423a3",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

" Numbers\n",

"0 36\n",

"1 29\n",

"2 32\n",

"3 38\n",

"4 43\n",

"5 28\n",

"6 42\n",

"7 26\n",

"8 24\n",

"9 33\n",

"10 45\n",

"11 41\n",

"12 35\n",

"13 31\n",

"14 44\n",

"15 19\n",

"16 25\n",

"17 46\n",

"18 22\n",

"19 23\n",

"20 30\n",

"21 21\n",

"22 47\n",

"23 48\n",

"24 20\n",

"25 37\n",

"26 40\n",

"27 18\n",

"28 34\n"

]

}

],

"source": [

"print (df)"

]

},

{

"cell\_type": "code",

"execution\_count": 308,

"id": "a998f244",

"metadata": {},

"outputs": [],

"source": [

"Mydict1={\"Place Country\":\"The Country\",\"Population\":\"The Number of people in the Country\",\"Yearly Change\": \"The Increase/Decrease in People\", \"Net Change\":\"The difference between days 1.3% roughly\",\"Density (P/KmÂ²)\":\"The number of people in one area (km)\",\"Land Area (KmÂ²)\": \"The space, used to calculate Density\",\"Migrants\":\"People moving to area\",\"Fert. Rate\" :\"How may babies born\",\"Med. Age\":\"Divide total into 2 groups\",\"Urban Pop %\":\"People in built up areas\",\"World Share\":\"what % of the total in the World\"}\n"

]

},

{

"cell\_type": "code",

"execution\_count": 309,

"id": "4c526e8a",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

"{'Place Country': 'The Country', 'Population': 'The Number of people in the Country', 'Yearly Change': 'The Increase/Decrease in People', 'Net Change': 'The difference between days 1.3% roughly', 'Density (P/KmÂ²)': 'The number of people in one area (km)', 'Land Area (KmÂ²)': 'The space, used to calculate Density', 'Migrants': 'People moving to area', 'Fert. Rate': 'How may babies born', 'Med. Age': 'Divide total into 2 groups', 'Urban Pop %': 'People in built up areas', 'World Share': 'what % of the total in the World'}\n"

]

}

],

"source": [

"print (Mydict1) "

]

},

{

"cell\_type": "code",

"execution\_count": 310,

"id": "04eb5b7c",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

"Place Country\n",

"Population\n",

"Yearly Change\n",

"Net Change\n",

"Density (P/KmÂ²)\n",

"Land Area (KmÂ²)\n",

"Migrants\n",

"Fert. Rate\n",

"Med. Age\n",

"Urban Pop %\n",

"World Share\n"

]

}

],

"source": [

"for x in Mydict1:\n",

" print(x)"

]

},

{

"cell\_type": "code",

"execution\_count": 330,

"id": "fd75c98a",

"metadata": {},

"outputs": [

{

"name": "stdout",

"output\_type": "stream",

"text": [

"The Country\n",

"The Number of people in the Country\n",

"The Increase/Decrease in People\n",

"The difference between days 1.3% roughly\n",

"The number of people in one area (km)\n",

"The space, used to calculate Density\n",

"People moving to area\n",

"How may babies born\n",

"Divide total into 2 groups\n",

"People in built up areas\n",

"what % of the total in the World\n"

]

}

],

"source": [

"for x in Mydict1:\n",

" print(Mydict1[x])"

]

},

{

"cell\_type": "markdown",

"id": "4dce7f9e",

"metadata": {},

"source": [

"#### 5. Visualisation\n",

"o Generate at least two charts using Matplotlib or Seaborn"

]

},

{

"cell\_type": "code",

"execution\_count": 347,

"id": "0374dc27",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"plt =merged\_data.plot(kind='scatter',x='Yearly Change', y='Fert. Rate')"

]

},

{

"cell\_type": "code",

"execution\_count": 86,

"id": "84582e30",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"Text(0.5, 1.0, 'World Wide Med. Age')"

]

},

"execution\_count": 86,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 612x396 with 2 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"plt =merged\_data.plot(kind='hexbin',x='World Share', bins='log',y='Med. Age',)\n",

"plt.set\_title(\"World Wide Med. Age\")"

]

},

{

"cell\_type": "code",

"execution\_count": 365,

"id": "84abf513",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Was the plan originally Approved?', ylabel='count'>"

]

},

"execution\_count": 365,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"import seaborn as sns\n",

"sns.countplot(x='Was the plan originally Approved?',data=merged\_data)"

]

},

{

"cell\_type": "code",

"execution\_count": 374,

"id": "0bbc6ea8",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Plan Status', ylabel='count'>"

]

},

"execution\_count": 374,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.countplot(x='Plan Status',data=merged\_data)"

]

},

{

"cell\_type": "code",

"execution\_count": 93,

"id": "7bafd986",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 612x396 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"ax = sns.countplot(x=\"Plan Status\", hue=\"Geo\", data=merged\_data )\n"

]

},

{

"cell\_type": "code",

"execution\_count": 401,

"id": "2df5b71e",

"metadata": {

"scrolled": true

},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {},

"output\_type": "display\_data"

}

],

"source": [

"ax = sns.countplot(y=\"Plan Status\",hue=\"Geo\", data=merged\_data,)\n",

" "

]

},

{

"cell\_type": "code",

"execution\_count": 101,

"id": "22a08d09",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 1504.62x360 with 2 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"g = sns.catplot(x=\"Plan Status\", hue=\"Geo\", col=\"Year\",\n",

" data=merged\_data, kind=\"count\",\n",

" height=5, aspect=2);"

]

},

{

"cell\_type": "code",

"execution\_count": 97,

"id": "81ab9355",

"metadata": {},

"outputs": [],

"source": [

"import seaborn as sns"

]

},

{

"cell\_type": "code",

"execution\_count": 49,

"id": "ffb61e1c",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 1447.03x288 with 6 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"g = sns.catplot(y=\"Plan Status\", hue=\"Geo\", col=\"Plan Status\",\n",

" data=merged\_data, kind=\"count\",\n",

" height=4, aspect=.8);"

]

},

{

"cell\_type": "code",

"execution\_count": 51,

"id": "683b8ae5",

"metadata": {},

"outputs": [

{

"data": {

"image/png": "\n",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"ax = sns.countplot(y=\"Plan Status\",hue=\"Geo\", data=merged\_data,)"

]

},

{

"cell\_type": "code",

"execution\_count": 55,

"id": "aa4d0427",

"metadata": {},

"outputs": [

{

"data": {

"text/plain": [

"Approved 1296\n",

"Draft 269\n",

"Submitted to Customer 211\n",

"Pending Internal Approval 14\n",

"Pending re-approval 4\n",

"Rejected 1\n",

"Name: Plan Status, dtype: int64"

]

},

"execution\_count": 55,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"merged\_data['Plan Status'].value\_counts()[:20]"

]

},

{

"cell\_type": "markdown",

"id": "3265d8da",

"metadata": {},

"source": [

"#### 6. Insights\n",

"o Derive five valuable insights from the analysis\n",

"o Justify your insights with reference to the charts or analysis"

]

},

{

"cell\_type": "markdown",

"id": "15d2fe3c",

"metadata": {},

"source": [

"##### Insight 1.0 \n",

"\n",

"Using the Pair plot Graph g = sns.pairplot, In line 25 Graph ;with a limited number of datapoints. This allowed me to investigate and analyze which Quarter are the highest number of Plans created? Q1 is the main Quarter for Plan creation. This aligns to the current Planning Program guidelines. However, when taken in conjunction with Plan creation decreases, this may present an opportunity to review guidelines. Do all Customers Plan to the same Fiscal year dates? Further analysis will be carried out on Planning over a 2- or 3-year period. Is Plan creation increasing or decreasing? In this scenario, decreasing. As customer attrition is an element of any transition to SaaS business models, this is an expected result. One of the new program metrics, informed by this will be to increase Customer retention. New customer development is an existing company OKR. (Objective & Key Result)\n"

]

},

{

"cell\_type": "markdown",

"id": "4fca89ab",

"metadata": {},

"source": [

"\n",

"##### Insight 2.0 \n",

"\n",

"Using the Pandas.DataFrame.info; df1.info() In 160 ; allowed me to review the Data by Datatype, index dtype and columns and non-null values.\n",

"This highlighted Data issues Missing values in row 4,21,22. These missing values were not picked up in monthly Data hygiene audits, as we have not looked at the data in this manner before. This provided for a new perspective and these data points will be added to future Dashboards. This will also be added to Customer Sales Teams, accountability matrix. We can now see the score, add the score to a compelling dashboard and act on it. \n",

"\n",

"4: D&B Address info.\n",

"Leaving as is. This shows where there may be opportunities for data hygiene improvements in Customer mapping operations. We are missing data about our customers and in todayâ€™s transition to SaaS, this needs to be rectified.\n",

"\n",

"21.Define Customerâ€™s marketing capability. \n",

"Leaving as is. This is not a mandatory field in CRM (Salesforce). In this scenario, Null references are valuable. This Null reference will be interrupted as Limited Marketing Capability. Data will be shared with Marketing Operations, to follow up and define measures to handle new visibility. Do new values need to be made available to Sales Operations, why is the value not added? Once we ascertain the why, this will become actionable.\n",

"22.Define Customerâ€™s pre-sales scalability\n",

"This field will be made mandatory in CRM (Salesforce). Data will be retrospectively added, at this yearâ€™s Customer contract renewal. On reviewing the options available in the field, I will also add a new value to account for No Pre-Sales Scalability.\n"

]

},

{

"cell\_type": "markdown",

"id": "48d6d170",

"metadata": {},

"source": [

"##### Insight 3.0\n",

"g = sns.catplot(y=\"Plan Status\", hue=\"Geo\", col=\"Plan Status\", In 406\n",

"This data visualization highlights Plan Status by Geo. \n",

"Pan Status-= Submitted to Customer: as the majority of Plans are created in Q1, at this stage in the year (Q3), we should not have more than a no less than 20 Plans awaiting Customer Approval. This highlights where we may have Customer Relationship issues or Sales Rep issues. It may also be an early indicator of Customer Attrition/Churn. This warrants further investigation. The highest number of Plans are from Geo = AMER, with a Frequency of 76.36% of all Plans, Submitted to Customers, are in AMER.\n",

" "

]

},

{

"cell\_type": "markdown",

"id": "39c8331b",

"metadata": {},

"source": [

"##### Insight 4\n",

"df1.groupby('Plan Status')['Geo'].describe() In 163\n",

"53% of Customer Plans Worldwide have failed Kpis Compliant audits, this will need to be highlighted to each Geos Executive Teams. Customer satisfaction is a major contributor to attrition/Churn. This early indicator allows for action to be taken to remedy the issues. \n",

"New Training, product developments can be informed with these findings. \n",

"\n"

]

},

{

"cell\_type": "markdown",

"id": "d4b27316",

"metadata": {},

"source": [

"##### Insight 5\n",

"\n",

"The Chart World Wide Med. Age (86) identifies the Median Age across the World Share of population. This infers where my company can focus efforts to attract new Customers. Although certain Geos/Regions/Countries may have higher World Share of population, they also have a higher Median Age. Opportunities in the new era of Cloud, growth in countries with a lower age cohort. There are markets not yet fully explored as new Regions to focus on. \n",

"Ref Gartner, â€œDespite a cloud spending compounded growth rate of 34.8% through 2022, China remains a lagging country (the group lagging by four or more years behind the U.S.). Beyond 2023, China is predicted to become a tracking country.â€\n",

"\*Contributor: Laurence Goasduff, Published August 19, 2019, Available at: < https://www.gartner.com/smarterwithgartner/cloud-adoption-where-does-your-country-rank (Accessed 25/08/2022).\n"

]

},

{

"cell\_type": "markdown",

"id": "5d2e0ed4",

"metadata": {},

"source": [

"#### 7. Machine Learning\n",

"o Describe what kind of prediction you could perform in future using\n",

"machine learning and/or deep learning.\n",

"o Would you use classification or regression methods?"

]

},

{

"cell\_type": "markdown",

"id": "2922790b",

"metadata": {},

"source": [

"Customer Churn : the value is obtained by dividing the total number of clients present at the beginning of a time frame by that of the end.This would fall into Classification. Churn = Yes/No. Using Machine Learning and analysis of the exisiting Data , we can apply Machine Learning to predict future Churn. And prehaps limit it with increased excellence in Customer Service and Product Developments. Number of Customers will become the Discreate value. Classification would be used to predict Customers leaving the company, in the future. This data couls then be built into exisiting Tableau Dashboards and aplplied, when reviewing data. Example Pipeline for next year. By Customers with an Apporved Plan in APAC. Adding the predicted Chrun value. "

]

},

{

"cell\_type": "markdown",

"id": "ed9e92f8",

"metadata": {},

"source": [

"Github Link : https://github.com/Dream43210/UCDPA\_Mary-Drea-Project-for-Data-Analysis-Course"

]

},

{

"cell\_type": "markdown",

"id": "848be0e9",

"metadata": {},

"source": []

},

{

"cell\_type": "markdown",

"id": "3bef53e1",

"metadata": {},

"source": []

}

],

"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": "3.9.12"

}

},

"nbformat": 4,

"nbformat\_minor": 5

}