```
{
 "nbformat": 4,
 "nbformat_minor": 0,
 "metadata": {
  "colab": {
   "provenance": [],
   "collapsed_sections": []
  },
  "kernelspec": {
   "name": "python3",
   "display_name": "Python 3"
  },
  "language_info": {
   "name": "python"
  }
 },
 "cells": [
  {
   "cell_type": "code",
   "execution_count": null,
   "metadata": {
    "id": "fvfwGvWkJFhg"
   },
   "outputs": [],
   "source": [
    "import pandas as pd\n",
    "import numpy as np\n",
```

```
    "import matplotlib.pyplot as plt"
   ]
  },
  {
   "cell_type": "code",
   "source": [
    "from google.colab import drive\n",
    "drive.mount('/content/drive')"
   ],
   "metadata": {
    "colab": {
     "base_uri": "https://localhost:8080/"
    },
    "id": "xVgKGPOnJLH2",
    "outputId": "b8fa85ed-62c4-4e9f-be22-46726b67bc00"
   },
   "execution_count": null,
   "outputs": [
    {
     "output_type": "stream",
     "name": "stdout",
     "text": [
      "Mounted at /content/drive\n"
     ]
    }
   ]
  },
```

```
{
 "cell_type": "code",
 "source": [
  "data = pd.read_csv(\"/content/drive/MyDrive/Dataset/Assignment 2/Churn_Modelling.csv\")"
 ],
 "metadata": {
  "id": "H_jg37-GKnEJ"
 },
 "execution_count": null,
 "outputs": []
},
{
 "cell_type": "code",
 "source": [
  "data.head() #univariate analysis"
 ],
 "metadata": {
  "colab": {
   "base_uri": "https://localhost:8080/",
   "height": 270
  },
  "id": "IEn5cIWtN8Xl",
  "outputId": "4e2bd6e7-8d53-4eb0-d022-eaa1d9940391"
 },
 "execution_count": null,
 "outputs": [
  {
```

```
"output_type": "execute_result",
"data": {
 "text/plain": [
  "   RowNumber  CustomerId   Surname  CreditScore Geography  Gender  Age  \\\n",
  "0          1    15634602  Hargrave          619    France  Female   42  \n",
  "1          2    15647311      Hill          608     Spain  Female   41  \n",
  "2          3    15619304      Onio          502    France  Female   42  \n",
  "3          4    15701354      Boni          699    France  Female   39  \n",
  "4          5    15737888  Mitchell          850     Spain  Female   43  \n",
  "\n",
  "   Tenure    Balance  NumOfProducts  HasCrCard  IsActiveMember  \\\n",
  "0       2       0.00              1          1               1  \n",
  "1       1   83807.86              1          0               1  \n",
  "2       8  159660.80              3          1               0  \n",
  "3       1       0.00              2          0               0  \n",
  "4       2  125510.82              1          1               1  \n",
  "\n",
  "   EstimatedSalary  Exited  \n",
  "0        101348.88       1  \n",
  "1        112542.58       0  \n",
  "2        113931.57       1  \n",
  "3         93826.63       0  \n",
  "4         79084.10       0  "
 ],
 "text/html": [
  "\n",
  "  <div id=\"df-9cc50632-9db8-4335-b67a-b85254c851cf\">\n",
```

```
"    <div class=\"colab-df-container\">\n",
"     <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
```

```
"     <th>Balance</th>\n",
"     <th>NumOfProducts</th>\n",
"     <th>HasCrCard</th>\n",
"     <th>IsActiveMember</th>\n",
"     <th>EstimatedSalary</th>\n",
"     <th>Exited</th>\n",
"   </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"   <tr>\n",
"     <th>0</th>\n",
"     <td>1</td>\n",
"     <td>15634602</td>\n",
"     <td>Hargrave</td>\n",
"     <td>619</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42</td>\n",
"     <td>2</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>101348.88</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
```

```
"     <th>1</th>\n",
"     <td>2</td>\n",
"     <td>15647311</td>\n",
"     <td>Hill</td>\n",
"     <td>608</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>41</td>\n",
"     <td>1</td>\n",
"     <td>83807.86</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>112542.58</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>2</th>\n",
"     <td>3</td>\n",
"     <td>15619304</td>\n",
"     <td>Onio</td>\n",
"     <td>502</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42</td>\n",
"     <td>8</td>\n",
"     <td>159660.80</td>\n",
```

```
"     <td>3</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>113931.57</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>3</th>\n",
"     <td>4</td>\n",
"     <td>15701354</td>\n",
"     <td>Boni</td>\n",
"     <td>699</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>39</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>93826.63</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>5</td>\n",
"     <td>15737888</td>\n",
```

```
"     <td>Mitchell</td>\n",

"     <td>850</td>\n",

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

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

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

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

"     <td>125510.82</td>\n",

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

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

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

"     <td>79084.10</td>\n",

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

"   </tr>\n",

"  </tbody>\n",

"</table>\n",

"</div>\n",

"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-9cc50632-9db8-4335-b67a-b85254c851cf')\"\n",

"             title=\"Convert this dataframe to an interactive table.\"\n",

"             style=\"display:none;\">\n",

"       \n",

"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",

"      width=\"24px\">\n",

"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",

"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
```

```
"  </svg>\n",
"    </button>\n",
"    \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
```

```
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-9cc50632-9db8-4335-b67a-b85254c851cf button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-9cc50632-9db8-4335-b67a-b85254c851cf');\n",
"          const dataTable =\n",
"            await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                     [key], {});\n",
"          if (!dataTable) return;\n",
```

```
          "\n",
          "         const docLinkHtml = 'Like what you see? Visit the ' +\n",
          "          '<a target=\"_blank\"
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
          "          + ' to learn more about interactive tables.';\n",
          "         element.innerHTML = '';\n",
          "         dataTable['output_type'] = 'display_data';\n",
          "         await google.colab.output.renderOutput(dataTable, element);\n",
          "         const docLink = document.createElement('div');\n",
          "         docLink.innerHTML = docLinkHtml;\n",
          "         element.appendChild(docLink);\n",
          "       }\n",
          "     </script>\n",
          "    </div>\n",
          "  </div>\n",
          "  "
         ]
        },
        "metadata": {},
        "execution_count": 6
      }
     ]
    },
    {
     "cell_type": "code",
     "source": [
       "pd.plotting.scatter_matrix(data.loc[:, \"Age\":\"Balance\"],
diagonal=\"kde\",figsize=(20,15))\n",
```

```
      "plt.show() #multivariate analysis"
     ],
     "metadata": {
      "colab": {
       "base_uri": "https://localhost:8080/",
       "height": 658
      },
      "id": "xNPxRy6lPtvn",
      "outputId": "7f4f32fc-13d5-4d80-f59a-2349a2be2887"
     },
     "execution_count": null,
     "outputs": [
      {
       "output_type": "display_data",
       "data": {
        "text/plain": [
         "<Figure size 1440x1080 with 9 Axes>"
        ],
        "image/png":
},
       "metadata": {
        "needs_background": "light"
       }
      }
     ]
    },
    {
```

```
"cell_type": "code",
"source": [
 "data.mean() #discriptive analysis"
],
"metadata": {
 "colab": {
  "base_uri": "https://localhost:8080/"
 },
 "id": "0SpPqrktuUOU",
 "outputId": "6788574c-8f93-4190-9a00-47896bf76c91"
},
"execution_count": 13,
"outputs": [
 {
  "output_type": "stream",
  "name": "stderr",
  "text": [
   "/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select only valid columns before calling the reduction.\n",
   "  \"\"\"Entry point for launching an IPython kernel.\n"
  ]
 },
 {
  "output_type": "execute_result",
  "data": {
   "text/plain": [
    "RowNumber         5.000500e+03\n",
```

```
        "CustomerId        1.569094e+07\n",
        "CreditScore       6.505288e+02\n",
        "Age               3.892180e+01\n",
        "Tenure            5.012800e+00\n",
        "Balance           7.648589e+04\n",
        "NumOfProducts     1.530200e+00\n",
        "HasCrCard         7.055000e-01\n",
        "IsActiveMember    5.151000e-01\n",
        "EstimatedSalary   1.000902e+05\n",
        "Exited            2.037000e-01\n",
        "dtype: float64"
       ]
      },
      "metadata": {},
      "execution_count": 13
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "data.isnull().sum() #missing values"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/"
   },
```

```
  "id": "uSm3ckJqvOV3",
  "outputId": "9999156b-bda8-4966-c175-e28d0f109943"
},
"execution_count": 17,
"outputs": [
 {
   "output_type": "execute_result",
   "data": {
     "text/plain": [
       "RowNumber         0\n",
       "CustomerId        0\n",
       "Surname           0\n",
       "CreditScore       0\n",
       "Geography         0\n",
       "Gender            0\n",
       "Age               0\n",
       "Tenure            0\n",
       "Balance           0\n",
       "NumOfProducts     0\n",
       "HasCrCard         0\n",
       "IsActiveMember    0\n",
       "EstimatedSalary   0\n",
       "Exited            0\n",
       "dtype: int64"
     ]
   },
   "metadata": {},
```

```
      "execution_count": 17
    }
  ]
},
{
  "cell_type": "code",
  "source": [
    "import seaborn as sns"
  ],
  "metadata": {
    "id": "2hyhODzv8YYi"
  },
  "execution_count": 18,
  "outputs": []
},
{
  "cell_type": "code",
  "source": [
    "q = data.quantile(q=[0.25,0.75]) #0utlier detection using upper and lower extreme\n",
    "q"
  ],
  "metadata": {
    "colab": {
      "base_uri": "https://localhost:8080/",
      "height": 112
    },
    "id": "GAZ1_rbB9OWu",
```

```
  "outputId": "f821796b-765f-4a11-8de3-2d60a5fd0492"
},
"execution_count": 20,
"outputs": [
 {
   "output_type": "execute_result",
   "data": {
     "text/plain": [
       "      RowNumber   CustomerId  CreditScore   Age  Tenure    Balance  \\\n",
       "0.25    2500.75  15628528.25       584.0  32.0     3.0       0.00   \n",
       "0.75    7500.25  15753233.75       718.0  44.0     7.0  127644.24   \n",
       "\n",
       "      NumOfProducts  HasCrCard  IsActiveMember  EstimatedSalary  Exited  \n",
       "0.25          1.0       0.0           0.0      51002.1100     0.0  \n",
       "0.75          2.0       1.0           1.0     149388.2475     0.0  "
     ],
     "text/html": [
       "\n",
       "  <div id=\"df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c\">\n",
       "    <div class=\"colab-df-container\">\n",
       "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0.25</th>\n",
```

```
"     <td>2500.75</td>\n",
"     <td>15628528.25</td>\n",
"     <td>584.0</td>\n",
"     <td>32.0</td>\n",
"     <td>3.0</td>\n",
"     <td>0.00</td>\n",
"     <td>1.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>51002.1100</td>\n",
"     <td>0.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>0.75</th>\n",
"     <td>7500.25</td>\n",
"     <td>15753233.75</td>\n",
"     <td>718.0</td>\n",
"     <td>44.0</td>\n",
"     <td>7.0</td>\n",
"     <td>127644.24</td>\n",
"     <td>2.0</td>\n",
"     <td>1.0</td>\n",
"     <td>1.0</td>\n",
"     <td>149388.2475</td>\n",
"     <td>0.0</td>\n",
"   </tr>\n",
" </tbody>\n",
```

```
"</table>\n",

"</div>\n",

"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c')\"\n",

"             title=\"Convert this dataframe to an interactive table.\"\n",

"             style=\"display:none;\">\n",

"       \n",

"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",

"      width=\"24px\">\n",

"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",

"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",

"  </svg>\n",

"     </button>\n",

"     \n",

"  <style>\n",

"    .colab-df-container {\n",

"      display:flex;\n",

"      flex-wrap:wrap;\n",

"      gap: 12px;\n",

"    }\n",

"\n",

"    .colab-df-convert {\n",

"      background-color: #E8F0FE;\n",

"      border: none;\n",

"      border-radius: 50%;\n",
```

```
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
```

```
"     <script>\n",
"       const buttonEl =\n",
"         document.querySelector('#df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c button.colab-df-convert');\n",
"       buttonEl.style.display =\n",
"         google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"       async function convertToInteractive(key) {\n",
"         const element = document.querySelector('#df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c');\n",
"         const dataTable =\n",
"           await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                    [key], {});\n",
"         if (!dataTable) return;\n",
"\n",
"         const docLinkHtml = 'Like what you see? Visit the ' +\n",
"           '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
"           + ' to learn more about interactive tables.';\n",
"         element.innerHTML = '';\n",
"         dataTable['output_type'] = 'display_data';\n",
"         await google.colab.output.renderOutput(dataTable, element);\n",
"         const docLink = document.createElement('div');\n",
"         docLink.innerHTML = docLinkHtml;\n",
"         element.appendChild(docLink);\n",
"       }\n",
"     </script>\n",
"   </div>\n",
"  </div>\n",
```

```
        "  "

       ]

      },

      "metadata": {},

      "execution_count": 20

    }

  ]

},

{

  "cell_type": "code",

  "source": [

    "IQR = q.loc[0.75]-q.loc[0.25]\n",

    "IQR"

  ],

  "metadata": {

    "colab": {

      "base_uri": "https://localhost:8080/"

    },

    "id": "EDXNvLEK-Sxt",

    "outputId": "de3dc014-6e40-4e05-ac34-4412f418397f"

  },

  "execution_count": 23,

  "outputs": [

    {

      "output_type": "execute_result",

      "data": {

        "text/plain": [
```

```
        "RowNumber           4999.5000\n",
        "CustomerId        124705.5000\n",
        "CreditScore          134.0000\n",
        "Age                  12.0000\n",
        "Tenure                4.0000\n",
        "Balance          127644.2400\n",
        "NumOfProducts         1.0000\n",
        "HasCrCard             1.0000\n",
        "IsActiveMember        1.0000\n",
        "EstimatedSalary    98386.1375\n",
        "Exited                0.0000\n",
        "dtype: float64"
       ]
     },
     "metadata": {},
     "execution_count": 23
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "upper_ex = q.loc[0.75]+1.5*IQR\n",
   "upper_ex"
  ],
  "metadata": {
   "colab": {
```

```
    "base_uri": "https://localhost:8080/"
   },
   "id": "P-2-du8y_NzJ",
   "outputId": "0efd5c62-7318-4e1a-8054-47a277024672"
  },
  "execution_count": 24,
  "outputs": [
   {
    "output_type": "execute_result",
    "data": {
     "text/plain": [
      "RowNumber         1.499950e+04\n",
      "CustomerId        1.594029e+07\n",
      "CreditScore       9.190000e+02\n",
      "Age               6.200000e+01\n",
      "Tenure            1.300000e+01\n",
      "Balance           3.191106e+05\n",
      "NumOfProducts     3.500000e+00\n",
      "HasCrCard         2.500000e+00\n",
      "IsActiveMember    2.500000e+00\n",
      "EstimatedSalary   2.969675e+05\n",
      "Exited            0.000000e+00\n",
      "dtype: float64"
     ]
    },
    "metadata": {},
    "execution_count": 24
```

```
   }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "lower_ex = q.loc[0.25]-1.5*IQR\n",
   "lower_ex"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/"
   },
   "id": "BXrzmTpK_9Hg",
   "outputId": "e2ed9f06-7b3c-4ad3-efd9-cf587f11dad1"
  },
  "execution_count": 25,
  "outputs": [
   {
    "output_type": "execute_result",
    "data": {
     "text/plain": [
      "RowNumber        -4.998500e+03\n",
      "CustomerId        1.544147e+07\n",
      "CreditScore       3.830000e+02\n",
      "Age               1.400000e+01\n",
      "Tenure           -3.000000e+00\n",
```

```
        "Balance          -1.914664e+05\n",
        "NumOfProducts     -5.000000e-01\n",
        "HasCrCard         -1.500000e+00\n",
        "IsActiveMember    -1.500000e+00\n",
        "EstimatedSalary   -9.657710e+04\n",
        "Exited            0.000000e+00\n",
        "dtype: float64"
       ]
      },
      "metadata": {},
      "execution_count": 25
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "data[data['Age']>62]"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 424
   },
   "id": "zls7NnYvB2wQ",
   "outputId": "c7c0f883-5710-4740-e75f-613a0d3b79cc"
  },
```

```
"execution_count": 31,
"outputs": [
 {
  "output_type": "execute_result",
  "data": {
   "text/plain": [
    "     RowNumber  CustomerId    Surname  CreditScore Geography  Gender  Age  \\\n",
    "58        59   15623944      T'ien        511    Spain  Female   66   \n",
    "85        86   15805254     Ndukaku       652    Spain  Female   75   \n",
    "104       105   15804919    Dunbabin       670    Spain  Female   65   \n",
    "158       159   15589975     Maclean       646   France  Female   73   \n",
    "181       182   15789669       Hsia        510   France    Male   65   \n",
    "...       ...       ...        ...        ...      ...     ...  ...   \n",
    "9753     9754   15705174    Chiedozie       656  Germany    Male   68   \n",
    "9765     9766   15777067     Thomas        445   France    Male   64   \n",
    "9832     9833   15814690  Chukwujekwu       595  Germany  Female   64   \n",
    "9894     9895   15704795      Vagin        521   France  Female   77   \n",
    "9936     9937   15653037      Parks        609   France    Male   77   \n",
    "\n",
    "     Tenure   Balance  NumOfProducts  HasCrCard  IsActiveMember  \\\n",
    "58       4      0.00              1          1               0   \n",
    "85      10      0.00              2          1               1   \n",
    "104      1      0.00              1          1               1   \n",
    "158      6  97259.25              1          0               1   \n",
    "181      2      0.00              2          1               1   \n",
    "...    ...       ...            ...        ...             ...   \n",
    "9753     7  153545.11              1          1               1   \n",
```

```
 "9765      2  136770.67           1        0           1  \n",
 "9832      2  105736.32           1        1           1  \n",
 "9894      6       0.00           2        1           1  \n",
 "9936      1       0.00           1        0           1  \n",
 "\n",
 "      EstimatedSalary  Exited  \n",
 "58           1643.11       1  \n",
 "85         114675.75       0  \n",
 "104        177655.68       1  \n",
 "158        104719.66       0  \n",
 "181         48071.61       0  \n",
 "...              ...     ...  \n",
 "9753       186574.68       0  \n",
 "9765        43678.06       0  \n",
 "9832        89935.73       1  \n",
 "9894        49054.10       0  \n",
 "9936        18708.76       0  \n",
 "\n",
 "[359 rows x 14 columns]"
],
"text/html": [
 "\n",
 "  <div id=\"df-9e085c94-3e79-436b-ac60-18162a30037b\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
```

```
"     <th>EstimatedSalary</th>\n",
"     <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"     <th>58</th>\n",
"     <td>59</td>\n",
"     <td>15623944</td>\n",
"     <td>T'ien</td>\n",
"     <td>511</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>66</td>\n",
"     <td>4</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1643.11</td>\n",
"     <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"     <th>85</th>\n",
"     <td>86</td>\n",
"     <td>15805254</td>\n",
"     <td>Ndukaku</td>\n",
```

```
"     <td>652</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>75</td>\n",
"     <td>10</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>114675.75</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>104</th>\n",
"     <td>105</td>\n",
"     <td>15804919</td>\n",
"     <td>Dunbabin</td>\n",
"     <td>670</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>65</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>177655.68</td>\n",
```

```
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>158</th>\n",
"     <td>159</td>\n",
"     <td>15589975</td>\n",
"     <td>Maclean</td>\n",
"     <td>646</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>73</td>\n",
"     <td>6</td>\n",
"     <td>97259.25</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>104719.66</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>181</th>\n",
"     <td>182</td>\n",
"     <td>15789669</td>\n",
"     <td>Hsia</td>\n",
"     <td>510</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
```

```
"    <td>65</td>\n",
"    <td>2</td>\n",
"    <td>0.00</td>\n",
"    <td>2</td>\n",
"    <td>1</td>\n",
"    <td>1</td>\n",
"    <td>48071.61</td>\n",
"    <td>0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>...</th>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"    <td>...</td>\n",
"  </tr>\n",
"  <tr>\n",
```

```
"     <th>9753</th>\n",
"     <td>9754</td>\n",
"     <td>15705174</td>\n",
"     <td>Chiedozie</td>\n",
"     <td>656</td>\n",
"     <td>Germany</td>\n",
"     <td>Male</td>\n",
"     <td>68</td>\n",
"     <td>7</td>\n",
"     <td>153545.11</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>186574.68</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9765</th>\n",
"     <td>9766</td>\n",
"     <td>15777067</td>\n",
"     <td>Thomas</td>\n",
"     <td>445</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
"     <td>64</td>\n",
"     <td>2</td>\n",
"     <td>136770.67</td>\n",
```

```
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>43678.06</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9832</th>\n",
"     <td>9833</td>\n",
"     <td>15814690</td>\n",
"     <td>Chukwujekwu</td>\n",
"     <td>595</td>\n",
"     <td>Germany</td>\n",
"     <td>Female</td>\n",
"     <td>64</td>\n",
"     <td>2</td>\n",
"     <td>105736.32</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>89935.73</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9894</th>\n",
"     <td>9895</td>\n",
"     <td>15704795</td>\n",
```

```
"     <td>Vagin</td>\n",
"     <td>521</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>77</td>\n",
"     <td>6</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>49054.10</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9936</th>\n",
"     <td>9937</td>\n",
"     <td>15653037</td>\n",
"     <td>Parks</td>\n",
"     <td>609</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
"     <td>77</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
```

```
"     <td>18708.76</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"  </tbody>\n",
"</table>\n",
"<p>359 rows Ã— 14 columns</p>\n",
"</div>\n",
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-9e085c94-3e79-436b-ac60-18162a30037b')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"     display:flex;\n",
"     flex-wrap:wrap;\n",
"     gap: 12px;\n",
"    }\n",
```

```
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
```

```
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-9e085c94-3e79-436b-ac60-18162a30037b button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-9e085c94-3e79-436b-ac60-18162a30037b');\n",
"          const dataTable =\n",
"            await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                     [key], {});\n",
"          if (!dataTable) return;\n",
"\n",
"          const docLinkHtml = 'Like what you see? Visit the ' +\n",
"            '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
"            + ' to learn more about interactive tables.';\n",
"          element.innerHTML = '';\n",
"          dataTable['output_type'] = 'display_data';\n",
"          await google.colab.output.renderOutput(dataTable, element);\n",
"          const docLink = document.createElement('div');\n",
"          docLink.innerHTML = docLinkHtml;\n",
```

```
        "         element.appendChild(docLink);\n",
        "       }\n",
        "     </script>\n",
        "   </div>\n",
        "  </div>\n",
        "  "
       ]
      },
      "metadata": {},
      "execution_count": 31
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "data[data['Age']<14]"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 49
   },
   "id": "m5PNqd0dCBxm",
   "outputId": "f255f515-142a-46ac-a116-799403812d23"
  },
  "execution_count": 32,
```

```
"outputs": [
 {
  "output_type": "execute_result",
  "data": {
   "text/plain": [
    "Empty DataFrame\n",
    "Columns: [RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary, Exited]\n",
    "Index: []"
   ],
   "text/html": [
    "\n",
    "  <div id=\"df-779b7cfd-007e-427b-82f0-663a72630694\">\n",
    "    <div class=\"colab-df-container\">\n",
    "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"      <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-779b7cfd-007e-427b-82f0-663a72630694')\"\n",
"              title=\"Convert this dataframe to an interactive table.\"\n",
"              style=\"display:none;\">\n",
```

```
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
```

```
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-779b7cfd-007e-427b-82f0-663a72630694 button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
```

```
        "         google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
        "\n",
        "       async function convertToInteractive(key) {\n",
        "         const element = document.querySelector('#df-779b7cfd-007e-427b-82f0-663a72630694');\n",
        "         const dataTable =\n",
        "           await google.colab.kernel.invokeFunction('convertToInteractive',\n",
        "                                                    [key], {});\n",
        "         if (!dataTable) return;\n",
        "\n",
        "         const docLinkHtml = 'Like what you see? Visit the ' +\n",
        "           '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
        "           + ' to learn more about interactive tables.';\n",
        "         element.innerHTML = '';\n",
        "         dataTable['output_type'] = 'display_data';\n",
        "         await google.colab.output.renderOutput(dataTable, element);\n",
        "         const docLink = document.createElement('div');\n",
        "         docLink.innerHTML = docLinkHtml;\n",
        "         element.appendChild(docLink);\n",
        "       }\n",
        "     </script>\n",
        "   </div>\n",
        "  </div>\n",
        "  "
      ]
    },
    "metadata": {},
```

```
      "execution_count": 32
     }
    ]
   },
   {
    "cell_type": "code",
    "source": [
     "data['Age'] = np.where(data['Age']>62,data['Age'].mean(),data['Age']) #replacing the outliers using mean value"
    ],
    "metadata": {
     "id": "5mJNa6rtLJbP"
    },
    "execution_count": 34,
    "outputs": []
   },
   {
    "cell_type": "code",
    "source": [
     "data[data['Age']>62]"
    ],
    "metadata": {
     "colab": {
      "base_uri": "https://localhost:8080/",
      "height": 49
     },
     "id": "XQcSNYuyL7Pk",
     "outputId": "89651c63-7904-4782-da33-2c38ff96faf2"
```

```
},
"execution_count": 35,
"outputs": [
 {
  "output_type": "execute_result",
  "data": {
   "text/plain": [
    "Empty DataFrame\n",
    "Columns: [RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary, Exited]\n",
    "Index: []"
   ],
   "text/html": [
    "\n",
    "  <div id=\"df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd\">\n",
    "    <div class=\"colab-df-container\">\n",
    "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"      <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd')\"\n",
```

```
"            title=\"Convert this dataframe to an interactive table.\"\n",
"            style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
```

```
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
```

```
        "         document.querySelector('#df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd
button.colab-df-convert');\n",
        "       buttonEl.style.display =\n",
        "         google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
        "\n",
        "       async function convertToInteractive(key) {\n",
        "         const element = document.querySelector('#df-b69ea559-76e1-44bd-bb60-
ca7e0c66c7bd');\n",
        "         const dataTable =\n",
        "           await google.colab.kernel.invokeFunction('convertToInteractive',\n",
        "                                                    [key], {});\n",
        "         if (!dataTable) return;\n",
        "\n",
        "         const docLinkHtml = 'Like what you see? Visit the ' +\n",
        "           '<a target=\"_blank\"
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
        "           + ' to learn more about interactive tables.';\n",
        "         element.innerHTML = '';\n",
        "         dataTable['output_type'] = 'display_data';\n",
        "         await google.colab.output.renderOutput(dataTable, element);\n",
        "         const docLink = document.createElement('div');\n",
        "         docLink.innerHTML = docLinkHtml;\n",
        "         element.appendChild(docLink);\n",
        "       }\n",
        "     </script>\n",
        "   </div>\n",
        "  </div>\n",
        "  "
       ]
```

```
    },
    "metadata": {},
    "execution_count": 35
   }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "data.head() #encoding"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 206
   },
   "id": "8FzUUUAnMbU2",
   "outputId": "fd237dce-7253-4523-e010-eb8662b16acf"
  },
  "execution_count": 36,
  "outputs": [
   {
    "output_type": "execute_result",
    "data": {
     "text/plain": [
      "   RowNumber  CustomerId   Surname  CreditScore Geography  Gender   Age  \\\n",
      "0          1    15634602  Hargrave          619    France  Female  42.0   \n",
```

```
 "1        2    15647311     Hill        608     Spain  Female  41.0   \n",
 "2        3    15619304     Onio        502     France  Female  42.0   \n",
 "3        4    15701354     Boni        699     France  Female  39.0   \n",
 "4        5    15737888  Mitchell       850     Spain  Female  43.0   \n",
 "\n",
 "   Tenure    Balance  NumOfProducts  HasCrCard  IsActiveMember  \\\n",
 "0      2       0.00           1        1           1   \n",
 "1      1   83807.86           1        0           1   \n",
 "2      8  159660.80           3        1           0   \n",
 "3      1       0.00           2        0           0   \n",
 "4      2  125510.82           1        1           1   \n",
 "\n",
 "   EstimatedSalary  Exited  \n",
 "0       101348.88       1  \n",
 "1       112542.58       0  \n",
 "2       113931.57       1  \n",
 "3        93826.63       0  \n",
 "4        79084.10       0  "
],
"text/html": [
 "\n",
 "  <div id=\"df-7c4f6fec-89db-40d5-9055-1d0dac2fc948\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
```

```
"     <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"     <th>0</th>\n",
"     <td>1</td>\n",
"     <td>15634602</td>\n",
"     <td>Hargrave</td>\n",
"     <td>619</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42.0</td>\n",
"     <td>2</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>101348.88</td>\n",
"     <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"     <th>1</th>\n",
"     <td>2</td>\n",
"     <td>15647311</td>\n",
"     <td>Hill</td>\n",
"     <td>608</td>\n",
```

```
"    <td>Spain</td>\n",
"    <td>Female</td>\n",
"    <td>41.0</td>\n",
"    <td>1</td>\n",
"    <td>83807.86</td>\n",
"    <td>1</td>\n",
"    <td>0</td>\n",
"    <td>1</td>\n",
"    <td>112542.58</td>\n",
"    <td>0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>2</th>\n",
"    <td>3</td>\n",
"    <td>15619304</td>\n",
"    <td>Onio</td>\n",
"    <td>502</td>\n",
"    <td>France</td>\n",
"    <td>Female</td>\n",
"    <td>42.0</td>\n",
"    <td>8</td>\n",
"    <td>159660.80</td>\n",
"    <td>3</td>\n",
"    <td>1</td>\n",
"    <td>0</td>\n",
"    <td>113931.57</td>\n",
"    <td>1</td>\n",
```

```
"   </tr>\n",
"   <tr>\n",
"     <th>3</th>\n",
"     <td>4</td>\n",
"     <td>15701354</td>\n",
"     <td>Boni</td>\n",
"     <td>699</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>39.0</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>93826.63</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>5</td>\n",
"     <td>15737888</td>\n",
"     <td>Mitchell</td>\n",
"     <td>850</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>43.0</td>\n",
```

```
"     <td>2</td>\n",
"     <td>125510.82</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>79084.10</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-7c4f6fec-89db-40d5-9055-1d0dac2fc948')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
```

```
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
```

```
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-7c4f6fec-89db-40d5-9055-1d0dac2fc948 button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-7c4f6fec-89db-40d5-9055-1d0dac2fc948');\n",
"          const dataTable =\n",
"            await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                     [key], {});\n",
"          if (!dataTable) return;\n",
"\n",
"          const docLinkHtml = 'Like what you see? Visit the ' +\n",
"            '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
"            + ' to learn more about interactive tables.';\n",
"          element.innerHTML = '';\n",
```

```
        "         dataTable['output_type'] = 'display_data';\n",
        "         await google.colab.output.renderOutput(dataTable, element);\n",
        "         const docLink = document.createElement('div');\n",
        "         docLink.innerHTML = docLinkHtml;\n",
        "         element.appendChild(docLink);\n",
        "       }\n",
        "     </script>\n",
        "    </div>\n",
        "  </div>\n",
        "  "
       ]
      },
      "metadata": {},
      "execution_count": 36
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "pd.get_dummies(data,columns=['Surname']) #encoding method-1"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 488
   },
```

```
 "id": "i2CPqBA9MaGt",
 "outputId": "0badeada-4bb1-44df-dffe-c9340fea2d4d"
},
"execution_count": 37,
"outputs": [
 {
  "output_type": "execute_result",
  "data": {
   "text/plain": [
    "     RowNumber  CustomerId  CreditScore Geography  Gender   Age  Tenure  \\\n",
    "0          1    15634602        619    France  Female  42.0      2  \n",
    "1          2    15647311        608     Spain  Female  41.0      1  \n",
    "2          3    15619304        502    France  Female  42.0      8  \n",
    "3          4    15701354        699    France  Female  39.0      1  \n",
    "4          5    15737888        850     Spain  Female  43.0      2  \n",
    "...      ...         ...        ...       ...     ...   ...    ...  \n",
    "9995    9996    15606229        771    France    Male  39.0      5  \n",
    "9996    9997    15569892        516    France    Male  35.0     10  \n",
    "9997    9998    15584532        709    France  Female  36.0      7  \n",
    "9998    9999    15682355        772   Germany    Male  42.0      3  \n",
    "9999   10000    15628319        792    France  Female  28.0      4  \n",
    "\n",
    "       Balance  NumOfProducts  HasCrCard  ...  Surname_Zinachukwudi  \\\n",
    "0         0.00              1          1  ...                     0  \n",
    "1     83807.86              1          0  ...                     0  \n",
    "2    159660.80              3          1  ...                     0  \n",
    "3         0.00              2          0  ...                     0  \n",
```

"4 125510.82 1 1 ... 0 \n",

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

"9995 0.00 2 1 ... 0 \n",

"9996 57369.61 1 1 ... 0 \n",

"9997 0.00 1 0 ... 0 \n",

"9998 75075.31 2 1 ... 0 \n",

"9999 130142.79 1 1 ... 0 \n",

"\n",

" Surname_Zito Surname_Zotov Surname_Zotova Surname_Zox \\\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",

"9995 0 0 0 0 \n",

"9996 0 0 0 0 \n",

"9997 0 0 0 0 \n",

"9998 0 0 0 0 \n",

"9999 0 0 0 0 \n",

"\n",

" Surname_Zubarev Surname_Zubareva Surname_Zuev Surname_Zuyev \\\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",
  "9995            0              0           0            0  \n",
  "9996            0              0           0            0  \n",
  "9997            0              0           0            0  \n",
  "9998            0              0           0            0  \n",
  "9999            0              0           0            0  \n",
  "\n",
  "     Surname_Zuyeva  \n",
  "0                 0  \n",
  "1                 0  \n",
  "2                 0  \n",
  "3                 0  \n",
  "4                 0  \n",
  "...             ...  \n",
  "9995              0  \n",
  "9996              0  \n",
  "9997              0  \n",
  "9998              0  \n",
  "9999              0  \n",
  "\n",
  "[10000 rows x 2945 columns]"
 ],
 "text/html": [
  "\n",
  "  <div id=\"df-a4595f0d-c407-4432-8dea-3affcf313c96\">\n",
  "    <div class=\"colab-df-container\">\n",
  "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
```

```
"     <th>...</th>\n",
"     <th>Surname_Zinachukwudi</th>\n",
"     <th>Surname_Zito</th>\n",
"     <th>Surname_Zotov</th>\n",
"     <th>Surname_Zotova</th>\n",
"     <th>Surname_Zox</th>\n",
"     <th>Surname_Zubarev</th>\n",
"     <th>Surname_Zubareva</th>\n",
"     <th>Surname_Zuev</th>\n",
"     <th>Surname_Zuyev</th>\n",
"     <th>Surname_Zuyeva</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"     <th>0</th>\n",
"     <td>1</td>\n",
"     <td>15634602</td>\n",
"     <td>619</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42.0</td>\n",
"     <td>2</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>...</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>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>1</th>\n",
"     <td>2</td>\n",
"     <td>15647311</td>\n",
"     <td>608</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>41.0</td>\n",
"     <td>1</td>\n",
"     <td>83807.86</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>...</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
```

```
          "     <td>0</td>\n",
          "    </tr>\n",
          "    <tr>\n",
          "     <th>3</th>\n",
          "     <td>4</td>\n",
          "     <td>15701354</td>\n",
          "     <td>699</td>\n",
          "     <td>France</td>\n",
          "     <td>Female</td>\n",
          "     <td>39.0</td>\n",
          "     <td>1</td>\n",
          "     <td>0.00</td>\n",
          "     <td>2</td>\n",
          "     <td>0</td>\n",
          "     <td>...</td>\n",
          "     <td>0</td>\n",
          "     <td>0</td>\n",
          "     <td>0</td>\n",
          "     <td>0</td>\n",
          "     <td>0</td>\n",
          …{
"nbformat": 4,
"nbformat_minor": 0,
"metadata": {
  "colab": {
    "provenance": [],
    "collapsed_sections": []
```

```
  },
  "kernelspec": {
    "name": "python3",
    "display_name": "Python 3"
  },
  "language_info": {
    "name": "python"
  }
 },
 "cells": [
  {
    "cell_type": "code",
    "execution_count": null,
    "metadata": {
      "id": "fvfwGvWkJFhg"
    },
    "outputs": [],
    "source": [
      "import pandas as pd\n",
      "import numpy as np\n",
      "import matplotlib.pyplot as plt"
    ]
  },
  {
    "cell_type": "code",
    "source": [
      "from google.colab import drive\n",
```

```
    "drive.mount('/content/drive')"
   ],
   "metadata": {
    "colab": {
     "base_uri": "https://localhost:8080/"
    },
    "id": "xVgKGPOnJLH2",
    "outputId": "b8fa85ed-62c4-4e9f-be22-46726b67bc00"
   },
   "execution_count": null,
   "outputs": [
    {
     "output_type": "stream",
     "name": "stdout",
     "text": [
      "Mounted at /content/drive\n"
     ]
    }
   ]
  },
  {
   "cell_type": "code",
   "source": [
    "data = pd.read_csv(\"/content/drive/MyDrive/Dataset/Assignment 2/Churn_Modelling.csv\")"
   ],
   "metadata": {
    "id": "H_jg37-GKnEJ"
```

```
 },
 "execution_count": null,
 "outputs": []
},
{
 "cell_type": "code",
 "source": [
  "data.head() #univariate analysis"
 ],
 "metadata": {
  "colab": {
   "base_uri": "https://localhost:8080/",
   "height": 270
  },
  "id": "IEn5cIWtN8Xl",
  "outputId": "4e2bd6e7-8d53-4eb0-d022-eaa1d9940391"
 },
 "execution_count": null,
 "outputs": [
  {
   "output_type": "execute_result",
   "data": {
    "text/plain": [
     "   RowNumber  CustomerId   Surname  CreditScore Geography  Gender  Age  \\\n",
     "0          1    15634602  Hargrave          619    France  Female   42  \n",
     "1          2    15647311      Hill          608     Spain  Female   41  \n",
     "2          3    15619304      Onio          502    France  Female   42  \n",
```

```
 "3        4    15701354     Boni       699    France  Female   39   \n",
 "4        5    15737888  Mitchell       850    Spain  Female   43   \n",
 "\n",
 "   Tenure    Balance  NumOfProducts  HasCrCard  IsActiveMember  \\\n",
 "0      2       0.00           1        1           1   \n",
 "1      1   83807.86           1        0           1   \n",
 "2      8  159660.80           3        1           0   \n",
 "3      1       0.00           2        0           0   \n",
 "4      2  125510.82           1        1           1   \n",
 "\n",
 "   EstimatedSalary  Exited  \n",
 "0       101348.88       1  \n",
 "1       112542.58       0  \n",
 "2       113931.57       1  \n",
 "3        93826.63       0  \n",
 "4        79084.10       0  "
],
"text/html": [
 "\n",
 "  <div id=\"df-9cc50632-9db8-4335-b67a-b85254c851cf\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
```

```
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0</th>\n",
"      <td>1</td>\n",
"      <td>15634602</td>\n",
"      <td>Hargrave</td>\n",
"      <td>619</td>\n",
"      <td>France</td>\n",
"      <td>Female</td>\n",
"      <td>42</td>\n",
"      <td>2</td>\n",
"      <td>0.00</td>\n",
"      <td>1</td>\n",
"      <td>1</td>\n",
"      <td>1</td>\n",
"      <td>101348.88</td>\n",
"      <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>1</th>\n",
"      <td>2</td>\n",
"      <td>15647311</td>\n",
"      <td>Hill</td>\n",
"      <td>608</td>\n",
"      <td>Spain</td>\n",
"      <td>Female</td>\n",
```

```
"     <td>41</td>\n",
"     <td>1</td>\n",
"     <td>83807.86</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>112542.58</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>2</th>\n",
"     <td>3</td>\n",
"     <td>15619304</td>\n",
"     <td>Onio</td>\n",
"     <td>502</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42</td>\n",
"     <td>8</td>\n",
"     <td>159660.80</td>\n",
"     <td>3</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>113931.57</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
```

```
"     <th>3</th>\n",
"     <td>4</td>\n",
"     <td>15701354</td>\n",
"     <td>Boni</td>\n",
"     <td>699</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>39</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>93826.63</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>5</td>\n",
"     <td>15737888</td>\n",
"     <td>Mitchell</td>\n",
"     <td>850</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>43</td>\n",
"     <td>2</td>\n",
"     <td>125510.82</td>\n",
```

```
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>79084.10</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-9cc50632-9db8-4335-b67a-b85254c851cf')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
```

```
"     gap: 12px;\n",
"   }\n",
"\n",
"   .colab-df-convert {\n",
"     background-color: #E8F0FE;\n",
"     border: none;\n",
"     border-radius: 50%;\n",
"     cursor: pointer;\n",
"     display: none;\n",
"     fill: #1967D2;\n",
"     height: 32px;\n",
"     padding: 0 0 0 0;\n",
"     width: 32px;\n",
"   }\n",
"\n",
"   .colab-df-convert:hover {\n",
"     background-color: #E2EBFA;\n",
"     box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
"     fill: #174EA6;\n",
"   }\n",
"\n",
"   [theme=dark] .colab-df-convert {\n",
"     background-color: #3B4455;\n",
"     fill: #D2E3FC;\n",
"   }\n",
"\n",
"   [theme=dark] .colab-df-convert:hover {\n",
```

```
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-9cc50632-9db8-4335-b67a-b85254c851cf button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-9cc50632-9db8-4335-b67a-b85254c851cf');\n",
"          const dataTable =\n",
"            await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                     [key], {});\n",
"          if (!dataTable) return;\n",
"\n",
"          const docLinkHtml = 'Like what you see? Visit the ' +\n",
"            '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
"            + ' to learn more about interactive tables.';\n",
"          element.innerHTML = '';\n",
"          dataTable['output_type'] = 'display_data';\n",
"          await google.colab.output.renderOutput(dataTable, element);\n",
```

```
          "         const docLink = document.createElement('div');\n",
          "         docLink.innerHTML = docLinkHtml;\n",
          "         element.appendChild(docLink);\n",
          "       }\n",
          "     </script>\n",
          "    </div>\n",
          "  </div>\n",
          "  "
         ]
        },
        "metadata": {},
        "execution_count": 6
       }
      ]
     },
     {
      "cell_type": "code",
      "source": [
       "pd.plotting.scatter_matrix(data.loc[:, \"Age\":\"Balance\"], diagonal=\"kde\",figsize=(20,15))\n",
       "plt.show() #multivariate analysis"
      ],
      "metadata": {
       "colab": {
        "base_uri": "https://localhost:8080/",
        "height": 658
       },
       "id": "xNPxRy6lPtvn",
```

```
      "outputId": "7f4f32fc-13d5-4d80-f59a-2349a2be2887"
     },
     "execution_count": null,
     "outputs": [
      {
       "output_type": "display_data",
       "data": {
        "text/plain": [
         "<Figure size 1440x1080 with 9 Axes>"
        ],
        "image/png":
},
       "metadata": {
        "needs_background": "light"
       }
      }
     ]
    },
    {
     "cell_type": "code",
     "source": [
      "data.mean() #discriptive analysis"
     ],
     "metadata": {
      "colab": {
       "base_uri": "https://localhost:8080/"
      },
```

```
  "id": "0SpPqrktuUOU",
  "outputId": "6788574c-8f93-4190-9a00-47896bf76c91"
 },
 "execution_count": 13,
 "outputs": [
  {
   "output_type": "stream",
   "name": "stderr",
   "text": [
    "/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select only valid columns before calling the reduction.\n",
    "  \"\"\"Entry point for launching an IPython kernel.\n"
   ]
  },
  {
   "output_type": "execute_result",
   "data": {
    "text/plain": [
     "RowNumber         5.000500e+03\n",
     "CustomerId        1.569094e+07\n",
     "CreditScore       6.505288e+02\n",
     "Age               3.892180e+01\n",
     "Tenure            5.012800e+00\n",
     "Balance           7.648589e+04\n",
     "NumOfProducts     1.530200e+00\n",
     "HasCrCard         7.055000e-01\n",
     "IsActiveMember    5.151000e-01\n",
```

```
        "EstimatedSalary    1.000902e+05\n",
        "Exited              2.037000e-01\n",
        "dtype: float64"
       ]
      },
      "metadata": {},
      "execution_count": 13
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "data.isnull().sum() #missing values"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/"
   },
   "id": "uSm3ckJqvOV3",
   "outputId": "9999156b-bda8-4966-c175-e28d0f109943"
  },
  "execution_count": 17,
  "outputs": [
   {
    "output_type": "execute_result",
    "data": {
```

```
      "text/plain": [
       "RowNumber         0\n",
       "CustomerId        0\n",
       "Surname           0\n",
       "CreditScore       0\n",
       "Geography         0\n",
       "Gender            0\n",
       "Age               0\n",
       "Tenure            0\n",
       "Balance           0\n",
       "NumOfProducts     0\n",
       "HasCrCard         0\n",
       "IsActiveMember    0\n",
       "EstimatedSalary   0\n",
       "Exited            0\n",
       "dtype: int64"
      ]
     },
     "metadata": {},
     "execution_count": 17
    }
   ]
  },
  {
   "cell_type": "code",
   "source": [
    "import seaborn as sns"
```

```
 ],
 "metadata": {
  "id": "2hyhODzv8YYi"
 },
 "execution_count": 18,
 "outputs": []
},
{
 "cell_type": "code",
 "source": [
  "q = data.quantile(q=[0.25,0.75]) #0utlier detection using upper and lower extreme\n",
  "q"
 ],
 "metadata": {
  "colab": {
   "base_uri": "https://localhost:8080/",
   "height": 112
  },
  "id": "GAZ1_rbB9OWu",
  "outputId": "f821796b-765f-4a11-8de3-2d60a5fd0492"
 },
 "execution_count": 20,
 "outputs": [
  {
   "output_type": "execute_result",
   "data": {
    "text/plain": [
```

```
  "     RowNumber   CustomerId  CreditScore   Age  Tenure    Balance  \\\n",
  "0.25    2500.75  15628528.25       584.0  32.0     3.0       0.00   \n",
  "0.75    7500.25  15753233.75       718.0  44.0     7.0  127644.24   \n",
  "\n",
  "     NumOfProducts  HasCrCard  IsActiveMember  EstimatedSalary  Exited  \n",
  "0.25          1.0       0.0             0.0      51002.1100     0.0  \n",
  "0.75          2.0       1.0             1.0     149388.2475     0.0  "
 ],
 "text/html": [
  "\n",
  "  <div id=\"df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c\">\n",
  "    <div class=\"colab-df-container\">\n",
  "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0.25</th>\n",
"      <td>2500.75</td>\n",
"      <td>15628528.25</td>\n",
"      <td>584.0</td>\n",
"      <td>32.0</td>\n",
"      <td>3.0</td>\n",
"      <td>0.00</td>\n",
"      <td>1.0</td>\n",
"      <td>0.0</td>\n",
```

```
"     <td>0.0</td>\n",
"     <td>51002.1100</td>\n",
"     <td>0.0</td>\n",
"    </tr>\n",
"    <tr>\n",
"     <th>0.75</th>\n",
"     <td>7500.25</td>\n",
"     <td>15753233.75</td>\n",
"     <td>718.0</td>\n",
"     <td>44.0</td>\n",
"     <td>7.0</td>\n",
"     <td>127644.24</td>\n",
"     <td>2.0</td>\n",
"     <td>1.0</td>\n",
"     <td>1.0</td>\n",
"     <td>149388.2475</td>\n",
"     <td>0.0</td>\n",
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-fc9f5a4f-ccad-
4dc9-8a4b-e1d26d6b439c')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
```

```
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",

"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",

"  </svg>\n",

"    </button>\n",

"    \n",

"  <style>\n",

"    .colab-df-container {\n",

"      display:flex;\n",

"      flex-wrap:wrap;\n",

"      gap: 12px;\n",

"    }\n",

"\n",

"    .colab-df-convert {\n",

"      background-color: #E8F0FE;\n",

"      border: none;\n",

"      border-radius: 50%;\n",

"      cursor: pointer;\n",

"      display: none;\n",

"      fill: #1967D2;\n",

"      height: 32px;\n",

"      padding: 0 0 0 0;\n",

"      width: 32px;\n",

"    }\n",

"\n",
```

```
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-fc9f5a4f-ccad-4dc9-8a4b-e1d26d6b439c button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
```

```
          "         const element = document.querySelector('#df-fc9f5a4f-ccad-4dc9-8a4b-
e1d26d6b439c');\n",

          "         const dataTable =\n",

          "           await google.colab.kernel.invokeFunction('convertToInteractive',\n",

          "                                                    [key], {});\n",

          "         if (!dataTable) return;\n",

          "\n",

          "         const docLinkHtml = 'Like what you see? Visit the ' +\n",

          "           '<a target=\"_blank\"
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",

          "           + ' to learn more about interactive tables.';\n",

          "         element.innerHTML = '';\n",

          "         dataTable['output_type'] = 'display_data';\n",

          "         await google.colab.output.renderOutput(dataTable, element);\n",

          "         const docLink = document.createElement('div');\n",

          "         docLink.innerHTML = docLinkHtml;\n",

          "         element.appendChild(docLink);\n",

          "       }\n",

          "     </script>\n",

          "    </div>\n",

          "  </div>\n",

          "  "

        ]

       },

       "metadata": {},

       "execution_count": 20

      }

     ]
```

```
},
{
 "cell_type": "code",
 "source": [
  "IQR = q.loc[0.75]-q.loc[0.25]\n",
  "IQR"
 ],
 "metadata": {
  "colab": {
   "base_uri": "https://localhost:8080/"
  },
  "id": "EDXNvLEK-Sxt",
  "outputId": "de3dc014-6e40-4e05-ac34-4412f418397f"
 },
 "execution_count": 23,
 "outputs": [
  {
   "output_type": "execute_result",
   "data": {
    "text/plain": [
     "RowNumber          4999.5000\n",
     "CustomerId       124705.5000\n",
     "CreditScore         134.0000\n",
     "Age                 12.0000\n",
     "Tenure               4.0000\n",
     "Balance          127644.2400\n",
     "NumOfProducts         1.0000\n",
```

```
        "HasCrCard              1.0000\n",
        "IsActiveMember         1.0000\n",
        "EstimatedSalary     98386.1375\n",
        "Exited                  0.0000\n",
        "dtype: float64"
       ]
      },
      "metadata": {},
      "execution_count": 23
    }
  ]
},
{
  "cell_type": "code",
  "source": [
    "upper_ex = q.loc[0.75]+1.5*IQR\n",
    "upper_ex"
  ],
  "metadata": {
    "colab": {
      "base_uri": "https://localhost:8080/"
    },
    "id": "P-2-du8y_NzJ",
    "outputId": "0efd5c62-7318-4e1a-8054-47a277024672"
  },
  "execution_count": 24,
  "outputs": [
```

```
   {
    "output_type": "execute_result",
    "data": {
     "text/plain": [
      "RowNumber         1.499950e+04\n",
      "CustomerId        1.594029e+07\n",
      "CreditScore       9.190000e+02\n",
      "Age               6.200000e+01\n",
      "Tenure            1.300000e+01\n",
      "Balance           3.191106e+05\n",
      "NumOfProducts     3.500000e+00\n",
      "HasCrCard         2.500000e+00\n",
      "IsActiveMember    2.500000e+00\n",
      "EstimatedSalary   2.969675e+05\n",
      "Exited            0.000000e+00\n",
      "dtype: float64"
     ]
    },
    "metadata": {},
    "execution_count": 24
   }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "lower_ex = q.loc[0.25]-1.5*IQR\n",
```

```
  "lower_ex"
 ],
 "metadata": {
  "colab": {
   "base_uri": "https://localhost:8080/"
  },
  "id": "BXrzmTpK_9Hg",
  "outputId": "e2ed9f06-7b3c-4ad3-efd9-cf587f11dad1"
 },
 "execution_count": 25,
 "outputs": [
  {
   "output_type": "execute_result",
   "data": {
    "text/plain": [
     "RowNumber        -4.998500e+03\n",
     "CustomerId        1.544147e+07\n",
     "CreditScore       3.830000e+02\n",
     "Age               1.400000e+01\n",
     "Tenure           -3.000000e+00\n",
     "Balance          -1.914664e+05\n",
     "NumOfProducts    -5.000000e-01\n",
     "HasCrCard        -1.500000e+00\n",
     "IsActiveMember   -1.500000e+00\n",
     "EstimatedSalary  -9.657710e+04\n",
     "Exited            0.000000e+00\n",
     "dtype: float64"
```

```
       ]
      },
      "metadata": {},
      "execution_count": 25
    }
  ]
},
{
  "cell_type": "code",
  "source": [
    "data[data['Age']>62]"
  ],
  "metadata": {
    "colab": {
      "base_uri": "https://localhost:8080/",
      "height": 424
    },
    "id": "zls7NnYvB2wQ",
    "outputId": "c7c0f883-5710-4740-e75f-613a0d3b79cc"
  },
  "execution_count": 31,
  "outputs": [
    {
      "output_type": "execute_result",
      "data": {
        "text/plain": [
          "      RowNumber  CustomerId     Surname  CreditScore Geography  Gender  Age  \\\n",
```

```
"58 59 15623944 T'ien 511 Spain Female 66 \n",
"85 86 15805254 Ndukaku 652 Spain Female 75 \n",
"104 105 15804919 Dunbabin 670 Spain Female 65 \n",
"158 159 15589975 Maclean 646 France Female 73 \n",
"181 182 15789669 Hsia 510 France Male 65 \n",
"... ... ... ... ... ... ... ... \n",
"9753 9754 15705174 Chiedozie 656 Germany Male 68 \n",
"9765 9766 15777067 Thomas 445 France Male 64 \n",
"9832 9833 15814690 Chukwujekwu 595 Germany Female 64 \n",
"9894 9895 15704795 Vagin 521 France Female 77 \n",
"9936 9937 15653037 Parks 609 France Male 77 \n",
"\n",
" Tenure Balance NumOfProducts HasCrCard IsActiveMember \\\n",
"58 4 0.00 1 1 0 \n",
"85 10 0.00 2 1 1 \n",
"104 1 0.00 1 1 1 \n",
"158 6 97259.25 1 0 1 \n",
"181 2 0.00 2 1 1 \n",
"... ... ... ... ... ... \n",
"9753 7 153545.11 1 1 1 \n",
"9765 2 136770.67 1 0 1 \n",
"9832 2 105736.32 1 1 1 \n",
"9894 6 0.00 2 1 1 \n",
"9936 1 0.00 1 0 1 \n",
"\n",
" EstimatedSalary Exited \n",
"58 1643.11 1 \n",
```

```
 "85         114675.75       0  \n",
 "104        177655.68       1  \n",
 "158        104719.66       0  \n",
 "181         48071.61       0  \n",
 "...              ...     ...  \n",
 "9753       186574.68       0  \n",
 "9765        43678.06       0  \n",
 "9832        89935.73       1  \n",
 "9894        49054.10       0  \n",
 "9936        18708.76       0  \n",
 "\n",
 "[359 rows x 14 columns]"
],
"text/html": [
 "\n",
 "  <div id=\"df-9e085c94-3e79-436b-ac60-18162a30037b\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>58</th>\n",
```

```
"     <td>59</td>\n",
"     <td>15623944</td>\n",
"     <td>T'ien</td>\n",
"     <td>511</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>66</td>\n",
"     <td>4</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1643.11</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>85</th>\n",
"     <td>86</td>\n",
"     <td>15805254</td>\n",
"     <td>Ndukaku</td>\n",
"     <td>652</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>75</td>\n",
"     <td>10</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
```

```
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>114675.75</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>104</th>\n",
"     <td>105</td>\n",
"     <td>15804919</td>\n",
"     <td>Dunbabin</td>\n",
"     <td>670</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>65</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>177655.68</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>158</th>\n",
"     <td>159</td>\n",
"     <td>15589975</td>\n",
"     <td>Maclean</td>\n",
```

```
"     <td>646</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>73</td>\n",
"     <td>6</td>\n",
"     <td>97259.25</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>104719.66</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>181</th>\n",
"     <td>182</td>\n",
"     <td>15789669</td>\n",
"     <td>Hsia</td>\n",
"     <td>510</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
"     <td>65</td>\n",
"     <td>2</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>48071.61</td>\n",
```

```
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>...</th>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9753</th>\n",
"     <td>9754</td>\n",
"     <td>15705174</td>\n",
"     <td>Chiedozie</td>\n",
"     <td>656</td>\n",
"     <td>Germany</td>\n",
"     <td>Male</td>\n",
```

```
"     <td>68</td>\n",
"     <td>7</td>\n",
"     <td>153545.11</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>186574.68</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9765</th>\n",
"     <td>9766</td>\n",
"     <td>15777067</td>\n",
"     <td>Thomas</td>\n",
"     <td>445</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
"     <td>64</td>\n",
"     <td>2</td>\n",
"     <td>136770.67</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>43678.06</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
```

```
"     <th>9832</th>\n",
"     <td>9833</td>\n",
"     <td>15814690</td>\n",
"     <td>Chukwujekwu</td>\n",
"     <td>595</td>\n",
"     <td>Germany</td>\n",
"     <td>Female</td>\n",
"     <td>64</td>\n",
"     <td>2</td>\n",
"     <td>105736.32</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>89935.73</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9894</th>\n",
"     <td>9895</td>\n",
"     <td>15704795</td>\n",
"     <td>Vagin</td>\n",
"     <td>521</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>77</td>\n",
"     <td>6</td>\n",
"     <td>0.00</td>\n",
```

```
"     <td>2</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>49054.10</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>9936</th>\n",
"     <td>9937</td>\n",
"     <td>15653037</td>\n",
"     <td>Parks</td>\n",
"     <td>609</td>\n",
"     <td>France</td>\n",
"     <td>Male</td>\n",
"     <td>77</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>1</td>\n",
"     <td>18708.76</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"  </tbody>\n",
"</table>\n",
"<p>359 rows Ã— 14 columns</p>\n",
"</div>\n",
```

```
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-9e085c94-3e79-436b-ac60-18162a30037b')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
```

```
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
```

```
        "         document.querySelector('#df-9e085c94-3e79-436b-ac60-18162a30037b
button.colab-df-convert');\n",
        "       buttonEl.style.display =\n",
        "         google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
        "\n",
        "       async function convertToInteractive(key) {\n",
        "         const element = document.querySelector('#df-9e085c94-3e79-436b-ac60-
18162a30037b');\n",
        "         const dataTable =\n",
        "           await google.colab.kernel.invokeFunction('convertToInteractive',\n",
        "                                                    [key], {});\n",
        "         if (!dataTable) return;\n",
        "\n",
        "         const docLinkHtml = 'Like what you see? Visit the ' +\n",
        "           '<a target=\"_blank\"
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
        "           + ' to learn more about interactive tables.';\n",
        "         element.innerHTML = '';\n",
        "         dataTable['output_type'] = 'display_data';\n",
        "         await google.colab.output.renderOutput(dataTable, element);\n",
        "         const docLink = document.createElement('div');\n",
        "         docLink.innerHTML = docLinkHtml;\n",
        "         element.appendChild(docLink);\n",
        "       }\n",
        "     </script>\n",
        "   </div>\n",
        "  </div>\n",
        "  "
       ]
```

```
      },
      "metadata": {},
      "execution_count": 31
     }
    ]
   },
   {
    "cell_type": "code",
    "source": [
     "data[data['Age']<14]"
    ],
    "metadata": {
     "colab": {
      "base_uri": "https://localhost:8080/",
      "height": 49
     },
     "id": "m5PNqd0dCBxm",
     "outputId": "f255f515-142a-46ac-a116-799403812d23"
    },
    "execution_count": 32,
    "outputs": [
     {
      "output_type": "execute_result",
      "data": {
       "text/plain": [
        "Empty DataFrame\n",
        "Columns: [RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary, Exited]\n",
```

```
  "Index: []"
 ],
 "text/html": [
  "\n",
  "  <div id=\"df-779b7cfd-007e-427b-82f0-663a72630694\">\n",
  "    <div class=\"colab-df-container\">\n",
  "      <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>RowNumber</th>\n",
  "      <th>CustomerId</th>\n",
  "      <th>Surname</th>\n",
```

```
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"      <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-779b7cfd-007e-427b-82f0-663a72630694')\"\n",
"              title=\"Convert this dataframe to an interactive table.\"\n",
"              style=\"display:none;\">\n",
"        \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"       width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-
```

```
.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41
20z\"/>\n",
"  </svg>\n",
"    </button>\n",
"    \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
```

```
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-779b7cfd-007e-427b-82f0-663a72630694 button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-779b7cfd-007e-427b-82f0-663a72630694');\n",
"          const dataTable =\n",
"            await google.colab.kernel.invokeFunction('convertToInteractive',\n",
```

```
          "                                        [key], {});\n",
          "        if (!dataTable) return;\n",
          "\n",
          "        const docLinkHtml = 'Like what you see? Visit the ' +\n",
          "          '<a target=\"_blank\"
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
          "          + ' to learn more about interactive tables.';\n",
          "        element.innerHTML = '';\n",
          "        dataTable['output_type'] = 'display_data';\n",
          "        await google.colab.output.renderOutput(dataTable, element);\n",
          "        const docLink = document.createElement('div');\n",
          "        docLink.innerHTML = docLinkHtml;\n",
          "        element.appendChild(docLink);\n",
          "       }\n",
          "     </script>\n",
          "    </div>\n",
          "  </div>\n",
          "  "
        ]
       },
       "metadata": {},
       "execution_count": 32
      }
     ]
    },
    {
     "cell_type": "code",
     "source": [
```

```
      "data['Age'] = np.where(data['Age']>62,data['Age'].mean(),data['Age']) #replacing the outliers using mean value"
     ],
     "metadata": {
      "id": "5mJNa6rtLJbP"
     },
     "execution_count": 34,
     "outputs": []
    },
    {
     "cell_type": "code",
     "source": [
      "data[data['Age']>62]"
     ],
     "metadata": {
      "colab": {
       "base_uri": "https://localhost:8080/",
       "height": 49
      },
      "id": "XQcSNYuyL7Pk",
      "outputId": "89651c63-7904-4782-da33-2c38ff96faf2"
     },
     "execution_count": 35,
     "outputs": [
      {
       "output_type": "execute_result",
       "data": {
        "text/plain": [
```

```
 "Empty DataFrame\n",
 "Columns: [RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary, Exited]\n",
 "Index: []"
],
"text/html": [
 "\n",
 "  <div id=\"df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
```

```
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"      <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd')\"\n",
"              title=\"Convert this dataframe to an interactive table.\"\n",
"              style=\"display:none;\">\n",
"        \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"       width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
```

```
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"    </button>\n",
"    \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
```

```
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67, 0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
"  </style>\n",
"\n",
"      <script>\n",
"        const buttonEl =\n",
"          document.querySelector('#df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd button.colab-df-convert');\n",
"        buttonEl.style.display =\n",
"          google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"        async function convertToInteractive(key) {\n",
"          const element = document.querySelector('#df-b69ea559-76e1-44bd-bb60-ca7e0c66c7bd');\n",
```

```
          "        const dataTable =\n",
          "          await google.colab.kernel.invokeFunction('convertToInteractive',\n",
          "                                                   [key], {});\n",
          "        if (!dataTable) return;\n",
          "\n",
          "        const docLinkHtml = 'Like what you see? Visit the ' +\n",
          "          '<a target=\"_blank\" 
href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
          "          + ' to learn more about interactive tables.';\n",
          "        element.innerHTML = '';\n",
          "        dataTable['output_type'] = 'display_data';\n",
          "        await google.colab.output.renderOutput(dataTable, element);\n",
          "        const docLink = document.createElement('div');\n",
          "        docLink.innerHTML = docLinkHtml;\n",
          "        element.appendChild(docLink);\n",
          "       }\n",
          "     </script>\n",
          "   </div>\n",
          "  </div>\n",
          "  "
        ]
       },
       "metadata": {},
       "execution_count": 35
      }
     ]
    },
    {
```

```
"cell_type": "code",
"source": [
  "data.head() #encoding"
],
"metadata": {
  "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 206
  },
  "id": "8FzUUUAnMbU2",
  "outputId": "fd237dce-7253-4523-e010-eb8662b16acf"
},
"execution_count": 36,
"outputs": [
 {
   "output_type": "execute_result",
   "data": {
     "text/plain": [
       "   RowNumber  CustomerId   Surname  CreditScore Geography  Gender   Age  \\\n",
       "0          1    15634602  Hargrave          619    France  Female  42.0   \n",
       "1          2    15647311      Hill          608     Spain  Female  41.0   \n",
       "2          3    15619304      Onio          502    France  Female  42.0   \n",
       "3          4    15701354      Boni          699    France  Female  39.0   \n",
       "4          5    15737888  Mitchell          850     Spain  Female  43.0   \n",
       "\n",
       "   Tenure    Balance  NumOfProducts  HasCrCard  IsActiveMember  \\\n",
       "0       2       0.00              1          1               1   \n",
```

```
 "1      1   83807.86           1        0           1   \n",
 "2      8  159660.80           3        1           0   \n",
 "3      1       0.00           2        0           0   \n",
 "4      2  125510.82           1        1           1   \n",
 "\n",
 "   EstimatedSalary  Exited  \n",
 "0       101348.88       1  \n",
 "1       112542.58       0  \n",
 "2       113931.57       1  \n",
 "3        93826.63       0  \n",
 "4        79084.10       0  "
],
"text/html": [
 "\n",
 "  <div id=\"df-7c4f6fec-89db-40d5-9055-1d0dac2fc948\">\n",
 "    <div class=\"colab-df-container\">\n",
 "      <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>Surname</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>IsActiveMember</th>\n",
"      <th>EstimatedSalary</th>\n",
"      <th>Exited</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0</th>\n",
"      <td>1</td>\n",
```

```
"     <td>15634602</td>\n",
"     <td>Hargrave</td>\n",
"     <td>619</td>\n",
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>42.0</td>\n",
"     <td>2</td>\n",
"     <td>0.00</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>101348.88</td>\n",
"     <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>1</th>\n",
"     <td>2</td>\n",
"     <td>15647311</td>\n",
"     <td>Hill</td>\n",
"     <td>608</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>41.0</td>\n",
"     <td>1</td>\n",
"     <td>83807.86</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
```

```
"    <td>1</td>\n",
"    <td>112542.58</td>\n",
"    <td>0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>2</th>\n",
"    <td>3</td>\n",
"    <td>15619304</td>\n",
"    <td>Onio</td>\n",
"    <td>502</td>\n",
"    <td>France</td>\n",
"    <td>Female</td>\n",
"    <td>42.0</td>\n",
"    <td>8</td>\n",
"    <td>159660.80</td>\n",
"    <td>3</td>\n",
"    <td>1</td>\n",
"    <td>0</td>\n",
"    <td>113931.57</td>\n",
"    <td>1</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>3</th>\n",
"    <td>4</td>\n",
"    <td>15701354</td>\n",
"    <td>Boni</td>\n",
"    <td>699</td>\n",
```

```
"     <td>France</td>\n",
"     <td>Female</td>\n",
"     <td>39.0</td>\n",
"     <td>1</td>\n",
"     <td>0.00</td>\n",
"     <td>2</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>93826.63</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>5</td>\n",
"     <td>15737888</td>\n",
"     <td>Mitchell</td>\n",
"     <td>850</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>43.0</td>\n",
"     <td>2</td>\n",
"     <td>125510.82</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>1</td>\n",
"     <td>79084.10</td>\n",
"     <td>0</td>\n",
```

```
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>\n",
"     <button class=\"colab-df-convert\" onclick=\"convertToInteractive('df-7c4f6fec-89db-40d5-9055-1d0dac2fc948')\"\n",
"             title=\"Convert this dataframe to an interactive table.\"\n",
"             style=\"display:none;\">\n",
"       \n",
"  <svg xmlns=\"http://www.w3.org/2000/svg\" height=\"24px\"viewBox=\"0 0 24 24\"\n",
"      width=\"24px\">\n",
"    <path d=\"M0 0h24v24H0V0z\" fill=\"none\"/>\n",
"    <path d=\"M18.56 5.44l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94zm-11 1L8.5 8.5l.94-2.06 2.06-.94-2.06-.94L8.5 2.5l-.94 2.06-2.06.94zm10 10l.94 2.06.94-2.06 2.06-.94-2.06-.94-.94-2.06-.94 2.06-2.06.94z\"/><path d=\"M17.41 7.96l-1.37-1.37c-.4-.4-.92-.59-1.43-.59-.52 0-1.04.2-1.43.59L10.3 9.45l-7.72 7.72c-.78.78-.78 2.05 0 2.83L4 21.41c.39.39.9.59 1.41.59.51 0 1.02-.2 1.41-.59l7.78-7.78 2.81-2.81c.8-.78.8-2.07 0-2.86zM5.41 20L4 18.59l7.72-7.72 1.47 1.35L5.41 20z\"/>\n",
"  </svg>\n",
"     </button>\n",
"     \n",
"  <style>\n",
"    .colab-df-container {\n",
"      display:flex;\n",
"      flex-wrap:wrap;\n",
"      gap: 12px;\n",
"    }\n",
"\n",
"    .colab-df-convert {\n",
"      background-color: #E8F0FE;\n",
```

```
"      border: none;\n",
"      border-radius: 50%;\n",
"      cursor: pointer;\n",
"      display: none;\n",
"      fill: #1967D2;\n",
"      height: 32px;\n",
"      padding: 0 0 0 0;\n",
"      width: 32px;\n",
"    }\n",
"\n",
"    .colab-df-convert:hover {\n",
"      background-color: #E2EBFA;\n",
"      box-shadow: 0px 1px 2px rgba(60, 64, 67, 0.3), 0px 1px 3px 1px rgba(60, 64, 67,
0.15);\n",
"      fill: #174EA6;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert {\n",
"      background-color: #3B4455;\n",
"      fill: #D2E3FC;\n",
"    }\n",
"\n",
"    [theme=dark] .colab-df-convert:hover {\n",
"      background-color: #434B5C;\n",
"      box-shadow: 0px 1px 3px 1px rgba(0, 0, 0, 0.15);\n",
"      filter: drop-shadow(0px 1px 2px rgba(0, 0, 0, 0.3));\n",
"      fill: #FFFFFF;\n",
"    }\n",
```

```
"  </style>\n",
"\n",
"    <script>\n",
"      const buttonEl =\n",
"        document.querySelector('#df-7c4f6fec-89db-40d5-9055-1d0dac2fc948 button.colab-df-convert');\n",
"      buttonEl.style.display =\n",
"        google.colab.kernel.accessAllowed ? 'block' : 'none';\n",
"\n",
"      async function convertToInteractive(key) {\n",
"        const element = document.querySelector('#df-7c4f6fec-89db-40d5-9055-1d0dac2fc948');\n",
"        const dataTable =\n",
"          await google.colab.kernel.invokeFunction('convertToInteractive',\n",
"                                                   [key], {});\n",
"        if (!dataTable) return;\n",
"\n",
"        const docLinkHtml = 'Like what you see? Visit the ' +\n",
"          '<a target=\"_blank\" href=https://colab.research.google.com/notebooks/data_table.ipynb>data table notebook</a>'\n",
"          + ' to learn more about interactive tables.';\n",
"        element.innerHTML = '';\n",
"        dataTable['output_type'] = 'display_data';\n",
"        await google.colab.output.renderOutput(dataTable, element);\n",
"        const docLink = document.createElement('div');\n",
"        docLink.innerHTML = docLinkHtml;\n",
"        element.appendChild(docLink);\n",
"      }\n",
"    </script>\n",
```

```
        "    </div>\n",
        "  </div>\n",
        "  "
       ]
      },
      "metadata": {},
      "execution_count": 36
    }
  ]
 },
 {
  "cell_type": "code",
  "source": [
   "pd.get_dummies(data,columns=['Surname']) #encoding method-1"
  ],
  "metadata": {
   "colab": {
    "base_uri": "https://localhost:8080/",
    "height": 488
   },
   "id": "i2CPqBA9MaGt",
   "outputId": "0badeada-4bb1-44df-dffe-c9340fea2d4d"
  },
  "execution_count": 37,
  "outputs": [
   {
    "output_type": "execute_result",
```

```
"data": {
 "text/plain": [
  "     RowNumber  CustomerId  CreditScore Geography  Gender   Age  Tenure  \\\n",
  "0          1    15634602        619    France  Female  42.0      2  \n",
  "1          2    15647311        608     Spain  Female  41.0      1  \n",
  "2          3    15619304        502    France  Female  42.0      8  \n",
  "3          4    15701354        699    France  Female  39.0      1  \n",
  "4          5    15737888        850     Spain  Female  43.0      2  \n",
  "...      ...         ...        ...       ...     ...   ...    ...  \n",
  "9995    9996    15606229        771    France    Male  39.0      5  \n",
  "9996    9997    15569892        516    France    Male  35.0     10  \n",
  "9997    9998    15584532        709    France  Female  36.0      7  \n",
  "9998    9999    15682355        772   Germany    Male  42.0      3  \n",
  "9999   10000    15628319        792    France  Female  28.0      4  \n",
  "\n",
  "       Balance  NumOfProducts  HasCrCard  ...  Surname_Zinachukwudi  \\\n",
  "0         0.00              1          1  ...                     0  \n",
  "1     83807.86              1          0  ...                     0  \n",
  "2    159660.80              3          1  ...                     0  \n",
  "3         0.00              2          0  ...                     0  \n",
  "4    125510.82              1          1  ...                     0  \n",
  "...        ...            ...        ...  ...                   ...  \n",
  "9995      0.00              2          1  ...                     0  \n",
  "9996  57369.61              1          1  ...                     0  \n",
  "9997      0.00              1          0  ...                     0  \n",
  "9998  75075.31              2          1  ...                     0  \n",
  "9999 130142.79              1          1  ...                     0  \n",
```

"\n",

"   Surname_Zito Surname_Zotov Surname_Zotova Surname_Zox \\\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",

"9995   0   0   0   0 \n",

"9996   0   0   0   0 \n",

"9997   0   0   0   0 \n",

"9998   0   0   0   0 \n",

"9999   0   0   0   0 \n",

"\n",

"   Surname_Zubarev Surname_Zubareva Surname_Zuev Surname_Zuyev \\\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",

"9995   0   0   0   0 \n",

"9996   0   0   0   0 \n",

"9997   0   0   0   0 \n",

"9998   0   0   0   0 \n",

"9999   0   0   0   0 \n",

"\n",

```
  "     Surname_Zuyeva  \n",
  "0               0  \n",
  "1               0  \n",
  "2               0  \n",
  "3               0  \n",
  "4               0  \n",
  "...           ...  \n",
  "9995            0  \n",
  "9996            0  \n",
  "9997            0  \n",
  "9998            0  \n",
  "9999            0  \n",
  "\n",
  "[10000 rows x 2945 columns]"
 ],
 "text/html": [
  "\n",
  "  <div id=\"df-a4595f0d-c407-4432-8dea-3affcf313c96\">\n",
  "    <div class=\"colab-df-container\">\n",
  "     <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>RowNumber</th>\n",
"      <th>CustomerId</th>\n",
"      <th>CreditScore</th>\n",
"      <th>Geography</th>\n",
"      <th>Gender</th>\n",
"      <th>Age</th>\n",
"      <th>Tenure</th>\n",
"      <th>Balance</th>\n",
"      <th>NumOfProducts</th>\n",
"      <th>HasCrCard</th>\n",
"      <th>...</th>\n",
"      <th>Surname_Zinachukwudi</th>\n",
"      <th>Surname_Zito</th>\n",
"      <th>Surname_Zotov</th>\n",
"      <th>Surname_Zotova</th>\n",
"      <th>Surname_Zox</th>\n",
"      <th>Surname_Zubarev</th>\n",
```

```
"    <th>Surname_Zubareva</th>\n",
"    <th>Surname_Zuev</th>\n",
"    <th>Surname_Zuyev</th>\n",
"    <th>Surname_Zuyeva</th>\n",
"   </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"   <tr>\n",
"    <th>0</th>\n",
"    <td>1</td>\n",
"    <td>15634602</td>\n",
"    <td>619</td>\n",
"    <td>France</td>\n",
"    <td>Female</td>\n",
"    <td>42.0</td>\n",
"    <td>2</td>\n",
"    <td>0.00</td>\n",
"    <td>1</td>\n",
"    <td>1</td>\n",
"    <td>...</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>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>1</th>\n",
"     <td>2</td>\n",
"     <td>15647311</td>\n",
"     <td>608</td>\n",
"     <td>Spain</td>\n",
"     <td>Female</td>\n",
"     <td>41.0</td>\n",
"     <td>1</td>\n",
"     <td>83807.86</td>\n",
"     <td>1</td>\n",
"     <td>0</td>\n",
"     <td>...</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"     <td>0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>3</th>\n",
"     <td>4</td>\n",
"     <td>15701354</td>\n",
"     <td>699</td>\n",
```

```
"     <td>France</td>\n",

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

"     <td>39.0</td>\n",

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

"     <td>0.00</td>\n",

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

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

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

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

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

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

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

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

...v
```