;;;;;;;;;;;;;;;;;;;;{

"cells": [

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "rm4n-J12yRxN"

},

"outputs": [],

"source": [

"import pandas as pd\n",

"import number as np\n",

"import sea prediction as sms\n",

"from matplotlib import py plot as plt\n",

"import warnings\n",

"warnings.filterwarnings('ignore')"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "loV7U9AI9Smw"

},

"source": [

"##Load model"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "IR0edMsi0hNd"

},

"outputs": [],

"source": [

"data=pd.read\_csv(\"Churn\_Modelling.csv\")"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 363

},

"id": "8FJAYsyU0uF6",

"outputId": "07467116-8323-474e-fec6-fca9da2ded6e"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

" <th>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",

" <tr>\n",

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

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

" <td>15574012</td>\n",

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

" <td>645</td>\n",

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

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

" <td>44</td>\n",

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

" <td>113755.78</td>\n",

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

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

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

" <td>149756.71</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>7</td>\n",

" <td>15592531</td>\n",

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

" <td>822</td>\n",

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

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

" <td>50</td>\n",

" <td>7</td>\n",

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

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

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

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

" <td>10062.80</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>7</th>\n",

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

" <td>15656148</td>\n",

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

" <td>376</td>\n",

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

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

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

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

" <td>115046.74</td>\n",

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

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

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

" <td>119346.88</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15792365</td>\n",

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

" <td>501</td>\n",

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

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

" <td>44</td>\n",

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

" <td>142051.07</td>\n",

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

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

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

" <td>74940.50</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15592389</td>\n",

" <td>H?</td>\n",

" <td>684</td>\n",

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

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

" <td>27</td>\n",

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

" <td>134603.88</td>\n",

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

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

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

" <td>71725.73</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"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",

"5 6 15574012 Chu 645 Spain Male 44 \n",

"6 7 15592531 Bartlett 822 France Male 50 \n",

"7 8 15656148 Obinna 376 Germany Female 29 \n",

"8 9 15792365 He 501 France Male 44 \n",

"9 10 15592389 H? 684 France Male 27 \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",

"5 8 113755.78 2 1 0 \n",

"6 7 0.00 2 1 1 \n",

"7 4 115046.74 4 1 0 \n",

"8 4 142051.07 2 0 1 \n",

"9 2 134603.88 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 \n",

"5 149756.71 1 \n",

"6 10062.80 0 \n",

"7 119346.88 1 \n",

"8 74940.50 0 \n",

"9 71725.73 0 "

]

},

"execution\_count": 6,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.head(10)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "NHVadOGn8LAo",

"outputId": "85c9450d-0311-4307-e336-0cda5efa65c7"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

" <th>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>9990</th>\n",

" <td>9991</td>\n",

" <td>15798964</td>\n",

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

" <td>714</td>\n",

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

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

" <td>33</td>\n",

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

" <td>35016.60</td>\n",

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

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

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

" <td>53667.08</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9991</th>\n",

" <td>9992</td>\n",

" <td>15769959</td>\n",

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

" <td>597</td>\n",

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

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

" <td>53</td>\n",

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

" <td>88381.21</td>\n",

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

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

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

" <td>69384.71</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9992</th>\n",

" <td>9993</td>\n",

" <td>15657105</td>\n",

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

" <td>726</td>\n",

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

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

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

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

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

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

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

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

" <td>195192.40</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9993</th>\n",

" <td>9994</td>\n",

" <td>15569266</td>\n",

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

" <td>644</td>\n",

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

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

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

" <td>7</td>\n",

" <td>155060.41</td>\n",

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

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

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

" <td>29179.52</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9994</th>\n",

" <td>9995</td>\n",

" <td>15719294</td>\n",

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

" <td>800</td>\n",

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

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

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

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

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

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

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

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

" <td>167773.55</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9995</th>\n",

" <td>9996</td>\n",

" <td>15606229</td>\n",

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

" <td>771</td>\n",

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

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

" <td>39</td>\n",

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

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

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

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

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

" <td>96270.64</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9996</th>\n",

" <td>9997</td>\n",

" <td>15569892</td>\n",

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

" <td>516</td>\n",

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

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

" <td>35</td>\n",

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

" <td>57369.61</td>\n",

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

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

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

" <td>101699.77</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9997</th>\n",

" <td>9998</td>\n",

" <td>15584532</td>\n",

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

" <td>709</td>\n",

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

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

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

" <td>7</td>\n",

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

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

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

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

" <td>42085.58</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9998</th>\n",

" <td>9999</td>\n",

" <td>15682355</td>\n",

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

" <td>772</td>\n",

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

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

" <td>42</td>\n",

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

" <td>75075.31</td>\n",

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

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

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

" <td>92888.52</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9999</th>\n",

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

" <td>15628319</td>\n",

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

" <td>792</td>\n",

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

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

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

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

" <td>130142.79</td>\n",

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

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

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

" <td>38190.78</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId Surname CreditScore Geography Gender Age \\\n",

"9990 9991 15798964 Nkemakonam 714

japan Male 33 \n",

"9991 9992 15769959 Ajuluchukwu 597 France Female 53 \n",

"9992 9993 15657105 Chukwualuka 726 Spain Male 36 \n",

"9993 9994 15569266 Rahman 644 France Male 28 \n",

"9994 9995 15719294 Wood 800 France Female 29 \n",

"9995 9996 15606229 Obijiaku 771 italy Male 39 \n",

"9996 9997 15569892 Johnstone 516 France Male 35 \n",

"9997 9998 15584532 Liu 709 France Female 36 \n",

"9998 9999 15682355 Sabbatini 772 Germany Male 42 \n",

"9999 10000 15628319 Walker 792 France Female 28 \n",

"\n",

" Tenure Balance NumOfProducts HasCrCard IsActiveMember \\\n",

"9990 3 35016.60 1 1 0 \n",

"9991 4 88381.21 1 1 0 \n",

"9992 2 0.00 1 1 0 \n",

"9993 7 155060.41 1 1 0 \n",

"9994 2 0.00 2 0 0 \n",

"9995 5 0.00 2 1 0 \n",

"9996 10 57369.61 1 1 1 \n",

"9997 7 0.00 1 0 1 \n",

"9998 3 75075.31 2 1 0 \n",

"9999 4 130142.79 1 1 0 \n",

"\n",

" EstimatedSalary Exited \n",

"9990 53667.08 0 \n",

"9991 69384.71 1 \n",

"9992 195192.40 0 \n",

"9993 29179.52 0 \n",

"9994 167773.55 0 \n",

"9995 96270.64 0 \n",

"9996 101699.77 0 \n",

"9997 42085.58 1 \n",

"9998 92888.52 1 \n",

"9999 38190.78 0 "

]

},

"execution\_count": 7,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.tail(10)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 344

},

"id": "jAMhixkp02\_T",

"outputId": "4996ca1f-c1c0-4a57-a506-e0bbfff89c22"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

" <td>10000.00000</td>\n",

" <td>1.000000e+04</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.00000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" <td>10000.000000</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>5000.50000</td>\n",

" <td>1.569094e+07</td>\n",

" <td>650.528800</td>\n",

" <td>38.921800</td>\n",

" <td>5.012800</td>\n",

" <td>76485.889288</td>\n",

" <td>1.530200</td>\n",

" <td>0.70550</td>\n",

" <td>0.515100</td>\n",

" <td>100090.239881</td>\n",

" <td>0.203700</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>2886.89568</td>\n",

" <td>7.193619e+04</td>\n",

" <td>96.653299</td>\n",

" <td>10.487806</td>\n",

" <td>2.892174</td>\n",

" <td>62397.405202</td>\n",

" <td>0.581654</td>\n",

" <td>0.45584</td>\n",

" <td>0.499797</td>\n",

" <td>57510.492818</td>\n",

" <td>0.402769</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>1.00000</td>\n",

" <td>1.556570e+07</td>\n",

" <td>350.000000</td>\n",

" <td>18.000000</td>\n",

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

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

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

" <td>0.00000</td>\n",

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

" <td>11.580000</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>2500.75000</td>\n",

" <td>1.562853e+07</td>\n",

" <td>584.000000</td>\n",

" <td>32.000000</td>\n",

" <td>3.000000</td>\n",

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

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

" <td>0.00000</td>\n",

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

" <td>51002.110000</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>5000.50000</td>\n",

" <td>1.569074e+07</td>\n",

" <td>652.000000</td>\n",

" <td>37.000000</td>\n",

" <td>5.000000</td>\n",

" <td>97198.540000</td>\n",

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

" <td>1.00000</td>\n",

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

" <td>100193.915000</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>7500.25000</td>\n",

" <td>1.575323e+07</td>\n",

" <td>718.000000</td>\n",

" <td>44.000000</td>\n",

" <td>7.000000</td>\n",

" <td>127644.240000</td>\n",

" <td>2.000000</td>\n",

" <td>1.00000</td>\n",

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

" <td>149388.247500</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>10000.00000</td>\n",

" <td>1.581569e+07</td>\n",

" <td>850.000000</td>\n",

" <td>92.000000</td>\n",

" <td>10.000000</td>\n",

" <td>250898.090000</td>\n",

" <td>4.000000</td>\n",

" <td>1.00000</td>\n",

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

" <td>199992.480000</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId CreditScore Age Tenure \\\n",

"count 10000.00000 1.000000e+04 10000.000000 10000.000000 10000.000000 \n",

"mean 5000.50000 1.569094e+07 650.528800 38.921800 5.012800 \n",

"std 2886.89568 7.193619e+04 96.653299 10.487806 2.892174 \n",

"min 1.00000 1.556570e+07 350.000000 18.000000 0.000000 \n",

"25% 2500.75000 1.562853e+07 584.000000 32.000000 3.000000 \n",

"50% 5000.50000 1.569074e+07 652.000000 37.000000 5.000000 \n",

"75% 7500.25000 1.575323e+07 718.000000 44.000000 7.000000 \n",

"max 10000.00000 1.581569e+07 850.000000 92.000000 10.000000 \n",

"\n",

" Balance NumOfProducts HasCrCard IsActiveMember \\\n",

"count 10000.000000 10000.000000 10000.00000 10000.000000 \n",

"mean 76485.889288 1.530200 0.70550 0.515100 \n",

"std 62397.405202 0.581654 0.45584 0.499797 \n",

"min 0.000000 1.000000 0.00000 0.000000 \n",

"25% 0.000000 1.000000 0.00000 0.000000 \n",

"50% 97198.540000 1.000000 1.00000 1.000000 \n",

"75% 127644.240000 2.000000 1.00000 1.000000 \n",

"max 250898.090000 4.000000 1.00000 1.000000 \n",

"\n",

" EstimatedSalary Exited \n",

"count 10000.000000 10000.000000 \n",

"mean 100090.239881 0.203700 \n",

"std 57510.492818 0.402769 \n",

"min 11.580000 0.000000 \n",

"25% 51002.110000 0.000000 \n",

"50% 100193.915000 0.000000 \n",

"75% 149388.247500 0.000000 \n",

"max 199992.480000 1.000000 "

]

},

"execution\_count": 8,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"#describe statistics\n",

"data.describe()"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "lMRduK\_hAG1J"

},

"source": [

"#Univariate Analysis "

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "JnBk7-BqOFZK",

"outputId": "ad1855f1-b680-412b-b8e5-3a04ddab1545"

},

"outputs": [

{

"data": {

"text/plain": [

"RowNumber -1.200000\n",

"CustomerId -1.196113\n",

"CreditScore -0.425726\n",

"Age 1.395347\n",

"Tenure -1.165225\n",

"Balance -1.489412\n",

"NumOfProducts 0.582981\n",

"HasCrCard -1.186973\n",

"IsActiveMember -1.996747\n",

"EstimatedSalary -1.181518\n",

"Exited 0.165671\n",

"dtype: float64"

]

},

"execution\_count": 9,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.kurt(axis=0,skipna=True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "\_4hL\_my0OWq9",

"outputId": "3e5a8fe4-b416-4f91-a71e-f5977f3cd689"

},

"outputs": [

{

"data": {

"text/plain": [

"0 10.998778\n",

"1 10.997909\n",

"2 10.995886\n",

"3 10.998962\n",

"4 10.997675\n",

" ... \n",

"9995 10.998908\n",

"9996 10.998551\n",

"9997 10.999788\n",

"9998 10.998530\n",

"9999 10.997973\n",

"Length: 10000, dtype: float64"

]

},

"execution\_count": 10,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.kurt(axis=1,skipna=True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "0enR86F41AdN",

"outputId": "9032f29d-7473-4536-bdcb-741c0db2c333"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age', ylabel='Density'>"

]

},

"execution\_count": 11,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.distplot(data['Age'])"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "7qa\_qjtqAilY"

},

"source": [

"#Bi - Variate Analysis"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 300

},

"id": "AZY0yxpQ1s56",

"outputId": "5b18d04d-afb4-4fd9-8cd2-8c2de30e50d1"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age', ylabel='count'>"

]

},

"execution\_count": 12,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.countplot(data[\"Age\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "tITnp0r52a0W",

"outputId": "74687936-ca41-4868-9a04-1d130651e103"

},

"outputs": [

{

"data": {

"text/plain": [

"RowNumber 0.000000\n",

"CustomerId 0.001149\n",

"CreditScore -0.071607\n",

"Age 1.011320\n",

"Tenure 0.010991\n",

"Balance -0.141109\n",

"NumOfProducts 0.745568\n",

"HasCrCard -0.901812\n",

"IsActiveMember -0.060437\n",

"EstimatedSalary 0.002085\n",

"Exited 1.471611\n",

"dtype: float64"

]

},

"execution\_count": 13,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.skew(axis=0,skipna=True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "ywhDer-92rpQ",

"outputId": "b36da3c5-a639-423c-ccea-b35d50fc3e60"

},

"outputs": [

{

"data": {

"text/plain": [

"0 3.316373\n",

"1 3.316193\n",

"2 3.315777\n",

"3 3.316411\n",

"4 3.316145\n",

" ... \n",

"9995 3.316399\n",

"9996 3.316325\n",

"9997 3.316581\n",

"9998 3.316321\n",

"9999 3.316207\n",

"Length: 10000, dtype: float64"

]

},

"execution\_count": 14,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.skew(axis=1,skipna=True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "BLHj9QQT47is",

"outputId": "505d17a4-3b5d-4edd-dca7-e15b8e58b865"

},

"outputs": [

{

"data": {

"text/plain": [

"RowNumber False\n",

"CustomerId False\n",

"Surname False\n",

"CreditScore False\n",

"Geography False\n",

"Gender False\n",

"Age False\n",

"Tenure False\n",

"Balance False\n",

"NumOfProducts False\n",

"HasCrCard False\n",

"IsActiveMember False\n",

"EstimatedSalary False\n",

"Exited False\n",

"dtype: bool"

]

},

"execution\_count": 15,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.isnull().any()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "dxQt8hSf5IOi",

"outputId": "808ed8e6-ec73-4b58-fa28-559cc84e4fb8"

},

"outputs": [

{

"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"

]

},

"execution\_count": 16,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.isnull().sum()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "T-Ohc2ZJ5RiX",

"outputId": "32518f68-689f-439e-cf97-d2ca1e6bfade"

},

"outputs": [

{

"data": {

"text/plain": [

"0 False\n",

"1 False\n",

"2 False\n",

"3 False\n",

"4 False\n",

" ... \n",

"9995 False\n",

"9996 False\n",

"9997 False\n",

"9998 False\n",

"9999 False\n",

"Length: 10000, dtype: bool"

]

},

"execution\_count": 17,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.duplicated()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "5B\_MpgIl5WWL",

"outputId": "7246c92d-f695-4f3f-c781-e1dee5fdfdbe"

},

"outputs": [

{

"data": {

"text/plain": [

"0"

]

},

"execution\_count": 18,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.duplicated().sum()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "ApVJa6pf6sQA"

},

"outputs": [],

"source": [

"###VISUALISATION\n"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 282

},

"id": "2KKa9tDo62sL",

"outputId": "b46892ca-ef56-433c-af8d-fdf100e0c81c"

},

"outputs": [

{

"data": {

"text/plain": [

"<matplotlib.collections.PathCollection at 0x2226e809b20>"

]

},

"execution\_count": 20,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"plt.scatter(data.Age,data.Balance)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 286

},

"id": "4b5yco727TZg",

"outputId": "2de27243-d5f1-493c-a417-6708e5a840ea"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age', ylabel='Balance'>"

]

},

"execution\_count": 21,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.scatterplot(data.Age,data.Balance)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "C1NMpNKmOoPL",

"outputId": "7c445afe-1444-411a-b447-099b0f3b5518"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age'>"

]

},

"execution\_count": 22,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.barplot(data['Age'])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "Pri1TiSzO4Vu",

"outputId": "792b5d22-c759-4315-b788-6b16d05c1f3d"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age'>"

]

},

"execution\_count": 23,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.boxplot(data['Age'])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 394

},

"id": "1iYYmQF7PBRg",

"outputId": "1a187304-c0a8-4005-81c8-65f456972b44"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

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

" <td>0.004202</td>\n",

" <td>0.005840</td>\n",

" <td>0.000783</td>\n",

" <td>-0.006495</td>\n",

" <td>-0.009067</td>\n",

" <td>0.007246</td>\n",

" <td>0.000599</td>\n",

" <td>0.012044</td>\n",

" <td>-0.005988</td>\n",

" <td>-0.016571</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.004202</td>\n",

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

" <td>0.005308</td>\n",

" <td>0.009497</td>\n",

" <td>-0.014883</td>\n",

" <td>-0.012419</td>\n",

" <td>0.016972</td>\n",

" <td>-0.014025</td>\n",

" <td>0.001665</td>\n",

" <td>0.015271</td>\n",

" <td>-0.006248</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.005840</td>\n",

" <td>0.005308</td>\n",

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

" <td>-0.003965</td>\n",

" <td>0.000842</td>\n",

" <td>0.006268</td>\n",

" <td>0.012238</td>\n",

" <td>-0.005458</td>\n",

" <td>0.025651</td>\n",

" <td>-0.001384</td>\n",

" <td>-0.027094</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.000783</td>\n",

" <td>0.009497</td>\n",

" <td>-0.003965</td>\n",

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

" <td>-0.009997</td>\n",

" <td>0.028308</td>\n",

" <td>-0.030680</td>\n",

" <td>-0.011721</td>\n",

" <td>0.085472</td>\n",

" <td>-0.007201</td>\n",

" <td>0.285323</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>-0.006495</td>\n",

" <td>-0.014883</td>\n",

" <td>0.000842</td>\n",

" <td>-0.009997</td>\n",

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

" <td>-0.012254</td>\n",

" <td>0.013444</td>\n",

" <td>0.022583</td>\n",

" <td>-0.028362</td>\n",

" <td>0.007784</td>\n",

" <td>-0.014001</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>-0.009067</td>\n",

" <td>-0.012419</td>\n",

" <td>0.006268</td>\n",

" <td>0.028308</td>\n",

" <td>-0.012254</td>\n",

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

" <td>-0.304180</td>\n",

" <td>-0.014858</td>\n",

" <td>-0.010084</td>\n",

" <td>0.012797</td>\n",

" <td>0.118533</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.007246</td>\n",

" <td>0.016972</td>\n",

" <td>0.012238</td>\n",

" <td>-0.030680</td>\n",

" <td>0.013444</td>\n",

" <td>-0.304180</td>\n",

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

" <td>0.003183</td>\n",

" <td>0.009612</td>\n",

" <td>0.014204</td>\n",

" <td>-0.047820</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.000599</td>\n",

" <td>-0.014025</td>\n",

" <td>-0.005458</td>\n",

" <td>-0.011721</td>\n",

" <td>0.022583</td>\n",

" <td>-0.014858</td>\n",

" <td>0.003183</td>\n",

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

" <td>-0.011866</td>\n",

" <td>-0.009933</td>\n",

" <td>-0.007138</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>0.012044</td>\n",

" <td>0.001665</td>\n",

" <td>0.025651</td>\n",

" <td>0.085472</td>\n",

" <td>-0.028362</td>\n",

" <td>-0.010084</td>\n",

" <td>0.009612</td>\n",

" <td>-0.011866</td>\n",

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

" <td>-0.011421</td>\n",

" <td>-0.156128</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>-0.005988</td>\n",

" <td>0.015271</td>\n",

" <td>-0.001384</td>\n",

" <td>-0.007201</td>\n",

" <td>0.007784</td>\n",

" <td>0.012797</td>\n",

" <td>0.014204</td>\n",

" <td>-0.009933</td>\n",

" <td>-0.011421</td>\n",

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

" <td>0.012097</td>\n",

" </tr>\n",

" <tr>\n",

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

" <td>-0.016571</td>\n",

" <td>-0.006248</td>\n",

" <td>-0.027094</td>\n",

" <td>0.285323</td>\n",

" <td>-0.014001</td>\n",

" <td>0.118533</td>\n",

" <td>-0.047820</td>\n",

" <td>-0.007138</td>\n",

" <td>-0.156128</td>\n",

" <td>0.012097</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId CreditScore Age Tenure \\\n",

"RowNumber 1.000000 0.004202 0.005840 0.000783 -0.006495 \n",

"CustomerId 0.004202 1.000000 0.005308 0.009497 -0.014883 \n",

"CreditScore 0.005840 0.005308 1.000000 -0.003965 0.000842 \n",

"Age 0.000783 0.009497 -0.003965 1.000000 -0.009997 \n",

"Tenure -0.006495 -0.014883 0.000842 -0.009997 1.000000 \n",

"Balance -0.009067 -0.012419 0.006268 0.028308 -0.012254 \n",

"NumOfProducts 0.007246 0.016972 0.012238 -0.030680 0.013444 \n",

"HasCrCard 0.000599 -0.014025 -0.005458 -0.011721 0.022583 \n",

"IsActiveMember 0.012044 0.001665 0.025651 0.085472 -0.028362 \n",

"EstimatedSalary -0.005988 0.015271 -0.001384 -0.007201 0.007784 \n",

"Exited -0.016571 -0.006248 -0.027094 0.285323 -0.014001 \n",

"\n",

" Balance NumOfProducts HasCrCard IsActiveMember \\\n",

"RowNumber -0.009067 0.007246 0.000599 0.012044 \n",

"CustomerId -0.012419 0.016972 -0.014025 0.001665 \n",

"CreditScore 0.006268 0.012238 -0.005458 0.025651 \n",

"Age 0.028308 -0.030680 -0.011721 0.085472 \n",

"Tenure -0.012254 0.013444 0.022583 -0.028362 \n",

"Balance 1.000000 -0.304180 -0.014858 -0.010084 \n",

"NumOfProducts -0.304180 1.000000 0.003183 0.009612 \n",

"HasCrCard -0.014858 0.003183 1.000000 -0.011866 \n",

"IsActiveMember -0.010084 0.009612 -0.011866 1.000000 \n",

"EstimatedSalary 0.012797 0.014204 -0.009933 -0.011421 \n",

"Exited 0.118533 -0.047820 -0.007138 -0.156128 \n",

"\n",

" EstimatedSalary Exited \n",

"RowNumber -0.005988 -0.016571 \n",

"CustomerId 0.015271 -0.006248 \n",

"CreditScore -0.001384 -0.027094 \n",

"Age -0.007201 0.285323 \n",

"Tenure 0.007784 -0.014001 \n",

"Balance 0.012797 0.118533 \n",

"NumOfProducts 0.014204 -0.047820 \n",

"HasCrCard -0.009933 -0.007138 \n",

"IsActiveMember -0.011421 -0.156128 \n",

"EstimatedSalary 1.000000 0.012097 \n",

"Exited 0.012097 1.000000 "

]

},

"execution\_count": 24,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.corr()"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "pebXg1h5Azd5"

},

"source": [

"#Multivariate Analysis"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 358

},

"id": "7HabKLduPIUI",

"outputId": "e77d604d-461c-4aaf-eb8f-113085bfc755"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:>"

]

},

"execution\_count": 25,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 2 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.heatmap(data.corr(),annot=True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 1000

},

"id": "WTXodwZwPeZ\_",

"outputId": "4486b459-00e0-4e05-8cea-7be981b586bf"

},

"outputs": [

{

"data": {

"text/plain": [

"<seaborn.axisgrid.PairGrid at 0x222285c5580>"

]

},

"execution\_count": 36,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 1980x1980 with 132 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.pairplot(data)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "zHC6y\_ONTGlN"

},

"outputs": [],

"source": [

"from scipy.stats import spearmanr"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "gZUtVGkmRHhY",

"outputId": "7adc5404-6db9-4ea7-9322-3fcbf6f44ec7"

},

"outputs": [

{

"data": {

"text/plain": [

"SpearmanrResult(correlation=array([[ 1.00000000e+00, 4.18684789e-03, 1.82537815e-03,\n",

" 5.13017187e-03, -1.01176571e-02, 1.81963613e-02,\n",

" 4.76064421e-04, -6.93433206e-03, -9.01325568e-03,\n",

" 8.30510741e-03, 5.98746525e-04, 1.20443901e-02,\n",

" -6.00682958e-03, -1.65713715e-02],\n",

" [ 4.18684789e-03, 1.00000000e+00, 5.31564210e-03,\n",

" 5.96746465e-03, 6.03529435e-03, -2.62440728e-03,\n",

" 8.77466555e-03, -1.50720283e-02, -1.39321914e-02,\n",

" 1.92970188e-02, -1.40233299e-02, 1.68193033e-03,\n",

" 1.52457829e-02, -6.26374782e-03],\n",

" [ 1.82537815e-03, 5.31564210e-03, 1.00000000e+00,\n",

" 6.68503170e-03, -2.26792517e-02, -2.14337922e-03,\n",

" 1.37678535e-03, -1.70916721e-02, -8.00358124e-04,\n",

" -1.72831393e-02, -8.93818901e-03, 1.37684719e-03,\n",

" 1.17949476e-02, -1.09832944e-02],\n",

" [ 5.13017187e-03, 5.96746465e-03, 6.68503170e-03,\n",

" 1.00000000e+00, 6.10527978e-03, -3.01144279e-03,\n",

" -7.97404431e-03, 1.13317419e-03, 5.68657057e-03,\n",

" 1.25677271e-02, -3.80181966e-03, 2.42623407e-02,\n",

" 1.23652438e-03, -2.32893966e-02],\n",

" [-1.01176571e-02, 6.03529435e-03, -2.26792517e-02,\n",

" 6.10527978e-03, 1.00000000e+00, 2.05197803e-03,\n",

" 3.53513965e-02, 3.76366156e-03, 9.94871724e-02,\n",

" 7.69108918e-04, -7.22407343e-03, 4.44007080e-03,\n",

" -1.94818567e-04, 5.30920641e-02],\n",

" [ 1.81963613e-02, -2.62440728e-03, -2.14337922e-03,\n",

" -3.01144279e-03, 2.05197803e-03, 1.00000000e+00,\n",

" -2.97848194e-02, 1.50959348e-02, 1.35043861e-02,\n",

" -1.28505367e-02, 5.76612437e-03, 2.25443247e-02,\n",

" -8.26853704e-03, -1.06512488e-01],\n",

" [ 4.76064421e-04, 8.77466555e-03, 1.37678535e-03,\n",

" -7.97404431e-03, 3.53513965e-02, -2.97848194e-02,\n",

" 1.00000000e+00, -1.04049493e-02, 3.33043436e-02,\n",

" -5.85664619e-02, -1.52782371e-02, 3.98391734e-02,\n",

" -2.43149876e-03, 3.23967912e-01],\n",

" [-6.93433206e-03, -1.50720283e-02, -1.70916721e-02,\n",

" 1.13317419e-03, 3.76366156e-03, 1.50959348e-02,\n",

" -1.04049493e-02, 1.00000000e+00, -9.51289512e-03,\n",

" 1.29080538e-02, 2.23540939e-02, -2.86732861e-02,\n",

" 7.77808376e-03, -1.39780555e-02],\n",

" [-9.01325568e-03, -1.39321914e-02, -8.00358124e-04,\n",

" 5.68657057e-03, 9.94871724e-02, 1.35043861e-02,\n",

" 3.33043436e-02, -9.51289512e-03, 1.00000000e+00,\n",

" -3.16626558e-01, -9.83460270e-03, -1.14965258e-02,\n",

" 1.17780035e-02, 1.11110193e-01],\n",

" [ 8.30510741e-03, 1.92970188e-02, -1.72831393e-02,\n",

" 1.25677271e-02, 7.69108918e-04, -1.28505367e-02,\n",

" -5.85664619e-02, 1.29080538e-02, -3.16626558e-01,\n",

" 1.00000000e+00, 3.85886031e-03, 1.62917706e-02,\n",

" 1.25698129e-02, -1.25282063e-01],\n",

" [ 5.98746525e-04, -1.40233299e-02, -8.93818901e-03,\n",

" -3.80181966e-03, -7.22407343e-03, 5.76612437e-03,\n",

" -1.52782371e-02, 2.23540939e-02, -9.83460270e-03,\n",

" 3.85886031e-03, 1.00000000e+00, -1.18656369e-02,\n",

" -1.00409074e-02, -7.13776560e-03],\n",

" [ 1.20443901e-02, 1.68193033e-03, 1.37684719e-03,\n",

" 2.42623407e-02, 4.44007080e-03, 2.25443247e-02,\n",

" 3.98391734e-02, -2.86732861e-02, -1.14965258e-02,\n",

" 1.62917706e-02, -1.18656369e-02, 1.00000000e+00,\n",

" -1.14690521e-02, -1.56128278e-01],\n",

" [-6.00682958e-03, 1.52457829e-02, 1.17949476e-02,\n",

" 1.23652438e-03, -1.94818567e-04, -8.26853704e-03,\n",

" -2.43149876e-03, 7.77808376e-03, 1.17780035e-02,\n",

" 1.25698129e-02, -1.00409074e-02, -1.14690521e-02,\n",

" 1.00000000e+00, 1.20805366e-02],\n",

" [-1.65713715e-02, -6.26374782e-03, -1.09832944e-02,\n",

" -2.32893966e-02, 5.30920641e-02, -1.06512488e-01,\n",

" 3.23967912e-01, -1.39780555e-02, 1.11110193e-01,\n",

" -1.25282063e-01, -7.13776560e-03, -1.56128278e-01,\n",

" 1.20805366e-02, 1.00000000e+00]]), pvalue=array([[0.00000000e+000, 6.75483429e-001, 8.55178468e-001,\n",

" 6.07981798e-001, 3.11698199e-001, 6.88261457e-002,\n",

" 9.62034639e-001, 4.88086885e-001, 3.67465405e-001,\n",

" 4.06300660e-001, 9.52261425e-001, 2.28461236e-001,\n",

" 5.48097586e-001, 9.75106276e-002],\n",

" [6.75483429e-001, 0.00000000e+000, 5.95071292e-001,\n",

" 5.50722932e-001, 5.46203060e-001, 7.93006618e-001,\n",

" 3.80283664e-001, 1.31785022e-001, 1.63585747e-001,\n",

" 5.36514208e-002, 1.60847582e-001, 8.66447868e-001,\n",

" 1.27389774e-001, 5.31116466e-001],\n",

" [8.55178468e-001, 5.95071292e-001, 0.00000000e+000,\n",

" 5.03861020e-001, 2.33332702e-002, 8.30304249e-001,\n",

" 8.90508036e-001, 8.74364323e-002, 9.36216720e-001,\n",

" 8.39475437e-002, 3.71469037e-001, 8.90503148e-001,\n",

" 2.38243578e-001, 2.72106138e-001],\n",

" [6.07981798e-001, 5.50722932e-001, 5.03861020e-001,\n",

" 0.00000000e+000, 5.41558890e-001, 7.63332662e-001,\n",

" 4.25266703e-001, 9.09790109e-001, 5.69634028e-001,\n",

" 2.08874884e-001, 7.03844602e-001, 1.52541637e-002,\n",

" 9.01602674e-001, 1.98609526e-002],\n",

" [3.11698199e-001, 5.46203060e-001, 2.33332702e-002,\n",

" 5.41558890e-001, 0.00000000e+000, 8.37437458e-001,\n",

" 4.06537979e-004, 7.06678685e-001, 2.01668047e-023,\n",

" 9.38702072e-001, 4.70093788e-001, 6.57076013e-001,\n",

" 9.84458653e-001, 1.08256524e-007],\n",

" [6.88261457e-002, 7.93006618e-001, 8.30304249e-001,\n",

" 7.63332662e-001, 8.37437458e-001, 0.00000000e+000,\n",

" 2.89407525e-003, 1.31173411e-001, 1.76909716e-001,\n",

" 1.98811127e-001, 5.64246762e-001, 2.41686809e-002,\n",

" 4.08370570e-001, 1.25850456e-026],\n",

" [9.62034639e-001, 3.80283664e-001, 8.90508036e-001,\n",

" 4.25266703e-001, 4.06537979e-004, 2.89407525e-003,\n",

" 0.00000000e+000, 2.98157345e-001, 8.65526378e-004,\n",

" 4.60240532e-009, 1.26581605e-001, 6.74797620e-005,\n",

" 8.07912562e-001, 4.60367975e-243],\n",

" [4.88086885e-001, 1.31785022e-001, 8.74364323e-002,\n",

" 9.09790109e-001, 7.06678685e-001, 1.31173411e-001,\n",

" 2.98157345e-001, 0.00000000e+000, 3.41506861e-001,\n",

" 1.96808492e-001, 2.53904935e-002, 4.13650739e-003,\n",

" 4.36732384e-001, 1.62203448e-001],\n",

" [3.67465405e-001, 1.63585747e-001, 9.36216720e-001,\n",

" 5.69634028e-001, 2.01668047e-023, 1.76909716e-001,\n",

" 8.65526378e-004, 3.41506861e-001, 0.00000000e+000,\n",

" 1.12319427e-231, 3.25429744e-001, 2.50330560e-001,\n",

" 2.38918636e-001, 7.64706959e-029],\n",

" [4.06300660e-001, 5.36514208e-002, 8.39475437e-002,\n",

" 2.08874884e-001, 9.38702072e-001, 1.98811127e-001,\n",

" 4.60240532e-009, 1.96808492e-001, 1.12319427e-231,\n",

" 0.00000000e+000, 6.99615740e-001, 1.03295766e-001,\n",

" 2.08799333e-001, 2.85374243e-036],\n",

" [9.52261425e-001, 1.60847582e-001, 3.71469037e-001,\n",

" 7.03844602e-001, 4.70093788e-001, 5.64246762e-001,\n",

" 1.26581605e-001, 2.53904935e-002, 3.25429744e-001,\n",

" 6.99615740e-001, 0.00000000e+000, 2.35441825e-001,\n",

" 3.15383179e-001, 4.75414918e-001],\n",

" [2.28461236e-001, 8.66447868e-001, 8.90503148e-001,\n",

" 1.52541637e-002, 6.57076013e-001, 2.41686809e-002,\n",

" 6.74797620e-005, 4.13650739e-003, 2.50330560e-001,\n",

" 1.03295766e-001, 2.35441825e-001, 0.00000000e+000,\n",

" 2.51464473e-001, 1.34826852e-055],\n",

" [5.48097586e-001, 1.27389774e-001, 2.38243578e-001,\n",

" 9.01602674e-001, 9.84458653e-001, 4.08370570e-001,\n",

" 8.07912562e-001, 4.36732384e-001, 2.38918636e-001,\n",

" 2.08799333e-001, 3.15383179e-001, 2.51464473e-001,\n",

" 0.00000000e+000, 2.27067756e-001],\n",

" [9.75106276e-002, 5.31116466e-001, 2.72106138e-001,\n",

" 1.98609526e-002, 1.08256524e-007, 1.25850456e-026,\n",

" 4.60367975e-243, 1.62203448e-001, 7.64706959e-029,\n",

" 2.85374243e-036, 4.75414918e-001, 1.34826852e-055,\n",

" 2.27067756e-001, 0.00000000e+000]]))"

]

},

"execution\_count": 40,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"\n",

"corr=spearmanr(data)\n",

"corr"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "1\_SfXMd62kQa"

},

"outputs": [],

"source": [

"import statsmodels.api as sm"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "QX-ml5WV3slY"

},

"outputs": [],

"source": [

"x=data[[\"EstimatedSalary\"]]\n",

"y=data[\"CreditScore\"]\n"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 448

},

"id": "f\_s\_S1jR4arm",

"outputId": "d2185511-8cdc-4552-dd76-930f2db78bb5"

},

"outputs": [

{

"data": {

"text/html": [

"<table class=\"simpletable\">\n",

"<caption>OLS Regression Results</caption>\n",

"<tr>\n",

" <th>Dep. Variable:</th> <td>CreditScore</td> <th> R-squared (uncentered):</th> <td> 0.735</td> \n",

"</tr>\n",

"<tr>\n",

" <th>Model:</th> <td>OLS</td> <th> Adj. R-squared (uncentered):</th> <td> 0.735</td> \n",

"</tr>\n",

"<tr>\n",

" <th>Method:</th> <td>Least Squares</td> <th> F-statistic: </th> <td>2.779e+04</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Date:</th> <td>Sat, 24 Sep 2022</td> <th> Prob (F-statistic):</th> <td> 0.00</td> \n",

"</tr>\n",

"<tr>\n",

" <th>Time:</th> <td>15:56:14</td> <th> Log-Likelihood: </th> <td> -72429.</td> \n",

"</tr>\n",

"<tr>\n",

" <th>No. Observations:</th> <td> 10000</td> <th> AIC: </th> <td>1.449e+05</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Df Residuals:</th> <td> 9999</td> <th> BIC: </th> <td>1.449e+05</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Df Model:</th> <td> 1</td> <th> </th> <td> </td> \n",

"</tr>\n",

"<tr>\n",

" <th>Covariance Type:</th> <td>nonrobust</td> <th> </th> <td> </td> \n",

"</tr>\n",

"</table>\n",

"<table class=\"simpletable\">\n",

"<tr>\n",

" <td></td> <th>coef</th> <th>std err</th> <th>t</th> <th>P>|t|</th> <th>[0.025</th> <th>0.975]</th> \n",

"</tr>\n",

"<tr>\n",

" <th>EstimatedSalary</th> <td> 0.0049</td> <td> 2.93e-05</td> <td> 166.705</td> <td> 0.000</td> <td> 0.005</td> <td> 0.005</td>\n",

"</tr>\n",

"</table>\n",

"<table class=\"simpletable\">\n",

"<tr>\n",

" <th>Omnibus:</th> <td>1758.359</td> <th> Durbin-Watson: </th> <td> 1.554</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Prob(Omnibus):</th> <td> 0.000</td> <th> Jarque-Bera (JB): </th> <td> 376.161</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Skew:</th> <td> 0.004</td> <th> Prob(JB): </th> <td>2.08e-82</td>\n",

"</tr>\n",

"<tr>\n",

" <th>Kurtosis:</th> <td> 2.050</td> <th> Cond. No. </th> <td> 1.00</td>\n",

"</tr>\n",

"</table><br/><br/>Notes:<br/>[1] R² is computed without centering (uncentered) since the model does not contain a constant.<br/>[2] Standard Errors assume that the covariance matrix of the errors is correctly specified."

],

"text/plain": [

"<class 'statsmodels.iolib.summary.Summary'>\n",

"\"\"\"\n",

" OLS Regression Results \n",

"=======================================================================================\n",

"Dep. Variable: CreditScore R-squared (uncentered): 0.735\n",

"Model: OLS Adj. R-squared (uncentered): 0.735\n",

"Method: Least Squares F-statistic: 2.779e+04\n",

"Date: Sat, 24 Sep 2022 Prob (F-statistic): 0.00\n",

"Time: 15:56:14 Log-Likelihood: -72429.\n",

"No. Observations: 10000 AIC: 1.449e+05\n",

"Df Residuals: 9999 BIC: 1.449e+05\n",

"Df Model: 1 \n",

"Covariance Type: nonrobust \n",

"===================================================================================\n",

" coef std err t P>|t| [0.025 0.975]\n",

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

"EstimatedSalary 0.0049 2.93e-05 166.705 0.000 0.005 0.005\n",

"==============================================================================\n",

"Omnibus: 1758.359 Durbin-Watson: 1.554\n",

"Prob(Omnibus): 0.000 Jarque-Bera (JB): 376.161\n",

"Skew: 0.004 Prob(JB): 2.08e-82\n",

"Kurtosis: 2.050 Cond. No. 1.00\n",

"==============================================================================\n",

"\n",

"Notes:\n",

"[1] R² is computed without centering (uncentered) since the model does not contain a constant.\n",

"[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.\n",

"\"\"\""

]

},

"execution\_count": 43,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"model=sm.OLS(y,x)\n",

"result=model.fit()\n",

"result.summary()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "HuC1A7Cj7AOH",

"outputId": "c10a5dd1-f179-4dba-feb6-54480372eb21"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[ 0.02188649],\n",

" [ 0.21653375],\n",

" [ 0.2406869 ],\n",

" ...,\n",

" [-1.00864308],\n",

" [-0.12523071],\n",

" [-1.07636976]])"

]

},

"execution\_count": 44,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"from sklearn.preprocessing import scale\n",

"x=scale(x)\n",

"x\n"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 386

},

"id": "dWFwbEtZ491N",

"outputId": "53b70ab4-0b46-46a4-cd5a-f0a1c3f47ae1"

},

"outputs": [

{

"data": {

"text/plain": [

"<seaborn.axisgrid.FacetGrid at 0x2223192f220>"

]

},

"execution\_count": 45,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 360x360 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.lmplot(x='Age',y='Balance',data=data)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "zK0yVgXSK8AJ",

"outputId": "e63e4ea8-2f72-4618-c4ba-9f0cf19477c6"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age', ylabel='CreditScore'>"

]

},

"execution\_count": 46,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.barplot(x=\"Age\",y=\"CreditScore\",data=data)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "y1r6s4B5K8Ku"

},

"outputs": [],

"source": [

"###outier detection"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 112

},

"id": "xlWqMUEgLdQr",

"outputId": "8967e1f6-f40d-4ac1-9058-baea7d64403e"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

" <th>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>0.75</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",

" <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",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId CreditScore Age Tenure Balance \\\n",

"0.75 7500.25 15753233.75 718.0 44.0 7.0 127644.24 \n",

"0.25 2500.75 15628528.25 584.0 32.0 3.0 0.00 \n",

"\n",

" NumOfProducts HasCrCard IsActiveMember EstimatedSalary Exited \n",

"0.75 2.0 1.0 1.0 149388.2475 0.0 \n",

"0.25 1.0 0.0 0.0 51002.1100 0.0 "

]

},

"execution\_count": 47,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"qnt = data.quantile(q=[0.75,0.25])\n",

"qnt"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "5ZNZCwzMLrLh",

"outputId": "74714037-4913-4477-a11f-272ce4b4251e"

},

"outputs": [

{

"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"

]

},

"execution\_count": 48,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"iqr=qnt.loc[0.75]-qnt.loc[0.25]\n",

"iqr"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "vGJy3DAF8LA4",

"outputId": "8d6f1b49-069b-4617-89e3-7eda47e29cc3"

},

"outputs": [

{

"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"

]

},

"execution\_count": 51,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"upper= qnt.loc[0.75]+1.5\*iqr\n",

"upper"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "a1Z2vOPqL7fG",

"outputId": "4dda8c0e-a5f6-4c17-94c1-1037c8c115c7"

},

"outputs": [

{

"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"

]

},

"execution\_count": 52,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"lower= qnt.loc[0.25]-1.5\*iqr\n",

"lower"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "Xy5UdGDQMLII"

},

"outputs": [],

"source": [

"###rplacing outlier"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "VB\_ntaajMPt1",

"outputId": "3a296682-06d0-4949-b115-f337656957df"

},

"outputs": [

{

"data": {

"text/plain": [

"<matplotlib.axes.\_subplots.AxesSubplot at 0x7fb1fdb656d0>"

]

},

"execution\_count": 37,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.boxplot(data[\"Age\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "M\_sj9xCKMaVr"

},

"outputs": [],

"source": [

"data[\"Age\"]= np.where(data[\"Age\"]>45,31,data[\"Age\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "bDkBKT0DMw\_V",

"outputId": "6a650ecc-6972-4dc8-eb20-e6e9f6b3bf36"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Age'>"

]

},

"execution\_count": 54,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.boxplot(data[\"Age\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "MO79hX5qM0it"

},

"outputs": [],

"source": [

"data[\"Balance\"]= np.where(data[\"Balance\"]>618,316,data[\"Balance\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 296

},

"id": "4GHkA2L7NTW5",

"outputId": "b6d596b5-1e44-4484-a817-db5c68a8ba86"

},

"outputs": [

{

"data": {

"text/plain": [

"<AxesSubplot:xlabel='Balance'>"

]

},

"execution\_count": 56,

"metadata": {},

"output\_type": "execute\_result"

},

{

"data": {

"image/png": "",

"text/plain": [

"<Figure size 432x288 with 1 Axes>"

]

},

"metadata": {

"needs\_background": "light"

},

"output\_type": "display\_data"

}

],

"source": [

"sns.boxplot(data[\"Balance\"])"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 206

},

"id": "V1R\_bUm0N9et",

"outputId": "4a15c8c0-d7ac-4e40-ebe1-ee5e770893a9"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

" <th>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>0.0</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>316.0</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>316.0</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>0.0</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>316.0</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>"

],

"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 EstimatedSalary \\\n",

"0 2 0.0 1 1 1 101348.88 \n",

"1 1 316.0 1 0 1 112542.58 \n",

"2 8 316.0 3 1 0 113931.57 \n",

"3 1 0.0 2 0 0 93826.63 \n",

"4 2 316.0 1 1 1 79084.10 \n",

"\n",

" Exited \n",

"0 1 \n",

"1 0 \n",

"2 1 \n",

"3 0 \n",

"4 0 "

]

},

"execution\_count": 57,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.head()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "fJd659yYNeJI"

},

"outputs": [],

"source": [

"data[\"Gender\"].replace({\"Female\":0, \"Male\":1},inplace = True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 363

},

"id": "UJBJma7sOXs4",

"outputId": "48efd9bc-cefc-4942-ffa3-e03ddf0e42a9"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

" <td>42</td>\n",

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

" <td>0.0</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>0</td>\n",

" <td>41</td>\n",

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

" <td>316.0</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>0</td>\n",

" <td>42</td>\n",

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

" <td>316.0</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>0</td>\n",

" <td>39</td>\n",

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

" <td>0.0</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>0</td>\n",

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

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

" <td>316.0</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",

" <tr>\n",

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

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

" <td>15574012</td>\n",

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

" <td>645</td>\n",

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

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

" <td>44</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>149756.71</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>7</td>\n",

" <td>15592531</td>\n",

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

" <td>822</td>\n",

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

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

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

" <td>7</td>\n",

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

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

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

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

" <td>10062.80</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>7</th>\n",

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

" <td>15656148</td>\n",

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

" <td>376</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>119346.88</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15792365</td>\n",

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

" <td>501</td>\n",

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

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

" <td>44</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>74940.50</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15592389</td>\n",

" <td>H?</td>\n",

" <td>684</td>\n",

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

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

" <td>27</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>71725.73</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId Surname CreditScore Geography Gender Age \\\n",

"0 1 15634602 Hargrave 619 France 0 42 \n",

"1 2 15647311 Hill 608 Spain 0 41 \n",

"2 3 15619304 Onio 502 France 0 42 \n",

"3 4 15701354 Boni 699 France 0 39 \n",

"4 5 15737888 Mitchell 850 Spain 0 43 \n",

"5 6 15574012 Chu 645 Spain 1 44 \n",

"6 7 15592531 Bartlett 822 France 1 31 \n",

"7 8 15656148 Obinna 376 Germany 0 29 \n",

"8 9 15792365 He 501 France 1 44 \n",

"9 10 15592389 H? 684 France 1 27 \n",

"\n",

" Tenure Balance NumOfProducts HasCrCard IsActiveMember EstimatedSalary \\\n",

"0 2 0.0 1 1 1 101348.88 \n",

"1 1 316.0 1 0 1 112542.58 \n",

"2 8 316.0 3 1 0 113931.57 \n",

"3 1 0.0 2 0 0 93826.63 \n",

"4 2 316.0 1 1 1 79084.10 \n",

"5 8 316.0 2 1 0 149756.71 \n",

"6 7 0.0 2 1 1 10062.80 \n",

"7 4 316.0 4 1 0 119346.88 \n",

"8 4 316.0 2 0 1 74940.50 \n",

"9 2 316.0 1 1 1 71725.73 \n",

"\n",

" Exited \n",

"0 1 \n",

"1 0 \n",

"2 1 \n",

"3 0 \n",

"4 0 \n",

"5 1 \n",

"6 0 \n",

"7 1 \n",

"8 0 \n",

"9 0 "

]

},

"execution\_count": 59,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.head(10)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "T2oVz12gOc0Z"

},

"outputs": [],

"source": [

"data[\"HasCrCard\"].replace({1:\"yes\",0:\"no\"},inplace = True)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 363

},

"id": "pGxj-YS4PaQx",

"outputId": "3e59afb6-842b-4090-d1b5-55ab8c289ddf"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

" <td>42</td>\n",

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

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

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

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

" <td>41</td>\n",

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

" <td>316.0</td>\n",

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

" <td>no</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>0</td>\n",

" <td>42</td>\n",

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

" <td>316.0</td>\n",

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

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

" <td>39</td>\n",

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

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

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

" <td>no</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>0</td>\n",

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

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

" <td>316.0</td>\n",

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

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

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

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

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15574012</td>\n",

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

" <td>645</td>\n",

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

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

" <td>44</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>149756.71</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>7</td>\n",

" <td>15592531</td>\n",

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

" <td>822</td>\n",

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

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

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

" <td>7</td>\n",

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

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

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

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

" <td>10062.80</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>7</th>\n",

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

" <td>15656148</td>\n",

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

" <td>376</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>119346.88</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15792365</td>\n",

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

" <td>501</td>\n",

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

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

" <td>44</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>74940.50</td>\n",

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

" </tr>\n",

" <tr>\n",

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

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

" <td>15592389</td>\n",

" <td>H?</td>\n",

" <td>684</td>\n",

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

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

" <td>27</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>71725.73</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId Surname CreditScore Geography Gender Age \\\n",

"0 1 15634602 Hargrave 619 France 0 42 \n",

"1 2 15647311 Hill 608 Spain 0 41 \n",

"2 3 15619304 Onio 502 France 0 42 \n",

"3 4 15701354 Boni 699 France 0 39 \n",

"4 5 15737888 Mitchell 850 Spain 0 43 \n",

"5 6 15574012 Chu 645 Spain 1 44 \n",

"6 7 15592531 Bartlett 822 France 1 31 \n",

"7 8 15656148 Obinna 376 Germany 0 29 \n",

"8 9 15792365 He 501 France 1 44 \n",

"9 10 15592389 H? 684 France 1 27 \n",

"\n",

" Tenure Balance NumOfProducts HasCrCard IsActiveMember EstimatedSalary \\\n",

"0 2 0.0 1 yes 1 101348.88 \n",

"1 1 316.0 1 no 1 112542.58 \n",

"2 8 316.0 3 yes 0 113931.57 \n",

"3 1 0.0 2 no 0 93826.63 \n",

"4 2 316.0 1 yes 1 79084.10 \n",

"5 8 316.0 2 yes 0 149756.71 \n",

"6 7 0.0 2 yes 1 10062.80 \n",

"7 4 316.0 4 yes 0 119346.88 \n",

"8 4 316.0 2 no 1 74940.50 \n",

"9 2 316.0 1 yes 1 71725.73 \n",

"\n",

" Exited \n",

"0 1 \n",

"1 0 \n",

"2 1 \n",

"3 0 \n",

"4 0 \n",

"5 1 \n",

"6 0 \n",

"7 1 \n",

"8 0 \n",

"9 0 "

]

},

"execution\_count": 61,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.head(10)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "hIi3DPsvQV7s"

},

"outputs": [],

"source": [

"#label encoding"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "c1m6l4OaQj59"

},

"outputs": [],

"source": [

"from sklearn.preprocessing import LabelEncoder\n",

"le=LabelEncoder()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "di\_OMjcARUK3"

},

"outputs": [],

"source": [

"data[\"Age\"]=le.fit\_transform(data[\"Age\"])\n"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "Vc6auq5WSaA3",

"outputId": "be48a858-50f9-491f-e8f4-2058675704b2"

},

"outputs": [

{

"data": {

"text/plain": [

"array([24, 23, 21, 25, 26, 13, 11, 9, 6, 16, 7, 17, 27, 14, 20, 18, 15,\n",

" 22, 19, 1, 8, 3, 4, 12, 10, 2, 5, 0], dtype=int64)"

]

},

"execution\_count": 76,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.Age.unique()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "iI9Xs76kSgjt",

"outputId": "e0e47459-80f1-4987-884f-ada795de369f"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[1, 15634602, 'Hargrave', ..., 'yes', 1, 101348.88],\n",

" [2, 15647311, 'Hill', ..., 'no', 1, 112542.58],\n",

" [3, 15619304, 'Onio', ..., 'yes', 0, 113931.57],\n",

" ...,\n",

" [9998, 15584532, 'Liu', ..., 'no', 1, 42085.58],\n",

" [9999, 15682355, 'Sabbatini', ..., 'yes', 0, 92888.52],\n",

" [10000, 15628319, 'Walker', ..., 'yes', 0, 38190.78]], dtype=object)"

]

},

"execution\_count": 77,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x=data.iloc[:,0:13].values\n",

"x"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "6D4aVa\_XSv\_Q",

"outputId": "7bf8074c-8ad8-45a0-8fd4-1f9a6a2f368f"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[1],\n",

" [0],\n",

" [1],\n",

" ...,\n",

" [1],\n",

" [1],\n",

" [0]], dtype=int64)"

]

},

"execution\_count": 78,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y=data.iloc[:,13:14].values\n",

"y"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 206

},

"id": "kJC6vPxHTBY9",

"outputId": "c2357b90-9ab5-4542-f603-33b9bb6d4441"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

" <td>24</td>\n",

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

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

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

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

" <td>23</td>\n",

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

" <td>316.0</td>\n",

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

" <td>no</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>0</td>\n",

" <td>24</td>\n",

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

" <td>316.0</td>\n",

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

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

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

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

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

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

" <td>no</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>0</td>\n",

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

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

" <td>316.0</td>\n",

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

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

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

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId Surname CreditScore Geography Gender Age \\\n",

"0 1 15634602 Hargrave 619 France 0 24 \n",

"1 2 15647311 Hill 608 Spain 0 23 \n",

"2 3 15619304 Onio 502 France 0 24 \n",

"3 4 15701354 Boni 699 France 0 21 \n",

"4 5 15737888 Mitchell 850 Spain 0 25 \n",

"\n",

" Tenure Balance NumOfProducts HasCrCard IsActiveMember EstimatedSalary \\\n",

"0 2 0.0 1 yes 1 101348.88 \n",

"1 1 316.0 1 no 1 112542.58 \n",

"2 8 316.0 3 yes 0 113931.57 \n",

"3 1 0.0 2 no 0 93826.63 \n",

"4 2 316.0 1 yes 1 79084.10 \n",

"\n",

" Exited \n",

"0 1 \n",

"1 0 \n",

"2 1 \n",

"3 0 \n",

"4 0 "

]

},

"execution\_count": 79,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.head()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "gV\_xRVTcTtAq"

},

"outputs": [],

"source": [

"from sklearn.preprocessing import OneHotEncoder"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "MqN2RTiLVXnF"

},

"outputs": [],

"source": [

"ohe= OneHotEncoder()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "o3x-RweqVcha",

"outputId": "afb77ff6-02a6-43b1-db9e-9b96bc681265"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[1., 0., 0., ..., 0., 0., 0.],\n",

" [0., 1., 0., ..., 0., 0., 0.],\n",

" [0., 0., 1., ..., 0., 0., 0.],\n",

" ...,\n",

" [0., 0., 0., ..., 0., 0., 0.],\n",

" [0., 0., 0., ..., 0., 0., 0.],\n",

" [0., 0., 0., ..., 0., 0., 0.]])"

]

},

"execution\_count": 82,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"z=ohe.fit\_transform(x[:,0:14]).toarray()\n",

"z"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "9yXDMMbXBW4p"

},

"source": [

"##Split the data into training and testing"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "H8TpvBeUV6s\_"

},

"outputs": [],

"source": [

"from sklearn.model\_selection import train\_test\_split"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "M\_BmRh9\_WZZq"

},

"outputs": [],

"source": [

"x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=0)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "LPeAiLz0WzYZ",

"outputId": "4c3b3513-6d3e-45d4-cb26-747a9c876eb7"

},

"outputs": [

{

"data": {

"text/plain": [

"((8000, 13), (2000, 13), (8000, 1), (2000, 1))"

]

},

"execution\_count": 86,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x\_train.shape,x\_test.shape,y\_train.shape,y\_test.shape"

]

},

{

"cell\_type": "markdown",

"metadata": {

"id": "G4zI-4cIYgXj"

},

"source": []

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "oQBg1h8fXJVM",

"outputId": "ebb50436-2f4b-409b-e7e1-e2f65f8c1453"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[7390, 15676909, 'Mishin', ..., 'yes', 0, 163830.64],\n",

" [9276, 15749265, 'Carslaw', ..., 'yes', 1, 57098.0],\n",

" [2996, 15582492, 'Moore', ..., 'yes', 0, 185630.76],\n",

" ...,\n",

" [3265, 15574372, 'Hoolan', ..., 'yes', 0, 181429.87],\n",

" [9846, 15664035, 'Parsons', ..., 'yes', 1, 148750.16],\n",

" [2733, 15592816, 'Udokamma', ..., 'yes', 0, 118855.26]],\n",

" dtype=object)"

]

},

"execution\_count": 87,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x\_train"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "xvW-p4jCXM2z",

"outputId": "25af08eb-11fd-4eec-fa14-b4c3a4338e97"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[9395, 15615753, 'Upchurch', ..., 'yes', 1, 192852.67],\n",

" [899, 15654700, 'Fallaci', ..., 'yes', 0, 128702.1],\n",

" [2399, 15633877, 'Morrison', ..., 'yes', 1, 75732.25],\n",

" ...,\n",

" [9550, 15772604, 'Chiemezie', ..., 'yes', 0, 141533.19],\n",

" [2741, 15787699, 'Burke', ..., 'yes', 1, 11276.48],\n",

" [6691, 15579223, 'Niu', ..., 'yes', 0, 192950.6]], dtype=object)"

]

},

"execution\_count": 88,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x\_test"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "VtE1knPHXQ0O",

"outputId": "246b26ea-6a5e-4eb5-8a9a-6ee6622872db"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[0],\n",

" [0],\n",

" [0],\n",

" ...,\n",

" [0],\n",

" [0],\n",

" [1]], dtype=int64)"

]

},

"execution\_count": 89,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y\_train"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "WBSduoG\_XVCu",

"outputId": "0fc73334-14f1-4db3-eb45-815a6db7d35e"

},

"outputs": [

{

"data": {

"text/plain": [

"array([[0],\n",

" [1],\n",

" [0],\n",

" ...,\n",

" [0],\n",

" [0],\n",

" [0]], dtype=int64)"

]

},

"execution\_count": 90,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y\_test"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "UB8I18h\_XpnN"

},

"outputs": [],

"source": [

"from sklearn.preprocessing import scale"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "RtGluk5Far9n",

"outputId": "76dea6b3-39eb-4814-92df-343fb06b2703"

},

"outputs": [

{

"data": {

"text/plain": [

"array([-0.32622142, -0.44003595, -1.53679418, ..., 0.60498839,\n",

" 1.25683526, 1.46377078])"

]

},

"execution\_count": 92,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x=data[\"CreditScore\"]\n",

"S=scale(x)\n",

"S"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "PSpSJEkxa3\_9"

},

"outputs": [],

"source": [

"###INDEPENDENT VARIABLE"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "nDR1l9DmbArD",

"outputId": "9c962a53-e3b1-40f1-ea59-9927c692bf72"

},

"outputs": [

{

"data": {

"text/plain": [

"0 24\n",

"1 23\n",

"2 24\n",

"3 21\n",

"4 25\n",

" ..\n",

"9995 21\n",

"9996 17\n",

"9997 18\n",

"9998 24\n",

"9999 10\n",

"Name: Age, Length: 10000, dtype: int64"

]

},

"execution\_count": 94,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y=data[\"Age\"]\n",

"y"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "oTX0y1DtbfFq"

},

"outputs": [],

"source": []

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 488

},

"id": "H-BQVn0SbXCl",

"outputId": "b953618c-7b6f-4e80-eb84-ba3cc667cc9c"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

" <th>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>28</th>\n",

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

" <td>15728693</td>\n",

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

" <td>574</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>100187.43</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>30</td>\n",

" <td>15656300</td>\n",

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

" <td>411</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>53483.21</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>30</th>\n",

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

" <td>15589475</td>\n",

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

" <td>591</td>\n",

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

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

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

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

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

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

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

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

" <td>140469.38</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>31</th>\n",

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

" <td>15706552</td>\n",

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

" <td>533</td>\n",

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

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

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

" <td>7</td>\n",

" <td>316.0</td>\n",

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

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

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

" <td>156731.91</td>\n",

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

" </tr>\n",

" <tr>\n",

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

" <td>33</td>\n",

" <td>15750181</td>\n",

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

" <td>553</td>\n",

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

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

" <td>23</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>81898.81</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",

" </tr>\n",

" <tr>\n",

" <th>9995</th>\n",

" <td>9996</td>\n",

" <td>15606229</td>\n",

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

" <td>771</td>\n",

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

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

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

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

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

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

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

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

" <td>96270.64</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9996</th>\n",

" <td>9997</td>\n",

" <td>15569892</td>\n",

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

" <td>516</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>101699.77</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9997</th>\n",

" <td>9998</td>\n",

" <td>15584532</td>\n",

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

" <td>709</td>\n",

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

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

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

" <td>7</td>\n",

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

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

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

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

" <td>42085.58</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9998</th>\n",

" <td>9999</td>\n",

" <td>15682355</td>\n",

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

" <td>772</td>\n",

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

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

" <td>24</td>\n",

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

" <td>316.0</td>\n",

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

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

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

" <td>92888.52</td>\n",

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

" </tr>\n",

" <tr>\n",

" <th>9999</th>\n",

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

" <td>15628319</td>\n",

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

" <td>792</td>\n",

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

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

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

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

" <td>316.0</td>\n",

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

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

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

" <td>38190.78</td>\n",

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"<p>9972 rows × 14 columns</p>\n",

"</div>"

],

"text/plain": [

" RowNumber CustomerId Surname CreditScore Geography Gender \\\n",

"28 29 15728693 McWilliams 574 Germany 0 \n",

"29 30 15656300 Lucciano 411 France 1 \n",

"30 31 15589475 Azikiwe 591 Spain 0 \n",

"31 32 15706552 Odinakachukwu 533 France 1 \n",

"32 33 15750181 Sanderson 553 Germany 1 \n",

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

"9995 9996 15606229 Obijiaku 771 France 1 \n",

"9996 9997 15569892 Johnstone 516 France 1 \n",

"9997 9998 15584532 Liu 709 France 0 \n",

"9998 9999 15682355 Sabbatini 772 Germany 1 \n",

"9999 10000 15628319 Walker 792 France 0 \n",

"\n",

" Age Tenure Balance NumOfProducts HasCrCard IsActiveMember \\\n",

"28 25 3 316.0 1 yes 1 \n",

"29 11 0 316.0 2 yes 1 \n",

"30 21 3 0.0 3 yes 0 \n",

"31 18 7 316.0 1 no 1 \n",

"32 23 9 316.0 2 no 0 \n",

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

"9995 21 5 0.0 2 yes 0 \n",

"9996 17 10 316.0 1 yes 1 \n",

"9997 18 7 0.0 1 no 1 \n",

"9998 24 3 316.0 2 yes 0 \n",

"9999 10 4 316.0 1 yes 0 \n",

"\n",

" EstimatedSalary Exited \n",

"28 100187.43 0 \n",

"29 53483.21 0 \n",

"30 140469.38 1 \n",

"31 156731.91 0 \n",

"32 81898.81 0 \n",

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

"9995 96270.64 0 \n",

"9996 101699.77 0 \n",

"9997 42085.58 1 \n",

"9998 92888.52 1 \n",

"9999 38190.78 0 \n",

"\n",

"[9972 rows x 14 columns]"

]

},

"execution\_count": 95,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"x=data.drop(data[\"Age\"],axis=0)\n",

"x"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "z\_fUC\_R-cqR8"

},

"outputs": [],

"source": [

"###spiliting dependent variable"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "9xsISAWgcCxG",

"outputId": "0af62a12-e6eb-4961-a135-a795b9dcf9ee"

},

"outputs": [

{

"data": {

"text/plain": [

"array([1, 0, 1, ..., 1, 1, 0], dtype=int64)"

]

},

"execution\_count": 97,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"y=data.iloc[:,-1].values\n",

"y"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 143

},

"id": "HdwTGyPEe0j2",

"outputId": "9d35daa1-286b-443c-a573-ff5b2082df37"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

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

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

" </tr>\n",

" <tr>\n",

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

" <td>2.0</td>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Age CreditScore Balance\n",

"0 1.0 1.0 1\n",

"1 2.0 NaN 2\n",

"2 NaN NaN 3"

]

},

"execution\_count": 98,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data=pd.DataFrame({\"Age\":[1,2,np.nan],\"CreditScore\":[1,np.nan,np.nan],\"Balance\":[1,2,3]})\n",

"data"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "Jz6wJwL9ffjM",

"outputId": "e9314cbe-d429-45f9-eb57-b944f6a06eef"

},

"outputs": [

{

"data": {

"text/plain": [

"Age True\n",

"CreditScore True\n",

"Balance False\n",

"dtype: bool"

]

},

"execution\_count": 99,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.isnull().any()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "snVB0qtEfrcW",

"outputId": "2928545b-cf45-485e-f72c-d2a06d0d4ba8"

},

"outputs": [

{

"data": {

"text/plain": [

"Age 1\n",

"CreditScore 2\n",

"Balance 0\n",

"dtype: int64"

]

},

"execution\_count": 100,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.isnull().sum()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 81

},

"id": "xuGAVQ2GfrIZ",

"outputId": "9c105914-0c72-420f-b548-5def00568f92"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Age CreditScore Balance\n",

"0 1.0 1.0 1"

]

},

"execution\_count": 101,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.dropna()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/",

"height": 143

},

"id": "I97C45G-f3TO",

"outputId": "ef0701fe-681a-4221-ccfb-fb94e9de5204"

},

"outputs": [

{

"data": {

"text/html": [

"<div>\n",

"<style scoped>\n",

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

" vertical-align: middle;\n",

" }\n",

"\n",

" .dataframe tbody tr th {\n",

" vertical-align: top;\n",

" }\n",

"\n",

" .dataframe thead th {\n",

" text-align: right;\n",

" }\n",

"</style>\n",

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

" <thead>\n",

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

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

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

" </tr>\n",

" </thead>\n",

" <tbody>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" <tr>\n",

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

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

" </tr>\n",

" </tbody>\n",

"</table>\n",

"</div>"

],

"text/plain": [

" Balance\n",

"0 1\n",

"1 2\n",

"2 3"

]

},

"execution\_count": 102,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data.dropna(axis=1)"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"colab": {

"base\_uri": "https://localhost:8080/"

},

"id": "eQbAgX0jf\_ns",

"outputId": "8b757b68-0040-40f3-c9d5-bcd010bc523e"

},

"outputs": [

{

"data": {

"text/plain": [

"1.5"

]

},

"execution\_count": 103,

"metadata": {},

"output\_type": "execute\_result"

}

],

"source": [

"data[\"Age\"].mean()"

]

},

{

"cell\_type": "code",

"execution\_count": null,

"metadata": {

"id": "iJxahw6x8LBJ"

},

"outputs": [],

"source": []

}

],

"metadata": {

"colab": {

"collapsed\_sections": [],

"provenance": []

},

"kernelspec": {

"display\_name": "Python 3.9.1 64-bit",

"language": "python",

"name": "python3"

},

"language\_info": {

"codemirror\_mode": {

"name": "ipython",

"version": 3

},

"file\_extension": ".py",

"mimetype": "text/x-python",

"name": "python",

"nbconvert\_exporter": "python",

"pygments\_lexer": "ipython3",

"version": "3.9.1"

},

"vscode": {

"interpreter": {

"hash": "3eb5c436f6c629c85c5d36456d7868b63ffe24c1f53ee212367a20d7901b329c"

}

}

},

"nbformat": 4,

"nbformat\_minor": 0

}