**Basic operations**

In [1]:

**import** numpy **as** np

In [2]:

**import** pandas **as** pd

**Import dataset**

In [3]:

data **=** pd**.**read\_csv(r"C:\Users\user\Desktop\dataset.csv")

**Select first 10 rows**

In [4]:

data**.**head(10)

Out[4]:

|  | **ID** | **model** | **engine\_power** | **age\_in\_days** | **km** | **previous\_owners** | **lat** | **lon** | **price** | **Unnamed: 9** | **Unnamed: 10** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1.0 | lounge | 51.0 | 882.0 | 25000.0 | 1.0 | 44.907242 | 8.611559868 | 8900 | NaN | NaN |
| **1** | 2.0 | pop | 51.0 | 1186.0 | 32500.0 | 1.0 | 45.666359 | 12.24188995 | 8800 | NaN | NaN |
| **2** | 3.0 | sport | 74.0 | 4658.0 | 142228.0 | 1.0 | 45.503300 | 11.41784 | 4200 | NaN | NaN |
| **3** | 4.0 | lounge | 51.0 | 2739.0 | 160000.0 | 1.0 | 40.633171 | 17.63460922 | 6000 | NaN | NaN |
| **4** | 5.0 | pop | 73.0 | 3074.0 | 106880.0 | 1.0 | 41.903221 | 12.49565029 | 5700 | NaN | NaN |
| **5** | 6.0 | pop | 74.0 | 3623.0 | 70225.0 | 1.0 | 45.000702 | 7.68227005 | 7900 | NaN | NaN |
| **6** | 7.0 | lounge | 51.0 | 731.0 | 11600.0 | 1.0 | 44.907242 | 8.611559868 | 10750 | NaN | NaN |
| **7** | 8.0 | lounge | 51.0 | 1521.0 | 49076.0 | 1.0 | 41.903221 | 12.49565029 | 9190 | NaN | NaN |
| **8** | 9.0 | sport | 73.0 | 4049.0 | 76000.0 | 1.0 | 45.548000 | 11.54946995 | 5600 | NaN | NaN |
| **9** | 10.0 | sport | 51.0 | 3653.0 | 89000.0 | 1.0 | 45.438301 | 10.99170017 | 6000 | NaN | NaN |

**select last 10 rows**

In [5]:

data**.**tail(10)

Out[5]:

|  | **ID** | **model** | **engine\_power** | **age\_in\_days** | **km** | **previous\_owners** | **lat** | **lon** | **price** | **Unnamed: 9** | **Unnamed: 10** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1539** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | avg | 8576.003901 | NaN | NaN |
| **1540** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | count | 1538 | NaN | NaN |
| **1541** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | countif | 377 | NaN | >10000 |
| **1542** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | sumif | 4017825 | NaN | NaN |
| **1543** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | counta (not empty) | 1538 | NaN | NaN |
| **1544** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | length | 5 | NaN | NaN |
| **1545** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | concat | lonprice | NaN | NaN |
| **1546** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | Null values | NO | NaN | NaN |
| **1547** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | find | 1 | NaN | NaN |
| **1548** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | search | 1 | NaN | NaN |

**to display the statistical data**

In [7]:

data**.**describe()

Out[7]:

|  | **ID** | **engine\_power** | **age\_in\_days** | **km** | **previous\_owners** | **lat** | **Unnamed: 9** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **count** | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 1538.000000 | 0.0 |
| **mean** | 769.500000 | 51.904421 | 1650.980494 | 53396.011704 | 1.123537 | 43.541361 | NaN |
| **std** | 444.126671 | 3.988023 | 1289.522278 | 40046.830723 | 0.416423 | 2.133518 | NaN |
| **min** | 1.000000 | 51.000000 | 366.000000 | 1232.000000 | 1.000000 | 36.855839 | NaN |
| **25%** | 385.250000 | 51.000000 | 670.000000 | 20006.250000 | 1.000000 | 41.802990 | NaN |
| **50%** | 769.500000 | 51.000000 | 1035.000000 | 39031.000000 | 1.000000 | 44.394096 | NaN |
| **75%** | 1153.750000 | 51.000000 | 2616.000000 | 79667.750000 | 1.000000 | 45.467960 | NaN |
| **max** | 1538.000000 | 77.000000 | 4658.000000 | 235000.000000 | 4.000000 | 46.795612 | NaN |

**to display rows and cols**

In [12]:

data**.**shape

Out[12]:

(1549, 11)

**to display size of table**

In [13]:

data**.**size

Out[13]:

17039

**to get the misssing values**

In [14]:

data**.**isna()

Out[14]:

|  | **ID** | **model** | **engine\_power** | **age\_in\_days** | **km** | **previous\_owners** | **lat** | **lon** | **price** | **Unnamed: 9** | **Unnamed: 10** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | False | False | False | False | False | False | False | False | False | True | True |
| **1** | False | False | False | False | False | False | False | False | False | True | True |
| **2** | False | False | False | False | False | False | False | False | False | True | True |
| **3** | False | False | False | False | False | False | False | False | False | True | True |
| **4** | False | False | False | False | False | False | False | False | False | True | True |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **1544** | True | True | True | True | True | True | True | False | False | True | True |
| **1545** | True | True | True | True | True | True | True | False | False | True | True |
| **1546** | True | True | True | True | True | True | True | False | False | True | True |
| **1547** | True | True | True | True | True | True | True | False | False | True | True |
| **1548** | True | True | True | True | True | True | True | False | False | True | True |

1549 rows × 11 columns

**fill data in missing**

In [16]:

data**.**fillna("ts")

Out[16]:

|  | **ID** | **model** | **engine\_power** | **age\_in\_days** | **km** | **previous\_owners** | **lat** | **lon** | **price** | **Unnamed: 9** | **Unnamed: 10** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 1.0 | lounge | 51.0 | 882.0 | 25000.0 | 1.0 | 44.907242 | 8.611559868 | 8900 | ts | ts |
| **1** | 2.0 | pop | 51.0 | 1186.0 | 32500.0 | 1.0 | 45.666359 | 12.24188995 | 8800 | ts | ts |
| **2** | 3.0 | sport | 74.0 | 4658.0 | 142228.0 | 1.0 | 45.5033 | 11.41784 | 4200 | ts | ts |
| **3** | 4.0 | lounge | 51.0 | 2739.0 | 160000.0 | 1.0 | 40.633171 | 17.63460922 | 6000 | ts | ts |
| **4** | 5.0 | pop | 73.0 | 3074.0 | 106880.0 | 1.0 | 41.903221 | 12.49565029 | 5700 | ts | ts |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **1544** | ts | ts | ts | ts | ts | ts | ts | length | 5 | ts | ts |
| **1545** | ts | ts | ts | ts | ts | ts | ts | concat | lonprice | ts | ts |
| **1546** | ts | ts | ts | ts | ts | ts | ts | Null values | NO | ts | ts |
| **1547** | ts | ts | ts | ts | ts | ts | ts | find | 1 | ts | ts |
| **1548** | ts | ts | ts | ts | ts | ts | ts | search | 1 | ts | ts |

1549 rows × 11 columns

**dataset 2**

**import dataset**

In [28]:

data1 **=** pd**.**read\_csv(r"C:\Users\user\Desktop\2015.csv")

**select first 10 rows**

In [29]:

data1**.**head(10)

Out[29]:

|  | **Country** | **Region** | **Happiness Rank** | **Happiness Score** | **Standard Error** | **Economy (GDP per Capita)** | **Family** | **Health (Life Expectancy)** | **Freedom** | **Trust (Government Corruption)** | **Generosity** | **Dystopia Residual** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | Switzerland | Western Europe | 1 | 7.587 | 0.03411 | 1.39651 | 1.34951 | 0.94143 | 0.66557 | 0.41978 | 0.29678 | 2.51738 |
| **1** | Iceland | Western Europe | 2 | 7.561 | 0.04884 | 1.30232 | 1.40223 | 0.94784 | 0.62877 | 0.14145 | 0.43630 | 2.70201 |
| **2** | Denmark | Western Europe | 3 | 7.527 | 0.03328 | 1.32548 | 1.36058 | 0.87464 | 0.64938 | 0.48357 | 0.34139 | 2.49204 |
| **3** | Norway | Western Europe | 4 | 7.522 | 0.03880 | 1.45900 | 1.33095 | 0.88521 | 0.66973 | 0.36503 | 0.34699 | 2.46531 |
| **4** | Canada | North America | 5 | 7.427 | 0.03553 | 1.32629 | 1.32261 | 0.90563 | 0.63297 | 0.32957 | 0.45811 | 2.45176 |
| **5** | Finland | Western Europe | 6 | 7.406 | 0.03140 | 1.29025 | 1.31826 | 0.88911 | 0.64169 | 0.41372 | 0.23351 | 2.61955 |
| **6** | Netherlands | Western Europe | 7 | 7.378 | 0.02799 | 1.32944 | 1.28017 | 0.89284 | 0.61576 | 0.31814 | 0.47610 | 2.46570 |
| **7** | Sweden | Western Europe | 8 | 7.364 | 0.03157 | 1.33171 | 1.28907 | 0.91087 | 0.65980 | 0.43844 | 0.36262 | 2.37119 |
| **8** | New Zealand | Australia and New Zealand | 9 | 7.286 | 0.03371 | 1.25018 | 1.31967 | 0.90837 | 0.63938 | 0.42922 | 0.47501 | 2.26425 |
| **9** | Australia | Australia and New Zealand | 10 | 7.284 | 0.04083 | 1.33358 | 1.30923 | 0.93156 | 0.65124 | 0.35637 | 0.43562 | 2.26646 |

**last 10 rows**

In [30]:

data1**.**tail(10)

Out[30]:

|  | **Country** | **Region** | **Happiness Rank** | **Happiness Score** | **Standard Error** | **Economy (GDP per Capita)** | **Family** | **Health (Life Expectancy)** | **Freedom** | **Trust (Government Corruption)** | **Generosity** | **Dystopia Residual** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **148** | Chad | Sub-Saharan Africa | 149 | 3.667 | 0.03830 | 0.34193 | 0.76062 | 0.15010 | 0.23501 | 0.05269 | 0.18386 | 1.94296 |
| **149** | Guinea | Sub-Saharan Africa | 150 | 3.656 | 0.03590 | 0.17417 | 0.46475 | 0.24009 | 0.37725 | 0.12139 | 0.28657 | 1.99172 |
| **150** | Ivory Coast | Sub-Saharan Africa | 151 | 3.655 | 0.05141 | 0.46534 | 0.77115 | 0.15185 | 0.46866 | 0.17922 | 0.20165 | 1.41723 |
| **151** | Burkina Faso | Sub-Saharan Africa | 152 | 3.587 | 0.04324 | 0.25812 | 0.85188 | 0.27125 | 0.39493 | 0.12832 | 0.21747 | 1.46494 |
| **152** | Afghanistan | Southern Asia | 153 | 3.575 | 0.03084 | 0.31982 | 0.30285 | 0.30335 | 0.23414 | 0.09719 | 0.36510 | 1.95210 |
| **153** | Rwanda | Sub-Saharan Africa | 154 | 3.465 | 0.03464 | 0.22208 | 0.77370 | 0.42864 | 0.59201 | 0.55191 | 0.22628 | 0.67042 |
| **154** | Benin | Sub-Saharan Africa | 155 | 3.340 | 0.03656 | 0.28665 | 0.35386 | 0.31910 | 0.48450 | 0.08010 | 0.18260 | 1.63328 |
| **155** | Syria | Middle East and Northern Africa | 156 | 3.006 | 0.05015 | 0.66320 | 0.47489 | 0.72193 | 0.15684 | 0.18906 | 0.47179 | 0.32858 |
| **156** | Burundi | Sub-Saharan Africa | 157 | 2.905 | 0.08658 | 0.01530 | 0.41587 | 0.22396 | 0.11850 | 0.10062 | 0.19727 | 1.83302 |
| **157** | Togo | Sub-Saharan Africa | 158 | 2.839 | 0.06727 | 0.20868 | 0.13995 | 0.28443 | 0.36453 | 0.10731 | 0.16681 | 1.56726 |

**statistical data**

In [31]:

data1**.**describe()

Out[31]:

|  | **Happiness Rank** | **Happiness Score** | **Standard Error** | **Economy (GDP per Capita)** | **Family** | **Health (Life Expectancy)** | **Freedom** | **Trust (Government Corruption)** | **Generosity** | **Dystopia Residual** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **count** | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 |
| **mean** | 79.493671 | 5.375734 | 0.047885 | 0.846137 | 0.991046 | 0.630259 | 0.428615 | 0.143422 | 0.237296 | 2.098977 |
| **std** | 45.754363 | 1.145010 | 0.017146 | 0.403121 | 0.272369 | 0.247078 | 0.150693 | 0.120034 | 0.126685 | 0.553550 |
| **min** | 1.000000 | 2.839000 | 0.018480 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.328580 |
| **25%** | 40.250000 | 4.526000 | 0.037268 | 0.545808 | 0.856823 | 0.439185 | 0.328330 | 0.061675 | 0.150553 | 1.759410 |
| **50%** | 79.500000 | 5.232500 | 0.043940 | 0.910245 | 1.029510 | 0.696705 | 0.435515 | 0.107220 | 0.216130 | 2.095415 |
| **75%** | 118.750000 | 6.243750 | 0.052300 | 1.158448 | 1.214405 | 0.811013 | 0.549092 | 0.180255 | 0.309883 | 2.462415 |
| **max** | 158.000000 | 7.587000 | 0.136930 | 1.690420 | 1.402230 | 1.025250 | 0.669730 | 0.551910 | 0.795880 | 3.602140 |

**dropthe missing data**

In [33]:

data1**.**dropna()

Out[33]:

|  | **Country** | **Region** | **Happiness Rank** | **Happiness Score** | **Standard Error** | **Economy (GDP per Capita)** | **Family** | **Health (Life Expectancy)** | **Freedom** | **Trust (Government Corruption)** | **Generosity** | **Dystopia Residual** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | Switzerland | Western Europe | 1 | 7.587 | 0.03411 | 1.39651 | 1.34951 | 0.94143 | 0.66557 | 0.41978 | 0.29678 | 2.51738 |
| **1** | Iceland | Western Europe | 2 | 7.561 | 0.04884 | 1.30232 | 1.40223 | 0.94784 | 0.62877 | 0.14145 | 0.43630 | 2.70201 |
| **2** | Denmark | Western Europe | 3 | 7.527 | 0.03328 | 1.32548 | 1.36058 | 0.87464 | 0.64938 | 0.48357 | 0.34139 | 2.49204 |
| **3** | Norway | Western Europe | 4 | 7.522 | 0.03880 | 1.45900 | 1.33095 | 0.88521 | 0.66973 | 0.36503 | 0.34699 | 2.46531 |
| **4** | Canada | North America | 5 | 7.427 | 0.03553 | 1.32629 | 1.32261 | 0.90563 | 0.63297 | 0.32957 | 0.45811 | 2.45176 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **153** | Rwanda | Sub-Saharan Africa | 154 | 3.465 | 0.03464 | 0.22208 | 0.77370 | 0.42864 | 0.59201 | 0.55191 | 0.22628 | 0.67042 |
| **154** | Benin | Sub-Saharan Africa | 155 | 3.340 | 0.03656 | 0.28665 | 0.35386 | 0.31910 | 0.48450 | 0.08010 | 0.18260 | 1.63328 |
| **155** | Syria | Middle East and Northern Africa | 156 | 3.006 | 0.05015 | 0.66320 | 0.47489 | 0.72193 | 0.15684 | 0.18906 | 0.47179 | 0.32858 |
| **156** | Burundi | Sub-Saharan Africa | 157 | 2.905 | 0.08658 | 0.01530 | 0.41587 | 0.22396 | 0.11850 | 0.10062 | 0.19727 | 1.83302 |
| **157** | Togo | Sub-Saharan Africa | 158 | 2.839 | 0.06727 | 0.20868 | 0.13995 | 0.28443 | 0.36453 | 0.10731 | 0.16681 | 1.56726 |

158 rows × 12 columns