

```
In [6]: #1,2,...read csv file.....for pandas#
import pandas as pd
d=pd.read_csv("eda_data.csv")
```

```
In [7]: print(d)

   Unnamed: 0      Job Title      Salary Estimate      Job Description      Rating      Company Name      Location      Headquarters      Size      Founded      ...      age      python_yn      R_yn      spark      aws      excel      job_simp      seniority
0           0           Data Scientist      53K-91K (Glassdoor est.)      Data Scientist\nLocation: Albuquerque, NM\nEdu...      3.8      Tecolote Research\n3.8      Albuquerque, NM      Goleta, CA      501 to 1000 employees      1973      ...      47      1      0      0      0      1      data scientist      na
1           1      Healthcare Data Scientist      63K-112K (Glassdoor est.)      What You Will Do:\n\nI. General Summary\n\nThe...      3.4      University of Maryland Medical System\n3.4      Linthicum, MD      Baltimore, MD      10000+ employees      1984      ...      36      1      0      0      0      0      data scientist      na
2           2           Data Scientist      80K-90K (Glassdoor est.)      KnowBe4, Inc. is a high growth information sec...      4.8      KnowBe4\n4.8      Clearwater, FL      Clearwater, FL      501 to 1000 employees      2010      ...      10      1      0      1      0      1      data scientist      na
3           3           Data Scientist      56K-97K (Glassdoor est.)      *Organization and Job ID**\nJob ID: 310709\n\n...      3.8      PNNL\n3.8      Richland, WA      Richland, WA      1001 to 5000 employees      1965      ...      55      1      0      0      0      0      data scientist      na
4           4           Data Scientist      86K-143K (Glassdoor est.)      Data Scientist\nAffinity Solutions / Marketing...      2.9      Affinity Solutions\n2.9      New York, NY      New York, NY      51 to 200 employees      1998      ...      22      1      0      0      0      1      data scientist      na
...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...
737       737      Sr Scientist, Immuno-Oncology - Oncology      58K-111K (Glassdoor est.)      Site Name: USA - Massachusetts - Cambridge\nPo...      3.9      GSK\n3.9      Cambridge, MA      Brentford, United Kingdom      10000+ employees      1830      ...      190      0      0      0      0      1      0
738       738      Senior Data Engineer      72K-133K (Glassdoor est.)      THE CHALLENGE\nEventbrite has a world-class da...      4.4      Eventbrite\n4.4      Nashville, TN      San Francisco, CA      1001 to 5000 employees      2006      ...      14      1      0      1      1      0      0
739       739      Project Scientist - Auton Lab, Robotics Institute      56K-91K (Glassdoor est.)      The Auton Lab at Carnegie Mellon University is...      2.6      Software Engineering Institute\n2.6      Pittsburgh, PA      Pittsburgh, PA      501 to 1000 employees      1984      ...      36      0      0      0      0      0      1
740       740      Data Science Manager      95K-160K (Glassdoor est.)      Data Science ManagerResponsibilities:\n\nOvers...      3.2      Numeric, LLC\n3.2      Allentown, PA      Chadds Ford, PA      1 to 50 employees      -1      ...      -1      0      0      0      0      0      1      manager
741       741      Research Scientist - Security and Privacy      61K-126K (Glassdoor est.)      Returning Candidate? Log back in to the Career...      3.6      Riverside Research Institute\n3.6      Beavercreek, OH      Arlington, VA      501 to 1000 employees      1967      ...      53      1      0      0      0      0      0

   num_comp
0           0
1           0
2           0
3           3
4           3
...      ...
737         3
738         3
739         0
740         0
741         0

[742 rows x 33 columns]
```

```
In [8]: d.head(10)
```

	Unnamed: 0	Job Title	Salary Estimate	Job Description	Rating	Company Name	Location	Headquarters	Size	Founded	...	age	python_yn	R_yn	spark	aws	excel	job_simp	seniority
0	0	Data Scientist	53K-91K (Glassdoor est.)	Data Scientist\nLocation: Albuquerque, NM\nEdu...	3.8	Tecolote Research\n3.8	Albuquerque, NM	Goleta, CA	501 to 1000 employees	1973	...	47	1	0	0	0	1	data scientist	na
1	1	Healthcare Data Scientist	63K-112K (Glassdoor est.)	What You Will Do:\n\nI. General Summary\n\nThe...	3.4	University of Maryland Medical System\n3.4	Linthicum, MD	Baltimore, MD	10000+ employees	1984	...	36	1	0	0	0	0	data scientist	na
2	2	Data Scientist	80K-90K (Glassdoor est.)	KnowBe4, Inc. is a high growth information sec...	4.8	KnowBe4\n4.8	Clearwater, FL	Clearwater, FL	501 to 1000 employees	2010	...	10	1	0	1	0	1	data scientist	na
3	3	Data Scientist	56K-97K (Glassdoor est.)	*Organization and Job ID**\nJob ID: 310709\n\n...	3.8	PNNL\n3.8	Richland, WA	Richland, WA	1001 to 5000 employees	1965	...	55	1	0	0	0	0	data scientist	na
4	4	Data Scientist	86K-143K (Glassdoor est.)	Data Scientist\nAffinity Solutions / Marketing...	2.9	Affinity Solutions\n2.9	New York, NY	New York, NY	51 to 200 employees	1998	...	22	1	0	0	0	1	data scientist	na
5	5	Data Scientist	71K-119K (Glassdoor est.)	CyrusOne is seeking a talented Data Scientist ...	3.4	CyrusOne\n3.4	Dallas, TX	Dallas, TX	201 to 500 employees	2000	...	20	1	0	0	1	1	data scientist	na
6	6	Data Scientist	54K-93K (Glassdoor est.)	Description\n\nJob Description\n\nPlease only local candida...	4.1	ClearOne Advantage\n4.1	Baltimore, MD	Baltimore, MD	501 to 1000 employees	2008	...	12	0	0	0	0	1	data scientist	na
7	7	Data Scientist	86K-142K (Glassdoor est.)	Advanced Analytics - Lead Data Scientist\nOver...	3.8	Logic20/20\n3.8	San Jose, CA	Seattle, WA	201 to 500 employees	2005	...	15	1	0	1	1	1	data scientist	na
8	8	Research Scientist	38K-84K (Glassdoor est.)	SUMMARY\n\nThe Research Scientist I will be ta...	3.3	Rochester Regional Health\n3.3	Rochester, NY	Rochester, NY	10000+ employees	2014	...	6	0	0	0	0	0	na	na
9	9	Data Scientist	120K-160K (Glassdoor est.)	isn't your usual company. Our work is powered ...	4.6	<intent>\n4.6	New York, NY	New York, NY	51 to 200 employees	2009	...	11	1	0	1	0	0	data scientist	na

10 rows x 33 columns

```
In [9]: #####3 3 clean empty cell#####
d.dropna()
```

	Unnamed: 0	Job Title	Salary Estimate	Job Description	Rating	Company Name	Location	Headquarters	Size	Founded	...	age	python_yn	R_yn	spark	aws	excel	job_simp	seniority
0	0	Data Scientist	53K-91K (Glassdoor est.)	Data Scientist\nLocation: Albuquerque, NM\nEdu...	3.8	Tecolote Research\n3.8	Albuquerque, NM	Goleta, CA	501 to 1000 employees	1973	...	47	1	0	0	0	1	d scien	
1	1	Healthcare Data Scientist	63K-112K (Glassdoor est.)	What You Will Do:\n\nI. General Summary\n\nThe...	3.4	University of Maryland Medical System\n3.4	Linthicum, MD	Baltimore, MD	10000+ employees	1984	...	36	1	0	0	0	0	d scien	
2	2	Data Scientist	80K-90K (Glassdoor est.)	KnowBe4, Inc. is a high growth information sec...	4.8	KnowBe4\n4.8	Clearwater, FL	Clearwater, FL	501 to 1000 employees	2010	...	10	1	0	1	0	1	d scien	
3	3	Data Scientist	56K-97K (Glassdoor est.)	*Organization and Job ID**\nJob ID: 310709\n\n...	3.8	PNNL\n3.8	Richland, WA	Richland, WA	1001 to 5000 employees	1965	...	55	1	0	0	0	0	d scien	
4	4	Data Scientist	86K-143K (Glassdoor est.)	Data Scientist\nAffinity Solutions / Marketing...	2.9	Affinity Solutions\n2.9	New York, NY	New York, NY	51 to 200 employees	1998	...	22	1	0	0	0	1	d scien	
...	
737	737	Sr Scientist, Immuno-Oncology - Oncology	58K-111K (Glassdoor est.)	Site Name: USA - Massachusetts - Cambridge\nPo...	3.9	GSK\n3.9	Cambridge, MA	Brentford, United Kingdom	10000+ employees	1830	...	190	0	0	0	0	1	0	
738	738	Senior Data Engineer	72K-133K (Glassdoor est.)	THE CHALLENGE\nEventbrite has a world-class da...	4.4	Eventbrite\n4.4	Nashville, TN	San Francisco, CA	1001 to 5000 employees	2006	...	14	1	0	1	1	0	d engin	
739	739	Project Scientist - Auton Lab, Robotics Institute	56K-91K (Glassdoor est.)	The Auton Lab at Carnegie Mellon University is...	2.6	Software Engineering Institute\n2.6	Pittsburgh, PA	Pittsburgh, PA	501 to 1000 employees	1984	...	36	0	0	0	0	0	1	
740	740	Data Science Manager	95K-160K (Glassdoor est.)	Data Science ManagerResponsibilities:\n\nOvers...	3.2	Numeric, LLC\n3.2	Allentown, PA	Chadds Ford, PA	1 to 50 employees	-1	...	-1	0	0	0	0	1	mana	
741	741	Research Scientist - Security and Privacy	61K-126K (Glassdoor est.)	Returning Candidate? Log back in to the Career...	3.6	Riverside Research Institute\n3.6	Beavercreek, OH	Arlington, VA	501 to 1000 employees	1967	...	53	1	0	0	0	0	0	

742 rows x 33 columns

```
In [10]: #.....#####remove duplicates#####.....#
d.drop_duplicates()
```

	Unnamed: 0	Job Title	Salary Estimate	Job Description	Rating	Company Name	Location	Headquarters	Size	Founded	...	age	python_yn	R_yn	spark	aws	excel	job_simp	seniority
0	0	Data Scientist	53K-91K (Glassdoor est.)	Data Scientist\nLocation: Albuquerque, NM\nEdu...	3.8	Tecolote Research\n3.8	Albuquerque, NM	Goleta, CA	501 to 1000 employees	1973	...	47	1	0	0	0	1	d scien	
1	1	Healthcare Data Scientist	63K-112K (Glassdoor est.)	What You Will Do:\n\nI. General Summary\n\nThe...	3.4	University of Maryland Medical System\n3.4	Linthicum, MD	Baltimore, MD	10000+ employees	1984	...	36	1	0	0	0	0	d scien	
2	2	Data Scientist	80K-90K (Glassdoor est.)	KnowBe4, Inc. is a high growth information sec...	4.8	KnowBe4\n4.8	Clearwater, FL	Clearwater, FL	501 to 1000 employees	2010	...	10	1	0	1	0	1	d scien	
3	3	Data Scientist	56K-97K (Glassdoor est.)	*Organization and Job ID**\nJob ID: 310709\n\n...	3.8	PNNL\n3.8	Richland, WA	Richland, WA	1001 to 5000 employees	1965	...	55	1	0	0	0	0	d scien	
4	4	Data Scientist	86K-143K (Glassdoor est.)	Data Scientist\nAffinity Solutions / Marketing...	2.9	Affinity Solutions\n2.9	New York, NY	New York, NY	51 to 200 employees	1998	...	22	1	0	0	0	1	d scien	
...	
737	737	Sr Scientist, Immuno-Oncology - Oncology	58K-111K (Glassdoor est.)	Site Name: USA - Massachusetts - Cambridge\nPo...	3.9	GSK\n3.9	Cambridge, MA	Brentford, United Kingdom	10000+ employees	1830	...	190	0	0	0	0	1	0	
738	738	Senior Data Engineer	72K-133K (Glassdoor est.)	THE CHALLENGE\nEventbrite has a world-class da...	4.4	Eventbrite\n4.4	Nashville, TN	San Francisco, CA	1001 to 5000 employees	2006	...	14	1	0	1	1	0	d engin	
739	739	Project Scientist - Auton Lab, Robotics Institute	56K-91K (Glassdoor est.)	The Auton Lab at Carnegie Mellon University is...	2.6	Software Engineering Institute\n2.6	Pittsburgh, PA	Pittsburgh, PA	501 to 1000 employees	1984	...	36	0	0	0	0	0	1	
740	740	Data Science Manager	95K-160K (Glassdoor est.)	Data Science ManagerResponsibilities:\n\nOvers...	3.2	Numeric, LLC\n3.2	Allentown, PA	Chadds Ford, PA	1 to 50 employees	-1	...	-1	0	0	0	0	1	mana	
741	741	Research Scientist - Security and Privacy	61K-126K (Glassdoor est.)	Returning Candidate? Log back in to the Career...	3.6	Riverside Research Institute\n3.6	Beavercreek, OH	Arlington, VA	501 to 1000 employees	1967	...	53	1	0	0	0	0	0	

742 rows x 33 columns

```
In [11]: d.duplicated().sum()
```

```
Out[11]: 0
```

```
In [4]: #for numpys.....
import numpy as np
from array import*
a= np.array([10,22,47,89,242,100,22,20,21,43])
print(a[1:6])
print(a[-5:-2])

ar= np.array([[22,24,35,26,18],[29,55,13,17,16],[4,22,30,46,64]])
print(ar.shape)

for i in range(0,3):
    for j in range(0,5):
        print(ar[i][j], end=" ")
        print()
#.....split the array.....
print(np.array_split(a, 4))

[ 22  47  89 242 100]
[100  22  20]
(3, 5)
22  24  35  26  18
29  55  13  17  16
4  22  30  46  64
[array([10, 22, 47]), array([ 89, 242, 100]), array([22, 20]), array([21, 43])]
```

```
In [ ]:
```