In [1]:
```
import pandas as pd
df = pd.read_csv("D:\\bizschoolpython\\Sample.csv")
df
```

Out[1]:

	First Name	Gender	Start Date	Last Login Time	Salary	Bonus %	Senior Management	Team
0	Douglas	Male	08-06-1993	12:42 PM	97308.0	6.945	True	Marketing
1	Thomas	Male	3/31/1996	6:53 AM	61933.0	4.170	True	NaN
2	Maria	Female	4/23/1993	11:17 AM	130590.0	11.858	False	Finance
3	Jerry	Male	03-04-2005	1:00 PM	NaN	9.340	True	Finance
4	Larry	Male	1/24/1998	4:47 PM	101004.0	1.389	True	Client Services
5	Dennis	Male	4/18/1987	1:35 AM	115163.0	10.125	False	Legal
6	Ruby	Female	8/17/1987	4:20 PM	65476.0	10.012	True	Product
7	NaN	Female	7/20/2015	10:43 AM	45906.0	11.598	NaN	Finance
8	Angela	Female	11/22/2005	6:29 AM	NaN	18.523	True	Engineering
9	Frances	Female	08-08-2002	6:51 AM	139852.0	7.524	True	Business Development
10	Louise	Female	08-12-1980	9:01 AM	63241.0	15.132	True	NaN
11	Julie	Female	10/26/1997	3:19 PM	102508.0	12.637	True	Legal
12	Brandon	Male	12-01-1980	1:08 AM	112807.0	17.492	True	Human Resources
13	Gary	Male	1/27/2008	11:40 PM	109831.0	5.831	False	Sales
14	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance
15	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product
16	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources
17	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product
18	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services
19	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product
20	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance
21	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product
22	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product
24	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services
25	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product

In [2]:
```
df.head(6)
```

Out[2]:

	First Name	Gender	Start Date	Last Login Time	Salary	Bonus %	Senior Management	Team
0	Douglas	Male	08-06-1993	12:42 PM	97308.0	6.945	True	Marketing
1	Thomas	Male	3/31/1996	6:53 AM	61933.0	4.170	True	NaN
2	Maria	Female	4/23/1993	11:17 AM	130590.0	11.858	False	Finance
3	Jerry	Male	03-04-2005	1:00 PM	NaN	9.340	True	Finance
4	Larry	Male	1/24/1998	4:47 PM	101004.0	1.389	True	Client Services
5	Dennis	Male	4/18/1987	1:35 AM	115163.0	10.125	False	Legal

In [3]:
```
df.tail()
```

Out[3]:

	First Name	Gender	Start Date	Last Login Time	Salary	Bonus %	Senior Management	Team
21	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product
22	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product
24	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services
25	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product

In [4]:
```
df.columns
```

Out[4]:
```
Index(['First Name', 'Gender', 'Start Date', 'Last Login Time', 'Salary',
       'Bonus %', 'Senior Management', 'Team'],
      dtype='object')
```

In [5]:
```
df[["First Name","Gender","Team"]]
```

Out[5]:

	First Name	Gender	Team
0	Douglas	Male	Marketing
1	Thomas	Male	NaN
2	Maria	Female	Finance
3	Jerry	Male	Finance
4	Larry	Male	Client Services
5	Dennis	Male	Legal
6	Ruby	Female	Product
7	NaN	Female	Finance
8	Angela	Female	Engineering
9	Frances	Female	Business Development
10	Louise	Female	NaN
11	Julie	Female	Legal
12	Brandon	Male	Human Resources
13	Gary	Male	Sales
14	Kimberly	Female	Finance
15	Lillian	Female	Product
16	Jeremy	Male	Human Resources
17	Shawn	Male	Product
18	Diana	Female	Client Services
19	Donna	Female	Product
20	Kimberly	Female	Finance
21	Lillian	Female	Product
22	Jeremy	Male	Human Resources
23	Shawn	Male	Product
24	Diana	Female	Client Services
25	Donna	Female	Product

In [6]:
```
df = df.rename(columns={"First Name" : "Name", "Bonus %":"Bonus"})
```

In [7]:
```
df
```

Out[7]:

	Name	Gender	Start Date	Last Login Time	Salary	Bonus	Senior Management	Team
0	Douglas	Male	08-06-1993	12:42 PM	97308.0	6.945	True	Marketing
1	Thomas	Male	3/31/1996	6:53 AM	61933.0	4.170	True	NaN
2	Maria	Female	4/23/1993	11:17 AM	130590.0	11.858	False	Finance
3	Jerry	Male	03-04-2005	1:00 PM	NaN	9.340	True	Finance
4	Larry	Male	1/24/1998	4:47 PM	101004.0	1.389	True	Client Services
5	Dennis	Male	4/18/1987	1:35 AM	115163.0	10.125	False	Legal
6	Ruby	Female	8/17/1987	4:20 PM	65476.0	10.012	True	Product
7	NaN	Female	7/20/2015	10:43 AM	45906.0	11.598	NaN	Finance
8	Angela	Female	11/22/2005	6:29 AM	NaN	18.523	True	Engineering
9	Frances	Female	08-08-2002	6:51 AM	139852.0	7.524	True	Business Development
10	Louise	Female	08-12-1980	9:01 AM	63241.0	15.132	True	NaN
11	Julie	Female	10/26/1997	3:19 PM	102508.0	12.637	True	Legal
12	Brandon	Male	12-01-1980	1:08 AM	112807.0	17.492	True	Human Resources
13	Gary	Male	1/27/2008	11:40 PM	109831.0	5.831	False	Sales
14	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance
15	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product
16	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources
17	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product
18	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services
19	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product
20	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance
21	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product
22	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product
24	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services
25	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product

In [8]:
```
df["BonusAmt"] = (df["Bonus"] * df["Salary"])/100
```

In [9]:
```
df
```

Out[9]:

	Name	Gender	Start Date	Last Login Time	Salary	Bonus	Senior Management	Team	BonusAmt
0	Douglas	Male	08-06-1993	12:42 PM	97308.0	6.945	True	Marketing	6758.04060
1	Thomas	Male	3/31/1996	6:53 AM	61933.0	4.170	True	NaN	2582.60610
2	Maria	Female	4/23/1993	11:17 AM	130590.0	11.858	False	Finance	15485.36220
3	Jerry	Male	03-04-2005	1:00 PM	NaN	9.340	True	Finance	NaN
4	Larry	Male	1/24/1998	4:47 PM	101004.0	1.389	True	Client Services	1402.94556
5	Dennis	Male	4/18/1987	1:35 AM	115163.0	10.125	False	Legal	11660.25375
6	Ruby	Female	8/17/1987	4:20 PM	65476.0	10.012	True	Product	6555.45712
7	NaN	Female	7/20/2015	10:43 AM	45906.0	11.598	NaN	Finance	5324.17788
8	Angela	Female	11/22/2005	6:29 AM	NaN	18.523	True	Engineering	NaN
9	Frances	Female	08-08-2002	6:51 AM	139852.0	7.524	True	Business Development	10522.46448
10	Louise	Female	08-12-1980	9:01 AM	63241.0	15.132	True	NaN	9569.62812
11	Julie	Female	10/26/1997	3:19 PM	102508.0	12.637	True	Legal	12953.93596
12	Brandon	Male	12-01-1980	1:08 AM	112807.0	17.492	True	Human Resources	19732.20044
13	Gary	Male	1/27/2008	11:40 PM	109831.0	5.831	False	Sales	6404.24561
14	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance	NaN
15	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product	746.23984
16	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources	6659.36530
17	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118
18	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services	25367.61080
19	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product	1534.40516
20	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance	NaN
21	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product	746.23984
22	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources	6659.36530
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118
24	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services	25367.61080
25	Donna	Female	7/22/2010	3:48 AM	81014.0	1.894	False	Product	1534.40516

In [10]:
```
df[(df["Gender"] == "Male") & (df["Team"]=="Product")]
```

Out[10]:

	Name	Gender	Start Date	Last Login Time	Salary	Bonus	Senior Management	Team	BonusAmt
17	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118

In [16]:
```
#Get all female --columns-- Name and Team
df.loc[(df["Gender"] == "Female"),['Name','Team']]
```

Out[16]:

	Name	Team
2	Maria	Finance
6	Ruby	Product
7	NaN	Finance
8	Angela	Engineering
9	Frances	Business Development
10	Louise	NaN
11	Julie	Legal
14	Kimberly	Finance
15	Lillian	Product
18	Diana	Client Services
19	Donna	Product
20	Kimberly	Finance
21	Lillian	Product
24	Diana	Client Services
25	Donna	Product

In [19]:
```
#Get Name and Bonus Amount of a Male employee with min BonusAmt
df[(df.Gender == 'Male')][['Name','BonusAmt']].min()
```

Out[19]:
```
Name         Brandon
BonusAmt    1402.94556
dtype: object
```

In [20]:
```
#Group By Gender - Average Salary
df.groupby('Gender')['Salary'].mean()
```

Out[20]:
```
Gender
Female     91192.416667
Male      100226.000000
Name: Salary, dtype: float64
```

In [21]:
```
df.drop('Last Login Time', axis=1)
```

Out[21]:

	Name	Gender	Start Date	Salary	Bonus	Senior Management	Team	BonusAmt
0	Douglas	Male	08-06-1993	97308.0	6.945	True	Marketing	6758.04060
1	Thomas	Male	3/31/1996	61933.0	4.170	True	NaN	2582.60610
2	Maria	Female	4/23/1993	130590.0	11.858	False	Finance	15485.36220
3	Jerry	Male	03-04-2005	NaN	9.340	True	Finance	NaN
4	Larry	Male	1/24/1998	101004.0	1.389	True	Client Services	1402.94556
5	Dennis	Male	4/18/1987	115163.0	10.125	False	Legal	11660.25375
6	Ruby	Female	8/17/1987	65476.0	10.012	True	Product	6555.45712
7	NaN	Female	7/20/2015	45906.0	11.598	NaN	Finance	5324.17788
8	Angela	Female	11/22/2005	NaN	18.523	True	Engineering	NaN
9	Frances	Female	08-08-2002	139852.0	7.524	True	Business Development	10522.46448
10	Louise	Female	08-12-1980	63241.0	15.132	True	NaN	9569.62812
11	Julie	Female	10/26/1997	102508.0	12.637	True	Legal	12953.93596
12	Brandon	Male	12-01-1980	112807.0	17.492	True	Human Resources	19732.20044
13	Gary	Male	1/27/2008	109831.0	5.831	False	Sales	6404.24561
14	Kimberly	Female	1/14/1999	NaN	14.543	True	Finance	NaN
15	Lillian	Female	06-05-2016	59414.0	1.256	False	Product	746.23984
16	Jeremy	Male	9/21/2010	90370.0	7.369	False	Human Resources	6659.36530
17	Shawn	Male	12-07-1986	111737.0	6.414	False	Product	7166.81118
18	Diana	Female	10/23/1981	132940.0	19.082	False	Client Services	25367.61080
19	Donna	Female	7/22/2010	81014.0	1.894	False	Product	1534.40516
20	Kimberly	Female	1/14/1999	NaN	14.543	True	Finance	NaN
21	Lillian	Female	06-05-2016	59414.0	1.256	False	Product	746.23984
22	Jeremy	Male	9/21/2010	90370.0	7.369	False	Human Resources	6659.36530
23	Shawn	Male	12-07-1986	111737.0	6.414	False	Product	7166.81118
24	Diana	Female	10/23/1981	132940.0	19.082	False	Client Services	25367.61080
25	Donna	Female	7/22/2010	81014.0	1.894	False	Product	1534.40516

In [22]:
```
df = df.loc[df['Name']!='Donna']
```

In [23]:
```
df
```

Out[23]:

	Name	Gender	Start Date	Last Login Time	Salary	Bonus	Senior Management	Team	BonusAmt
0	Douglas	Male	08-06-1993	12:42 PM	97308.0	6.945	True	Marketing	6758.04060
1	Thomas	Male	3/31/1996	6:53 AM	61933.0	4.170	True	NaN	2582.60610
2	Maria	Female	4/23/1993	11:17 AM	130590.0	11.858	False	Finance	15485.36220
3	Jerry	Male	03-04-2005	1:00 PM	NaN	9.340	True	Finance	NaN
4	Larry	Male	1/24/1998	4:47 PM	101004.0	1.389	True	Client Services	1402.94556
5	Dennis	Male	4/18/1987	1:35 AM	115163.0	10.125	False	Legal	11660.25375
6	Ruby	Female	8/17/1987	4:20 PM	65476.0	10.012	True	Product	6555.45712
7	NaN	Female	7/20/2015	10:43 AM	45906.0	11.598	NaN	Finance	5324.17788
8	Angela	Female	11/22/2005	6:29 AM	NaN	18.523	True	Engineering	NaN
9	Frances	Female	08-08-2002	6:51 AM	139852.0	7.524	True	Business Development	10522.46448
10	Louise	Female	08-12-1980	9:01 AM	63241.0	15.132	True	NaN	9569.62812
11	Julie	Female	10/26/1997	3:19 PM	102508.0	12.637	True	Legal	12953.93596
12	Brandon	Male	12-01-1980	1:08 AM	112807.0	17.492	True	Human Resources	19732.20044
13	Gary	Male	1/27/2008	11:40 PM	109831.0	5.831	False	Sales	6404.24561
14	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance	NaN
15	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product	746.23984
16	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources	6659.36530
17	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118
18	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services	25367.61080
20	Kimberly	Female	1/14/1999	7:13 AM	NaN	14.543	True	Finance	NaN
21	Lillian	Female	06-05-2016	6:09 AM	59414.0	1.256	False	Product	746.23984
22	Jeremy	Male	9/21/2010	5:56 AM	90370.0	7.369	False	Human Resources	6659.36530
23	Shawn	Male	12-07-1986	7:45 PM	111737.0	6.414	False	Product	7166.81118
24	Diana	Female	10/23/1981	10:27 AM	132940.0	19.082	False	Client Services	25367.61080

In [26]:
```
pivot_table = df.pivot_table(index = 'Gender',columns=['Team'],aggfunc = 'count')
pivot_table
```

Out[26]:

	Bonus		BonusAmt	...	Senior Management		
Team	Business Development	Client Services	Engineering	Finance	Human Resources	Legal	Marketing	Product	Sales	Business Development	...	Sales	Bus Develop
Gender													
Female	1.0	2.0	1.0	4.0	NaN	1.0	NaN	3.0	NaN	1.0	...	NaN	
Male	NaN	1.0	NaN	1.0	3.0	1.0	1.0	2.0	1.0	NaN	...	1.0	

2 rows × 63 columns

In [28]:
```
pivot_table.to_excel("D:\\bizschoolpython\\Pivot_Table_Team_Gender.xlsx")
```

In [ ]:
```
```