pandas-notebook-practice-2

April 17, 2024

1 Pandas Practice Questions

2 Pandas Notebook

[2]:
```
import pandas as pd
```

2.0.1 31. Read the CSV file.

[3]:
```
file_name = "employee_dataset.csv"

df = pd.read_csv(file_name)
```

2.0.2 32. Print the first 5 rows.

[ ]:
```
df_temp = df.head(5)

df_temp
```

[ ]:
```
              Name                  Company_Name  \
0  Spencer Adkins                James and Sons
1    Julie Morton                 Nichols-James
2    Matthew Hall                     Scott Inc
3      Brad Scott  Johnston, Fleming and Tanner
4   Theresa Owens      Baker, Allen and Edwards

                        Employee_Job_Title      Employee_City  \
0                        Equities trader     New Russellton
1  Diplomatic Services operational officer  North Melissafurt
2               Regulatory affairs officer           Wardfort
3                      Production engineer     West Jamesview
4                      Production engineer          Whiteside

         Employee_Country  Employee_Salary Employment_Status  Employee_Rating
0  Palestinian Territory           321520         Full Time             3.9
1       Marshall Islands           589090         Full Time             4.3
2               Anguilla           630890         Full Time             3.1
3   Syrian Arab Republic           116400         Full Time             3.1
```

```
4              Dominica          523499         Full Time              4.8
```

2.0.3 33. Print the last 10 rows.

[ ]:
```
df_temp = df.tail(10)
df_temp
```

[ ]:
```
                    Name                  Company_Name  \
299990     Monica Bender    Wallace, Smith and Shepard
299991  William Rodriguez                     Scott Inc
299992    Steven Thornton      Baker, Allen and Edwards
299993         Terry Hill       White, Mcclain and Cobb
299994      Kelly Bennett       White, Mcclain and Cobb
299995         Nancy Neal              Bullock-Carrillo
299996     Michele Butler  Johnston, Fleming and Tanner
299997        Lynn Wilson                Nichols-James
299998      Lindsey Keith                Nichols-James
299999      Karen Delgado                James and Sons

                 Employee_Job_Title      Employee_City Employee_Country  \
299990           Production engineer   New Cindychester       Bangladesh
299991                    Ergonomist          Whiteside      New Zealand
299992                Patent examiner         Aliciafort            Benin
299993  Investment banker, corporate   New Cindychester           Sweden
299994     Radiographer, therapeutic           Wardfort             Fiji
299995                   Optometrist  North Melissafurt            Samoa
299996                Administrator         Aliciafort             Cuba
299997                Administrator   New Cindychester          Bolivia
299998                         Make       Whitakerbury   Western Sahara
299999                      Actuary     New Russellton   United Kingdom

        Employee_Salary Employment_Status  Employee_Rating
299990           852660         Full Time             3.0
299991           214400         Full Time             1.5
299992            60240         Full Time             4.3
299993           414720            Intern             2.3
299994           182270         Full Time             3.6
299995            99670            Intern             3.3
299996           949580            Intern             2.6
299997           802830         Full Time             0.6
299998           257240         Full Time             2.4
299999           575770            Intern             3.3
```

2.0.4 34. Print the number of rows.

[ ]:
```
rows = df.shape[0]

print(rows)
```

```
300000
```

2.0.5 35. Print the number of columns.

[ ]:
```
columns = df.shape[1]

print(columns)
```

```
8
```

2.0.6 36. Print all column names.

[ ]:
```
column_names = df.columns

print(column_names)
```

```
Index(['Name', 'Company_Name', 'Employee_Job_Title', 'Employee_City',
       'Employee_Country', 'Employee_Salary', 'Employment_Status',
       'Employee_Rating'],
      dtype='object')
```

2.0.7 37. Print the mean of Employee_Salary.

[ ]:
```
df["Employee_Salary"].mean().round()
```

[ ]:
```
500224.0
```

2.0.8 38. Print the mean of Employee_Rating.

[ ]:
```
df["Employee_Rating"].mean().round(2)
```

[ ]:
```
2.51
```

2.0.9 39. Print the number of distinct Company_Name.

[ ]:
```
len(df["Company_Name"].unique())
```

[ ]:
```
15
```

2.0.10 40. Print all the distinct Company_Name.

[ ]:
```
df["Company_Name"].unique()
```

[ ]:
```
array(['James and Sons', 'Nichols-James', 'Scott Inc',
       'Johnston, Fleming and Tanner', 'Baker, Allen and Edwards',
       'Andrade LLC', 'Matthews Inc', 'Taylor-Ramos',
       'Campos, Reynolds and Mccormick', 'White, Mcclain and Cobb',
       'Thomas-Spencer', 'Bullock-Carrillo', 'Marshall-Holloway',
       'Wallace, Smith and Shepard', 'Nelson-Li'], dtype=object)
```

[ ]:
```
df.columns
```

[ ]:
```
Index(['Name', 'Company_Name', 'Employee_Job_Title', 'Employee_City',
       'Employee_Country', 'Employee_Salary', 'Employment_Status',
       'Employee_Rating'],
      dtype='object')
```

[ ]:
```
df
```

[ ]:

	Name	Company_Name	\
0	Spencer Adkins	James and Sons	
1	Julie Morton	Nichols-James	
2	Matthew Hall	Scott Inc	
3	Brad Scott	Johnston, Fleming and Tanner	
4	Theresa Owens	Baker, Allen and Edwards	
…	…	…	
299995	Nancy Neal	Bullock-Carrillo	
299996	Michele Butler	Johnston, Fleming and Tanner	
299997	Lynn Wilson	Nichols-James	
299998	Lindsey Keith	Nichols-James	
299999	Karen Delgado	James and Sons	

	Employee_Job_Title	Employee_City	\
0	Equities trader	New Russellton	
1	Diplomatic Services operational officer	North Melissafurt	
2	Regulatory affairs officer	Wardfort	
3	Production engineer	West Jamesview	
4	Production engineer	Whiteside	
…	…	…	
299995	Optometrist	North Melissafurt	
299996	Administrator	Aliciafort	
299997	Administrator	New Cindychester	
299998	Make	Whitakerbury	
299999	Actuary	New Russellton	

	Employee_Country	Employee_Salary	Employment_Status	\

```
0       Palestinian Territory           321520         Full Time
1            Marshall Islands           589090         Full Time
2                    Anguilla           630890         Full Time
3         Syrian Arab Republic           116400         Full Time
4                    Dominica           523499         Full Time
…                         …              …              …
299995                  Samoa            99670            Intern
299996                   Cuba           949580            Intern
299997                Bolivia           802830         Full Time
299998         Western Sahara           257240         Full Time
299999         United Kingdom           575770            Intern

        Employee_Rating
0                   3.9
1                   4.3
2                   3.1
3                   3.1
4                   4.8
…                   …
299995              3.3
299996              2.6
299997              0.6
299998              2.4
299999              3.3

[300000 rows x 8 columns]
```

2.0.11 41. Print the number of employees working in the company "Nichols-James".

[ ]:
```
df["Company_Name"].value_counts()["Nichols-James"]
```

[ ]:
```
19911
```

2.0.12 42-44. Print the maximum, minimum and median Employee_Salary.

[ ]:
```
median=df["Employee_Salary"].median()
min=df["Employee_Salary"].min()
max=df["Employee_Salary"].max()
print("median:",median)
print("min:",min)
print("max:",max)
```

```
median: 500780.0
min: 0
max: 999990
```

2.0.13 45-49. Print the distribution of the following columns: (the frequency of individual entries).

1. Company_Name

```
[ ]: df["Company_Name"]
```

```
[ ]: 0                        James and Sons
     1                         Nichols-James
     2                             Scott Inc
     3           Johnston, Fleming and Tanner
     4               Baker, Allen and Edwards
                            …
     299995                  Bullock-Carrillo
     299996      Johnston, Fleming and Tanner
     299997                    Nichols-James
     299998                    Nichols-James
     299999                   James and Sons
     Name: Company_Name, Length: 300000, dtype: object
```

2. Employee_Job_Title

```
[ ]: df["Employee_Job_Title"]
```

```
[ ]: 0                                  Equities trader
     1           Diplomatic Services operational officer
     2                        Regulatory affairs officer
     3                               Production engineer
     4                               Production engineer
                                 …
     299995                                  Optometrist
     299996                                Administrator
     299997                                Administrator
     299998                                        Make
     299999                                     Actuary
     Name: Employee_Job_Title, Length: 300000, dtype: object
```

3. Employee_City

```
[ ]: df["Employee_City"]
```

```
[ ]: 0              New Russellton
     1           North Melissafurt
     2                    Wardfort
     3              West Jamesview
     4                   Whiteside
                     …
     299995      North Melissafurt
     299996             Aliciafort
```

```
299997     New Cindychester
299998         Whitakerbury
299999       New Russellton
Name: Employee_City, Length: 300000, dtype: object
```

4. Employee_Country

[ ]:
```
df["Employee_Country"]
```

[ ]:
```
0         Palestinian Territory
1              Marshall Islands
2                      Anguilla
3           Syrian Arab Republic
4                      Dominica
                   …
299995                    Samoa
299996                     Cuba
299997                  Bolivia
299998           Western Sahara
299999           United Kingdom
Name: Employee_Country, Length: 300000, dtype: object
```

5. Employment_Status

[ ]:
```
df["Employment_Status"]
```

[ ]:
```
0         Full Time
1         Full Time
2         Full Time
3         Full Time
4         Full Time
             …
299995       Intern
299996       Intern
299997    Full Time
299998    Full Time
299999       Intern
Name: Employment_Status, Length: 300000, dtype: object
```

2.0.14 50. Print the company with the most number of employees.

[ ]:
```
df["Company_Name"].value_counts().nlargest(1)
```

[ ]:
```
Company_Name
Scott Inc    20390
Name: count, dtype: int64
```

2.0.15 51. Print the number of employees in the above company.

[ ]:
```
df["Company_Name"].value_counts()["Scott Inc"]
```

[ ]:
```
20390
```

2.0.16 52. Print the company with the least number of employees.

[ ]:
```
df["Company_Name"].value_counts().nsmallest(1)
```

[ ]:
```
Company_Name
Wallace, Smith and Shepard    19364
Name: count, dtype: int64
```

2.0.17 53. Print the number of employees in the above company.

[ ]:
```
df["Company_Name"].value_counts()["Wallace, Smith and Shepard"]
```

[ ]:
```
19364
```

2.0.18 54. Print the employee details with the maximum salary

[ ]:
```
df.sort_values(by="Employee_Salary",ascending=False).drop_duplicates().head(1)
```

[ ]:
```
              Name  Company_Name   Employee_Job_Title Employee_City  \
70356  Anna Lawson  Taylor-Ramos  Production engineer   Kristaburgh

      Employee_Country  Employee_Salary Employment_Status  Employee_Rating
70356          Lesotho           999990         Full Time              4.0
```

2.0.19 55. Print the employee details with the maximum rating

[ ]:
```
df.nlargest(1,"Employee_Rating")
```

[ ]:
```
               Name    Company_Name            Employee_Job_Title  \
124  Joshua Huffman  James and Sons  Garment/textile technologist

      Employee_City Employee_Country  Employee_Salary Employment_Status  \
124  New Russellton     Cook Islands           694760         Full Time

     Employee_Rating
124              5.0
```

2.0.20 56. Print the Company_Name with most number of employees in 'Wardfort' city.

[ ]:
```
#Filter the DataFrame to include only employees in 'Wardfort' city
wardfort_employees = df[df['Employee_City'] == 'Wardfort']

#Group the employees by 'Company_Name' and count the number of employees in␣
 ↪each company
company_employee_count = wardfort_employees['Company_Name'].value_counts()

# Get the company with the most number of employees
company_with_most_employees = company_employee_count.idxmax()

print("The company with the most number of employees in Wardfort is:",␣
 ↪company_with_most_employees)
```

```
The company with the most number of employees in Wardfort is: White, Mcclain and
Cobb
```

[ ]:
```
df.columns
```

[ ]:
```
Index(['Name', 'Company_Name', 'Employee_Job_Title', 'Employee_City',
       'Employee_Country', 'Employee_Salary', 'Employment_Status',
       'Employee_Rating'],
      dtype='object')
```

2.0.21 57. Change the Data type of 'Employee_Salary' column to float.

[ ]:
```
df["Employee_Salary"].astype(float)
```

[ ]:
```
0         321520.0
1         589090.0
2         630890.0
3         116400.0
4         523499.0
             …
299995     99670.0
299996    949580.0
299997    802830.0
299998    257240.0
299999    575770.0
Name: Employee_Salary, Length: 300000, dtype: float64
```

2.0.22 58. Print the Employee_City with the most number of 'Production engineer'.

[ ]:
```
#df[df["Employee_Job_Title"]=='Production engineer']["Employee_City"].
 ↪value_counts().sort_values(ascending=False)
df[df["Employee_Job_Title"]=='Production engineer']["Employee_City"].
 ↪value_counts().idxmax()
```

[ ]:
```
'Aliciafort'
```

[ ]:
```
df
```

[ ]:

	Name	Company_Name	Employee_Job_Title	Employee_City	Employee_Country	Employee_Salary	Employment_Status
0	Spencer Adkins	James and Sons	Equities trader	New Russellton	Palestinian Territory	321520	Full Time
1	Julie Morton	Nichols-James	Diplomatic Services operational officer	North Melissafurt	Marshall Islands	589090	Full Time
2	Matthew Hall	Scott Inc	Regulatory affairs officer	Wardfort	Anguilla	630890	Full Time
3	Brad Scott	Johnston, Fleming and Tanner	Production engineer	West Jamesview	Syrian Arab Republic	116400	Full Time
4	Theresa Owens	Baker, Allen and Edwards	Production engineer	Whiteside	Dominica	523499	Full Time
…	…	…	…	…	…	…	…
299995	Nancy Neal	Bullock-Carrillo	Optometrist	North Melissafurt	Samoa	99670	Intern
299996	Michele Butler	Johnston, Fleming and Tanner	Administrator	Aliciafort	Cuba	949580	Intern
299997	Lynn Wilson	Nichols-James	Administrator	New Cindychester	Bolivia	802830	Full Time
299998	Lindsey Keith	Nichols-James	Make	Whitakerbury			
299999	Karen Delgado	James and Sons	Actuary	New Russellton			

```
299998         Western Sahara           257240         Full Time
299999         United Kingdom           575770            Intern

        Employee_Rating
0                   3.9
1                   4.3
2                   3.1
3                   3.1
4                   4.8
…                   …
299995              3.3
299996              2.6
299997              0.6
299998              2.4
299999              3.3

[300000 rows x 8 columns]
```

2.0.23 59. Print the Company_Name with the most number of Full-time Employees.

[ ]:
```
df[df["Employment_Status"]=="Full Time"][["Company_Name","Employment_Status"]].
↪value_counts().sort_values(ascending=False).head(1)
```

[ ]:
```
Company_Name  Employment_Status
Scott Inc     Full Time           16396
Name: count, dtype: int64
```

2.0.24 60. Print the Company_Name with the highest average 'Employee_Rating'.

[ ]:
```
df_60=df[["Company_Name","Employee_Rating"]].groupby("Company_Name").mean().
↪round(2).reset_index()
df_60.nlargest(1,"Employee_Rating")
```

[ ]:
```
   Company_Name  Employee_Rating
10    Scott Inc             2.53
```

2.0.25 61. Print the number of employees working in 'Ricardomouth' and 'Kristaburgh' location combined.

[ ]:
```
df["Employee_City"].value_counts()[["Ricardomouth","Kristaburgh"]].sum()
```

[ ]:
```
35355
```

2.0.26 62. Print the distinct Company_Name corresponding to the 5 highest paid employees in the dataset.

```
[ ]: top_5_employees = df[["Company_Name", "Employee_Salary"]].
      ↪sort_values(by="Employee_Salary", ascending=False).head(6)
     distinct_company_names = top_5_employees["Company_Name"].unique()

     print(distinct_company_names)
```

```
['Taylor-Ramos' 'Thomas-Spencer' 'White, Mcclain and Cobb'
 'Campos, Reynolds and Mccormick' 'Wallace, Smith and Shepard']
```

2.0.27 63. Check if there are any duplicate rows in the DataFrame.

```
[ ]: duplicated=df.duplicated()
     if duplicated.any():
       print("There are duplicate rows")
     else:
       print("No duplicates found!")
```

```
No duplicates found!
```

2.0.28 64. Check if any of the columns has NaN values.

```
[ ]: df.isna().any()
```

```
[ ]: Name                 False
     Company_Name         False
     Employee_Job_Title   False
     Employee_City        False
     Employee_Country     False
     Employee_Salary      False
     Employment_Status    False
     Employee_Rating       True
     dtype: bool
```

2.0.29 65. Print the data type of every column in the DataFrame.

```
[ ]: df.dtypes
```

```
[ ]: Name                 object
     Company_Name         object
     Employee_Job_Title   object
     Employee_City        object
     Employee_Country     object
     Employee_Salary       int64
     Employment_Status    object
```

```
Employee_Rating       float64
dtype: object
```

2.0.30 66. Print the Company_Name column only as a Series

[ ]:
```
df["Company_Name"]
```

[ ]:
```
0                       James and Sons
1                        Nichols-James
2                            Scott Inc
3         Johnston, Fleming and Tanner
4             Baker, Allen and Edwards
                        …
176038      Wallace, Smith and Shepard
176039                   Matthews Inc
176040                  Nichols-James
176041                   Matthews Inc
176042        Baker, Allen and Edwards
Name: Company_Name, Length: 176043, dtype: object
```

2.0.31 67. Print the Company_Name column only as a DataFrame

[ ]:
```
df[["Company_Name"]]
```

[ ]:
```
                        Company_Name
0                       James and Sons
1                        Nichols-James
2                            Scott Inc
3         Johnston, Fleming and Tanner
4             Baker, Allen and Edwards
…                                  …
176038      Wallace, Smith and Shepard
176039                   Matthews Inc
176040                  Nichols-James
176041                   Matthews Inc
176042        Baker, Allen and Edwards

[176043 rows x 1 columns]
```

2.0.32 68. Select the 'Employee_Job_Title' and 'Employee_City' column.

[ ]:
```
df[["Employee_Job_Title","Employee_City"]]
```

[ ]:
```
                              Employee_Job_Title      Employee_City
0                                Equities trader     New Russellton
1       Diplomatic Services operational officer  North Melissafurt
```

```
2                    Regulatory affairs officer           Wardfort
3                           Production engineer     West Jamesview
4                           Production engineer          Whiteside
…                                           …                  …
176038             Investment banker, corporate           Wardfort
176039                                     Make        Kristaburgh
176040                                  Actuary          Whiteside
176041  Diplomatic Services operational officer     West Jamesview
176042                            Naval architect     New Russellton

[176043 rows x 2 columns]
```

2.0.33 69. Print the number of employees with Employee_Rating greater than the average Employee_Rating

[ ]:
```
greater_than_average = df[df["Employee_Rating"] > df["Employee_Rating"].mean()]
num_employees = greater_than_average.shape[0]

print("The number of employees with an Employee Rating greater than the average␣
 ↪is:", num_employees)
```

```
The number of employees with an Employee Rating greater than the average is:
86528
```

2.0.34 70. Find the employee which has the maximum salary among the ones with the minimum Employee_Rating

[8]:
```
Lowest_Rating=df[df["Employee_Rating"]==df["Employee_Rating"].min()]
Max_Salary_in_Lowest_Rating=Lowest_Rating[Lowest_Rating["Employee_Salary"]==Lowest_Rating["Emp
 ↪max()]
Max_Salary_in_Lowest_Rating
```

[8]:
```
                Name             Company_Name Employee_Job_Title  \
240878  Austin Poole  White, Mcclain and Cobb      Administrator

       Employee_City       Employee_Country  Employee_Salary  \
240878     Whiteside  Saint Kitts and Nevis           999690

       Employment_Status  Employee_Rating
240878            Intern             0.0
```

2.0.35 71. Sort the DataFrame in ascending order of Employee_Salary

[ ]:
```
df.sort_values(by="Employee_Salary")
```

[ ]:
```
                        Name                    Company_Name  \
54531        Kristen Vasquez         White, Mcclain and Cobb
94582        Laurie Williams         White, Mcclain and Cobb
139347       Benjamin Flores         White, Mcclain and Cobb
72786   Christopher Townsend               Marshall-Holloway
119422           Amber Blake                  James and Sons
…                        …                               …
23536          Dennis Sexton  Campos, Reynolds and Mccormick
153993         Jason Jenkins                  Taylor-Ramos
51642        Richard Kennedy                Thomas-Spencer
50461          Logan Jackson         White, Mcclain and Cobb
70356            Anna Lawson                  Taylor-Ramos

                                      Employee_Job_Title   Employee_City  \
54531   Armed forces logistics/support/administrative …        Wardfort
94582                                        Optometrist       Whiteside
139347                                   Naval architect        Wardfort
72786                                  Retail merchandiser    Ricardomouth
119422                           Trading standards officer  New Russellton
…                                                      …               …
23536              Diplomatic Services operational officer     Kristaburgh
153993                   Sales promotion account executive      Aliciafort
51642                                      Energy manager      Aliciafort
50461                                          Ergonomist        Wardfort
70356                                  Production engineer     Kristaburgh

       Employee_Country  Employee_Salary Employment_Status  Employee_Rating
54531           Bermuda                0         Full Time              3.4
94582           Bermuda               20         Full Time              4.3
139347        Australia               30         Full Time              1.3
72786           Moldova               30         Full Time              1.5
119422          Belarus               30         Full Time              2.7
…                   …                …                 …              …
23536    United Kingdom           999970         Full Time              0.9
153993         Pakistan           999980         Full Time              2.6
51642       Cook Islands           999980         Full Time              2.4
50461       Saint Helena           999980         Full Time              2.0
70356            Lesotho           999990         Full Time              4.0

[176043 rows x 8 columns]
```

2.0.36 72. Sort the DataFrame in descending order of Employee_Rating

[ ]:
```
df.sort_values(by="Employee_Rating",ascending=False)
```

[ ]:
```
                    Name                    Company_Name  \
42227      Heather Allen         White, Mcclain and Cobb
```

```
123497  Stephanie Romero                    Matthews Inc
51301      Blake Clayton                    Matthews Inc
130351  Stephanie Fuller                  James and Sons
130363      Nathan Walls                Marshall-Holloway
…                  …                              …
142462        Kelsey Lee          White, Mcclain and Cobb
43090         Jenny Diaz   Campos, Reynolds and Mccormick
158050     Sandra Taylor                         Nelson-Li
99062        Kathy Hayes         Baker, Allen and Edwards
176042    Whitney Roberts         Baker, Allen and Edwards

                                      Employee_Job_Title      Employee_City  \
42227                                          Ergonomist     New Russellton
123497            Diplomatic Services operational officer     New Russellton
51301             Diplomatic Services operational officer       Ricardomouth
130351                 Sales promotion account executive        Kristaburgh
130363  Armed forces logistics/support/administrative …         Aliciafort
…                                                     …              …
142462                                  Equities trader        Kristaburgh
43090             Diplomatic Services operational officer   New Cindychester
158050                        Garment/textile technologist   New Cindychester
99062                         Investment banker, corporate  North Melissafurt
176042                                  Naval architect     New Russellton

       Employee_Country  Employee_Salary Employment_Status  Employee_Rating
42227           Nigeria            78340         Full Time             5.0
123497          Cyprus            122340         Full Time             5.0
51301           Malawi            508569         Full Time             5.0
130351       Seychelles           450080         Full Time             5.0
130363           Poland           111960         Full Time             5.0
…                   …                …                …               …
142462             Niue           952880         Full Time             0.0
43090             Korea           705400         Full Time             0.0
158050            Palau           113910            Intern             0.0
99062              Togo           766050            Intern             0.0
176042          Myanmar           210780               Ful             NaN

[176043 rows x 8 columns]
```

2.0.37 73. Print the name of 100th employee after sorting on Name

[ ]:
```
df.sort_values(by="Name").iloc[99]["Name"]
```

[ ]:
```
'Aaron Cook'
```

2.0.38 74. Print the first 5 rows of the first 5 columns.

```
[ ]: df.iloc[0:5,0:5]
```

```
[ ]:              Name                  Company_Name  \
     0  Spencer Adkins                James and Sons
     1    Julie Morton                 Nichols-James
     2    Matthew Hall                     Scott Inc
     3      Brad Scott  Johnston, Fleming and Tanner
     4   Theresa Owens      Baker, Allen and Edwards

                          Employee_Job_Title      Employee_City  \
     0                        Equities trader     New Russellton
     1  Diplomatic Services operational officer  North Melissafurt
     2               Regulatory affairs officer           Wardfort
     3                      Production engineer     West Jamesview
     4                      Production engineer          Whiteside

          Employee_Country
     0  Palestinian Territory
     1       Marshall Islands
     2                Anguilla
     3   Syrian Arab Republic
     4                Dominica
```

2.0.39 75. Print the transpose of the first 5 rows

```
[ ]: df.iloc[0:5,0:5].T
```

```
[ ]:                                      0  \
     Name                       Spencer Adkins
     Company_Name                James and Sons
     Employee_Job_Title         Equities trader
     Employee_City               New Russellton
     Employee_Country     Palestinian Territory

                                                       1  \
     Name                                     Julie Morton
     Company_Name                            Nichols-James
     Employee_Job_Title  Diplomatic Services operational officer
     Employee_City                       North Melissafurt
     Employee_Country                     Marshall Islands

                                          2                           3  \
     Name                       Matthew Hall                   Brad Scott
     Company_Name                  Scott Inc  Johnston, Fleming and Tanner
     Employee_Job_Title  Regulatory affairs officer          Production engineer
```

```
Employee_City                          Wardfort                West Jamesview
Employee_Country                       Anguilla          Syrian Arab Republic

                                              4
Name                             Theresa Owens
Company_Name          Baker, Allen and Edwards
Employee_Job_Title         Production engineer
Employee_City                       Whiteside
Employee_Country                     Dominica
```

2.0.40 76. Print the number of employees whose first name starts with the letter 'V'.

[22]:
```
df["Name"].str.startswith("V").value_counts()[1]
```

[22]:
```
3683
```

[23]:
```
#Way-2
count_v_employees = df[df['Name'].str.startswith('V')].shape[0]
count_v_employees
```

[23]:
```
3683
```

2.0.41 77. Print the number of employees whose last name starts with the letter 'R'.

[28]:
```
count_r_employees = df[df['Name'].str.startswith('R')].shape[0]
print("The number of employees whose name starts with the letter 'R' is:",␣
 ↪count_r_employees)
```

```
The number of employees whose name starts with the letter 'R' is: 17940
```

[34]:
```
df["Name"]
```

[34]:
```
0         Spencer Adkins
1           Julie Morton
2           Matthew Hall
3             Brad Scott
4          Theresa Owens
                ...
299995        Nancy Neal
299996    Michele Butler
299997       Lynn Wilson
299998     Lindsey Keith
299999     Karen Delgado
Name: Name, Length: 300000, dtype: object
```

[38]:
```
df[df["Name"].str.split().str[-1].str.startswith("R")].shape[0]
```

[38]: 20712

[39]:
```
#Way2
count_r_employees = df[df['Name'].str.split().str[-1].str.startswith('R')].
 ↪shape[0]
print("The number of employees whose last name starts with the letter 'R' is:",␣
 ↪count_r_employees)
```

```
The number of employees whose last name starts with the letter 'R' is: 20712
```

2.0.42 78. Select the rows 2 to 7 and the columns 3 to 7 (both included)

[ ]:
```
df.iloc[2:8,3:8]
```

[ ]:
```
    Employee_City      Employee_Country  Employee_Salary Employment_Status  \
2        Wardfort             Anguilla           630890         Full Time
3  West Jamesview  Syrian Arab Republic           116400         Full Time
4       Whiteside              Dominica           523499         Full Time
5    Ricardomouth                  Mali           850140         Full Time
6        Wardfort                 Aruba           711410         Full Time
7     Kristaburgh        Western Sahara           777000         Full Time

   Employee_Rating
2              3.1
3              3.1
4              4.8
5              2.6
6              2.0
7              3.3
```

2.0.43 79. Select every row after the 10th row and select all columns.

[ ]:
```
df.iloc[10:]
```

[ ]:
```
                    Name                    Company_Name  \
10        Victoria Sutton         White, Mcclain and Cobb
11        Timothy Johnson        Baker, Allen and Edwards
12         Tiffany Galvan  Campos, Reynolds and Mccormick
13           David Duran                 Nichols-James
14            Julie Cook                  Matthews Inc
…                    …                             …
176038    Candice Johnson      Wallace, Smith and Shepard
176039  Rebecca Hernandez                  Matthews Inc
176040       Mark Newman                 Nichols-James
176041    Robert Phillips                  Matthews Inc
176042    Whitney Roberts        Baker, Allen and Edwards
```

```
                           Employee_Job_Title   Employee_City  \
10                            Naval architect       Whiteside
11                 Regulatory affairs officer    Ricardomouth
12     Diplomatic Services operational officer  West Jamesview
13                Investment banker, corporate  New Russellton
14                            Equities trader  New Russellton
…                                         …             …
176038            Investment banker, corporate       Wardfort
176039                                   Make     Kristaburgh
176040                                Actuary       Whiteside
176041  Diplomatic Services operational officer  West Jamesview
176042                            Naval architect  New Russellton

                Employee_Country  Employee_Salary Employment_Status  \
10                        Poland           656260         Full Time
11                       Georgia           503610         Full Time
12                         Niger           786430            Intern
13                       Georgia           910210         Full Time
14                   Puerto Rico           328860         Full Time
…                            …                …                …
176038                     Gabon           415220            Intern
176039  Bouvet Island (Bouvetoya)           103630         Full Time
176040              Cook Islands           336040         Full Time
176041                Luxembourg           763770         Full Time
176042                   Myanmar           210780               Ful

        Employee_Rating
10                  4.9
11                  2.9
12                  3.6
13                  4.9
14                  4.8
…                   …
176038              4.7
176039              1.5
176040              4.4
176041              0.9
176042              NaN

[176033 rows x 8 columns]
```

2.0.44 80. Select every row up to the 10th row and select all columns.

[ ]:
```
df.iloc[:10]
```

[ ]:

	Name	Company_Name	Employee_Job_Title	Employee_City	Employee_Country	Employee_Salary	Employment_Status	Employee_Rating
0	Spencer Adkins	James and Sons	Equities trader	New Russellton	Palestinian Territory	321520	Full Time	3.9
1	Julie Morton	Nichols-James	Diplomatic Services operational officer	North Melissafurt	Marshall Islands	589090	Full Time	4.3
2	Matthew Hall	Scott Inc	Regulatory affairs officer	Wardfort	Anguilla	630890	Full Time	3.1
3	Brad Scott	Johnston, Fleming and Tanner	Production engineer	West Jamesview	Syrian Arab Republic	116400	Full Time	3.1
4	Theresa Owens	Baker, Allen and Edwards	Production engineer	Whiteside	Dominica	523499	Full Time	4.8
5	Vanessa Allen	Andrade LLC	Make	Ricardomouth	Mali	850140	Full Time	2.6
6	Kelly Brown	Andrade LLC	Naval architect	Wardfort	Aruba	711410	Full Time	2.0
7	Vicki Beard	Matthews Inc	Diplomatic Services operational officer	Kristaburgh	Western Sahara	777000	Full Time	3.3
8	Lisa French	Taylor-Ramos	Radiographer, therapeutic	West Jamesview	Kuwait	870580	Full Time	1.3
9	Michael Edwards	Campos, Reynolds and Mccormick	Ergonomist	New Cindychester	Lao People's Democratic Republic	516950	Full Time	2.3

2.0.45 81. Select rows with employee rating > 4.5.

[ ]:
```
df[df["Employee_Rating"]>4.5]
```

[ ]:
```
                    Name                    Company_Name  \
4          Theresa Owens        Baker, Allen and Edwards
10       Victoria Sutton         White, Mcclain and Cobb
13           David Duran                 Nichols-James
14            Julie Cook                  Matthews Inc
49       Gerald Santiago                 Thomas-Spencer
…                    …                             …
175999   Joshua Holloway  Campos, Reynolds and Mccormick
176006  Michael Williams    Johnston, Fleming and Tanner
176025    Dana Rodriguez    Johnston, Fleming and Tanner
176030       Alexis Hill                 James and Sons
176038   Candice Johnson      Wallace, Smith and Shepard

                  Employee_Job_Title     Employee_City  Employee_Country  \
4                Production engineer         Whiteside          Dominica
10                   Naval architect         Whiteside            Poland
13       Investment banker, corporate    New Russellton           Georgia
14                   Equities trader    New Russellton       Puerto Rico
49       Garment/textile technologist       Kristaburgh       El Salvador
…                                …               …                 …
175999                   Optometrist  New Cindychester      Sierra Leone
176006     Trading standards officer          Wardfort  French Polynesia
176025                Equities trader        Aliciafort          Djibouti
176030    Regulatory affairs officer      Whitakerbury        Mauritania
176038  Investment banker, corporate          Wardfort              Gabon

        Employee_Salary Employment_Status  Employee_Rating
4                523499         Full Time             4.8
10               656260         Full Time             4.9
13               910210         Full Time             4.9
14               328860         Full Time             4.8
49               873320         Full Time             4.7
…                   …                 …               …
175999           464350            Intern             4.6
176006           441620         Full Time             4.9
176025           974380         Full Time             4.8
176030           339810         Full Time             4.8
176038           415220            Intern             4.7

[15847 rows x 8 columns]
```

2.0.46 82. Select rows with employee rating > 4.5 and < 4.8.

```
[ ]: df[(df["Employee_Rating"]>4.5) & (df["Employee_Rating"]<4.8)]
```

```
[ ]:                    Name                   Company_Name  \
49         Gerald Santiago                 Thomas-Spencer
88           Julie Richard                Bullock-Carrillo
94          Natasha Harmon                    Matthews Inc
119          Jason Parrish                   Andrade LLC
129         Connie Coleman                   Andrade LLC
…                      …                             …
175886     Jeffrey Lindsey               Marshall-Holloway
175903  Whitney Jefferson                    Matthews Inc
175969          Wendy Ross     Wallace, Smith and Shepard
175999     Joshua Holloway  Campos, Reynolds and Mccormick
176038     Candice Johnson     Wallace, Smith and Shepard

                                      Employee_Job_Title     Employee_City  \
49                          Garment/textile technologist       Kristaburgh
88                                      Naval architect          Wardfort
94                         Radiographer, therapeutic    West Jamesview
119                         Garment/textile technologist        Aliciafort
129                                        Optometrist       Kristaburgh
…                                                  …               …
175886                     Radiographer, therapeutic        Aliciafort
175903  Armed forces logistics/support/administrative …      Ricardomouth
175969                                  Patent examiner      Whitakerbury
175999                                      Optometrist  New Cindychester
176038                      Investment banker, corporate          Wardfort

       Employee_Country  Employee_Salary Employment_Status  Employee_Rating
49          El Salvador           873320         Full Time             4.7
88              Burundi           222850         Full Time             4.6
94              Bolivia           790000         Full Time             4.7
119       Cote d'Ivoire           581130         Full Time             4.7
129           Indonesia           899260         Full Time             4.7
…                   …                …                 …               …
175886            Niger           865230         Full Time             4.6
175903           Jordan           784940         Full Time             4.7
175969    Guinea-Bissau           858610         Full Time             4.7
175999     Sierra Leone           464350            Intern             4.6
176038            Gabon           415220            Intern             4.7

[7016 rows x 8 columns]
```

2.0.47 83. Print the name of the company with the maximum employees having rating > 4.

[55]:
```
df_83=df[["Company_Name","Employee_Rating"]]
EmployeeWithAbove4Rating=df_83[df_83["Employee_Rating"]>4]
company_with_the_maximum_employees_having_rating_Above_4=EmployeeWithAbove4Rating.
 ↪groupby("Company_Name").count().
 ↪sort_values(by="Employee_Rating",ascending=False).head(1)
print("company_with_the_maximum_employees_having_rating_Above_4",company_with_the_maximum_empl
```

```
company_with_the_maximum_employees_having_rating_Above_4
Employee_Rating
Company_Name
Matthews Inc              3922
```

2.0.48 84. Print the rating of the employee named 'Julie Morton'.

[57]:
```
df[df["Name"]=="Julie Morton"]["Employee_Rating"]
```

[57]:
```
1    4.3
Name: Employee_Rating, dtype: float64
```

2.0.49 85. Print the last 3rd entry in the column Employee_City

[64]:
```
df.iloc[-4:-1]["Employee_City"]
```

[64]:
```
299996          Aliciafort
299997    New Cindychester
299998        Whitakerbury
Name: Employee_City, dtype: object
```

2.0.50 86. Print the Employee_City column as a list

[77]:
```
Employee_City=list(df["Employee_City"].values)
print(Employee_City[1:10])
type(Employee_City)
```

```
['North Melissafurt', 'Wardfort', 'West Jamesview', 'Whiteside', 'Ricardomouth',
'Wardfort', 'Kristaburgh', 'West Jamesview', 'New Cindychester']
```

[77]:
```
list
```

2.0.51 87. Are the number of employees in 'Scott Inc' company greater than that in 'Andrade LLC'? Print True/False

[88]:
```
Employee_in_Scott_Inc=df['Company_Name'].value_counts()["Scott Inc"]
Employee_in_Andrade_LLC=df['Company_Name'].value_counts()["Andrade LLC"]
Employee_in_Scott_Inc>Employee_in_Andrade_LLC
```

[88]:
```
True
```

2.0.52 88. Which is the most common first name in the dataframe?

[104]:
```
from collections import Counter
Name_list=list(df['Name'].str.split())

#Extract first names from the "Name" column
first_names = [name[0] for name in Name_list]

# Count the occurrences of each first name
name_counts = Counter(first_names)

#Find the most common first name
most_common_name = name_counts.most_common(1)[0][0]
repeatition=name_counts.most_common(1)[0][1]
print("The most common first name in the dataframe is:", most_common_name,", It␣
 ↪occured",repeatition,"Times")
```

```
The most common first name in the dataframe is: Michael , It occured 6892 Times
```

2.0.53 89. Which is the least common last name in the dataframe?

[106]:
```
from collections import Counter
Name_list=list(df['Name'].str.split())

#Extract first names from the "Name" column
last_names = [name[1] for name in Name_list]

# Count the occurrences of each first name
name_counts = Counter(last_names)

#Find the most common first name
most_common_name = name_counts.most_common(1)[0][0]
repeatition=name_counts.most_common(1)[0][1]
print("The most common last name in the dataframe is:", most_common_name,", It␣
 ↪occured",repeatition,"Times")
```

```
The most common last name in the dataframe is: Smith , It occured 6510 Times
```

2.0.54 90. What is the average name length?

[117]:
```
import numpy as np
Name_list=list(df["Name"].values)
length_list=[]
for i in Name_list:
  length_list.append(len(i))
average_length=np.mean(length_list)
print("Average Name length is:",average_length)
```

```
Average Name length is: 13.1079
```

[ ]:
```
#Done Till Now...Follow "Govind Kumar" over linkedIn for more...
```