Problem 1: Import Tips dataset and store as the pandas dataframe with the name tips.

```
import pandas as pd

tips =
pd.read_csv("https://github.com/YBI-Foundation/Dataset/raw/main/Tips
%20Payment%20Data.csv")
```

Problem 2: Display the first 5 rows of the tips dataframe.

```
tips.head()
```

|   | Total Bill | Tip | Gender | Smoker | Day | Time | Size | Bill Per Person |
|---|---|---|---|---|---|---|---|---|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 | 8.49 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 | 3.45 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 | 7.00 |
| 3 | 23.68 | 3.31 | Male | No | Sun | Dinner | 2 | 11.84 |
| 4 | 24.59 | 3.61 | Female | No | Sun | Dinner | 4 | 6.15 |

|   | Payer Name | CC Number | Payment ID |
|---|---|---|---|
| 0 | Christy Cunningham | 3560325168603410 | Sun2959 |
| 1 | Douglas Tucker | 4478071379779230 | Sun4608 |
| 2 | Travis Walters | 6011812112971320 | Sun4458 |
| 3 | Nathaniel Harris | 4676137647685990 | Sun5260 |
| 4 | Tonya Carter | 4832732618637220 | Sun2251 |

Problem 3: Calculate percentage of tip to total bill.

Formula : (tip/Total Bill) * 100

```
tips["Tip"]/tips["Total Bill"]*100
```

```
0        5.944673
1       16.054159
2       16.658734
3       13.978041
4       14.680765
          ...
239     20.392697
240      7.358352
241      8.822232
242      9.820426
243     15.974441
Length: 244, dtype: float64
```

Problem 4: Create a new column of percentage tip

```
tip_percentage = tips["Tip"]/tips["Total Bill"]*100

tip_percentage
```

```
0         5.944673
1        16.054159
2        16.658734
3        13.978041
4        14.680765
           ...
239      20.392697
240       7.358352
241       8.822232
242       9.820426
243      15.974441
Length: 244, dtype: float64
```

Problem 5: Inserting tips_percentage col in tips dataframe

```
tips["tip_percentage"] = tips["Tip"]/tips["Total Bill"]*100

tips.head()
```

```
   Total Bill   Tip  Gender Smoker  Day    Time  Size  Bill Per Person
\
0       16.99  1.01  Female     No  Sun  Dinner     2             8.49

1       10.34  1.66    Male     No  Sun  Dinner     3             3.45

2       21.01  3.50    Male     No  Sun  Dinner     3             7.00

3       23.68  3.31    Male     No  Sun  Dinner     2            11.84

4       24.59  3.61  Female     No  Sun  Dinner     4             6.15


           Payer Name        CC Number Payment ID  tip_percentage
0  Christy Cunningham  3560325168603410    Sun2959        5.944673
1      Douglas Tucker  4478071379779230    Sun4608       16.054159
2      Travis Walters  6011812112971320    Sun4458       16.658734
3    Nathaniel Harris  4676137647685990    Sun5260       13.978041
4        Tonya Carter  4832732618637220    Sun2251       14.680765
```

Problem 6: Round upto one decimal place the tip_percentage column values.

```
tips["tip_percentage"] = tips["tip_percentage"].round(1)

tips.head()
```

|   | Total Bill | Tip | Gender | Smoker | Day | Time | Size | Bill Per Person |
|---|---|---|---|---|---|---|---|---|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 | 8.49 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 | 3.45 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 | 7.00 |
| 3 | 23.68 | 3.31 | Male | No | Sun | Dinner | 2 | 11.84 |
| 4 | 24.59 | 3.61 | Female | No | Sun | Dinner | 4 | 6.15 |

|   | Payer Name | CC Number | Payment ID | tip_percentage |
|---|---|---|---|---|
| 0 | Christy Cunningham | 3560325168603410 | Sun2959 | 5.9 |
| 1 | Douglas Tucker | 4478071379779230 | Sun4608 | 16.1 |
| 2 | Travis Walters | 6011812112971320 | Sun4458 | 16.7 |
| 3 | Nathaniel Harris | 4676137647685990 | Sun5260 | 14.0 |
| 4 | Tonya Carter | 4832732618637220 | Sun2251 | 14.7 |

Problem 7: Drop column Payer Number.

```python
tips = tips.drop(["Payer Name"], axis = 1)

tips.head()
```

|   | Total Bill | Tip | Gender | Smoker | Day | Time | Size | Bill Per Person |
|---|---|---|---|---|---|---|---|---|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 | 8.49 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 | 3.45 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 | 7.00 |
| 3 | 23.68 | 3.31 | Male | No | Sun | Dinner | 2 | 11.84 |
| 4 | 24.59 | 3.61 | Female | No | Sun | Dinner | 4 | 6.15 |

|   | CC Number | Payment ID | tip_percentage |
|---|---|---|---|
| 0 | 3560325168603410 | Sun2959 | 5.9 |
| 1 | 4478071379779230 | Sun4608 | 16.1 |
| 2 | 6011812112971320 | Sun4458 | 16.7 |
| 3 | 4676137647685990 | Sun5260 | 14.0 |
| 4 | 4832732618637220 | Sun2251 | 14.7 |

Problem 8: Index tips dataframe as per Payment ID

```python
tips.set_index("Payment ID")
```

```
                Total Bill    Tip  Gender Smoker    Day    Time  Size  \
Payment ID
Sun2959               16.99  1.01  Female     No    Sun  Dinner     2
Sun4608               10.34  1.66    Male     No    Sun  Dinner     3
Sun4458               21.01  3.50    Male     No    Sun  Dinner     3
Sun5260               23.68  3.31    Male     No    Sun  Dinner     2
Sun2251               24.59  3.61  Female     No    Sun  Dinner     4
...                     ...   ...     ...    ...    ...     ...   ...
Sat2657               29.03  5.92    Male     No    Sat  Dinner     3
Sat1766               27.18  2.00  Female    Yes    Sat  Dinner     2
Sat3880               22.67  2.00    Male    Yes    Sat  Dinner     2
Sat17                 17.82  1.75    Male     No    Sat  Dinner     2
Thur672               18.78  3.00  Female     No   Thur  Dinner     2

                Bill Per Person         CC Number  tip_percentage
Payment ID
Sun2959                    8.49  3560325168603410             5.9
Sun4608                    3.45  4478071379779230            16.1
Sun4458                    7.00  6011812112971320            16.7
Sun5260                   11.84  4676137647685990            14.0
Sun2251                    6.15  4832732618637220            14.7
...                         ...               ...             ...
Sat2657                    9.68  5296068606052840            20.4
Sat1766                   13.59  3506806155565400             7.4
Sat3880                   11.34  6011891618747190             8.8
Sat17                      8.91     4375220550950             9.8
Thur672                    9.39  3511451626698130            16.0

[244 rows x 10 columns]

tips.head()

   Total Bill    Tip  Gender Smoker  Day    Time  Size  Bill Per Person
\
0       16.99  1.01  Female     No  Sun  Dinner     2             8.49

1       10.34  1.66    Male     No  Sun  Dinner     3             3.45

2       21.01  3.50    Male     No  Sun  Dinner     3             7.00

3       23.68  3.31    Male     No  Sun  Dinner     2            11.84

4       24.59  3.61  Female     No  Sun  Dinner     4             6.15


          CC Number Payment ID  tip_percentage
0  3560325168603410    Sun2959             5.9
1  4478071379779230    Sun4608            16.1
2  6011812112971320    Sun4458            16.7
```

```
3   4676137647685990     Sun5260                14.0
4   4832732618637220     Sun2251                14.7
```

Problem 9: Change index tips dataframe as per Payment ID

```
tips = tips.set_index("Payment ID")
```

```
tips.head()
```

```
            Total Bill   Tip  Gender Smoker  Day    Time  Size  \
Payment ID
Sun2959          16.99  1.01  Female     No  Sun  Dinner     2
Sun4608          10.34  1.66    Male     No  Sun  Dinner     3
Sun4458          21.01  3.50    Male     No  Sun  Dinner     3
Sun5260          23.68  3.31    Male     No  Sun  Dinner     2
Sun2251          24.59  3.61  Female     No  Sun  Dinner     4

            Bill Per Person           CC Number  tip_percentage
Payment ID
Sun2959                8.49  3560325168603410             5.9
Sun4608                3.45  4478071379779230            16.1
Sun4458                7.00  6011812112971320            16.7
Sun5260               11.84  4676137647685990            14.0
Sun2251                6.15  4832732618637220            14.7
```

Eg for Locating row by payment id.

```
tips.loc["Sun4608"]
```

```
Total Bill                   10.34
Tip                           1.66
Gender                        Male
Smoker                          No
Day                            Sun
Time                        Dinner
Size                             3
Bill Per Person               3.45
CC Number         4478071379779230
tip_percentage                16.1
Name: Sun4608, dtype: object
```

Problem 10: Reset index of tips dataframe to row index

```
tips = tips.reset_index()
```

```
tips.head()
```

```
   Payment ID  Total Bill   Tip  Gender Smoker  Day    Time  Size  \
0     Sun2959       16.99  1.01  Female     No  Sun  Dinner     2
1     Sun4608       10.34  1.66    Male     No  Sun  Dinner     3
2     Sun4458       21.01  3.50    Male     No  Sun  Dinner     3
3     Sun5260       23.68  3.31    Male     No  Sun  Dinner     2
```

```
4     Sun2251        24.59  3.61  Female      No  Sun  Dinner      4

    Bill Per Person          CC Number  tip_percentage
0              8.49  3560325168603410             5.9
1              3.45  4478071379779230            16.1
2              7.00  6011812112971320            16.7
3             11.84  4676137647685990            14.0
4              6.15  4832732618637220            14.7
```