```
   id        date    NH4  BSK5  Suspended    O2    NO3    NO2    SO4    PO4     CL
0   1  17.02.2000  0.330  2.77       12.0  12.30   9.50  0.057  154.0  0.454  289.5
1   1  11.05.2000  0.044  3.00       51.6  14.61  17.75  0.034  352.0  0.090 1792.0
2   1  11.09.2000  0.032  2.10       24.5   9.87  13.80  0.173  416.0  0.200 2509.0
3   1  13.12.2000  0.170  2.23       35.6  12.40  17.13  0.099  275.2  0.377 1264.0
4   1  02.03.2001  0.000  3.03       48.8  14.69  10.00  0.065  281.6  0.134 1462.0
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
   id        date    NH4  BSK5  Suspended    O2    NO3    NO2    SO4    PO4     CL
0   1  17.02.2000  0.330  2.77       12.0  12.30   9.50  0.057  154.0  0.454  289.5
1   1  11.05.2000  0.044  3.00       51.6  14.61  17.75  0.034  352.0  0.090 1792.0
2   1  11.09.2000  0.032  2.10       24.5   9.87  13.80  0.173  416.0  0.200 2509.0
3   1  13.12.2000  0.170  2.23       35.6  12.40  17.13  0.099  275.2  0.377 1264.0
4   1  02.03.2001  0.000  3.03       48.8  14.69  10.00  0.065  281.6  0.134 1462.0
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2861 entries, 0 to 2860
Data columns (total 11 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   id         2861 non-null   int64
 1   date       2861 non-null   object
 2   NH4        2858 non-null   float64
 3   BSK5       2860 non-null   float64
 4   Suspended  2845 non-null   float64
 5   O2         2858 non-null   float64
 6   NO3        2860 non-null   float64
 7   NO2        2858 non-null   float64
 8   SO4        2812 non-null   float64
 9   PO4        2833 non-null   float64
 10  CL         2812 non-null   float64
dtypes: float64(9), int64(1), object(1)
memory usage: 246.0+ KB
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
(2861, 11)
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
             count       mean         std   min      25%     50%       75%      max
id          2861.0  12.397064    6.084226  1.00   8.0000  14.000  16.00000   22.000
NH4         2858.0   0.758734    2.486247  0.00   0.0800   0.220   0.50000   39.427
BSK5        2860.0   4.316182    2.973997  0.00   2.1600   3.800   5.80000   50.900
Suspended   2845.0  12.931905   16.543097  0.00   6.0000  10.000  15.00000  595.000
O2          2858.0   9.508902    4.428260  0.00   7.0925   8.995  11.52000   90.000
NO3         2860.0   4.316846    6.881188  0.00   1.3900   2.800   5.58250  133.400
NO2         2858.0   0.246128    2.182777  0.00   0.0300   0.059   0.12575  109.000
SO4         2812.0  59.362313   96.582641  0.00  27.0525  37.800  64.64000 3573.400
PO4         2833.0   0.418626    0.771326  0.00   0.1300   0.270   0.47000   13.879
CL          2812.0  93.731991  394.512184  0.02  26.8000  33.900  45.60750 5615.280
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
id              0
date            0
NH4             3
BSK5            1
Suspended      16
O2              3
NO3             1
NO2             3
SO4            49
PO4            28
CL             49
dtype: int64
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

	id	date	NH4	BSK5	Suspended	O2	NO3	NO2	SO4	PO4	CL
0	1	2000-02-17	0.330	2.77	12.0	12.30	9.50	0.057	154.0	0.454	289.5
1	1	2000-05-11	0.044	3.00	51.6	14.61	17.75	0.034	352.0	0.090	1792.0
2	1	2000-09-11	0.032	2.10	24.5	9.87	13.80	0.173	416.0	0.200	2509.0
3	1	2000-12-13	0.170	2.23	35.6	12.40	17.13	0.099	275.2	0.377	1264.0
4	1	2001-03-02	0.000	3.03	48.8	14.69	10.00	0.065	281.6	0.134	1462.0

```
Data columns (total 11 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   id          2861 non-null   int64
 1   date        2861 non-null   datetime64[ns]
 2   NH4         2858 non-null   float64
 3   BSK5        2860 non-null   float64
 4   Suspended   2845 non-null   float64
 5   O2          2858 non-null   float64
 6   NO3         2860 non-null   float64
 7   NO2         2858 non-null   float64
 8   SO4         2812 non-null   float64
 9   PO4         2833 non-null   float64
 10  CL          2812 non-null   float64
dtypes: datetime64[ns](1), float64(9), int64(1)
memory usage: 246.0 KB
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
   id       date    NH4  BSK5  Suspended     O2    NO3    NO2    SO4    PO4      CL
0   1 2000-02-17  0.330  2.77       12.0  12.30   9.50  0.057  154.0  0.454   289.5
1   1 2000-05-11  0.044  3.00       51.6  14.61  17.75  0.034  352.0  0.090  1792.0
2   1 2000-09-11  0.032  2.10       24.5   9.87  13.80  0.173  416.0  0.200  2509.0
3   1 2000-12-13  0.170  2.23       35.6  12.40  17.13  0.099  275.2  0.377  1264.0
4   1 2001-03-02  0.000  3.03       48.8  14.69  10.00  0.065  281.6  0.134  1462.0
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
   id       date    NH4  BSK5  Suspended     O2    NO3    NO2    SO4    PO4      CL  year  month
0   1 2000-02-17  0.330  2.77       12.0  12.30   9.50  0.057  154.0  0.454   289.5  2000      2
1   1 2000-05-11  0.044  3.00       51.6  14.61  17.75  0.034  352.0  0.090  1792.0  2000      5
2   1 2000-09-11  0.032  2.10       24.5   9.87  13.80  0.173  416.0  0.200  2509.0  2000      9
3   1 2000-12-13  0.170  2.23       35.6  12.40  17.13  0.099  275.2  0.377  1264.0  2000     12
4   1 2001-03-02  0.000  3.03       48.8  14.69  10.00  0.065  281.6  0.134  1462.0  2001      3
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```

```
Index(['id', 'date', 'NH4', 'BSK5', 'Suspended', 'O2', 'NO3', 'NO2', 'SO4',
       'PO4', 'CL', 'year', 'month'],
      dtype='object')
PS C:\Users\Krisha\Desktop\New folder\Water_Quality_Prediction>
```