# EDA Report

## df.head

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SINO | TotalBill | Tips | Smoker | Day | Time | Size |
| 1 | 16.99 | 1.01 | No | Sun | Dinner | 2.0 |
| 2 | 10.34 | 1.66 | No | Sun | Dinner | 3.0 |
| 3 | 21.01 | 3.5 | No | Sun | Dinner | 3.0 |
| 4 | 23.68 | 3.31 | No | Sun | Dinner | 2.0 |
| 5 | 24.59 | 3.61 | No | Sun | Dinner | 4.0 |

## df.describe

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistic | SINO | TotalBill | Tips | Size |
| count | 100.0 | 98.0 | 94.0 | 99.0 |
| mean | 50.5 | 19.805408163265305 | 3.036170212765957 | 2.525252525252525 |
| std | 29.011491975882016 | 7.950816747044754 | 1.201806083854297 | 0.8249110569672484 |
| min | 1.0 | 3.07 | 1.0 | 1.0 |
| 25% | 25.75 | 15.0175 | 2.035 | 2.0 |
| 50% | 50.5 | 18.285 | 3.0 | 2.0 |
| 75% | 75.25 | 23.964999999999996 | 3.5775 | 3.0 |
| max | 100.0 | 48.27 | 6.73 | 4.0 |

## df.info

<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 100 entries, 0 to 99  
Data columns (total 7 columns):  
 # Column Non-Null Count Dtype   
--- ------ -------------- -----   
 0 SINO 100 non-null int64   
 1 TotalBill 98 non-null float64  
 2 Tips 94 non-null float64  
 3 Smoker 93 non-null object   
 4 Day 93 non-null object   
 5 Time 100 non-null object   
 6 Size 99 non-null float64  
dtypes: float64(3), int64(1), object(3)  
memory usage: 5.6+ KB

## df.nunique

|  |  |
| --- | --- |
| Column | Number of unique values |
| SINO | 100 |
| TotalBill | 94 |
| Tips | 68 |
| Smoker | 2 |
| Day | 4 |
| Time | 2 |
| Size | 4 |

## df.isnull

|  |  |
| --- | --- |
| Column | Number of missing values values |
| SINO | 0 |
| TotalBill | 2 |
| Tips | 6 |
| Smoker | 7 |
| Day | 7 |
| Time | 0 |
| Size | 1 |