* 1. How to replace NULL in a Koalas Dataframe with specific value?  
     [**DataFrame.fillna**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.DataFrame.fillna.html#databricks.koalas.DataFrame.fillna)([value, method, axis, …])  
     [**DataFrame.bfill**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.DataFrame.bfill.html#databricks.koalas.DataFrame.bfill)([axis, inplace, limit])  
     [**DataFrame.ffill**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.DataFrame.ffill.html#databricks.koalas.DataFrame.ffill)([axis, inplace, limit])  
     same with series
  2. How to quickly collect stats for Pandas/Koalas Dataframe?  
     [**Series.mean**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.Series.mean.html#databricks.koalas.Series.mean)([axis, numeric\_only])  
     [**GroupBy.mean**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.groupby.GroupBy.mean.html#databricks.koalas.groupby.GroupBy.mean)()  
     [**DataFrame.mean**](https://koalas.readthedocs.io/en/latest/reference/api/databricks.koalas.DataFrame.mean.html#databricks.koalas.DataFrame.mean)([axis, numeric\_only])  
     and other stat functions in the same manner
  3. How to perform left outer join between Koalas dataframes df1 and df2 by id column?  
     df1.join(df2, how=’left’, on=’id’)
  4. What are use cases for Pandas Dataframe and for Koalas Dataframe?  
      Pandas is the standard tool for data science and it is typically the first step to explore and manipulate a data set, but pandas does not scale well to big data. With Koalas, data scientist can use the same APIs as pandas’ but at scale with PySpark.
  5. What Koalas Dataframe’s Index is used for? What is Spark analogue of Koalas Dataframe’s Index (if any)?  
     In Koalas, the default index is used in several cases, for instance, when Spark DataFrame is converted into Koalas DataFrame. In this case, internally Koalas attaches a default index into Koalas DataFrame.  
     There are several types of the default index that can be configured by compute.default\_index\_type: sequence, distributed\_sequence, distributed  
       
     The PySpark DataFrame, on the other hand, tends to be more compliant with the relations/tables in relational databases, and does not have unique row identifiers.