Pandas Functions -

1. Create Data Frame - pd.DataFrame()
2. Size of Data Frame - df.shape
3. Display First N Rows - df.head(n) OR df[:n]
4. Select Data - df.loc[df[‘id’]==1,[“columns”] #df.loc[rows,columns]
5. Create New Column - df[‘new column name”] = df[“old column name”] * 2
6. Drop Duplicates - df.drop_duplicates(subset=”col1”, keep =”first/last/False”, inplace=True)
7. Drop Missing Data - df.dropna(subset=”col1”, axis=0/1, how=”any”/”all”, inplace=True)
8. Modify Column Name - df[“col1”] = df[“col1”] * 2
9. Rename Columns - df.rename(columns = { “col1” : “newcol1”})
10. Change Data Type - df.astype({“col1” : datatype})
11. Fill Missing Data - df.fillna(value= 0, method = ffill/ bfill/etc, axis=0/1, inplace=True)
12. Concatenate/ COmbine Data - pd.concat([df1,df2], axis=0)
13. Pivot Table - df.pivot(index=”col1”,columns=”col2”, values=”col3”)
14. Reshape Data Using Melt : df = df.melt(id_vars = “unchanged column name”, value_vars=[“col1”,”col2”,”col3”] (name of column to be made into rows), var_name=”name of column after adding rows”, value_name = “col4”)
15. Method Chaining : df = df.sort_values(by=”col1”, ascending=True/False).loc[[“col2”] >/</== val, [“col3”]]