* **Index**

df.set\_index(‘C’,IPT)

df.sort\_index(IPT)

df.reset\_index(,IPT)

* **Subset**

df[‘C]

df.loc[]

* **Filter**

filter = df[‘C’]>10

.str.contains()

df[filter]

(f1) & (f2), (f1) | (f2)

* **Count values in a series**

df[‘C’].value\_counts()

* **Grouping**

grp= df.groupby(‘C’)

grp[‘C1’,’C2’].count()

* **Aggregate functions**

.nunique(),.sum(),.mean(),.max(),.min()

grp[‘C’].agg([‘min’,’max’,’mean’])

* **Join**

pd.merge(df1,df2,how='outer',on=’C’)

**self-join**

df.join(df.drop(‘C’,1).set\_index(‘Cn’),on=’Cn’,rsuffix=’Cn’)

* **Sort**

df.sort\_values(by=[‘Cn’])

sorted(“series”)

* **Find, Fill, & Drop NaN**

df[df.isna().any(axis=1)]

df.dropna(axis=0,subset=[‘C’], IPT)

df[‘C’].fillna(value, IPT)

* **Duplicates**

df[df[‘C’].duplicated(keep=False)==True]

df.drop([row #],axis=0, IPT)

df.drop\_duplicates(subset=’C’,keep=’first’, IPT)

* **Create a unique key**

df[‘C’].astype(str)

* **Concat data sources**

pd.concat([df1,df2])