

```
In [ ]: # Grouping
```

```
In [ ]: import pandas as pd
```

```
In [ ]: pd.read_csv('路径')
df.head(2) #head 2 rows
df.tail(2) #tail 2 rows
```

```
In [ ]: df.groupby('Sex').replace('?.np.nan).dropna()
```

```
In [ ]: for k,g in df.groupby('Sex'):
        print(k,g) #give tuples of unique values
```

```
In [ ]: df.groupby('Education').min()
```

```
In [ ]: # agg
df[['Education', 'Year', 'Salary']].groupby('Education').agg(['mean', 'mi
```

```
In [ ]: df.groupby('Education').agg({'Department': ['size'], 'Sex': ['size'], 'Year': ['size']})
# 每个col都有自己的想要的东西
```

```
In [ ]: # 这个还是一个dataframe type
```

```
In [ ]: def myfunc(X):
        return np.mean(X[X!=X.max()])
df[['Department', 'Salary', 'Year']].groupby('Department').agg(myfunc)
```

```
In [ ]: # Example
df=pd.read_csv('.../asd/asda/...', sep=";")
df.head()
```

```
In [ ]: # choose only the following columns: ...
# change all the columns to lower cases
# drop all columns/rows with missing values
# limit the dataset to countries with populations over 20 million
# reset the index to start from 0
# add a new column (coastal) that is true for countries with coastline
df=df[['Country', 'Continent'...]]
df.columns=[i.lower() for i in df.columns]
# df.columns=df.columns.str.lower()
df=df.dropna()
df=df.loc[df.Population>20,: ]
df=df.reset_index()
# df=df.reset_index().drop('index',axis=1)
del df['index']
df['coastal'] = df['coastline']>0
'''
new_col=pd.DataFrame([[True] if i=='yes' else [False] for i in df.Coastline])
df=df.concat([df,new_col],axis=1)
'''
```

```
In [ ]: df.groupby(['coastal', [continent]]).agg(['min', 'max'])
```

```
In [ ]: # Other see the uploaded video
```

```
In [ ]: df['dense']=df['density']>df.density.mean()  
def max_min_diff(x):  
    return np.max(x)-np.min(x)  
df[['dense', 'birthrate']].groupby('dense').agg(['mean', max_min_diff])
```

```
In [ ]: # pivot table
```

```
In [ ]: df.pivot_table(index='continent', columns='coastal', values='birthrate',
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```