import pandas as pd

import seaborn as sns

link = "https://s3-ap-southeast-1.amazonaws.com/intro-to-ml-minhdh/Salaries.csv"

sal = pd.read\_csv(link)

sal.head()

sal.info()

sal['BasePay'].mean()

sal['BasePay'].min()

sal['BasePay'].max()

sal['OvertimePay'].describe()

sal[sal['EmployeeName']=='JOSEPH DRISCOLL']['JobTitle']

sal[sal['EmployeeName']=='GARY JIMENEZ']['JobTitle']

sal[sal['EmployeeName']=='GARY JIMENEZ']['TotalPayBenefits']

sal[sal['TotalPayBenefits']==sal['TotalPayBenefits'].max()]

sal.groupby('Year').mean()['BasePay']

sal['JobTitle'].nunique()

sal['JobTitle'].value\_counts().head(5)

(sal[sal['Year']==2013]['JobTitle'].value\_counts()==1).sum()

sal = sns.countplot(x = 'Year',data=sal)