Assignment 9

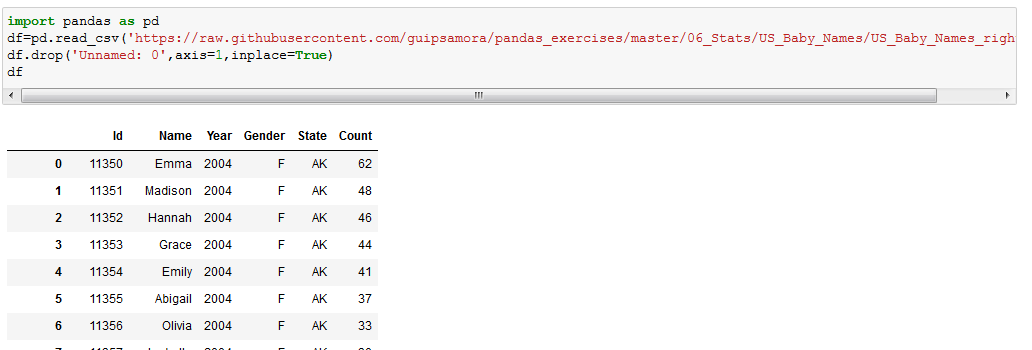
Q1. Delete unnamed columns

Answer:

import pandas as pd

df=pd.read\_csv('https://raw.githubusercontent.com/guipsamora/pandas\_exercises/master/06\_Stats/US\_Baby\_Names/US\_Baby\_Names\_right.csv')

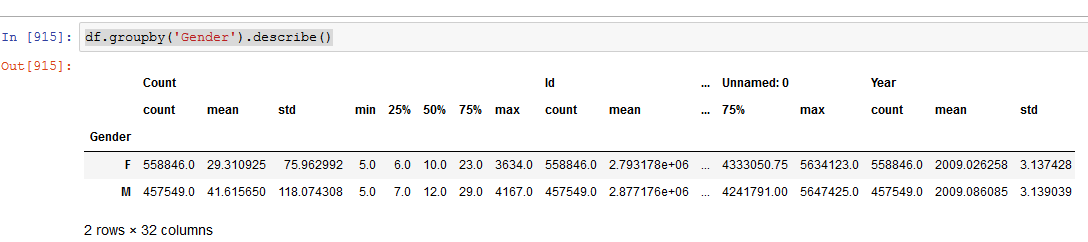
df.drop('Unnamed: 0',axis=1,inplace=True)



Q2. Show the distribution of male and female

Answer:

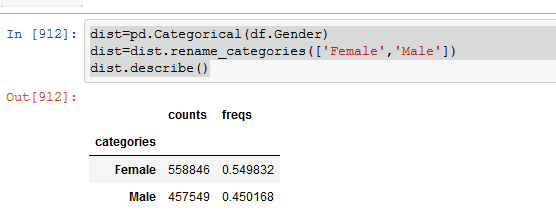
1. df.groupby('Gender').describe()



1. dist=pd.Categorical(df.Gender)

dist=dist.rename\_categories(['Female','Male'])

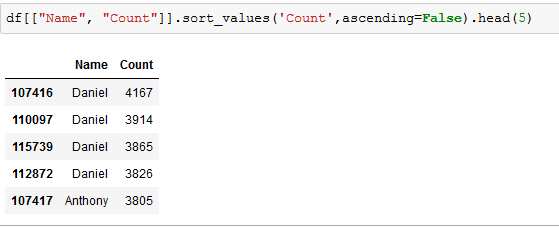
dist.describe()



Q3. Show the top 5 most preferred names

Answer :

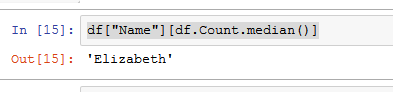
df[["Name", "Count"]].sort\_values('Count',ascending=False).head(5)



Q4. What is the median name occurrence in the dataset

Answer:

df["Name"][df.Count.median()]



5. Distribution of male and female born count by states

Answer:

df.groupby(['Gender','State']).describe()

