import numpy as np import pandas as pd import seaborn as sns import scipy.stats as stats from scipy.stats import linregress from scipy.stats import norm from matplotlib import pyplot as plt cities = pd.read csv('C:/Users/Home/Desktop/bigcity.csv') In [4]: cities.head(7) cities.tail(2) Out[4]: Unnamed: 0 47 232 161 48 36 54 cities.columns Out[5]: Index(['Unnamed: 0', 'u', 'x'], dtype='object') cities = cities.drop(columns = "id") cities.shape (49, 3) cities.describe() id X **count** 49.00000 49.000000 49.000000 **mean** 25.00000 103.142857 127.795918 14.28869 104.405061 123.121211 min 1.00000 2.000000 46.000000 **25%** 13.00000 43.000000 58.000000 **50%** 25.00000 64.000000 79.000000 **75%** 37.00000 120.000000 130.000000 max 49.00000 507.000000 634.000000 plt.plot(cities) Out[22]: [<matplotlib.lines.Line2D at 0x25d29fb2670>, <matplotlib.lines.Line2D at 0x25d29fb2730>, <matplotlib.lines.Line2D at 0x25d29fb27f0>] 600 500 400 300 200 100 20 30 x = cities['x']u = cities['u'] plt.scatter(u, x) m, b = np.polyfit(u, x, 1)plt.plot(u, m*u + b)Out[26]: [<matplotlib.lines.Line2D at 0x25d2a076490>] 600 500 400 300 200 100 0 100 300 400 500 In [45]: sns.boxplot(x) C:\Users\Home\anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other ar guments without an explicit keyword will result in an error or misinterpretation. warnings.warn(Out[45]: <AxesSubplot:xlabel='x'> 100 200 300 500 600 400 sns.barplot(u,x)C:\Users\Home\anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning: Pass the following variable s as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing othe r arguments without an explicit keyword will result in an error or misinterpretation. warnings.warn(Out[38]: <AxesSubplot:xlabel='u', ylabel='x'> 600 500 400 300 200 100 22.25.29.06.7840.3445468.5660.164677.17467787934.1007.11639.11729759985507 sns.violinplot(x) C:\Users\Home\anaconda3\lib\site-packages\seaborn\ decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other ar guments without an explicit keyword will result in an error or misinterpretation. warnings.warn(Out[39]: <AxesSubplot:xlabel='x'> -100 100 200 300 400 500 600 700 Х In [43]: pd.DataFrame(cities).boxplot() Out[43]: <AxesSubplot:> 600 500 ٥ 400 300 8 200 100 In [46]: pd.DataFrame(cities).hist() Out[46]: array([[<AxesSubplot:title={'center':'id'}>, <AxesSubplot:title={'center':'u'}>], [<AxesSubplot:title={'center':'x'}>, <AxesSubplot:>]], dtype=object) id 20 10 2 0 0 ²⁰ x 40 200 400 30 20 10 200 400 600 pd.DataFrame(cities).plot.kde() Out[67]: <AxesSubplot:ylabel='Density'> 0.006 0.005 0.004 0.003 0.002 0.001 0.000 -200 600 800 200 400 pd.DataFrame(cities).plot.bar() head Traceback (most recent call last) <ipython-input-56-ac8715ad3775> in <module> 1 pd.DataFrame(cities).plot.bar() ---> 2 limit(5) NameError: name 'limit' is not defined 600 500 400 300 200 100 cities.plot.line() Out[57]: <AxesSubplot:> 600 500 400 300 200 100 0 30 g = cities.head(5)In [64]: g.plot.bar() Out[64]: <AxesSubplot:> 250 200 150 100 50 dist = norm(u.mean(), u.std()) xs = np.linspace(-200,800)ys = dist.pdf(xs)plt.clf() plt.plot(xs, ys) sns.kdeplot(u) Out[70]: <AxesSubplot:xlabel='u', ylabel='Density'> 0.006 0.005 0.004 0.003 0.002 0.001 0.000 -200200 400 600 800 0 linregress(xs, ys) Out[71]: LinregressResult(slope=-2.2161684554950963e-06, intercept=0.0016435536429545683, rvalue=-0.5020601352474419, pv alue=0.00020332570688773712, stderr=5.510087811430419e-07) In [87]: cities['x'].replace([48], 100, inplace = True) cities.min() 2 Out[88]: u 50 dtype: int64