In [5]:

```
print("innovation in education across USA, China, and Germany: world need ideas and innovation to deal p

#literacy rates across USA, China, and Germany
print(' the following data frames and graphs below show how the literacy rates vary between USA,China an
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd


dataframe = pd.read_csv('cross-country-literacy-rates two.csv')

dataframe
```

innovation in education across USA, China, and Germany: world need ideas and innovation to deal progessively with the major problems the humanity face.Countries should create opportunities and conditions to cultivate such creative and talented people who contribute to this important work. let us look at the human resources'quality of those three countries
 the following data frames and graphs below show how the literacy rates vary between USA,China and Germany

Out[5]:

	COUNTRY_Code	Year	rates_of_Literacy
0	Germany DEU	1475	9.000000
1	Germany DEU	1550	16.000000
2	Germany DEU	1650	31.000000
3	Germany DEU	1750	38.000000
4	Germany DEU	2003	99.000000
5	China CHN	1982	65.505089
6	China CHN	1990	77.785057
7	China CHN	2000	90.920212
8	China CHN	2010	95.124481
9	China CHN	2015	96.357452
10	United States USA	1870	80.000000
11	United States USA	1880	83.000000
12	United States USA	1890	86.000000
13	United States USA	1900	89.300003
14	United States USA	1910	92.300003
15	United States USA	1920	94.000000
16	United States USA	1930	95.699997
17	United States USA	1940	97.099998
18	United States USA	1947	97.300003
19	United States USA	1950	96.800003
20	United States USA	1952	97.500000
21	United States USA	1959	97.800003
22	United States USA	1969	99.000000
23	United States USA	1979	99.400002
24	United States USA	2003	99.000000

In [6]:

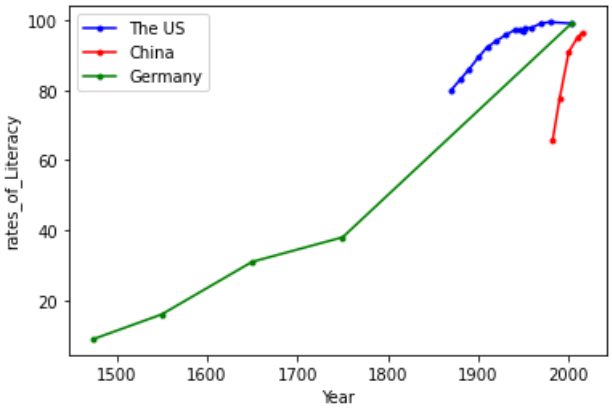
```
# let us Compare Literacy rates across the US, China and Germany
US = dataframe[ dataframe.COUNTRY_Code == 'United States USA' ]
China = dataframe[ dataframe.COUNTRY_Code == 'China CHN' ]
Germany = dataframe[ dataframe.COUNTRY_Code == 'Germany DEU' ]
Germany
```

Out[6]:

	COUNTRY_Code	Year	rates_of_Literacy
0	Germany DEU	1475	9.0
1	Germany DEU	1550	16.0
2	Germany DEU	1650	31.0
3	Germany DEU	1750	38.0
4	Germany DEU	2003	99.0

In [7]:

```
plt.plot(US.Year, US.rates_of_Literacy,'b.-')
plt.plot(China.Year, China.rates_of_Literacy,'r.-')
plt.plot(Germany.Year, Germany.rates_of_Literacy,'g.-')
plt.legend(['The US', 'China', 'Germany'])
plt.xlabel('Year')
plt.ylabel('rates_of_Literacy')
plt.show()
```



In [8]:

```
#as we can see from these graphs, US was leader for gaining higher literacy rates within the past decade.
```

In [9]:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
researchers = pd.read_csv('researchers-in-rd-per-million-peopleII.csv')

researchers
```

Out[9]:

	country	Code	Year	Researchers_in_R_and_D_(per_million_people)
0	China	CHN	1996	442.56548
1	China	CHN	1997	471.99494
2	China	CHN	1998	386.77167
3	China	CHN	1999	420.60209
4	China	CHN	2000	547.30387
5	China	CHN	2001	581.53188
6	China	CHN	2002	631.07801
7	China	CHN	2003	667.53208
8	China	CHN	2004	713.28449
9	China	CHN	2005	856.84548
10	China	CHN	2006	932.31417
11	China	CHN	2007	1078.62521
12	China	CHN	2008	1200.29490
13	China	CHN	2009	863.92630
14	China	CHN	2010	902.95975
15	China	CHN	2011	977.68206
16	China	CHN	2012	1035.87909

	country	Code	Year	Researchers_in_R_and_D_(per_million_people)
17	China	CHN	2013	1089.19205
18	China	CHN	2014	1113.07185
19	China	CHN	2015	1176.57712
20	Germany	DEU	1996	2811.61389
21	Germany	DEU	1997	2875.74074
22	Germany	DEU	1998	2898.56684
23	Germany	DEU	1999	3107.28263
24	Germany	DEU	2000	3148.80136
25	Germany	DEU	2001	3231.71759
26	Germany	DEU	2002	3253.51966
27	Germany	DEU	2003	3297.09098
28	Germany	DEU	2004	3318.86921
29	Germany	DEU	2005	3349.64573
30	Germany	DEU	2006	3452.20997
31	Germany	DEU	2007	3597.23882
32	Germany	DEU	2008	3751.78161
33	Germany	DEU	2009	3940.73576
34	Germany	DEU	2010	4077.76719
35	Germany	DEU	2011	4211.25463
36	Germany	DEU	2012	4379.07515
37	Germany	DEU	2013	4399.67239
38	Germany	DEU	2014	4363.77950
39	Germany	DEU	2015	4431.08151
40	United States	USA	1996	3122.57281
41	United States	USA	1997	3224.23913
42	United States	USA	1998	3388.20888
43	United States	USA	1999	3445.30961
44	United States	USA	2000	3475.69505
45	United States	USA	2001	3545.83215
46	United States	USA	2002	3630.61832
47	United States	USA	2003	3870.56966
48	United States	USA	2004	3765.10261
49	United States	USA	2005	3718.19505
50	United States	USA	2006	3781.80788
51	United States	USA	2007	3757.86688
52	United States	USA	2008	3911.53450
53	United States	USA	2009	4073.17633
54	United States	USA	2010	3868.56644
55	United States	USA	2011	4011.32862
56	United States	USA	2012	4015.88708
57	United States	USA	2013	4117.67409
58	United States	USA	2014	4231.98928

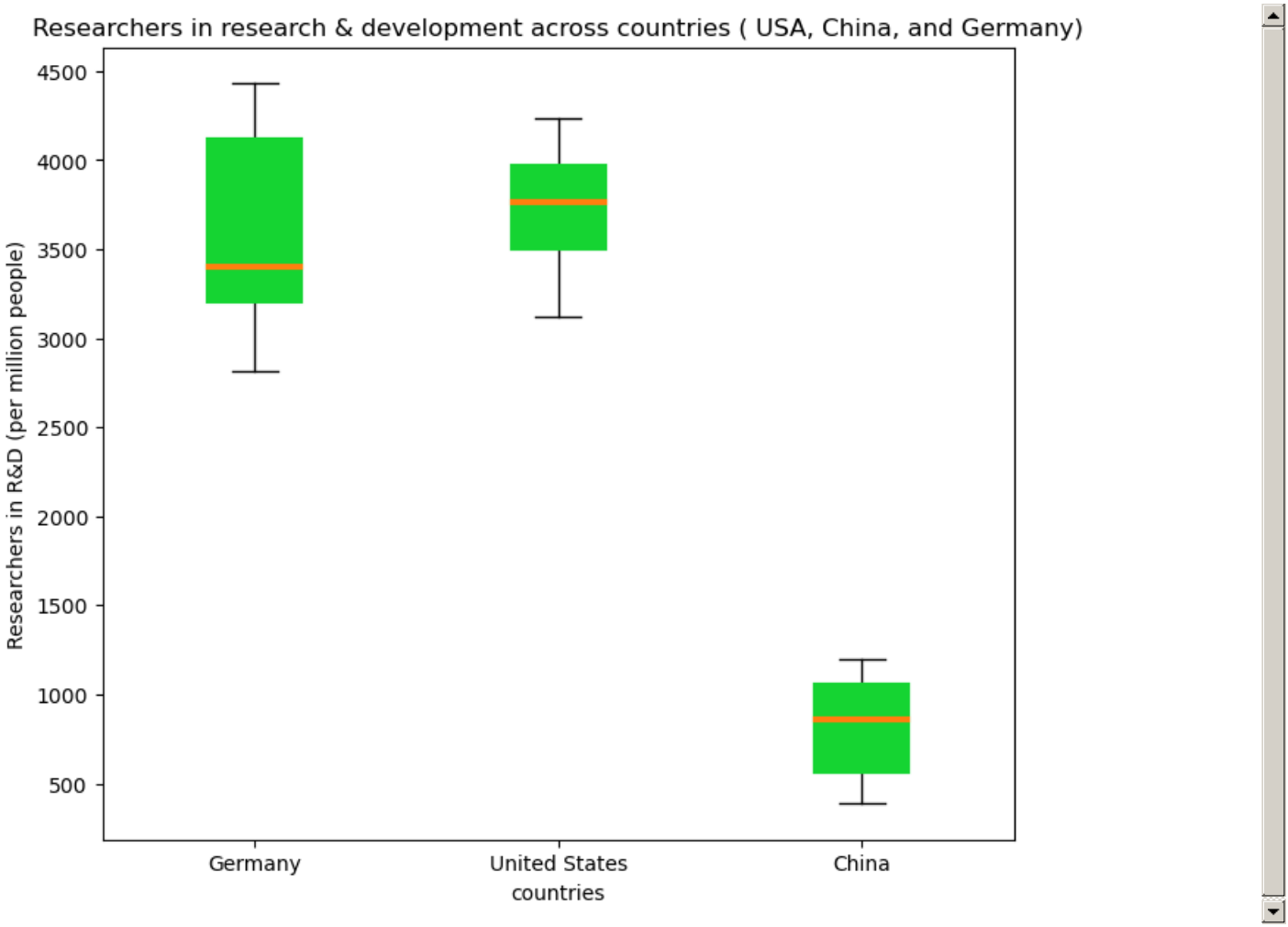
In [162]:

```
plt.style.use('default')

plt.figure(figsize=(8,7))
Germany = researchers.loc[researchers.country =='Germany']['Researchers_in_R_and_D_(per_million_people)']
United_States = researchers.loc[researchers.country == 'United States']['Researchers_in_R_and_D_(per_mill
```

```
China = researchers.loc[researchers.country == 'China']['Researchers_in_R_and_D_(per_million_people)']

labels = ['Germany','United States','China']
boxes = plt.boxplot([Germany, United_States,China],labels= labels,patch_artist=True, medianprops={'linewi
for box in boxes['boxes']:
    box.set(color='#15d432',linewidth =3)
plt.title('Researchers in research & development across countries ( USA, China, and Germany)')
plt.ylabel('Researchers in R&D (per million people)')
plt.xlabel('countries')
plt.show()
```



In [36]:

```
#The OECD Programme for International Student Assessment (PISA) examines what students know in reading, 
#It provides the most comprehensive and rigorous international assessment of student learning outcomes t
#Results from PISA indicate the quality and equity of learning outcomes attained around the world, and a

#NOTE: The Reading, Mathematics and Science scale ranges from 0 to 1000. Some apparent differences betwe
#SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student

print("women vs men PISA reading performances across countries")

PISA = pd.read_csv('Averages for age 15 years PISA reading scale.csv')
PISA
```

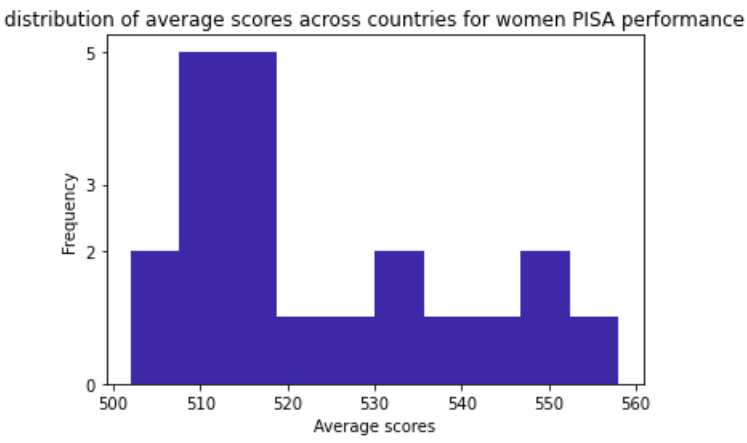
```
women vs men PISA reading performance across countries
```

Out[36]:

	Year	Country	average	Standard_Error
0	2018	Germany	512	(3.2)
1	2018	United States	517	(3.6)
2	2018	Hong Kong (China)	542	(2.8)
3	2015	Germany	520	(3.1)
4	2015	United States	507	(3.9)
5	2015	Hong Kong (China)	541	(3.6)
6	2012	Germany	530	(3.1)
7	2012	United States	513	(3.8)
8	2012	Hong Kong (China)	558	(3.3)
9	2009	Germany	518	(2.9)
10	2009	United States	513	(3.8)
11	2009	Hong Kong (China)	550	(2.8)
12	2006	Germany	517	(4.4)
13	2006	United States	515	(3.8)
14	2006	Hong Kong (China)	551	(3.0)
15	2003	Germany	513	(3.9)
16	2003	United States	511	(3.5)
17	2003	Hong Kong (China)	525	(3.5)
18	2000	Germany	502	(3.9)
19	2000	United States	518	(6.2)
20	2000	Hong Kong (China)	533	(3.6)

In [26]:

```
plt.hist(PISA.average,color='#3e28a8')
plt.yticks([0,2,3,5])
plt.xlabel('Average scores')
plt.ylabel('Frequency')
plt.title('distribution of average scores across countries for women PISA performance')
plt.show()
```



In [43]:

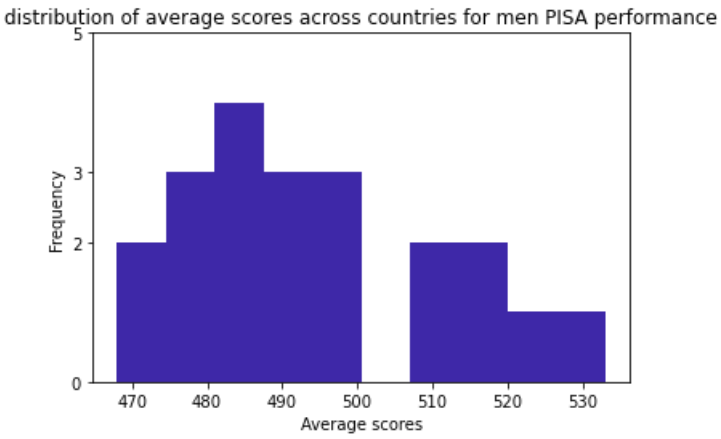
```
PISAmen = pd.read_csv('Averages for age 15 years PISA reading scale ( male).csv')
PISAmen
```

Out[43]:

	Year	Country	average	Standard Error
0	2018	Germany	486	(3.4)
1	2018	United States	494	(4.2)
2	2018	Hong Kong (China)	507	(3.5)
3	2015	Germany	499	(3.7)
4	2015	United States	487	(3.7)
5	2015	Hong Kong (China)	513	(3.4)
6	2012	Germany	486	(2.9)
7	2012	United States	482	(4.1)
8	2012	Hong Kong (China)	533	(3.8)
9	2009	Germany	478	(3.6)
10	2009	United States	488	(4.2)
11	2009	Hong Kong (China)	518	(3.3)
12	2006	Germany	475	(5.3)
13	2006	United States	490	(4.1)
14	2006	Hong Kong (China)	520	(3.5)
15	2003	Germany	471	(4.2)
16	2003	United States	479	(3.7)
17	2003	Hong Kong (China)	494	(5.3)
18	2000	Germany	468	(3.2)
19	2000	United States	490	(8.4)
20	2000	Hong Kong (China)	518	(4.8)

In [44]:

```
plt.hist(PISAmen.average, color='#3e28a8')
plt.yticks([0,2,3,5])
plt.xlabel('Average scores')
plt.ylabel('Frequency')
plt.title('distribution of average scores across countries for men PISA performance')
plt.show()
```



In [49]:

```
# mean PISA performance in The science, math and reading across countries
print('mean reading performance across countries')
mean_PISA_reading = pd.read_csv('PISA mean reading performance.csv')
mean_PISA_reading
```

```
mean math performance
```

Out[49]:

	country	Year	PISA: Mean performance on the reading scale
0	Germany	2000	483.990824
1	Germany	2003	491.357999
2	Germany	2006	494.944418
3	Germany	2009	497.305058
4	Germany	2012	507.676530
5	Germany	2015	509.104100
6	United States	2000	504.419691
7	United States	2003	495.182412
8	United States	2009	499.826821
9	United States	2012	497.581718
10	United States	2015	496.935100
11	China	2009	555.828113
12	China	2012	569.588408
13	China	2015	493.941200

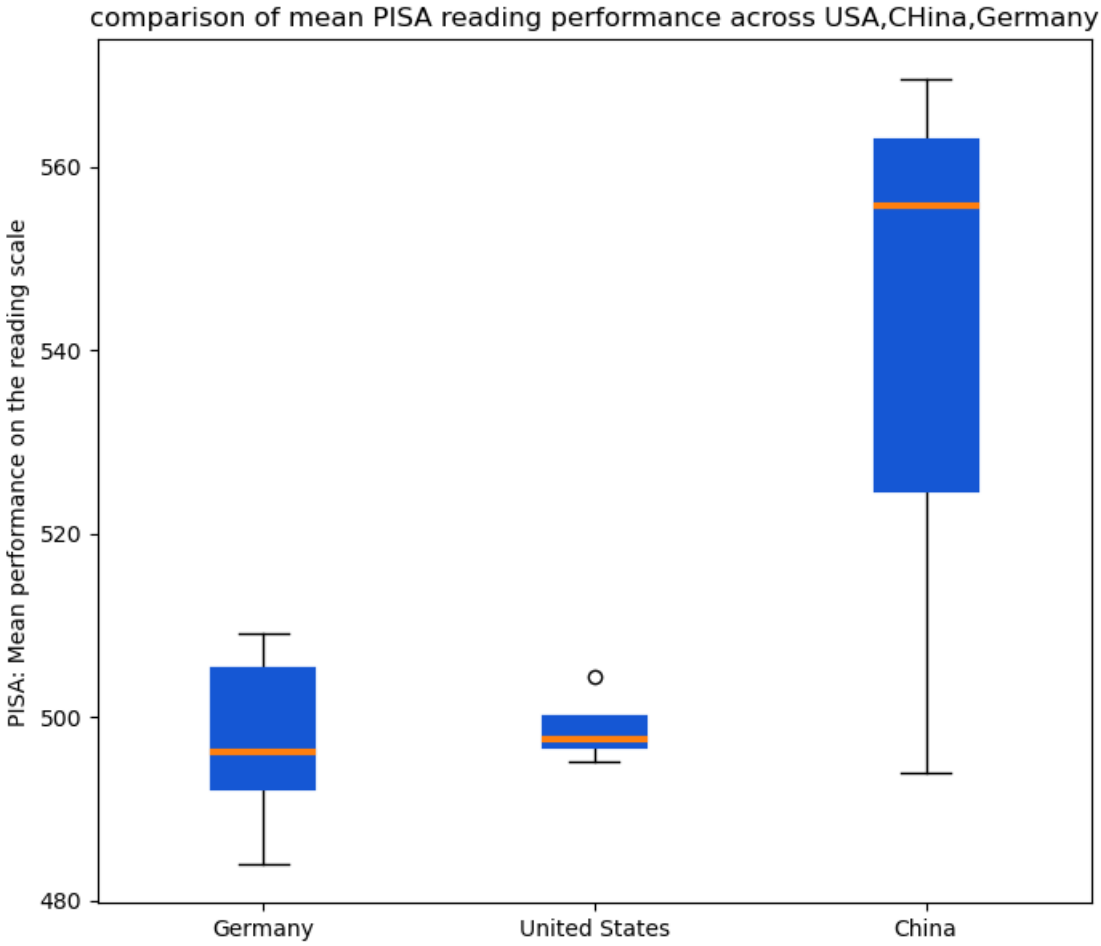
In [88]:

```
plt.style.use('default')

plt.figure(figsize=(8,7))
Germany = mean_PISA_reading.loc[mean_PISA_reading.country =='Germany']['PISA: Mean performance on the rea
United_States = mean_PISA_reading.loc[mean_PISA_reading.country == 'United States']['PISA: Mean performar
China = mean_PISA_reading.loc[mean_PISA_reading.country == 'China']['PISA: Mean performance on the readir

labels = ['Germany','United States','China']
boxes = plt.boxplot([Germany, United_States,China],labels= labels,patch_artist=True, medianprops={'linewi
for box in boxes['boxes']:
    box.set(color='#1557d4',linewidth =3)
plt.title('comparison of mean PISA reading performance across USA,CHina,Germany')
plt.ylabel('PISA: Mean performance on the reading scale')

plt.show()
```



In [90]:

```
print('mean math performance across countries')
mean_PISA_math = pd.read_csv('PISA mean math performance.csv')
mean_PISA_math
```

```
mean math performance across countries
```

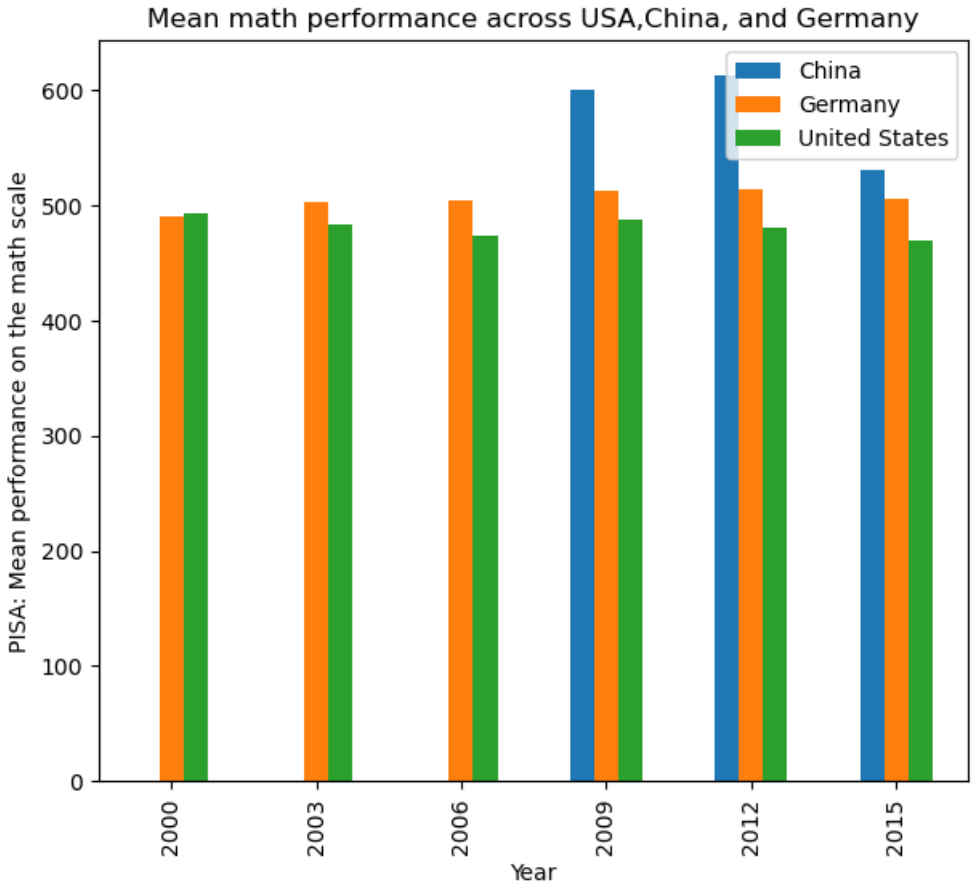
Out[90]:

	country	Year	PISA: Mean performance on the mathematics scale
0	United States	2000	493.000000
1	United States	2003	482.885038
2	United States	2006	474.352158
3	United States	2009	487.396520
4	United States	2012	481.366786
5	United States	2015	469.628500
6	China	2009	600.076253
7	China	2012	612.675536
8	China	2015	531.296100
9	Germany	2000	490.000000
10	Germany	2003	502.985532
11	Germany	2006	503.790859
12	Germany	2009	512.777643
13	Germany	2012	513.525056
14	Germany	2015	505.971300

In [181]:

```
mean_PISA_math.pivot(index='Year', columns ='country', values = 'PISA: Mean performance on the mathematic
plt.xlabel('Year')
plt.ylabel('PISA: Mean performance on the math scale')
plt.title('Mean math performance across USA,China, and Germany')
```

```
plt.legend(loc='upper right')
plt.show()
```



In [164]:

```
print('mean science performance across countries')
mean_PISA_science = pd.read_csv('PISA mean science performance.csv')
mean_PISA_science
```

```
mean science performance across countries
```

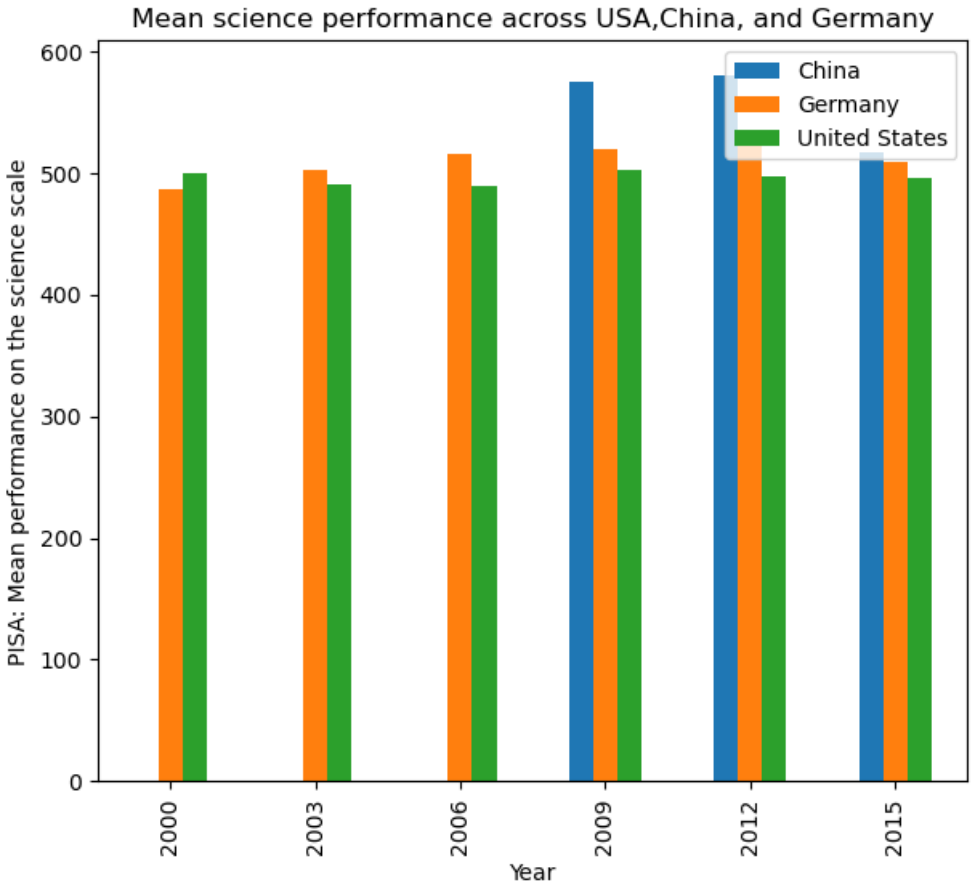
Out[164]:

	country	Year	PISA: Mean performance on the science scale
0	United States	2000	499.460152
1	United States	2003	491.263463
2	United States	2006	488.906837
3	United States	2009	502.002265
4	United States	2012	497.409811
5	United States	2015	496.242400
6	China	2009	574.617355
7	China	2012	580.117831
8	China	2015	517.779300
9	Germany	2000	487.105836
10	Germany	2003	502.336476
11	Germany	2006	515.649130
12	Germany	2009	520.405349
13	Germany	2012	524.120799
14	Germany	2015	509.140600

In [179]:

```
mean_PISA_science.pivot(index='Year', columns ='country', values = 'PISA: Mean performance on the science
plt.xlabel('Year')
plt.ylabel('PISA: Mean performance on the science scale')
plt.title('Mean science performance across USA,China, and Germany')
```

```
plt.legend(loc='upper right')
plt.show()
```



In [191]:

```
Digital_device_usage = pd.read_csv('using digital devices.csv')
Digital_device_usage
```

Out[191]:

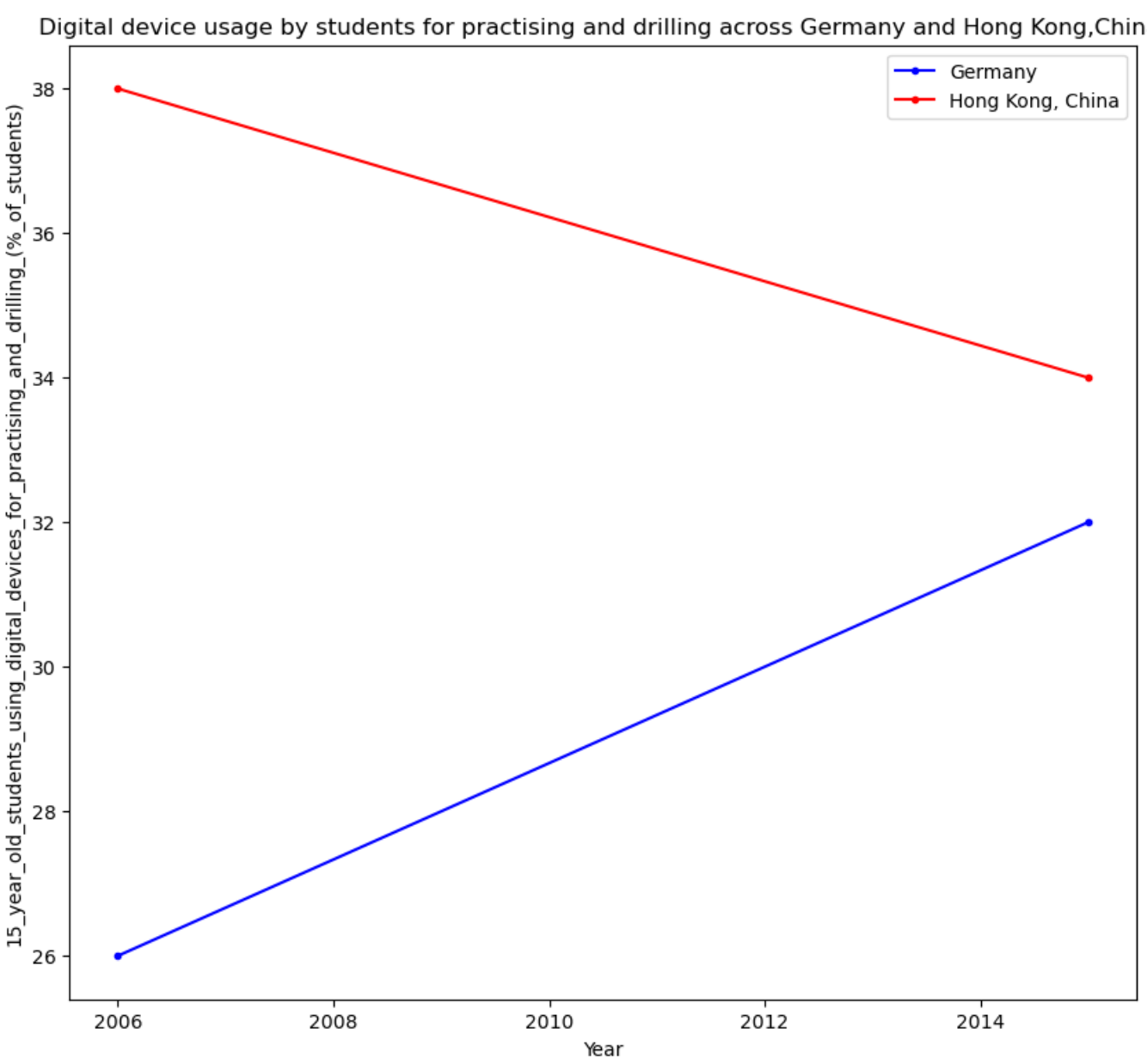
	Country	Year	15_year_old_students_using_digital_devices_for_practising_and_drilling_(%_of_students)
0	Germany	2006	26
1	Germany	2015	32
2	Hong Kong China	2006	38
3	Hong Kong China	2015	34

In [200]:

```
Germany = Digital_device_usage[ Digital_device_usage.Country == 'Germany' ]
Hong_Kong_China = Digital_device_usage[ Digital_device_usage.Country == 'Hong Kong China' ]

Germany_students= Digital_device_usage.loc[Digital_device_usage.Country =='Germany']['15_year_old_student
Hong_Kong_China_students= Digital_device_usage.loc[Digital_device_usage.Country =='Hong Kong China']['15_
plt.figure(figsize=(10,9))
plt.plot(Germany.Year,Germany_students,'b.-')
plt.plot(Hong_Kong_China.Year,Hong_Kong_China_students,'r.-')

plt.title('Digital device usage by students for practising and drilling across Germany and Hong Kong,Chir
plt.legend(['Germany', 'Hong Kong, China', ])
plt.xlabel('Year')
plt.ylabel('15_year_old_students_using_digital_devices_for_practising_and_drilling_(%_of_students)')
plt.show()
```


Digital device usage by students for practising and drilling across Germany and Hong Kong,China
Germany
Hong Kong, China
15_year_old_students_using_digital_devices_for_practising_and_drilling_(%_of_students)
38
36
34
32
30
28
26
2006
2008
2010
2012
2014
Year

In [ ]: