NAME:JACKLINE KIMORGO

REG NO:SCT211-0039/2017

ASSIGNMENT 2

In [144]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

In [145]:

```python
a=pd.read_csv("happy.csv")#pandas will read data into this file
a
```

Out[145]:

	Country (region)	Ladder	SD of Ladder	Positive affect	Negative affect	Social support	Freedom	Corruption	Generosity
0	Finland	1	4	41.0	10.0	2.0	5.0	4.0	47.0
1	Denmark	2	13	24.0	26.0	4.0	6.0	3.0	22.0
2	Norway	3	8	16.0	29.0	3.0	3.0	8.0	11.0
3	Iceland	4	9	3.0	3.0	1.0	7.0	45.0	3.0
4	Netherlands	5	1	12.0	25.0	15.0	19.0	12.0	7.0
5	Switzerland	6	11	44.0	21.0	13.0	11.0	7.0	16.0
6	Sweden	7	18	34.0	8.0	25.0	10.0	6.0	17.0
7	New Zealand	8	15	22.0	12.0	5.0	8.0	5.0	8.0
8	Canada	9	23	18.0	49.0	20.0	9.0	11.0	14.0
9	Austria	10	10	64.0	24.0	31.0	26.0	19.0	25.0
10	Australia	11	26	47.0	37.0	7.0	17.0	13.0	6.0
11	Costa Rica	12	62	4.0	87.0	42.0	16.0	58.0	75.0
12	Israel	13	14	104.0	69.0	38.0	93.0	74.0	24.0
13	Luxembourg	14	3	62.0	19.0	27.0	28.0	9.0	30.0
14	United Kingdom	15	16	52.0	42.0	9.0	63.0	15.0	4.0
15	Ireland	16	34	33.0	32.0	6.0	33.0	10.0	9.0
16	Germany	17	17	65.0	30.0	39.0	44.0	17.0	19.0
17	Belgium	18	7	57.0	53.0	22.0	53.0	20.0	44.0
18	United States	19	49	35.0	70.0	37.0	62.0	42.0	12.0
19	Czech Republic	20	20	74.0	22.0	24.0	58.0	121.0	117.0
20	United Arab Emirates	21	65	43.0	56.0	72.0	4.0	NaN	15.0
21	Malta	22	42	83.0	103.0	16.0	12.0	32.0	5.0
22	Mexico	23	76	6.0	40.0	67.0	71.0	87.0	120.0
23	France	24	19	56.0	66.0	32.0	69.0	21.0	68.0
24	Taiwan	25	37	17.0	1.0	48.0	102.0	56.0	56.0
25	Chile	26	61	15.0	78.0	58.0	98.0	99.0	45.0
26	Guatemala	27	136	8.0	85.0	78.0	25.0	82.0	78.0
27	Saudi Arabia	28	93	49.0	82.0	62.0	68.0	NaN	82.0

	Country (region)	Ladder	SD of Ladder	Positive affect	Negative affect	Social support	Freedom	Corruption	Generosity
28	Qatar	29	86	NaN	NaN	NaN	NaN	NaN	NaN
29	Spain	30	21	107.0	107.0	26.0	95.0	78.0	50.0
...	...	...	...	...	...	...	...	...	...
126	Congo (Kinshasa)	127	78	125.0	95.0	107.0	125.0	106.0	127.0
127	Mali	128	96	48.0	122.0	112.0	110.0	107.0	138.0
128	Sierra Leone	129	153	139.0	149.0	135.0	116.0	112.0	79.0
129	Sri Lanka	130	91	32.0	81.0	80.0	55.0	111.0	35.0
130	Myanmar	131	70	45.0	86.0	96.0	29.0	24.0	1.0
131	Chad	132	139	136.0	151.0	141.0	142.0	80.0	106.0
132	Ukraine	133	69	131.0	44.0	56.0	141.0	143.0	66.0
133	Ethiopia	134	38	100.0	74.0	119.0	106.0	53.0	99.0
134	Swaziland	135	104	26.0	57.0	103.0	113.0	41.0	145.0
135	Uganda	136	148	91.0	139.0	114.0	99.0	95.0	74.0
136	Egypt	137	66	146.0	124.0	118.0	129.0	89.0	132.0
137	Zambia	138	145	84.0	128.0	115.0	73.0	69.0	53.0
138	Togo	139	103	123.0	147.0	149.0	120.0	72.0	131.0
139	India	140	41	93.0	115.0	142.0	41.0	73.0	65.0
140	Liberia	141	156	103.0	146.0	127.0	94.0	126.0	110.0
141	Comoros	142	143	67.0	114.0	143.0	148.0	81.0	62.0
142	Madagascar	143	77	46.0	96.0	128.0	146.0	116.0	136.0
143	Lesotho	144	150	72.0	64.0	98.0	97.0	59.0	151.0
144	Burundi	145	138	98.0	126.0	152.0	135.0	23.0	149.0
145	Zimbabwe	146	123	63.0	34.0	110.0	96.0	63.0	141.0
146	Haiti	147	111	142.0	119.0	146.0	152.0	48.0	20.0
147	Botswana	148	125	87.0	65.0	105.0	60.0	54.0	150.0
148	Syria	149	137	155.0	155.0	154.0	153.0	38.0	69.0
149	Malawi	150	132	129.0	110.0	150.0	65.0	64.0	109.0
150	Yemen	151	85	153.0	75.0	100.0	147.0	83.0	155.0
151	Rwanda	152	63	54.0	102.0	144.0	21.0	2.0	90.0
152	Tanzania	153	122	78.0	50.0	131.0	78.0	34.0	49.0
153	Afghanistan	154	25	152.0	133.0	151.0	155.0	136.0	137.0
154	Central African Republic	155	117	132.0	153.0	155.0	133.0	122.0	113.0

	Country (region)	Ladder	SD of Ladder	Positive affect	Negative affect	Social support	Freedom	Corruption	Generosity
155	South Sudan	156	140	127.0	152.0	148.0	154.0	61.0	85.0

156 rows × 11 columns

In [146]:
```
b=a[:19]#slicing my data into 19 rows

x=b['Social support']#selecting this as a column to use to analyze data
x.values
```

Out[146]:
```
array([ 2.,  4.,  3.,  1., 15., 13., 25.,  5., 20., 31.,  7., 42., 38.,
       27.,  9.,  6., 39., 22., 37.])
```

In [147]:
```
y=b['Generosity']
y.values
```

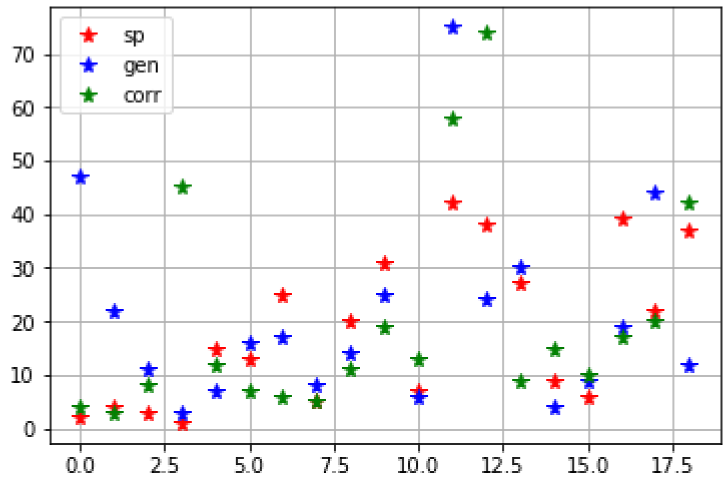
Out[147]:
```
array([47., 22., 11.,  3.,  7., 16., 17.,  8., 14., 25.,  6., 75., 24.,
       30.,  4.,  9., 19., 44., 12.])
```

In [148]:
```
v=b['Corruption']
v.values
```

Out[148]:
```
array([ 4.,  3.,  8., 45., 12.,  7.,  6.,  5., 11., 19., 13., 58., 74.,
        9., 15., 10., 17., 20., 42.])
```

scatter plot

In [149]:
```
#scatter plot for x.values and y.values
x=b['Social support']
y=b['Generosity']
plt.plot(x,'r*',markersize=8,label='sp')
plt.plot(y,'b*',markersize=8,label='gen')
plt.plot(v,'g*',markersize=8,label='corr')
plt.legend(loc='upper left')
plt.grid()
```



In [150]:
```
#the ouput should be same
x.shape
```

Out[150]: (19,)

In [151]:
```
y.shape
```

Out[151]: (19,)

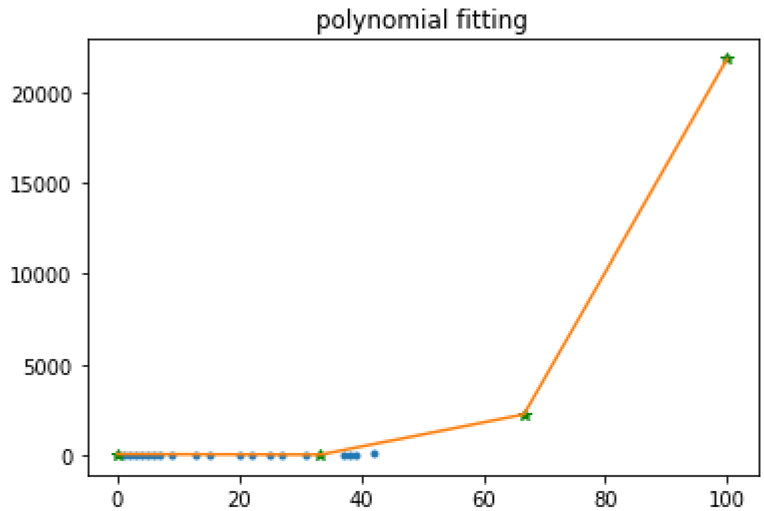
In [152]:
```
x.ndim
```

Out[152]: 1

Plotting the polynomial

In [153]:
```
p1=np.polyfit(x,y,4)#polynomial of degree 4
xp1=np.poly1d(p1)#creates an abstraction for the polynomial math operations
p2=np.polyfit(x,y,4)#polynomial of degree 4
xp2=np.poly1d(p2)#encapsulate to solve the polynomial
z_array=np.linspace(0,100,4)
plt.plot(x,y,".",z_array,xp1(z_array),'g*',z_array,xp2(z_array),"-")
plt.title('polynomial fitting')
```

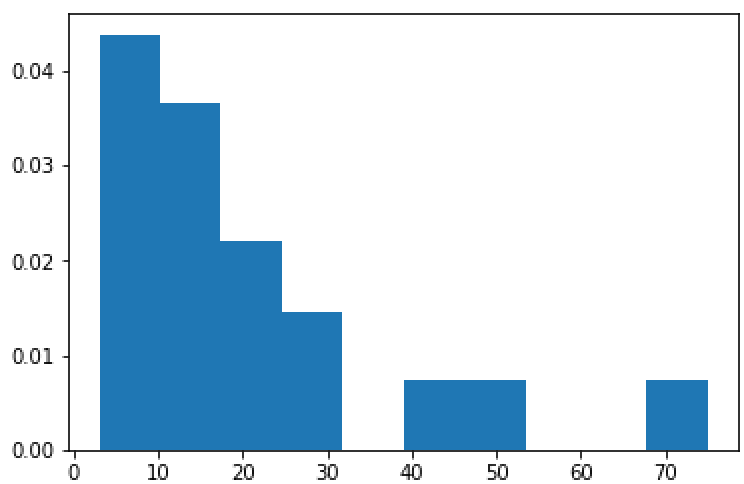
Out[153]:
```
Text(0.5, 1.0, 'polynomial fitting')
```



Histogram

In [154]:
```
plt.hist(y,normed=True)
```

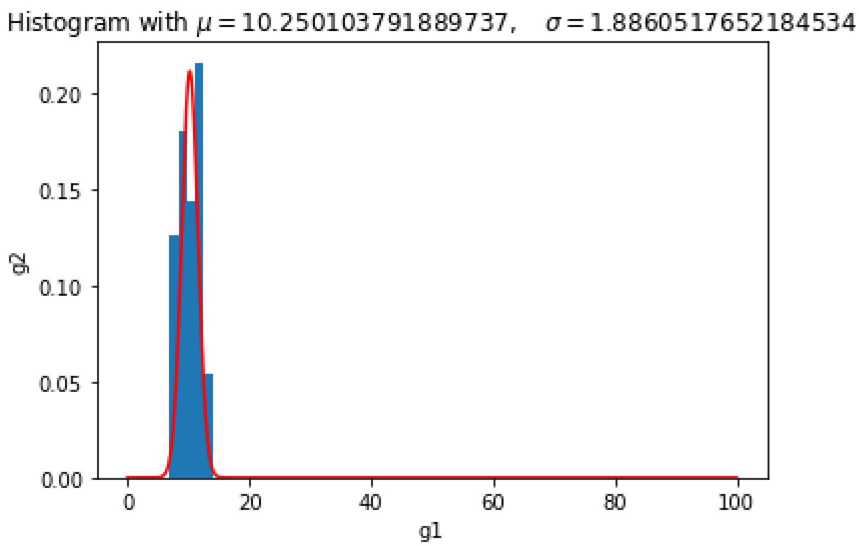
Out[154]:
```
(array([0.04385965, 0.03654971, 0.02192982, 0.01461988, 0.        ,
        0.00730994, 0.00730994, 0.        , 0.        , 0.00730994]),
 array([ 3. , 10.2, 17.4, 24.6, 31.8, 39. , 46.2, 53.4, 60.6, 67.8, 75. ]),
 <a list of 10 Patch objects>)
```



plotting the normal

In [196]:
```
sorted_x=np.sort(x)
mu=np.mean(sorted_x)#findimg the mean of x values sorted
sigma=np.std(sorted_x)#finding the standard deviation
plt.hist(sorted_x,5,normed=True)#5 is the number of bins
y=np.linspace(0,100,100)
plt.plot(y,(1/np.sqrt(2*np.pi*sigma**2))*np.exp(-(y-mu)**2/(sigma**2)),'r')
plt.xlabel('g1')
plt.ylabel('g2')
plt.title('Histogram with $\mu={},\quad \sigma={}$'.format(mu,sigma))
```

Out[196]:
```
Text(0.5, 1.0, 'Histogram with $\\mu=10.250103791889737,\\quad \\sigma=1.886051
7652184534$')
```



In [ ]: