

データ解析への入門（大和 元） - 講義資料  
< 幹葉図：日本の山を「山として」見る >

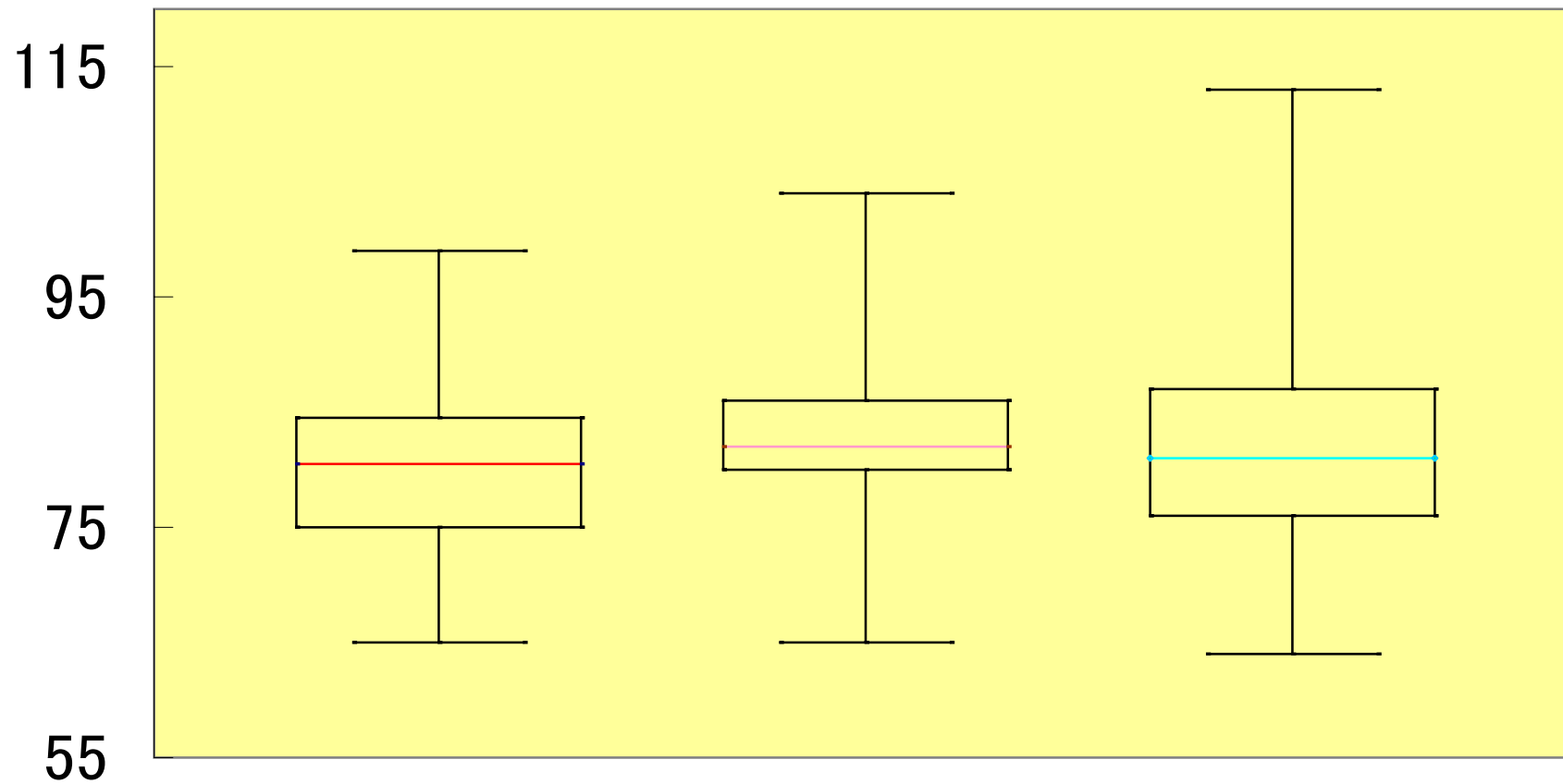
日本の主な山、幹葉図

(データ(単位 m)の数値は10の位で四捨五入し、1桁目は省略)

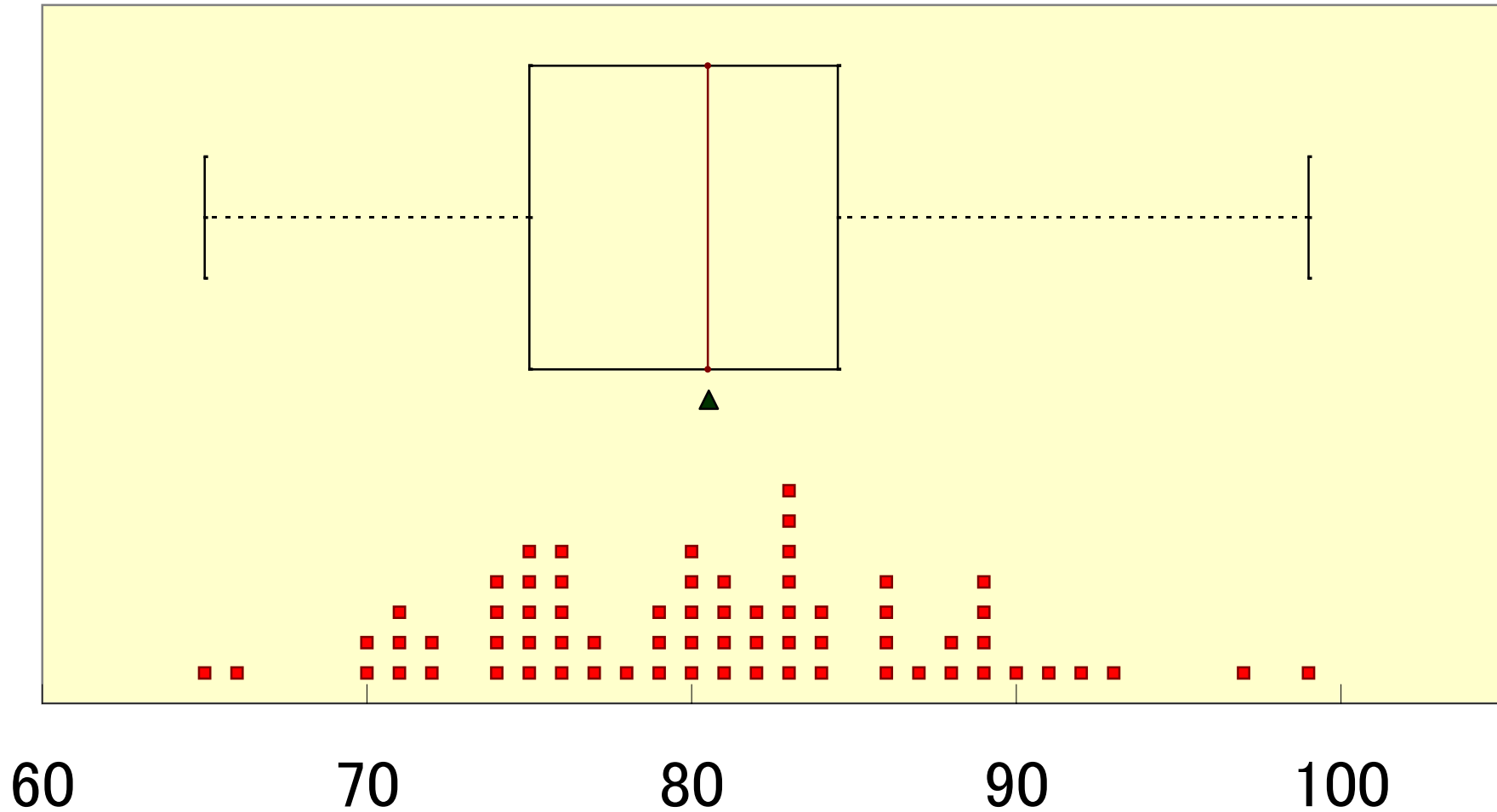
[データは理科年表（丸善）2006による]

1	456	
2	023444	
3	223889	
4	01124445566779	
5	0012334788	
6	00112334455689	
7	0002336889	
8	0235555668	
9	122223346789	(開聞岳、924m)
10	2236678	
11	0223335779	
12	00112334455557	
13	0012445557788	
14	12224455999	
15	00012345666889	(阿蘇山<高岳> 1592m)
16	0123346778	
17	0001223346899	
18	2347899	
19	0012223346678888	(宮之浦岳、1936m)
20	0112334456689	
21	1133579	
22	34914556	
24	0056688889	
25	37899	
26	00	
27	078	
28	000011111122234444556666677789	
29	0000122233679	
30	012333557889	
31	011248999	
32		
33		
34		
35		
36		
37	8	(富士山<剣ヶ峯>、3776m)

## 箱ヒゲ図 (06年選手体重) 阪神 と 読売 そして ソフトバンク



## 箱ヒゲ図+ドットプロット(阪神06年支配下選手体重)

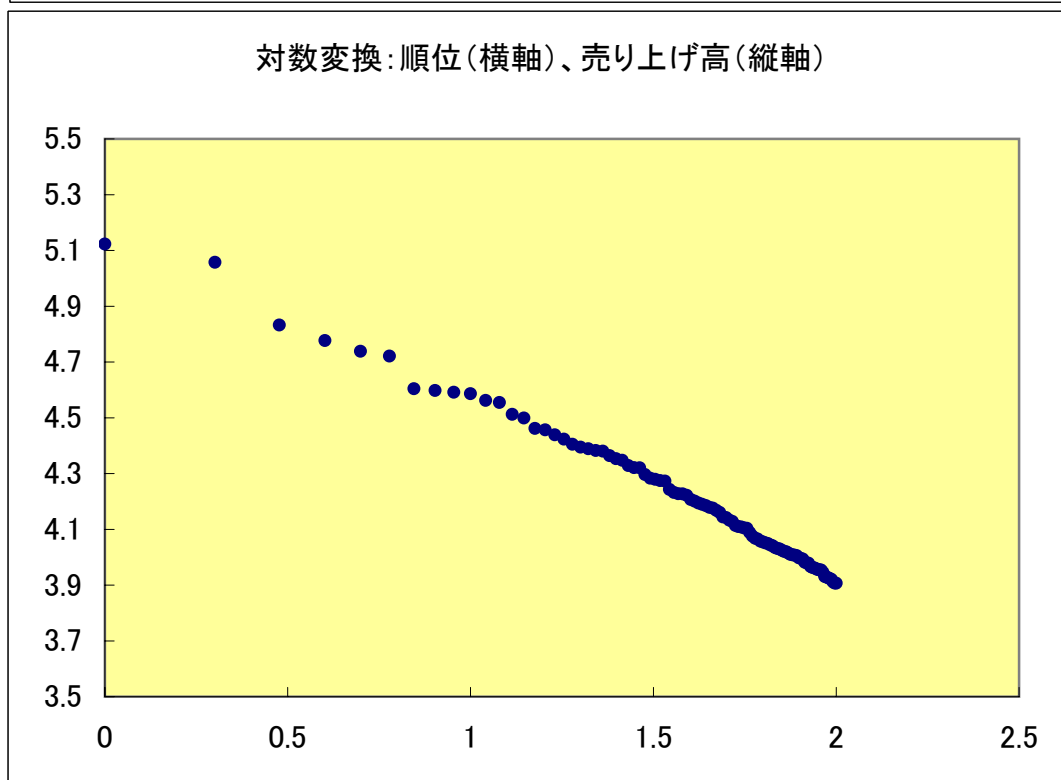
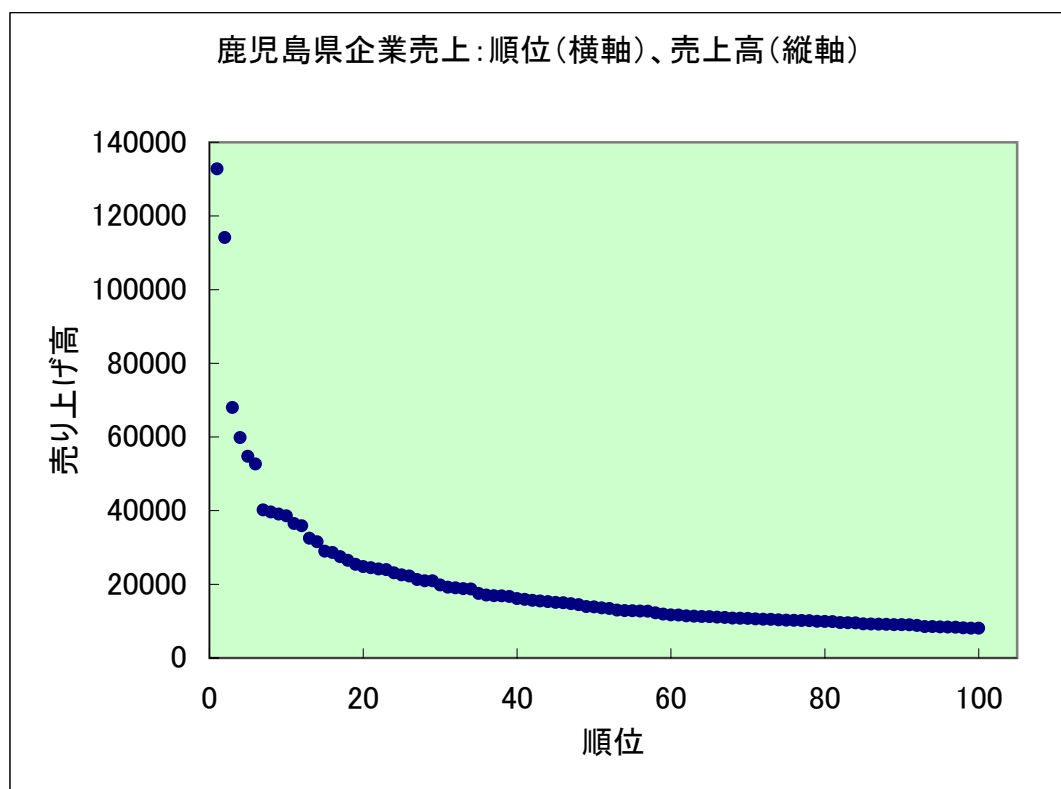


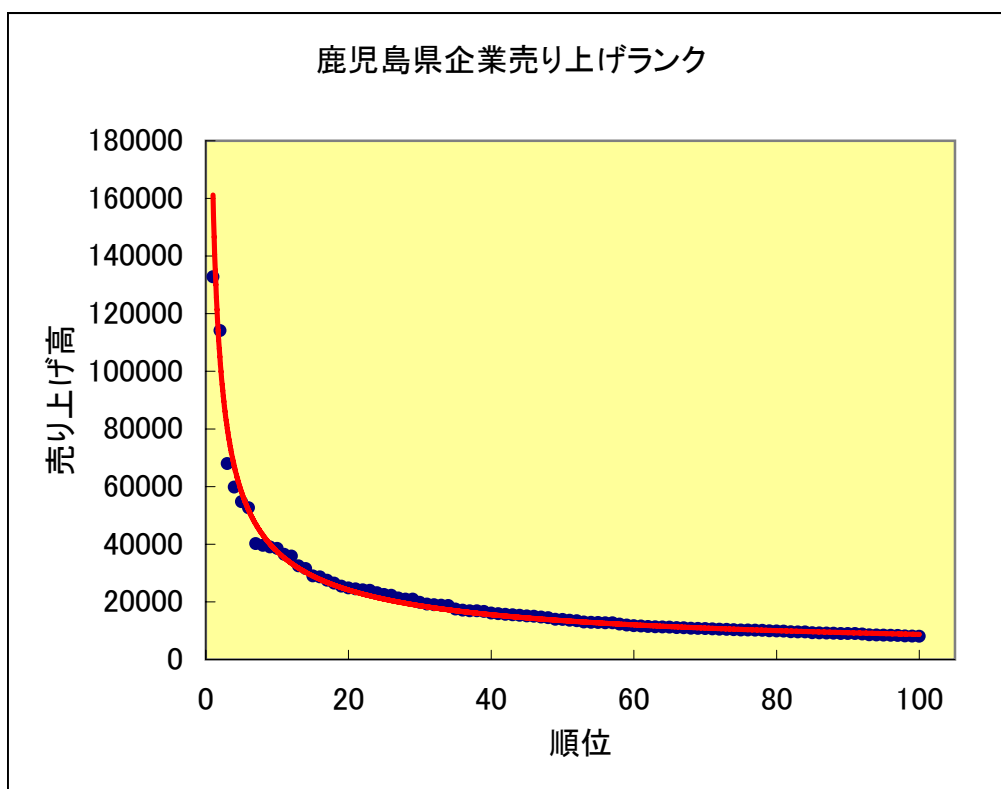
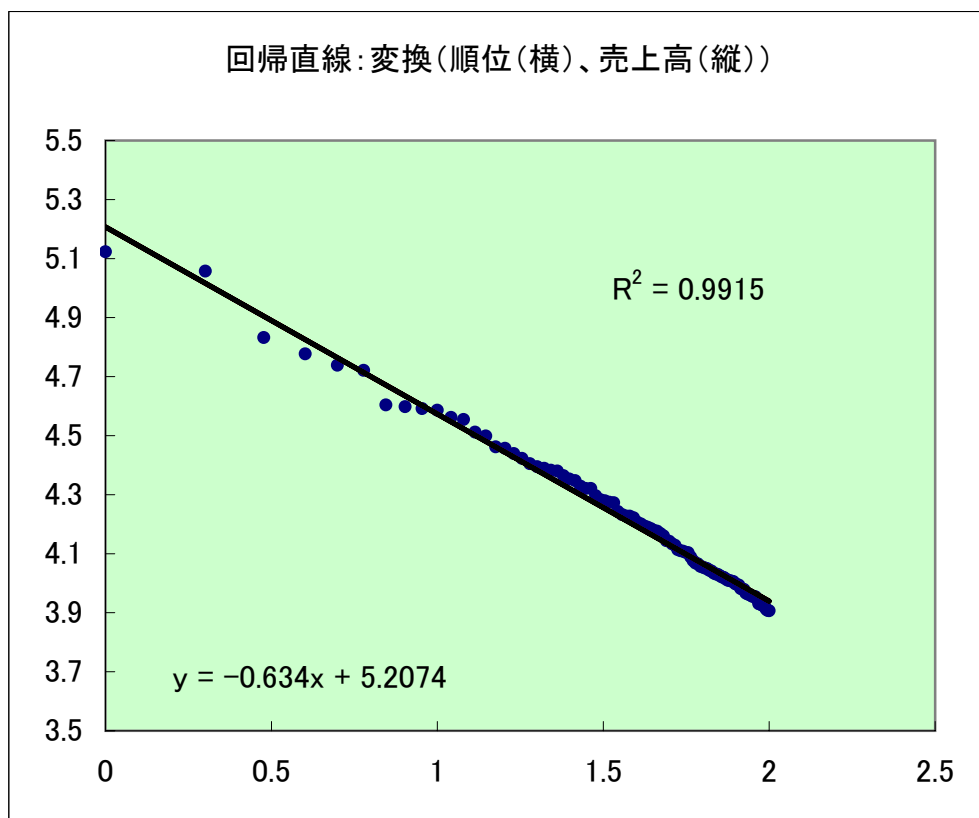
は平均を示す。（データは日本野球機構のホームページから。）

<データ解析への入門（大和 元）>

2007/2/24

データ解析への入門（大和 元） 講義資料  
ーZipfの法則、鹿児島県2005年度企業売り上げ上位100社へー  
（データは南日本新聞8月18日（2006年）から）

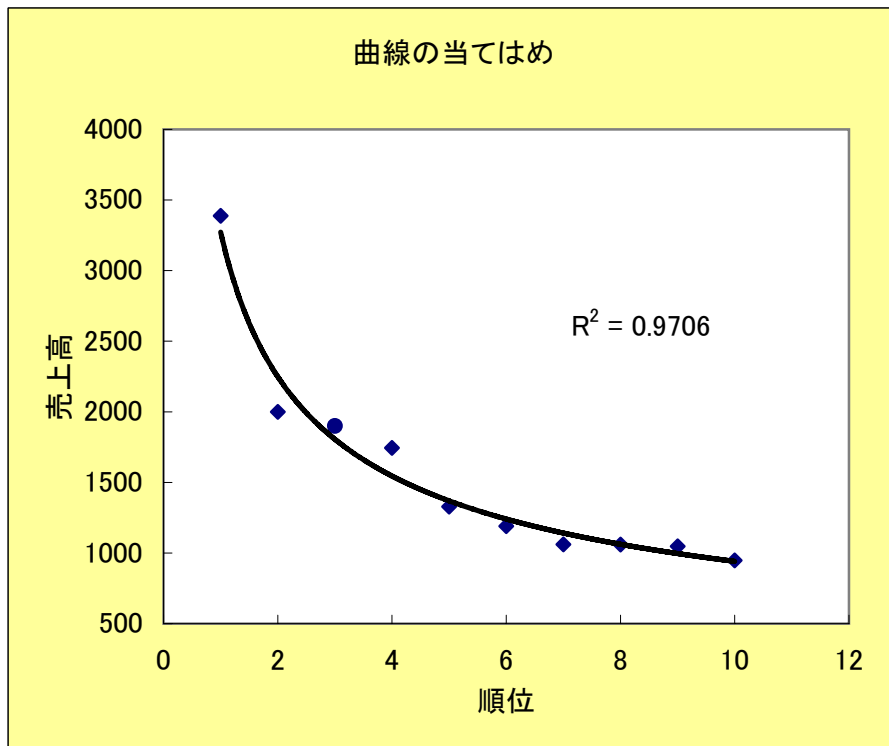
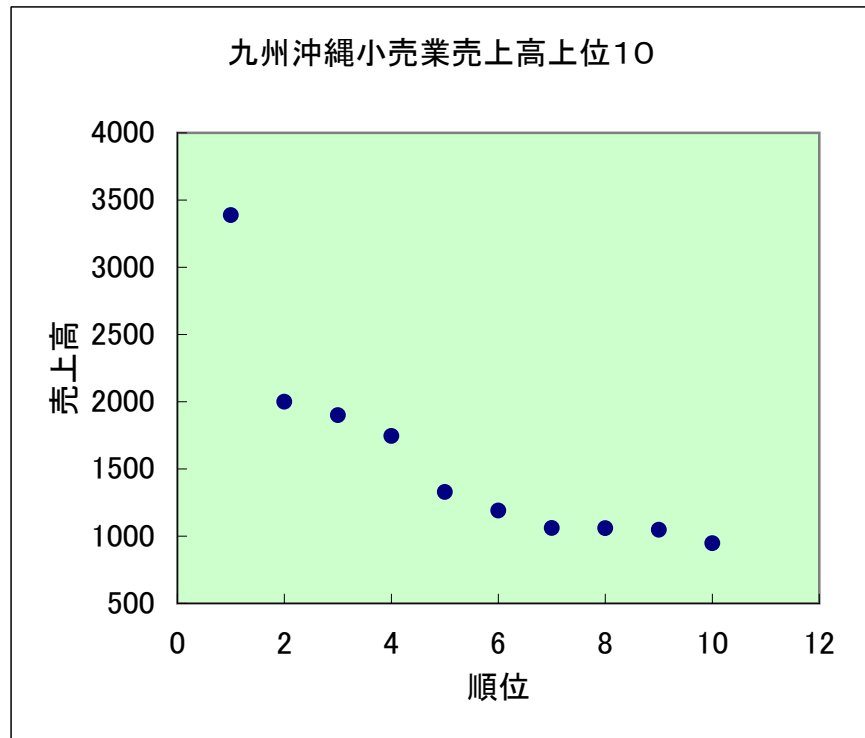




データ解析への入門（大和 元）- 講義資料 -  
- Zipfの法則、回帰の例として - 新規テキスト

九州沖縄 小売業売上高ランキング（南日本9月29日朝刊：2006年）

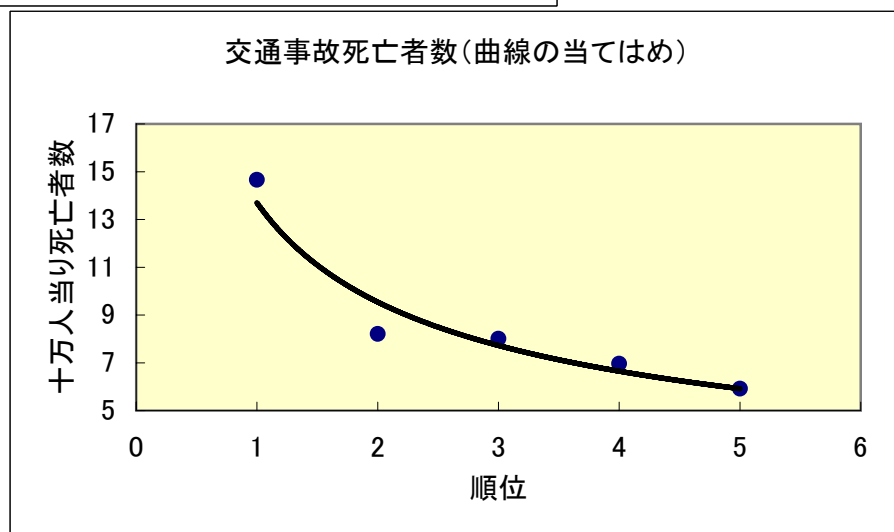
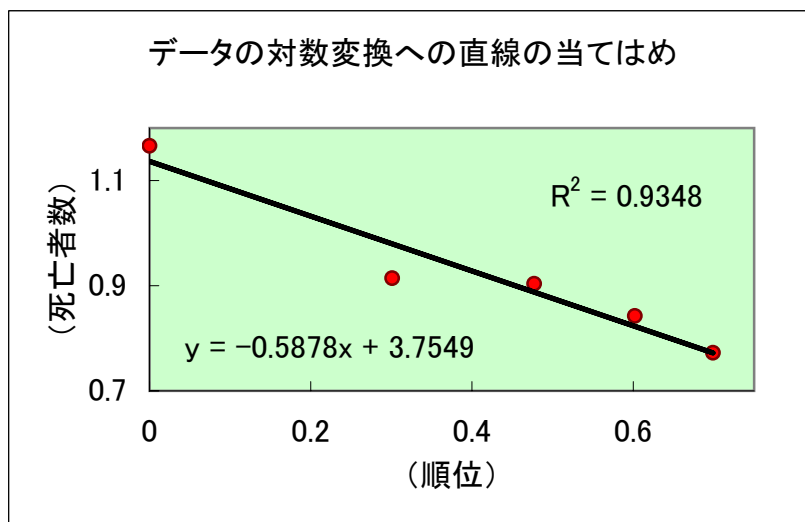
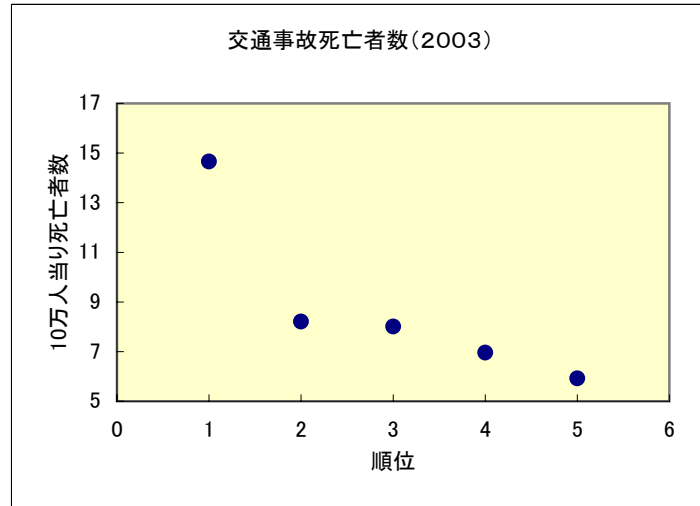
順位	売上高(億円)	
1	3388	ベスト電器
2	2000	
3	1900	
4	1745	
5	1328	タイヨー
6	1190	
7	1061	
8	1060	
9	1048	
10	948	



データ解析への入門（大和 元） - 講義資料 -  
 - Zipfの法則、国別交通事故死亡者数への当てはめ

順位	10万人当り死亡者数		対数変換	順位	死亡者数
1	14.66	U.S.A.		0	1.16613397
2	8.21	Australia	0.30103		0.91434316
3	8.01	Germany	0.477121		0.90363252
4	6.96	Japan	0.60206		0.84260924
5	5.92	Sweden	0.69897		0.77232171

(朝日データ年鑑 2006 から)

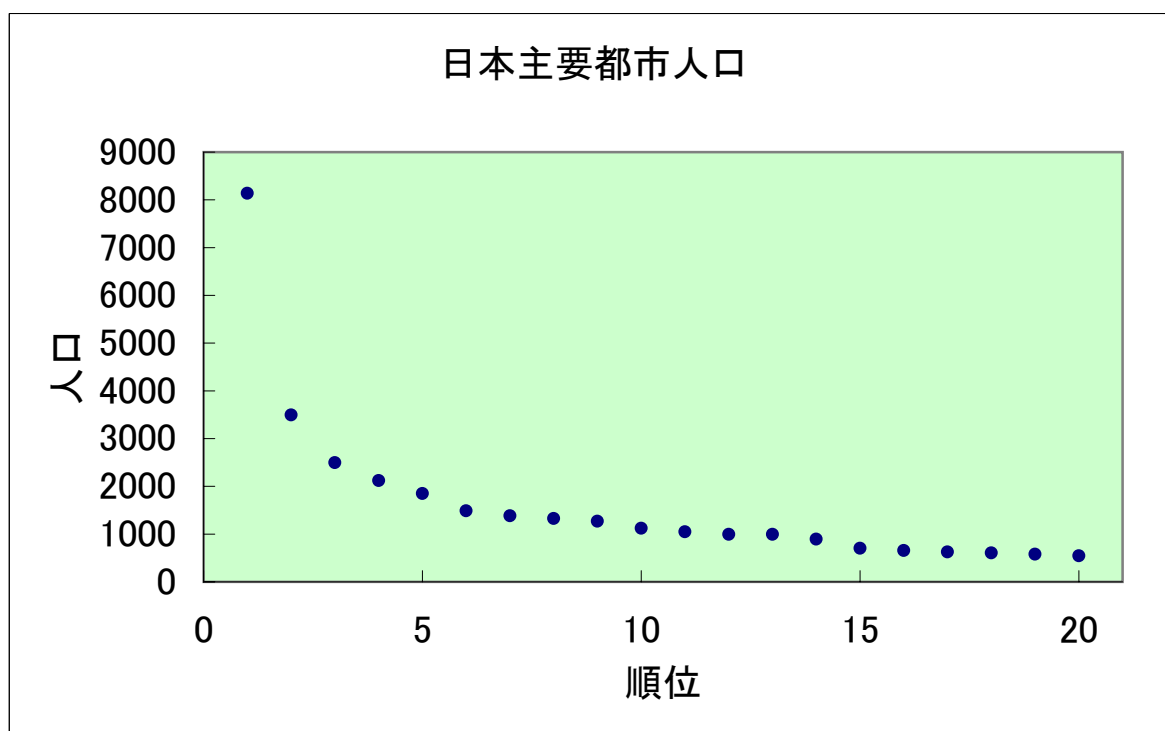


データ解析への入門（大和 元）：講義資料  
 -Zipfの法則、日本の主要都市人口への考察-

主要都市 順位	別人口 人口・単位 (1,000人)	H16年3月	対数 (順位)	変換 (人口)
1	8138	東京特別区部	0	3.9105177
2	3495	横浜	0.30103	3.5434472
3	2496	大阪	0.4771213	3.3972446
4	2123	名古屋	0.60206	3.32695
5	1850	札幌	0.69897	3.2671717
6	1489	神戸	0.7781513	3.1728947
7	1386	京都	0.845098	3.1417632
8	1327	福岡	0.90309	3.1228709
9	1270	川崎	0.9542425	3.1038037
10	1123	広島	1	3.0503798
11	1048	さいたま	1.0413927	3.0203613
12	994	北九州	1.0791812	2.9973864
13	994	仙台	1.1139434	2.9973864
14	895	千葉	1.146128	2.951823
15	703	静岡	1.1760913	2.8469553
16	657	熊本	1.20412	2.8175654
17	628	岡山	1.2304489	2.7979596
18	608	相模原	1.2552725	2.7839036
19	579	浜松	1.2787536	2.7626786
20	547	鹿児島	1.30103	2.7379873

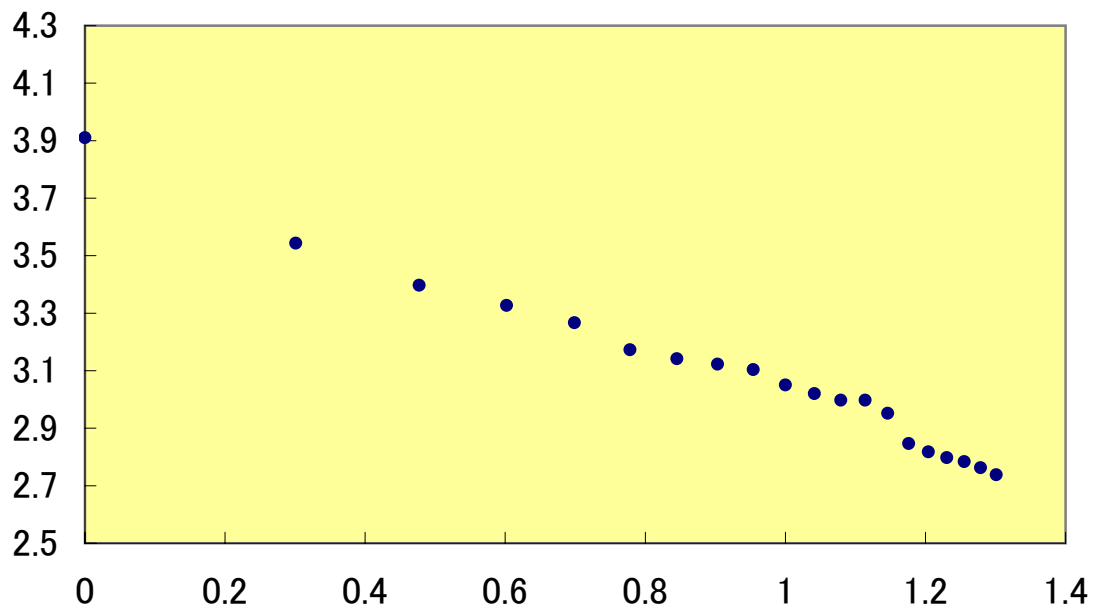
(日本の統計 2005 総務省統計局から)

(寄与率) 0.9811136

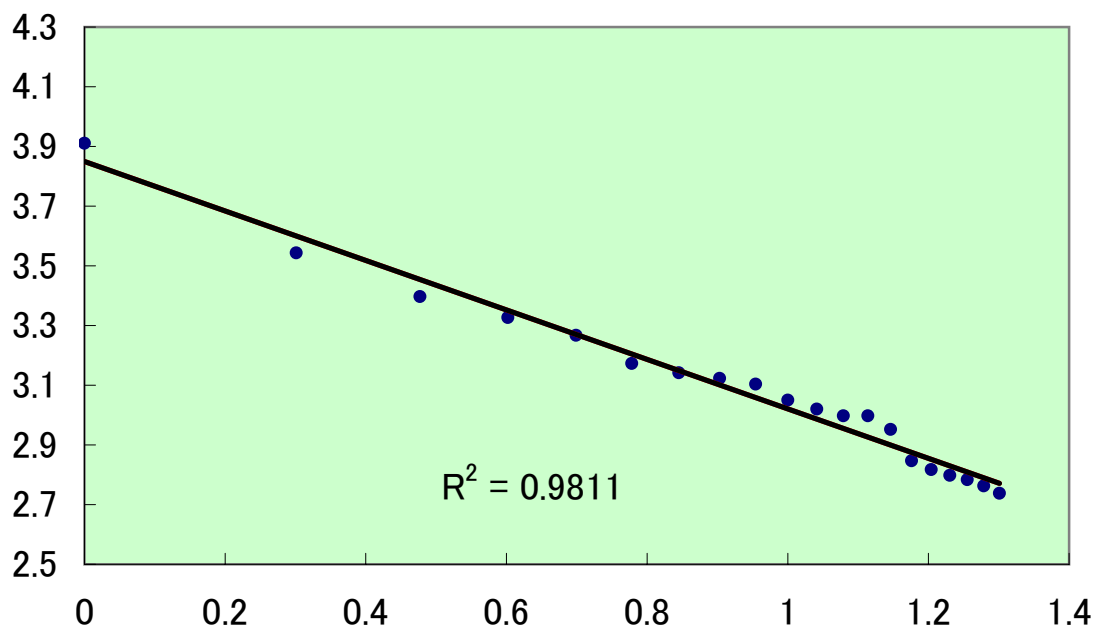




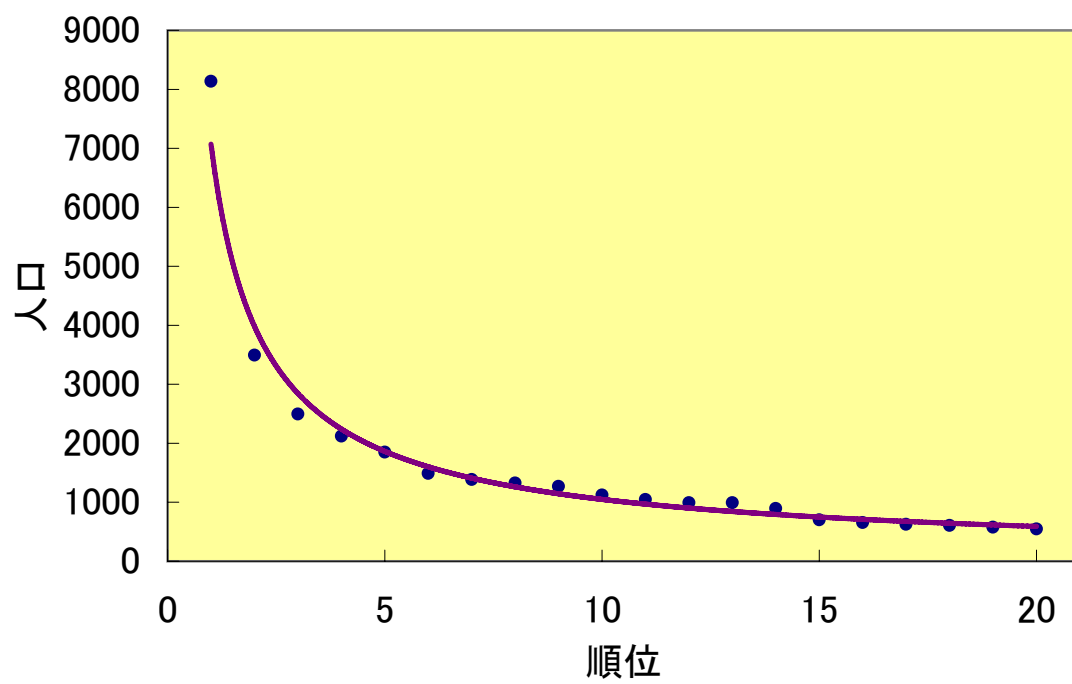
対数変換: 順位(横軸)、人口(縦軸)



直線の当てはめ(変換: 順位(横)、人口(縦))



### 元のデータに戻って

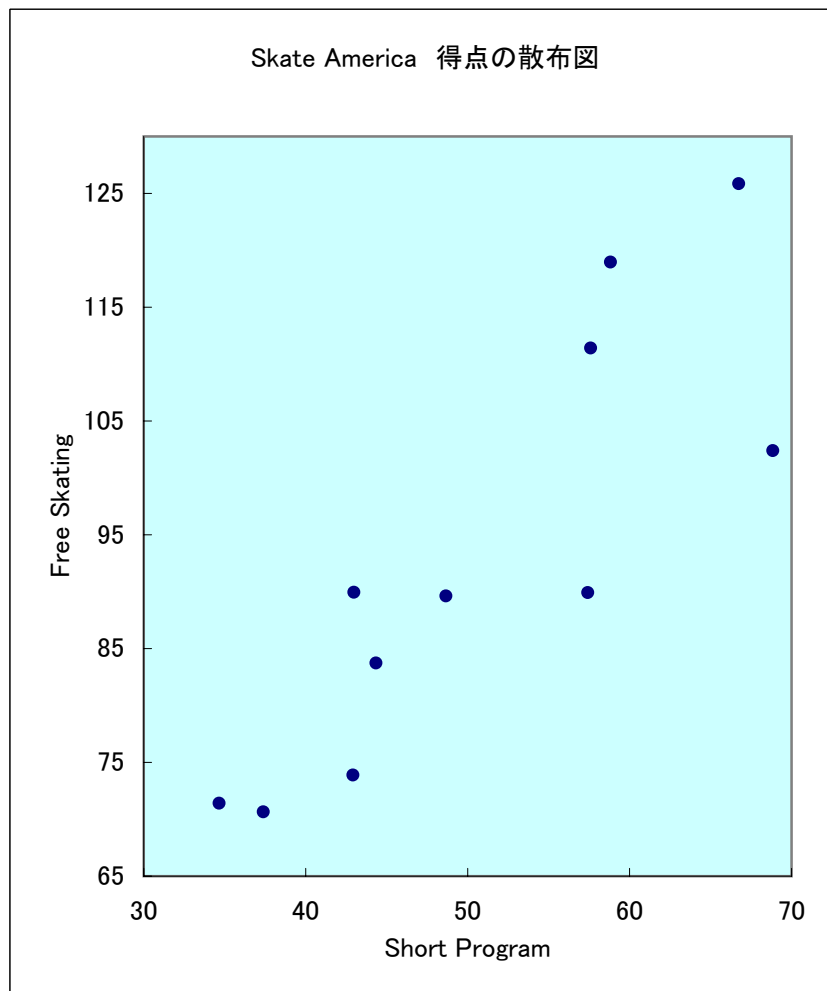


2007/2/24

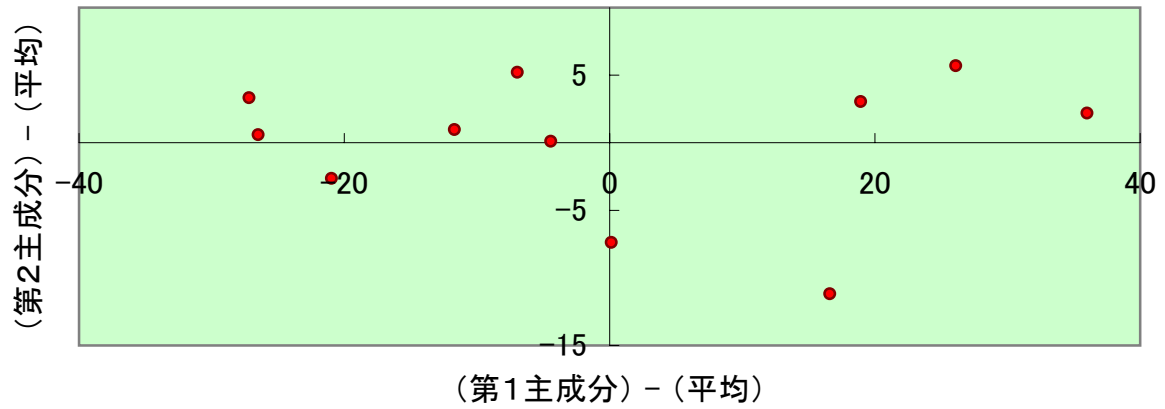
データ解析への入門（大和 元）：講義資料  
スケートアメリカ(ISU GRAND PRIX Oct-06)への主成分分析

ISU	GRAND Oct-06 x1 Short Program	PRIX Skate x2 Free Skating	Figure Skating America (Total)	第一 主成分	第二 主成分
Miki ANDO	66.74	125.85	192.59	142.39983	3.839117
Kimmie MEISSNER	58.82	118.96	177.78	132.50419	7.342653
Mao ASADA	68.84	102.39	171.23	123.01116	-9.53478
Sarah MEIER	57.6	111.41	169.01	125.33141	4.688948
Emily HUGHES	57.42	89.92	147.34	106.53555	-5.73065
Mai ASADA	48.66	89.63	138.29	101.97187	1.7523
Kiira KORPI	42.98	89.95	132.93	99.455023	6.854288
Mira LEUNG	44.34	83.73	128.07	94.70977	2.609224
Valentina MARCHEL	42.92	73.89	116.81	85.445092	-0.9974
Michelle CANTU	37.38	70.66	108.04	79.906834	2.235586
Katy TAYLOR	34.66	71.41	106.07	79.221071	4.972488

（データはISUのホームページから）



主成分散布図



Skate America (主成分説明)

