

欢迎访问 6sigma 品质网

site: <http://www.6sq.net/>

bbs: <http://bbs.6sq.net/>

6 品质网，原名龙在天涯成立于 2001 年 10 月，我们定位于一个关于 SPC, FMEA,DOE, Six Sigma,品质统计，品质体系，品质管理,企业管理的综合网站！

我们致力于提高国内企业界对品质的认识，以及为中国企业培养大量优秀的品质人才，为改善中国企业产品品质，打造中国企业卓越品质形象而努力。目前它的发展良好，吸引了一大批优秀的品质，统计,企业管理，6sigma 爱好者和专业人士！

我们的目标是建立国内最好的品质/质量管理的专业网站。本站及时提供大量的最新专业文章，以及相关的资料下载，内容丰富，我们提供的就是你想要的，另外本站为企业和个人提供各种有偿服务与无偿服务（如销售各种统计软件和相关书籍，企业咨询，培训等）。站内还设立了一个 6 品质论坛供大家交流，大批的质量工程师，企业管理人士和 6sigma 黑带不断加盟其中。

图的力量

作者: **ilikespc**

定量分析技术由一些统计手法组成，它会产生数字或表格形式的输出。通常包括以下内容：

- 假设检验
- 方差分析
- 点的估计和置信区间
- 最小二乘回归

这些技术以及类似的技术都是非常有效的，并且是经典分析中的主流。尤其是计算机迅速发展的今天，广泛的统计软件的应用，大大降低了定量分析的难度，从而使其得到普遍的应用，极大提高了我们的质量分析工作的水平。而且很多时候，分析工作的具体过程是由计算机自动完成，我们要做的任务只是简单地理解它的输出的含义。

但是事物总有两面性的，在下面的例子中我们将展示它的不足的地方。这是Anscombe构造的一个简单、经典的关于图的用途的例子，用来展示它洞察数据的功效。

我们先看第一组数据：

X1	Y1
10.00	8.04
8.00	6.95
13.00	7.58
9.00	8.81
11.00	8.33
14.00	9.96
6.00	7.24
4.00	4.26
12.00	10.84
7.00	4.82
5.00	5.68

如果我们分析的目标是计算简单的统计量以及测定Y和X之间的最佳线性拟合，那么统计软件输出可能的结果如下：

$N = 11$

Mean of X = 9.0

Mean of $Y = 7.5$

Intercept = 3

Slope = 0.5

Residual standard deviation = 1.237

Correlation = 0.816

以上的简单定量分析在一般场合非常有用，但请注意它只给了我们关于数据的有限观察信息。

我们再来看其他三组数据[Anscombe data sets 2, 3, and 4]:

X2	Y2	X3	Y3	X4	Y4
10.00	9.14	10.00	7.46	8.00	6.58
8.00	8.14	8.00	6.77	8.00	5.76
13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.77	9.00	7.11	8.00	8.84
11.00	9.26	11.00	7.81	8.00	8.47
14.00	8.10	14.00	8.84	8.00	7.04
6.00	6.13	6.00	6.08	8.00	5.25
4.00	3.10	4.00	5.39	19.00	12.50
12.00	9.13	12.00	8.15	8.00	5.56
7.00	7.26	7.00	6.42	8.00	7.91
5.00	4.74	5.00	5.73	8.00	6.89

对第2组数据进行定量分析，结果如下：

$N = 11$

Mean of $X = 9.0$

Mean of $Y = 7.5$

Intercept = 3

Slope = 0.5

Residual standard deviation = 1.237

Correlation = 0.816

看起来和第1组数据是一样的。有些人就会天真地以为按统计告诉我们的结果，两组数据看起来是一样的；但你是否注意到还存在统计没有告诉我们的信息呢？

让我们再来看对第3、4组数据的有趣的定量分析结果

$N = 11$

Mean of $X = 9.0$

Mean of $Y = 7.5$

Intercept = 3

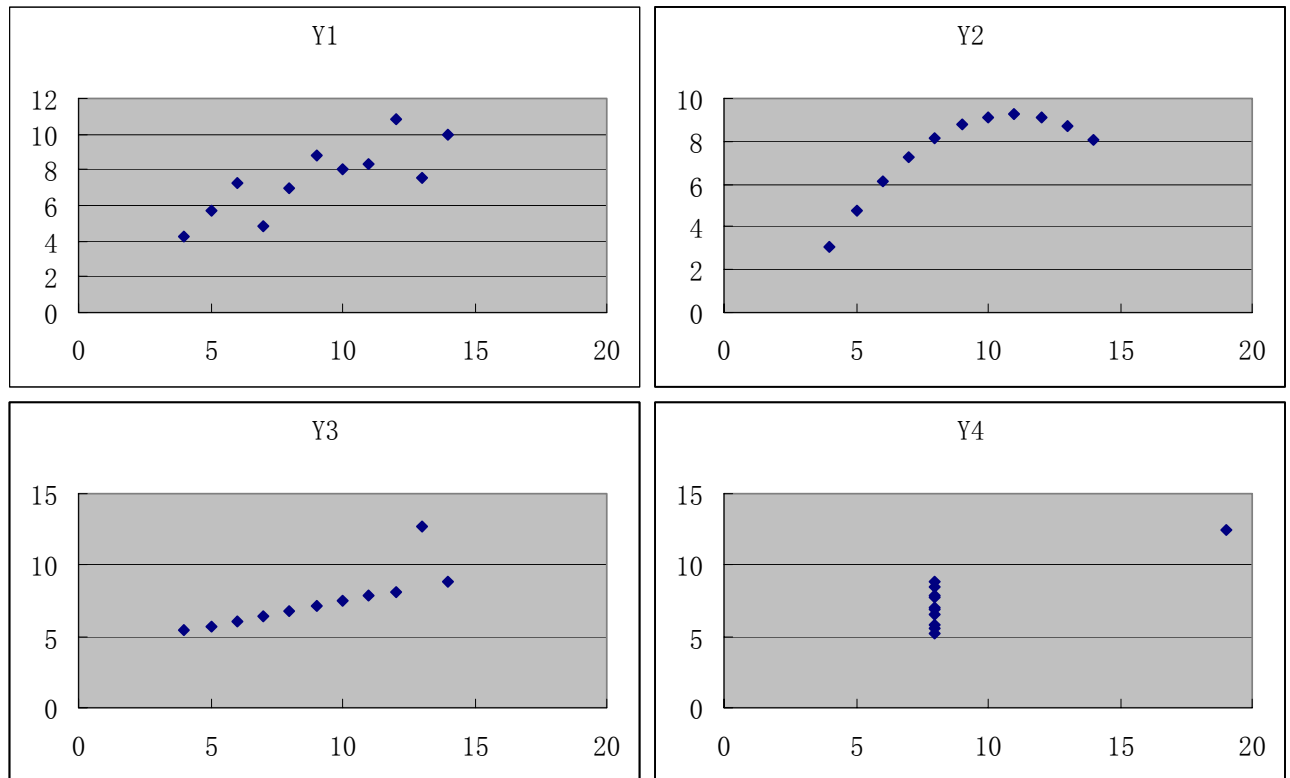
Slope = 0.5

Residual standard deviation = 1.236

Correlation = 0.816 (0.817 for data set 4)

这意味着按定量的结果，似乎所有四组数据是差不多的。但实际上

这些数据是非常不一样的。通过散点图我们可以看的很清楚。



定量分析本质上来说没有错，但它可能是不够完全的。因为它们是对数据进行汇总，而汇总的操作过程一般来说是通过数据的计算处理，将其转化为有限的几个数字输出，而这些数字在把握数据的某些特征方面（如平均值、斜率、截距、相关性等）做的很棒。不过这样做同时也是对数据进行了筛选，它必然抛弃了一些东西，而有时也许是很重要的信息。这会导致我们对数据观察不够全面，糟糕的时候甚至是误导。对第2、3、4组数据来说，截距（=3）和斜率（=0.5）就是一种误导，因为这样的估计只是对线性模型有效，但在以上场合这样的分析中却是致命的错误。

简单地说定量分析是非常有用的工具，但只有你将它用在正确的地方，才能真正发挥它的用途。单纯地迷信它，为了计算而计算是非常危险的。而图的技术是原始数据信息的完整体现，就能帮助我们进行整体把握，初步得到直观的印象，建立比较正确的数据模型，然后在此基础上在结合定量分析技术，才能有效地完成品质分析工作。