

# 不同属性的 Pokémon 强度分析

计 91 刘程华 20180116897

清华大学计算机系

日期：2021 年 1 月 9 日

# 目录

<b>1</b>	<b>统计问题的背景与提出</b>	<b>3</b>
<b>2</b>	<b>数据收集与预处理</b>	<b>3</b>
2.1	原始数据 . . . . .	3
2.2	数据预处理 . . . . .	3
<b>3</b>	<b>数据预分析</b>	<b>4</b>
3.1	定量展示 . . . . .	4
3.2	定性分析 . . . . .	4
3.3	数据再处理 . . . . .	7
<b>4</b>	<b>数据建模与模型诊断</b>	<b>9</b>
4.1	假设检验 . . . . .	9
4.1.1	对称性检验 . . . . .	9
4.1.2	正态性检验 . . . . .	10
4.1.3	等方差检验 . . . . .	11
4.1.4	数据特征分析 . . . . .	11
4.2	单总体分析 . . . . .	12
4.3	配对比较 . . . . .	12
<b>5</b>	<b>总结</b>	<b>13</b>
<b>6</b>	<b>附录</b>	<b>14</b>
6.1	代码 . . . . .	14
6.2	参考资料 . . . . .	14

# 1 统计问题的背景与提出

宝可梦（日语：ポケットモンスター<sup>①</sup>、ポケモン，英语：Pokémon、Pocket Monsters，旧译“神奇宝贝”，“宠物小精灵”，“口袋妖怪”等）是一个跨媒体制作的作品系列。其作品中的各种各样可爱的宠物小精灵有着我们童年美好的回忆，也给我们留下了深刻的印象。在精灵宝可梦的世界中，所有的宝可梦都用属性来划分。比如雷系属性的皮卡丘，水系属性的杰尼龟，草系属性的妙蛙种子，火系属性的小火龙。至最新的宝可梦游戏第八世代“宝可梦剑/盾扩展票”，属性一共有 18 个类型，有的宝可梦是单属性，有的宝可梦是双属性。这些属性一定程度上反映了宝可梦的特点和性格，在对战中也起到了关键的作用。可以说，属性是兼具趣味性和竞技性的分类，是宝可梦世界观中的重要组成部分。令我好奇的是，不同属性之间的宝可梦强度差异大吗？是否存在某种属性的宝可梦比于其他属性的宝可梦强？

注意到宝可梦精灵的总数并不多但种类却多达 18 种，因此每种属性的宝可梦数据量并不大；宝可梦各种数据的分布情况对我们来说也是未知的。出于以上两点考虑，本文采用非参的方法对宝可梦属性的强度进行研究和分析。

## 2 数据收集与预处理

### 2.1 原始数据

原始数据来自网站 <https://www.kaggle.com/abcsds/pokemon>。此数据集是关于神奇宝贝游戏的 721 个神奇宝贝，包括它们的编号，名称，第一和第二属性以及基本统计信息：HP，攻击，防御，特殊攻击，特殊防御和速度。对于这些数据的解释如下：

- ID：每个宠物小精灵的 ID，我们不使用这些数据。
- Name：每个宠物小精灵的名字，我们不使用这些数据。
- Type 1：每个宠物小精灵的属性。
- Type 2：某些宠物小精灵是双重属性。
- Total：此后得出的所有统计信息的总和，这是关于宠物小精灵强度的一般指南。
- HP：生命值，指宠物小精灵击倒前可以承受的伤害量。
- Attack：普通攻击的伤害值。
- Defense：抵抗普通攻击的基础伤害抵抗。
- Sp.atk：特殊攻击的伤害值。
- Sp.Def：抵抗特殊攻击的基础伤害。
- Speed：确定每轮首先攻击的是哪个宠物小精灵。
- Generation：宠物小精灵对应首次出现版本，我们不使用这些数据。
- Legendary：是否是传说级宝可梦。

### 2.2 数据预处理

在进行分析研究之前，我们需要对数据进行必要的预处理。预处理的主要流程包括：

- (1) 选用 total 作为我们衡量宠物小精灵强度的依据。一方面 total 是各种的属性值的加和，能够反应宠物小精灵各纬度的信息；另一方面 total 值是综合能力的体现，不受设计师设计不同属性对应不同特点的干扰。值得一提的是，在实际游戏和比赛中，total 值是衡量宠物小精灵强度极为重要的依据，有着很广的应用基础。
- (2) 删去传说级宝可梦的样本。在共 800 组数据中的，有 65 组数据是传说级宝可梦，他们的 Total 值都在 580 以上，显著高于其他样本。而且在游戏和动画中大多作为剧情角色出现起到推动游戏剧情作用，不参与玩家之间的战斗。除此之外每个属性都有传说级宝可梦，因此我们删去传说级宝可梦有助于我们问题的研究。
- (3) 双属性宠物小精灵的处理。我们注意到有的宠物小精灵具有双重属性，第一属性往往代表了其设计的原型与自身特点。相对的，第二属性往往是宝可梦随着自身进化/适应环境获得的全新特点。具有双重属性的宠物小精灵相较于单属性的宠物小精灵在战斗中的优劣情况有明显不同（优势/劣势对局中优势/劣势会放大，但总体仍均衡）。为了便于我们分析，我们把具有双重属性宠物小精灵当成两个单属性的宠物小精灵。这样处理是合理的，这是因为我们关心的是不同属性的宠物小精灵的强度差异。
- (4) 分类。我们把宠物小精灵按照属性值分成 18 类：Bug , Dark , Dragon , Electric , Fairy , Fighting , Fire , Flying , Ghost , Grass , Ground , Ice , Normal , Poison , Psychic , Rock , Steel , Water。

处理过后的数据见“Pokemon.txt”文件。

### 3 数据预分析

这里我们通过定量和定性的方法来展示数据，对数据的情况进行初步的了解。并对数据根据分析结果做进一步的调整。

#### 3.1 定量展示

对于定量展示，这里计算了每个组别的样本均值和样本标准差并给出了最小值、第 1 四分位数、中位数、第 3 四分位数和最大值。从表 (1) 中我们不难发现各属性 total 值标准差均在 100 左右差异较小，平均值在 400 上下差异较大。除此之外，存在不同属性的 total 的分位数有较明显的差异，例如钢系 (steel) 的各分位数都显著高于草系 (grass)。

#### 3.2 定性分析

我们分别画出 18 个组别的箱型图 (1)(2) 与直方图 (3)(4)(5)。

我们可以通过箱图和直方图对各个组别的分布获得直观认识：

- 异常值：我们可以通过箱型图来观察数据中异常值的情况。通过箱型图我们可以发现，样本中异常值非常少，在战斗系 (fighting) 和钢系 (steel) 中各有一个。由于我们已经在预处理中对传说级宝可梦进行了剔除，因此异常值很少是符合预期的。我们可以认为数据的质量比较好。

	size	mean	s.d.	min	1st quartile	median	3rd quartile	max
Bug	72	379.5	116.3557	194.0	281.0	392.5	477.2	600.0
Dark	48	448.9	104.9609	220.0	350.2	477.0	510.0	700.0
Dragon	34	479.1	129.7487	245.0	374.0	492.0	600.0	700.0
Electric	45	427.6	98.27314	205.0	330.0	460.0	515.0	610.0
Fairy	37	396.1	120.1615	190.0	300.0	420.0	480.0	618.0
Fighting	49	457.1	105.5779	210.0	405.0	474.0	525.0	630.0
Fire	56	442.9	103.3658	250.0	357.5	465.0	528.5	634.0
Flying	86	424.2	106.4319	244.0	345.8	430.0	499.8	700.0
Ghost	43	421.8	91.78853	236.0	335.0	455.0	494.0	600.0
Grass	92	412.2	99.78224	180.0	323.8	409.0	494.0	630.0
Ground	62	424.4	107.3015	210.0	328.5	427.5	507.2	700.0
Ice	33	440.7	98.56295	250.0	334.0	480.0	521.0	594.0
Normal	100	396.2	107.9686	190.0	297.5	414.0	480.2	670.0
Poison	56	397.7	99.75669	195.0	318.5	395.0	490.0	625.0
Psychic	86	433.8	112.442	198.0	322.5	455.0	505.0	700.0
Rock	43	436.9	92.48262	280.0	355.0	435.0	498.8	700.0
Steel	92	472.8	99.68055	300.0	423.0	489.5	525.0	700.0
Water	62	419.6	102.7619	200.0	325.0	430.0	500.0	640.0
Total	1109	423.7	107.3773	180.0	330.0	435.0	500.0	700.0

表 1: 均值方差分位表

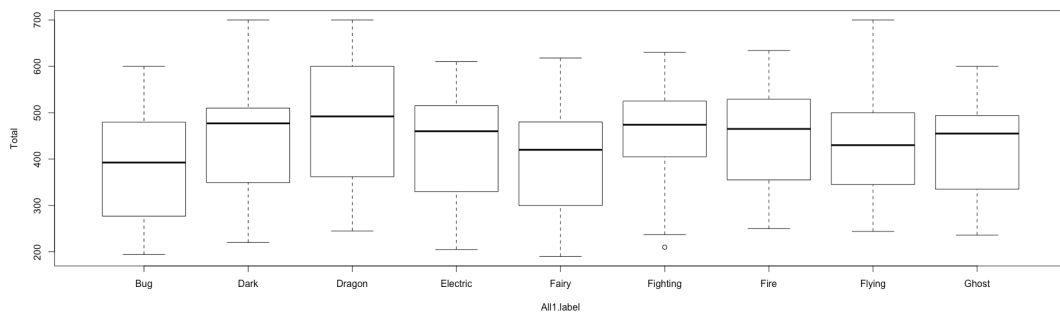


图 1: 箱型图 a

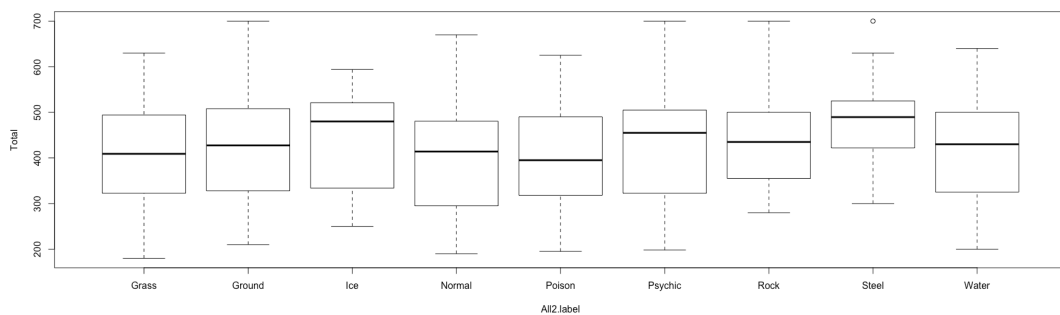


图 2: 箱型图 b

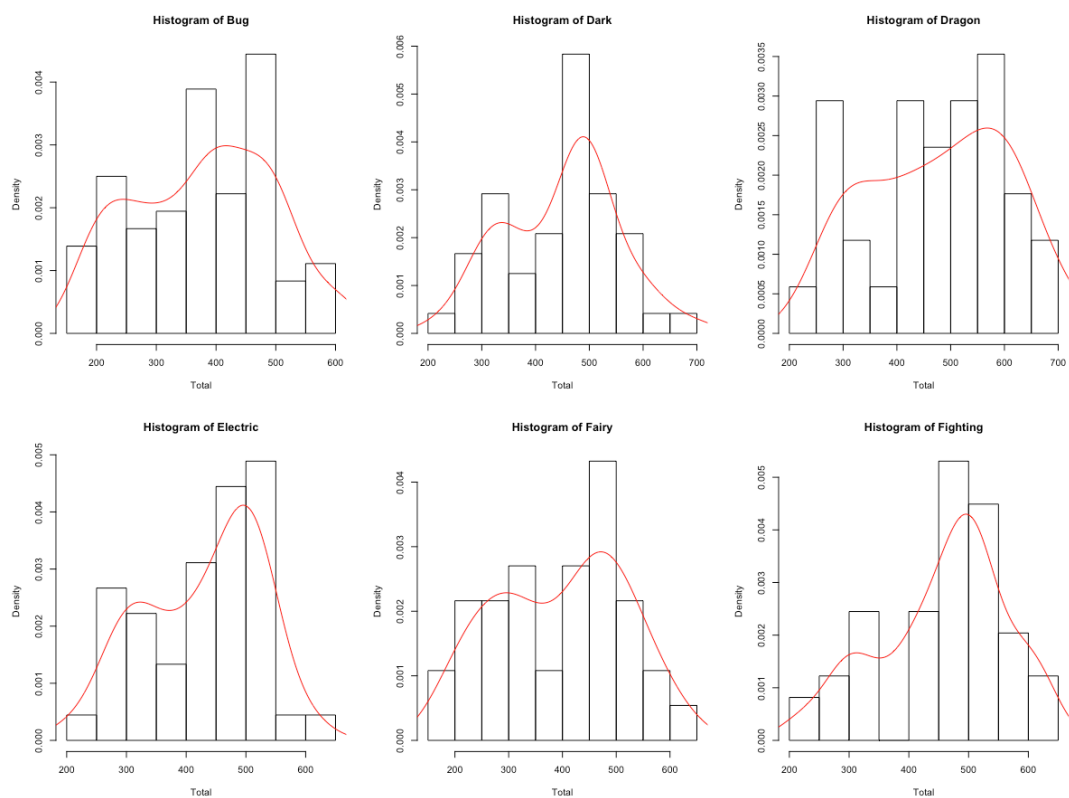


图 3: 直方图 a

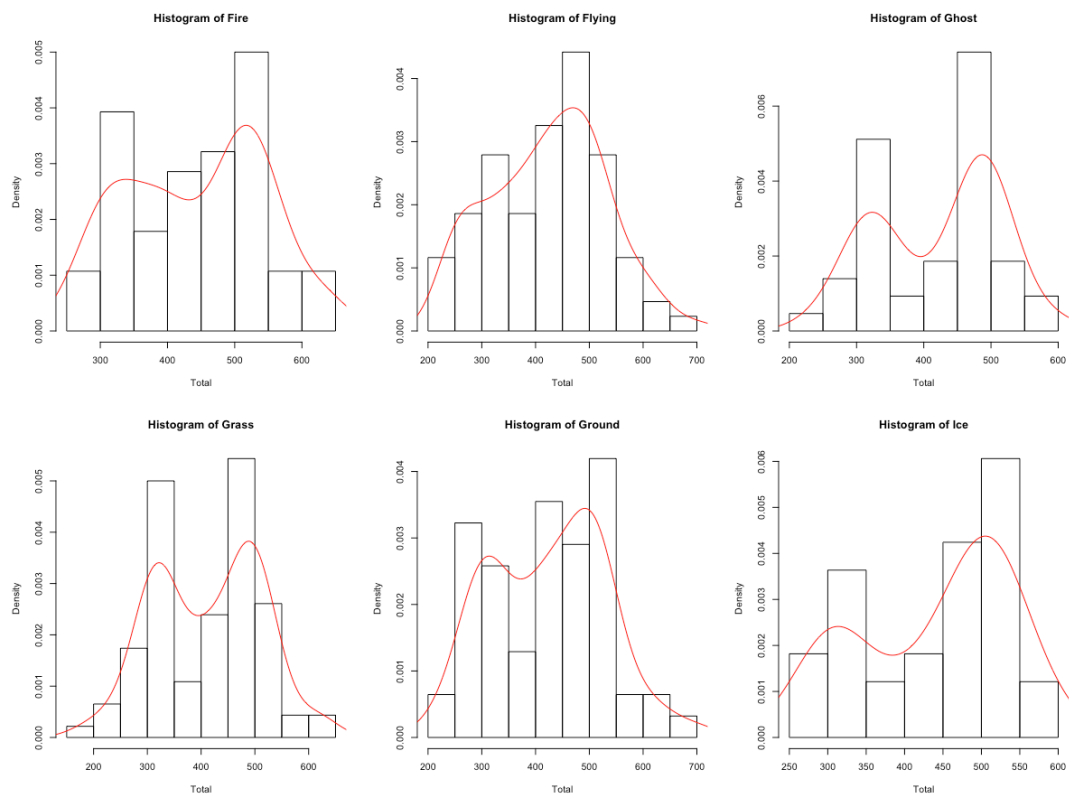


图 4: 直方图 b

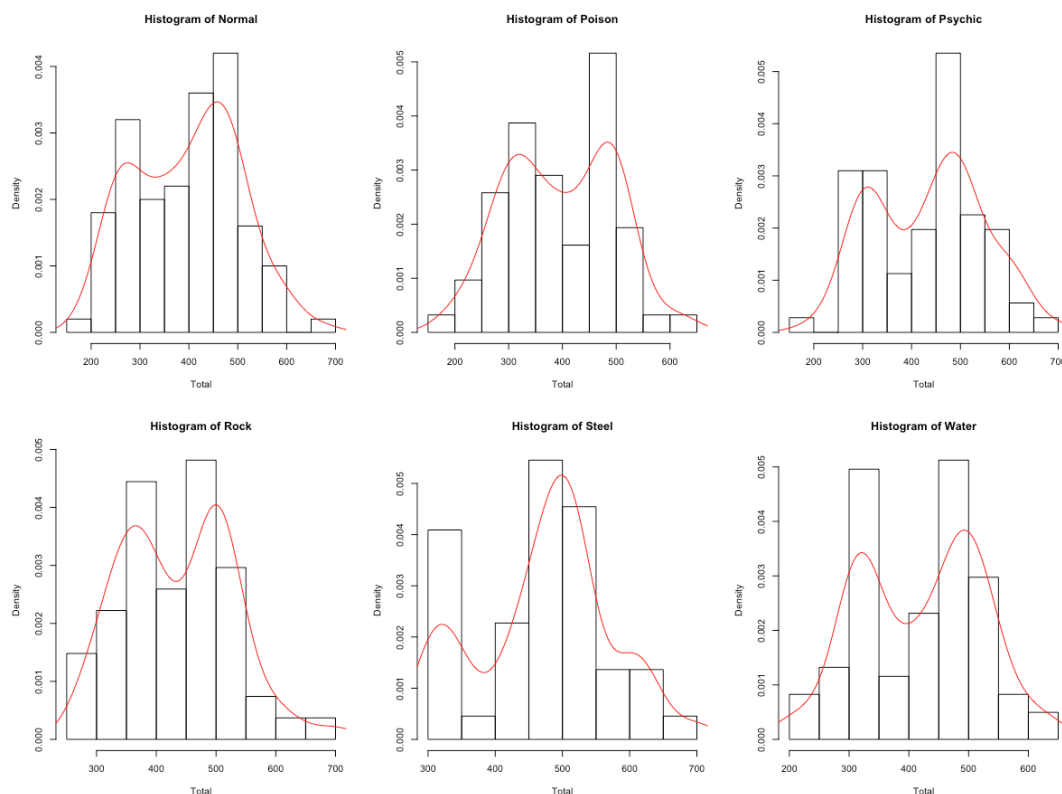


图 5: 直方图 c

- 相似性：从箱型图可以看出，有一些属性的 total 值是很相似的。例如草系（grass）、一般系（normal）、水系（water）和虫系（bug）。
- 对称性：在后面的非参数统计分析中，我们可能会用到分布的对称性信息。从箱型图可以看出，暗黑系（dark）、冰系（ice）、雷系（electric）、幽灵系（ghost）对称性肯恩较差，从直方图也可以验证这一点。另外从直方图中我们可以发现龙系（dragon）、地面系（ground）和钢系（steel）的对称性也较差。
- 正态性：从直方图直观来看，大部分数据呈双峰分布，不符合正态分布的特征。同时，我们在一张图上比较 18 个属性的密度曲线 (6)。可以较为直观的印证上面的观察。

### 3.3 数据再处理

通过上面的分析我们可以发现有多组数据具有较为相同的分布特征，注意到我们的属性数有 18 种，全部两两分析显然是不太现实的。回顾我们的的目标是对宝可梦各种属性的平衡性进行研究，因此我们可以结合 total 值的分布情况和实际讨论中各属性人气高低来选取六个代表的属性来进行平衡性的分析。我们选择如下六个属性来分析平衡性：龙系（dragon）、雷系（electric）、火系（fire）、草系（Grass）、钢系（steel）和水系（water）。

为了进一步了解数据的特点，验证我们选用 total 值作为强度合理性，我们尝试对所选数据进行主成分分析。考虑到宠物小精灵的各项其他数据数值（hp、attack、defense...）在数量级上并无较大差异，我们不对数据做标准化处理。我们得到崖底碎石图 (7)。

我们可以看到，第一主成分占据了总方差的一半，我们查看第一主成分的荷载：

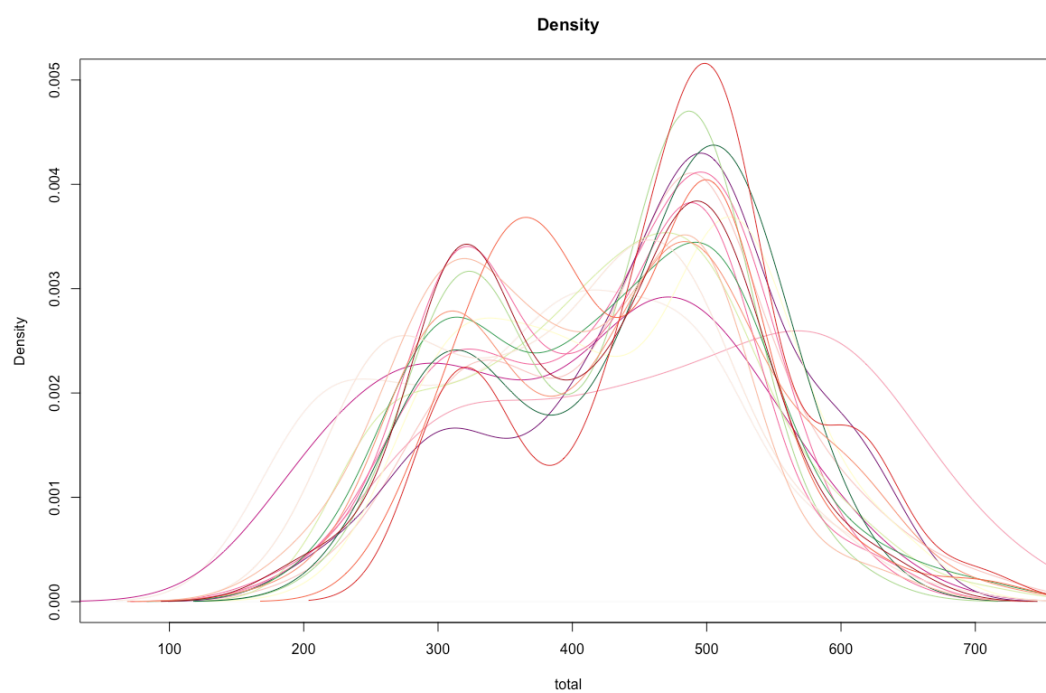


图 6: 密度图

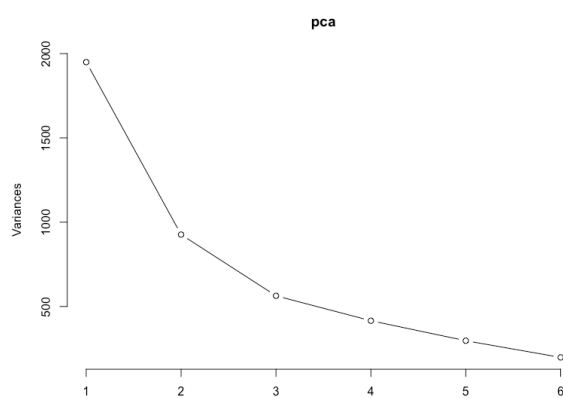


图 7: 崖底碎石图



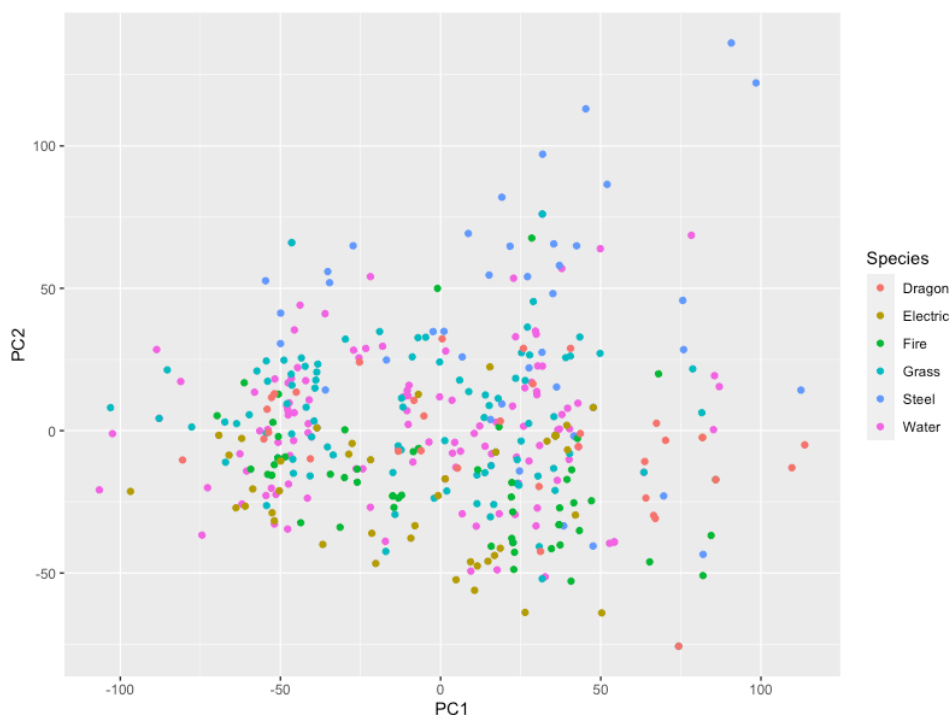


图 8: 分类图

$$e'_1 = [0.3045422, 0.4955158, 0.4858002, 0.4801963, 0.36705790, 0.2457565]$$

可以发现第一主成分的负载是较为平均的，这表明各个属性值都对差异性起到了作用，因此在此考虑总体强度差异时，我们需要较为平均的考虑宠物小精灵的各项数值。这和我们选择 `total` 作为数据是相吻合的。为了进一步分析，我们选取前两个主成分进行绘图得到图 (8)。发现前两者的对属性的区分度并不好，各种属性的点杂糅在一起分布，直观上来说，这说明不同属性的宠物小精灵强度上没有明显的巨大差异。这也与我们前面的结果相吻合。

## 4 数据建模与模型诊断

这一部分中，我们对处理后的数据进行统计分析。我们先用非参数检验的方法，预检验我们的数据是否能够满足各种后续检验所需要的假设：如正态性、对称性，然后选择合适的方法对六个群体进行单总体分析和两两比较，用图的方式给出这些总体之间的关系。

### 4.1 假设检验

由数据预分析我们知，各个属性的 `total` 值的标准差相差很小，为了验证这一点我们还会进行等方差检验。之后，我们将对模型的对称性和正态性进行检验。

#### 4.1.1 对称性检验

我们考虑将样本关于中位数进行切割，得到大于中位数和小于中位数的两个样本。然后我们对这两个中位数进行 Kolmogorov-Smirnov Test。我们选分布对称作为零假设，分布不对称作为备择假设。得到  $p$  值见表 (2)。我们注意到在显著性水平  $\alpha = 0.05$  下，除了雷系外都没有否定

type	p
Dragon	0.1715682039521903
Electric	0.008433935512456414
Fire	0.10497765593914699
Grass	0.6669293690780796
Steel	0.19074239756721156
Water	0.19074239756721156

表 2: 对称性检验 a

type	Modify test with $n = 1200$
Dragon	0.828
Electric	0.272
Fire	0.209
Grass	0.335
Steel	0.294
Water	0.796

表 3: 对称性检验 b

原假设。我们查看之前预分析的图像，发现我们大多是双峰的。考虑到对于双峰采取中位数作为切割点会有较大的输入敏感性，在借鉴了上一届优秀 project（自 63 鲍文轩学长）的对称性检验方法后，我们对检验方法进行了如下改进：在单谷区间的采用一维搜索算法选取使得 KS 距离  $D$  最小的对称中心  $m$ ，计算对应的 KS 距离  $D$ ，然后用蒙特卡洛方法选取合适的间值，来判断是否拒绝原假设。改进后的假设检验见表 (3)。在显著性水平  $\alpha = 0.05$  下，都没有否定原假设。

#### 4.1.2 正态性检验

下面我们采用 Lilliefors 和 Shapiro-Wilk 检验对模型的正态性进行检验，见表 (4)。从结果可以看到，在显著性水平  $\alpha = 0.05$  下，两种检验的结果是较为一致的。只有在火系上出现了差异。而雷系、草系、水系都显著偏离了正态性假设。为了进一步观察样本的正态性，我们决定绘制 QQ 图 (9)。

从 QQ 图我们可以验证我们前面两种检验的结果，与他们是相符的。对于前面存在差异的火系和未被拒绝原假设的龙系和钢系，我们从对应的 QQ 图可以看到，他们的左尾和右尾偏离

type	Lilliefors	Shapiro-Wilk
Dragon	0.1753	0.07722
Electric	0.02654	0.03663
Fire	0.008675	0.01007
Grass	0.000284	0.008296
Steel	0.06828	0.05627
Water	0.00009556	0.001056

表 4: 正态性检验

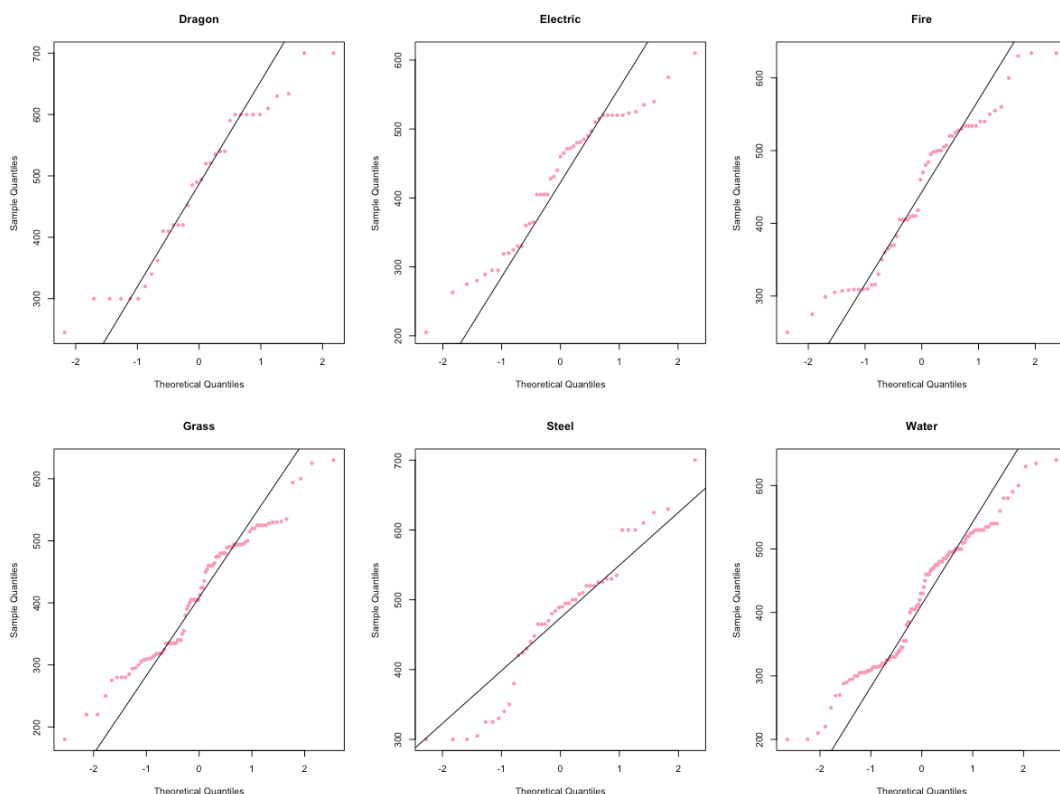


图 9: QQ 图

了正态性。

一般来说，我们可以对数据进行指数变换、对数变换或者 Box-Cox 变换使数据满足正态性假设。但是我们的数据有六组，并且偏离的情形不太一样，水系和钢系的左尾和右尾的偏离情形是相反的（见图 (9)）。因此没有办法对六组数据统一处理使其满足正态性。出于这一点考虑，在后面的单总体检验和双总体检验中，我们采取非参的办法。

#### 4.1.3 等方差检验

从前面预分析部分我们可以发现这六种属性 total 值的方差是相差不大的。我们来检验数据的方差齐性。由于偏离正态分布，我们放弃 Bartlett 检验。从直方图中我们可以看到多个属性呈双峰分布，同时借助 QQ 图我们有足够证据说明数据偏离正态但分布。所以这里我们非参的 Fligner-Killeen 检验不选择效果不如它的 Levene 检验。由于 Fligner-Killeen 检验需要样本量相同，我们对样本量大的一组做随机无放回抽样处理。我们得到表 (5)。从结果可以看到，在显著性水平  $\alpha = 0.05$  下，我们认为不同属性宝可梦的 total 值方差是相同的。

#### 4.1.4 数据特征分析

通过以上的检验，我们发现数据有着同方差性和对称性，但正态性并不满足。数据满足对称性和同方差性时，中位数和平均值相差较小，并且两者能较好的反映数据的整体水平。因此，在下面的单总体均值中位数分析时，均值和中位数是对某一种属性宝可梦整体强度非常有效的估计。

p.value	Dragon	Electric	Fire	Grass	Steel	Water
Dragon	/	0.4851	0.5598	0.3946	0.3322	0.408
Electric	/	/	0.2642	0.6342	0.5795	0.3079
Fire	/	/	/	0.8221	0.6132	0.2884
Grass	/	/	/	/	0.3874	0.5468
Steel	/	/	/	/	/	0.3202
Water	/	/	/	/	/	/

表 5: 方差齐性检验

	Sign Test		Wilcoxon Test		t Test	
	median	conf.int	(pseudo)median	conf.int	mean	conf.int
Dragon	492	[410 ,590]	476	[430,527.5]	479.0588	[ 433.7874,524.3303]
Electric	460	[405 ,485]	427	[399.5,465]	427.6	[398.0755 ,457.1245]
Fire	465	[ 405 ,500]	442.5	[414.5,471.5]	442.9464	[415.2649 ,470.6280]
Grass	409	[380 , 461]	410	[394,432.5]	412.2065	[ 391.5422 ,432.8708]
Steel	489.5	[465 ,510]	477.5	[440,505]	472.7955	[442.4898,503.1011]
Water	430	[400 ,470]	417.5	[402,437.5]	419.5537	[401.0572 ,438.0502]

表 6: 单总体分析

## 4.2 单总体分析

这一小节我们来研究每种属性自身，获取其 total 值的均值和中位数的置信区间。我们采用 Sign Test 和 Wilcoxon Test 得到各组 total 值中位数置信水平为 95% 的置信区间。另外我们通过 t Test 给出每个组平均值置信水平为 95% 的置信区间。但是正态性假设有的组并不满足，因此采用的 t Test 得到估计和置信区间可靠性较差。我们得到表 (6)。

根据检验结果我们可以认为各组数据的方差是相同的。综合三个检验的结果，我们可以发现六种属性的 total 值的大致关系 (Dragon, Steel) > (Electric, Fire) > (Grass, Water)。但是，这并不意味着他们之间具有显著性差异。为了明确差异是否显著，在下一节中我们会进行配对比较。

## 4.3 配对比较

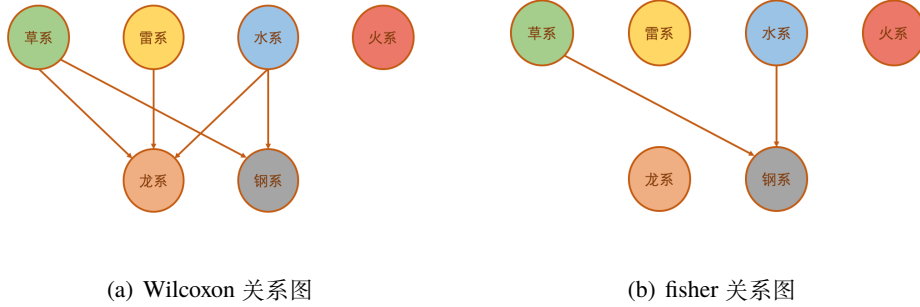
为了明确差异是否显著，我们采用 Wilcoxon Test 和 Fisher Test 对六组数据进行两两配对比较。我们计算每一对比较的 p 值，得到表 (7) 和表 (8)。

p.value	Dragon	Electric	Fire	Grass	Steel	Water
Dragon	/	0.04482483	0.1734147	0.007599996	0.7775488	0.01647009
Electric	/	/	0.372334	0.3849255	0.05315747	0.6641199
Fire	/	/	/	0.05026156	0.322271	0.1415031
Grass	/	/	/	/	0.002083804	0.5436796
Steel	/	/	/	/	/	0.008314149
Water	/	/	/	/	/	/

表 7: Wilcoxon Test

p.value	Dragon	Electric	Fire	Grass	Steel	Water
Dragon	/	0.4936	0.5146	0.3157	1	0.3314
Electric	/	/	1	0.3595	0.2864	0.7272
Fire	/	/	/	0.8655	0.5459	0.8719
Grass	/	/	/	/	0.01005	0.7815
Steel	/	/	/	/	/	0.02225
Water	/	/	/	/	/	/

表 8: Fisher Test



我们注意到两种方法得到的结果较为一致，草系和钢系、水系和钢系在两种检验中  $p$  值均小于 0.05。而在 Wilcoxon Test 中龙系和草系、龙系和雷系、龙系和水系的  $p$  也都值小于 0.05。如果我们关注的问题是整体是否有显著性差异，我们还需对显著性水平进行修正。由于我们不知道数据之间的关系，我们采用最保守的 Bonferroni 联合置信区间策略。修正后， $\alpha^* = \frac{\alpha}{6} = \frac{0.05}{6} = 0.0083$ 。我们可以得到结论，在显著性水平  $\alpha = 0.05$  下，从整体来看，不同属性之间 total 值存在显著性差异。值得一提的是，如果我们关注的是两两之间的差异，则不需对显著性水平进行修正。通过配对比较，我们可以给出六种属性宠物小精灵强度的偏序关系图 (10(a))(10(b))。可以看出，fisher 关系图是 Wilcoxon 关系图的支撑子图，这说明在这种情形下，Fisher Test 比 Wilcoxon Test 更加谨慎，这也与我们的预期相符。

## 5 总结

在明确研究内容后，我们首先对数据进行了预处理提高了数据的质量，在进行探索性分析后我们注意到有些属性具有相似的分布，结合我们的研究问题和实际应用中的情形，我们选取了六个代表属性作为我们的研究对象并对我们采用 total 值作为强度体现的合理性进行了分析和说明。为了明确模型的特征方便以后的检验，我们依次对模型的对称性、正态性和同方差性进行了检验。然后我们分别对单总体属性的均值和中位数进行了估计，由于模型的同方差假设和对称性假设成立，因此均值和中位数对强度的估计具有较好的参考作用。最后我们对不同属性的宠物小精灵两两配对进行显著性差异检验，并得到了显著性偏序关系图。

我们可以发现存在两种属性之间的 total 值有显著性差异，这很大程度上说明了存在两种属性他们的强度是不同的。回到宠物小精灵设定本身，有些属性是较为稀有的属性，例如龙系和钢系，他们的强度高于其他属性在也存在着一定的合理性。由于本报告没有考虑不同属性之间

的相互克制的复杂情况，在具体战斗中的平衡性仍有待明确。

## **6 附录**

### **6.1 代码**

程序主要由 R 语言编写，在进行对称性检验时采用了 python 语言。具体代码见文件。

### **6.2 参考资料**

- [1] 清华大学统计学研究中心《非参数统计导论》课程 PPT 周在莹 2020 年秋
- [2] 往届优秀 project 自 63 鲍文轩