

葡萄酒的评价

摘要

本文讨论的是葡萄酒的评价问题，通过建立不同模型，分析出了两组评酒员的评价结果有无显著性差异，对酿酒葡萄进行了分级，并找出了酿酒葡萄与葡萄酒的量化指标间的联系，最后把芳香物质加入指标对酿酒葡萄与葡萄酒的指标关系进行论证。

对于问题一，要求根据附件 1 的数据，判断两组评酒员的评价结果有无显著性差异。通过建立分组 T 检验和成对数据检验两个模型，得出若分别考虑红、白葡萄酒，则认为两组评酒员的评价结果都存在显著性差异，若对红、白葡萄酒综合考虑时，则认为两组评价员的评价结果无显著性差异。最后，以方差为参考指标，得出第二组评酒员的评价结果更可信。

对于问题二，要求根据酿酒葡萄的理化指标和葡萄酒的质量对酿酒葡萄进行分类，采用系统聚类分析法，运用 SPSS 软件得出谱系聚类图 and 不同分类数时的分类结果。分类完成后，以每个类所包含样品的质量得分的均值作为参考，对类进行分级，具体结果见正文。

对于问题三，要求分析酿酒葡萄与葡萄酒的理化指标间的联系，我们建立了偏最小二乘法回归分析模型——综合主成分分析与经典的多元回归分析。首先对附件 2 中酿酒葡萄的 30 个一级理化指标进行主成分分析，之后利用多元回归得出葡萄酒的 9 个一级理化指标关于酿酒葡萄的 30 指标的回归方程，具体结果见正文。

对于问题四，首先求出酿酒葡萄、葡萄酒的一级指标与葡萄酒质量的四个方面的相关系数，分析得出分别影响葡萄酒外观、口感、香气、整体的主要指标；为了论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量，建立加入芳香物质和未加入芳香物质时各指标与可信质量得分的多元统计回归模型，对回归值与评酒员的评分结果的差值是进行 T 检验，得出不能仅仅用葡萄和葡萄酒的理化指标来评价葡萄酒的质量的论证，同时，又给出了合理的评价的方法。

关键字： 分组数据检验 聚类分析 偏最小二乘模型 芳香物质

一、 问题重述

确定葡萄酒质量时一般是通过聘请一批有资质的评酒员进行品评。每个评酒员在对葡萄酒进行品尝后对其分类指标打分，然后求和得到其总分，从而确定葡萄酒的质量。酿酒葡萄的好坏与所酿葡萄酒的质量有直接的关系，葡萄酒和酿酒葡萄检测的理化指标会在一定程度上反映葡萄酒和葡萄的质量。附件 1 给出了某一年份一些葡萄酒的评价结果，附件 2 和附件 3 分别给出了该年份这些葡萄酒的和酿酒葡萄的成分数据。请尝试建立数学模型讨论下列问题：

1. 分析附件 1 中两组评酒员的评价结果有无显著性差异，哪一组结果更可信？
2. 根据酿酒葡萄的理化指标和葡萄酒的质量对这些酿酒葡萄进行分级。
3. 分析酿酒葡萄与葡萄酒的理化指标之间的联系。
4. 分析酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响，并论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量？

二、 问题假设

- 1.附件中的数据真实的反映葡萄酒和酿酒葡萄的变量值；
- 2.芳香物质对香气的贡献率相同；
- 3, 各个评分员的评分公平公正，对不同样本的标准相同。

三、 符号说明

a_{i0} ：葡萄酒 i 的质量分数， $i=1,2,\cdots,55$ ；
 a_i^{mn} ： $i=1,2,\cdots,55$; $m=1$ 或 2 ; $n=1,2,\cdots,10$ ；
 X_i ：第一组评酒员对样品 i 的评价结果均值， $i=1,2,\cdots,55$ ，前 27 个对应红葡萄酒，后 28 个对应白葡萄酒；
 Y_i ：第二组评酒员对样品 i 的评价结果均值， $i=1,2,\cdots,55$ ，前 27 个对应红葡萄酒，后 28 个对应白葡萄酒；
 D_i ：两组评酒员对样品 i 的评价结果均值的差值， $i=1,2,\cdots,55$ ；
 s_i ：第 i 组评酒员的评酒结果方差；
 d_{ij} ：样品 i 与样品 j 间的距离， $i,j=1,2,\cdots,55$ ；
 ω_i ：酿酒葡萄的第 i 个一级指标的距离权重， $i=1,2,\cdots,31$ ；

$x_{ij}^{(1)}$: 红葡萄样品 i 的第 j 个一级指标的数值, $i=1,2,\dots,27; j=1,2,\dots,30$;
 $x_{ij}^{(2)}$: 白葡萄样品 i 的第 j 个一级指标的数值, $i=1,2,\dots,28; j=1,2,\dots,30$;
 $y_{ij}^{(1)}$: 红葡萄酒样品 i 的第 j 个一级指标的数值, $i=1,2,\dots,27; j=1,2,\dots,9$;
 $y_{ij}^{(2)}$: 白葡萄酒样品 i 的第 j 个一级指标的数值, $i=1,2,\dots,28; j=1,2,\dots,8$;
 n_i : $n_1=27, n_2=28$ 分别表示红葡萄与白葡萄酒的样品数;
 $\hat{y}_i^{(1)}$: 未加入芳香物质指标时葡萄酒样品 i 质量得分拟合值, $i=1,2,\dots,55$;
 $\hat{y}_i^{(2)}$: 加入芳香物质指标时葡萄酒样品 i 质量得分拟合值, $i=1,2,\dots,55$ 。

四、 问题分析

(一) 问题一的分析

问题一要求我们分析附件 1 中两组评酒员的评价结果有无显著性差异。确定葡萄酒质量时, 每组评酒员有 10 名成员, 红葡萄酒和白葡萄酒各取 27 个样品, 每位评酒员分别对每个样品进行评分。采用两种模型对此问题进行分析。

模型一中, 考虑到若把 55 组评价结果整合在一起与另外一组评酒员比较, 进行显著性差异检验时, 由于每组评酒员内部的 55 组数据之间就存在显著性差异, 不服从正态分布, 必定会影响检验结果。为了排除样品实际质量不同而对显著性差异检验造成的影响, 把每组评酒员对一种样品的 10 个评价结果形成一个样本, 对两组样本进行显著性差异分析, 共检验 55 次。得出结果后, 只要其中存在显著性差异的样本达到一定比率时, 就认为两组评酒员的评价结果存在显著性差异。

模型二中, 运用成对数据的检验方法, 首先以两组 10 个评酒员的评价结果均值 X_i, Y_i 作为该组对一种酒的评分, 得到 55 组独立的值。对数据进行处理, 令 $D_i = X_i - Y_i$, 则 D_i 是由同一因素引起的, 服从正态分布。若两组评价结果无显著差异, 则 D_i 的均值为 0, 故检验时检验 D_i 的均值是否为 0 即可印证两组评酒员是否存在显著性差异。

若求出两组评酒员的评价结果有显著性差异, 则需比较哪一组更可信。设每一种葡萄酒样品的实际质量分数为 a_{i0} 。由于 a_{i0} 未知, 利用参数估计的替换原理: 用样本均值来代替总体均值的思想, 本题中每组 10 个评酒员的评价结果即为样本, 理论上评酒员对其的评分都应该在均值左右小范围内波动, 即数据越集中, 评价的可信度越高。故我们可以用比较两组评酒员评价结果的平均方差来分析它们的可信度, 方差越小, 评分的离散程度越低, 可信度越高。

(二) 问题二的分析

问题二要求我们根据酿酒葡萄的理化指标和葡萄酒的质量对这些酿酒葡萄进行分级, 我们采用聚类分析法解决此题。附件 2 中给出的理化指标分为一级指标和二级指标两种, 其中酿酒葡萄的一级指标为 28 个, 二级指标为 33 个。显然理化指标非常多, 对酿酒葡萄进行初步分类很难做到, 故采用常用的聚类分析方法——系统聚类法, 它开始时每个对象自成一类, 然后每次讲最相似两类合并。数据的变换方法采用标准化变换, 样品间的距离采用最广泛的距离——欧式距离, 类之间的距离采用最短距离法、最长距离法, 中心距离法、离差平方和法分别计算比较后, 采用其中分类均匀相对合理的一种。最后给出谱系聚类图和分类个数不同时的具体分法。

对于一级和二级指标, 我们发现二级指标是一级指标的具体内容, 如对于一级理化

指标氨基酸总量，其后紧跟的 17 个二类指标为各种氨基酸，它们的信息已经合并体现在一级指标中了。若一级、二级指标同时采用，不但增加了变量的个数，问题复杂，而且由于收指标间相关性的影响，样品距离的计算值会与实际值产生偏差，故操作时我们只采用 28 个一级理化指标。

另外，在计算样品间的距离时，我们必须考虑到酿酒葡萄的理化指标与葡萄酒质量的权重问题。题目中明确提到：酿酒葡萄的好坏与所酿葡萄酒的质量有直接的关系。而反应葡萄酒的质量的只有一个值，即两组评酒员对葡萄酒评价结果的均值，而酿酒葡萄涉及到 28 个理化指标，多所有数据标准化处理之后，若无权重限制，葡萄酒质量对距离的贡献率很小，不符合题意。

（三）问题三的分析

问题三要求分析酿酒葡萄与葡萄酒的理化指标之间的联系，采用偏最小二乘回归模型。由于酿酒葡萄的理化指标太多，造成信息有所重叠，并且使得在高维空间中研究样本的分布规律比较复杂，为了用最少的指标尽可能多的反映原来变量的信息，需要用到主成分分析的思想。其次，为了建立酿酒葡萄与理化指标之间的联系，需要建立两者间的回归关系。

（三）问题四的分析

要求我们论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量，对于这个问题，我们想通过对葡萄和葡萄酒的理化指标和葡萄酒质量的得分进行拟合，酿酒葡萄的理化指标比较多，既有一级指标，又有二级指标，其中一级指标的数目大约是 30 个，考虑到一级理化指标的数目很多，而且二级理化指标对于葡萄酒质量影响不大，所以我们在分析的过程中只取酿酒葡萄理化指标中的所有的一级指标，葡萄酒的理化指标相对的少一点，也是既有一级指标，又有二级指标，其中一级指标大约是 9 个，在分析的过程中，我们也是只取葡萄酒理化指标中的所有的一级指标，并且，我们又对这些一级理化指标进行主成分分析，得到了较少的变量，然后我们利用这些较少的变量替代原来的一级理化指标进行下面的回归拟合。

五、模型建立与求解

（一）问题一的模型建立与求解

5.1.1 异常数据的处理

观察分析附件 1 中的数据，发现第一组白葡萄酒品尝评分中有 3 处数据出现异常。我们对它们做近似处理如下：

（1）酒样品 3 中，评酒员 7 对口感分析的持久性表中显示数据为 77，而此指标的满分为 8，显然错误。认为是录入错误，更改为 7。

（2）酒样品 8 中，评酒员 9 对口感分析的浓度表中显示数据为 16，而此指标的满分为 8，显然错误。我们以其他评酒员对此指标的评分均值 6 填补此值。

（3）样品酒 14 中，评酒员 8 对口感分析的持久性评分显示为空，我们以其他评酒员对此指标的评分均值 6 填补此值。

5.1.2 各个样品检验

问题一要求我们分析附件 1 中两组评酒员的评价结果有无显著性差异。从附件 1 中可以看出，确定葡萄酒质量时，每组评酒员有 10 名成员，红葡萄酒和白葡萄酒各取 27

个样品，每位评酒员分别对每个样品进行评分。

首先，需要确定进行显著性差异检验时的样本。按照直接理解的结果，若需要对两组评酒员的评价结果进行显著性差异的判断，则需要与每组评酒员对应各提取一个样本，即需要每组评价的 55 组结果进行综合。但是，由于每个样品的实际质量存在差别，必定导致同一评酒员对不同样品的评价结果有明显差异，从而一组评酒员对每种样品的总体评价结果存在差异，进行显著性差异检验时，由于每组评酒员内部的 27×2 组数据之间就存在显著性差异，必定会影响检验结果。如第一组评酒员对红葡萄样品 1 和红葡萄样品 23 的评分结果，均值分别为 62.7、85.6，方差为 92.9、32.5 显然两组数据在离散程度和均值方面都存在明显差异性。而评酒员在品评时，不同评酒员对同一样品的评分围绕该样品的实际质量波动，并无显著性差异。为了排除样品实际质量不同而对显著性差异检验造成的影响，把每组评酒员对一种样品的 10 个评价结果形成一个样本，对两组样本进行显著性差异分析，共检验 55 次。

根据评酒规则：每个评酒员对葡萄酒进行品尝后对其分类指标打分，然后求和得其总分。故 m 组的评酒员 i 对红葡萄酒样品 j 的评价结果计算式为：

$$a_i^{mn} = \sum_{j=1}^{10} a_{ij}^{mn}, i=1,2,\dots,55; m=1 \text{ 或 } 2; n=1,2,\dots,10; \quad (1.1)$$

j 表示 10 个分类指标。则每次显著性差异检验分析时的样本为：

$$(a_i^{11}, a_i^{12}, \dots, a_i^{110}) \text{ 和 } (a_i^{21}, a_i^{22}, \dots, a_i^{210}) i=1,2,\dots,55$$

其次，需要确定进行显著性差异检验的方法。由于两组评酒员的评价结果是相互独立的，且两组评价结果的方差不能确定，故可以采用两独立样本的 t 检验进行分析。运用此方法需特别注意：样本来自的总体必须服从正态分布。实际中，样本来自的总体可以看成无数个评酒员对每个样品葡萄酒的评分，由于葡萄酒自身的指标一定，其质量有一真实值即为所有评分的期望，所有评酒员对一个样品的评分在其左右波动，若评酒员足够多，所有评分均匀分布在期望的两边，即认为样本来自的总体服从正态分布。

设组 1 对葡萄样品 i 的评分均值为 μ_{1i} ，组 2 对葡萄样品 i 的评分均值为 μ_{2i} ， $i=1,2,\dots,54$ 则运用 T 检验时的检验假设为：

$$H_0: \mu_{1i} = \mu_{2i}, H_1: \mu_{1i} \neq \mu_{2i} \quad (I)$$

然后，分别利用 SPSS 对两组评酒员的评分对象——54 组样品的评价结果进行显著性检验分析。对于不同的样品，两组评酒员的评价结果有无显著性差异的结果可以得出。

最后，需要根据检验结果判断两组评酒员的评价结果有无显著性差异，认为两组评酒员的评价结果存在显著性差异的样品占有比例不超过一定比例时认为他们之间是不存在显著性差异的。

算法步骤如下：

第一步：评价结果汇总。根据附件 1 中每个评酒员对每个样品的分类指标分数求其最后得分，在 EXCEL 中用函数 sum 即可求得，结果在附录表 1。

第二步：运用 SPSS 软件，对 54 组评价结果进行两独立样本显著性差异检验。SPSS 的输出结果中首先是采用 F 检验对方差的齐次性。若方差想相等，直接采用一般的 T 检验，即第一行输出结果；若方差不等，采用校正的 T 检验，即第二行输出结果。

根据以上输出结果取舍原则，两组评酒员对 54 组样品酒的评分显著性差异检验结果见附录，列出红葡萄酒前 9 个样品显著性水平如下：

表 1 两组评酒员对前 9 个红葡萄酒评分显著性差异减压结果

样品号	1	2	3	4	5	6	7	8	9
-----	---	---	---	---	---	---	---	---	---

显著性水平 sig 值	0.213	0.016	0.05	0.511	0.67	0.056	0.146	0.073	0.19
-------------	-------	-------	------	-------	------	-------	-------	-------	------

检验过程中设置的置信区间为 95%，且此为双侧检验，故当 sig 小于 0.05 时，认为两组评价结果存在显著性差异。如表中两组评酒员对 2、3 号红葡萄酒的评价结果都存在显著性差异。

第三步：计算两组评分存在显著性差异的样品比例。统计结果可知，对 27 个红葡萄酒样品和 28 个白葡萄酒而言，27 组红葡萄酒样品中有 7 组，而白葡萄酒中只有 2 组，则差异性占有率约为 1/6，两组评价结果的差异性不大。

下面用另一种模型进一步分析。

5.1.3 成对数据的检验

有时为了比较两种产品、两种方法等的差异，我们常在相同的条件下做对比试验，得到一批成对的观察值，然后分析观察数据，这种方法称为逐步对比法。虽然在模型 (I) 中我们讨论不能使用不同样品的评分直接分析两组评酒员的显著性差异，但我们可以对数据预先处理，时得出的数据服从正态分布，继而使用 T 检验进行显著性差异分析。

设组 1 对葡萄酒 i 的评分为 X_i ，组 2 对葡萄酒 i 的评分为 Y_i ，其中：

$$X_i = \frac{1}{10} \sum_{n=1}^{10} a_i^{1n}, \quad Y_i = \frac{1}{10} \sum_{n=1}^{10} a_i^{2n}, \quad i=1,2,\dots,55$$

则 27 组样品红葡萄酒和 28 组样品白葡萄酒最后得出 55 组相对独立的评价结果：

$(X_1, Y_1), (X_2, Y_2), \dots, (X_{27}, Y_{27}), (X_{28}, Y_{28}), \dots, (X_{55}, Y_{55})$ 。

令 $D_1 = X_1 - Y_1, D_2 = X_2 - Y_2, \dots, D_{55} = X_{55} - Y_{55}$ ，则 D_1, D_2, \dots, D_{55} 相互独立。又由于 D_1, D_2, \dots, D_{55} 是由于同一原因：组 1 和组 2 的评酒员的评酒结果差值引起的，可认为他们服从同一分布。假设 $D_i \sim (\mu_D, \sigma_D^2), i=1,2,\dots,55$ 。这就是说 D_1, D_2, \dots, D_{55} 构成正态总体 $N(\mu_D, \sigma_D^2)$ 的一个样本，其中 μ_D, σ_D^2 未知。

根据我们需要求解的内容：两组评价员的评价结果有无显著性差异。当它们之间无显著性差异时，认为 X_i 与 Y_i 相等，则 $D_i = 0$ ，故我们的假设性检验为：

$$H_0: \mu_D = 0, \quad H_1: \mu_D \neq 0; \quad (\text{II})$$

分别记 D_1, D_2, \dots, D_{55} 得样本均值和样本方差为 \bar{d}, s_D^2 。则此检验问题的拒绝域为（检验水平为 α ）：

$$|t| = \left| \frac{\bar{d}}{s_D / \sqrt{n}} \right| \geq t_{\alpha/2}(n-1)$$

计算 t 时白葡萄酒与红葡萄酒分别单独和综合起来计算，数据如下：

表 1 成对数据检验法结果

	\bar{d}	s_D^2	t	n	$t_{0.025}(n-1)$	是否接受 H_0
红葡萄酒	2.833	5.83	2.53	27	2.056	否
白葡萄酒	-2.557	5.792	-2.336	28	2.0518	否
综合	-0.089	5.65	-0.117	55	2.014	是

从上表可以得出结论，当分别考虑红葡萄酒与白葡萄酒酒时，得出的 $|t|$ 落在对应的拒绝域内，故拒绝 H_0 ，认为两组评酒员的评价结果有显著性差异；当综合红葡萄酒和白葡萄酒的的评分结果得出的 $|t|$ 都没有落在对应的拒绝域内，故接受 H_0 ，认为两组评酒员

的评价结果无显著性差异。

5.1.4 可信度分析

用模型 (I)、(II) 得出结论都为两组评酒员的评价结果存在显著性差异，故我们需进一步确定哪一组结果更可信。设每一种葡萄酒样品的实际质量分数为 a_{i0} 。由于 a_{i0} 未知，利用参数估计的替换原理：用样本均值来代替总体均值的思想，本题中 2 组评酒员的评价结果即为样本，则 a_{i0} 可用此每组 10 位成员的分数的均值替代即：

$$a_{i0}^{(1)} = \sum_{n=1}^{10} a_i^{1n} \quad a_{i0}^{(2)} = \sum_{n=1}^{10} a_i^{2n}$$

理论上评酒员对其的评分都应该在 a_{i0} 左右小范围内波动，即数据越集中，评价的可信度越高。故我们可以用比较两组评酒员评价结果的与 a_{i0} 的偏离程度即方差来分析它们的可信度，方差越小，评分的离散程度越低，可信度越高。

设组 1 和组 2 与 a_{i0} 的偏离方差分别为 s_1 、 s_2 ，则他们的计算式为：

$$s_1 = \frac{1}{27-1} \sum_{i=1}^{27} (X_i - a_{i0}^{(1)})^2; \quad s_2 = \frac{1}{28-1} \sum_{i=1}^{28} (Y_i - a_{i0}^{(2)})^2$$

计算可得 $s_1 = 892.39, s_2 = 565.75$ ，即 $s_1 > s_2$ 。

所以，我们认为在以偏离方差来分析可信度时，第二组评价结果的可信度更高。

(二) 问题二的模型建立与求解

5.2.1 模型建立

问题二要求我们根据酿酒葡萄的理化指标和葡萄酒的质量对这些酿酒葡萄进行分级，我们采用聚类分析法解决此题。附件 2 中给出的理化指标分为一级指标和二级指标两种，其中酿酒葡萄的一级指标为 30 个，二级指标为 32 个。显然理化指标非常多，对酿酒葡萄进行初步分类很难做到，故采用常用的聚类分析方法——系统聚类法，它开始时每个对象自成一类，然后每次讲最相似两类合并，合并后重新计算新类与其他类的距离，这一过程一直继续到所有队形归为一类为止。样品间的距离采用最广泛的距离——欧式距离，类之间的距离采用最短距离法、中心距离法、离差平方和法分别计算比较后，采用其中分类均匀相对合理的一种。最后给出谱系聚类图和分类个数不同的具体分法。

设在一个统计问题中有 m 个变量， n 个样本，则有

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (1)$$

为统计样本资料矩阵。其中， x_{ij} 为第 i 个样本第 j 个变量的观测值，对应每个样本的观测值 $x_i = (x_{i1}, x_{i2}, \cdots, x_{ik}, \cdots, x_{im})$ 可以看成是 m 维空间的一个点，而各样本之间的亲疏程度可用 m 维空间中点与点之间的距离 d_{ij} 表示，也可以用 Pearson 相关系数、So sine 相

似度等表示. d_{ij} 常用的计算方法有：欧氏距离、欧氏距离平方、马氏距离、切比雪夫距离、闵科夫斯基距离等. 本文使用欧氏距离公式计算 d_{ij} ，其公式为

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}$$

系统聚类法的聚类原则决定于样品间的距离及类间距离的定义，类间距离采用不同的定义方法时，可能得出的分类结果会有很大差异，故我们计算时首先采用了三种方法：重心法、离差平方和法、最短距离法。最短距离法是把两类中相聚最近的样品间的距离作为类间距离；重心法是考虑到了每类中包含样品个数，将两类重心间的距离作为类间距离；离差平方和法基于方差分析的思想，如果类分的正确，则类之间的离差平方和应该较小，不同类样品之间的离差平方和应该较大。

对于一级和二级指标，我们发现二级指标是一级指标的具体内容，如对于一级理化指标氨基酸总量，其后紧跟的 17 个二级指标为各种氨基酸，它们的信息已经合并体现在一级指标中了。若一级、二级指标同时采用，不但增加了变量的个数，问题复杂，而且由于收指标间相关性的影响，样品距离的计算值会与实际值产生偏差，故操作时我们只采用 30 个一级理化指标。

另外，在计算样品间的距离时，我们必须考虑到酿酒葡萄的理化指标与葡萄酒质量的权重问题。题目中明确提到：酿酒葡萄的好坏与所酿葡萄酒的质量有直接的关系。而反应葡萄酒的质量的只有一个值，即两组评酒员对葡萄酒评价结果的均值，而酿酒葡萄涉及到 30 个理化指标，多所有数据标准化处理之后，若无权重限制，葡萄酒质量对距离的贡献率很小，不符合题意。故加过权重后的欧氏距离公式变为

$$d_{ij} = \sqrt{\sum_{k=1}^m \varpi_k (x_{ik} - x_{jk})^2}$$

其中满足 $\sum_{k=1}^m \varpi_k = 1$ ， m 为变量个数。

为了体现葡萄酒的质量与酿酒葡萄的好坏有很大关系，且酿酒葡萄自身的理化指标与葡萄酒质量一样很大程度上体现其质量，我们取 $\varpi_{29} = 0.5$, $\varpi_i = 0.5/28, i = 1, 2, \dots, 28$ 。

5.2.3 模型求解

a. 数据标准化变换

在本题中，酿酒葡萄分类时共有 31 个指标，不同的指标有不同的量纲，不同的数量级单位，不同的取值范围。若把它们直接放在一起计算距离，距离必会偏向取值范围偏大的指标。为了使不同量纲、不同取值范围的数据能够放在一起进行比较，需要对数据进行变换处理，这里我们采用标准化变换。

设红葡萄样品 i 一级理化指标的观测数据为 $x_{ij}^{(1)}, i = 1, 2, \dots, 27; j = 1, 2, \dots, 28$ (j 对应的理化指标编号为附件中从前到后一级指标的依次顺序)，由于每种葡萄酒样品的质量 $a_{i0}, i = 1, 2, \dots, 55$ 在问题一中已经求解过，故对应红葡萄酒的质量为 $x_{i29}^{(1)} = a_{i0}, i = 1, 2, \dots, 27$ 白葡萄样品 i 一级理化指标观测数据为 $x_{ij}^{(2)}, i = 1, 2, \dots, 28, j = 1, 2, \dots, 28$ ，对应葡萄酒的质量为 $x_{i29}^{(2)} = a_{(i+28)0}, i = 1, 2, \dots, 27$ 。数据标准化变换公式为：

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j}$$

式中： \bar{x}_j 为第 j 个变量的平均数； s_j 为第 j 个变量的标准差。

b.问题求解

根据上面的介绍，现对酿酒葡萄进行分级，算法步骤如下：

- 第一步：根据数据标准化变换公式对各指标的观测值进行处理；
- 第二步：根据各标量指标间的权重欧式距离计算方法得出样品间的距离矩阵；
- 第三步：用不同的类间距离分类。
- 第四步：用较优的类间距离方法得出分类结果；
- 第五步：

用 SPSS 软件可以很方便的得出结果，当把 27 个红葡萄分成,5 类时，三种类间距离法得到的结果如下：

表 3 三种类间距离分成 4 类时的结果

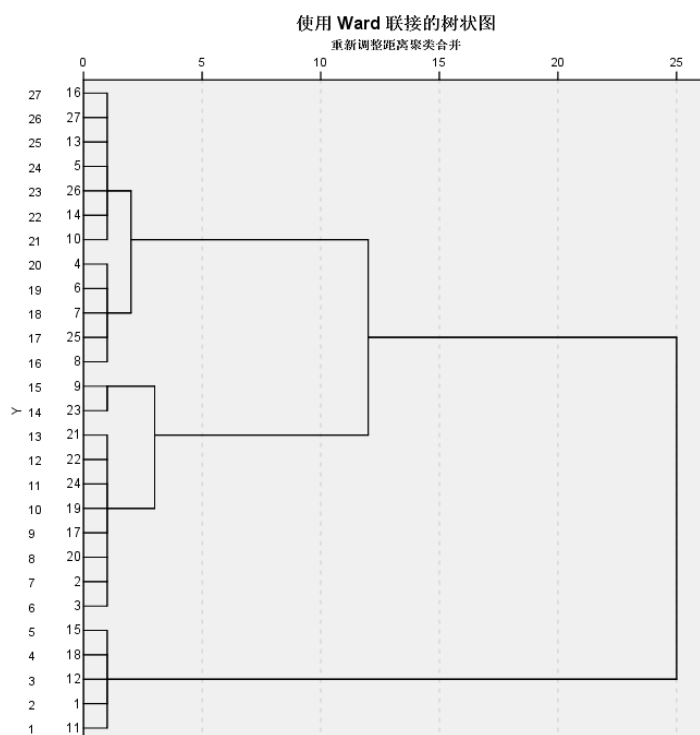
	第 1 类	第二类	第三类	第四类	第五类
最短距离法	1	2,3,4,5,6,7,8,10,13,14,16,17,19,20,21,22,24,25,26,27	9,23	11	12,15,18
重心法	1,11	2,3,17,19,20,21,22,24	4,5,6,7,8,10,13,14,16,25,26,27	9,23	12,15,18
离差平方和法	1,11,12,15,18	2,3,17,19,20,21,22,24	4,6,7,8,25,	5,10,13,14,16,26,27	9,23

我们从两个方面分析上表：

（1）分类的平均性。从上表可以看出用最短距离法把酿酒红葡萄分成 5 类时，28 个样品中有 20 个集中在了同一类，显然分类效果不佳。而重心法分类结果中各类包含样品数分别, 2、8、12、2、3，比最短距离法由一定改善。用离差平方和法分出的 5 类包含样品数分别为 5、8、5、7、2，可见此种方法分出的各类包含品数最为平均。

（2）数据的聚拢性。在用重心法与离差平方和法（Ward 法）分出的类别中，很多样品的集中程度是很相似的。例如样品 9 和 23，样品 4、6、7、8 和 25 等用两种类间距离时都是集中在同一类的。分析上表可以发现，重心法中第二、三类共 20 个样本可以合并为最短距离的第二类，而离差平方和（Ward 法）的第二、三、四类可以合并为最短距离的第二类，即对于同样的样本，离差平方和法可以把分类更精细，灵敏度更高。

综上，离差平方和法（Ward 法）的分类效果最好，我们以下着重分析用此方法对酿酒葡萄进行分类的结果。首先对 27 个红葡萄进行分类，用 SPSS 得出的谱系聚类树状图



如下：

图 1

从谱系聚类图中可以得到类数为任意数时的分类结果，分类个数分别为 3、4、5、6 的结果如下：

表 4 类间距离 WARD 法时红葡萄的聚类分析结果

	第一类	第二类	第三类	第四类	第五类	第六类
分成 3 类	1,11,12,15,18	2,3,9,17,19,20,21,22,23,24	4,5,6,7,8,10,13,14,16,25,26,27			
分成 4 类	1,11,12,15,18,	2,3,17,19,20,21,22,24	4,5,6,7,8,10,13,14,16,25,26,27	9,23		
分成 5 类	1,11,12,15,18	2,3,17,19,20,21,22,24	4,6,7,8,25	5,10,13,14,16,26,27	9,23	
分成 6 类	1,11	2,3,17,19,20,21,22,24	4,6,7,8,25	5,10,13,14,16,26,27	9,23	12,15,18

对 28 个白葡萄进行分类，得到谱系聚类树状图如下：

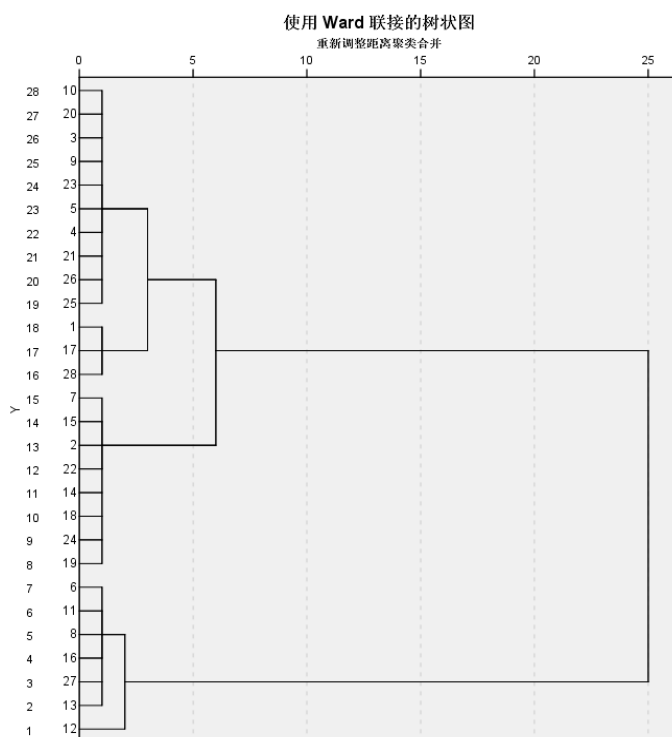


图 2

从谱系聚类图中可以得到类数为任意数时的分类结果，分类个数分别为 3、4、5、6 的结果如下：

表 5 类间距离 WARD 法时白葡萄的聚类分析结果

	第一类	第二类	第三类	第四类	第五类	第六类
3 类	1,3,4,5,9,10,17,20,21,23,25,26,28	2,7,14,15,18,19,22,24	6,8,11,12,13,16,27			
4 类	1,17,28	2,7,14,15,18,19,22,24	3,4,5,9,10,20,21,23,25,26	8,11,12,13,16,27		
5 类	1,17,28	2,7,14,15,18,19,22,24	3,4,5,9,10,17,20,21,23,25,26	6,8,11,13,6,27	12	
6 类	1,17,28	2,7,14,15,18,19,22,24	3,5,9,10,20,23	4,21,25,26	6,8,11,13,16,27	12

表 4 和表 5 具体列出所分类数不同时的具体分类结果，但如果要对他们进行分级，必须要有一个衡量指标，实际中当衡量指标不同时最后的分级结果可能会有一定的差异。由于酿酒葡萄的好坏与所酿葡萄酒的质量直接相关，所以我们考虑根据对应葡萄酒的质量对葡萄酒分级。具体操作为：计算每类包含的样本葡萄对应的葡萄酒的质量均值，比较它们大小，越大则对应的酿酒葡萄的级别越高，最大值对应的级别定为第 I 级。我们给出分为 4 类时级别的确定结果如下：

表 6 对 Ward 分为 4 类后的级别定义

	第一类	第二类	第三类	第四类
红葡萄质量均值	62.77	76.06	70.96	80.6

红葡萄级别	IV	II	III	I
白葡萄质量均值	79.98	74.99	77.63	67.85
白葡萄级别	I	III	II	IV

从上表看出不同级别间的质量均值差值比较大且均匀，即印证了我们在分类时达到了不同类间的距离很大，同类间的距离很小的原则，分了比较成功。类数为 3、5、6 的级别分析可用相似的方法得到。

（三） 问题三的建模与求解

5.3.1 问题三的模型建立

问题三要求分析酿酒葡萄与葡萄酒的理化指标之间的联系。首先对附件 2 中的数据进行分析，发现酿酒葡萄与葡萄酒的一级指标分别有 30 个和 9 个，为了建立它们间的联系，我们必须减少指标的数量，之后用精简后的指标建立回归关系即得出他们间的联系。为了满足降维和回归两方面的要求，采用多元统计分析中的偏最小二乘分析。

偏最小二乘回归分析的基本做法是，考虑 p 个因变量 Y_1, \dots, Y_p 与 m 个自变量 X_1, \dots, X_m 的关系。首先在自变量中提取第一成分 T_1 （ T_1 是 X_1, \dots, X_m 的线性组合，且尽可能多的提取原变量集中的变异信息）；同时在因变量集中也提取第一成分 U_1 ，并要求 T_1 与 U_1 相关程度达到最大；然后建立因变量 Y_1, \dots, Y_p 与 T_1 的回归方程。如果回归方程已达到满意的精度，则算法终止；否则继续对第二成分的提取，知道能达到满意的精度为止。若最终对自变量集提取 r 个成分 T_1, T_2, \dots, T_r ，偏最小二乘回归分析法将通过建立 Y_1, \dots, Y_p 与 T_1, T_2, \dots, T_r 的回归方程，然后再表示为 Y_1, \dots, Y_p 与原自变量的回归方程，即偏最小二乘回归方程。为了比较方便，因变量与自变量均为标准化变量， n 次标准化观测数据记为：

$$Y_0 = \begin{bmatrix} y_{11} & \cdots & y_{1p} \\ y_{21} & \cdots & y_{2p} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix}, \quad X_0 = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ x_{21} & \cdots & x_{2p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

偏最小二乘回归分析建模的具体步骤为：

（1）分别提取两变量组的第一主成分 T_1 和 U_1 ，并使之相关性达到最大。 T_1 是自变量 $X = (X_1, \dots, X_m)'$ 的线性组合：

$$T_1 = \varpi_{11}X_1 + \cdots + \varpi_{1m}X_m = \varpi_1'X$$

是因变量集 $Y = (Y_1, \dots, Y_p)'$ 的线性组合：

$$U_1 = v_{11}X_1 + \cdots + v_{1m}X_m = v_1'Y$$

为了回归分析的需求，要求：

- ① T_1 和 U_1 尽可能多地提取所在变量组的变异信息；
- ② T_1 和 U_1 的相关程度达到最大。

（2）建立 Y_1, \dots, Y_p 对 T_1 的回归，以及 X_1, \dots, X_m 对 T_1 的回归方程。回归模型为

$$\begin{cases} X_0 = t_1 \alpha_1' + E_1, \\ Y_0 = t_1 \beta_1' + F_1, \end{cases}$$

其中 t_1 为 n 维得分向量, $\alpha_1' = (\alpha_{11}, \dots, \alpha_{1m})$, $\beta_1' = (\beta_{11}, \dots, \beta_{1m})$ 分别是多因变量而只有一个自变量的回归模型中的参数向量, E_1 和 F 分别为 $n \times m$ 和 $n \times p$ 残差阵, 回归系数向量的最小二乘估计为

(3) 用残差阵 E_1 和 F 代替 X_0 和 Y_0 , 然后再重复以上步骤。记 $\hat{X}_0 = t_1 \alpha_1'$, $\hat{Y}_0 = t_1 \beta_1'$, 则残差阵 $E_1 = X_0 - \hat{X}_0$, $F_1 = Y_0 - \hat{Y}_0$ 。当残差阵 F_1 中元素的绝对值近似为 0 时停止抽取成分。

(4) 设 $n \times m$ 数据阵 X_0 的秩为 $r \leq \min(n-1, m)$, 则存在 r 个成分 t_1, t_2, \dots, t_r , 使得

$$\begin{cases} X_0 = t_1 \alpha_1' + \dots + t_r \alpha_r' + E_r, \\ Y_0 = t_1 \beta_1' + \dots + t_r \beta_r' + F_r, \end{cases}$$

设 $X_i^* (i=1, \dots, m), Y_j^* (j=1, \dots, p)$ 表示标准变量, 把

$$T_k = \varpi_{k1} X_1^* + \dots + \varpi_{km} X_m^* (k=1, \dots, r)$$

代入

$$Y_j^* = \beta_{1j} T_1 + \beta_{2j} T_2 + \dots + \beta_{rj} T_r (j=1, \dots, p)$$

即得 p 个标准化因变量的偏最小二乘回归方程:

$$\hat{Y}_j^* = a_{j1}^* X_1^* + \dots + a_{jm}^* X_m^* (j=1, \dots, p)$$

然后再还原原始变量的偏最小二乘回归方程:

$$\hat{Y}_j = \hat{a}_{j0} + \hat{a}_{j1} X_1 + \dots + \hat{a}_{jm} X_m (j=1, \dots, p)$$

(5) 确定抽取成分的个数 l 。则只需在 r 个成分 t_1, t_2, \dots, t_r 中选用前 l 个 ($l \leq r$) 建立回归方程, 即可得到预测能力较好的回归方程模型。

5.3.2 问题三模型的求解

此问题中, 自变量为 30 个酿酒葡萄的一级理化指标, 因变量为 9 个葡萄酒一级指标, 每个指标有 55 组观测数据。在使用偏最小二乘分析计算之前, 首先还是把各指标的数据标准化处理, 标准化变换公式为:

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j}$$

式中: \bar{x}_j 为第 j 个变量的平均数; s_j 为第 j 个变量的标准差。

使用 *Matlab* 工具箱中最小二乘回归分析命令 *plsregress*, 使用格式

$$[XL, YL, XS, YS, BETA, PCTVAR, MSE, STATS] = plsregress(X, Y, ncomp)$$

对 30 个酿酒葡萄指标和 9 个葡萄酒指标进行分析。首先对红葡萄酒进行分析。

从输出结果看出, 由酿酒葡萄的 30 个指标中抽取的前 10 个成分对葡萄酒的量化指标的解释解释率达到了 84.64%, 因此取前 10 对成分就已经足够了。

求出自变量组和因变量组与 10 个成分间的回归方程, 系数阵如附录所示。

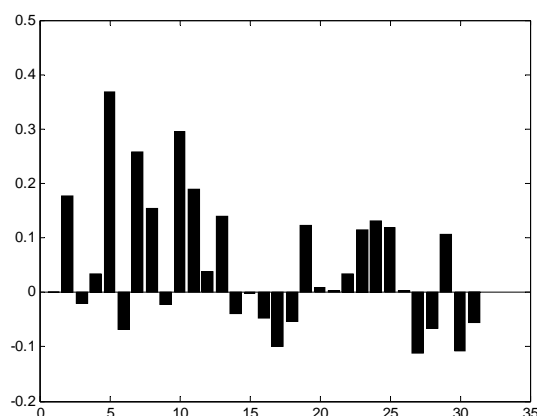
之后得出因变量与自变量之间的回归方程, 即 9 个葡萄酒的理化指标与酿酒葡萄的 30 个理化指标之间的回归方程。其中, 红葡萄酒的第一个量化指标氨基酸总量与酿酒葡萄的 30 个量化指标的系数如下表所示:

表 7 红葡萄酒氨基酸总量与酿酒葡萄的 30 个量化指标的系数									
-746.62	0.02063	-0.4426	0.65609	0.85536	-7.1587	16.7522	49.31712	0.64369	0.13176
6	3	9	8	1	8	6	31	9	7
256.370	2.06857	4.78185	-0.6081	-0.3848	0.04060	-0.2547	-0.647877	1.19323	44.0727
7	1	4	6	5	1	4	4	2	2
-2.642	1.78788	4.86222	0.18791	0.31411	19.8118	-1.9454	-110.7104	15.2857	-11.227
	7	4	5	8	6	3	4	5	1

表中从前到后一次为红葡萄酒氨基酸总量与酿酒葡萄从前到后依从的 30 个指标的回归系数。可以看出，系数的绝对值最大达到 746.62，最小只有 0.13176，可见红葡萄酒中各项指标不是与所有的酿酒葡萄都有很强的关系的，同指标对葡萄酒的影响差距很大。

为了更直观、迅速的观察酿酒葡萄指标在解释葡萄酒的指标时的边际作用，绘制红葡萄酒第一个指标——总氨基酸总量的回归系数图如下：

图 3 红葡萄酒总氨基酸量的回归系数图



从图中可以看出，酿酒葡萄的第 1、3、5、7、9、10 个指标对总氨基酸量的正边际作用最大，酿酒葡萄的第 17、27、28 个指标对其的负边际作用最大。

另外，为了考察回归方程的模型精度，对样本绘制预测图，如果所有点都能在图的对角线走位均为分布，则方程的拟合值与原值差异很小。我们画出红葡萄 9 个一级指标中的 4 个体能训练预测图来大概了解回归方程的精度。图形如下：

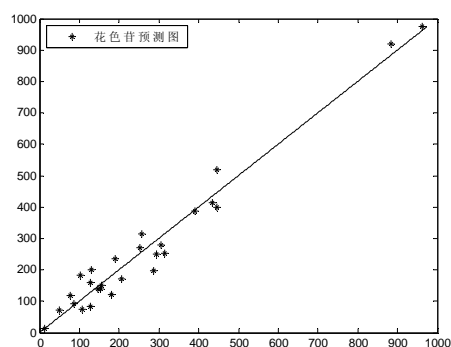


图 4 花色苷预测图

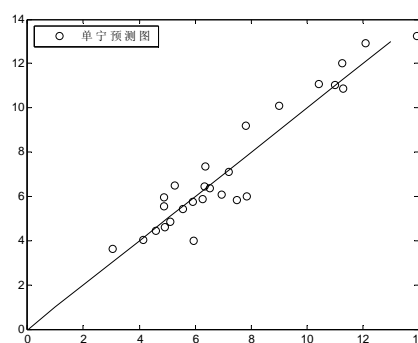


图 5 单宁预测图

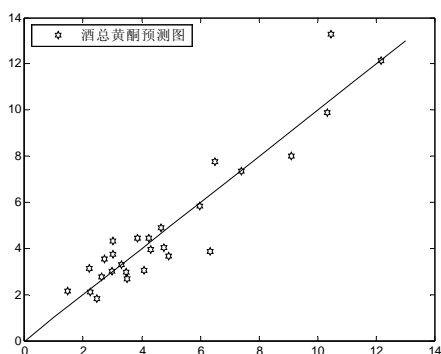


图 6 总黄酮预测图

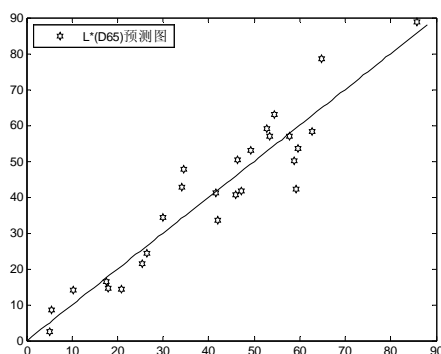


图 7 $L^*(D65)$ 预测图

从上面四个图，我们看出数据基本均匀分布在对角线两边，预测模型的精度较高。

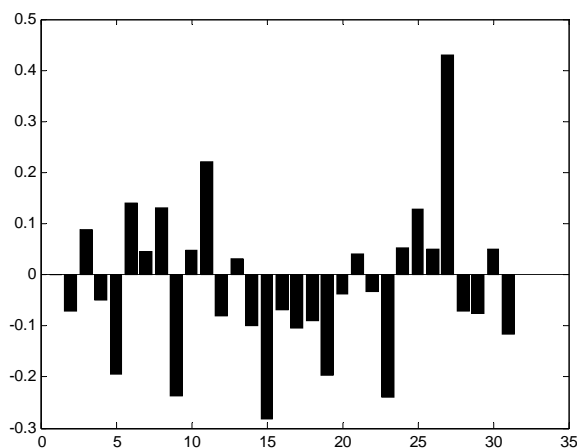
红葡萄酒的其他 8 个一级量化指标和白葡萄的一级量化指标的回归系数图，预测图可用同样的方法得出。其中，白葡萄酒的第一个量化指标氨基酸总量与酿酒葡萄的 30 个量化指标的系数如下表所示：

表 8 白葡萄酒氨基酸总量与酿酒葡萄的 30 个量化指标的系数

-2.35	0.00024	0.00257	0.44025	-0.0487	0.08725	-0.0658	0.0998	-0.005	-0.00
554	9	5	3	7	7	4	16	4	029
2.374	-0.0214	0.08614	0.03948	-0.0800	0.01389	0.01256	-0.003	0.00056	-0.70
52	9	9	2	5	9	6	61	1	896
0.001	-0.0013	0.01094	0.00068	0.00184		0.02029	0.5782	0.00955	0.050
909	9	3	6	9	-0.037	1	5	4	237

同时，绘制白葡萄酒第一个指标——总氨基酸总量的回归系数图如下：

图 8 白葡萄酒总氨基酸量的回归系数图



（四） 问题四的建模与求解

a. 酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响

酿酒葡萄的理化指标比较多，既有一级指标，又有二级指标，其中一级指标的数目是 30 个，二级指标 8 个。考虑到一级理化指标的数量明显较多，而且二级理化指标对于葡萄酒质量影响集中体现在一级指标，所以我们在分析的过程中只取酿酒葡萄理化指标中的所有一级指标，葡萄酒的主要评价指标只取指标中的所有一级指标。设所有一

级理化指标中的数目为 s ，所用的一级理化指标记为 $X_i (i=1, \dots, s_1)$ 。

考虑到每个样品的葡萄酒质量与所有的一级指标都有一定的影响，而葡萄酒的质量得分又分为外观分析，香气分析，口感分析，整体评价等 4 个方面，分别记为 Y_1, \dots, Y_4 ，一级理化指标影响质量的四个方面的侧重点会有所不同，本文主要分析一级理化指标对于四个方面的不同的影响。

于是，我们利用相关系数进行分析，求出 $Y_j (j=1, \dots, 4)$ 与 $X_i (i=1, \dots, n_1)$ 的相关系数记为 $\rho(i, j)$ ，则

$$\rho(i, j) = \frac{Cov(i, j)}{\sqrt{D(i)} \sqrt{D(j)}}$$

其中 $Cov(i, j)$ 表示 X_i 和 Y_j 的协方差， $D(i)$ 表示 X_i 的方差， $D(j)$ 表示 Y_j 的方差。

如果 $\rho(i, j) = 0$ ，说明 X_i 和 Y_j 不相关，即 X_i 理化指标对 Y_j 没有任何影响，如果 $\rho(i, j)$ 较大，则说明 X_i 对 Y_j 影响很大，如果 $\rho(i, j)$ 较小，说明 X_i 对 Y_j 影响很小，这样通过分析矩阵 $\rho(i, j)$ 就可以得到一级理化指标中对于质量的哪个具体方面的影响。

考虑到每个样品的葡萄酒质量与所有的一级指标都有一定的影响和关联，而葡萄酒的质量得分又分为外观分析，香气分析，口感分析，整体评价 4 个方面，分别记为 Y_1, \dots, Y_4 ，而一级理化指标影响质量的四个方面的侧重点会有所不同，我们在这里主要分析一级理化指标对于四个方面的不同的影响。

于是，我们利用相关系数进行分析，求出 $Y_j (j=1, \dots, 4)$ 与 $X_i (i=1, \dots, n_1)$ 的相关系数，记为 $\rho(i, j)$ ，则

$$\rho(i, j) = \frac{Cov(i, j)}{\sqrt{D(i)} \sqrt{D(j)}}$$

其中 $Cov(i, j)$ 表示 X_i 和 Y_j 的协方差， $D(i)$ 表示 X_i 的方差， $D(j)$ 表示 Y_j 的方差

如果 $\rho(i, j) = 0$ ，说明 X_i 和 Y_j 不相关，即 X_i 理化指标对 Y_j 没有任何影响，如果 $\rho(i, j)$ 较大，则说明 X_i 对 Y_j 影响很大，如果 $\rho(i, j)$ 较小，说明 X_i 对 Y_j 影响很小，这样通过分析矩阵 $\rho(i, j)$ 就可以得到一级理化指标中对于质量的哪个具体方面的影响。

利用 matlab 中的 corrcoef 函数计算出了样本数据的相关系数矩阵 $\rho(i, j)$ ，下面分析计算酿酒红葡萄和红葡萄酒的理化指标对于红葡萄酒的质量的影响。

红葡萄酒中，一级理化指标数目 $n_1=39$ 。

我们得到的相关系数矩阵如下表：

表 9 红葡萄与红葡萄酒的理化指标之间的相关系数

	1(外观)	2(香气)	3(香气)	4(平衡评价)
1	0.247734	0.199226	0.040966	0.036329
2	0.264041	0.471595	0.448658	0.42725
3	0.304869	0.497258	0.402108	0.415023
4	0.227136	0.526062	0.513833	0.495606
5	0.496205	0.426058	0.514056	0.491406

6	0.319503	0.537346	0.549471	0.517984
7	-0.51242	-0.25934	-0.27066	-0.29524
8	0.448717	-0.16128	0.119013	0.122817
9	0.069299	-0.01425	0.330493	0.347778
10	0.04901	0.076605	0.264875	0.175562
11	0.390916	0.510448	0.398977	0.50106
12	0.09818	0.03391	-0.05767	-0.06203
13	0.281679	0.249979	0.102227	0.124422
14	0.190041	0.216704	0.354348	0.385396
15	-0.22661	-0.15487	-0.35373	-0.35221
16	0.002964	-0.2304	-0.24665	-0.24414
17	0.034616	-0.24394	-0.23565	-0.22324
18	0.081093	0.033661	-0.0004	0.001264
19	0.478124	0.673585	0.467789	0.535783
20	0.379473	0.521786	0.493315	0.495451
21	0.446713	0.207599	0.173892	0.111175
22	0.420739	0.646489	0.586771	0.613942
23	-0.11406	0.247308	-0.02378	0.00924
24	0.333711	0.239956	0.29933	0.271037
25	-0.32897	-0.06945	0.175707	0.101044
26	-0.34025	-0.18817	0.148967	0.048739
27	-0.34049	-0.20445	0.029533	-0.07861
28	0.22678	0.512847	0.508506	0.506071
29	-0.18817	-0.45672	-0.2686	-0.35692
30	0.050128	0.338374	0.293148	0.33035
31	-0.28483	-0.21167	0.078141	-0.03638
32	0.327979	0.179908	0.029262	0.099428
33	0.044049	0.136858	0.00026	0.01508
34	0.297008	0.34146	0.215672	0.222082
35	0.239431	0.274322	0.281603	0.272297
36	0.261843	0.005262	0.090049	0.053624
37	-0.1067	0.054174	-0.02508	-0.03944
38	-0.50877	-0.10858	-0.30792	-0.29413
39	-0.4616	-0.16843	-0.37146	-0.34812

由上表可知：

影响红葡萄酒质量中的外观的理化指标：

正向贡献的主要是：白藜芦醇，DPPH 自由基 1/IC50, a*(D65), 单宁，葡萄总黄酮。负向贡献的主要是：L* (D65)，a*(+红；-绿)，b*(+黄；-蓝)，可溶性固形物，还原糖。

影响红葡萄酒质量的香气的主要理化指标：

正向贡献的主要是：DPPH 自由基 1/IC50, 葡萄总黄酮，DPPH 半抑制体积，酒总黄酮，总酚。负向贡献的主要是：可滴定酸，L* (D65)，多酚氧化酶活力，柠檬酸，干物质含量。

影响红葡萄酒质量的口感的主要理化指标

正向贡献的主要是：葡萄总黄酮，DPPH 半抑制体积，白藜芦醇，酒总黄酮，PH 值，总

酚。负向贡献的主要是： b^* (+黄；-蓝)，苹果酸， a^* (+红；-绿)， L^* (D65)，可滴定酸。

影响红葡萄酒质量的平衡的主要理化指标：

正向贡献的主要是：葡萄酒总黄酮，DPPH 自由基 1/IC50，DPPH 半抑制体积, PH 值，蛋白质。负向贡献的主要是：可滴定酸，苹果酸， b^* (+黄；-蓝)， L^* (D65)， a^* (+红；-绿)。

对于酿酒白葡萄和白葡萄酒的理化指标对于白葡萄酒的质量的影响，由于白葡萄酒和红葡萄酒的分析方法，模型的建立和数据的处理完全一样，限于论文的篇幅有限，我们在这里就不再一一叙述了，仅仅把处理后的结果写在下面：

影响白葡萄酒质量中的外观的理化指标：

正向贡献的主要是：白藜芦醇，DPPH 自由基 1/IC50，单宁，葡萄总黄酮。负向贡献的主要是： L^* (D65)， b^* (+黄；-蓝)，可溶性固形物，还原糖。

影响白葡萄酒质量的香气的主要理化指标：

正向贡献的主要是：DPPH 自由基 1/IC50, 葡萄总黄酮，DPPH 半抑制体积，酒总黄酮，总酚。负向贡献的主要是：可滴定酸，多酚氧化酶活力，柠檬酸，干物质含量。

影响白葡萄酒质量的口感的主要理化指标

正向贡献的主要是：葡萄总黄酮，DPPH 半抑制体积，白藜芦醇，酒总黄酮，PH 值，总酚。负向贡献的主要是： b^* (+黄；-蓝)，苹果酸， L^* (D65)，可滴定酸。

影响白葡萄酒质量的平衡的主要理化指标：

正向贡献的主要是：葡萄酒总黄酮，DPPH 自由基 1/IC50，DPPH 半抑制体积, PH 值。负向贡献的主要是：可滴定酸，苹果酸， L^* (D65)。

b. 能否用葡萄酒和葡萄理化指标评价葡萄酒质量的论证

此小问是要求我们论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量，对于这个问题，我们想通过对葡萄和葡萄酒的理化指标和葡萄酒质量的得分进行拟合，酿酒葡萄的理化指标比较的多，既有一级指标，又有二级指标，其中一级指标的数目大约是 30 个，考虑到一级理化指标的数目很多，而且二级理化指标对于葡萄酒质量影响不大，所以我们在分析的过程中只取酿酒葡萄理化指标中的所有的一级指标，葡萄酒的理化指标相对的少一点，也是既有一级指标，又有二级指标，其中一级指标大约是 9 个，在分析的过程中，我们也是只取葡萄酒理化指标中的所有的一级指标，并且，我们又对这些一级理化指标进行主成分分析，得到了较少的变量，然后我们利用这些较少的变量替代原来的一级理化指标进行下面的回归拟合。

设评酒员给葡萄酒的质量打的分数为 $y_i (i=1, \dots, n_2)$ ，其中 s_2 是葡萄酒的样品数。

然后我们又通过拟合的结果求出每个葡萄酒样品的得分记为 $\hat{y}_i^{(1)} (i=1, \dots, n_2)$ ，由于 $(y_i - \hat{y}_i^{(1)}) (i=1, \dots, n_2)$ 符合正态分布，所以我们可以对 $(y_i - \hat{y}_i^{(1)}) (i=1, \dots, n_2)$ 进行 T 检验，如果 T 检验通过则说明能用葡萄和葡萄酒的理化指标来评价葡萄酒的质量，如果 T 检验没能通过，则说明不能用葡萄和葡萄酒的理化指标进行评价。

接着我们进一步探讨合理地评价葡萄酒质量的各项因素和指标。由于附录三中给出了芳香物质的信息，而芳香物质本身确实会通过影响葡萄酒的口感来影响葡萄酒的质量，所以我们现在把芳香物质和理化指标一起考虑，进行回归拟合葡萄酒的得分，由于葡萄酒和葡萄中的芳香物质非常的多，我们同样采用主成分分析的方法，提取出一些较少的变量来替换全部的芳香物质。最后，我们把芳香物质和理化指标一起考虑，进行回归拟合葡萄酒的得分，记求出的每个葡萄酒样品的得分记为 $\hat{y}_i^{(2)} (i=1, \dots, n_2)$ ，同样

$(y_i - \hat{y}_i^{(2)})(i=1, \dots, n_2)$ 符合正态分布, 对 $(y_i - \hat{y}_i^{(2)})(i=1, \dots, n_2)$ 进行 T 检验, 如果检验通过, 则说明我们建立的评价葡萄酒质量需要把理化指标和芳香物质综合考虑的合理性。

5.4.2 模型的建立和求解:

对于红葡萄酒进行数据处理, 设评酒员给葡萄酒的质量打的分数为 $y_i (i=1, \dots, 27)$, 其中 s_2 是葡萄酒的样品数, 得到 y_i 的一些列的值, 如下:

67.11038	78.44535	78.21597	69.29178	74.73933	68.98883	64.61187
80.69194	71.90545	66.4726	67.70201	74.80556	71.92547	65.58584
72.36689	66.17466	73.99297	73.85815	74.81568	70.53351	80.72251
71.16193	75.42549	69.9611				

我们又对这些一级理化指标进行主成分分析, 得到了较少的变量, 然后我们利用这些较少的变量替代原来的一级理化指标进行多元线性回归拟合,

$$y_i = \beta_i^{(1)(0)} + \beta_i^{(1)(1)} x_i^{(1)(1)} + \dots + \beta_i^{(1)(9)} x_i^{(1)(9)} + \varepsilon_i^{(1)} \\ (i=1, \dots, 27)$$

于是我们可以得到参数系数 $\beta_i^{(1)(0)}, \dots, \beta_i^{(1)(9)}$ 的值

然后我们再利用参数系数 $\beta_i^{(1)(0)}, \dots, \beta_i^{(1)(9)}$ 和 $x_i^{(1)(1)}, \dots, x_i^{(1)(9)}$ 的值, 得出 $\hat{y}_i^{(1)}$ 的值

65.56 792	76.50 977	77.69 094	70.28 82	74.75 628	69.73 056	68.00 854	70.27 4	65.56 792	76.50 977
80.28 654	71.38 916	65.18 281	60.62 393	73.47 821	70.06 187	64.48 69	71.26 474	80.28 654	71.38 916
74.86 886	65.21 809	73.73 924	77.67 058	75.86 061	73.99 464	81.74 658	71.06 978	74.86 886	65.21 809
67.99 883	74.48 622	71.94 621							

然后我们计算 $(y_i - \hat{y}_i^{(1)})(i=1, \dots, 27)$, 由于得到的一组值符合正态分布, 我们把得到的一组值进行 T 检验:

先计算出均值

$$\bar{y}_i = \frac{\sum_{i=1}^{s_2} (y_i - \hat{y}_i^{(1)})}{27}$$

假设样本均值与总体均值相等, 即 $u = u_0$,

$$T = \sqrt{s_2} \times \frac{\bar{y} - u_0}{s}, \quad s \text{ 是样本的方差。}$$

通过 MATLAB 的 ttest 函数进行 T 检验, 得出判断仅仅用葡萄和葡萄酒的理化指标来评价葡萄酒的质量不合理。

接着我们进一步探讨合理地评价葡萄酒质量的各项因素和指标。由于附录三中给出了芳香物质的信息, 而芳香物质本身确实会通过影响葡萄酒的口感来影响葡萄酒的质量, 所以现在我们把芳香物质和理化指标一起考虑, 进行回归拟合葡萄酒的得分, 由于

葡萄酒和葡萄中的芳香物质非常的多，我们同样采用主成分分析的方法，提取出一些较少的变量来替换全部的芳香物质。最后，我们把芳香物质和理化指标一起考虑，进行回归拟合葡萄酒的得分，记求出的每个葡萄酒样品的得分记为 $\hat{y}_i^{(2)} (i=1, \dots, 27)$ ，同样 $(y_i - \hat{y}_i^{(2)}) (i=1, \dots, 27)$ 符合正态分布，对 $(y_i - \hat{y}_i^{(2)}) (i=1, \dots, 27)$ 进行 T 检验：返回值为 1，拒绝零假设，接受备择假设 $u \neq u_0$ 。

再利用主成分分析后的一级理化指标和主成分分析后的芳香物质进行多元线性回归拟合，

$$y_i = \beta_i^{(2)(0)} + \beta_i^{(2)(1)} x_i^{(2)(1)} + \dots + \beta_i^{(2)(9)} x_i^{(2)(20)} + \varepsilon_i^{(2)} \quad (i=1, \dots, 27)$$

于是我们可以得到参数系数 $\beta_i^{(2)(0)}, \dots, \beta_i^{(2)(9)}$ 的值然后我们再利用参数系数 $\beta_i^{(2)(0)}, \dots, \beta_i^{(2)(9)}$ 和 $x_i^{(2)(1)}, \dots, x_i^{(2)(9)}$ 的值，得出 $\hat{y}_i^{(2)}$ 的值

65.56 792	76.50 977	77.69 094	70.28 82	74.75 628	69.73 056	68.00 854	70.27 4	65.56 792	76.50 977
80.28 654	71.38 916	65.18 281	60.62 393	73.47 821	70.06 187	64.48 69	71.26 474	80.28 654	71.38 916
74.86 886	65.21 809	73.73 924	77.67 058	75.86 061	73.99 464	81.74 658	71.06 978	74.86 886	65.21 809
67.99 883	74.48 622	71.94 621							

然后我们计算 $(y_i - \hat{y}_i^{(2)}) (i=1, \dots, 27)$ ，由于得到的一组值符合正态分布，我们把得到的一组值进行 T 检验：

$$\text{先计算出均值 } \bar{y}_i = \frac{\sum_{i=1}^{s_2} (y_i - \hat{y}_i^{(2)})}{n_2} \text{ 其中 } n_2 = 27$$

假设样本均值与总体均值相等，即 $u = u_0$ ，

$$T = \sqrt{n_2} \cdot \frac{\bar{y} - u_0}{s}, \quad s \text{ 是样本的方差。返回值为 0，接受零假设，}$$

现在我们对于白葡萄酒进行建模和求解：

由于白葡萄酒和红葡萄酒的分析方法，模型的建立和数据的处理完全一样，限于论文的篇幅有限，我们在这里就不再一一叙述了，仅仅把处理后的结果写在下面：

当我们仅仅用白葡萄和白葡萄酒的理化指标来评价葡萄酒的质量时，不满足 T 检验，说明仅仅用白葡萄和白葡萄酒的理化指标来评价白葡萄酒的质量不合理。同时，当我们我们把芳香物质和理化指标一起考虑，得打满足 T 检验，说明，用理化指标和和芳香物质评价白葡萄酒的质量是合理的。

综上所述：无论对于红葡萄酒还是白葡萄酒，当我们不能仅仅用葡萄和葡萄酒的理化指标来评价葡萄酒的质量时，然而，当我们用理化指标和和芳香物质评价白葡萄酒的质量是合理的

六、模型的改进方向

1. 我们在问题二中采用加权的欧式距离法进行聚类分析，对于酿酒葡萄的量化指标和葡萄酒的质量权重值各取为 50%。实际研究中，可以根据不同的要求、不同的偏爱程度把

权重取成不同的值，得出对应要求的分类结果；

2. 问题二中，对分好类的酿酒葡萄进行级别定义时，仅参考了每个类中所包含样品对应 20 个评酒员的评分均值。实际分级时指标时受多方面因素影响的，可考虑加入更多的参考变量使分级更合理。

3. 问题四中，把芳香物质中的醇、醛、酯等相同的化合物种类的数据相加合并。实际中，可以根据每种芳香物质单位量呈现的香味效果，以加权相加合并的方式对芳香物质并类，使模型更精确。

七、参考文献

- [1] 苏金明 阮沈勇 王永利，《MATLAB 工程数学》，北京：电子工业出版社，2005；
- [2] 高惠璇《应用多元统计分析》，北京大学出版社，370 页至 374 页，2005；
- [3] 司守奎 孙玺菁，《数学建模算法与应用》，国防工业出版社，2012；
- [4] 杜强，贾丽艳 SPSS 统计分析从入门到精通》北京：人们邮电出版地，2009；
- [5] 王静龙 《多元统计分析》，北京：科学出版社，2008。

八、附录

SPSS 两独立样本 T 检验：

红葡萄酒样品酒的的检验值

0.213	0.016	0.05	0.511	0.67	0.056	0.146	0.073	0.19
0.051	0.019	0	0.03	0.871	0.989	0.016	0.152	0.099
0.77	0.166	0.229	0.056	0.002	0.047	0.765	0.513	0.578

红葡萄酒样品酒的的检验值

0.248	0.753	0.562	0.407	0.015	0.117	0.262	0.672	0.110	0.315
0.803	0.089	0.103	0.174	0.181	0.206	0.730	0.416	0.136	0.727
0.572	0.071	0.053	0.479	0.530	0.112	0.010	0.608		

红葡萄酒自变量、因变量与成分的系数阵：

成分	1	2	3	4	5	6	7	8	9	10
1	1.77 8347	-1.3098 1	-3.4296 2	0.1382 54	0.1131 83	0.2009 5	1.7130 84	1.7061 61	0.8931 62	0.0615 21
2	3.05 4797	-0.5602 8	0.8100 8	-0.1192 5	-3.0053 87	0.0091 7	-1.8506 7	0.9971 57	-0.5906 1	0.1032 56
3	-0.6 9077	0.1328 2	2.1070 2	-1.4755 1	-1.0908 4	-1.4832 6	1.8122 39	1.2674 16	-0.1288 8	1.1425 91
4	4.40 1347	1.2197 47	1.1818 3	0.1483 31	0.9200 39	-0.3700 9	-0.1708 6	-0.2164 3	0.8862 24	0.5074 98

5	1.66	-0.2894	-2.2008	0.5971	-2.4803	1.8144	-0.1759	0.0529	0.4742	1.1541
	9862	2	2	09	7	81	2	43	16	93
6	1.92	3.7465	0.0687	0.2815	2.2536	-0.4302	0.4950	-0.6153	0.4096	0.4416
	5696	29	5	35	61	7	65		41	18
7	1.27	2.6358	-0.7397	1.9793	-0.5406	1.2194	0.1995	1.0522	0.2824	1.2558
	5627	17	7	45	1	41	4	67	21	04
8	1.60	1.1612	0.8777	3.0553	1.2587	0.3064	-0.7939	-0.5702	-0.4396	-0.3894
	2882	77	81	07	3	4	4	7	5	6
9	3.00		1.7650	1.4973	0.7162	0.0868	0.2261	-0.3552	-0.4205	-0.7156
	4678	2.6506	89	64	95	08	01	1	5	8
10	4.00	-0.8259	1.1816	-1.7579	-0.8816	1.1910	-0.7043	0.4629	-0.2695	0.0891
	986	7	02	1	1	66	5	84	7	14
11	4.56	-1.0293	0.2382	-1.4716	0.1390	0.2502	-0.3921	-0.5017	0.0066	-0.6067
	7798	8	8	1	15	33	8	6	33	9
12	4.05	-0.5348	1.1779	-0.6563	0.8841	-0.2974	0.8670	0.5803	-0.8522	0.4064
	0114	7	02	4	87	8	32	62	1	41
13	3.89	-1.4529	0.5687	-2.0567	0.0095	0.9507	-0.6606	-1.0225	-0.6036	-0.0328
	3787	3	31	9	01	94	8	2	4	2
14	0.04		-0.9881	-2.6404	-1.9433	1.2294	0.4904	0.9614	-0.7548	0.4647
	6375	2.6888		5	5	3	44	13	1	68
15	2.79	-0.0664	-0.4832	1.2393	-0.2424	1.2982	-0.0008	2.6185	-1.2707	-0.1255
	7736	3	4	42	4	13	3	71	5	6
16	1.23	-0.5058	-3.8350	0.3049	1.7873	-0.9625	0.0648	-0.6081	0.3422	-0.2064
	8069	1	5	91	75	3	39	7	64	1
17	0.26	-0.1895	-3.6728	1.4111	1.3100	-0.5148		-0.9842	-0.1016	0.6232
	6992	5	4	22	6	4	-0.4174	7	7	58
18	1.29	-0.6592	-3.1708	0.6151		-1.2027	0.0326	0.3001	0.4740	-0.8677
	3319	3	5	65	2.7875	9	88	57	27	7
19	1.31	-1.4301	-0.7033	-1.7297	-2.4643	-1.1847	-1.4772	0.8407	2.4210	-0.2618
	2071	4	2	7	6	2	5	95	9	3
20	-1.2		-0.8212	0.2064	3.1545	-0.5342	1.9729	0.0879	-1.6498	0.7836
	8244	-1.936	4	88	07	5	01	64	8	49
21	1.80	2.0686	-0.6982	0.0450	-1.2348	0.0059	-2.4742	-0.3144		-0.7368
	4677	3	1	86	7	49	6	9	1.9141	7
22	1.81	-0.1377	-3.7563	1.2757	2.0232	-0.8401	0.4082	-0.3206	-0.1147	0.2947
	6565	8	6	69	09	9	65	4	9	39
23	-1.4	-1.0882	1.8751	0.0140	0.3754	1.2352	-2.5979	0.3822	1.6119	2.5143
	9849	7	01	65	12	43	6	4	76	05
24	-2.31	-1.8904	1.9394	-0.8648	1.8478	0.3430	-2.1565	0.8578	0.3255	0.6841
	159	5	36	9	98	75	6	44	43	31
25	2.86	0.6622	0.9861	0.7507	-1.2242	1.2387	1.8213	1.2127	0.0030	-0.8478
	6488	57	63	96	4	95	15	69	11	2
26	2.90	-0.9722	0.8179	-0.8424	-0.3764		-0.1245	-1.4471	-0.0172	2.1308
	8812	3	55	5	7	-1.9343	3	3	1	1
27	-0.8	-1.9492	1.8174	0.7757	2.3583	0.9083	-2.1563	1.1793	-0.9036	0.8076
	9518	5	26	77	43	04	9	73	5	01
28	-2.8	0.5262	1.0585	-0.8085	-0.7708	2.0516	-0.3404	0.4981	1.1157	-1.7351
	9685	74	41	1	7	42	8	84	46	6
29	-1.9	2.8674	-2.2057	-2.2270	-0.7166	0.5453	-0.2672	0.5946	-1.1392	0.2604
	7941	47	2	3	3	79	7	58	5	73
30	-0.9	3.1056	-2.7905	-1.6701	0.4916	0.3511	-0.5953	0.4403	-0.8112	1.4446
	5291	82	1	4	22	61	5	59	2	65

红葡萄酒与酿酒葡萄得到的回归方程系数:

	1	2	3	4	5	6	7	8	9
1	-746 .626 0.02	-16.38 97	-3.163 18	-14.73 41	15.845 99	-0.449 61	41.568 21	177.347 227	66.221 51
2	0.063 3	0.0003 54	0.0002 33	2.07E- 05	0.0005 6	1.24E- 05	-0.001 77	-7.48E-0 5	-0.000 19
3	-0.4 426 9	0.0043 88	-0.008 32	-0.001 73	-0.039 92	-0.000 36	0.0496 8	-0.03432 49	-0.011 43
4	0.65 609 8	-0.055 67	-0.066 16	-0.150 48	-0.048 96	-0.004 75	0.0736 19	0.22864 336	-0.714 45
5	0.85 536 1	0.0052 78	0.0051 89	0.0042 14	0.0044 65	0.0002 22	-0.048 87	-0.02492 39	-0.021 36
6	-7.1 587 8	-0.018 11	0.0546 63	-0.017 93	0.3147 99	0.0025 38	-0.765 2	0.97342 845	0.9950 3
7	16.7 522 6	0.0616 57	0.0681 31	0.0379 84	-0.028 87	0.0020 66	-0.519 84	-0.85360 1	-0.526 37
8	49.3 171 2	0.1660 01	-0.174 15	-0.683 21	-0.571 37	-0.016 41	-2.585 99	-0.85018 23	-1.187 19
9	0.64 369 9	-0.029 35	-0.008 47	-0.015 84	-0.005 73	-0.000 85	-0.248 67	0.18050 889	0.1766 18
10	0.13 176 7	0.0007 48	0.0006 51	0.0012 63	-0.001 01	2.67E- 05	-0.003 27	-0.00672 66	-0.003 26
11	256. 370 7	5.1252 31	5.4376 57	5.4677 69	5.7253 57	0.2813 33	-24.93 46	-8.26974 61	-0.060 74
12	2.06 857 1	0.0624 53	0.0865 77	0.1100 47	0.1005 01	0.0046 39	-0.354 64	-0.06617 97	0.1136 69
13	4.78 185 4	0.0732 33	0.0484 36	0.0595 23	-0.007 56	0.0024 43	-0.218 23	-0.15482 27	-0.232 25
14	-0.6 081 6	0.0579 2	0.1435 75	0.1737 93	0.2743 83	0.0076 67	-0.502 81	0.20643 964	0.3532 42
15	-0.3 848 5	0.0070 12	0.0558 63	0.0240 85	0.1265 13	0.0023 97	-0.022 19	-0.19836 6	0.0784 64
16	0.04 060 1	0.0076 09	0.0005 48	-0.005 75	-0.013 16	0.00E+ 00	-0.036 9	0.01191 487	0.0242 73

17	-0.2 547 4	0.0035 55	0.0045 35	0.0041 16	0.0146 08	0.0003 09	-0.014 66	0.00350 103	0.0360 57
18	-0.6 478 8	-0.004 68	0.0005 01	0.0016 36	0.0068 04	8.14E- 05	-0.009 46	0.01190 874	0.0793 09
19	1.19 323 2	0.0273 05	0.0031 44	0.0121 42	-0.039 97	0.0003 83	0.0631 41	-0.15925 14	-0.078 37
20	44.0 727 2	0.8910 59	-0.571 02	0.0556 94	-2.110 3	-0.016 19	2.2485 19	-8.92130 03	-7.646 67
21	-2.6 42	0.1065 14	0.0828 8	-0.017 67	0.2494 68	0.0052 34	0.1315 85	0.85130 582	-0.083 6
22	1.78 788 7	-0.002 75	0.0099 05	0.0470 76	-0.038 86	0.0005 13	-0.038 33	-0.37960 3	0.0607 24
23	4.86 222 4	0.1144 83	0.0464 85	0.0781 48	-0.036 84	0.0033 66	-0.086 14	-0.51792 89	0.0812 93
24	0.18 791 5	-0.000 63	0.0015 27	-0.001 61	0.0075 55	5.00E- 05	-0.029 37	0.00998 145	0.0050 01
25	0.31 411 8	0.0052 34	0.0004 16	0.0026 59	-0.011 31	8.20E- 05	0.0278 43	-0.05257 11	-0.033 38
26	19.8 118 6	0.2446 72	0.1007 72	-0.062 41	0.0754 09	0.0024 91	-1.668 61	0.65152 779	-0.515 17
27	-1.9 454 3	-0.034 52	-0.018 03	-0.003 28	-0.008 5	-0.001 06	-0.117 26	0.19098 14	0.0156 94
28	-110 .71	1.0483 43	0.2919 74	-1.447 15	-5.869 49	0.0242 63	-8.800 68	11.3273 585	16.001 03
29	15.2 857 5	0.1405 41	0.2240 74	0.3184 85	0.6063 44	0.0131 35	0.8364 56	-1.15088 71	-0.933 49
30	-11. 227 1	-0.003 75	-0.068 39	0.0117 69	-0.409 4	-0.002 62	2.2064 69	-1.79058 03	-0.360 38
31	-5.6 371 3	-0.019 77	0.0001 46	-0.121 72	-0.417 93	-0.001 58	1.8299 65	-3.20198 51	-0.049 62

白葡萄酒与酿酒葡萄得到的回归方程系数:

-2.35554	-1.06943	1.598193	1.7134627	-0.03717	102.51484	-0.268536	0.566294
0.000249	0.000202	0.000408	1.31E-05	9.17E-06	-2.47E-05	4.25E-05	-5.37E-05
0.002575	0.001569	0.005291	-0.001206	8.66E-05	0.0004894	0.0012593	-0.003479
0.440253	0.414856	0.738669	0.0405926	0.0392644	-0.064807	0.0148607	0.1879728
-0.04877	-0.02057	0.018729	0.0023716	-0.0008	-0.058756	-0.025844	0.2147606
0.087257	0.031049	-0.05639	-0.045546	0.0032239	0.0206396	0.0088521	-0.069886
-0.06584	-0.02346	0.141894	-0.017414	-0.0035	0.0064529	0.016108	-0.062512
0.099816	0.002452	0.007289	0.0048199	0.0010427	0.0267132	-0.019839	-0.029347
-0.0054	-0.00442	-0.00379	-0.001739	-0.000493	-0.005704	-0.000823	0.0182558
-0.00029	-0.00017	0.000519	-1.40E-05	-6.90E-06	5.42E-05	1.56E-05	-0.000128
2.37452	1.355511	-0.35547	0.4739263	0.095714	0.787234	0.3771104	-2.804898
-0.02149	0.003973	0.107894	-0.003084	-0.000456	-0.008817	-0.003754	0.039439
0.086149	0.062045	0.04959	0.0166795	0.0032515	0.0057724	0.0041308	-0.012542
0.039482	0.040719	0.154574	-0.001544	0.0017197	-0.014832	0.0019668	0.047007
-0.08005	-0.03958	-0.27089	-0.084158	-0.001857	-0.093556	-7.27E-05	0.2726567
0.013899	0.007144	0.035597	0.0001671	0.0006616	-0.002295	-0.001868	0.0143916
0.012566	0.005846	-0.00814	-0.007441	0.0004112	-0.001479	0.0026138	0.0009289
-0.00361	-0.00133	0.003016	0.0048359	-0.000191	-0.001164	-0.001489	0.0053543
0.000561	0.001253	-0.00516	-0.001644	-0.000142	-0.003652	-0.001595	0.0137413
-0.70896	-0.37384	-1.40993	0.1391313	-0.032644	-0.060883	-0.339276	0.7899917
0.001909	-0.01294	-0.15263	0.0106869	-0.000945	0.0083794	-0.008126	-0.014789
-0.00139	0.00054	0.014524	-0.004656	-6.88E-05	-0.000921	0.0017482	0.0003995
0.010943	0.012266	0.053802	0.0214318	-0.000563	-0.040743	-0.014165	0.1419574
0.000686	0.000112	-0.00131	-0.000313	6.85E-06	0.0001741	0.000157	-0.000813
0.001849	0.000553	-0.00305	-0.001202	4.67E-05	0.0007053	0.0004735	-0.002625
-0.037	-0.05887	-0.30367	0.0216243	0.002665	0.017119	-0.061982	0.0699443
0.020291	0.009041	0.02893	0.0021799	0.0008776	0.0252627	0.0148456	-0.105997
0.57825	0.636348	1.453228	0.1759816	-0.035902	-0.379627	-0.541037	1.852209
0.009554	0.004276	0.023023	-0.014223	0.0014596	-0.007277	-0.005721	0.0360671
0.050237	0.029607	0.050871	0.0124737	0.001765	0.0040985	0.0086001	-0.029089
-0.04279	-0.02381	-0.02216	-0.003581	-0.001453	-0.00871	-0.009539	0.0407392

%第四问数据处理:

```
for i=1:27
```

```
xiangqi(i,1)=sum(sum(a(3:5,(i-1)*10+1:(i-1)*10+10)+b(3:5,(i-1)*10+1:(i-1)*10+10))/2)/10;
```

```
waiguan(i,1)=sum(sum(a(1:2,(i-1)*10+1:(i-1)*10+10)+b(1:2,(i-1)*10+1:(i-1)*10+10))/2)/10;
```

```
koukan(i,1)=sum(sum(a(6:9,(i-1)*10+1:(i-1)*10+10)+b(6:9,(i-1)*10+1:(i-1)*10+10))/2)/10;
```

```
pingheng(i,1)=sum(sum(a(10,(i-1)*10+1:(i-1)*10+10)+b(10,(i-1)*10+1:(i-1)*10+10))/2)/10;
```

```
end
```

```
relation_1=corrcoef([xiangqi hj_zhibiao hp_zhibiao]);
```

```
relation_xiangqidefen(1,:)=relation_1(1,:);
```

```
relation_xiangqi_pjun=sum(abs(relation_xiangqidefen))/39;
```

```

relation_2=corrcoef([xh(:,31), hj_zhibiao hp_zhibiao ]);
relation_zondefen(1,:)=relation_2(1,:);
relation_zondefen_pjun=sum(abs(relation_zondefen))/39;

relation_3=corrcoef([waiguan hj_zhibiao hp_zhibiao ]);
relation_waiguan(1,:)=relation_3(1,:);
relation_waiguan_pjun=sum(abs(relation_waiguan))/39;

relation_4=corrcoef([koukan hj_zhibiao hp_zhibiao ]);
relation_kougan(1,:)=relation_4(1,:);
relation_kougan_pjun=sum(abs(relation_kougan))/39;

relation_5=corrcoef([pingheng hj_zhibiao hp_zhibiao ]);
relation_pingheng(1,:)=relation_5(1,:);
relation_pingheng_pjun=sum(abs(relation_pingheng))/39;

```

%第四问，第二小问：评价模型法

%主成分评价模型分析法程序

%红葡萄酒评价模型：

```

gj=[hp_zhibiao hj_zhibiao fangxiang_j_zhibiao fangxiang_p_zhibiao];
gj=zscore(gj); %数据标准化
r=corrcoef(gj); %计算相关系数矩阵%下面利用相关系数矩阵进行主成分分析，vec1的列为r
的特征向量，即主成分的系数
[vec1,lamda,rate]=pcacov(r) %vec为r的特征向量，lamda为r的特征值，rate为各个主
成分的贡献率
f=repmat(sign(sum(vec1)),size(vec1,1),1);%构造与vec1同维数的元素为±1的矩阵
vec2=vec1.*f %修改特征向量的正负号，使得每个特征向量的分量和为正
num=input('请选项主成分的个数:'); %通过累积贡献率交互式选择主成分的个数
df=gj*vec2(:,1:num)/100; %计算各个主成分的得分
tf=df*rate(1:num); %计算综合得分
[stf,ind]=sort(tf,'descend'); %把得分按照从高到低的次序排列
stf=stf'; ind=ind';
sum=0;
for i =1:27
    [s t]=find(h0==ind(i));
    sum=sum+abs(s-i);
end
Sum

```

%白葡萄酒评价模型：

```

gj=[bp_zhibiao bj_zhibiao];
gj=zscore(gj); %数据标准化
r=corrcoef(gj); %计算相关系数矩阵%下面利用相关系数矩阵进行主成分分析，vec1的列为r
的特征向量，即主成分的系数
[vec1,lamda,rate]=pcacov(r) %vec为r的特征向量，lamda为r的特征值，rate为各个主
成分的贡献率
f=repmat(sign(sum(vec1)),size(vec1,1),1);%构造与vec1同维数的元素为±1的矩阵
vec2=vec1.*f %修改特征向量的正负号，使得每个特征向量的分量和为正
num=input('请选项主成分的个数:'); %通过累积贡献率交互式选择主成分的个数
df=gj*vec2(:,1:num)/100; %计算各个主成分的得分
tf=df*rate(1:num); %计算综合得分
[stf,ind]=sort(tf,'descend'); %把得分按照从高到低的次序排列

```

```

stf=stf';ind=ind';
sum=0;
for i =1:28
    [s t]=find(h0==ind(i));
    sum=sum+(s-i)^2;
end
Sum

```

%第四问，主成分分析代码

```

sn=hp_zhibiao;
[m,n]=size(sn);
x0=sn(:,[1:n-1]);y0=sn(:,n);
sn=hp_zhibiao;
[m,n]=size(sn);
x0=sn(:,[1:n-1]);y0=sn(:,n);
r=corrcoef(x0) %计算相关系数矩阵
xd=zscore(x0); %对设计矩阵进行标准化处理
yd=zscore(y0); %对y0进行标准化处理
[vec1,lamda,rate]=pcacov(r); %vec为r的特征向量，lamda为r的特征值，rate为各个主成分的贡献率
f= repmat(sign(sum(vec1)),size(vec1,1),1); %构造与vec同维数的元素为±1的矩阵
vec2=vec1.*f %修改特征向量的正负号，使得特征向量的所有分量和为正
contr=cumsum(rate)/sum(rate); %计算累积贡献率，第i个分量表示前i个主成分的贡献率
df=xd*vec2; %计算所有主成分的得分
num=input('请选项主成分的个数:') %通过累积贡献率交互式选择主成分的个数
hg21=df(:,[1:num])\yd %主成分变量的回归系数,这里由于数据标准化，回归方程的常数项为0
hg22=vec2(:,1:num)*hg21 %标准化变量的回归方程系数
hg23=[mean(y0)-std(y0)*mean(x0)./std(x0)*hg22, std(y0)*hg22'./std(x0)]
%计算原始变量回归方程的系数
fprintf('y=%f',hg23(1)); %开始显示主成分回归结果
for i=2:n
    if hg23(i)>0
        fprintf(' +%f*x%d',hg23(i),i-1);
    else
        fprintf(' %f*x%d',hg23(i),i-1);
    end
end
end

```

附录二：

分类后芳香物质：

2.223	182.54	256.71	2.141	7.73	3.723	64.748	1.	3.36	0.9
		3		8			21	5	18
							7		
2.413	215.56	181.64	4.123	4.93	0.508	54.107	8.	0.77	0.0
		9					13	3	00
							2		
2.645	185.75	121.29	1.225	1.21	0.510	38.598	1.	0.00	0.0
		6					83	0	00

							4		
1. 724	107. 50	242. 71	3. 477	3. 26	2. 361	37. 277	1. 39	0. 00	2. 3
	2	1		3				0	91
							7		
1. 921	124. 00	264. 94	6. 883	3. 37	2. 661	44. 577	4. 03	0. 00	1. 7
	7	1		1				0	21
							2		
4. 136	193. 21	209. 68	4. 322	4. 24	0. 312	50. 434	6. 19	0. 00	32.
	9	3		1				0	417
							3		
1. 405	100. 90	197. 00	4. 447	3. 55	0. 717	37. 011	3. 56	0. 00	42.
	1	8		2				0	692
							3		
3. 450	162. 94	146. 03	1. 268	2. 94	0. 340	34. 003	0. 08	0. 18	0. 0
	9	9		8				1	00
							5		
2. 091	137. 00	201. 18	4. 097	3. 35	1. 097	49. 427	1. 80	0. 00	0. 5
	1	2		7				0	85
							0		
1. 744	94. 293	185. 98	2. 312	3. 68	1. 091	27. 646	0. 03	0. 00	0. 3
		0		5				0	08
							9		
1. 341	124. 04	69. 687	0. 096	2. 59	0. 652	38. 897	1. 08	2. 21	1. 1
	1			1				4	55
							4		
2. 456	241. 42	148. 66	0. 216	2. 73	0. 279	101. 12	2. 94	1. 94	0. 0
	7	6		7		5		2	00
							4		
1. 580	121. 58	217. 24	2. 688	3. 18	0. 864	47. 128	1. 79	0. 00	26.
	0	5		3				0	264
							1		
0. 637	89. 601	84. 454	0. 000	2. 10	0. 813	30. 587	1. 30	0. 68	0. 3
				4				8	44
							0		
1. 444	180. 16	213. 14	4. 425	4. 76	0. 510	38. 793	9. 93	0. 52	0. 0
	0	2		6				1	00
							2		
4. 469	96. 608	238. 26	4. 330	3. 70	1. 378	33. 004	3. 39	0. 00	0. 1
		3		2				0	97
							0		
2. 185	148. 88	280. 62	6. 605	3. 45	0. 438	47. 958	10. 3	0. 30	0. 0
	7	2		2				0	00
							73		
4. 597	202. 26	75. 773	0. 154	2. 67	0. 511	61. 884	4.	0. 00	0. 0

	6			1			28	0	00
							3		
1. 987	126. 22	221. 60	3. 174	3. 78	1. 748	42. 061	1. 95	0. 00	0. 6
	8	0		1			0	0	35
1. 740	144. 63	126. 57	1. 543	2. 62	0. 412	30. 991	3. 61	0. 00	0. 4
	2	8		2			9	0	21
2. 176	144. 66	180. 94	6. 847	3. 90	0. 523	39. 859	1. 98	0. 00	1. 9
	9	1		1			9	0	02
2. 002	137. 05	376. 71	8. 605	4. 01	3. 488	49. 315	1. 69	0. 39	0. 7
	6	4		8			0	1	82
1. 780	138. 50	237. 79	6. 229	4. 59	2. 607	55. 236	9. 39	0. 00	0. 0
	1	3		1			3	0	00
5. 430	177. 40	229. 91	3. 593	4. 25	0. 507	38. 803	7. 79	0. 51	0. 5
	8	4		5			4	3	13
1. 402	88. 192	184. 31	2. 329	3. 58	0. 985	42. 787	2. 45	0. 46	29.
		2		6			2	6	341
2. 622	97. 577	223. 12	2. 479	3. 68	1. 402	33. 093	0. 99	0. 00	0. 7
		3		4			7	0	63
1. 565	133. 50	277. 88	3. 041	3. 73	0. 421	43. 668	2. 80	0. 00	87.
	9	6		3			3	0	627

附录三：

第三问偏最小二乘主成分回归系数：

-	0	-	-	-	-	0	-	0	0	-	0	0	-	-	0	-	0	0	-	-	0	0	0	-	0	-			
0	.	0	0	0	0	.	0	.	.	0	.	.	0	0	.	0	.	0	0	.	.	0	0	.	.	0	.	0	
.	2	1	.	1	0	.	2	0	.	.	1	.	0	.	.	2	0	.	.	3	0	2	.	0	.
1	4	0	2	1	2	4	0	0	8	3	7	5	0	2	6	3	5	1	2	1	3	1	0	2	1	5	1	2	0
4	4	0	6	6	2	0	0	6	1	0	7	6	5	4	5	2	2	9	5	5	2	6	3	3	5	1	0	1	0
2	5	7	9	9	5	4	7	0	2	1	3	1	5	4	3	1	5	8	7	6	7	8	4	8	0	2	3	3	6
0	1	6	7	8	6	7	9	5	2	9	7	7	4	6	3	5	6	4	4	8	3	0	2	2	7	2	6	5	7
4	6	9	6	8	6	1	4	3	6		2	4	4	9	8	5	7	9	2	4		4	5		4	2	7	1	
-	-	-	-	0	-	0	-	0	-	0	0	-	-	-	-	0	0	-	-	-	0	-	0	0	0	0	0	-	0
0	0	0	0	.	0	.	0	.	0	.	.	0	0	0	0	.	.	0	0	0	.	0	0	.
.	.	.	.	1	.	0	.	0	.	2	0	3	1	.	.	.	0	.	0	2	0	0	2	.	1
2	2	0	1	3	0	6	1	0	1	2	7	4	0	0	2	2	2	2	2	1	9	2	6	3	9	0	5	1	1
3	2	9	6	6	9	7	1	2	3	3	7	0	4	8	2	9	2	8	4	9	5	0	8	2	3	0	3	4	3
2	4	3	1	4	7	4	1	5	3	9	0	0	4	8	7	4	8	2	9	2	8	4	8	2	9	8	3	4	5

4	6	7	7	6	7	4	7	4	3	0	4	2	0	7	8	0	8	9	0	8	9	7	4	8	7		2	8	0
1	9		6	3	6	5	4	8		3	4	4	3		3	5	9		9	9	5	6	8	6	2		6	6	2
0	-	-	0	-	0	-	0	0	0	-	0	-	-	0	-	-	0	-	-	-	-	-	-	-	-	0	0	0	0
.	0	0	.	0	.	0	.	.	.	0	.	0	0	.	0	0	.	0	0	0	0	0	0	0
0	.	.	0	.	0	.	1	6	1	.	0	.	.	3	.	.	0	0	0	0	0	0
5	1	0	0	3	9	0	4	2	8	0	3	0	1	3	2	1	2	0	1	1	1	0	1	1	2	6	4	4	5
3	7	4	5	8	9	1	1	6	1	3	9	9	4	4	3	2	5	7	2	3	9	8	1	6	5	6	8	8	0
8	8	8	3	5	2	8	9	6	5	2	3	9	5	5	5	6	5	2	4	8	9	5	5	5	7	4	5	5	3
5	7	8	6	7	9	7	2	0	4	3	9	2	1	7	7	8	8	0	9	2	6	1	5	8	5	6	1	0	3
7	2	6	6	9	2	5	4	1	5	1	8	1	1	1	8	5	1	8	4	6	1		5	8	8	3	2	2	7
-	-	0	0	0	0	-	0	0	-	-	0	0	0	0	0	0	-	-	0	-	0	-	0	-	-	0	-	-	-
0	0	0	.	.	0	0	0	0	.	0	.	0	.	0	0	.	0	0	0
.	.	0	1	0	1	.	0	0	.	.	2	1	2	0	1	0	.	.	1	.	2	.	0	.	.	1	.	.	.
3	0	5	7	6	4	0	5	3	1	1	2	0	1	4	0	3	1	2	3	1	1	4	8	4	1	4	1	1	0
2	4	4	9	8	9	7	5	0	2	5	5	6	2	3	8	7	5	8	7	3	7	5	8	0	6	6	7	1	2
0	7	5	4	4	2	7	4	6	8	0	3	4	4	0	0	0	8	0	9	4	7	3	7	6	7	4	1	9	0
8	9	3	3	1	0	7	2	3	2	8	2	5	4	7	8	4	6	6	9	9	0	7	9	4	0	9	8	6	6
7	6		8	7	1	6	4	8	7	4	3	8	9	6	1	6	3	4	2		4	8	6	8	2	8	5	4	
-	0	-	-	0	-	0	-	0	0	0	0	0	0	-	0	0	0	-	0	0	0	0	-	-	0	-	-	-	-
0	.	0	0	.	0	.	0	0	.	.	.	0	0	0	.	0	0	0	0
.	0	.	.	2	.	1	.	0	0	0	0	1	.	4	0	1	.	0	1	1	1	.	.	0
1	4	1	2	2	1	6	4	8	9	4	2	8	3	5	2	0	0	7	4	2	3	0	2	0	0	0	1	1	0
4	4	9	2	0	1	8	5	2	6	0	6	4	0	2	7	2	3	9	1	3	7	7	8	0	9	9	8	6	0
4	5	0	9	8	0	6	8	7	3	5	3	0	9	2	8	5	1	8	8	0	2	4	3	7	1	1	0	7	8
1	8	0	0	7	7	2	1	7	8	5	9	4	4	6	4	6	4	6	6	7	2	8	9	5	8	3	7	5	7
5	2	8	6	5	2	5	6	2	9	5	3	1	5	3	2	7	8	2	4	5	9	7	4	3	1	8	7	1	9
-	0	-	0	0	0	-	0	0	-	-	0	0	0	0	-	0	-	0	-	-	0	-	-	0	0	-	0	0	0
0	.	0	.	.	.	0	.	.	0	0	0	.	0	.	0	0	.	0	0	.	.	0	.	.	.
.	1	.	3	0	2	.	1	0	.	.	1	1	0	0	.	1	.	2	.	.	0	.	.	4	2	.	0	0	0
1	4	0	9	6	7	0	0	0	0	2	4	6	2	2	2	7	0	0	0	0	3	1	3	1	8	0	5	1	0
4	4	8	2	1	8	9	3	6	5	7	9	7	8	1	0	7	7	1	4	7	6	1	9	9	3	6	9	4	5
8	4	6	2	1	5	4	1	4	6	7	3	8	2	6	5	9	2	5	9	9	1	0	5	7	6	2	2	5	0
3	4	1	0	6	1	7	3	6	9	4	3	0	3	7	9	1	9	1		3	3	5	5	1	9	6	7	5	1
	2		5	9		6		7	1	9	4	5	6	6	2	1	8	9		5	2	9	2	6	3	1	9		9
-	0	-	0	0	-	0	-	0	0	-	-	0	0	-	-	-	0	-	-	0	-	0	0	-	0	-	0	0	-
0	.	0	.	.	0	.	0	.	.	0	0	.	.	0	0	0	.	0	0	.	0	.	.	0	.	0	.	.	0
.	0	.	2	2	.	2	.	2	1	.	.	1	3	.	.	.	1	.	.	1	.	0	1	.	2	.	0	1	.
1	8	2	1	5	0	4	3	2	9	0	0	9	3	2	2	0	2	0	1	0	2	2	5	2	2	0	4	2	0
1	5	0	7	2	5	4	8	0	2	1	9	5	0	2	1	2	1	8	0	2	8	0	8	2	4	2	2	2	8
5	4	7	0	4	4	4	0	9	3	9	3	0	3	6	6	0	2	1	7	6	4	4	5	9	2	8	6	5	3
6	8	4	8		7	6	2	0	1	7	3	0	5	3	4	4	7	8	7	4	2	1	1	4	8	9	1	5	9
2	4	2	9		6	8	6	9	2	2	3	1	5	5	8		8	5		4	4	9	7	4	4	3	9	8	4
-	0	0	0	0	-	-	0	-	-	0	0	-	-	0	0	-	0	0	-	0	-	-	-	-	0	0	-	-	0
0	0	0	.	0	0	.	.	0	0	.	.	0	.	.	0	.	0	0	0	0	.	.	0	0	.
.	0	1	3	1	.	.	1	.	.	4	3	.	.	1	1	.	1	1	.	1	1	0	.	.	0
1	3	1	5	6	2	0	4	0	0	8	5	1	0	7	0	3	9	2	0	8	1	1	1	0	6	2	0	0	1
1	8	0	2	6	5	0	6	7	8	8	0	5	4	0	7	2	7	0	1	5	5	1	5	7	5	6	0	0	3
8	8	5	5	7	5	6	0	0	4	7	9	6	6	6	5	6	5	3	2	7	9	0	7	9	4	7	0	2	5
4	4	1	4	0	6	4	3	1	6	5	4	2	6	7	5	8	4	3	7	2	1	3	7	7	8	6	9	0	6

9	8	9	3	3	2	6	3	4	4		4	9	6	7	4	3	1	8	1	9	3	5	9	3	9	5	1	7	9
-	-	-	0	-	-	-	0	0	0	-	-	-	-	0	0	0	-	-	0	0	0	0	0	0	-	0	0	-	0
0	0	0	.	0	0	0	.	.	.	0	0	0	0	.	.	.	0	0	0	.	.	0	.
.	.	.	4	.	.	.	0	1	0	1	0	0	.	.	0	4	0	3	1	2	.	0	0	.	0
2	0	0	1	0	0	0	9	0	9	0	2	0	1	6	9	5	1	4	3	0	7	3	4	1	0	3	3	1	0
2	4	2	9	1	4	6	7	6	2	2	6	4	0	6	7	1	2	5	5	5	9	3	0	9	9	1	8	7	6
6	0	8	5	3	2	8	3	1	2	5	2	6	6	2	6	3	2	0	0	9	0	3	6	8	9	3	6	2	2
2	6	3	9	7	8	0	2	5	2	5	2	1	2	3	8	3	3	6	9	8	8	8	9	7	7	0	6	0	8
2	4	7	1	9	1	3	3		5	1	9	1	9	7	9	6	5	5	6	8	8	9	8		4	3	4	8	2
-	-	0	-	-	0	0	0	-	-	0	0	0	0	0	-	-	-	-	-	-	0	0	0	-	0	-	-	0	0
0	0	.	0	0	.	.	.	0	0	0	0	0	0	0	0	.	.	.	0	.	0	0	.	.
.	.	0	.	.	0	1	1	.	.	0	1	0	0	1	0	3	0	.	1	.	.	4	1
2	2	0	1	0	8	7	0	2	0	3	4	3	6	8	1	0	1	0	2	0	2	2	3	0	0	0	4	6	7
8	0	6	2	1	5	8	0	1	8	6	1	8	1	4	7	2	8	9	3	1	3	4	8	2	8	8	2	8	7
6	7	6	7	6	9	2	5	8	8	2	3	7	1	8	1	5	8	2	9	1	6	6	2	8	7	4	7	5	6
4	1	5	4	7	1	3	6	5	7	4	0	1	5	3	2	6	8	2	3	4	9	8	4	9	8	6	7	1	3
9	7	5	5	2	6	8		8	7	3	5	3	4	4		5	4	4	9	7	6	4	3	8	6	4		2	2
-	-	0	-	-	0	-	0	-	0	-	-	0	0	0	0	-	-	0	-	0	-	-	0	0	-	-	0	0	-
0	0	.	0	0	.	0	.	0	.	0	0	0	0	.	0	.	0	0	.	.	0	0	.	.	0
.	.	0	.	.	1	.	0	.	0	.	.	0	0	2	0	.	.	1	.	1	.	.	0	0	.	.	2	0	.
3	0	9	1	0	3	0	7	2	9	0	0	2	4	0	5	0	0	2	2	5	4	2	7	6	3	2	7	4	2
2	7	1	3	1	9	0	8	3	8	9	7	9	3	0	7	0	4	9	1	8	6	4	6	9	3	3	2	7	9
7	6	3	2	3	3	8	1	5	7	0	6	5	8	5	1	3	4	2	7	9	3	1	6	3	5	5	7	6	8
1	7	6	8	4	9	4	4	6	4	9	4	1	3	3	8	8	8	1	7	4	9	3	4	7	9	5	8	8	4
	5	5	3	6	8	8	4	1	9	6	7	3	3	5	8		1	7	8	1	5	1	2	9	1	7	8	2	6
-	-	0	0	-	0	0	-	0	0	0	-	-	0	-	0	-	0	0	0	-	0	0	0	0	0	-	-	-	0
0	0	.	.	0	.	.	0	.	.	.	0	0	.	0	.	0	.	.	.	0	0	0	0	.
.	.	1	0	.	1	2	.	0	3	0	.	.	1	.	3	.	2	1	0	.	0	1	0	1	0	.	.	.	1
2	0	4	4	1	8	0	0	8	2	6	0	0	9	0	9	0	8	1	8	3	1	6	6	9	4	2	1	2	0
8	6	4	0	1	6	3	5	1	5	3	0	4	3	1	1	8	4	0	0	9	2	3	8	8	4	0	6	6	2
6	8	6	3	7	1	8	1	8	6	8	5	5	9	5	7	9	9	4	9	8	8	3	9	7	7	1	9	1	9
2	5	8	4	9	9	3	4	5	5	0	3	6	7	5	2	0	6	0	3	1	0	0	7	4	9	2	7	5	1
6	1	1	3		7	7	7	2	6	8	2	4	9		8	3	1	7	6	4	4	6	1	4	6	3	2	4	3
-	-	0	-	0	0	0	0	-	0	-	-	0	-	-	-	0	-	0	-	0	-	-	0	-	0	0	0	-	0
0	0	.	0	.	.	.	0	.	0	0	.	0	0	0	0	.	0	.	0	0	.	0	0	.	.	.	0	.	.
.	.	1	.	0	2	1	0	.	0	.	.	1	0	.	1	.	.	0	.	.	3	3	2	.	0
2	1	0	1	2	2	0	5	2	7	0	2	2	2	0	1	2	2	0	7	0	0	4	1	1	4	9	6	1	6
7	2	1	6	1	5	4	0	8	1	2	7	0	3	5	2	8	7	6	1	2	3	9	7	8	3	5	7	4	8
2	8	6	8	3	0	3	0	0	6	2	9	9	1	9	2	9	2	4	0	3	1	6	0	3	5	8	2	5	7
3	7	1	6	9	4	9	6	4	1	9	2	7	4	6	2	5	2	9	5	0	5	0	9	9	5	2	0	3	0
3	5	3	7	3	4	7	7	1	5	9	4		3		6	1	7		8	9	3	1	4	4	3		3	7	9
-	-	-	-	-	0	0	0	0	-	0	-	0	0	0	-	-	-	0	-	0	0	-	0	-	0	-	0	-	-
0	0	0	0	0	0	.	0	.	.	.	0	0	0	.	0	.	.	0	.	0	.	0	.	0	0
.	1	2	2	0	.	1	.	1	1	0	.	.	.	2	.	1	4	.	2	.	0	.	1	.	.
0	0	4	0	1	2	4	5	2	0	8	0	0	1	2	0	2	0	5	1	9	8	0	1	0	7	1	1	1	1
2	2	2	4	5	4	6	9	3	4	3	5	3	3	7	8	0	3	8	6	7	6	8	5	4	1	6	6	9	6
4	6	3	4	3	5	3	9	1	3	5	4	8	8	1	7	2	6	8	7	3	1	3	0	1	8	7	3	2	4
1	6	3	6	7	3	1	9	7	3	3	6	5	3	6	8	5	7	7	5	5	5	4	5	6	2	5	3	8	6
7	6	6	3	7	1	5	1	3	4	9	8	8	6	4	2	7	1	1	7	7	7		8	5	5	4	8	4	3

```

- 0 - 0 - - 0 0 0 - - - - 0 - 0 0 - 0 0 - - 0 - - - - 0 0 -
0 . 0 . 0 0 . . . 0 0 0 0 . 0 . . 0 . . 0 0 . 0 0 0 0 . . 0
. 0 . 0 . . 3 1 1 . . . . 0 . 0 0 . 0 1 . . 0 . . . . 0 0 .
2 0 0 4 1 3 9 9 0 1 1 1 2 0 2 6 6 2 5 8 0 0 5 3 1 0 2 5 0 0
1 9 1 1 2 8 9 1 1 6 4 6 3 2 5 5 9 6 5 8 5 5 0 8 8 4 5 3 4 6
1 6 4 8 1 0 5 2 9 6 6 7 3 8 4 5 8 7 4 0 8 4 3 9 8 0 8 9 8 0
3 6 3 1 9 1 3 2 1 1 5 9 6 7 5 0 5 3 3 0 4 8 7 6 8 7 2 9 7
7 3 6 9 1 7 6 6 6 9 5 6 6 4 7 4 8 3 3 2 5 2 5 1 2 3
- 0 0 - 0 - - 0 0 0 0 - 0 - - - - - 0 - - - 0 0 0 - - - 0
0 . . 0 . 0 0 . . . . 0 . 0 0 0 0 0 0 . 0 0 0 . . . 0 0 0 .
. 3 0 . 0 . . 2 0 1 1 . 0 . . . . . 3 . . . 2 1 0 . . . 0
0 5 7 1 7 0 0 6 3 2 8 0 9 1 1 4 0 0 1 5 0 1 1 4 4 2 3 2 1 1
9 3 8 5 2 3 6 4 5 4 0 1 1 6 0 4 9 8 5 5 8 5 4 3 5 7 0 3 2 4
6 1 2 4 6 0 0 1 4 3 2 5 3 8 5 4 7 2 5 2 0 1 2 3 3 6 2 7 5 4
9 0 6 7 9 0 3 0 5 9 1 5 0 5 1 5 4 5 2 9 0 4 5 0 0 7 5 7 2 2
8 6 3 4 9 7 9 9 7 3 4 9 6 1 6 1 8 2 6 3 9 1 8 9 7 8 7
- 0 0 - 0 - - 0 0 - - - 0 0 0 0 - 0 - - - - - 0 - 0 - - - 0
0 . . 0 . 0 0 . . 0 0 0 . . . . 0 . 0 0 0 0 0 . 0 . 0 0 0 .
. 3 0 . 0 . . 0 1 . . . 1 0 2 1 . 0 . . . . 1 . 2 . . . 1
0 4 5 0 8 0 0 6 1 5 0 3 1 4 7 7 0 3 0 3 1 1 0 1 0 0 0 0 1 4
2 6 8 7 2 8 4 4 8 2 6 7 5 5 1 1 3 3 0 3 7 5 0 0 5 9 2 8 9 5
9 1 9 5 4 1 1 9 3 4 1 4 2 9 9 3 9 7 6 9 4 6 9 8 0 1 2 0 9 6
8 3 7 6 4 2 7 6 1 4 9 0 9 0 2 5 1 4 2 3 4 8 7 1 5 0 4 5 3
4 9 1 6 1 2 4 9 1 9 3 4 2 2 9 2 1 1 8 2 8 7 5 1 5 6
- 0 0 - 0 - - 0 - 0 0 0 - 0 - - 0 - 0 - 0 0 0 - - 0 - - -
0 . . 0 . 0 0 . 0 . . . 0 . 0 0 . 0 . 0 . . . 0 0 0 . 0 0 0
. 3 1 . 0 . . 2 . 3 0 1 . 0 . . 3 . 1 . 0 0 2 . . . 1 . . .
0 4 6 0 8 0 0 3 0 3 3 8 0 1 0 1 7 0 4 3 3 5 9 0 3 0 9 0 2 0
9 1 3 8 5 3 1 6 1 7 4 3 2 1 2 2 5 1 4 4 7 5 7 9 1 8 7 0 8 2
3 8 5 6 7 5 9 1 7 6 8 9 1 8 8 4 1 8 7 1 9 5 0 2 3 9 3 2 9 7
3 0 8 3 5 7 9 3 3 6 6 1 3 2 1 7 1 2 4 0 3 3 3 3 1 5 2 8 4 9
9 2 3 8 9 3 9 2 4 1 2 3 2 3 8 1 3 6 2 9 1 6 5 3 5 4 4
- - - - 0 - - 0 0 - - 0 0 0 0 0 0 0 0 0 0 - 0 - 0 0 0 0 - 0
0 0 0 0 . 0 0 . . 0 0 . . . . . . . . . 0 . 0 . . . 0 .
. . . . 0 . . 2 1 . . 1 0 4 1 1 1 0 0 3 2 . 0 . 0 2 0 1 . 0
1 1 0 4 9 0 2 1 9 0 0 2 1 7 4 3 7 0 4 4 8 0 9 0 6 8 3 1 0 8
0 2 9 1 2 8 4 3 7 2 4 4 4 2 6 3 2 8 0 7 0 9 2 2 0 2 1 5 1 4
2 5 3 3 4 5 0 1 2 1 5 2 8 5 0 3 9 2 8 7 8 9 7 8 6 7 2 7 8
2 3 5 3 1 5 3 2 5 0 4 4 3 9 4 3 2 6 0 7 7 2 7 8 9 0 8 5 9
9 1 2 7 4 1 6 2 3 4 4 2 7 3 5 1 4 8 5 7 9 7 8 6
0 0 0 0 - 0 0 - - 0 0 0 - - 0 0 0 - - 0 0 0 - 0 - 0 - 0 0 0
. . . . 0 . . 0 0 . . . 0 0 . . . 0 0 . . . 0 . 0 . 0 . .
1 2 3 0 . 1 2 . . 0 0 1 . . 0 2 2 . . 0 1 0 . 2 . 4 . 2 1 0
1 0 0 0 2 6 4 1 0 5 2 9 0 1 9 3 6 1 1 4 7 5 1 1 1 3 2 4 4 3
4 6 8 3 3 6 8 2 2 6 3 1 0 8 7 8 1 6 2 7 1 9 2 2 2 1 3 4 3 2
3 1 4 1 3 7 5 3 0 5 8 5 7 8 9 4 1 4 5 5 7 2 4 6 2 4 8 6 6 5
1 8 4 5 7 5 4 8 5 7 3 1 8 8 9 4 2 9 8 2 8 7 5 8 4 9 1 3 1 9
7 6 1 5 4 9 1 3 1 6 3 1 9 2 3 4 3 4 5 2 4 4 6 8 6 2 4 7
- - - 0 0 - - 0 0 0 - - - - - 0 - 0 0 - - 0 - 0 - 0 - 0 -

```


0	0	0	.	.	0	0	.	.	.	0	0	0	0	0	.	0	.	.	0	0	.	0	.	0	.	0	0	.	0	
.	.	.	0	3	.	.	1	0	2	2	.	1	1	.	.	1	.	1	.	1	.	.	2	.	
1	0	2	0	7	0	2	9	9	2	2	0	2	3	0	4	0	0	3	0	0	2	0	9	1	8	1	0	8	0	
4	2	2	0	8	7	6	9	5	9	8	8	0	9	5	5	2	6	5	5	9	8	5	6	5	5	3	0	3	2	
9	3	3	9	0	7	5	2	5	7	3	3	4	3	8	3	2	7	2	5	5	5	5	8	5	9	4	8	1	7	
7	1	1	3	2	5	6	4	2	4	0	4	6	2	9	5	7	8	5	9	8	1	4	3	9	4	6	1	4	8	
8	8	7	1			6	7			5	8	2		1	4	1	6	4	2	5	5	8	3	9	7	7	1	9	2	
-	0	0	-	0	-	0	0	0	-	0	-	0	0	0	-	-	0	-	0	-	0	0	-	0	-	0	0	0	-	
0	.	.	0	.	0	.	.	.	0	.	0	.	.	.	0	0	.	0	.	0	.	.	0	.	0	.	.	.	0	
.	3	0	.	0	.	0	0	1	.	1	.	0	1	1	.	.	2	.	1	.	3	1	.	0	.	0	3	4	.	
1	8	9	0	6	0	4	3	0	0	2	0	5	5	9	0	0	0	1	2	2	2	2	0	5	2	5	8	6	1	
4	4	7	5	7	1	5	0	5	1	6	6	9	5	9	3	8	1	4	8	2	2	4	9	9	3	9	5	8	2	
1	9	9	5	6	8	2	3	3	1	9	6	5	1	2	9	4	5	8	7	6	0	4	5	6	1	0	5	4	6	
9	6	7	8	8	2	7	7	1	0	6	2	8	1	0	9	0	2	7	5	1	7	0	9	3	1	1	4	6	3	
7	1			6	4	3	7	3	4	8		6	2	1	8	7	3	5	2	2	4	9	5	7		5	2	8	6	
0	-	0	-	0	0	0	0	0	-	-	0	0	-	-	0	-	-	-	-	-	-	0	0	0	-	0	0	-	-	
.	0	.	0	0	0	.	.	0	0	.	0	0	0	0	0	0	.	.	.	0	.	.	0	0	
1	.	1	.	4	0	1	0	2	.	.	3	2	.	.	0	2	1	1	.	0	2	.	.	
2	2	1	0	2	3	9	8	4	1	0	3	1	1	0	7	1	3	0	0	2	0	6	4	2	0	1	4	1	1	
9	0	3	3	2	4	0	2	2	3	2	3	6	2	2	3	7	6	1	2	3	7	7	9	1	5	6	5	4	6	
9	7	9	9	6	0	9	5	5	4	7	6	9	3	2	7	5	4	9	3	0	6	2	9	4	6	0	2	2	9	
3	1	4	5	2	1	8	3	2	4	8	9	7	5	8	3	3	7	9	3	8	7	4	8	6	6	7	9	5	5	
1	1	7	7	4	6	7	4	6	8	7	1	1	9	8	6	2	3		4	3	6	8	4	3	3	8	9	6	9	
0	-	0	-	0	0	0	0	-	-	-	0	0	-	0	-	-	0	-	-	0	0	0	-	-	-	-	-	-	-	
.	0	.	0	0	0	0	.	.	0	.	0	0	.	0	0	.	.	.	0	0	0	0	0	0	0
2	.	2	.	1	1	1	1	.	.	.	0	0	.	0	.	.	5	.	.	2	1	0	
0	1	4	0	8	1	8	7	0	1	3	6	1	0	4	1	0	7	2	0	0	0	5	1	1	0	2	0	1	1	
2	6	4	4	9	3	7	1	1	4	2	7	1	1	1	6	2	7	2	1	4	3	4	3	1	2	9	6	1	1	
0	0	3	6	4	6	5	6	4	9	9	5	0	4	5	4	3		4	3	8	0	3	9	6	2	8	8	3	1	
9	1	7	8	5	0	6	7	3	9	9	7	0	6	0	9	2		7	3	7	6	6	9	8	3	8	2	3	9	
5	1	7	8	5	9	8	3	2	4	6	8	9	9	2		7		6		3	1	8	5	1	9		7	8	1	
-	-	-	0	-	-	0	0	-	-	-	0	0	-	0	-	0	0	0	0	-	-	0	0	-	-	0	0	0	0	
0	0	0	.	0	0	.	.	0	0	0	.	.	0	.	0	0	0	.	.	0	0	
.	.	.	1	.	.	0	0	.	.	.	1	3	.	0	.	2	3	2	1	.	.	1	3	.	.	1	1	0	0	
2	0	0	2	2	3	7	3	0	2	1	9	3	2	2	0	5	2	1	9	0	0	2	4	0	0	5	0	1	4	
2	9	8	7	9	0	4	5	4	2	3	7	1	3	1	9	4	8	7	4	7	8	5	4	1	6	2	1	2	0	
0	5	8	9	0	7	1	8	7	4	9	0	8	6	5	2	0	2	7	8	8	9	3	4	2	8	5	0	6	4	
9	0	9	3	1	1	7	3	1	7	5	1	6	8	3	7	7	8	1	2	3	1	4	7	6	4	7	0	7	7	
1	4	2	1	6	7	7	9	8	2	2	3	1	6	8	9		2	6	8		2	3	3	1	4	7	6	6	9	
-	-	0	-	0	0	-	0	0	-	0	-	0	-	-	0	0	0	0	-	0	0	-	-	-	-	-	-	0	0	
0	0	.	0	.	.	0	.	.	0	.	0	.	0	0	0	.	.	0	0	0	0	0	0	.	.
.	.	1	.	0	3	.	0	3	.	3	.	2	.	.	1	1	1	0	.	2	0	1	0	
2	0	3	0	1	0	1	1	7	2	6	0	5	2	4	5	6	0	7	0	2	1	0	1	0	1	0	0	4	0	
0	8	9	9	0	1	1	2	1	1	5	9	6	1	2	6	4	2	2	7	5	0	0	4	1	2	3	9	4	2	
6	2	8	9	3	8	9	6	7	9	3	8	5	5	6	3	8	3	8	4	1	0	8	8	4	0	1	5	7	1	
7	1	9	8	4	9	7	8	6	8	2	7	2	3	4	0	1	2	9	7	5	7	8	1	0	1	5	2	3	5	
7	4	4	6	6	2	4		6	3	5	6		8	2	9	6	1	4	4	3	2	1	3	9	6	9		9	9	
0	-	0	0	0	-	0	0	0	0	-	-	-	0	0	-	0	-	0	0	0	0	-	0	0	-	0	-	0	0	
.	0	.	.	.	0	0	0	0	.	.	0	.	0	0	.	.	0	.	0	.	.	

0	.	3	0	2	.	4	1	1	1	.	.	.	0	1	.	1	.	2	0	1	0	.	1	2	.	3	.	1	0
9	1	1	6	3	0	0	9	3	1	0	2	0	6	1	0	7	0	7	6	1	6	3	6	1	0	6	1	4	6
7	1	6	7	0	5	1	5	9	4	0	7	6	1	9	5	2	2	1	1	4	2	0	8	6	3	8	7	3	0
3	1	4	3	2	9	0	1	3	6	8	7	5	9	3	1	2	0	9	9	9	8	3	9	2	8	5	0	4	4
3	8	9	2	7	9	0	6	1	3	3	9	4	9	2	7	8	8	8	1	2	6	0	7	4	9	0	7	3	1
2	3		6		2	9	8	1	5	3	1	1	2	6	7	9	3	4	4	5	4	4		6	3	1	4	3	3
0	-	-	0	0	-	0	0	-	0	0	-	0	0	-	0	0	0	-	-	-	-	-	-	0	-	-	0	0	0
.	0	0	.	.	0	.	.	0	.	.	0	.	.	0	.	.	.	0	0	0	0	0	0	.	0	0	.	.	.
2	.	.	0	0	.	0	3	.	2	1	.	5	0	.	1	1	0	0	.	.	0	0	2
1	1	1	2	3	2	6	0	1	7	5	0	1	0	0	9	3	2	2	1	1	0	1	1	0	0	1	9	5	7
3	4	5	3	5	3	4	6	5	6	7	8	2	9	0	0	9	7	8	4	3	4	9	9	9	1	1	9	7	6
7	8	7	1	6	1	6	5	2	5	6	7	8	9	5	7	7	8	6	7	7	7	3	2	0	3	1	9	0	1
8	4	6	0	5	6	4	1	1	1	4	4	0	3	8	9	2	7	5	9	7	0	8	1	6	9	5	5	7	1
3	9	7	7	9	1	9	6	1	2	3	7	9	8	8	3	8	1	6	4	6	6	6	7	6	9		4	5	
0	0	-	-	-	0	0	0	-	-	0	0	-	-	0	0	0	0	-	0	-	-	0	-	0	0	0	-	0	-
.	.	0	0	0	.	.	.	0	0	.	.	0	0	0	.	0	0	.	0	.	.	.	0	.	0
1	1	.	.	.	2	2	1	.	.	1	0	.	.	0	2	2	1	.	1	.	.	0	.	0	0	2	.	0	.
2	2	3	0	0	2	1	8	0	0	3	3	1	0	4	0	2	1	2	1	1	2	3	0	4	9	6	2	4	5
5	5	8	3	1	2	6	8	8	7	4	3	1	8	7	4	2	7	2	7	1	4	5	5	4	2	0	2	8	0
8	4	1	1	5	7	4	6	0	5	1	0	1	2	2	0	5	4	3	1	7	8	4	1	0	3	9	5	4	7
3	6	4	1	1	9	7	9	8	6	9	3		8	1	8	7	8	1	6	2	0	1	6	8	1	7	4	4	2
9	9	9	3	3	4	8	2	2	1	4	1		3	4	6	7	6	3	5	9	6	5	5	5		9	7	9	9
0	0	-	0	0	0	0	0	-	-	-	0	-	-	-	0	-	0	0	0	0	-	0	0	-	-	0	0	0	0
.	.	0	0	0	0	.	0	0	0	.	0	0	.	.	0	0
0	2	.	0	1	3	2	0	.	.	.	1	.	.	.	0	.	0	0	1	1	.	0	0	.	.	0	1	0	6
5	1	3	1	3	4	4	7	0	0	0	3	1	0	0	6	0	4	3	1	2	1	2	4	0	3	3	4	1	1
2	9	1	3	4	4	8	9	1	8	7	4	8	4	2	3	3	5	4	4	6	9	5	1	5	0	0	4	1	8
4	9	1	0	3	4	1	1	0	3	4	7	4	9	6	5	6	4	7	9	8	3	5	1	3	7	2	1	5	5
9	6	0	3	2	1	4	8	8	0	5	1	2	7	0	5	3	1	8	8	7	5	0	4	0	6	3	3	1	7
9	2	2		7	6	6	6	9	4	5	6	3	2	2		5	4	4	5	6	2	6	9	3		1	2	3	8