

南京审计大学16级物流管理系

作业1

本次作业覆盖课程1-4内容

截止时间：2018.10.11

需使用的数据集：

beer.Rda

googleplaystore.csv

数据集存放地址：

<https://github.com/exoplanetX/DataAnalytics/tree/master/%E7%AC%AC%E4%B8%89%E6%AC%A1%E8%AF%BE%E6%95%B0%E6%8D%AE>

习题1

创建矩阵 $A = \begin{bmatrix} 7 & 12 & 11 & 17 \\ 2 & 32 & 9 & 30 \\ 3 & 22 & 53 & 47 \end{bmatrix}$

取A矩阵的前三列生成新的矩阵，并对其做转置处理后求逆矩阵，将该逆矩阵记为Y

令X为未知向量 $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 求方程组 $YX = \begin{bmatrix} 17 \\ 9 \\ 3 \end{bmatrix}$ 的解

习题2

加载数据文件beer.Rda，该文件包含一个某厂按季度记录的啤酒销售量数据，

问题1:请将该数据做成时间序列格式，开始时间是1956年第一季度。

问题2:从该时间序列上截取2000年第一季度后的数据，并使用plot函数将数据以图形方式展现。

习题3

经理人	晋升日期	国籍	性别	年龄	Q1	Q2
William	10/24/08	US	M	32	5	4
Sarah	10/28/08	US	F	45	3	5
Debby	10/01/08	UK	F	25	3	5
Johnson	10/12/08	UK	M	39	3	3
Rosa	05/01/09	UK	F	99	2	2

上表是一张管理人员状况的记录表，请使用R语言创建该表，数据集名称自由设定，变量名称可更换

- 1. 请将“晋升日期”设为日期型变量，“国籍”“性别”设置为因子型变量
- 2. 将“经理人”属性设置为数据集的对象名称，即row.names
- 3. 删除“年龄”属性为99的对象
- 4. 在剩下的记录数据当中，计算晋升日期最大差距为多少天
- 5. 令Q1和Q2相加求出每个经理人的总分，并将总分变量插入数据集中

习题4.运输业务表

	南京	上海
南京	0	300
上海	500	0

右侧行名称代表出发地，列名称代表目的地，即300代表由南京运往上海的货运量为300吨，500则代表上海运往南京的货物为500吨

yunliang.Rda文件中存储量某物流企业的网络节点间的运输量

某物流企业在各省省会城市建立了物流中心，用于辐射省域范围内业务。各物流中心之间的业务量记录在yunliang.Rda文件中，请加载该文件中的数据，用合适的变量类型处理该数据集。

操作1:计算该企业一年全国范围内运输总量，区域中最大的运输量为多少。

操作2:现计划在省域物流中心基础上建立区域物流中心，合并地区范围内业务以提高运转效率降低成本。划分为

中部：武汉，成都，重庆，郑州，长沙，南昌

北部：北京，天津，石家庄，太原，呼和浩特，济南

东部：上海，南京，合肥，杭州，青岛，厦门，福州

东北：沈阳，长春，哈尔滨

南部：广州，海口，贵阳，昆明，深圳，南宁

西部：拉萨，乌鲁木齐，西宁，银川，兰州，西安


请合并矩阵为大区域间的运输业务表

最终形成如下合并后的运输表，'?'代表未知运输量

	北部	中部	东北	西部	南部	东部
北部	0	?	?	?	?	?
中部	?	0	?	?	?	?
东北	?	?	0	?	?	?
西部	?	?	?	0	?	?
南部	?	?	?	?	0	?
东部	?	?	?	?	?	0

习题5.记录数据操作

请以合适的方式和变量加载数据文件googleplaystore.csv。导入后的数据集将在环境中显示如右侧图，该数据集描述了当前google play上全部app的特征信息。

 googleplaystore

问题1:googleplaystore数据集包含多少变量多少对象？

问题2:目前免费和收费app各有多少？

问题3:所有app中评分超过5的占总数比例多少？

问题4:请查找所有app中阅读量最大的，并判断阅读量最大的app是否是评分最高的？