实 验 报 告

学 院：

课程名称：

专 业：

姓 名：

学 号：

1. 实验目的

掌握数据读入的各类函数；

简单的数据处理：新变量的生成，数据集的合并。

1. 实验内容
   1. 使用两种函数读入数据集“历年性别比.csv”、“各省人口基本信息.csv”这2个数据集；
   2. 根据数据集“历年性别比.csv”，计算出历次人口普查结果中男性和女性的人口数量，并在原数据中生成相应的新变量（单位：万人）；

（提示：男性人口数量=性别比/(性别比+100)\*人口总数

女性人口数量=100/(性别比+100)\*人口总数）

* 1. 首先将数据集“地区分类.csv”进行转换，然后与数据集“各省人口基本信息.csv”按“省份”这一变量进行合并，合并后形成新的数据集“total\_info”。

1. 实验要求

完成实验内容中的每一步，请附上代码，并使用“head()”函数对每一个数据集结果前15行进行部分展示，贴出结果。

rm(list = ls())

setwd('C:/Users/Desktop/R6,8')

Sys.setlocale("LC\_ALL", 'Chinese')

library(tidyverse)

#2.1

sexrat<- read.table("历年性别比.csv", sep = ",",header = T,encoding='UTF-8')

basinfo<-read.csv('各省人口基本信息.csv',sep = ",",skip = 1,header = T,encoding='UTF-8')

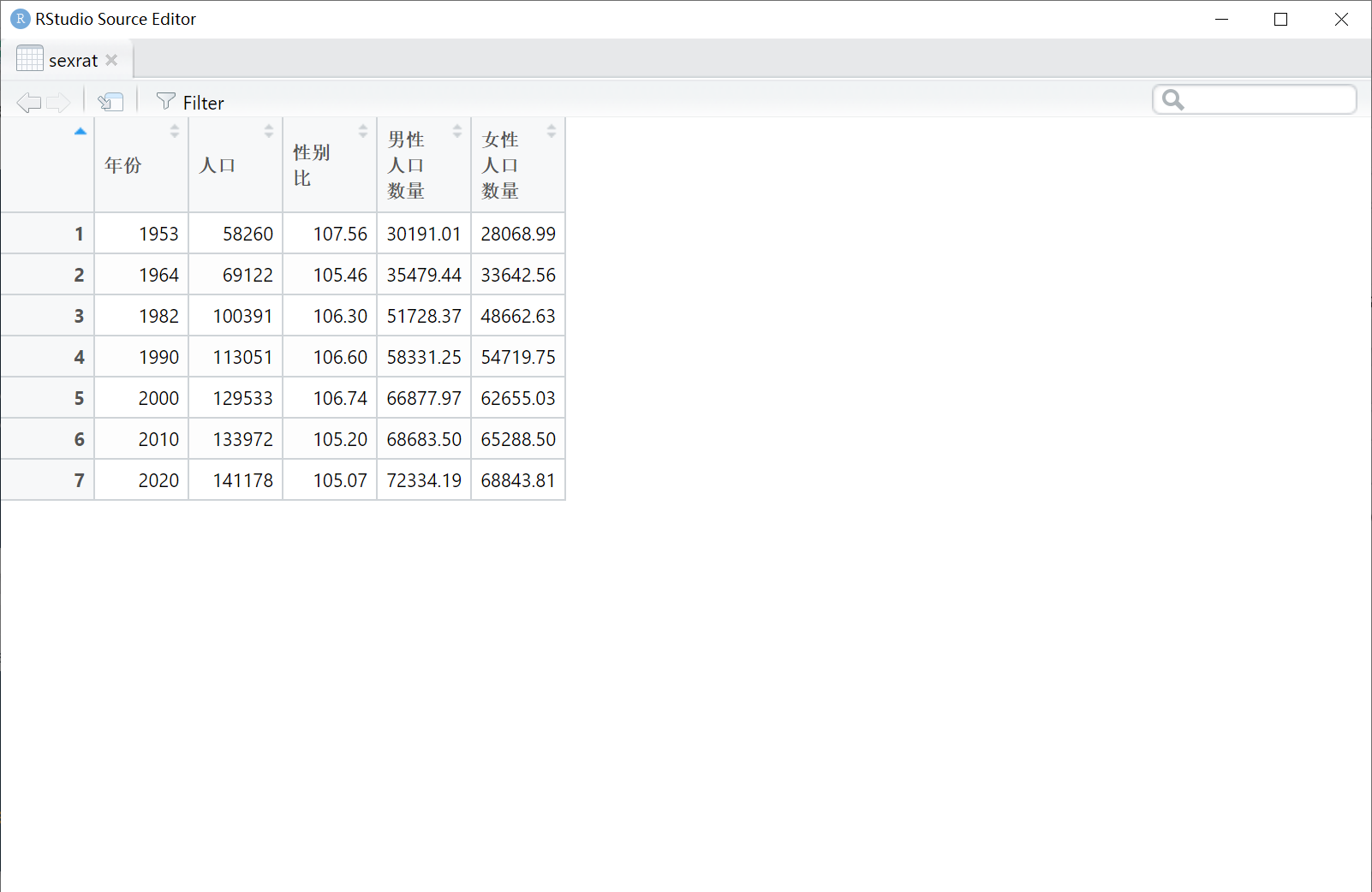
names(basinfo)[1]="province"

#2.2

sexrat$男性人口数量<-(sexrat[,3]/(100+sexrat[,3]))\*sexrat[,2]

sexrat$女性人口数量<-(100/(100+sexrat[,3]))\*sexrat[,2]

view(basinfo)



#2.3

loc\_c<-read\_csv('地区分类.csv')

newloc<-gather(loc\_c,na.rm = T,key = 'loc',value = "province")

for (i in c('省','市','自治区','维吾尔','回族','壮族')){

newloc$province<-str\_replace(newloc$province,i,'')}

basinfo$province=str\_remove\_all(basinfo$province,' ')

total\_info<-merge(basinfo,newloc,by="province",all = T)

view(total\_info)

