

بخش اول:

① ذخیره داده‌ها در یک متغیر: چون نوع متغیر آن ذکر نشده بود، کل داده‌ها را به یک شیء `data frame` ذخیره کردیم:

`all-exam-data <- read.csv("exam-data.csv", header=T)`

↓
رابطه نام فایل داده نظر بگیرد
↓
از نامش بخواند

② ابتدا سطر اول دوم `all-exam-data` را جدا کردیم هر کدام را در وکتورهای جداگانه ریختیم و بعد به دستورات تابع:

`mean()`, `min()`, `max()`, `var()`, `median()`, به ترتیب میانگین، بیشترین، کمترین، واریانس، و میانگین

واریانس و میانگین هر کدام را بدست آوردیم و بعد با استفاده از تابع `summary()` از درست بودن جواب‌ها اطمینان حاصل کردیم و اعداد بدین شکل اند:

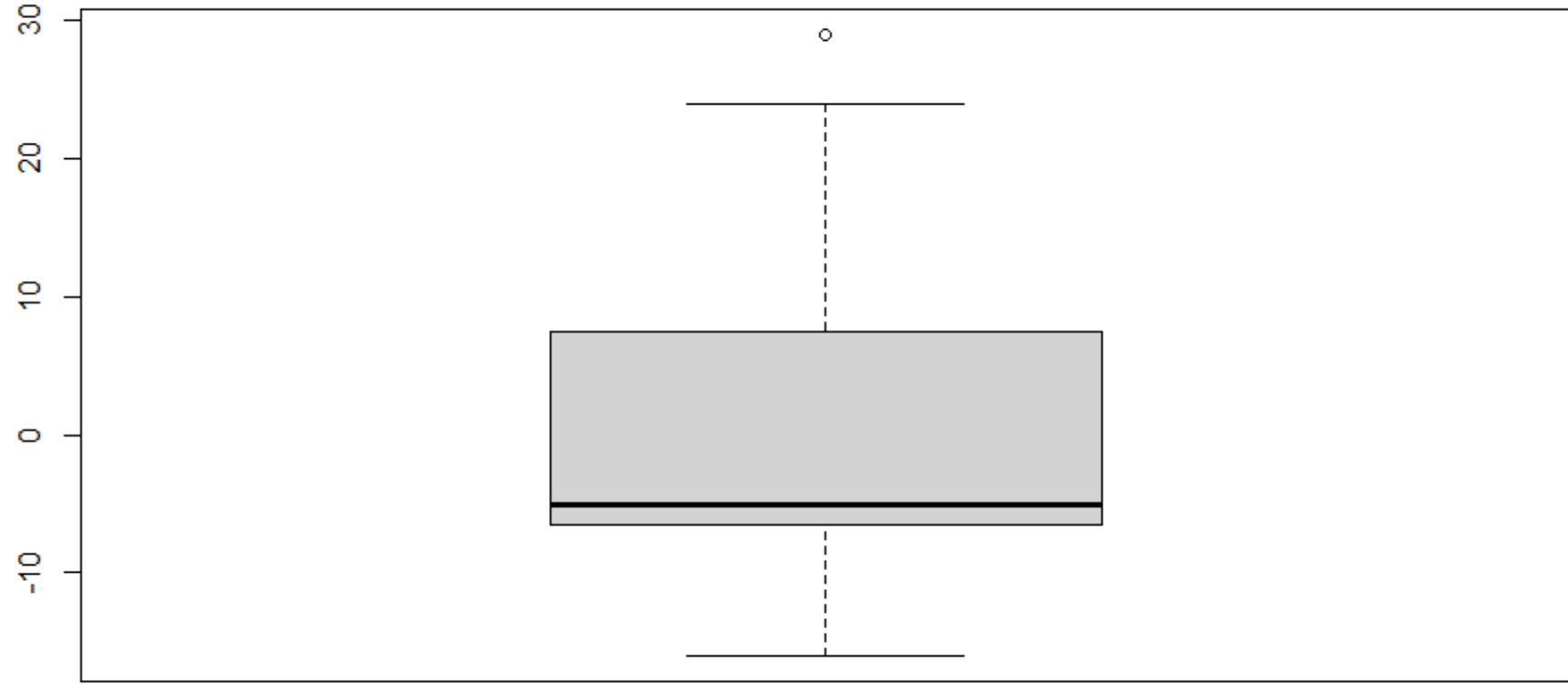
first-exam-mean = 79.25 second-exam-mean = 80.55
first-exam-min = 59.00 second-exam-min = 64.00
first-exam-max = 99.00 second-exam-max = 100.00
first-exam-var = 102.302631578947 second-exam-var = 160.786842105263
first-exam-median = 77.50 second-exam-median = 75.00

البته نکته جالب‌دهنده‌ای است که با استفاده از `c()` وکتور تولید کردیم که البته می‌توانستیم به‌دراهم، و دیگری اینکه با استفاده از `summary()` تنها واریانس را بدست می‌آوردیم.

③ در این سؤال من فرض را بر این گذاشتم که تفاوت امتحان دوم و اول، دوی مبنای اوی است و نه قدر مطلق یا به‌عکس آن و طبق آن یک وکتور `difference-vec` درست کردیم که اختلاف مقادیرهای `Index`‌های مساوی تغییر یافته آن‌ها در آن ذخیره کرده و بعد با استفاده از `boxplot()` نمودار جعبه‌ای آن را رسم کردیم (صن بعد).

در مورد چارک‌ها هم با استفاده از تابع `quantile()` چارک‌ها بدست می‌آید که البته نکته بود که صورت سؤال که باید چاپ شوند پس صرفاً همین جا نتایج را می‌نویسیم:

0%	25%	50%	75%	100%
-16.00	-6.25	-5.00	6.75	29.00
	↓	↓	↓	
	چارک اول	میان (چارک دوم)	چارک سوم	



بخش دوم

۱) دقیقاً همانند بخش قبلی با این تفاوت که یک argument جدید به تابع `read.csv()` اضافه شد به نام:

`na.strings = c(" ", NA)` که یعنی هر جا رشته خالی را دیدیم که بی یا `NULL` است به جای آن عبارت `<NA>` را ذخیره کنند

۲) ابتدا نام اسلایدها را به یک دستور ریخته و بعد با استفاده از `head()` در پایین آن argument ای که تعداد خطوطی که از ابتدای دستور می خواهم داشته باشیم است به دلتوی ۱۰ خط را انتخاب کردم و در `first-rows` ریخته و هر دو را چاپ کردم.

۳) ابتدا به دلتوی تعداد تکرار هر یک از فاکتورهای `workclass` را با استفاده از تابع `table()` بدست آورده و `repeated-input-num-workclass` ضمیمه کردم و آن را با `print()` چاپ کردم

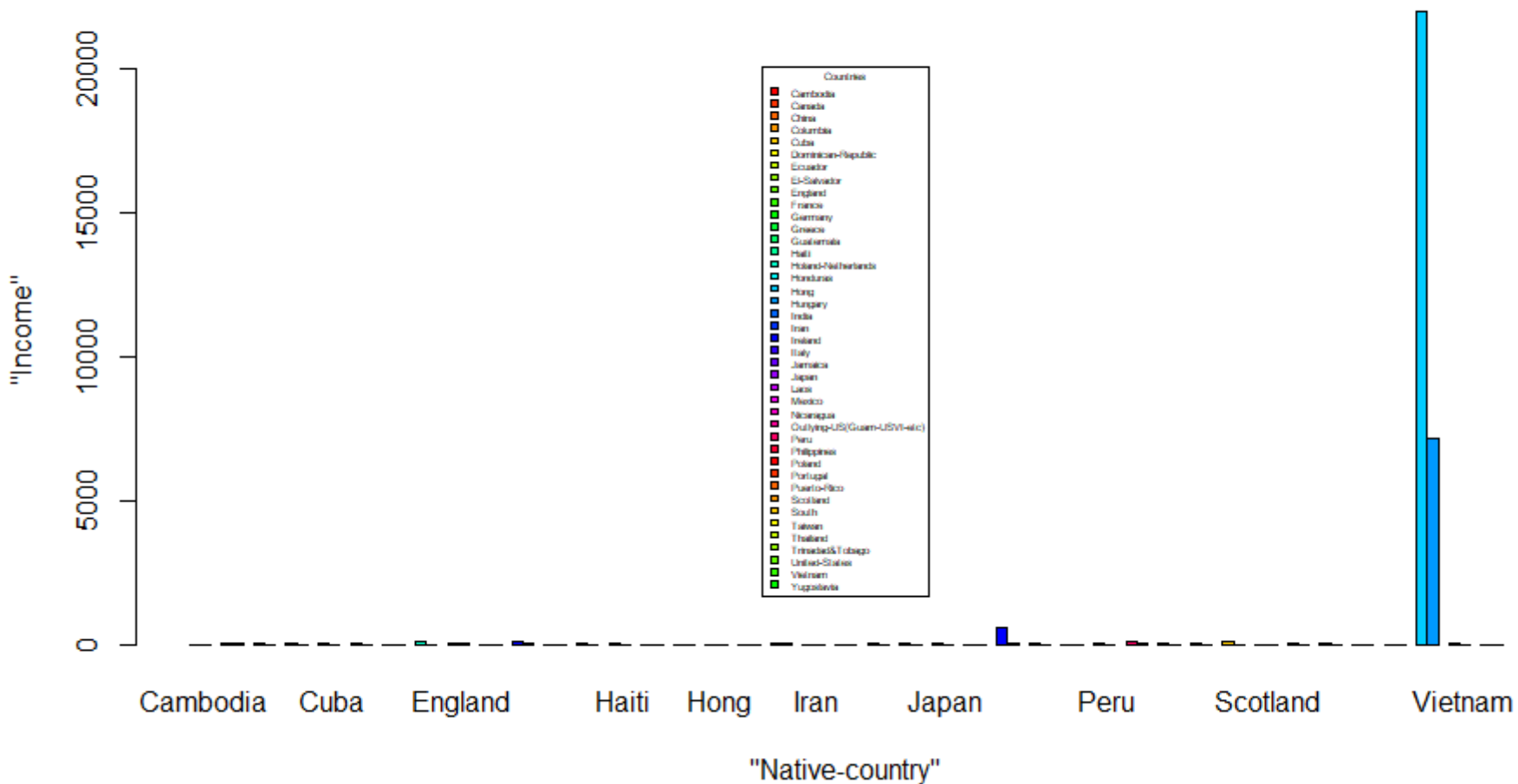
۴) برای قسمت بیکی دو جدول مقدار در نظر گرفته ام اولی که `repeated-country-and-income-1` است در آمد بر حسب کشورها و بیکی برعکس آن و هر دو را با استفاده از `table()` و پایین آن argument های که از اینترفیس برای مشخص کردن آن ها پیدا کردم کردم. (ص ۴ و ص ۵)

۵) نکته ای که به دست این است که کد های کوچک کردن نویسنده کشورها در محور x را به دستیم به و نام کشورها را چندتا در میان نوشته البته با `zoom` کردن و با بزرگ کردن های متفاوت تا جایی می توان تشخیص داد (در مقدار ص ۴)

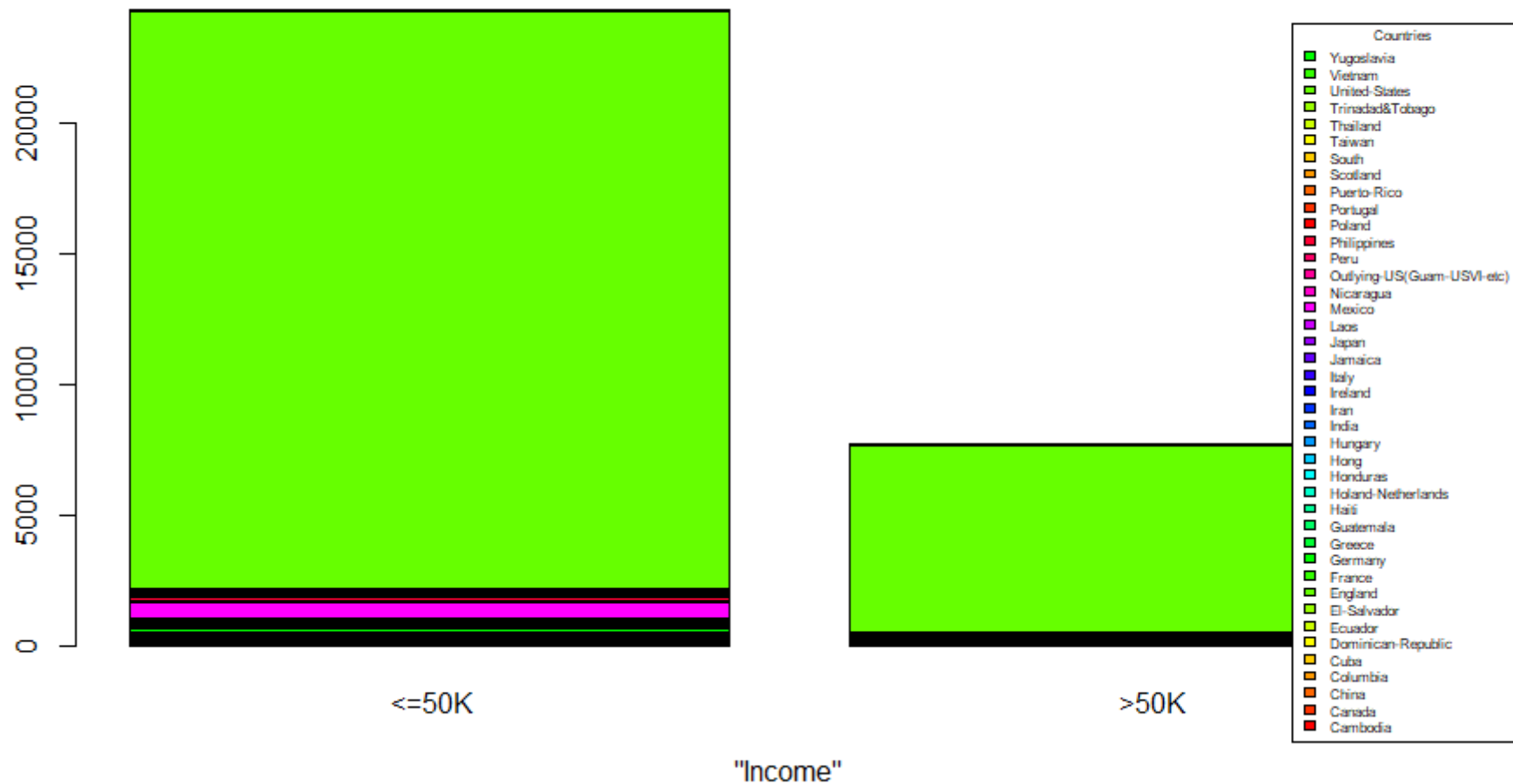
۶) برای حل این قسمت چون در دستور `education` ها زیاد بودند و اسم ها دلتوی و هم می شدیم تنها به یک `table()` آن دو را ریخته و بعد با استفاده از تابع `mosaicplot()` و یکی argument اضافه تر برای رنگ و فونت و ستریت سبکی در مشخص کردن مقدار داشته. (ص ۶)

۷) ابتدا با استفاده از تابع `age-frequency-table()` را ساخته و بعد با استفاده از `plot()` عکس آن را کشیدم البته با اضافه کردن رنگات (ص ۷)

"Grouped bar plot table 1"



"Grouped bar plot table 2"



"Income vs Education"

