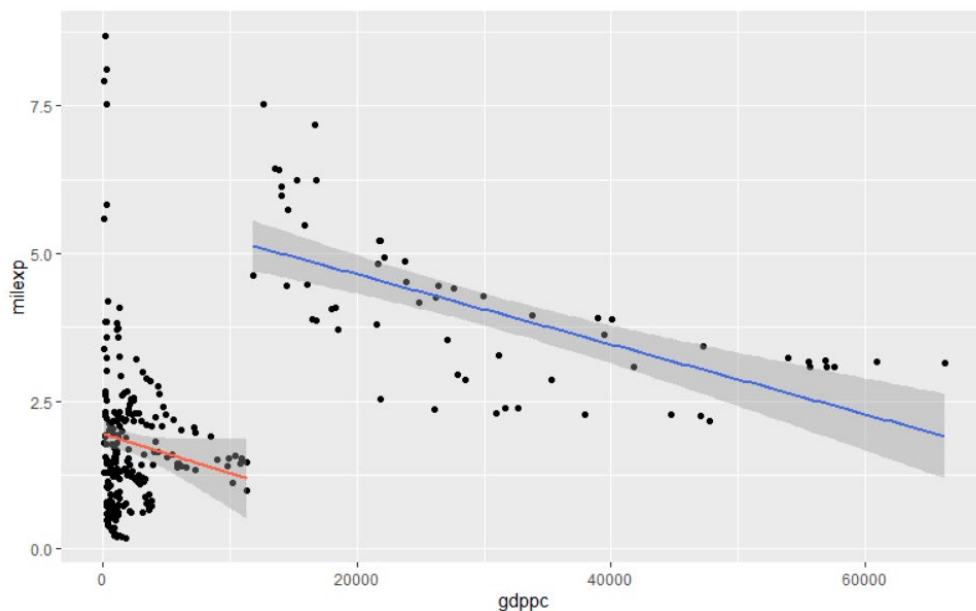


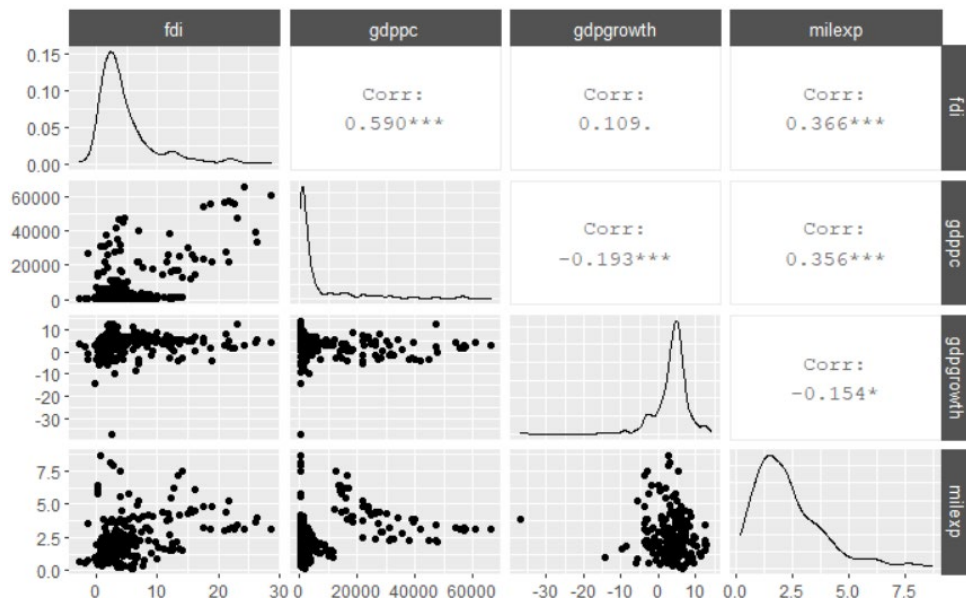
1. 第一题暂存数据维度用 `matrix` 就可以，不需要 `array`；
2. 在分类设置图层的颜色、线型等的时候用到的 `factor()`，老师意见是以后做以前可以先把数据集里面的数据类型转换好，就不需要在作图时候再套把数据转换成因子；
3. 第四题的小改动：东帝汶于 1999 年脱离印尼独立，2002 年正式建国、加入联合国，数据当中最早是有 2001 年的 GDP 增长率的。我感觉可以这样写：“东帝汶在亚洲金融危机期间仍然被印尼所占领，但金融危机对印尼的冲击间接促成了东帝汶在 1999 年通过全民公决独立。”
4. 第五题的 `gdppc` 和 `milexp` 之间的关系应该是有高收入国家和中低收入国家这个变量在影响的，为了便于分析部分的叙述，我觉得可以分组来回归：



参考了原数据，我把区分高收入与否的点设在 11500 美元，文莱和新加坡的 `gdppc` 都在这之上，而其余国家都不超过这个数。

关于分国家展示的小图，同样为了便于叙述负相关的结论，或许可以把回归的方法改成线性拟合。

5. 第七题如果需要提供显著性的话，感觉更适合用 `ggpairs`：



于是就可以发现 fdi 和 gdpgrowth 竟然相关性也不是很高，而且还不怎么显著……不过无论如何，从图上看，感觉 fdi 和 gdppc 的相关性以及 fdi 和 milexp 的相关性还可以，也许大家可以想一下分析里面要不要把它包括进去。

6. 第八题我觉得适宜把 polity score 处理成连续变量（而它反映出的政体类型是离散的），相应的用 histogram 会更加合适一些。

7. 第九题 tq1……不过小小在表述上较真一下，像印尼有一年的 Riot 数量翻了一倍（100 出头到 250），变化还是大的，只不过相对于 Protests 来说波动很小，也许可以说：各年间骚乱的数量变化起伏远远小于抗议数量的波动，主要突出是“相对比之下”。

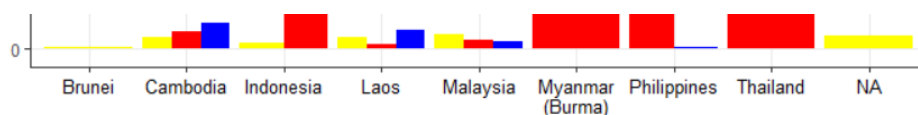
8. 第十题使横坐标不重叠，还可以让缅甸的名字分行显示，我这里是用个函数把空格替换成换行：

```

```{r q10, fig.cap='东南亚各国冲突数量（1946-2019）。'}
# 此题的部分代码如下
ggplot(ucdp_cnty, aes(x = countname, y = conflict, fill = type_of_conflict2)) +
  geom_bar(position="dodge", stat="identity") +
  scale_x_discrete(labels = function(x){sub("\\s", "\n", x)}) +
  scale_fill_manual(values=c("yellow", "red", "blue")) +
  labs(x = "", y = "Count", caption="Source: UCDP 2019") +
  theme_bw() +
  theme(axis.text.x = element_text(size = 10, color='black',angle = 0),
        legend.position = "bottom",
        legend.title = element_blank(),
        strip.background = element_rect(fill = "#525252", color='#525252'))

```

这样会让 labels 变成



或许更好。