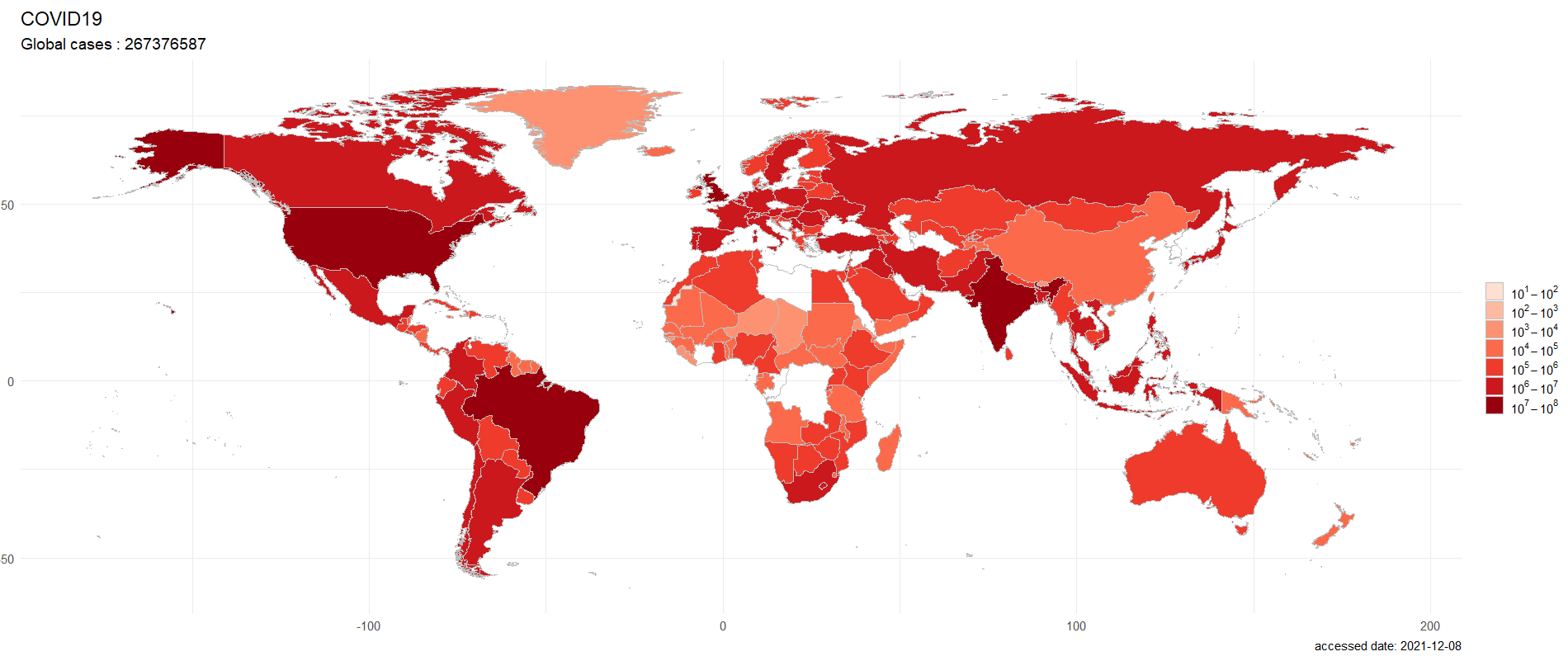
大家好，今天我们小组报告的主题是关于新冠疫情数据的一些思考。从疫情伊始至今过去了快两年了，可能两年前大家还在期待着疫情什么时候结束，两年后的现在已经有点习惯了吧。当然了，还是希望疫情能早日结束，让大家进出校园不用再扫健康码了。

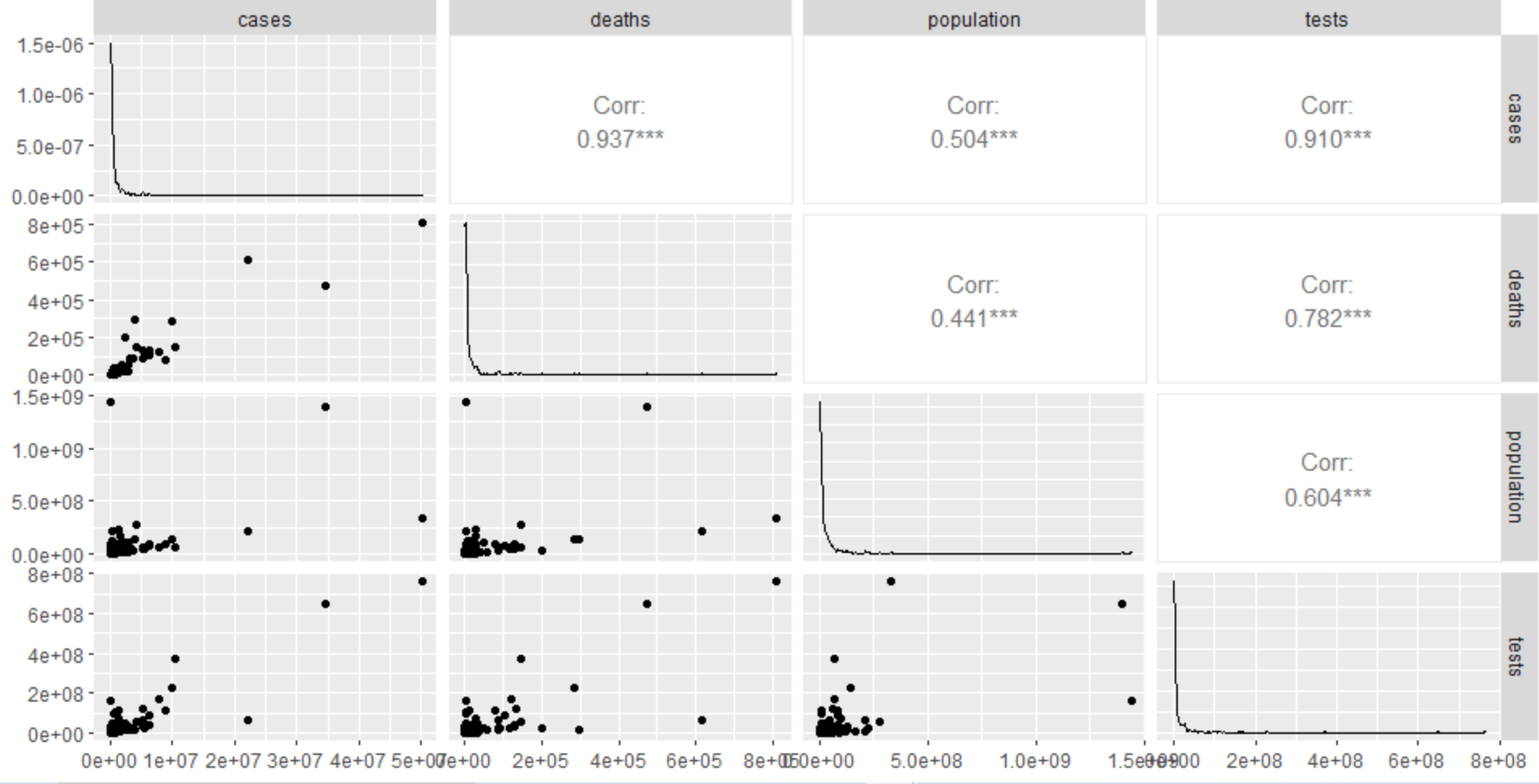
我们今天主要从两部分展开：一个是分析一下最新疫情数据，看一下目前最新的疫情形式；再然后是回顾一下过往的疫情数据，结合几个经典的病毒传播模型，配合一些典型国家的疫情发展趋势，来分析一下影响疫情传播的几个因素。

首先讲一下数据来源。疫情的数据还是比较好找的，这里推荐一下南方医科大学的余光创老师提供的nCov2019这个R包，可以很方便地爬取最新的疫情数据。不过余老师这个包缺失了自今年8月之后的部分数据，还有一些用到的数据可以从OurWorldinData里面获取。



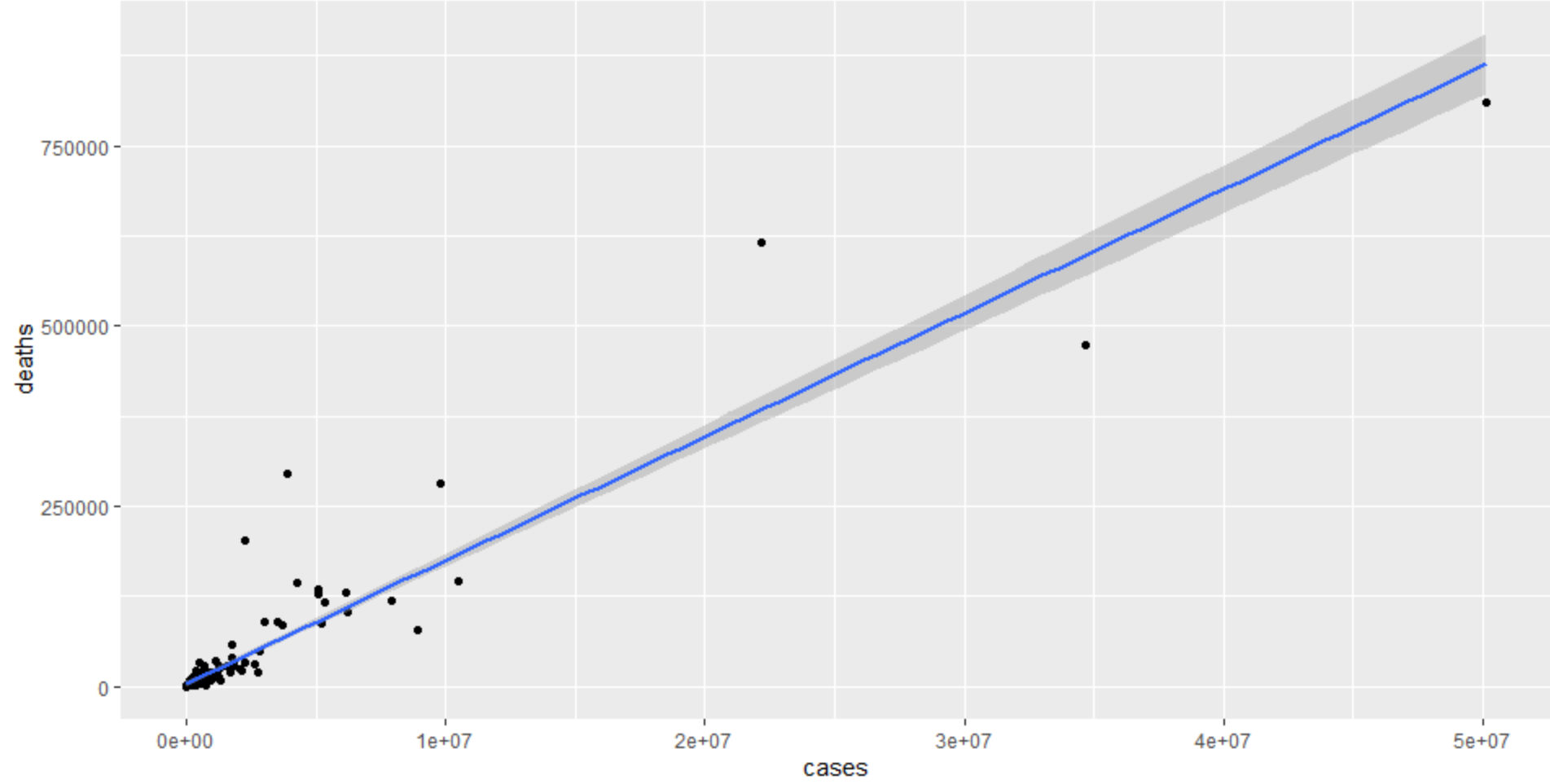
这是截止到昨天世界上依然确诊的病例数。国内最近内蒙古那边疫情比较严重，；全世界看，美国和非洲部分国家疫情依然很严重。

进一步，我们从数据集中找到了累计确诊、累计死亡、累计检测数量与人口的数据进行分析。这里为了便于分析，我们会选取一些比较具有代表性的国家。

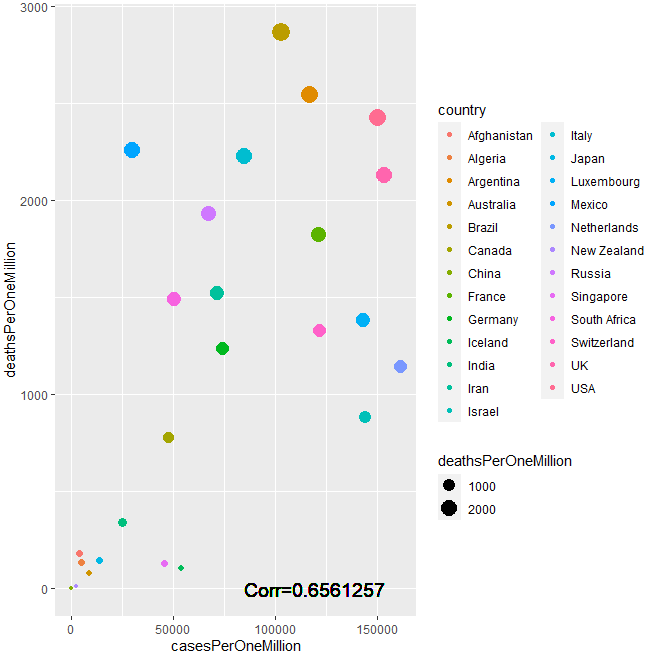


这张图是对各国的累计病例、累计死亡、累计检测数量与人口进行相关性分析。可以看到累计确诊cases和累计死亡、累计检测的相关性极高，在0.9以上。这也比较符合预期。

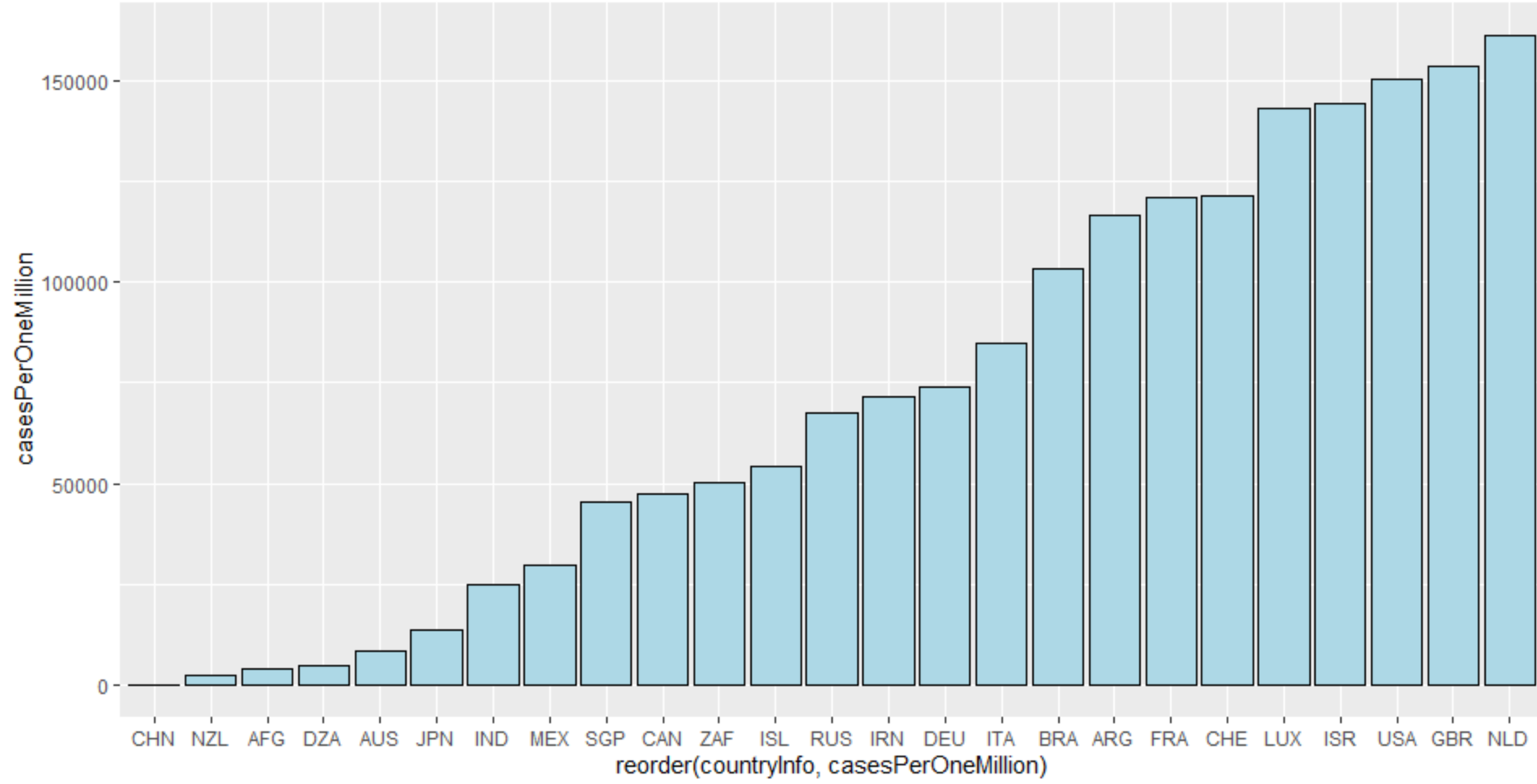
由于较强的相关性，我们尝试对累计病例和累计死亡做线性拟合。



我们不难理解他们之间的正相关性。但是随着目前世界上疫苗的广泛接种，累计死亡和确诊不应该再接近线性相关。考虑到各个国家的人口数量越大的话，累计死亡和累计确诊都有可能增加，即累计确诊和累计死亡的正相关性有一部分是通过人口数量population表现出来的，population在这里担当了confounder混淆因子的作用。我们再考虑去掉各国人口数量的影响，绘制了每百万人中的确诊数量与死亡数量的散点图。



每百万人的确诊数与死亡数总体上还是呈现正相关的，但相关系数并没有那么高，为0.66，实际上并没有很强烈的线性相关性。

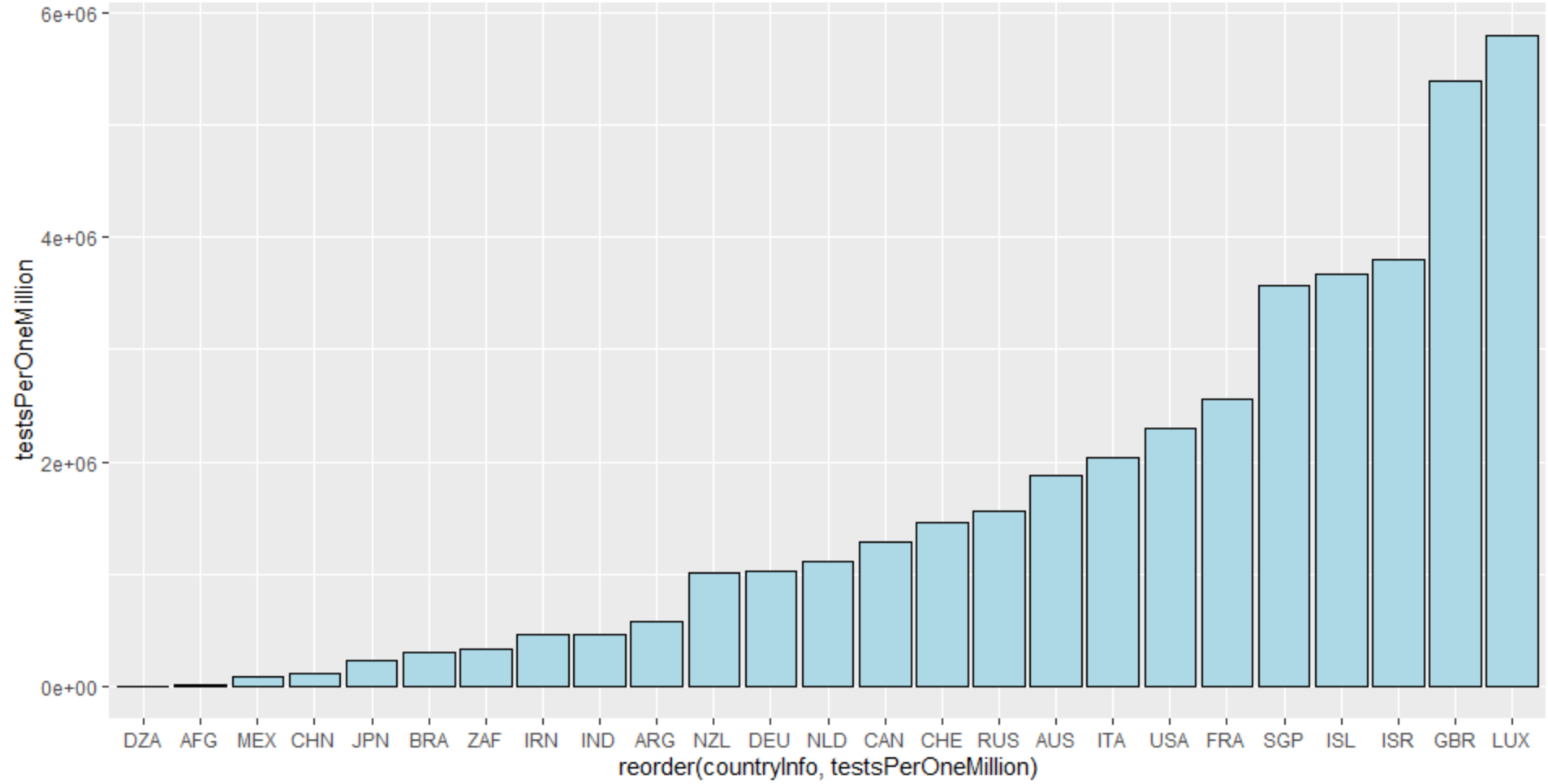


casesPerOneMillion最大的国家：荷兰，英国，美国，以色列，卢森堡

casesPerOneMillion较大的国家：瑞士，法国，阿根廷，巴西

casesPerOneMillion最低的国家：中国，新西兰，阿富汗，阿尔及利亚

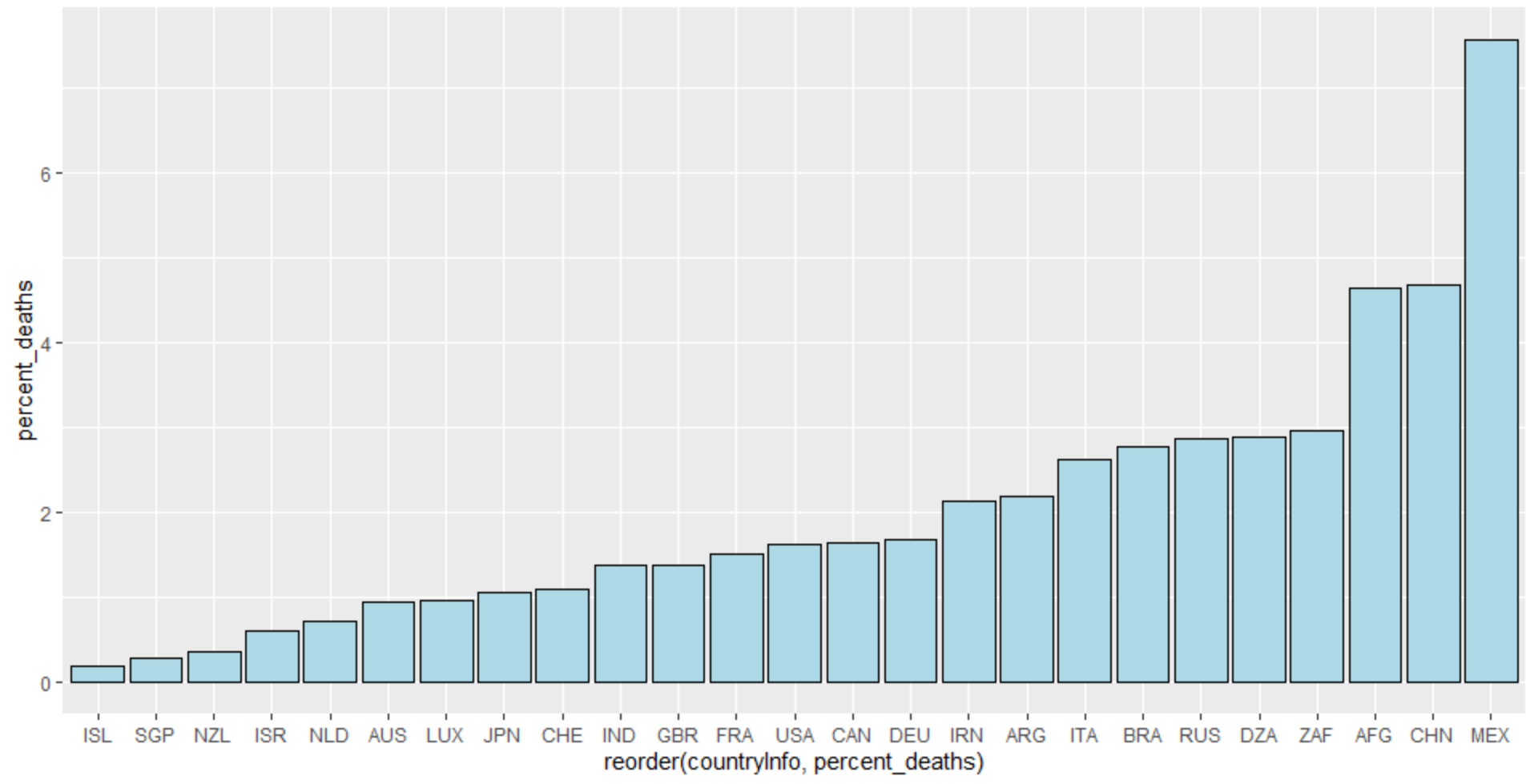
这张图是各国每百万人中累计确诊是数据。一些人口并不多的国家如荷兰、以色列、卢森堡，在去除人口的影响之后，发现他们每百万人口中确诊病例的个数十分多；而中国人口大国，但政府防控干预做的好，实际上每百万人中的确诊病例是在最低的国家中的；当然也还有其他的影响因素，阿富汗、阿尔及利亚的检测量极低，是每百万人中检测数量最低的两个国家(分别5132和19856)。



而在检测量上，可以看出各个国家检测数量差距很大。

我们计算得到变异系数为1677329，非常的大。正如之前所言，阿富汗、阿尔及利亚是最低的；卢森堡，德国检测数量最多。总的而言，经济发达的国家检测量更大，而贫困的国家人均检测量非常小。当然也和疫情严重程度有关：我国疫情控制的比较好，所以做检测的数量相对不高，而德国相对地疫情严重，检测数量也就会更多。

接下来我们再看一下各个国家的死亡率



死亡率最高的国家：墨西哥，中国，阿富汗

死亡率较高的国家：南非，阿尔及利亚，俄罗斯，巴西，意大利，阿根廷，伊朗

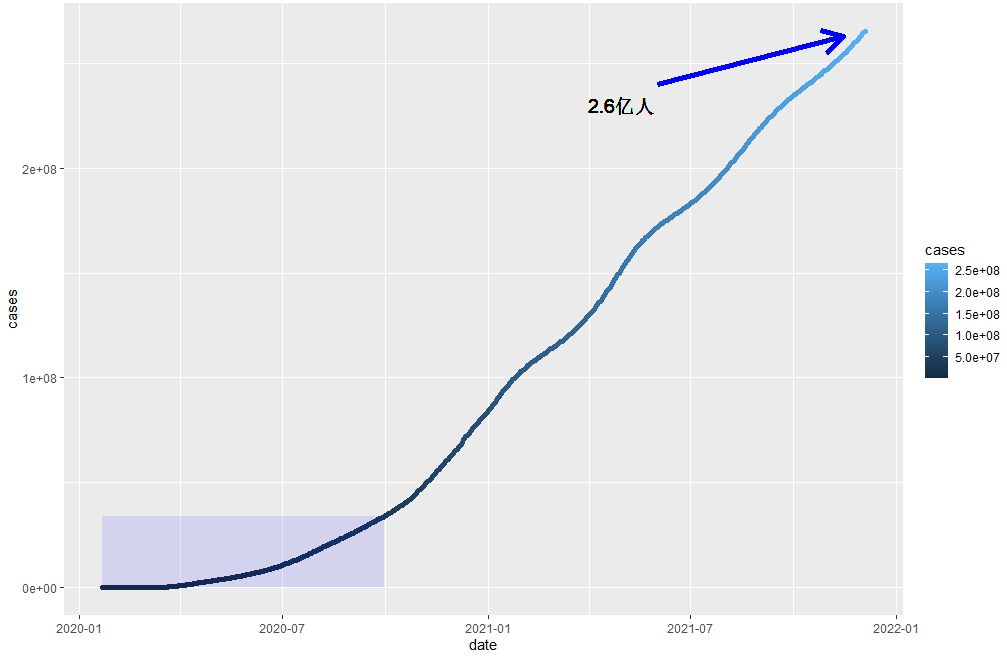
死亡率最低的国家：冰岛，新加坡，新西兰，以色列，荷兰

死亡率较低的国家：澳大利亚，卢森堡，日本，瑞士

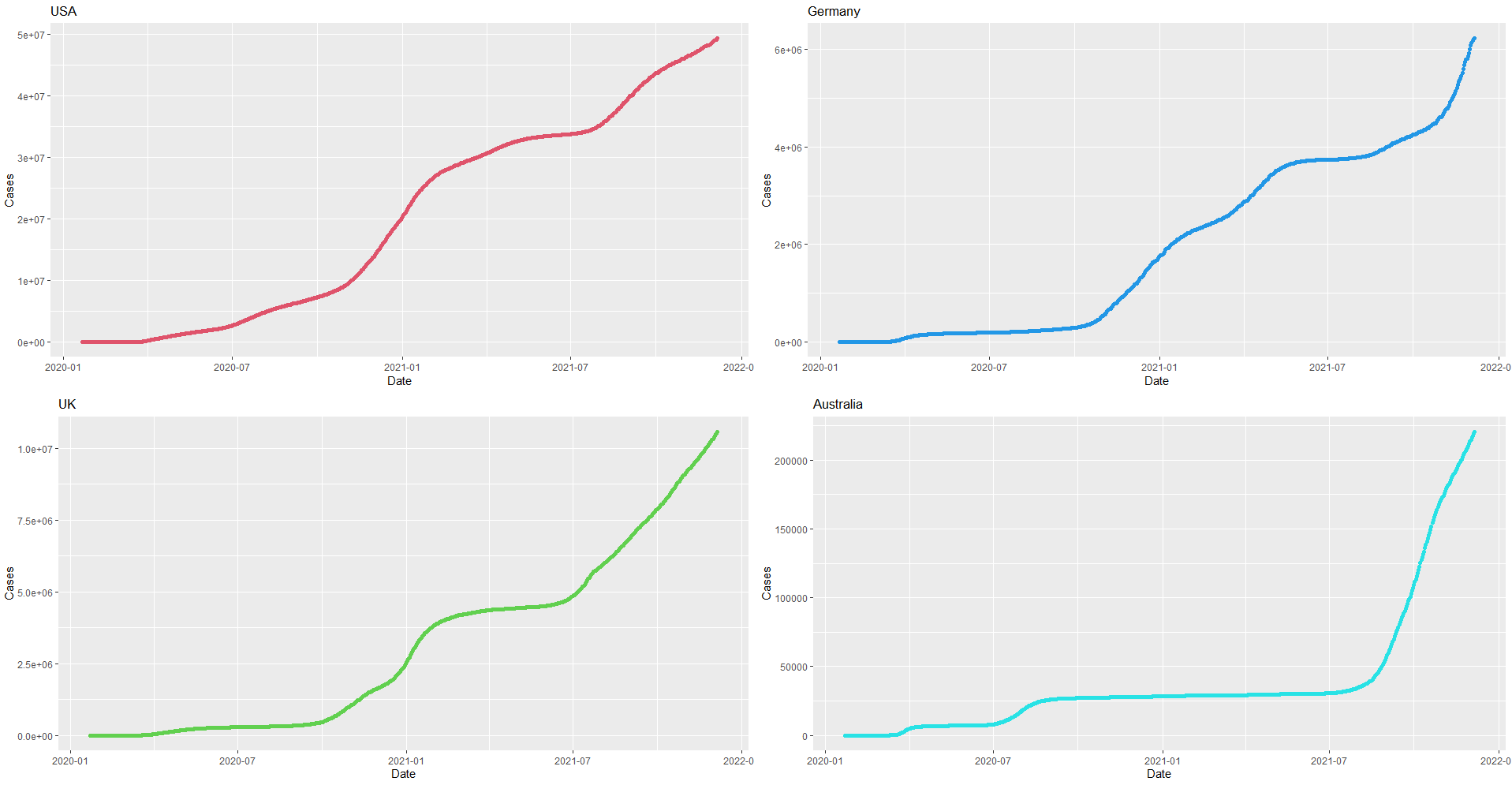
我国的话，去年在疫情初期，湖北的疫情确实比较严重，而往后疫情控制的比较好，分母较小使得总体的死亡率看上去比较高；而墨西哥、阿富汗确实整体疫情严重，管控不当，难以遏制死亡率的上升。死亡率较低的国家中，冰岛、新加坡、荷兰、日本、瑞士、澳大利亚都是医疗水平比较发达的国家；而以色列在今年3月就有半数以上的人口完全接种了疫苗，可以发现疫苗对于病毒还是有十分迷宫你先的效果的。

纵向数据：

自去年年初新冠疫情爆发以来，便在全世界蔓延开来。截止到今年11月底，世界总的确诊病例已经来到了2.6E人次。在2020年初期，受感染人数的上升还是比较缓和的；但目前，从图中的趋势来开，似乎还没有想要降下来的意思。



我们再看一下几个典型国家（美国、德国、英国、澳大利亚）的疫情走势。



我们不难发现，以上图像都呈现了先凸后凹再凸的趋势。20年底和21年8月时斜率激增。

20年底：德尔塔（Delta）

21年初~21年7月：增长率显著放缓

21年8月：奥密克戎（Omicron）

Delta和Omicron的出现让增长率显著提升，这个结果并不让人意外。

每一波变异病毒的来袭都让疫情走势有了新的变化，这里我们引入病毒传播模型来SIR来更好地理解这一现象。

在SIR模型中，我们假设总人数N不变，将所有人分为了健康者(Susceptible)、已感染者(Infected)和病愈免疫的移出者(Removed)。所谓移出者，是指获得了抗体具有免疫能力的人。实际上，大多数传染病如天花、流感、肝炎等治愈后均具有很强的免疫力，或者人们可以通过疫苗来获得了免疫力，受感染者最终应该会下降。这些人既不是健康者(Susceptible)、已感染者(Infective)，这里就称他们为移出者(Removed)。我们记某一个时刻t三类人在总人数N中的占比分别记作s(t), i(t), r(t)。每个病人每天有效接触的平均人数为λ，日治愈率为常数μ（每天被治愈的病人数占病人总数的比例），那么传染期可以理解为μ的倒数μ分之一。于市可以令σ = λ / μ表示传染期接触数。

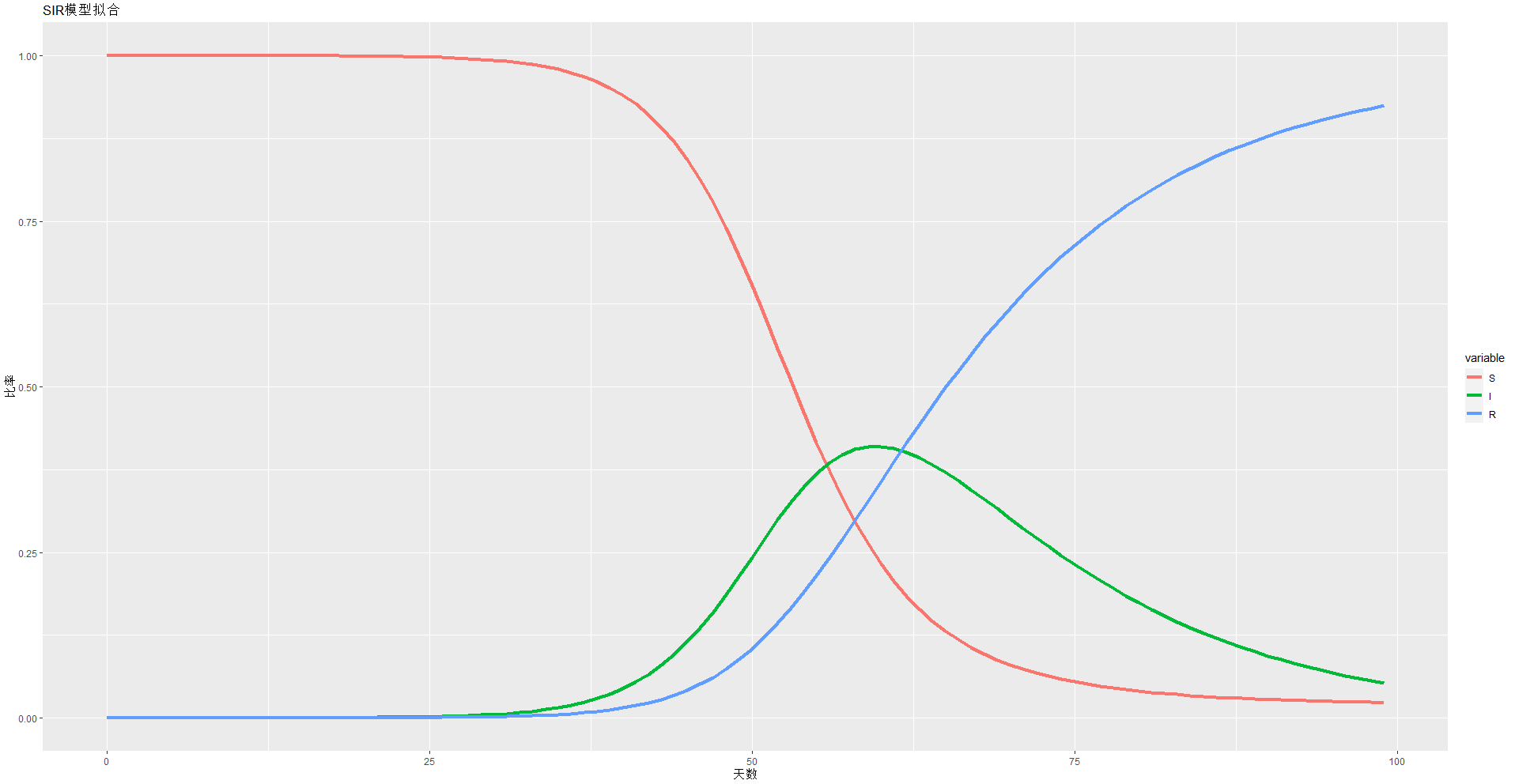
和之前简单的模型一样，我们根据假设建立两个微分方程。

di/dt = λsi – μi, i(0) = i0

ds/dt = -λsi, s(0) = s0

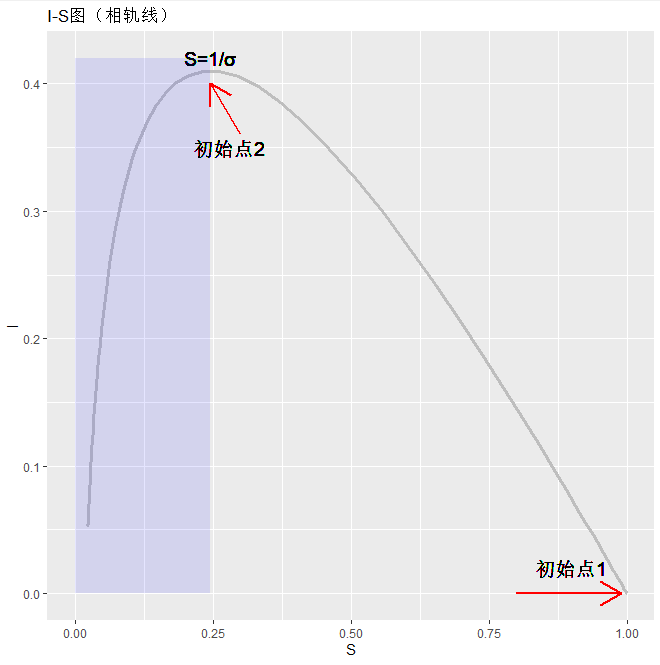
第一个式子表示受感染率的增长等于每个病人有效接触到的未感染人群S减去治愈人数R；第二个式子表示的是未感染人群的占比会因为每个受感染者每天的有效接触而降低。这个微分方程没有显式解，这里我们将湖北在疫情初期的数据带入模型进行拟合。

数据选择：由于自1月23日10时起湖北省关闭了省内的大部分出城通道，因此可将此阶段视为一个暂时封闭的整体；此外，自5月份后疫情没有大的变化，因此选择2020.1.22-2020.04.30的数据作为需要处理的数据。



这张图是拟合后的S I R三条曲线。随着时间的增长，易感人数比S逐渐从1降低至0；受感染人数比先增长，到达峰值后逐渐回落；而移出人数比R从0缓慢增长，最终会增长至1。这里我们带入的是疫情初期的数据，根据模型最终得到在50-60天之间，受感染人群的比例来到了高峰，达到了近40%。这显然和实际数据是不吻合的，因为SIR拟合的是不加干预的情况下疫情自然传播的情况。实际情况下，湖北省在疫情初期严格把控，新冠病毒感染人群增加的主要原为隔离人群在隔离期间确诊；且疫情防控期间，参数λ和μ都是变化的，而在模型中是一个定值，因此与实际情况会出现偏差。虽然SIR模型非常直观地展示了传染病的过程，但是该模型对人群的分类不够细致，未能清晰考虑出隔离这一重要因素，同时SIR对于模型的初值也十分敏感，使得模型的稳定性不佳。

**比起模型拟合的准确性，这里我们更加关心模型带来的意义，即理解曲线变化的原因。**



我们可以看一下受感染者I与其他易受感染者S之间的变化关系。首先，无论我们在从哪个初始点出发，最终都会到右下角的点，也就是疫情终将会结束。一般来说，在疫情初期，我们从右下角的初始点1出发，感染率I先增后减，在s=1/σ处到达峰值。在σ = λ/μ中，卫生水平越高，日接触率λ越小；医疗水平越高，日治愈率μ越大，于是σ越小，由此提高卫生水平和医疗水平有助于控制传染病的蔓延。

还有一种角度，如果我们把初始的S0就定在途中的峰值，即s=1/σ处，那么传染病也不会再蔓延。σs表示一个病人在传染期内传染的健康者的平均数，称为交换数，σs<=1，则疫情也不会再蔓延。这也就是群体免疫的理论依据。

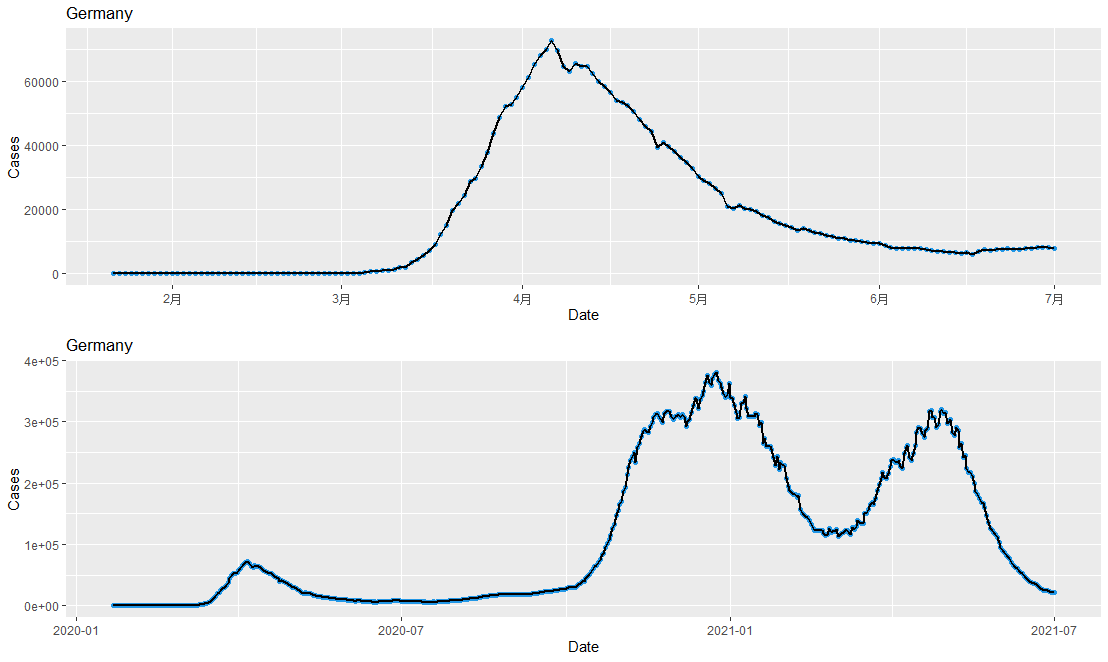
由于病毒只要接触有免疫力的人就会被消灭。病毒的大规模爆发就会被抑制。病毒即使传播到了没有免疫力的人身上也会因为大部分人有免疫力而无法传播更大范围，从而抑制病毒的传播，直到病人被治愈或死亡，病毒也被消灭。

英国、德国在疫情初期也尝试了群体免疫的做法，问题在于：

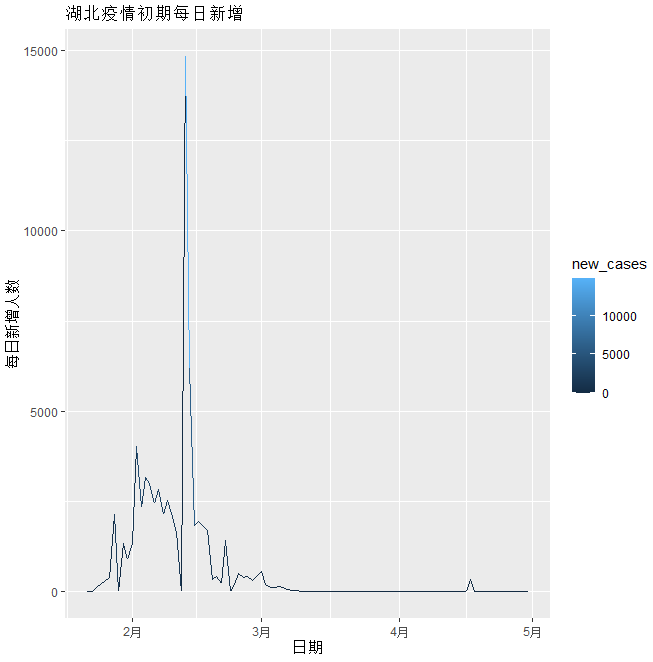
1）疫情初期不了解COVID-19的传播方式、感染率、变异率以及人体免疫系统能起到的作用；

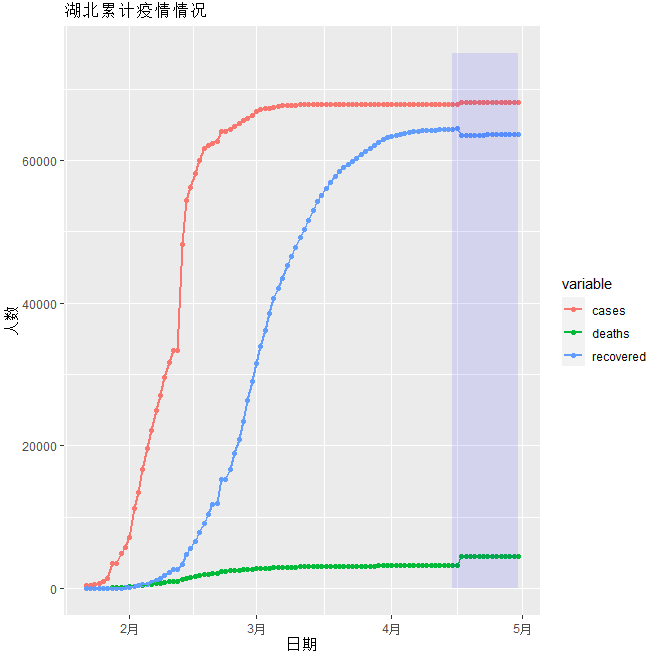
2）自然的群体免疫涌现大量患者，医疗系统难以维系；

3）无法证明治愈后终身免疫。



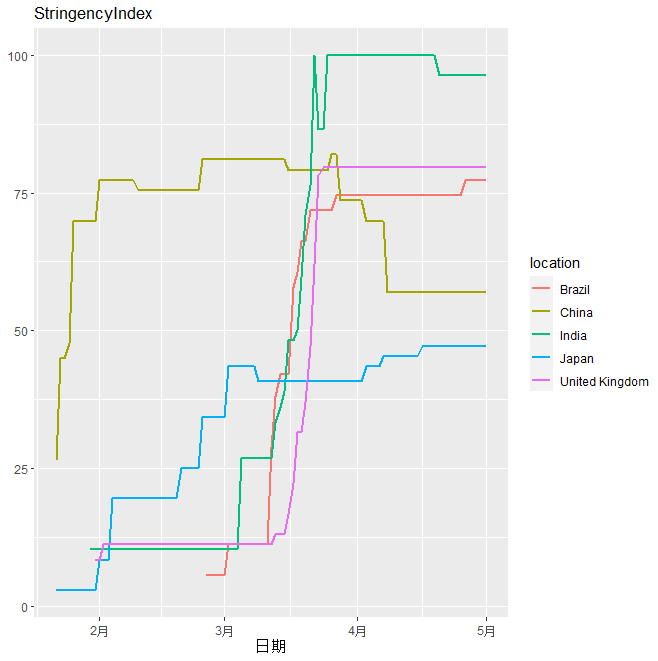
我们看一下德国的受感染人数图，这里不是累计感染人数而是各个时段的感染人数。可以发现去年到7月德国的疫情确实平稳了下来，但是随着变异毒株的袭来，德国接连经历了高峰——低谷——高峰——低谷的过程，如果要建模的话，就需要对这三个时间段分别用SIR模型尝试拟合，这正是病毒变异对疫苗及群体免疫带来的冲击。



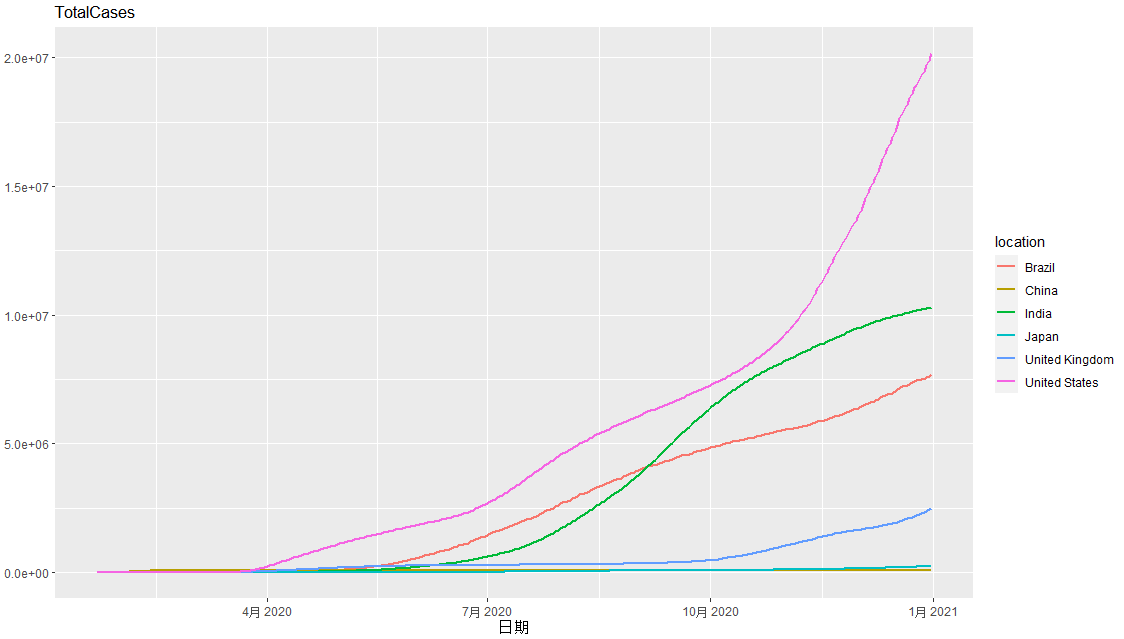


与之相对，通过口罩、隔离等提高卫生水平，降低日接触率λ；提升医疗水平，提高日治愈率μ，可以尽早引来感染者的高峰期，从而控制住传染病的蔓延。湖北在疫情初期便采取了封城、居家隔离等政策措施，有效限制了每个潜在的病人的有效接触人数。我们从湖北省去年上半年的疫情情况也可以看出来。

左边一个突然的高峰是去年2月12日是将临床诊断病例数纳入确诊病例统计，是实事求是、负责任的做法，既能加快患者收治，又能加大疫情防控力度，用更短时间切断传染链条。在2020年4月16日《关于武汉市新冠肺炎确诊病例数、确诊病例死亡数订正情况的通报》，对确证病例数、确证病例死亡数进行了订正，实际疫情已经得到了控制。

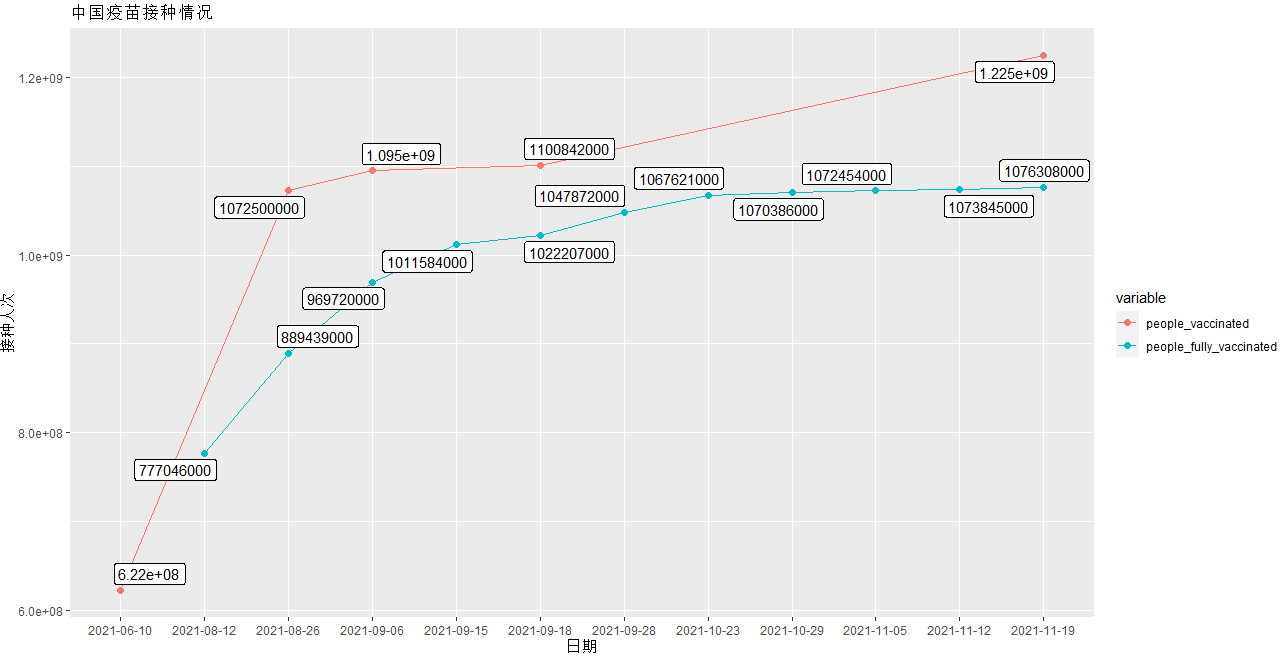


国家间不同的防疫政策最终会带来不同的结果。牛津大学给出了一个Stringency指标，结合了9项不同的数据来刻画不同国家对疫情管控的严格程度。这里选取了去年2月到5月部分国家的stringencyindex，可以看到中国在疫情之初就通过封城、限制出入境等政策进行了严格的防疫；日本也是比较早就采取了行动，但是比较温和；而美国、英国、巴西、印度等国家则行动迟缓，直到3月份乃至3月中旬才开始严加管控。政府管控的早晚与管控力度也对疫情带来了不同的影响。这张图可以很明晰地看到，国家防疫政策落实的快慢与严格程度对应了国家确诊病例的涨幅趋势。去年特朗普政府还执政的时候，声称很懂疫情，导致美国的确诊病例至今依然是世界第一；巴西总统博索纳罗，去年懂王的头号粉丝，带着巴西的疫情持续走高；大嘤帝国全体免疫失败后，疫情增长也不容小觑；阿三本身人口密度大，经济相对落后、医疗体系不健全的情况下，防疫政策慢人一步，也无法避免确诊病例的持续升高。考虑到政治、经济发展、社会制度等因素，确实不是所有国家都能像我国一样落实严格的防疫制度，但仍然希望各国领导能做出最好的决策。



小结：

目前，疫苗现在已经在许多国家开始广泛接种了，我国在今年9月15日，完成全程接种剂次的人数超过了10e人。尽管病毒在不断变异，疫苗仍然能够起到架墙抵抗力的作用，大家还需保持谨慎但不必悲观的态度，做好疫情防控。



如何能够恢复常态化？这里我们就不敢发表论断了。按照钟南山老先生的观点，在将病毒病死率控制在0.1%，复制指数（可以理解为之前模型中的σ）控制在1.0—1.5，传播系数降下来的前提下，可以完全开放。但要实现前提，需要全民接种疫苗、建立群体免疫，将社区群防群控常态化，并研制有效的药物，最后还是希望这一天能够早日到来。