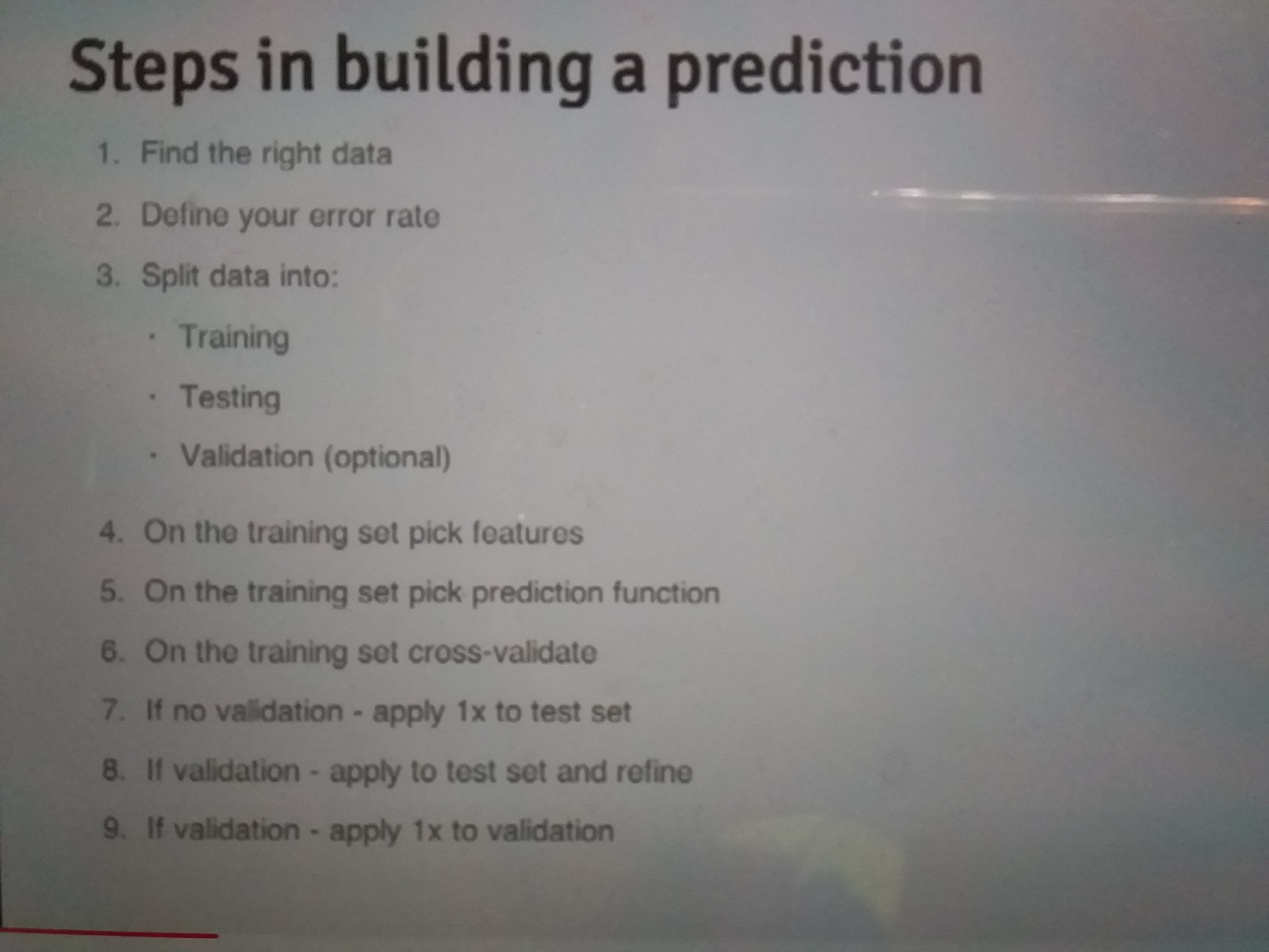
很多人以为做数据分析，预测模型是很高深的一件事，要从最基本的统计方程式开始，从温以前高中/大一的高数教材。不用怕！现在有了互联网，个人电脑，大量免费下载的程序，其实每个人，只要你乐于学习，都可以做数据分析，做预测。

例：大数据的目的是利用大量的数据分析，帮助我们可预测未来，通常的步骤如下：　，



要收集到合适并足够的数据，部分用来建模，部分用来验证／测试　已经不容易。

现在网上有了很多资源，提供可用来分析的数据，

甚至，有些网站是专门做预测模型的竞赛，有些公司希望有志愿者帮自己收集数据做分析，便可以把数据上传到这竞赛网站，提供奖金，公开让其大家来竞赛。

大家要做数据分析时，就可以不用等，直接从网站下载这些真实数据。好处是，可以与其他做数据分析的人进行比较，知道自己的分析水平。但是有了这些网站，就可以把自己的预测模型上传到网站上，它会帮您排名，立马知自己做出来的预测模型有多好，如果最终可以打败全世界其他竞赛者，还有可能获奖！

数据分析工具箱

起步阶段，可用现有的东西，再下载一些免费程序开始。

例，有些资深的数据分析师，列出一些要做分析的工具，

图

例如：excel就是很好的工具，单是靠它不够，再配合 R 。

如大家有兴趣，可以观看我们一系列视频，介绍如何利用Ｒ与　excel

用于分析铁达尼号数据

不要局限某一个工具／语言，它们各有长短，如要做关于语句或者文字的编辑调整，可以用VIM或者PERL等更直接。

如大家按视频做分析／建模，会更理解我常在一些企业看到的数据分析通病，立马去利用工具建立预测模型、　很多大师的分享，他们都要花70　～　80%的时间来展示，观察，　分析数据，建模是最后一步，因如果没有找到合适的关键因素（例如可从散点图、柱状图分析来看），最后的预测模型无效或很差。所以前面的工作要做好。

请不要误会，要成为一位优秀数据分析师还是很长时间的努力，但现代科技使入门的门槛降低了很多。

有了大的群体就比以前容易多了，而且很多要做数据分析不像以前要从最基本的统计方程式开始，有很多分析工具在网上都可以找到，可以去用，以数据分析为例，很多人觉得无法做出数据分析，就