2017(10-12月) 分析|数据科学类 硕士 全职@Facebook - 内推 - 技术电面 Onsite | Pass/Offer | 在职跳槽
小K数据科学培训课程里详细讲解的24道面试题目

题目来源：1point3acres.com 数科面经版 + personal collection from @小K

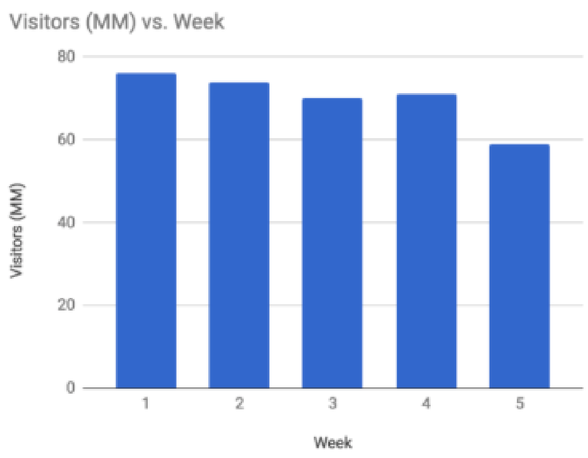
Note：

- 这24道题目是《Unit 1/DS501：Analytics & Product Sense》的面试实战例子，题目覆盖了analytics面试考察的方方面面。
- 其他Unit的面试题目，会随着课程的进行而逐渐推出。
- 小K会根据数科面经版的新题目以及工业界面试动态，对精选题目随时进行增改。
- 无论你是否打算参加小K的数科课程培训，都可以以精选题目作为面试准备材料。如果这些题目你能熟练应对，那analytics面试对你来说应该不难。

====== Analytics面试24题精选 ======

小K数据科学培训课程里详细讲解的24道面试题目

1. Like video: 添加给大家看朋友like的video， 如何评估这个feature 值不值得加
2. Close friend notification: 打算给用户的发text notification 告诉他们close friends 的Update, 如何评估这个feature 值不值得加, 注意是close
3. Prime Now, 就是prime的升级版，两小时送货上门，假设你提出来这个新的prime now的proposal，一周后给老板和其他部门老大分析这个事情可不可行，说说你需要哪些variable来进行分析，你怎么跟老板present
4. 1Point3Acres Smart Admit: 一亩三分地有很多录取数据：Build a data product “Smart Admit”: given an applicant’s background, predict chance of admission into a program
5. How do you explain this graph? (week 5 is this week)



6. Dashboard for Careers: Careers team’s mission is to help job seekers find their jobs (search, recommendation). For a weekly report of the team, what metrics are you going to present?
7. Feed diversity: Feed are overly optimized that users start to complain about feed having low diversity. How do you define diversity? (need to get a formula here) How do you measure and monitor to see whether diversity increased/decreased over time? How to decide on whether this is an issue to worry about?
8. Newsletter: The business sends two newsletters a week: one on Mondays for sales related contents, one on Wednesdays for explorational newsletters. Create a dashboard to monitor newsletter performance over time
9. Revenue goes flat yoy, what data would you use to explain why
10. AB test, slow roll out: How do you explain the result

Table 1: Conversion Rate for two days.
Each day has 1M customers, and the Treatment (T) is better than Control (C) on each day, yet worse overall

	Friday C/T split: 99%/1%	Saturday C/T split: 50%/50%	Total
C	$\frac{20,000}{990,000} = 2.02\%$	$\frac{5,000}{500,000} = 1.00\%$	$\frac{25,000}{1,490,000} = 1.68\%$
T	$\frac{230}{10,000} = 2.30\%$	$\frac{6,000}{500,000} = 1.20\%$	$\frac{6,230}{510,000} = 1.20\%$

11. Likes/DAU grows +10% year/ year. Explain why
12. Choosing the right analysis method:
    1. Know past shopping behavior, predict whether or not user will buy a new item
    2. See an image, tell whether it’s a cat/dog/human/etc
    3. Identify Platinum/Gold/Silver users
    4. Two new products, which earns more revenue?
    5. Given 10 slot machines with different payouts, get the best ROI for your $$$?
    6. Know past shopping behavior, predict price user is willing to pay for a new item
    7. We currently provide product X in country Y1 and earn $Z; what if we provide X in countries Y2, Y3, Y4?
    8. Is conversion rate consistent across Android/iOS/Web only users?
13. Interpret AB testing result, Treatment effect for each group as below:
    1. Trt1: -5, CI: (-7.5, -2.5)
    2. Trt2: -15, CI: (-17, -13)
    3. Trt3: -12, CI: (-28, -4)

    How to interpret C.I.? Which treatment to choose? Increase test power / accuracy?
14. Map: Where did I park? User开车出行,然后下车后可能还需要再走一段距离. 希望Map能告诉user自己的车的位置. 怎么实现?
15. Friend Recommendation: How do you test a new friend recommendation algorithm?
16. Given data
    1. User1 user2 #msg
    2. 001 002 2
    3. 001 003 4
    4. ...
    5. 003 001 4

    What is the distribution of #msgs? Who is the top exchanger for each user?
17. Roll out experiment, What’s the expected waiting time to receive treatment? 1K users
    1. Day 1: 10 users out of 1000 -> trt
    2. Day 2: 10 users out of 990 -> trt
    3. Day 3: 10 users out of 980 -> trt
    4. …
    5. Day 100: 10 users -> trt
18. Roll out experiment: What’s the expected waiting time for the user to first enroll in the treatment? 1K users
    1. Day 1: 10 users out of 1000 -> trt
    2. Day 2: 10 users out of 1000 -> trt
    3. Day 3: 10 users out of 1000 -> trt
    4. …
    5. Day 1000: 10 users -> trt
19. Optimal Dice Roll, you have a 6-sided fair dice, your reward is value of very last roll. You have a max of 3 rolls. What is optimal stopping strategy and expected reward amount.
20. X ~ N(0, sigma^2), Y ~ N(0, sigma^2), X & Y independent. P(X > 2Y)=? P(X > |2Y|)=?
21. Coding: find sqrt(x)
22. Coding: 某个小动物可以一次跳上一个台阶或者两个台阶，问爬上第N级台阶可以有几种方法，写个简单的code出来。
23. Coding: Random number generation: Give you a dice with 7 sides, how do you generate random numbers between 1 and 10 with equal probability?
24. Coding: Random sampling on streaming data: A stream of values x1,x2,x3 …; may or may not end. How do you sample k random items from the stream (without replacement), so that after seeing n data points, the probability that any individual data point is in the sample is k/n, with only one pass through the stream, and uses storage proportional to k (not the total, n, possibly infinite, size of the stream).

【最新更新】

小K之前的《数据科学面试40+真题讲解》课程视频上线，命名为《DS501： Analytics》

https://learn.1point3acres.com/courses/ds501-analytics/

【历史存档】-- inactive

数据科学培训课程介绍和报名页面：《数据科学面试40+真题讲解》
http://www.1point3acres.com/ds-courses/1-analytics/