

1.1. Frontend person wants to change color of the 'Go' button on a search bar. Will it increase conversion rate?

Answer: True

1.2. The data team created four versions of machine learning model for product recommendations to new users of an app. Which one is the best?

Answer: True เพราะ เราสามารถแบ่งกลุ่มได้ง่ายและเปรียบเทียบจาก user ซื้อ product ที่ recommend บ่อยแค่ไหน

1.3. Two managers from different factions have Layout A and Layout B for a physical convenience store. Which one should we use?

Answer: True เพราะใช้การหาร้านค้าที่มีอะไรหลายอย่างเหมือนกัน แล้วนำมาแบ่งเป็นกลุ่ม ที่ใช้ layout A และ B แล้วเปรียบเทียบกัน

1.4. Mr. Rabbito thinks offline stores are the best channel to distribute our products, whereas Ms. Rakko thinks online websites are the way to go. Who is right?

Answer: False เพราะการเปรียบเทียบยอดขายระหว่างสองแหล่งมันมี sample ที่มี properties ต่างกันทำให้เปรียบเทียบกันไม่ได้

1.5. Your boss wants to add a premium version to your freemium service. Is it a good idea?

Answer: False เพราะเดิมเป็น service free ถ้าลูกค้าต้องใช้งาน premium service อาจจะได้รับเสียงต่อต้านจากลูกค้าได้ อีกทั้งการทดสอบนี้อาจจะได้จำนวนเงินที่มากขึ้น แต่ในระยะยาวความน่าเชื่อถือของการทดลองจะน้อยลงเพราะลูกค้าน้อยลง

1.6. The backend team came up with a new setup that they think will speed up the website load time. Should we implement this change?

Answer: True

1.7. Kuruma Inc., a car dealer, wants to change the banner on their homepage to see if it will attract more repeated customers. Average time between purchase of the car company is 5 years. How do you know if the banner change has an effect?

Answer: False เพราะ Attribution period มีความเหมาะสมกับ product ที่จริงแต่ระยะเวลานานเกินไปก็จะมีปัจจัยอื่นๆมากระทบ ได้ทำให้เราไม่สามารถตัดสินใจได้ว่าการกลับมาซื้อนั้นเป็นผลมาจากการเปลี่ยน banner

1.8. Your company undergoes a total revamp of its corporate identity. Is it the right call?

Answer: False เพราะ ทาตัวชี้วัดยาก

1.9. Elastic ninja at your company wants to show 15 products on the first page of search results instead of 20 products. Should you allow them?

Answer: True

1.10. Marketing person wants to know who respond better to our ads campaigns between iOS users and Android users. How to tell?

Answer: False เพราะ กลุ่มของผู้ใช้งานทั้งสองกลุ่มมี Properties ที่แตกต่างกันไม่สามารถเอามาใช้วัดได้

1.

What are the metrics you should use for the following A/B tests? Assume that the granularities are: page views and unique visitors.

2.1. Which button colors will make customers find it more easily? clicks / __

Answer: clicks/views

2.2. Which sets of products on a landing page will make customers more likely to buy? purchases / __

Answer: purchases / unique visitors

2.3. Which types of promotion coupons will be more effective? purchases / __

Answer: purchases / unique visitors

2.4. Which website layouts will attract more customers to click on sign up button? clicks / __

Answer: clicks / unique visitors

2.

3.1. what are the event-based conversion rate of 2020-11?

3.2. what are cohort-based conversion rate of 2020-11?

Assume 7-day attribution period. Conversion rate is calculated by purchases / unique users.

date	user	event
2020-11-01	A	visit
2020-11-01	A	purchase
2020-11-05	B	visit
2020-11-13	B	visit
2020-11-30	C	visit
2020-12-05	C	purchase

3.1 Answer: 1/3

3.2 Answer: 2/3

3.

```

from scipy.stats import norm
from PIL import Image
import matplotlib.pyplot as plt
z_right = norm.ppf(0.85)
z_b = norm.ppf(0.5)
print(z_right)
print(z_b)

```

```

... 1.0364333894937898
    0.0

```

```

image = Image.open("9.jpg")
plt.figure(figsize=(10,10))
plt.imshow(image)
plt.axis("off")
plt.show()

```

9.

19:26 9.10.22 100%

Assume $\beta = 0.5$ and $p = 0.11$ (conversion ratio) for Bernoulli distribution
 then $\beta^2 = 0.11 \cdot (1 - 0.11) = 0.11 \cdot 0.89 = 0.0979$ with $m = \frac{p_0}{\beta_0} = 0.25$

Answer $n = \frac{m+1}{m} \left(\frac{(z_\alpha + z_\beta)\beta}{MOE} \right)^2$

$$= \frac{1.25}{0.25} \left(\frac{(z_{1-0.15} + z_{0.5}) \cdot 0.0979}{0.01} \right)^2$$

and $z_{0.85} = 1.03643$, $z_{0.5} = 0$

$$n = 5 \left(\frac{1.03643 \cdot 0.0979}{0.01} \right)^2$$

$$n = 614.772$$

$$\approx 616$$

```
z_r = norm.ppf(0.93)
z_l = norm.ppf(0.035)
print(z_r, z_l)
```

```
1.475791028179171 -1.8119106729525978
```

```
image = Image.open("10.jpg")
plt.figure(figsize=(10,10))
plt.imshow(image)
plt.axis("off")
plt.show()
```

10.

10.1 H_0 : Gold's Click rate \leq red Click rate
 H_1 : Gold's Click rate $>$ red Click rate

$$\alpha = 0.01 \text{ and } p = \frac{5904 + 6012}{59504 + 58944} = 0.1006$$

and z is approximately Bernoulli distributed

$$b = \sqrt{p(1-p)} = \sqrt{0.1006(1-0.1006)} = 0.3008$$

$$\text{and } p\text{-value} = \frac{(0.101995 - 0.099170) - 0}{0.3008 \cdot \sqrt{\frac{1}{59504} + \frac{1}{58944}}} = 1.61$$

$$\text{and } z_{score} \text{ and } z_{0.93} \text{ and } z_{0.93} = 1.4756$$

* Rejection region H_0 is $z_{score} > 1.475$ and since Gold's rate is 100%

10.2 Form CI interval and $CI = \hat{p} \pm \frac{z_{\alpha/2} \cdot b}{\sqrt{n}}$

$$\text{and } CI_{red} = 0.099170 \pm \frac{(-1.812) \cdot 0.3008}{\sqrt{59504}} = (0.0967, 0.1014) \quad \text{X}$$

$$CI_{gold} = 0.101995 \pm \frac{(-1.812) \cdot 0.3008}{\sqrt{58944}} = (0.0997, 0.1042) \quad \text{X}$$