Grant King

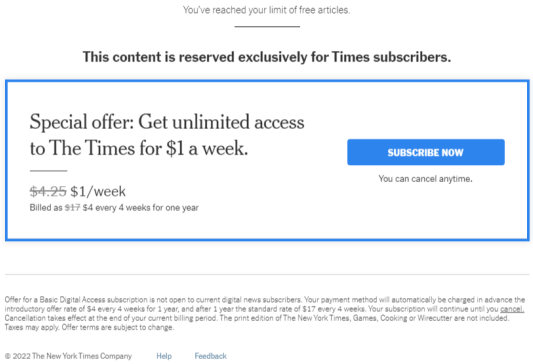
A/B test design #1: New York Times desktop website

Hypothesis

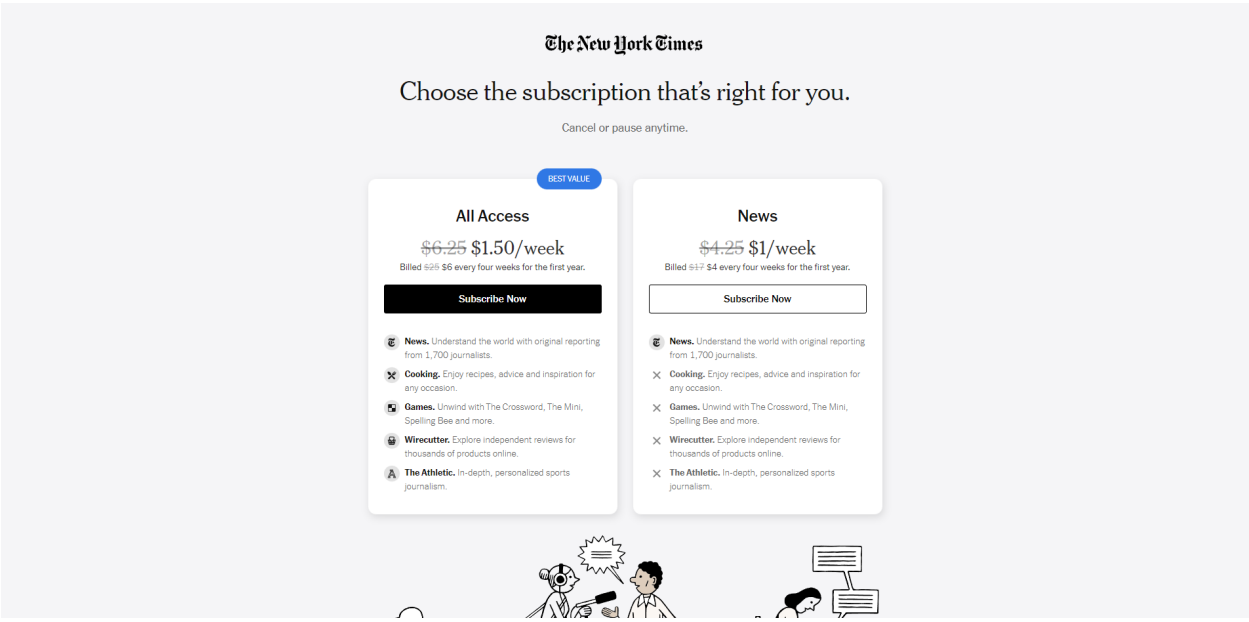
Online subscription purchases will increase by removing slashed prices appearing next to actual subscription prices, excluding those on the final subscription purchasing page (nytimes.com/subscription).

Rationale

Although the slashed prices are placed adjacent to the current prices in order to indicate a deal for the user, they also serve as a distraction from the simple and attractive price of $1. Upon receiving the following pop-up, for example, a new user who isn't paying attention may miss that the price is $1 before closing it:



Removing the slashed prices and isolating the $1 will further highlight it. However, the slashed prices should remain on the subscription purchasing page, as the user has already arrived at the subscription purchasing funnel, and they will enjoy the knowledge of a deal:



Experiment Prerequisites

Overall Evaluation Criterion/Key Metric: Online subscription purchases

Variants

Control Group: Current website design

Treatment Group: Slashed prices removed outside of subscription purchasing page

Randomization Units: Because the treatment is visible to the user, the randomization unit should be no less coarse than user-level: If the randomization unit were sessions or page visits, a single user would observe inconsistent website behavior. Additionally, the impact of the treatment may be most pronounced for new users who are unregistered. Therefore, the randomization unit should be cookies.

Target Population to be Sampled: All unsubscribed New York Times desktop website users

Experiment Design

Example Practical Significance Boundary: Online subscription purchases increase of 2%

Use industry standards for **power of the test** and **significance level** (α): 80%, 5%

Example Sample Size: $16\sigma^2/\delta^2 = 16\sigma^2/(2)^2 = 4\sigma^2$ (depends on assumed standard deviation)

Running Experiment

The A/B test should run for at least one week to account for any bias caused by specific weekdays and for the novelty effect.

Result to Decision

Conduct any necessary sanity checks before launching any practically and statistically significant positive changes, or retesting if results were inconclusive.

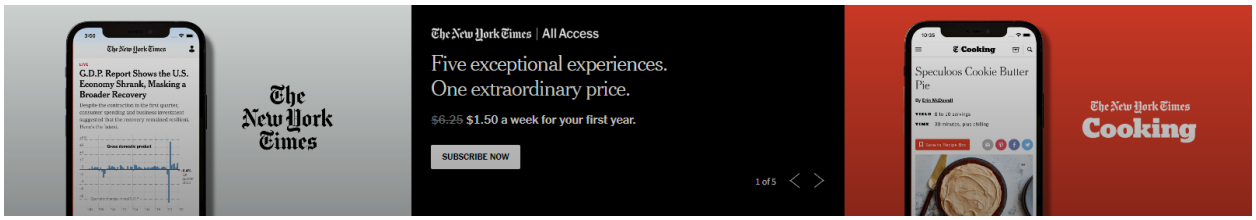
A/B test design #2: New York Times desktop website

Hypothesis

Online subscription purchases will increase by preventing promotions for the "All Access" subscription package from appearing in place of promotions for the "News" subscription on strictly "News" pages.

Rationale

Because the "All Access" subscription is priced at $1.50 per week compared to the "News" subscription's $1, a new user who is glancing at the prices but isn't paying attention may question the coverage of the $1 subscription and/or wonder if there are even more subscription tiers. More importantly, "All Access" promotions should not be shown to users on "News" pages, as those users have not demonstrated an interest in NYT Cooking, Games, Wirecutter, or The Athletic, and a promotion for the "News" subscription could've been shown instead:



Experiment Prerequisites

Overall Evaluation Criterion/Key Metric: Online subscription purchases

Variants

Control Group: Current website design

Treatment Group: "All Access" subscription promotions removed on "News" pages

Randomization Units: Because the treatment is visible to the user, the randomization unit should be no less coarse than user-level: If the randomization unit were sessions or page visits, a single user would observe inconsistent website behavior. Additionally, the impact of the treatment may be most pronounced for new users who are unregistered. Therefore, the randomization unit should be cookies.

Target Population to be Sampled: All unsubscribed New York Times desktop website users

Experiment Design

Example Practical Significance Boundary: Online subscription purchases increase of 3%

Use industry standards for **power of the test** and **significance level** (α): 80%, 5%

- **Example Sample Size:** $16\sigma^2/\delta^2 = 16\sigma^2/(3)^2 = 16\sigma^2/9$ (depends on assumed standard deviation)

Running Experiment

The A/B test should run for at least one week to account for any bias caused by specific weekdays and for the novelty effect.

Result to Decision

Conduct any necessary sanity checks before launching any practically and statistically significant positive changes, or retesting if results were inconclusive.

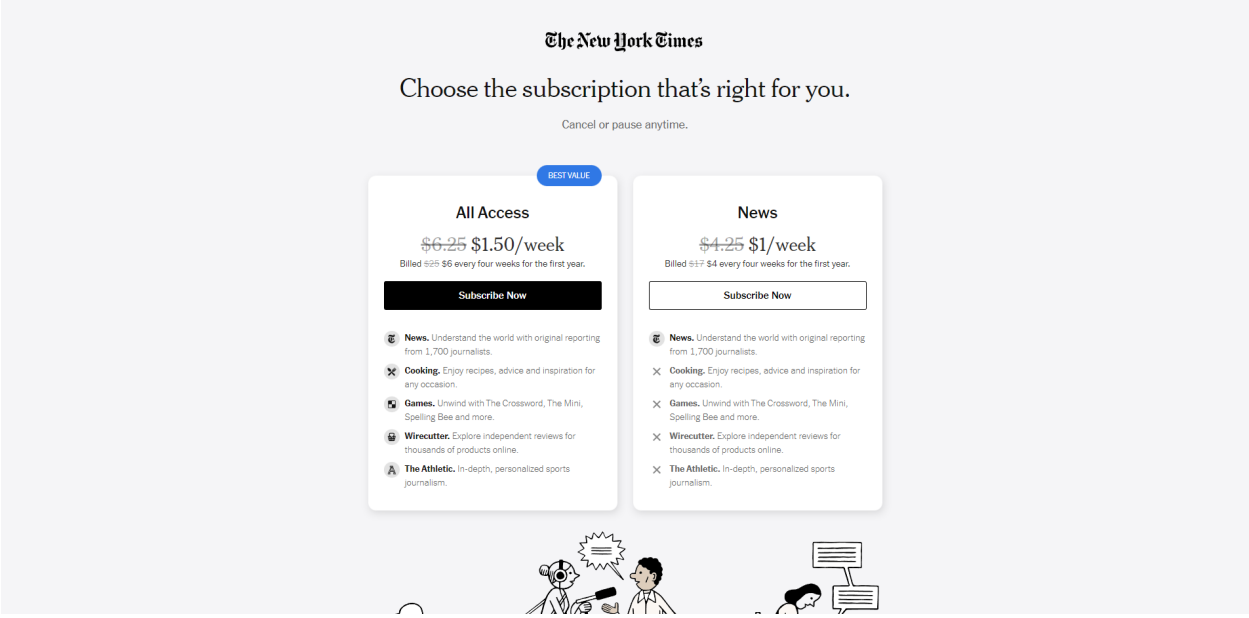
A/B test design #3: New York Times desktop website

Hypothesis

"All Access" subscription package purchases will increase by changing the color of the Xs in the "News" subscription column on the subscription purchasing page to red, or by making the Xs red and additionally changing the color of their adjacent text to red.

Rationale

Currently, the Xs are a gray that is slightly duller than the text in the "All Access" subscription column. The disadvantage of the "News" subscription compared to the "All Access" subscription package would be visually emphasized by red Xs and/or red adjacent text:



Experiment Prerequisites

Overall Evaluation Criterion/Key Metric: “"All Access"” subscription package purchases

Variants

Control Group: Current website design

Treatment Group I: Red "News" subscription column Xs

Treatment Group II: Red "News" subscription column Xs and red adjacent text

Randomization Units: Because the treatment is visible to the user, the randomization unit should be no less coarse than user-level: If the randomization unit were sessions or page visits, a single user would observe inconsistent website behavior. Therefore, the randomization unit should be user IDs.

Target Population to be Sampled: All unsubscribed or unupgraded New York Times desktop website users on the subscription purchasing page

Experiment Design

Example Practical Significance Boundary: Online subscription purchases increase of 4%

Use industry standards for **power of the test** and **significance level** (α): 80%, 5%

Example Sample Size: $16\sigma^2/\delta^2 = 16\sigma^2/(4)^2 = \sigma^2$ (depends on assumed standard deviation)

Running Experiment

The A/B test should run for at least one week to account for any bias caused by specific weekdays and for the novelty effect.

Result to Decision

Conduct any necessary sanity checks before launching any practically and statistically significant positive changes, or retesting if results were inconclusive.