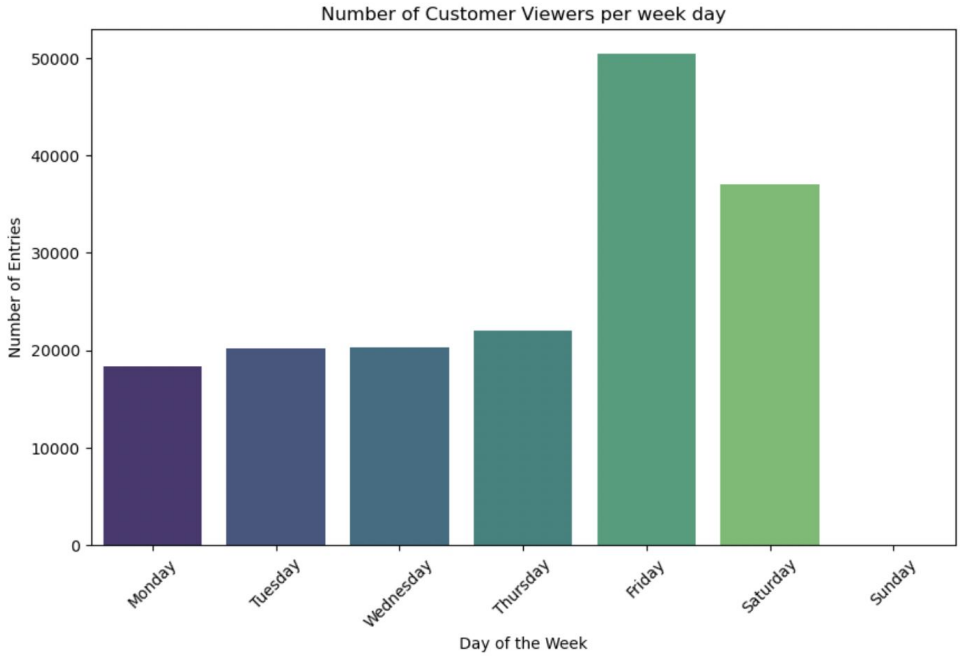
Sushi Star Takeaway

A/B Test Results

Data Exploration

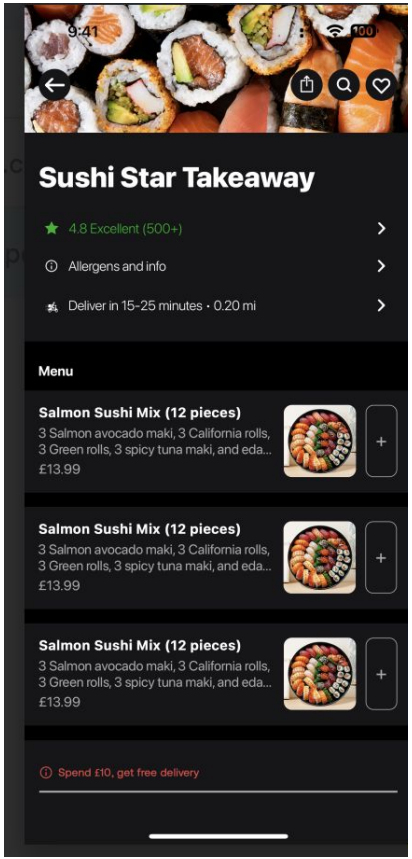
A deep dive into the data

Total Rows in Dataset	326921
# Unique Users	100,000
# Unique Users in Control Group (variation 1)	65,000
# Unique Users in Test Group (variation 2)	35,000

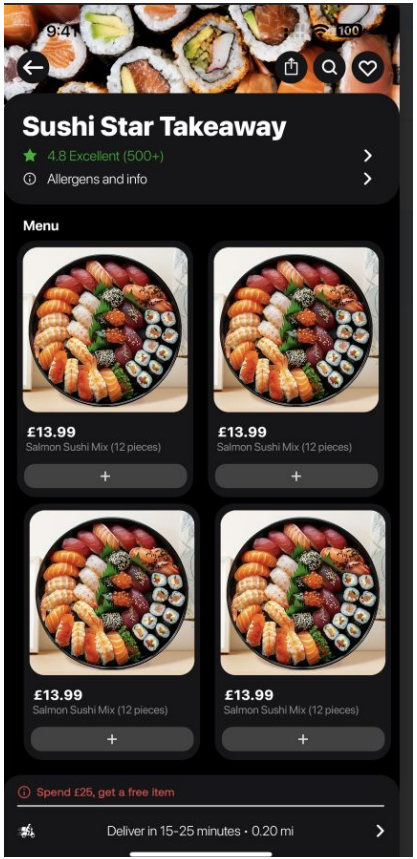


Methodology & Suggested Improvements

1. The mockup for the test design lacks item description and only shows images + item name. Should we consider a design which shows enlarged images as well as item description, and show fewer items per page in order to fit both?
2. Extend the study to capture a full week, including Sundays. Sunday is a weekend and the most popular day to order food in the UK per a 2023 study (https://www.tryotter.com/en-gb/blog/industry/online-food-delivery-statistics).
3. Include user context by looking up user ID to see if the user has ordered Sushi Star Takeaway previously. Then classify users into 3 or more groups - first time customer, previous but not frequent customer (<5 times), or frequent customer (>5 times). Determine frequent customers are significantly less influenced by photo size and if so, consider removing these users from the study.
4. In the UX mockups image A (control) has a promotion where spending £10 gets a free delivery, whereas image B (test) shows that spending £25 gets you a free item. To ensure that these promotions aren't influencing customer decisions, the promotions should be the same across all groups.

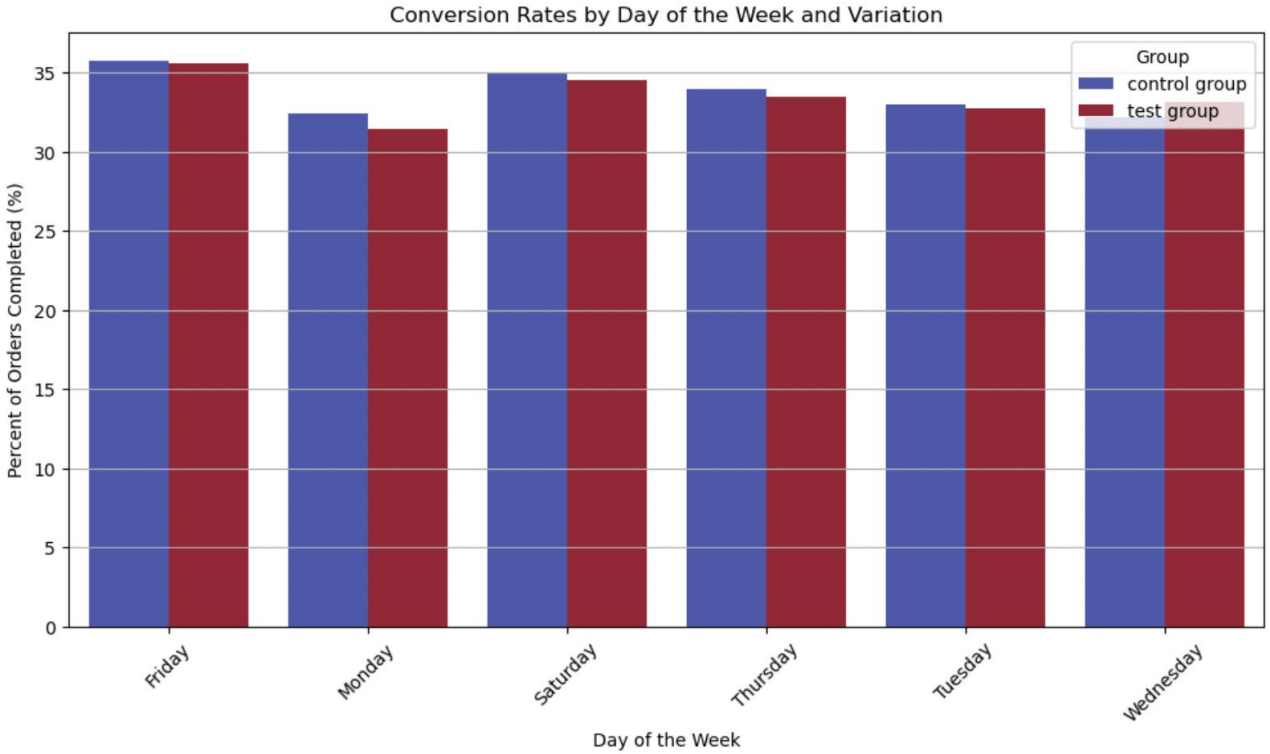


A



B

Test Results



- **Percent of Orders Completed = number of rows with 'order_paid' / number of rows with 'entry_to_shop'**
- **ALL % of completed orders range between 31%-35%, regardless of day of week or size of image.**

Final Results

- Control Group Order Completion Rate: 31.23%
- Test Group Order Completion Rate: 31.04%
- Z - Stat = 0.8
  - a z score of 0.8 means that the observed difference between the two groups (A and B) is **0.8 standard deviations** away from the null hypothesis, which is still small.

- P Value = .42
  - A p-value of 0.42 is quite large, meaning there is a **42% probability** that the observed difference is due to random chance and not an effect of enlarging the images.

- **A P value of > 0.05 means that we do not have enough evidence to determine that enlarging the images made a difference in customer order rate.**