salesforce

Commerce Cloud Platform Adoption Playbook:

# A/B Testing

Test site experiences to maximize results.

# What is an A/B Test?

## A/B Tests Defined

A/B testing is a way to test changes to your site against the current experience and determine which experience produces more positive results. You can test sorting rules, promotions, and content slots with Commerce Cloud out-of-the-box A/B testing functionality. You can also extend your testing beyond basic functionality using custom scripting.

A test segment is defined as the alternative experience. Each A/B test can have **up to five unique test segments**, **one of which is always the control group**. Each test segment targets a portion of the total test participants. Participation in the control group results in a storefront experience that is the same as if the A/B test did not exist.
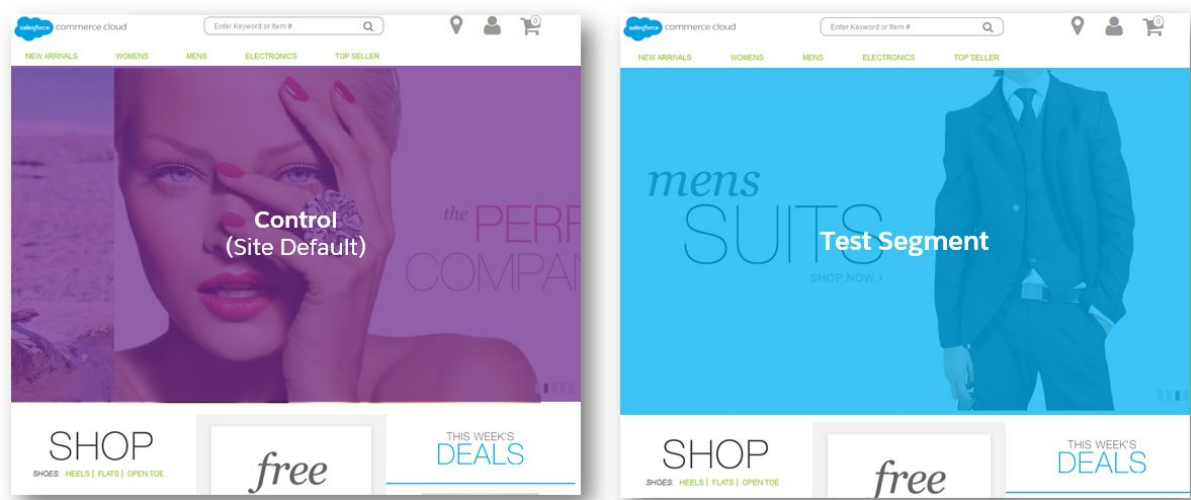


Figure 1. Control vs. Test Group Experience

## A/B Testing vs. Multivariate Testing

- A/B tests include one element (version A) against another element (version B).
- Multivariate tests (MVT) can include multiple variables of the same element.
- A/B tests usually require lower traffic in comparison to multivariate tests.

| A/B Tests | Multivariate Tests |
|---|---|
| Test one variable | Test multiple variables |
| Compare one element | Multiple page elements at once |
| Lower traffic required to run A vs. B | More traffic required to run all variables |

# Creating an A/B Test in Business Manager

The A/B Tests module is located in Business Manager in **Merchant Tools > Online Marketing > A/B Tests.**

In order to create a test, you have to fill out several fields. Some of them are mandatory, and some are optional. Mandatory fields are marked with a star.
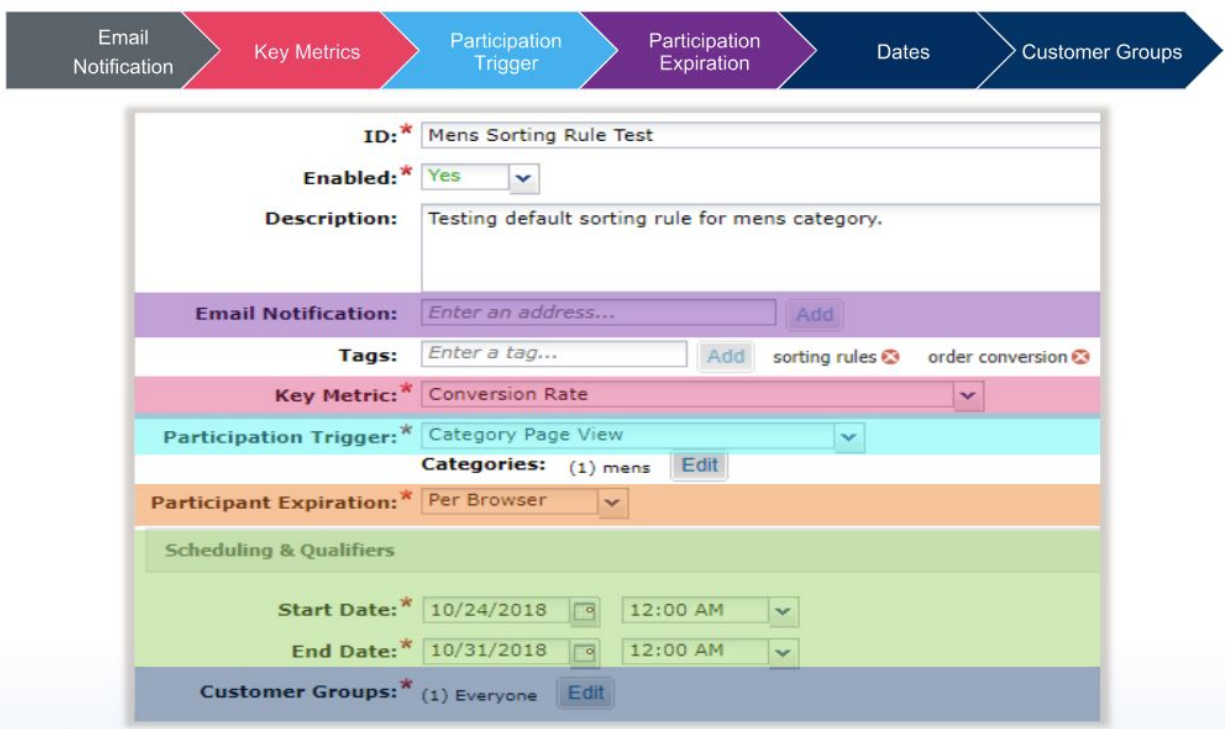


Figure 2. Creating an A/B test.

Fill in the following fields to set up your A/B test.

- ID - Can be any phrase that helps you to find your test among others. Do not use punctuation characters (",/, \ etc.).
- Enable/Disable - Provides you with control of running and pausing the test.
- Description - Details about the test that will help someone to understand quickly what the test is about, without scrolling down and analyzing the test set-up.
- Email notification - Include an email address to get notifications about when the test results become significantly significant, a test is paused, resumed or ends. You can enter more than one email address.
- Tags - Add tags that group your tests by topic.

- Key metrics - Select the measure on which you will base your decision about the efficacy of the changes you are testing.
- Participation trigger - Defines at which point in the customer journey customers are put into the test.
- Participation expiration - Defines when a participation in the test ends; per session, per customer and per browser.
- Start and end dates - define duration of the test.
- Customer groups - define which customer segments are eligible to take part in the test.

## Key Metrics

The key metrics drop-down includes a set of quantifiable measures used to gauge or compare performance.

Examples:
- Company KPI = Conversion Rate
- Metrics = Cart, Checkout & Order Conversion
- Measures = (Orders/Visits), (Orders/Unique Visitors) or (Orders/Sessions)

Different types of A/B tests should have a different key performance metrics. For example, efficacy of sorting rules should be measured by basket metrics, while promotions are best measured by revenue metrics.

The test report will list all the metrics defined in the dropdown, but the one you select as a key metric will be at the top of the report and highlighted.

To get meaningful results it is better to set the key metric as a calculated metric rather than a flat measure. **Calculated metrics are ratios.** More than one metric is used to create it, and it puts it on a scale that is normalized (i.e. conversion rate, which is the number of orders divided by number of visits).

**Flat metrics are single measures.** They are provided for reference, have so much variability and are not as quantifiable. In addition, test segments may have disproportionate traffic allocation. For these reasons, flat metrics are not an accurate indicator of success and you should not use them to make a test decision. You can look at the flat metrics to get an overall picture of numbers and volumes, but they are not good metrics to base overall test success on.

As you can see from the table below, most flat metrics have a corresponding calculated measure. So, if you want to measure baskets, but you are splitting the test 80/20, basket rate will better tell you how your changes are affecting baskets.
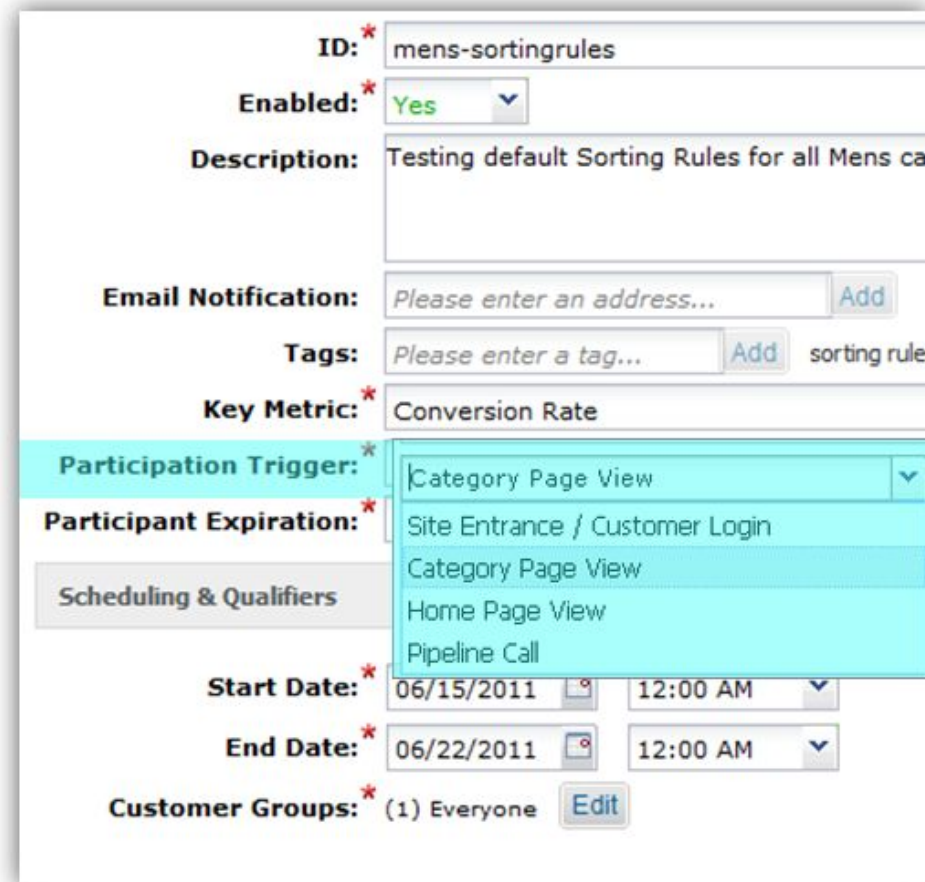
| Flat Metrics | Calculated Metrics |
|---|---|
| Baskets<br>Add to Baskets | Basket Rate |
| Units Added to Baskets | Average Units Added to Basket |
| Units | Conversion Rate |
| Checkouts | Checkout Rate |
| Orders | Sales Velocity |
| Revenue | Average Revenue per Visit<br>Average Order Value |
| Total Discount Amount | Average Discount Amount |
| Discount Amount with Coupon | Average Discount Amount with Coupon |
| Discount Amount without Coupon | Average Discount Amount without Coupon |

**Testing Tip:** Test results can only provide directional guidance. For example, add to baskets can be high but conversion rate low. In such a case, other factors such as checkout process, basket page design, products availability, etc. should be analyzed. There is not always a direct correlation between add to baskets and conversion rate metrics.

## Participation Trigger

Participation trigger defines when customers are put into a test. Not everyone accessing the site will become part of the test, only customers who pass certain trigger points are included in the test segment.

It is important to define the correct participation trigger in order to get reliable test results. Only relevant traffic should be allocated to the test.



Figure 3. Participation triggers.

There are four participation triggers to choose from to determine when site visitors should be added to the test.

- Site Entrance & Customer Login - triggers customers who come to the site on any page.
- Category Page View - triggers customers that view a certain category page.
- Home Page View - triggers customers that have hit the homepage.
- Pipeline Call - triggers customers that have hit a standard Commerce Cloud Pipeline.

> **Testing Tip:** Participation trigger should be based on the location of your test segment and control.
>
> If you are testing a content slot on a category banner, your test location is a category page, and your participation trigger would be a category page view. Do not select site entrance/customer login for this test, as some users may browse the site but never access the category page you are testing. With the wrong participation trigger you would be sending irrelevant traffic to the test.
>
> If you are testing a site wide header banner, your test location is site wide, and your participation trigger would be site entrance/customer login.

## Participation Expiration

Participation expiration defines when participation in a test ends for a user. There are three types of participation expirations - per session, per customer, and per browser.

**Per Session** - active for the duration of a session. Use when you want to provide a different experience each time the customer visits the site.

**Per Customer** - the experience will be active across all devices for a single customer. Lasts for the duration of the test. Use when you're only dealing with logged-in customers and want them to have the same experience on every device they use.

**Per Browser** - Active for all sessions on a browser. Lasts for the duration of the test. Use when you want a customer to have the same experience on a browser for the duration of the test, regardless of whether the customer is logged in.

## Scheduling and Qualifiers

**Dates**

In this section you are scheduling the test by defining a start and end date.

You should allow the test to run long enough to reach significant results. Timing varies by site and depends on a few factors, such as traffic volume, number of test segments, and differences in the test experiences. You can start with two weeks to test the waters, and extend the test if you want to. The test cannot be run longer than 90 days.

**Qualifiers**

Commerce Cloud A/B tests can be directly linked with the customer group functionality in Business Manager. This allows you to target your test to a specific customer segment. You can segment your customer by demographic features, customer expectations, or shopping needs.

There are a few considerations for setting up a test using customer groups:
- There can only be one A/B test at a time per participant. A participant will not see a "blend" of two tests.
- Which test customers from a group are included in is random.
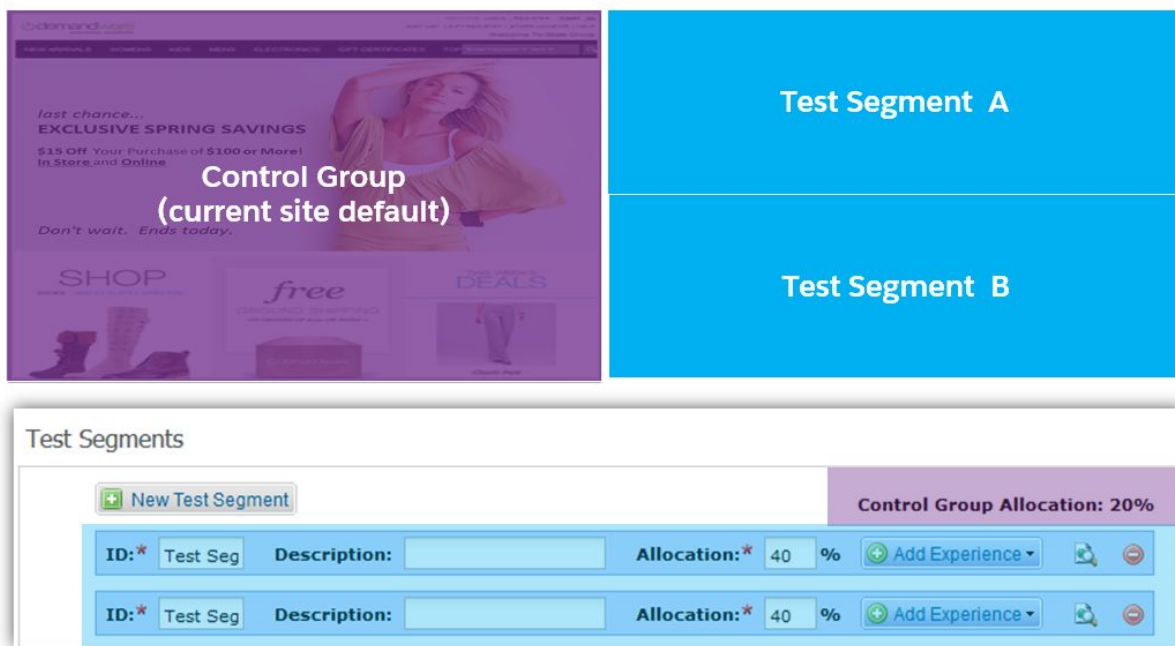- Depending on the trigger, the test may take more time than expected.

- Avoid using very fragmented customer groups, as these groups are normally smaller which can cause the test to not reach a statistically significant result.

> **Testing Tip:** Device specific customer groups are a great idea as customer behavior towards different site elements can significantly differ from one device to another. Desktop customers view content differently than mobile users, therefore the tested changes would influence KPIs differently. If you are making changes to the desktop site, you will only want to test your changes on desktop users otherwise the results may not give a realistic indication of customer behavior.

## Test Segments

A test segment is the set of experiences that a participant will see after qualifying for a test. The control group represents the existing site experience (you don't need to set up a separate test segment for the control group).



Figure 4. Test Segments.

Participation is selected at random by the A/B testing engine every time a participant qualifies for a test. After a participant qualifies for a test with expiration set to per browser, then the first segment experience will always be shown for that participant in that browser.

The allocation percentage indicates the percentage of visits you want routed to this segment. In order to make comparisons, a test requires at least two test segments, one of which must be the control group. If you don't use the control group, your test will be invalid. This is because Commerce Cloud compares segments to the control group, not segments to each other. Statistical significance will only be provided if traffic to the control group is allocated.

> **Testing Tip:** If you are testing promotions, the control group always has to be set to the lowest of the two offers. If not, it will cause the site to override it and give the best discount to the customer.

## Traffic Allocation

The control group represents the status quo experience currently deployed in production, and it is used to determine if the new segments introduce improvements over the existing experience. You start with allocating traffic percentage to the test segments. Any allocation percentage not taken up by the test segments will automatically go to the control group, so make sure you do not set up a segment for the control group.
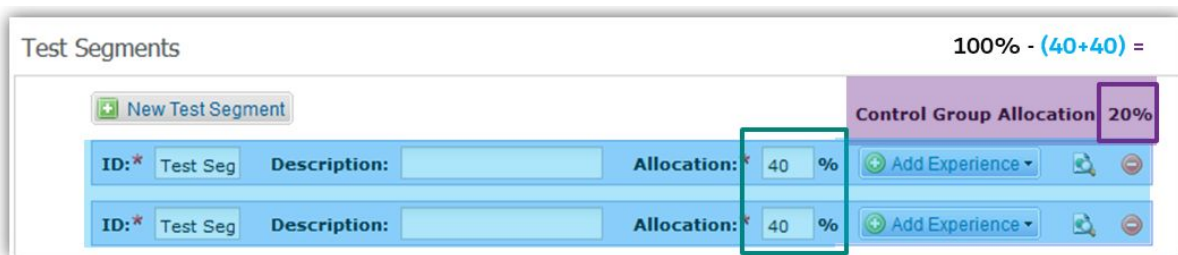


Figure 5. Allocating traffic to test segments and control group.

Traffic Allocation Example

| Segment A | 40% |
|-----------|-----|
| Segment B | 40% |
| Control Group | 20% |
| Total | 100% |

The following limitations will be enforced by quotas:
- Maximum of 3 simultaneous A/B Tests. You can create more, but only three will run at a time.
- Maximum of 5 test segments within a test, including the control group segment - so 4 user-defined segments in total.

> **Testing Tip:** Start with a 50/50 test split to get an understanding of how long it will take to get to statistical significance. A test split at 50/50 will conclude much faster than a test split 80/20.
>
> The higher the variation in allocation for the test segments, the longer the test will take to reach statistical significance. You may want to use a smaller test segment when the test experience varies drastically from the current experience in order to 'play it safe'. You can then refine your test split depending on the results you get.

## Add Experience

You need to add an experience to each test segment to define what you are testing and the experience the test segment will see.

There are 4 options available:
- Slot configuration
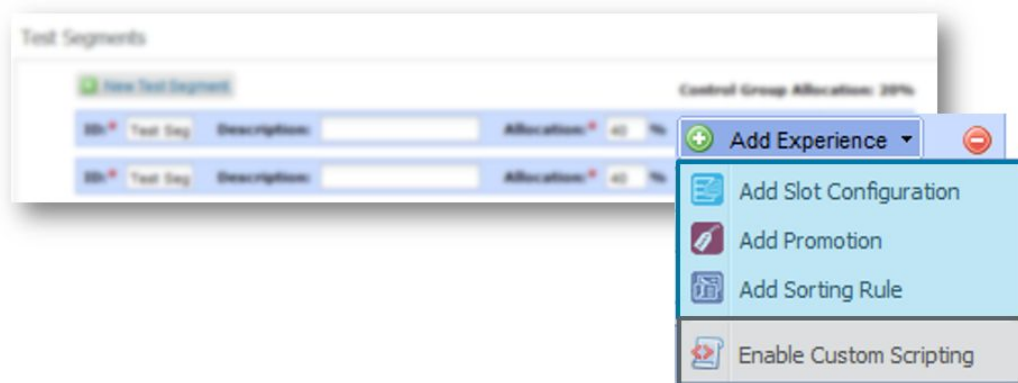- Promotion
- Sorting rule
- Custom scripting



Figure 6. Adding Experiences to a Test.

How to get started - test experiences from level of effort and knowledge. Take a crawl, walk. run approach to testing.

| Beginner | Test sorting rules - most Commerce Cloud clients start with tests on sorting rules, as they are the easiest tests to set up, site users cannot spot the difference easily, and there are no margin decisions involved. |
|---|---|
| Intermediate | Compare different content experiences and find out what content or copy generated higher conversion. You can add a promotion on top of the content test, or test a promotion by itself. This type of test is easy to set up in Commerce Cloud, but it requires creation of additional content, as well as agreeing within the organization what would be acceptable discounts to test. Please note that if you are testing promotions, you will need to set the control group with the lowest offer, otherwise the site will override it and give the best discount to the customer. |
| Advanced | Tests using custom scripting. These tests are more complicated and require developer work. You can still set up your test and monitor results in Business Manager, but you need a developer to create a custom script for you. These are covered in more detail below. |

## Preview the Test

Like all other changes in Business Manager, we recommend setting up A/B tests on Staging and replicating the tests to Production. Before replicating and running live tests, you should preview them and make sure they trigger the experience you intend. The Storefront Toolkit provides you with preview information for test segments. Click on Site Preview from the Toolkit, then select an A/B test segment you want to view on the Staging storefront.
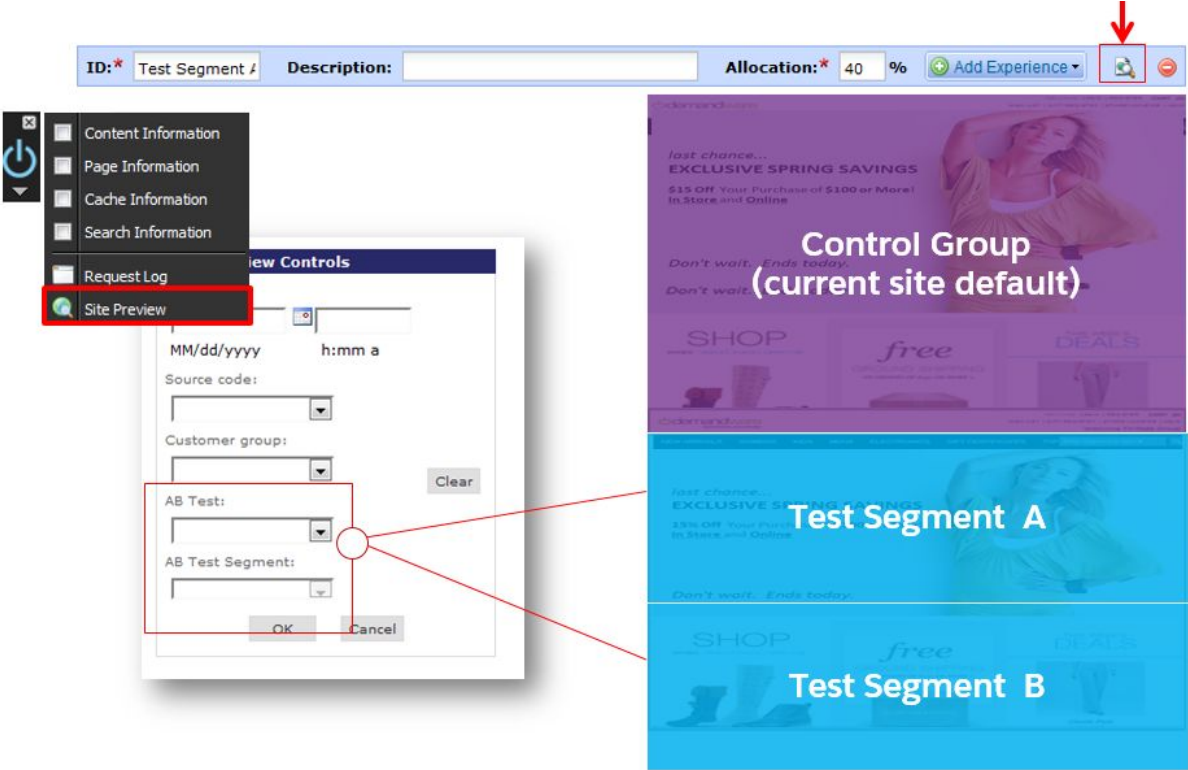
Figure 7. Using Storefront Toolkit to Preview Test Segments

# A/B Test Results

You will receive an A/B test email notification when an A/B test ends. At that point, you will analyze the test results. In some cases, the collected data will have statistically significant results. In other cases, it will not. It is up to you to determine if the test results are significant enough to make a decision, or if a modification needs to be made and the test rerun.

To access a test report, go to **Merchant Tools > Online Marketing > A/B Tests** and click on your test, then go to the **Test Results tab**.

In the first row of the report, you can see the key metric that you selected when you set up your test. Other KPI values for both the segments and control group are provided in separate columns, then delta and statistical confidence level calculated.

**Confidence Level** - the higher the percentage the more certain you can be that the results of the test are significant and not just random chance. When the confidence level reaches 95%, it is deemed a statistically significant result.

**Delta** - percentage change for each metric for each test segment compared to the control group. Delta shows how much better the test segment performed for each metric - so for the conversion rate metric if the delta is 7% then you know that the test metric had a seven percent increase in conversion than the test segment.

If the delta is negative, then it indicates that the test segment actually performed worse than the control, and no confidence level will be captured. Because Commerce Cloud A/B tests are two tailed, this means that you can only see significance in two directions. In order to prove that the control group actually performed better, you would need to re-run the same test but the opposite way around.

If, after analyzing your test results, you decide to make one of the test segments the default site experience, you can apply the changes using the 'deploy' button.
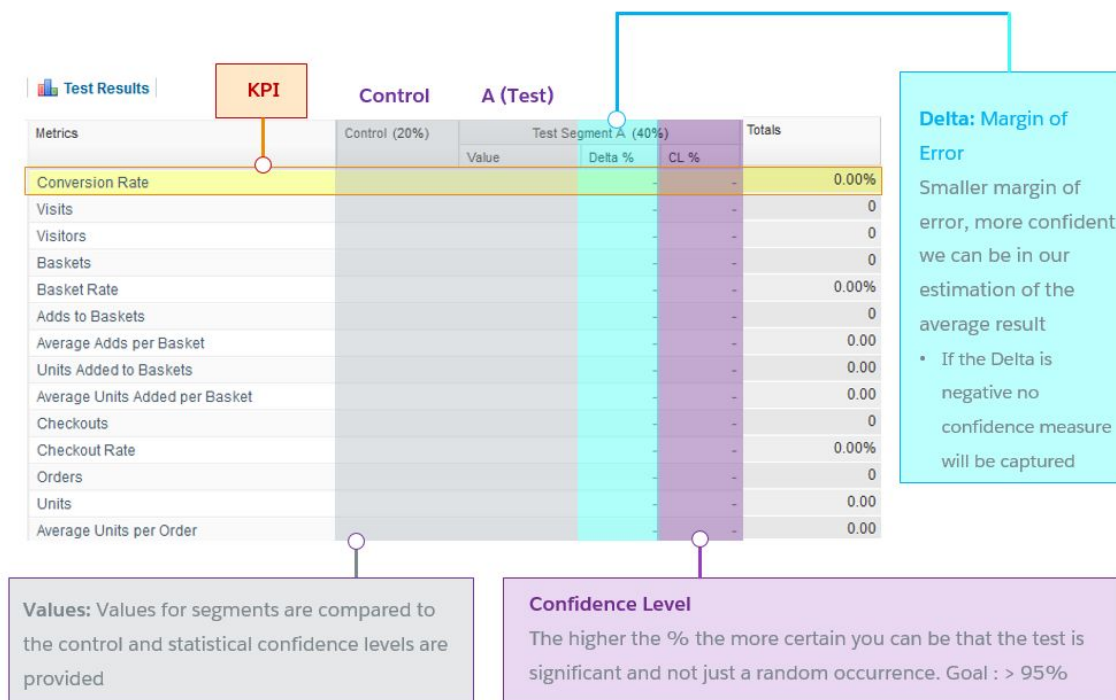
Figure 8. A/B Test Results.

Another important thing to note is that test results can only provide directional guidance. For example, add to baskets can be high but conversion rate low. In such a case, other factors should be analysed such as the checkout process, basket page design, product availability, etc. It is not necessary that there is a direct correlation between add to baskets and conversion rate.

Consider the following when interpreting your A/B test results:

- Welcome and accept failure as a reality – you will always have tests that don't reflect the outcome you expect and many tests that will fail.
- Data is updated in Production every two minutes (browser refresh) and flows down to Staging once a day.
- Be sure to exclude robots from the A/B test analytics. To do this, go to Production Business Manager > Administration > Global Preferences > Analytics, and then add the IP addresses of robots you do not want to be in there.
- Commerce Cloud A/B tests use single directional reporting. The end goal is to prove your hypothesis is the winner. These single tailed tests and can only tell you if your test was not successful. Once you know this, you would need to run the test the opposite way around in order to prove it was successful.
- Remember – there is no set formula for creating successful tests. You can test any one thing against any other thing and there is always the potential to find a winner.
- If you don't see a winner right away, that's ok! Let the test run its course. If this was due to low traffic, run it again sitewide, or in a more heavily viewed category.

# Custom Scripting

Testing using custom scripts extends the possibilities of what you can test but requires frontend development and custom scripts.

These are examples of changes you can test using custom scripts:
● User interface changes
● Fonts, colors, padding and styling changes
● Feature and functionality changes
● Adding or removing Quickview
● Showing or hiding the mini cart
● User experience changes
● Treatment of Quickview on desktop (button vs. rollover)
● Behavior of Quickview on tablet and mobile devices
● Process changes
● 5-step checkout vs. 1-step checkout
● Navigation - removing top tier category click-through (labels)

For more information on creating a test with custom scripting, visit this [guide](#) on XChange.

# Additional Resources

### On-Demand Videos
Listen to and watch [Platform Adoption Webinars](#).

### Ask the Community
Get answers to your questions from Commerce Cloud experts, partners, and other customers via the [Ask the Community](#) forum on XChange. Commerce Cloud XChange login required.

### 1-on-1 Help
Request a [Commerce Cloud Accelerator](#).

### Upcoming Events & Webinars
Review the [calendar](#) of upcoming Commerce Cloud learning opportunities.