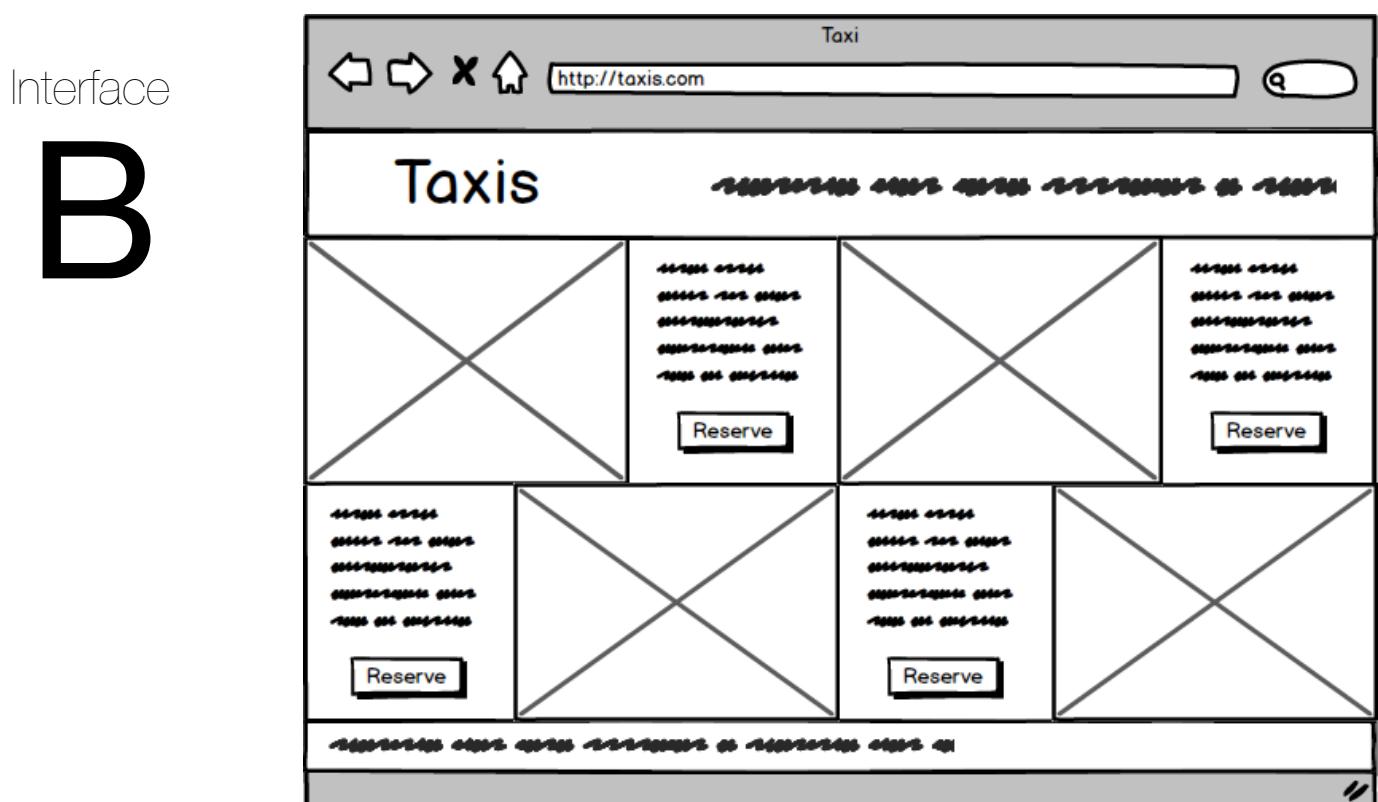
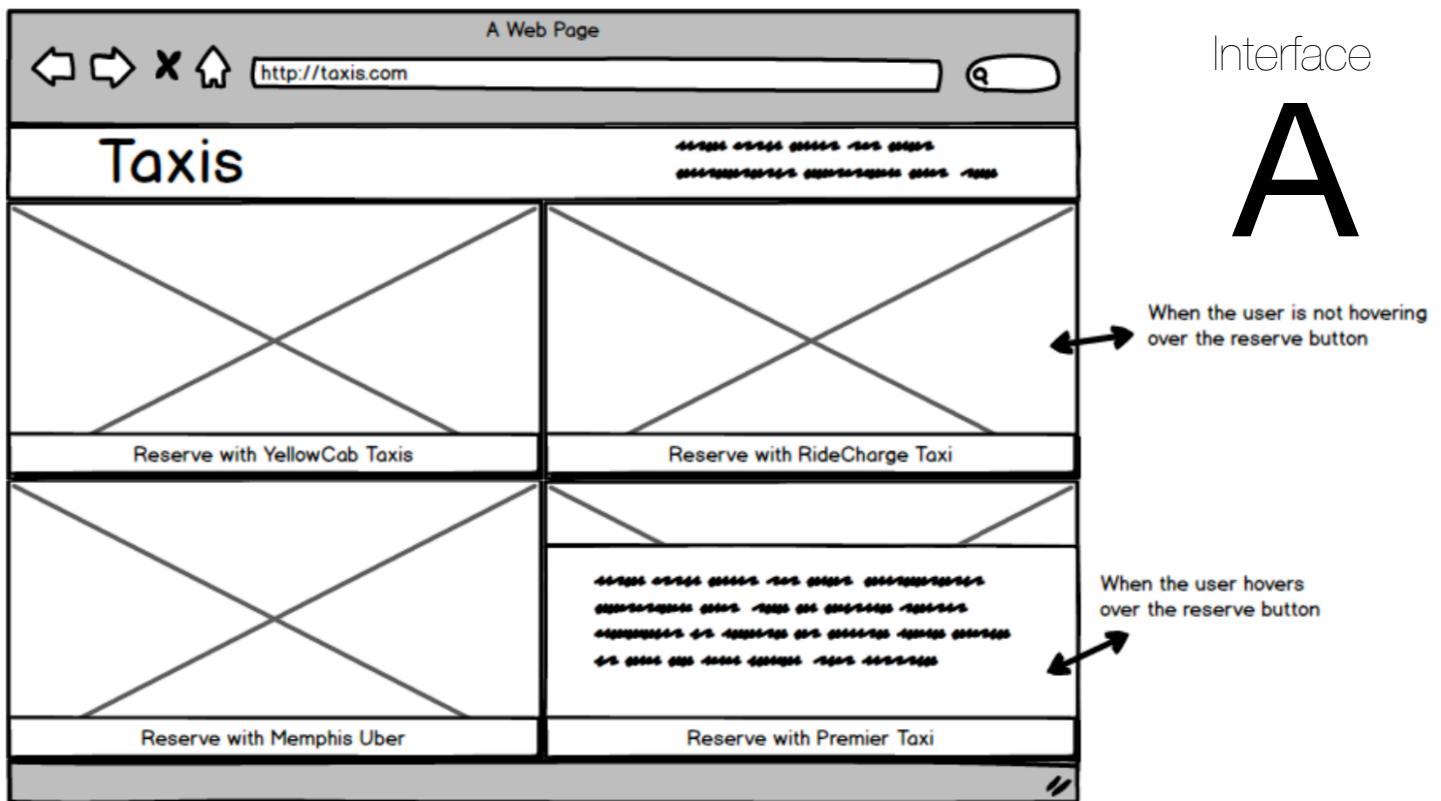


Part One: A/B Testing

Given two proposed interfaces for a Taxi Company, we will use statistical calculations to compare and contrast these two designs. The interfaces will be compared using click through rate, time to click, dwell time, and return rate.



HEROKU LINK: <https://stark-dawn-54724.herokuapp.com/>

Hypotheses

Click-through Rate

1. **Null:** There will be no difference in click rate between interface A and interface B.
2. **Alternative:** Interface B will have a larger click through rate than interface A, because interface B has more obvious clickable buttons.

Time to Click

1. **Null:** There will be no difference in time to click between interface A and interface B.
2. **Alternative:** Interface A will have a higher time to click than interface B because interface A has a hover feature that users will take time to find and then read the newer content.

Dwell Time

1. **Null:** There will be no difference in dwell time between interface A and interface B.
2. **Alternative:** Interface B will have a higher dwell time than interface A, since B has less emphasis on providing information about each taxi and therefore might drive users to gather this information on the linked page instead.

Return Rate

1. **Null:** There will be no difference in return rate between interface A and interface B.
2. **Alternative:** Interface B will have a higher return rate than interface A. With the hover feature of interface A, users are more likely to read the description and click on the links that they intend to use and therefore rarely return back to the site.

A/B Testing Results

Click through rate	12/13
Time to click	9127.625
Return rate	6/12
Dwell time	23795.81818

Statistical Tests

Click-Through Rate

= number of unique clicks / number of unique sessions

=COUNTA(UNIQUE(FILTER(H1:H36, NOT(G1:G36 = 0)))) divided by
=SUMPRODUCT(1/COUNTIF(H1:H6, H1:H6))

We chose to compare our click through data using the **Chi-squared test** of independence, since we were counting the actions of clicking and not clicking.

Chi Square: 3.530

P-value: 0.06025

A	B	C	D	E	F	G	
	Total Users A	Total Users B	Total Users	Total Clicks	Click percentage	No Click percentage:	
	23	26	49	36	0.7346938776	0.2653061224	
	Observed Data	Expected Data	difference()	Obs-Expected^2	quotient()		
A click	14	16.89795918	-2.897959184	8.39816743	0.4969930001		
A no click	9	6.102040816	2.897959184	8.39816743	1.376288308		
B click	22	19.10204082	2.897959184	8.39816743	0.4396476539		
B no click	4	6.897959184	-2.897959184	8.39816743	1.217485811		
					Sum:	3.530414773	

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34

The value in the statistic table for one degree of freedom and a probability of 0.05 is 3.84, which is greater than 3.530 so our result is **not statistically significant**. Therefore, there is no evidence that there is a difference in click-through rates between Versions A and B.

Time to Click

= click time - page load time

$$= (M51 - 9911.166667) / \text{SQRT}(((63-1)*T68^2) + ((83-1)*8544.945827^2)) / (63+83-2)*((1/63)+(1/83)))$$

We chose to use the **Independent Samples T-test** since we are looking in the difference of the means of both data sets. We want to see the difference, on average, of people's interactions and clicks of both of the site formats.

Time to click	avg	(avg- time to click) ^2
18651	10413.46154	67857039.88
24606	10413.46154	201428147.9
22006	10413.46154	134386947.9
2172	10413.46154	67921688.32
1389	10413.46154	81440906.09
3004	10413.46154	54900120.31
20215	10413.46154	96070156.18
8225	10413.46154	4789363.912
9825	10413.46154	346286.9841
5340	10413.46154	25740012
2759	10413.46154	58590781.47
7959	10413.46154	6024381.451
STDEV Time to Click A		7842.173131

Time to click	Avg	(avg- time to click) ^2
10291	9911.167	144273.1079
1302	9911.167	74117756.43
1105	9911.167	77548577.23
3284	9911.167	43919342.45
1473	9911.167	71202662.32
6953	9911.167	8750752
10257	9911.167	119600.4639
2750	9911.167	51282312.8
8067	9911.167	3400951.924
7851	9911.167	4244288.068
1316	9911.167	73876895.76
3695	9911.167	38640732.17
1375	9911.167	72866147.05
23295	9911.167	179126985.8
10829	9911.167	842417.4159
19857	9911.167	98919594.06
23769	9911.167	192039535.5
10857	9911.167	894600.0639
30820	9911.167	437179297.4
4325	9911.167	31205261.75
STDEV Time to Click B		8544.945827

Confidence Interval	difference of means +/- 1.710882* se
plus	161049129.7

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2} \right) \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}}$$

T-TEST (time to click)			
=	(M51-9911.166667)/SQRT(((63-1)*T68^2 + ((83-1)*8544.945827^2))/(63+83-2)*((1/63)+(1/83)))		

$$df = N_1 + N_2 - 2$$

Critical values of t distribution with v degrees of freedom

45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265

degrees (v): 47

t-value: 0.3643788763

Since the calculated t-value for Time to Click was less than the critical value for a t-distribution with 47 degrees of freedom, then our result is **not statistically significant**.

Return Rate

= number of returns / number of clicks

We chose the **Chi-squared test** of independence because we are counting the action of returning.

A	B	C	D	E	F	G	
	Total Users A	Total Users B	Total Users	Total Clicks	Click percentage	No Click percentage:	
	14	22	36	18	0.5	0.5	
	Observed Data	Expected Data	difference()	Obs-Expected^2	quotient()		
A click	7	7	0	0	0		
A no click	7	7	0	0	0		
B click	11	11	0	0	0		
B no click	11	11	0	0	0		
				Sum:			0

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34

Chi-Square: 0

P-value: 1

It is pretty clear that there is no difference between the tests (the click percentage is 50% for both), but the value in the statistic table for one degree of freedom and a probability of 0.5 is 3.84, which is greater than 0. Therefore, the result is **not statistically significant**.

Dwell time

= 2nd page load time - click time

We chose to use the **Independent Samples T-test**, since we are trying to find the difference in average dwell times of users on both website formats.

Dwell Time	Avg	(avg- dwell time) ^ 2
10548	12836.83333	5238758.028
5416	12836.83333	55068767.36
2017	12836.83333	117068793.4
2858	12836.83333	99577114.69
25509	12836.83333	160583808
30675	12836.83333	318200190
STDEV Dwell Time A	11223.0227	

Dwell Time	Avg	(avg- dwell time) ^ 2
4768	16811.1	145036257.6
5876	16811.1	119576412
3884	16811.1	167109914.4
36517	16811.1	388322494.8
8971	16811.1	61467168.01
58150	16811.1	1708904653
3784	16811.1	169705334.4
31345	16811.1	211234249.2
10166	16811.1	44157354.01
4650	16811.1	147892353.2
STDEV Dwell Time B	17785.96691	

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2} \right) \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}}$$

T-TEST (dwell time)	
=(\$M52-16811.1)/SQRT((((\$63-1)*\$77^2) + ((83-1)*17885.17701^2))/(\$63+83-2)*((1/\$63)+(1/83)))	

$$df = N_1 + N_2 - 2$$

Critical values of t distribution with ν degrees of freedom

45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265

degrees (ν): 47

t-value: -1.54696372

Since absolute value of the calculated t-value for Dwell Time was less than the critical value for a t-distribution with 47 degrees of freedom, then our result is **not statistically significant**.

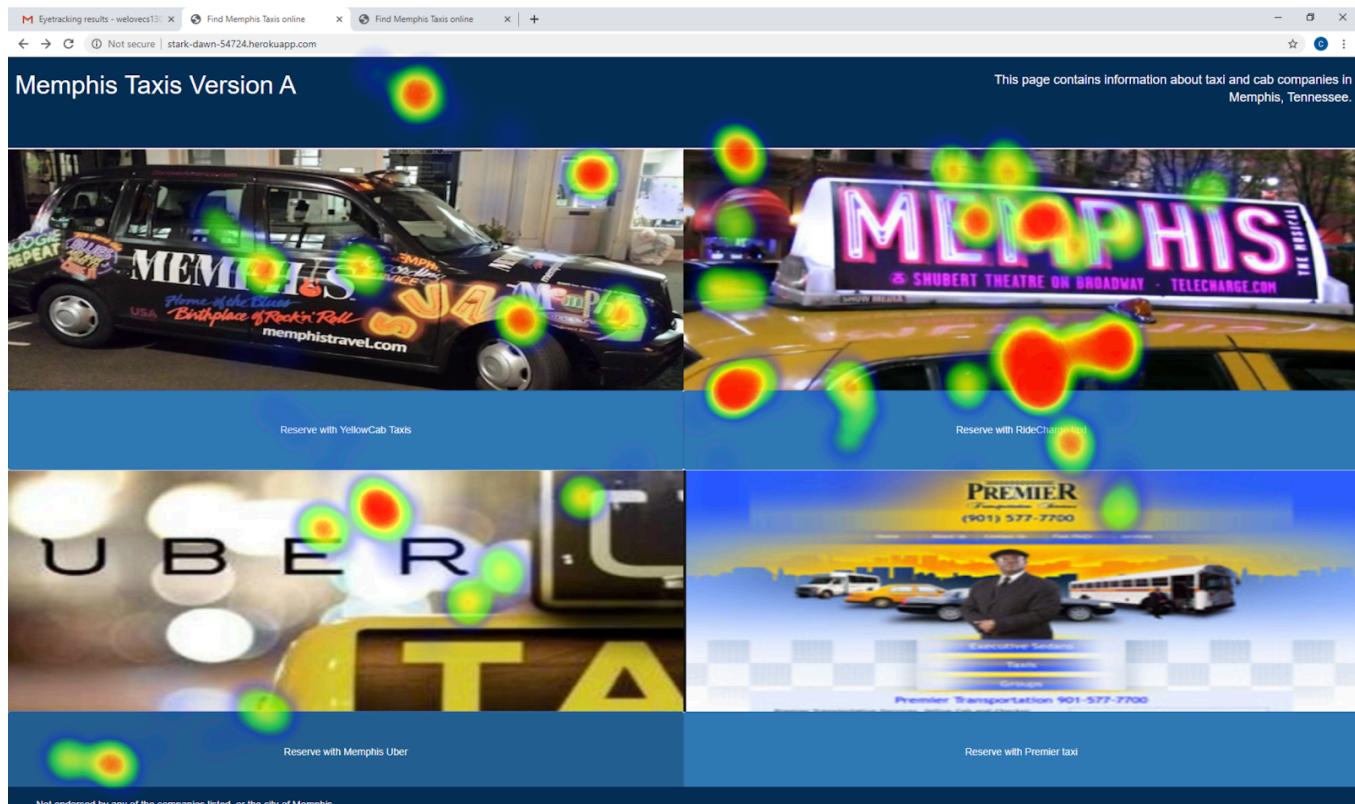
Part Two: Eye Tracking

Eye Tracking Hypothesis

Version A will have more eye gazes at the button because there is more emphasized information with movement, where Version B is more static and easier to ignore.

Session 1: Version A

Heatmap



Action Shot of Replay

Memphis Taxis Version A

This page contains information about taxi and cab companies in Memphis, Tennessee.

Reserve with YellowCab Taxis

Reserve with Ride a taxi

Reserve with Memphis Uber

Reserve with Premier taxi

Not endorsed by any of the companies listed, or the city of Memphis.

Finished Replay

Memphis Taxis Version A

This page contains information about taxi and cab companies in Memphis, Tennessee.

Reserve with YellowCab Taxis

Reserve with Ride a taxi

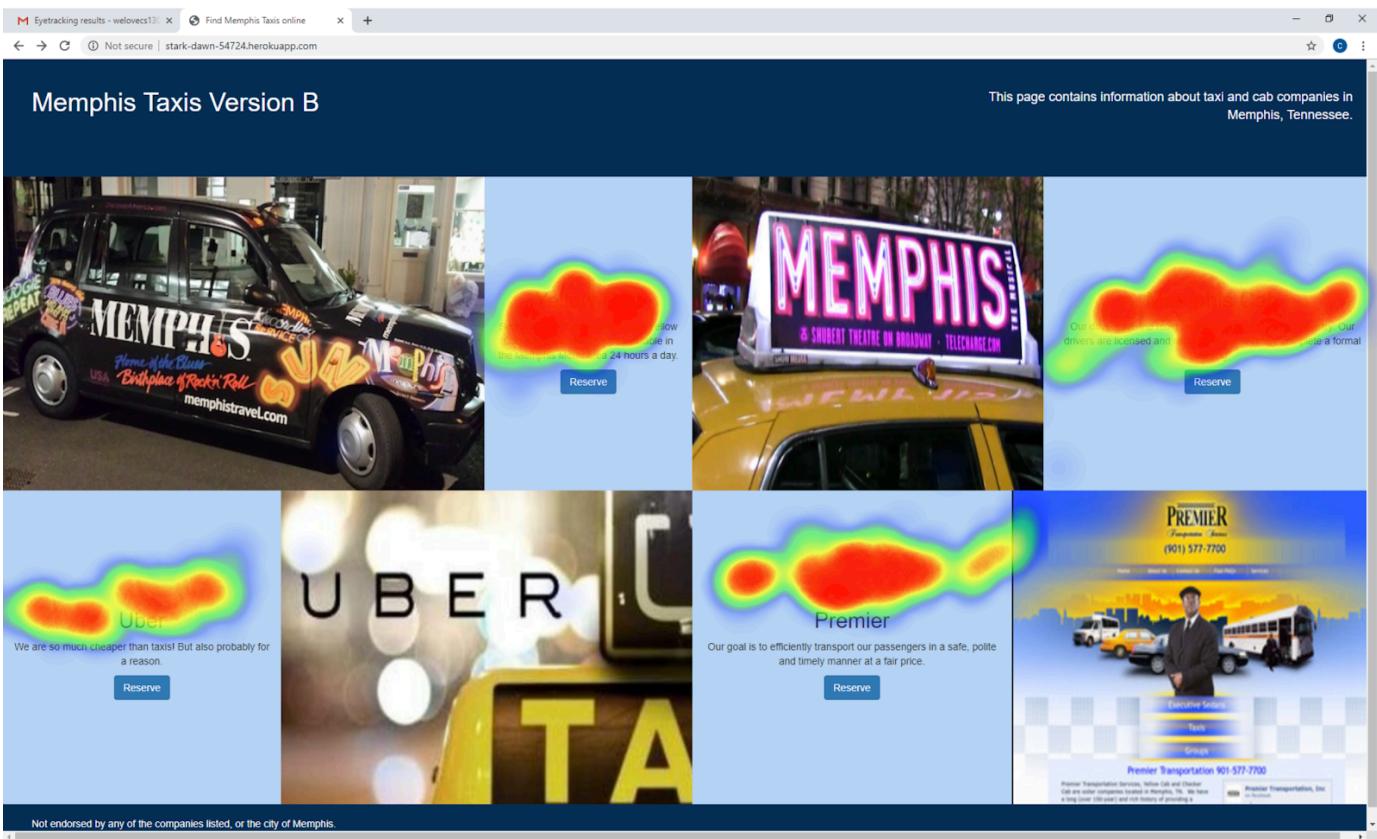
Reserve with Memphis Uber

Reserve with Premier taxi

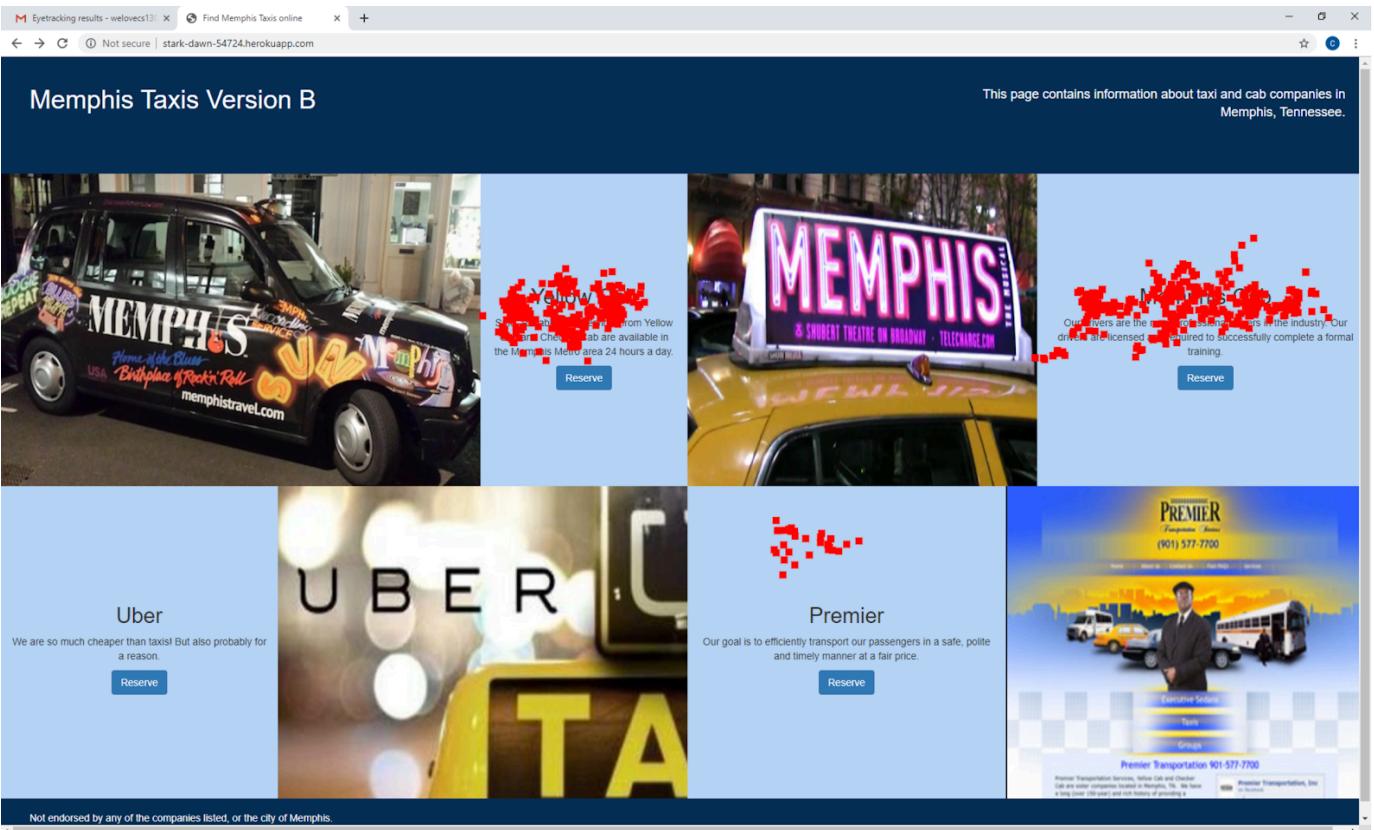
Not endorsed by any of the companies listed, or the city of Memphis.

Session 2: Version B

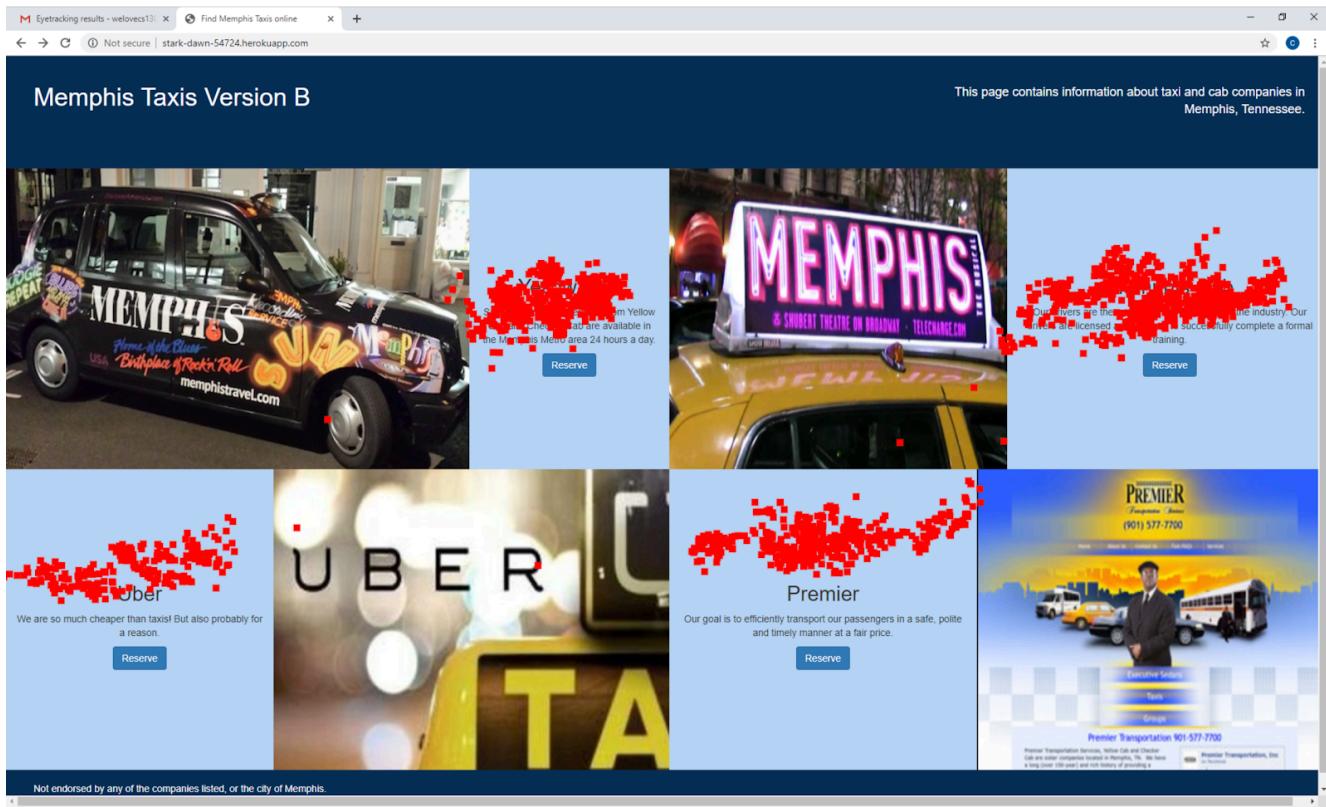
Heatmap



Action Shot of Replay



Finished Replay



Interpreting Results:

Contrary to our hypothesis, Version B had more eye gazes than Version A. While we had initially thought that Version A would attract more gaze due to its moving components, that movement ended up attracting the user's attention towards one taxi, and steering them away from the others. This can be seen through the approximately even amount of focus that the user gives to each of the taxis in Version B, which greatly contrasts with heavily skewed focus given to Memphis Taxi in Version A.

Part Three: Comparison

Question One: We would recommend that Memphis Taxis Co. uses Version B over Version A. According to our tests, users spent more time looking at each company in Version B, thus giving all companies a fair shot. On the contrary, in Version A, users just randomly chose the first company that attracted their attention. In Version B, the user's equal attention amongst the various taxis will allow Memphis Taxis Co. to use their reliable yet affordable pitch to convince users to reserve cars with their company.

Question Two: Our eye-tracking data found much more conclusive and visible results, whereas our A/B Testing data was rather inconclusive. In the A/B testing we felt that the data was rather trivial because in the speed data collection session we felt that the other students in the class were just randomly clicking buttons and had no legitimate goal of booking a taxi. In eye tracking, however, we gave a random user (who hadn't seen this assignment) a specific goal to book a taxi, thus allowing for more realistic and reliable data.

Question Three: A metric measuring dwell time can be used unethically if companies use addictive patterns and confusing return methods on the page that the user left to in order to maximize dwell time. Click through rate can also be used unethically, if companies choose to include misleading buttons and features that entice users to click without realizing what they are trying to accomplish.