

## **Executive Summary:**

This project aims to find the optimal Tile size, Match score, Preview length, and Preview type of the [www.netflix.com](http://www.netflix.com) homepage that minimizes browsing time. We started by running a screening experiment to observe how our factors interact with each other and our response variable and followed up by testing intervals for the global minimum. The final values we concluded were Tile size: 0.2, Match score: 69, Preview length: 75, and Preview type: TT.

**Introduction:**

Netflix is an on-demand Over The Top (OTT) video streaming service and production-based company where users log in and stream video content. The task at hand is to reduce the browsing time of users and avoid the problem of choice overload by choosing the optimal settings for the user landing page. Browsing time can be defined as the time between the users' login and the event of clicking on a content item. Various factors affect browsing time.

We chose to analyze four factors - Tile size, Match score, Preview length, and Preview type. Tile size, Match score, and Preview length are continuous variables, whereas Preview type is categorical. We made use of factorial experiments, hypothesis testing, linear regression, and OFAT(one factor at a time). In addition to the experiments, we used main effect plots and interaction plots to arrive at some conclusions.

In the upcoming pages of this report, we provide a detailed analysis of the design choices made to conduct the experiments, and how we have arrived at the optimal settings for Netflix landing page to reduce the browse time.

## The Experiments:

### Cycle 1:

Since we had no prior knowledge of what values to begin experimenting with, we began by performing a  $2^k$  factorial experiment ( $k=4$ ) for factor screening, to get a general idea of how our factors interact with each other and affect our response variable - Browsing time. We picked our low and high values of numerical variables (Match\_score and Preview\_length) to be at the 33rd and 67th percentiles to cut the range into even pieces and allow us to quickly pinpoint where the minimum would be in a binary-search-inspired style. This design also ensured that we selected values which are distant enough, so that if the factor is significant, the experiment (models and plots), will be able to detect them. The 16 conditions are the combinations of Match Score {33, 67}, Preview Length {60, 90}, Preview Type {TT, AC}, and Tile Size {0.2, 0.4}.

After collecting our data from these parameters, we ran a linear regression model to test the significance of our factors and any possible interactions, with results shown below:

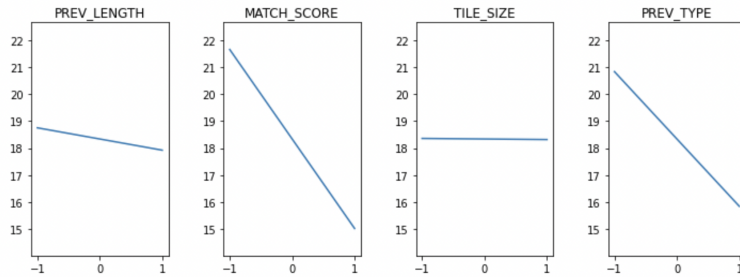
	coef	std err	t	P> t	[0.025	0.975]
Intercept	20.8287	0.035	589.669	0.000	20.759	20.898
Prev_Type[T.1]	-4.9791	0.050	-99.674	0.000	-5.077	-4.881
Prev_Length	-0.3642	0.035	-10.312	0.000	-0.434	-0.295
Prev_Length:Prev_Type[T.1]	-0.0984	0.050	-1.970	0.049	-0.196	-0.000
Match_Score	-3.3166	0.035	-93.895	0.000	-3.386	-3.247
Match_Score:Prev_Type[T.1]	0.0167	0.050	0.335	0.738	-0.081	0.115
Prev_Length:Match_Score	0.0142	0.035	0.402	0.688	-0.055	0.083
Prev_Length:Match_Score:Prev_Type[T.1]	0.0495	0.050	0.991	0.322	-0.049	0.147
Tile_Size	0.0049	0.035	0.138	0.890	-0.064	0.074
Tile_Size:Prev_Type[T.1]	-0.0511	0.050	-1.023	0.306	-0.149	0.047
Prev_Length:Tile_Size	0.0254	0.035	0.720	0.472	-0.044	0.095
Prev_Length:Tile_Size:Prev_Type[T.1]	-0.0401	0.050	-0.803	0.422	-0.138	0.058
Match_Score:Tile_Size	-0.0491	0.035	-1.390	0.165	-0.118	0.020
Match_Score:Tile_Size:Prev_Type[T.1]	0.0095	0.050	0.190	0.850	-0.089	0.107
Prev_Length:Match_Score:Tile_Size	-0.0577	0.035	-1.634	0.102	-0.127	0.012
Prev_Length:Match_Score:Tile_Size:Prev_Type[T.1]	0.0180	0.050	0.361	0.718	-0.080	0.116

### Main Effects:

After analyzing the p-values corresponding to the test statistic of the factor coefficients at a 5% significance level, we inferred that Tile size was not a significant factor, and we will not investigate it onward. Preview length, Preview type, and Match score were all significant factors (or active factors) Match\_Score had the highest main effect.

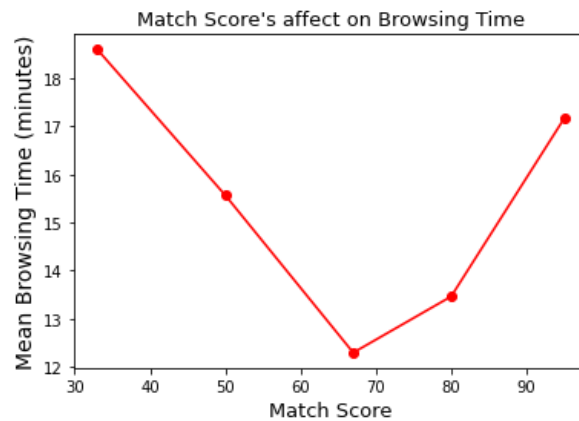
### Interaction Effects:

The only significant interaction we observed was between Preview length and Preview type. Since Match score did not have any significant interactions, we deemed it as an independent variable. As Match score had the highest main effect we decided to find the optimal value corresponding to the Match score first.



## Cycle 2:

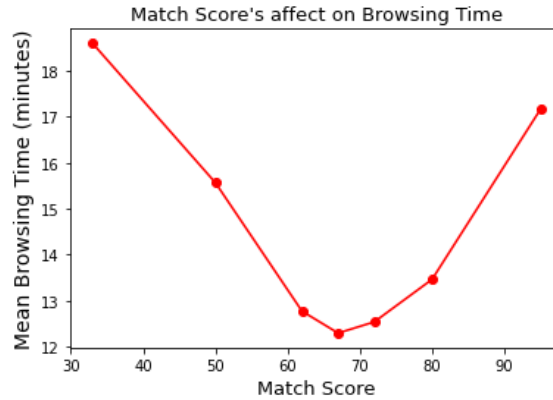
Since we concluded that Match score was independent, we wanted to find the optimal Match score value that minimized browsing time. To achieve this, we needed to narrow down the range since all we could conclude was that 67 was better than 33. So we decided to split up the range of [33,100] into 5 slices so we could get a general idea of the relationship between Match score and browsing time. The values we chose to test were 50, 80, and 95 which gave us values roughly 15 apart and would help us greatly narrow down where the optimal value of the Match score was. We made sure to keep all other variables constant (Preview length = 90, Preview type = TT, Tile size = 0.2), collected the data, and made the following plot.



We concluded that 67 is still the best Match score, and we will further narrow down the range of Match score in the next step.

## Cycle 3:

In the previous experiment, we found 50, 67, and 80 are the top 3 Match scores with 67 being the best. Then we aimed to narrow down the range to 5 around Match score 67. We chose to test Match scores 62 and 69. With all other variables the same, we collected the data and made the following plot.

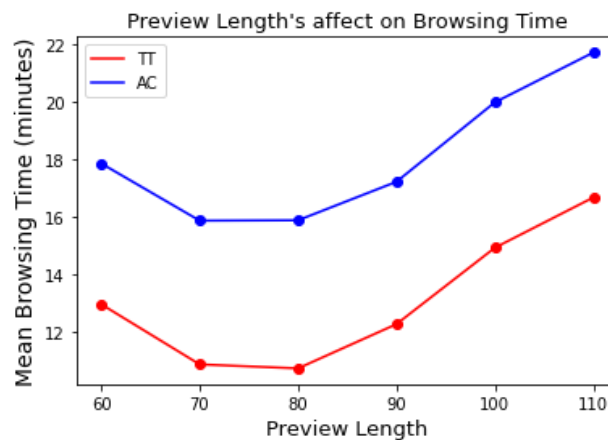


Again we found that 67 is still the best Match score. The differences in mean browsing time among Match scores 62, 67, and 69 are not large at this point. Hence we conclude that our Match score is close to the global optimum, and we will investigate Preview length and Preview type next.

#### Cycle 4:

In our first experiment, we found the interaction between Preview length and Preview type is somewhat significant with a p-value very close to 0.05. To investigate this interaction and hope to optimize either Preview length or Preview type, we chose to test 8 conditions from the combinations of Preview length {70, 80, 100, 110} and Preview type {TT, AC} with Match score 67 and Tile size 0.2. We ran a linear regression model to test the significance with results and interaction effect plot shown below:

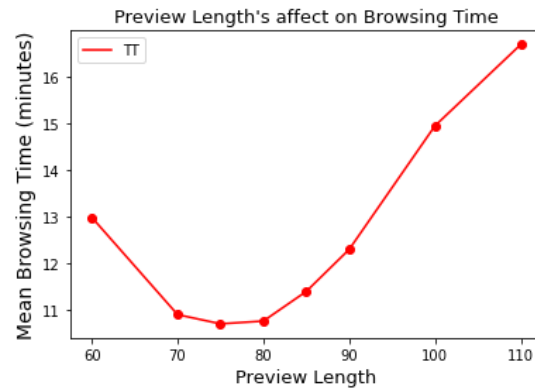
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.1710	0.334	12.481	0.000	3.515	4.827
Prev_Type[T.TT]	-5.0652	0.473	-10.718	0.000	-5.993	-4.138
Prev_Length	0.1579	0.004	43.166	0.000	0.151	0.165
Prev_Length:Prev_Type[T.TT]	7.37e-05	0.005	0.014	0.989	-0.010	0.010



The p-value is very large, suggesting that the interaction effect is not significant. We concluded that Preview type and Preview length are independent. TT was the optimal Preview type, and 80 was the best Preview length so far. We aimed to optimize Preview length in the next experiment.

### Cycle 5:

Now that our previous experiment concluded that Preview length does not depend on any other factor, we decided to treat it as independent and find its optimal value to minimize browsing time. Since our previous most optimal value was 80 and we had tested in intervals of 10, we decided to test values 5 (75 and 85). We concluded that 75 was the best but it was not much smaller than 70. We ran a hypothesis test to test if they were significantly different and we concluded that they weren't at a 5% significance level so we decided to use 75 as our final parameter for Preview length.



### Cycle 6:

Since we've greatly reduced the range of where the optimal point may be for all of our factors, we decided to revisit Match score. Previously we narrowed the range of the optimal Match score to be between 62 and 72 with 67 being the best so we tested around 67 by trying 65 and 69. Between these 3 values, we found a strong negative relationship which allowed us to conclude that 69 was the most optimal value here.

Match score	65	67	69
Browsing Time	11.112070	10.693082	10.323156

**Conclusion:**

In conclusion, our final parameters were 75, 69, 0.2, and TT for Preview length, Match score, Tile size, and Preview type respectively which resulted in an average browsing time of 10.323. It is possible that our initial experiment, which used a  $2^k$  factorial design, did not detect an interaction between factors such as Preview length and Match score, even though one may exist in reality. This is known as a Type II error, meaning a true effect is not detected. This may have occurred due to the particular values of the factors we selected as well as limitations in sample size.