MSDS 629: Final Project

Team: DSTEAM20

Members: Kang-Chi Ho, Ka Chun Chau, Byunghyun Kim, Preetham Pathi

Executive Summary

We are trying to find the optimal design factors (Tile Size, Match Score, Preview Length, and Preview Type) which produce the least browsing time. We performed three experiments, one to find the influential design factors and the rest to find the optimal conditions among those factors. It was found that at the following levels, the optimal browsing time is achieved:

- Preview Length = 75
- Match Score = 75
- Preview Type = TT
- Tile Size = 0.20

Introduction

Our goal is to help the users of Netflix to choose what to watch in the minimum amount of browsing time, which will lead to higher use of the service, and thus, more revenue for the company. While there are an infinite number of factors that could affect the browsing time of the users, we consider four specific factors: Tile Size, Match Score, Preview Length, and Preview Type. And even with just four factors, there are an infinite number of possible combinations of conditions of each factor, which are impossible to experiment with the limited resources. So we tried to find the combination that would be close enough to the optimal conditions through two phases of experiments.

In the first experiment, we find out which of the factors are the most important ones by constructing main/interaction effect plots and running partial F-tests. Here, we observe that Preview Length, Match Score, and Preview Type have a significant influence on the users' browsing time and that the interaction effects of Tile Size on other factors are negligible. Among two to three conditions of each factor, we find that 60 for Preview Length, 75 for Match Score, TT for Preview Type, and 0.25 for Tile Size result in the minimum expected browsing time of **11.643 minutes**.

In the second experiment, we experiment using main/interaction effect plots with new conditions over a reduced range based on the results of the first experiment. Here, we gain further insight into the main and interaction effect of Preview Length and Match Score. With that, our new optimal values are 60 for Preview Length, 75 for Match Score, TT for Preview Type, and 0.2 for Tile Size, which results in an expected browsing time of **11.9241 minutes**.

In the third experiment, we decided to turn in the opposite direction in Preview Length and Match Score with the small step of 5 based on the result of the second experiment. We observed a turning point in the interactive plot and obtained the minimum expected browsing time of **10.066 minutes** with the optimal values of 75 for Preview Length, 75 for Match Score, TT for Preview Type, and 0.2 for Tile Size.

Experiment

Objective

To mitigate the decision paralysis phenomenon among customers, Netflix actively devises mechanisms aimed at helping users choose quickly among its inventories. The objective of this study is to investigate the optimal design factor configurations that minimize average browsing time, i.e. metric of interest (MOI).

Plan

Instead of a full factorial experiment, which is costly, inefficient and might require more data than is practically feasible, the study is divided into 2 experimental phases.

First Experiment

The objective of the first experiment is factor screening, i.e. to identify the most important factors on the response variable. Despite not exploring the full condition space, by carefully selecting the levels for each factor, it is believed that this approach would allow us to establish the general relationship between design factors and the response variable.

Subsequent Experiments

Results of the screening experiment inform us of the significant design factors and the vicinity of the optimal levels thereof. With this information, subsequent experiments are designed to find the best condition within this vicinity.

Data

For both experiments, 100 units are assigned for every combination of design factor levels.

Results & Analysis

First Experiment

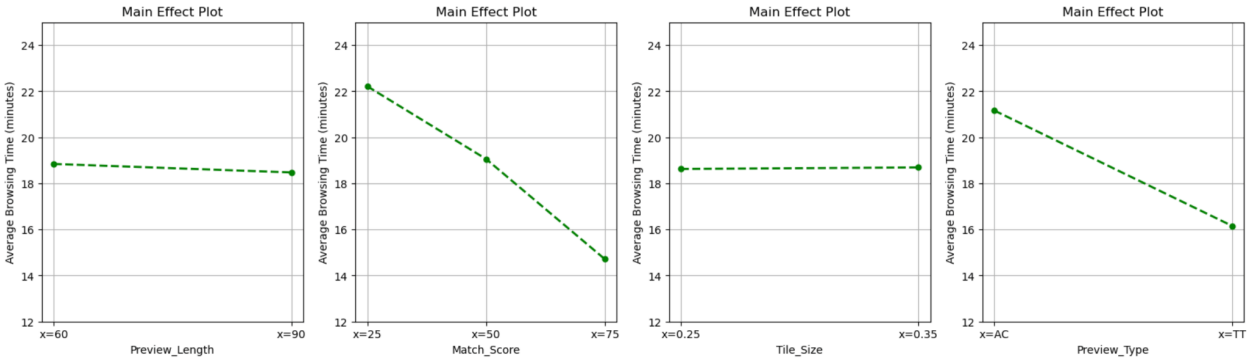
Experiment Parameters

The details of the configurations of design factor levels for the first experiment are stated below.

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	{60, 90}	{25, 50, 75}	{0.25, 0.35}	{TT, AC}

Main Effects

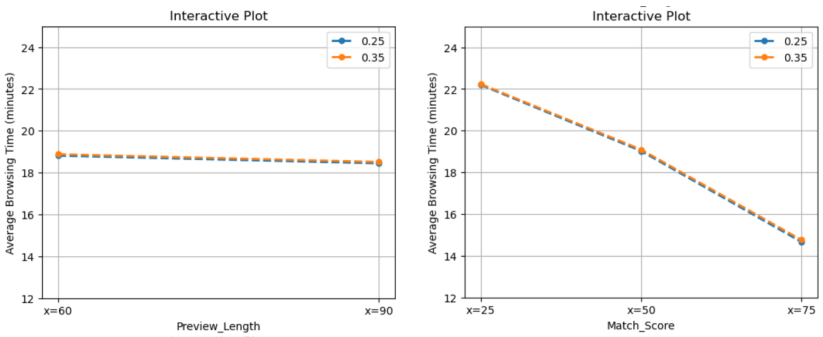
The main effect plots and partial F test results of each design factor with respect to the response variable are shown below.



Hypothesis	P-value
$H_0 : \beta_{Preview.Length} = 0$ vs $H_A : \beta_{Preview.Length} \neq 0$	5.25×10^{-14}
$H_0 : \beta_{Match.Score} = 0$ vs $H_A : \beta_{Match.Score} \neq 0$	0.00
$H_0 : \beta_{Tile.Size} = 0$ vs $H_A : \beta_{Tile.Size} \neq 0$	0.17
$H_0 : \beta_{Review.Type} = 0$ vs $H_A : \beta_{Review.Type} \neq 0$	0.00

1. **Preview Length**
   The main effect plot indicates that a longer Preview Length slightly reduces the average browsing time. The results of the Partial F Test suggest that Preview Length is a significant factor at a 1% significance level.
2. **Match Score**
   The main effect plot indicates that a higher Match Score noticeably reduces the average browsing time. The results of the Partial F Test suggest that the Match Score is a significant factor at a 1% significance level.
3. **Tile Size**
   The main effect plot indicates the response variable appears to be indifferent to the change in Tile Size. The results of the Partial F Test suggest that Tile Size is an insignificant factor at a 1% significance level
4. **Preview Type**
   The main effect plot indicates that changing the Preview Type from AC to TT significantly reduces the average browsing time. The results of the Partial F Test suggest that Preview Type is a significant factor at a 1% significance level.

Interaction Effects



From the interaction effect diagrams, it is observed that Tile Size has negligible interaction effects with other design factors.

Optimal Condition

The optimal values for each design factor in the first experiment:

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	60	75	0.25	TT

The expected browsing time is 11.643 minutes with a 95% confidence interval between (11.510, 11.775).

Conclusion

Based on the above analysis of the main effects of and interaction effects between different design factors, we conclude that Preview Length, Match Score, and Preview Type have a significant influence on user browsing time, whereas Tile Size is insignificant. Moreover, while both are significant, the difference in p-values suggests that Match Score and Preview Type are more significant than Preview Length.

Second Experiment

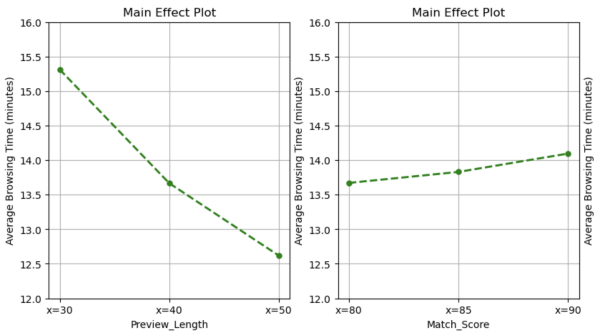
Based on the conclusion of the first experiment, we observed the significance of each factor and the direction to minimize the average browsing time.

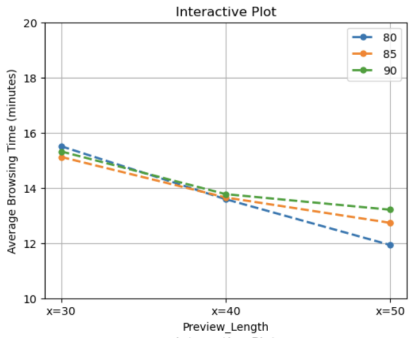
Experiment Parameters

The details of the configurations of design factor levels for the second experiment are stated below.

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	{30, 40, 50}	{80, 85, 90}	{0.2}	{TT}

Main Effects & Interaction Effects





1. **Preview Length:**
   The main effect plot indicates that increasing Preview Length in the three levels helps decrease browsing time. The optimal level for Preview Length is 50 in this case.
2. **Match Score:**
   The main effect plot indicates that decreasing Preview Length in the three levels helps decrease browsing time. The optimal level for Preview Length is 80 in this case.
3. **Match Score & Tile Size**
   The main effects of Match Score and Tile Size are not investigated since the factors are held at a constant level in this experiment.

We conclude from the main effect plots that Preview Length has a more significant influence on the response variable than Match Score within the selected levels.

We observe that only Preview Length and Match Score have meaningful interaction. As the three lines aren't parallel, we can conclude that the two factors have a slight dependency. The optimal level for Preview Length is 50, and Match Score 80.

Conclusion

The optimal values for each factor in the second experiment:

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	50	80	0.2	TT

The expected browsing time is **11.924 minutes** with a 95% confidence interval between (11.718, 12.131).

Third Experiment

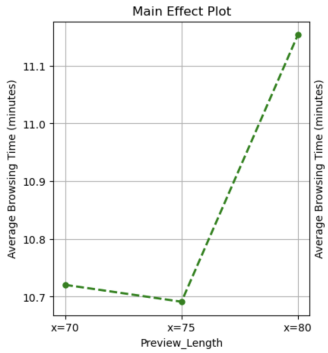
Based on the conclusion of the second experiment, we decided to turn in the opposite direction in Preview Length and Match Score with the small step of 5.

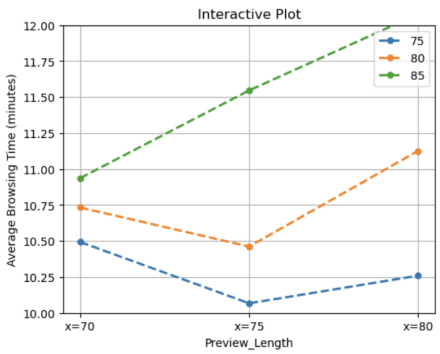
Experiment Parameters

The details of the configurations of design factor levels for the third experiment are stated below.

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	{70, 75, 80}	{65, 70, 75}	{0.2}	{TT}

Main Effects & Interaction Effects





From the main effects and interaction effect plot, we identify that when Preview Length=75 and Match Score=75 , the optimal browsing time is achieved.

Conclusion

The optimal values for each factor in the second experiment:

Design Factor	Preview Length	Match Score	Tile Size	Preview Type
Levels	75	75	0.2	TT

The expected browsing time is **10.066 minutes** with a 95% confidence interval between (9.852, 10.281).

Conclusion

Results

Experiment	Prev. Length	Match Score	Preview Type	Tile Size	Expected Browse Time
1	60	75	TT	0.25	11.643
2	50	80	TT	0.2	11.924
3	75	75	TT	0.2	10.066

As the average browsing time in the third experiment is the shortest among all experiments, the optimal condition for the Netflix project will be:

- Preview Length = 75
- Match Score = 75
- Preview Type = TT
- Tile Size = 0.20
- **Expected browsing time = 10.066 minutes**

Limitations

Response surface method can be used to find a better optimum than using a certain predefined set of experiments.