Final Project-Maribel Viveros

2023-11-1

Exploratory Eye Gaze Analysis

Introduction

Eye-based activity recognition has become a focal point in Human-Computer Interaction (HCI) and Ubiquitous Computing (UbiComp). Primarily centered on identifying reading behaviors and associated cognitive processes, this field leverages distinct eye movement patterns that vary across different activities and factors. In a study conducted by Srivastava and colleagues in 2018, the researchers aimed to identify sedentary activities marked by minimal physical movements by analyzing eye movements. The broader goal of their study was concerned with designing computing systems that can proactively monitor daily activities, providing assistance or encouragement towards a healthier lifestyle. Srivastava et al. (2018) expanded on prior research, emphasizing the potential of eye tracking as a promising method for activity recognition. Additionally, the researchers developed a classifier combining existing low-level gaze features with novel mid-level gaze features. When applied to the dataset that included 24 participants engaged in a range of desktop computing activities, the outcomes demonstrated an overall accuracy of activity recognition.

Prior studies have shown that first fixation duration can serve as a measure of visual information acquisition (Holmqvist, et al., 2018). First fixation, or initial gaze, is the amount of time an individual spends fixating their gaze on a specific area of interest during an eye-tracking experiment. Given that prior researchers have already identified eye movements that are operationalized into mid-level, low-level, and high-level categories suitable for classification, my study seeks to ascertain which features are most appropriate for this purpose.

The proposed study explores the dataset from Srivastava et al. (2018) containing specific eye behaviors, including initial gaze, and time spent on activity, and examine if inital gaze and duration can be used for desktop activity recognition.

The current research project posits: Can the application of first fixation metrics (a common eye-tracking measure) contribute to desktop activity recognition? This isn’t very precise, but that’s okay: Part of the goal of this EDA project is to clarify eye-metrics that contribute to understanding of visual perception.

Setup

```
suppressPackageStartupMessages({
  suppressWarnings(library(tidyverse))   # for working with the data
  suppressWarnings(library(skimr))       # generate a text-based overview of the data
  suppressWarnings(library(visdat))      # generate plots visualizing data types and missingness
  suppressWarnings(library(plotly))      # generate interactive plots
  suppressWarnings(library(readxl))
  library(downloader)
  library(openxlsx)
  library(viridis)
  library(tinytex)
  library(purrr)
  library(dplyr)
})
```

```
## Warning: package 'downloader' was built under R version 4.2.3
```

```
## Warning: package 'viridis' was built under R version 4.2.3
```

```
## Warning: package 'viridisLite' was built under R version 4.2.3
```

```
## Warning: package 'tinytex' was built under R version 4.2.3
```

Methods

Dataset

The Srivastava and colleagues (2018) gaze dataset used in this project is publicly accessible on Kaggle. It encompasses raw gaze coordinates (x-y) and timestamp data obtained from 24 participants actively engaged in eight specific desktop activities; Read, Browse, Play, Search, Watch, Write, Debug, and Interpret. Each individual has to go through 3 sets of the task.

```
#kaggle_url <- "https://www.kaggle.com/datasets/namratasri01/eye-movement-data-set-for-desktop-activities/do
wnload?datasetVersionNumber=1" Location of files for download
```

The initial gaze dataset was collected using the Tobii Pro X2-30 eye tracker. After visiting the kaggle website, the datafiles were downloaded into a 'data' folder. The variables 'participant', 'set', and 'activity' serve as identifiers. The data includes rows with unique x, y, and timestamp values, with each row identified by the raw_row_number.

```
file_paths <- file.path("data") # Set the path to the data folder where you downloaded the files from Kaggle
insert your file path here' with the actual path to the downloaded files
```

To facilitate the examination of desktop activity identification, these individual participant datasets were merged into a unified file, resulting in a dataset with approximately 1505813 rows (observations); 6 columns (variables), 3 variables are handled as characters, and 3 as numeric.This is necessary to support the robustness of predictions derived from the integrated dataset. Within the dataset, the variables x and y represent the spatial coordinates, while the timestamp indicates the specific time point when the gaze was at the corresponding x, y coordinate. The timestamp is instrumental in representing durations or the "time spent" on a particular activity.

```
output_file <- "combined_dataset.csv" # The combined resulting csv file  will be saved in your working direc
tory.
csv_files <- list.files(file_paths, pattern = "\\.csv$", full.names = TRUE)
list_data <- map(csv_files, read.csv)
combined_data <- do.call(rbind, list_data)
write.csv(combined_data, file = output_file, row.names = FALSE)
```

As part of the EDA process, I looked at dimensions of the dataframe and column (variable) types outlined by Peng and Matsui (2016). First, I confirmed the dataset aligned with the metadata description, and three variables, x, y, and timestamp were listed as detailed in the combined dataset below.

```
skim(combined_data)
```

Data summary

Name	combined_data
Number of rows	1505813
Number of columns	6

Column type frequency:	
character	3
numeric	3
________________________	
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
participant	0	1	3	3	0	24	0
set	0	1	1	1	0	3	0
activity	0	1	4	9	0	8	0

Variable type: numeric

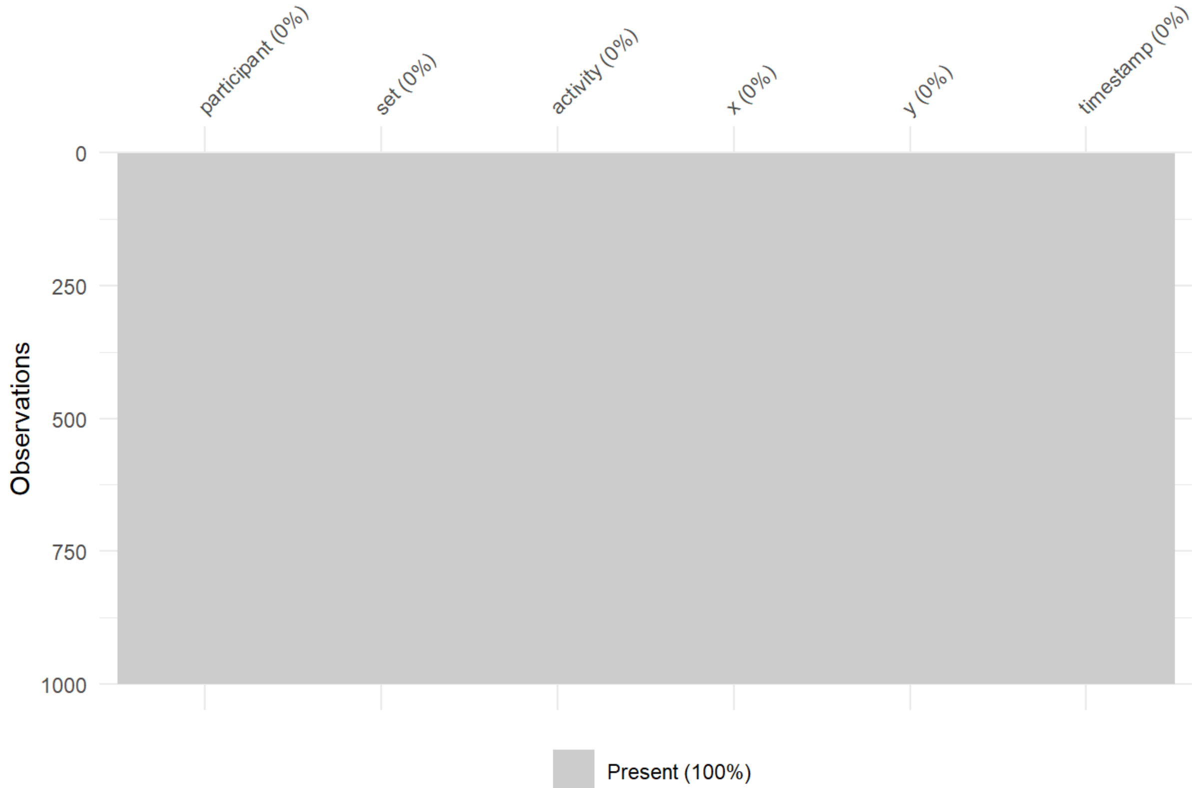
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
x	0	1	699.35	379.71	-1858	352	728	971	3574	
y	0	1	445.96	217.85	-1077	308	428	561	1917	
timestamp	0	1	153451.03	91141.00	0	75082	151728	228822	399184	

Missing values

Notably, the participant, set, and activity variables did not contain any missing values. Arguments in vis_miss() are useful for picking up patterns in missing values. However, because of large data which caused a error to examine the data set for missing values ‘sample_n’ to draw a subset was used.

```
set.seed(123)
dataf_smol = sample_n(combined_data, 1000) #sample to big so we'll draw a subset #but no missing data

vis_miss(dataf_smol)
```



Data Summary

The metadata specifies there should be 24 participants, so I confirmed in the resulting combined data set, the 24 participants, 3 tasks, and 8 different activities. Additionally, the metadata also specifies how many activities the participants have to complete. All the activities are included in the dataset. Thus, for the motivating question, the data set is appropriate as the eye gaze coordinates and timestamp is 100% complete. Also, the desktop activity data is 100% complete.

```
count(combined_data, participant)
```

```
##    participant     n
## 1          P01 61547
## 2          P02 59409
## 3          P03 68951
## 4          P04 65807
## 5          P05 61333
## 6          P06 61837
## 7          P07 73707
## 8          P08 64309
## 9          P09 62580
## 10         P10 68749
## 11         P11 65877
## 12         P12 53768
## 13         P13 68526
## 14         P14 60231
## 15         P15 66974
## 16         P16 68649
## 17         P17 52505
## 18         P18 44061
## 19         P19 54962
## 20         P20 72588
## 21         P21 64206
## 22         P22 65004
## 23         P23 56978
## 24         P24 63255
```

```
count(combined_data, set)
```

```
##   set      n
## 1   A 528567
## 2   B 507778
## 3   C 469468
```

```
count(combined_data, activity)
```

```
##    activity      n
## 1    BROWSE 214128
## 2     DEBUG 152997
## 3 INTERPRET 141485
## 4      PLAY 209728
## 5      READ 216161
## 6    SEARCH 183226
## 7     WATCH 249644
## 8     WRITE 138444
```

Validation

In the original methods section, the authors detail that all of the activities last about 5-6 minutes. Therefore the timestamps for the coordinates for each activity for all participants should not exceed this time, and indeed the max duration for one activity in going through all the values is ~6 minutes (399184 ms). A web search leads us to the website for the article where the data set is used: *https://www.researchgate.net/publication/329955224_Combining_Low_and_Mid-*

Level_Gaze_Features_for_Desktop_Activity_Recognition (https://www.researchgate.net/publication/329955224_Combining_Low_and_Mid-Level_Gaze_Features_for_Desktop_Activity_Recognition) and published.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    max_duration = max(timestamp - first(timestamp)),
    min_x = first(x),
    min_y = first(y)
  ) %>%
  arrange(desc(max_duration))
```

```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 192 × 5
## # Groups:   activity [8]
##    activity participant max_duration min_x min_y
##    <chr>    <chr>              <int> <int> <int>
##  1 WATCH    P24               399184   195  1092
##  2 WATCH    P07               396323   538   954
##  3 WATCH    P17               394897   527   325
##  4 WATCH    P01               394598   540  1067
##  5 WATCH    P03               394594   940   974
##  6 WATCH    P19               389977   188  1056
##  7 WATCH    P10               389362   720  1118
##  8 WATCH    P11               388227   691   436
##  9 WATCH    P15               382229   680  1065
## 10 WATCH    P08               379463  1098    56
## # ℹ 182 more rows
```

To verify that all the activities are within the expected time duration of the maximum limit of ~6 minutes (399184 ms) I identified the count of time stamps in the dataset less than 399184. A value of FALSE in this context indicates that there are entries in the combined_data dataset where the timestamp is not greater, so all of the entries were completed in the expected time.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(timestamp) %>%
  mutate(too_long = timestamp > 399184) %>%
  count(too_long)
```

```
##   too_long       n
## 1    FALSE 1505813
```

Results

This study focuses on discerning patterns in visual engagement durations for distinct activities, aiming to identify variations in the amount of time allocated to each activity. The investigation addresses the question: Can we distinguish activities requiring prolonged or shorter periods of visual engagement?

The first fixation, represents the immediate processing of the attended stimulus and serves as a measure of attention. This data point will align to our question of which eye-metric can help us detect visual stimuli and activity? There should be specific first landing coordinates that differ across activity, and I would expect the duration of first time to fixation to differ across activity.

In the table below, mean durations were computed for each group (activity) by analyzing the time elapsed from the first timestamp to subsequent timestamps within each activity. The analysis utilized the differences between timestamps to derive the average duration of the initial instance for each activity. This was needed to examine the first fixation duration.

Longer mean durations, are indicative of heightened attention or engagement, and were observed for ‘BROWSE’ (151699.8 ms) and ‘WATCH’ (182171.3 ms) activities. Conversely, ‘DEBUG’ exhibited a shorter mean duration (142958.9 ms). Examination of the min x and min y coordinates, representing the initial landing positions for activities, aimed to ascertain where and if there are distinctions in the earliest gaze locations. The similar x, y coordinates, indicate differences of initial landing point based on activity type. ‘READ’ and ‘WRITE’ activity x and y coordinates have similar inital glances.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(timestamp) %>%
  group_by(activity) %>%
  summarise(mean_duration = mean(timestamp - first(timestamp)),
            min_x = first(x),
            min_y = first(y))
```

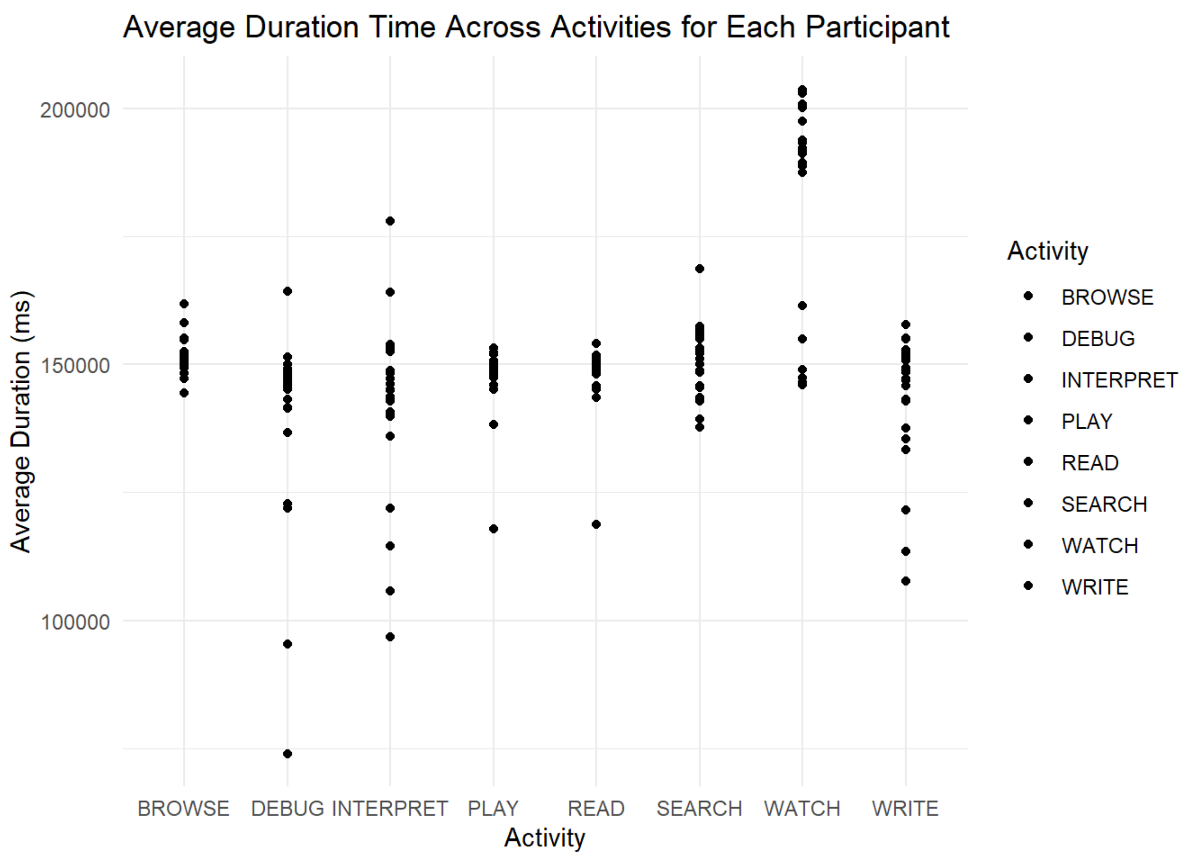
```
## # A tibble: 8 × 4
##   activity  mean_duration min_x min_y
##   <chr>             <dbl> <int> <int>
## 1 BROWSE          151700.  1016   459
## 2 DEBUG           142959.   341   319
## 3 INTERPRET       143692.   488   606
## 4 PLAY            148100.   160     3
## 5 READ            148165.   985   504
## 6 SEARCH          151359.   726   465
## 7 WATCH           182171.   540  1067
## 8 WRITE           145069.   973   481
```

Plot

The initial plot illustrates there is some variability in the location of the initial glances based on different activity. In review of the ‘time spent’ averages, the time participants spend observing a stimulus there is also a difference across activities. Additionally, there appears to be clear differences in average fixation times by activity. The plot shows the average duration (avg_duration),and minimum x-coordinate (min_x), and minimum y-coordinate (min_y) for each ‘activity’ and ‘participant.’

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    avg_duration = mean(timestamp - first(timestamp)),
    min_x = first(x),
    min_y = first(y)
  ) %>%
  ggplot(aes(x = activity, y = avg_duration, fill = activity)) +
  geom_point() +
  labs(title = "Average Duration Time Across Activities for Each Participant",
       x = "Activity",
       y = "Average Duration (ms)",
       fill = "Activity") +
  theme_minimal()
```

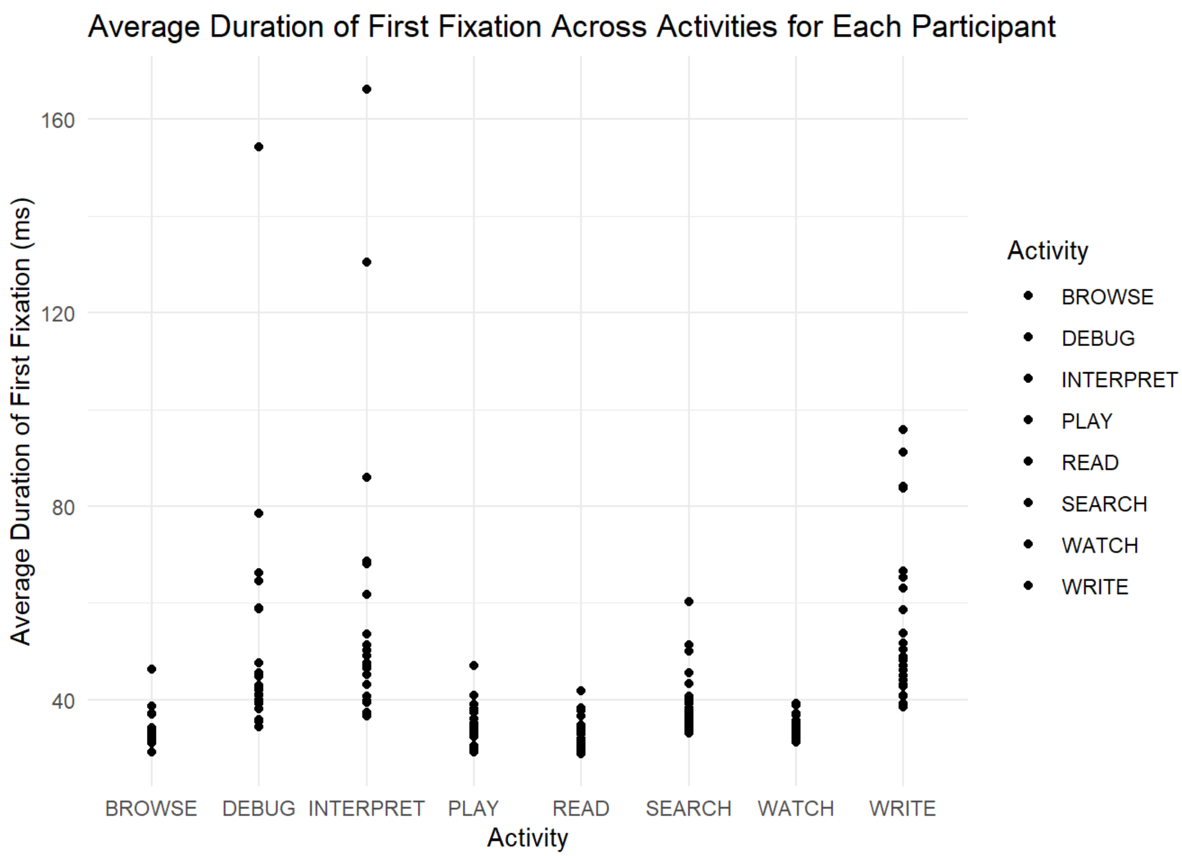
```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```



The second graph focuses on assessing the average duration of the first fixation-first glance, which represents the time difference between the first and start of timestamps. The differences in the time spent during the first glance aims to reveal more noticeable distinctions in first fixation times across various activities and participants. The code calculates the first fixation duration by determining the time difference between the timestamp of the first fixation and the start time (0).

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(participant, activity, timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    avg_duration = mean(timestamp - first(timestamp)),
    min_x = first(x),
    min_y = first(y),
    avg_first_fixation_duration = mean(c(0, diff(timestamp)))  # Calculate average duration of the first fix
ation
  ) %>%
  ggplot(aes(x = activity, y = avg_first_fixation_duration, fill = activity)) +
  geom_point() +
  labs(title = "Average Duration of First Fixation Across Activities for Each Participant",
       x = "Activity",
       y = "Average Duration of First Fixation (ms)",
       fill = "Activity") +
  theme_minimal()
```

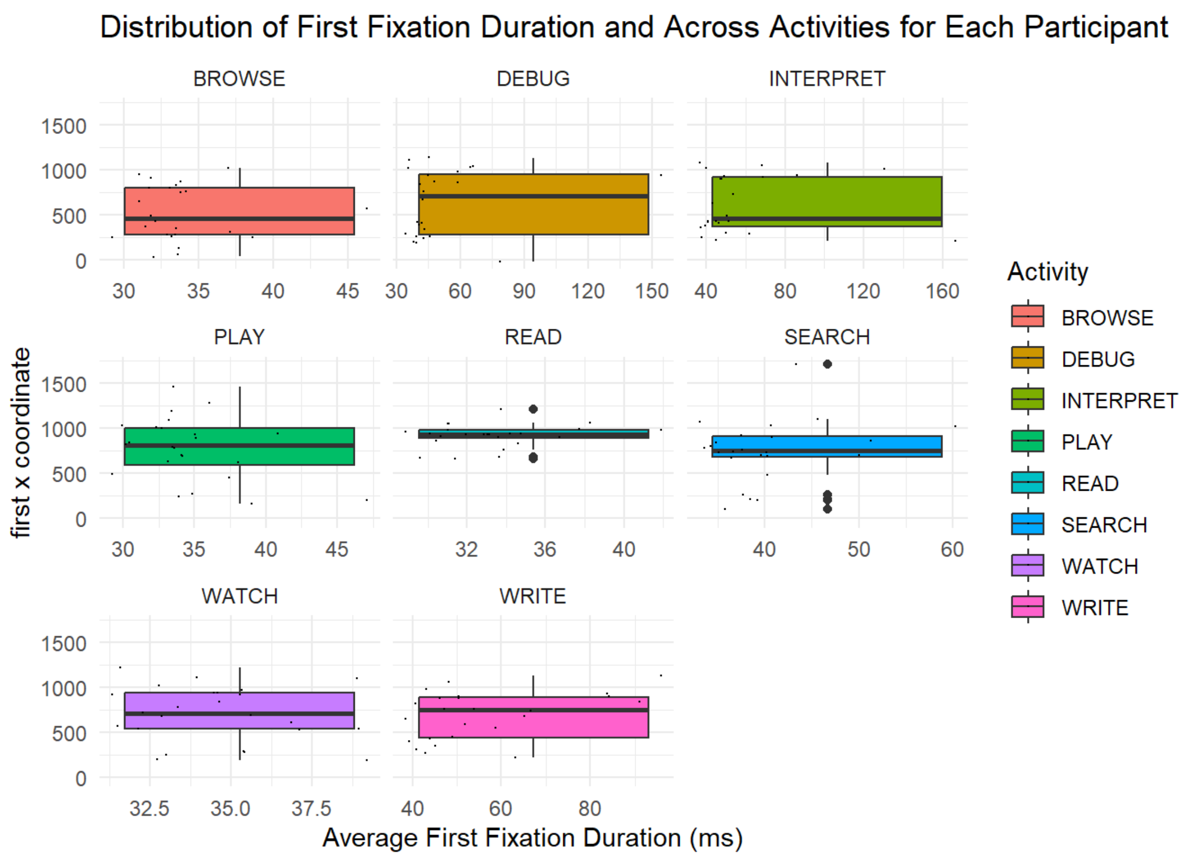
```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```



Finally, in looking at patterns of eye-behaviors Its important to identify outliers. This will give a better picture of if all the participants had particular regions of interest within a given activity. The box plot show the “avg_first_fixation_duration” variable for each “participant” categorized by “activity.” Outliers, exceeding the whiskers, are depicted as individual points. The box plot and the scattered points, show the average first fixation durations are spread across various activities for each participant. The individual participant data points, represented as dots, are plotted using coordinates derived from “first_x” on the x-axis and “avg_first_fixation_duration” on the y-axis. The graph show differences in average time spent per activity, and highlights individual differences across the first landing position (x) and average time spent, as noted by the outliers.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(participant, activity, timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    avg_duration = mean(timestamp - first(timestamp)),
    first_x = first(x),  # Take the first x-coordinate
    avg_first_fixation_duration = mean(c(0, diff(timestamp)))  # Calculate average duration of the first fix
ation
  ) %>%
  ggplot(aes(x = avg_first_fixation_duration, y = first_x , fill = activity)) +
  geom_boxplot() +
  geom_point(position = position_jitter(height = 1),
             size = .02,
             alpha = 1) +
  labs(title = "Distribution of First Fixation Duration and Across Activities for Each Participant",
       x = "Average First Fixation Duration (ms)",
       y = "first x coordinate",
       fill = "Activity") +
  theme_minimal()+
    facet_wrap(~ activity,scales = 'free_x')
```

```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```



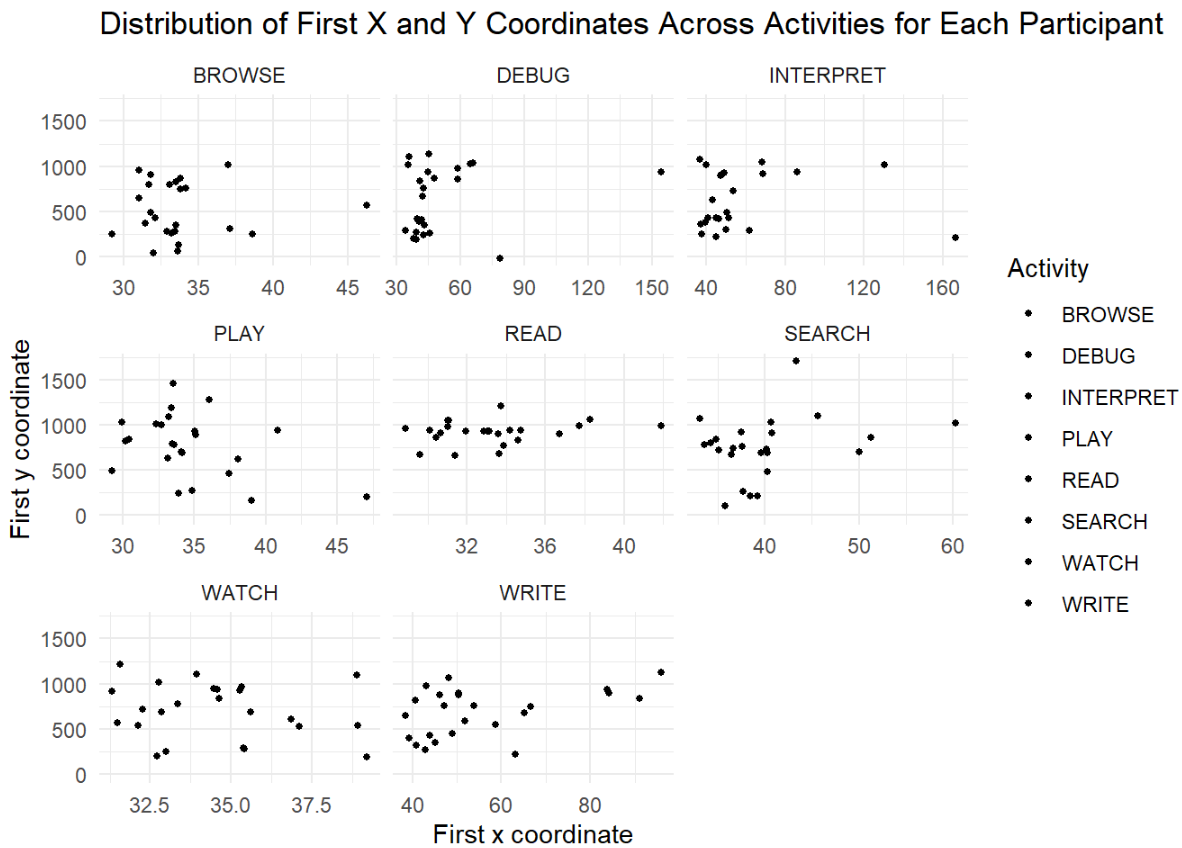

```
  scale_x_continuous(labels = scales::number_format(scale = 2))
```

```
## <ScaleContinuousPosition>
##  Range:
##  Limits:    0 --    1
```

To translate the research question: Do all participants stare in relatively the same positions across the activities? The easy solution is to estimate location by plotting the x, or y coordinates since eye-movements are symmetrical, for each participant across the average first fixation time stamp for each activity.Only the row with the minimum timestamp (i.e., the first fixation) is provided below by each activity. The preliminary results show, the first fixation across desktop activities is different but most participants look in the same area dependent on the activity. The plot below summarizes data related to participants' activities, then creates a boxplot, showing the distribution of the first x-coordinate across activities for each participant and the average duration of the first fixation. The plot is faceted by activity for comparison.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(participant, activity, timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    avg_duration = mean(timestamp - first(timestamp)),
    first_x = first(x),  # Take the first x-coordinate
    avg_first_fixation_duration = mean(c(0, diff(timestamp)))
  ) %>%
  ggplot(aes(x = avg_first_fixation_duration,
             y = first_x,
             fill = activity)) +
  geom_point(position = position_jitter(height = 2),
             size = 1,
             alpha = 2) +
  labs(
    title = "Distribution of First X and Y Coordinates Across Activities for Each Participant",
    x = "First x coordinate",
    y = "First y coordinate",
    fill = "Activity"  # Use fill instead of color for the legend title
  ) +
  theme_minimal() +
  facet_wrap(~ activity, scales = 'free_x')
```

```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```

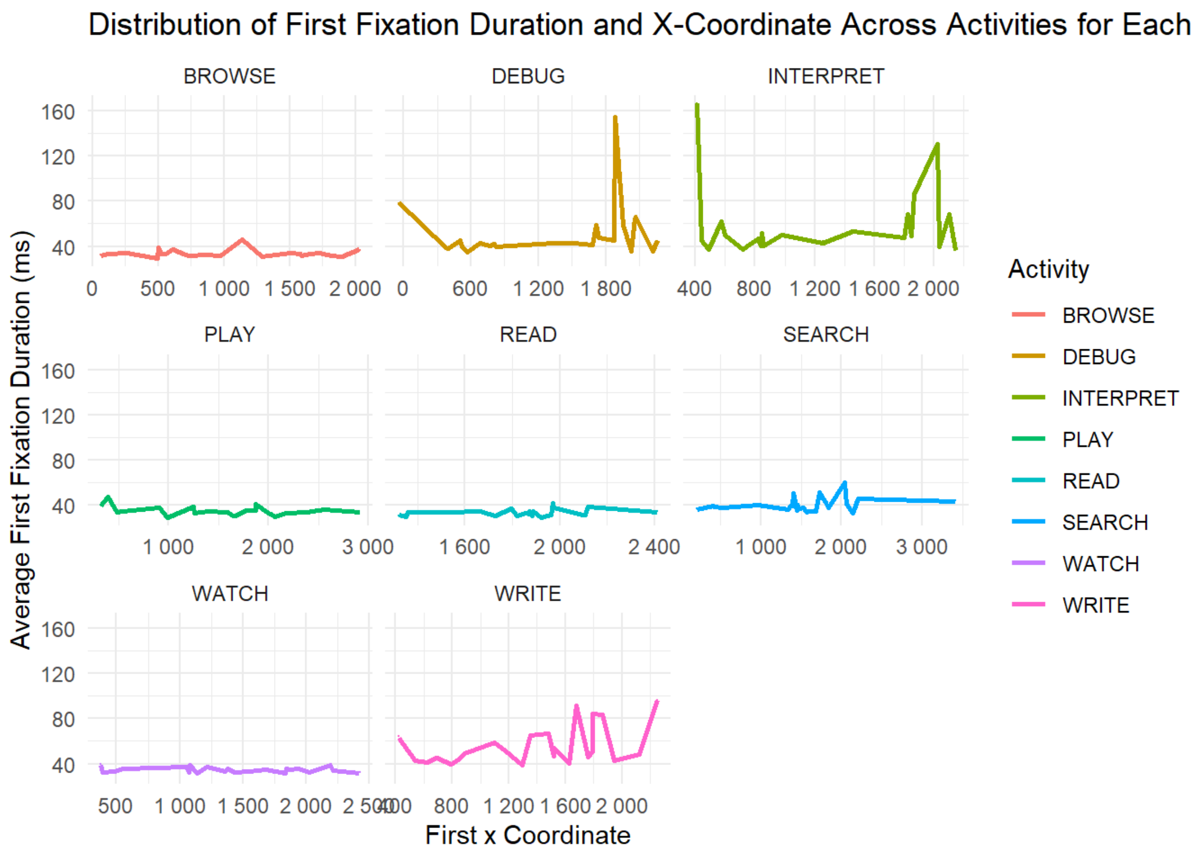


The graph below further demonstrates the relationships between the first x coordinates and average of first fixation duration on the y-axis. It plots the participants average first fixation time on the x-axis and utilizes facets to categorize the data points based on different activities.
There are distinct variations in the participants' duration gaze patterns across different activities and first time fixation across the stimulus area.

```
combined_data %>%
  filter(!is.na(activity)) %>%
  arrange(participant, activity, timestamp) %>%
  group_by(activity, participant) %>%
  summarise(
    avg_duration = mean(timestamp - first(timestamp)),
    first_x = first(x),  # Take the first x-coordinate
    min_y = first(y),
    avg_first_fixation_duration = mean(c(0, diff(timestamp)))  # Calculate average duration of the first fix
ation
  ) %>%
  ggplot(aes(x = first_x, y = avg_first_fixation_duration, color = activity)) +
  geom_line(size = 1) +  # Line graph for first x coordinates
  geom_point(position = position_jitter(height = 1), size = .02, alpha = 1) +  # Add individual points with
jitter
  labs(
    title = "Distribution of First Fixation Duration and X-Coordinate Across Activities for Each Participan
t",
    x = "First x Coordinate" ,
    y = "Average First Fixation Duration (ms)",
    color = "Activity"
  ) +
  theme_minimal() +
  scale_x_continuous(labels = scales::number_format(scale = 2)) +
  facet_wrap(~ activity, scales = 'free_x')
```

```
## `summarise()` has grouped output by 'activity'. You can override using the
## `.groups` argument.
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## ℹ Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



Limitations

Due to the substantial size of the data set per individual, particularly with eye-metrics generating extensive data, I anticipated challenges in the analysis. Determining which instance of "first time to fixation" to utilize proved difficult. Typically, eye tracking studies rely on the averages of first time to fixation for analysis so that is what I included. The complexity of the data aligned with my expectations and it was a challenging project. However, some notable differences in location of first fixation and time across activity were observed so the data fit the research question and I was able to make some conclusions.
In shaping the trajectory of future research, the focus should extend beyond the averages of first time to fixation. The box plot revealed outliers beyond the mean, upper, and lower quartiles for each activity, indicating a need for focused investigation in individual differences. Therefore, future studies should prioritize examining and understanding these individual differences, guiding the research community towards a more nuanced and impactful understanding of individual differences in eye-metrics.

Discussion

For each specific activity, both the average time spent on the first attended item and the location on the item where participants directed their gaze showed differences. The distinct differences in the plotted data points indicate potential variations in how individual participants engage with different desktop activities. These observations offer valuable insights into cognitive processes and task engagement strategies during the course of the study.

Regarding the initial research question, *Can the application of first fixation metrics (a common eye-tracking measure) contribute to desktop activity recognition?* Initial observations of the first glance and first time of fixation reveal differences across various activities. However, the data also emphasize individual disparities in eye behaviors, suggesting a need for further research to explore individual patterns of eye behaviors and the extent to which certain eye metrics can be generalized.

Citations

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Halszka, J., & van de Weijer, J. (2011). Eye Tracking : A Comprehensive Guide to Methods and Measures. Oxford University Press. http://ukcatalogue.oup.com/product/9780199697083.do (http://ukcatalogue.oup.com/product/9780199697083.do)

Srivastava, N., Newn, J., & Velloso, E. (2018). Combining Low and Mid-Level Gaze Features for Desktop Activity Recognition. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2(4), 189.