Block 4: Weeks 1-3 Reaction Time Report

1 Introduction

The data available at https://uoepsy.github.io/data/dapr1_reaction_time_2526.csv comprises of measurements of reaction times (RT) across two types of conditions - visual and auditory - from a sample of 179 Psychology 1A students at the University of Edinburgh. Data was collected across morning (AM) and afternoon (PM) lab sessions (`lab_session`) that took place on Wednesdays, Thursdays, and Fridays. Participants worked in pairs, and switched roles after one person had completed both conditions (`participant_order`), and noted which condition was completed first (`first_condition`). The participant was tasked with catching a falling ruler as quickly as possible across 5 trials in the visual condition (`visualcue_attempt1` - `visualcue_attempt5`) and 5 trials in the auditory condition (`auditorycue_attempt1` - `auditorycue_attempt5`). Measurements were originally recorded in distance (cm).

1.1 Research Questions

- RQ1: Did the average reaction time to catch a falling ruler in the visual condition significantly differ from 0.16 seconds?

- RQ2: Did the average reaction time to catch a falling ruler in the visual condition significantly differ between morning (AM) and afternoon (PM) labs?
- RQ3: Did the average reaction time to catch a falling ruler significantly differ between the visual and auditory cue conditions?

2 Analysis

First we transformed our distance into reaction time (t; measured in seconds) for each of the 5 trials in both the visual and auditory conditions. To do so, we applied the following conversion (where y is the distance the ruler fell (in cm), and g is the acceleration caused by gravity (which is known to be 980 cm/s^2)):

$$y = \frac{1}{2}gt^2$$

and so

$$t = \sqrt{\frac{2y}{g}}$$

We only retained cases where participants had average RTs for both the visual and auditory conditions. This resulted in a final sample size of 176.

2.1 Research Question 1

To investigate whether the average reaction time (RT) of those in the visual cue condition significantly differed from 0.16, we performed a one sample t-test (two-sided). The sample of Psych 1A students had a slower RT ($M = 0.17, SD = 0.02$) than the population mean ($M = 0.16$). The evidence suggested that this difference was statistically significant, $t(175) = 4.40, p < .001, two - sided$. The size of the effect was found to be small-medium $D = 0.33[0.18, 0.48]$.

The sample data did not raise any concerns of violations of independence (this can be assumed based on study design). There were no concerns regarding the normality of the sampling distribution of the mean. The average RTs in the Visual Cue condition appeared to follow a normal distribution based on our histogram and density plots (see Figure 1). Since there were slight deviations from normality observed in the QQPlot (see Figure 1), we statistically assessed the assumption of normality using a Shapiro-Wilk test. The sample data did not provide sufficient evidence to reject the null hypothesis that in the population, the data followed a normal distribution ($W = 0.99, p = .514$). Alongside these visual and statistical assessments of normality, the sample size was sufficiently large ($n = 176$), there were no concerns regarding skew (< 1), and there appeared to be no outliers, hence we concluded that the normality assumption had been met.

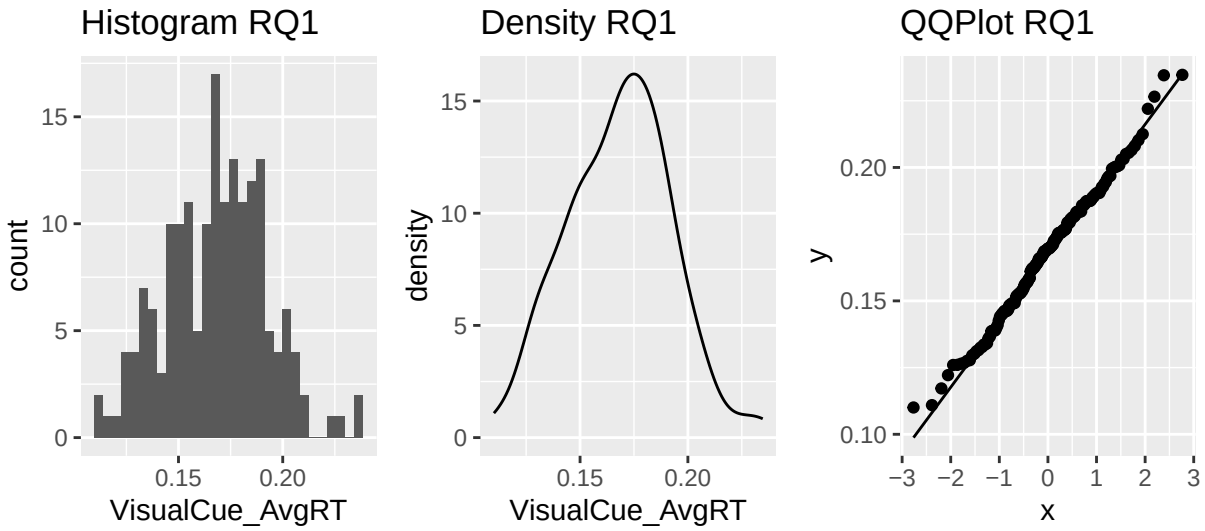

Figure 1: Distribution of Visual Cue Mean RTs

2.2 Research Question 2

To investigate whether the average visual cue RT for those in morning (AM) labs ($n = 116$) significantly differed from those in afternoon (PM) labs ($n = 60$), we performed a two-sided independent samples t-test. There was no significant difference in average reaction times between AM ($M = 0.17, SD = 0.02$), and PM ($M = 0.16, SD = 0.02$) lab groups ($t(174) = 1.89, p = .060, two - tailed$). Therefore, we failed to reject the null hypothesis.

The sample data did not raise any concerns of violations of independence (this can be assumed based on study design, since students could only attend one lab session). Although the histogram and density plots raised no concerns in either group, there were some concerns regarding the normality of the sampling distribution of the mean based on the QQplot (see Figure 2). We statistically assessed the assumption of normality with a Shapiro Wilks test. In both the AM ($W = 0.99, p = .798$) and PM ($W = 0.98, p = .392$) groups, the sample data did not provide sufficient evidence to reject the null hypothesis that in the population, the data follow a normal distribution. Furthermore, the sample size was sufficiently large in both groups ($n_{AM} = 60$;

$n_{PM} = 116$), there were no concerns regarding skew in either group (both < 1), and there did not appear to be any severe outliers. Therefore, we concluded that we had met the normality of the sampling distribution of the difference in means assumption. Based on the results of our F-test, there was no significant difference between the two population variances ($F(115, 59) = 0.89, p = .573$), and so the assumption of homogeneity of variances was met.

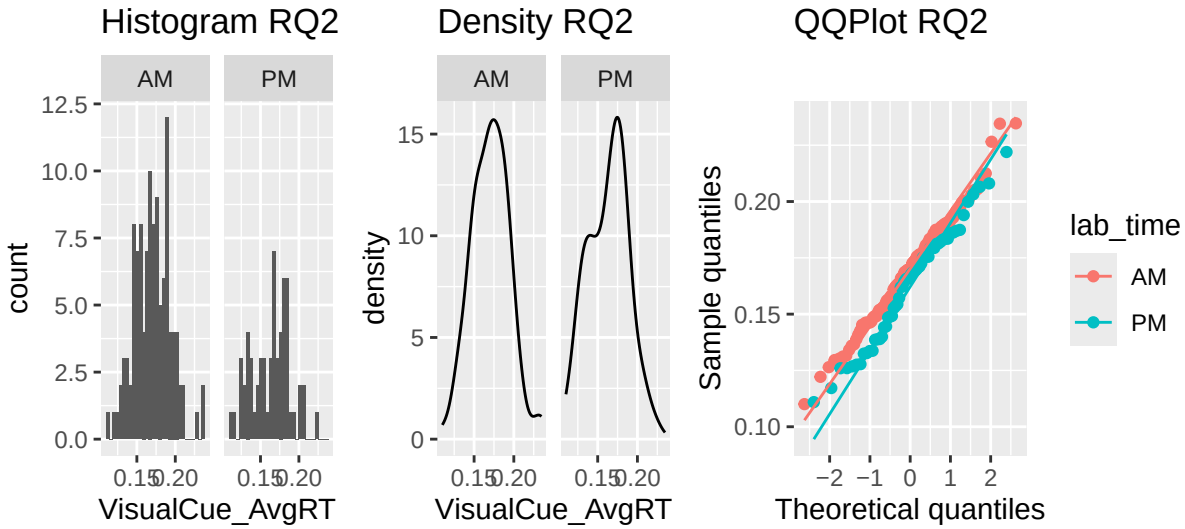

Figure 2: Distribution of Visual Cue Mean RTs by Lab Time

2.3 Research Question 3

To investigate whether the average RT in the visual cue condition significantly differed from that in the auditory cue condition, we performed a paired samples t-test. There was no significant difference in the average RT in the auditory cue condition ($M = 0.17, SD = 0.04$) and visual cue condition ($M = 0.17, SD = 0.02$), where $t(175) = -0.12, p = .907, two - tailed$. Therefore, we failed to reject the null hypothesis.

The data satisfied the independence assumptions, as this could be assumed on the study design (all participants completed both cue conditions within their one lab session). Furthermore, as noted above, we only retained data from participants who had data available for both conditions, and so our dataset comprised of matched pairs of continuous measures of RTs.

Based on the QQplot, there were some concerns regarding the normality of the sampling distribution of the difference in means (see Figure 3). We statistically assessed the assumption of normality using a Shapiro Wilk test. Based on this test result ($W = 0.96, p < .001$), we had sufficient evidence to reject the null hypothesis that in the population the difference scores followed a normal distribution. However, given that the number of matched pairs was sufficiently large ($n = 176$), and that there were no concerns regarding skew (<1) or extreme outliers (see Figure 3), normality of the sampling distribution of the mean difference could be assumed from the central limit theorem.

3 Discussion

In relation to our research questions, we concluded that: (1) the average reaction time to catch a falling ruler in the visual condition significantly differed from 0.16 seconds; (2) there was no significant difference in average reaction time to catch a falling ruler in the visual condition between morning (AM) and afternoon

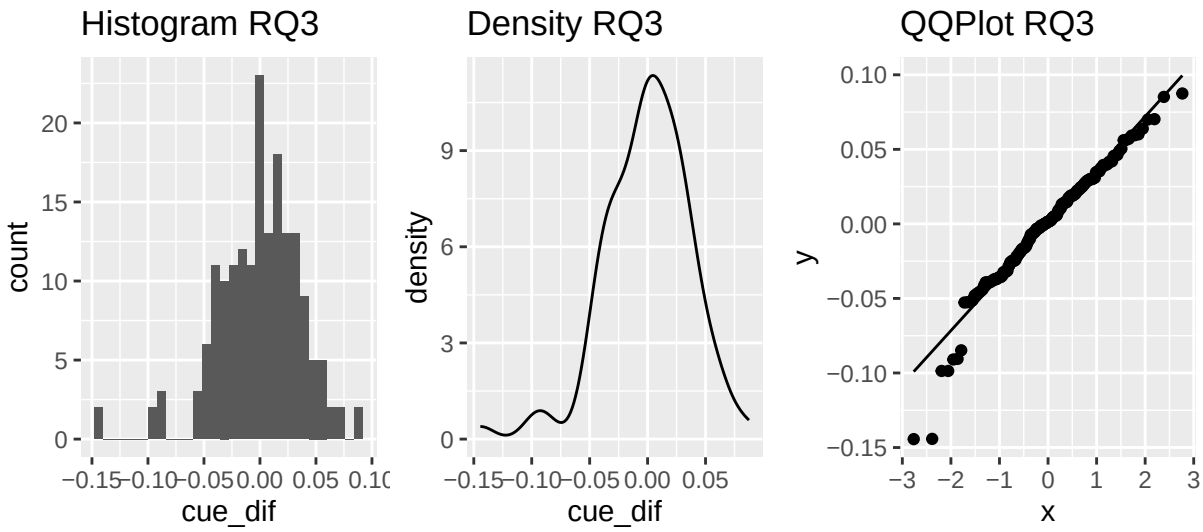

Figure 3: Distribution of the Mean Difference Scores between Visual and Audotory Conditions

(PM) labs; and (3) the average reaction time to catch a falling ruler did not significantly differ between the visual and auditory cue conditions.

4 Appendix

```
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE)
######## LOAD LIBRARIES & DATA ########

#load packages
library(tidyverse)
library(patchwork)
library(kableExtra)
library(psych)
library(effectsize)

#read in data
rt_data <- read_csv("https://uoepsy.github.io/data/dapr1_reaction_time_2526.csv")

############################ INSPECT, CLEAN, TIDY, & MANAGE DATA ############################


###### STEP 1: INSPECT AND CLEAN ######

# examine dataset
head(rt_data)
str(rt_data) # data that should be factors/numeric not currently coded as so
summary(rt_data) # we do not have values less than 0cm or over 30cm (as expected)

#make numeric / factors
rt_data <- rt_data |>
```

```
    mutate(age = as.numeric(age),
           sex = as.factor(sex),
           handedness = as.factor(handedness),
           participation_order = as.factor(participation_order),
           lab_session = as.factor(lab_session),
           first_condition = as.factor(first_condition),
           visualcue_attempt1 = as.numeric(visualcue_attempt1),
           visualcue_attempt2 = as.numeric(visualcue_attempt2),
           visualcue_attempt3 = as.numeric(visualcue_attempt3),
           visualcue_attempt4 = as.numeric(visualcue_attempt4),
           visualcue_attempt5 = as.numeric(visualcue_attempt5),
           auditorycue_attempt1 = as.numeric(auditorycue_attempt1),
           auditorycue_attempt2 = as.numeric(auditorycue_attempt2),
           auditorycue_attempt3 = as.numeric(auditorycue_attempt3),
           auditorycue_attempt4 = as.numeric(auditorycue_attempt4),
           auditorycue_attempt5 = as.numeric(auditorycue_attempt5)
           )

###### STEP 2: DISTANCE TO TIME CONVERSION (ON PSYCH 1A LAB SHEET) ######

#function to convert distance to time
rt_data <- rt_data |>
  mutate(across(visualcue_attempt1:auditorycue_attempt5, .fns = function(y) sqrt(2 * y / 980)) )

###### STEP 3: CREATE AVERAGE SCORES (Q ON PSYCH 1A LAB SHEET) AND SUBSET / FILTER ######

#create average scores
rt_data$VisualCue_AvgRT <- rowMeans(rt_data[ , c(8:12)], na.rm = TRUE)
rt_data$AuditoryCue_AvgRT <- rowMeans(rt_data[ , c(13:17)], na.rm = TRUE)

#subset
rt_data2 <- rt_data[ , c(1:7, 18:19)]

#retain only complete cases
rt_data2 <- rt_data2 |>
  filter(complete.cases(VisualCue_AvgRT, AuditoryCue_AvgRT))

############################ WEEK 1 ################################


###### STEP 1: DESCRIBE & VIZ ######

#descriptives
rt_viz_one <- describe(rt_data2$VisualCue_AvgRT)
rt_viz_one

###### STEP 2: CONDUCT ANALYSIS ######

# run our one sample t-test
t.test(rt_data2$VisualCue_AvgRT, mu = 0.16, alternative = "two.sided")

#effect size
cohens_d(rt_data2$VisualCue_AvgRT, mu = 0.16, alternative = "two.sided")
```

```
###### STEP 3: CHECK ASSUMPTIONS ######

###Normality

####Descriptives

#skew
rt_data2 |>
  summarise(
    viz_skew_one = round(skew(VisualCue_AvgRT), 2)
  )

####Visual Inspection

#hist
h_viz_one <- ggplot(rt_data2, aes(x = VisualCue_AvgRT)) +
  geom_histogram() +
  labs(title = "Histogram RQ1")
h_viz_one

#density
d_viz_one <- ggplot(rt_data2, aes(x = VisualCue_AvgRT)) +
  geom_density() +
  labs(title = "Density RQ1")
d_viz_one

#qq-plot
q_viz_one <- ggplot(rt_data2, aes(sample = VisualCue_AvgRT)) +
  geom_qq() +
  geom_qq_line() +
  labs(title = "QQPlot RQ1")
q_viz_one

####Testing

#shapiro-wilk

shapiro.test(rt_data2$VisualCue_AvgRT)

#################### WEEK 2 ####################################

###### STEP 1: INSPECT DATA ######

#check levels of lab_session - currently have day and time
levels(rt_data2$lab_session)

## we need to create a new variable of just lab_time
rt_data2 <- rt_data2 |>
  mutate(lab_time = factor(lab_session,
                           levels = c("Friday: 10am-12pm", "Thursday: 11am-1pm", "Wednesday: 10am-12pm"),
                           labels = c("AM", "PM", "AM")))

table(rt_data2$lab_time)
```

```r
###### STEP 2: DESCRIBE & VIZ ######

#boxplot
p2 <- ggplot(data = rt_data2, aes(x = lab_time, y = VisualCue_AvgRT, fill = lab_time)) +
  geom_boxplot() +
  geom_jitter(width = 0.1) #just demo 0.1 and 0.5 to link to yesterday then remove
p2

#create a table of descriptives, grouped by lab time
descr2 <- rt_data2 |>
  group_by(lab_time) |>
  summarise(
    M = mean(VisualCue_AvgRT),
    SD = sd(VisualCue_AvgRT),
    n = n()) |>
  kable(caption = "Descriptives RQ2", digits = 2) |>
  kable_styling()
descr2

###### STEP 3: CONDUCT ANALYSIS ######

##Statistical test
#run our independent samples t-test
res2 <- t.test(rt_data2$VisualCue_AvgRT ~ rt_data2$lab_time,
               alternative = "two.sided",
               mu = 0,
               var.equal = TRUE,
               conf.level = 0.95)
res2

### STEP 3.1: SIDE STEP - WHAT IF OUR DATA WAS IN A DIFFERENT FORMAT? #####

#what if our data were in wide format?

##pivot long data to make wide
data_wide <- rt_data2 |>
  pivot_wider(names_from = "lab_time", values_from = "VisualCue_AvgRT")


#how we set up t.test with wide data is slightly different
#wide data t.test
res_wide <- t.test(data_wide$AM,data_wide$PM,
               alternative = "two.sided",
               mu = 0,
               var.equal = TRUE,
               conf.level = 0.95)
res_wide

#IMPORTANT: We get exactly the same values regardless of wide or long data#

### Effect Size
#not calculated since we have a non-sig difference
```

```r
###### STEP 4: CHECK ASSUMPTIONS ######

###Normality

####Descriptives

#skew
rt_skew_indep <- rt_data2 |>
  group_by(lab_time) |>
  summarise(
    viz_skew = round(skew(VisualCue_AvgRT), 2)
  )
rt_skew_indep

####Visual Inspection

#hist
h_viz_indep <- ggplot(data = rt_data2, aes(x = VisualCue_AvgRT)) +
  geom_histogram() +
  facet_wrap(~lab_time) +
  labs(title = "Histogram RQ2")
h_viz_indep

#density
d_viz_indep <- ggplot(data = rt_data2, aes(x = VisualCue_AvgRT)) +
  geom_density() +
  facet_wrap(~lab_time) +
  labs(title = "Density RQ2")
d_viz_indep

#qqplot
q_viz_indep <- ggplot(data = rt_data2, aes(sample = VisualCue_AvgRT, colour = lab_time)) +
  geom_qq() +
  geom_qq_line() +
  labs(title = "QQPlot RQ2", x = "Theoretical quantiles", y = "Sample quantiles")
q_viz_indep

####Testing

#shapiro-wilk - recall that we need to do this for each level of our group variable, so need for AM and

#AM condition
rt_data2 |>
  filter(lab_time == "AM") |>
  pull(VisualCue_AvgRT) |>
  shapiro.test()

#PM condition
rt_data2 |>
  filter(lab_time == "PM") |>
  pull(VisualCue_AvgRT) |>
  shapiro.test()
```

```r
###Homogeneity of variances

# testing whether variances are equal
var.test(rt_data2$VisualCue_AvgRT ~ rt_data2$lab_time, ratio = 1)

#################### WEEK 3 ####################################

###### STEP 1: DESCRIBE DATA ######

desc_pair <- rt_data2 |>
  summarise(
    M_Viz = mean(VisualCue_AvgRT),
    M_Aud = mean(AuditoryCue_AvgRT),
    SD_Viz = sd(VisualCue_AvgRT),
    SD_Aud = sd(AuditoryCue_AvgRT),
    n = n()
  ) |>
  kable(caption = "Descriptives RQ3", digits = 2) |>
  kable_styling()
desc_pair

###### STEP 2: CONDUCT ANALYSIS ######


##Statistical test
#run our paired samples t-test
# t.test - rt_data2 = wide data format
res3 <- t.test(rt_data2$AuditoryCue_AvgRT, rt_data2$VisualCue_AvgRT,
               paired = TRUE,
               mu = 0,
               alternative = "two.sided")
res3


### Effect Size
#not calculated since we have a non-sig difference

###### STEP 3: CHECK ASSUMPTIONS ######

#first step: add difference scores to dataset
rt_data2 <- rt_data2 |>
  mutate(cue_dif = AuditoryCue_AvgRT - VisualCue_AvgRT)

###Normality

####Descriptives

#skew
rt_skew_pair <- rt_data2 |>
  summarise(
    skew = round(skew(cue_dif), 2)
  )
rt_skew_pair
```

```
####Visual Inspection

#hist
h_viz_pair <- ggplot(data = rt_data2, aes(x = cue_dif)) +
  geom_histogram() +
  labs(title = "Histogram RQ3")
h_viz_pair

#density
d_viz_pair <- ggplot(data = rt_data2, aes(x = cue_dif)) +
  geom_density() +
  labs(title = "Density RQ3")
d_viz_pair

#qqplot
q_viz_pair <- ggplot(data = rt_data2, aes(sample = cue_dif)) +
  geom_qq() +
  geom_qq_line() +
  labs(title = "QQPlot RQ3")
q_viz_pair

####Testing

#shapiro-wilk
shapiro.test(rt_data2$cue_dif)

h_viz_one | d_viz_one | q_viz_one
h_viz_indep | d_viz_indep | q_viz_indep
h_viz_pair | d_viz_pair | q_viz_pair
```