

# Video Track Analysis for Hero Rats Odor Preference

# AGENDA

Intra-rater and Inter-rater Analysis

Two Preference Indicator Variable (Difference & Ratio)

Analysis Based on Change of Preference Difference Regarding Duration:

- Normality Check - Paired T-test
- Bootstrapping

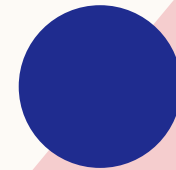
Principle Component Analysis (PCA)

Linear Regression

Decision Tree - Binary

Interesting Finding From Dataset

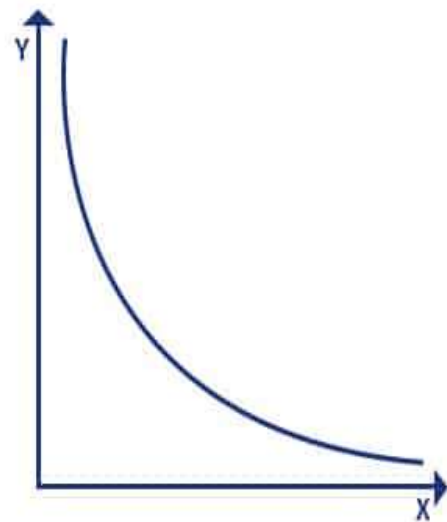
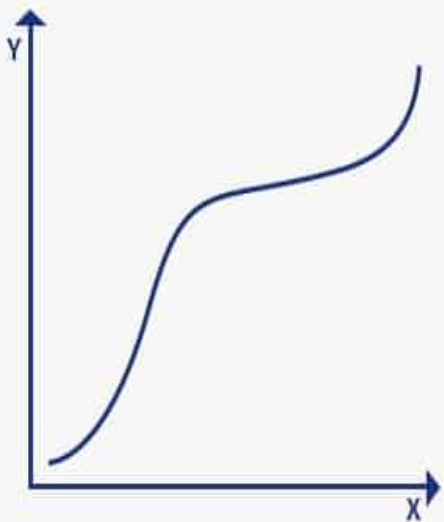
Summary



# INTRA- AND INTER-RATER ANALYSIS

## Method:

- Spearman rank correlation coefficient  
It measures the strength and direction of monotonic association between two variables
- Intraclass Correlation Coefficient (ICC)  
It measures the agreement between multiple raters rating in ordinal or continuous scales.



## Reason for choosing these method:

- Spearman rank correlation coefficient
  - Non –parametric
  - Used for continuous data and want to access the degree of linear or rank-based agreement between raters.
- Intraclass Correlation Coefficient
  - can be used to measure the agreement between multiple raters rating in ordinal or continuous scales.

# RESULTS OF POOR CORRELATION <sup>4</sup>

## Inter-rater reliability

	V1	spearman_rho	icc_value	icc_reliability
<b>rho.57</b>	R0	0.225374467927604	0.203821656050955	Poor
<b>rho.77</b>	RRP	0.364608187558621	0.433249370277078	Poor
<b>rho.82</b>	CRP	0.309035023100369	0.302788844621514	Poor
<b>rho.107</b>	R0.1	0.280691786106895	0.347826086956522	Poor
<b>rho.127</b>	RRP.1	0.346	0.405705229793978	Poor

R0: Counts of Rearing in the Middle of the Enclosure

RRP: Counts of Rearing within Right Proximal Zone

CRP: Counts of Interacting within the Right Proximal Zone

R0.1: Duration of Rearing in the Middle of the Enclosure

RRP.1: Duration of Rearing within Right Proximal Zone

# RESULTS OF POOR CORRELATION

## Intra-rater reliability

	V1	spearman_rho2	icc_value2	icc_reliability
<b>rho.37</b>	RLP	0.445454545454545	0.244186046511628	Poor
<b>rho.62</b>	C0	0.246376811594203	0.246376811594203	Poor
<b>rho.77</b>	RRP	0.217594108300831	0.341801385681293	Poor
<b>rho.87</b>	RLP.1	0.425054308472122	0.0736478711162256	Poor
<b>rho.127</b>	RRP.1	0.202978426306575	0.340807174887892	Poor
<b>rho.132</b>	CRP.1	0.569590036581684	0.418324849606664	Poor

RLP: Counts of Rearing within Left Proximal Zone

C0: Counts of Interaction with Cage within Middle of Enclosure

RRP: Counts of Rearing within Right Proximal Zone

RLP.1: Duration of Rearing within Left Proximal Zone

RRP.1: Duration of Rearing within Right Proximal Zone

CRP.1: Duration of Interacting within the Right Proximal Zone

# TWO PREFERENCE INDICATOR VARIABLES

- Difference (Univariate Analysis)
  - Difference of interaction duration between TNT area and Tea area

TEL\_duration-TER\_duration

- Ratio(PCA and Linear Regression)
  - Ratio of Proximal duration between TNT area and Tea area
  - Interaction Duration is Included in the Proximal Duration

$$\text{Ratio} = \frac{\text{Proximal duration of TNT area}}{\text{Proximal duration of Tea area}}$$

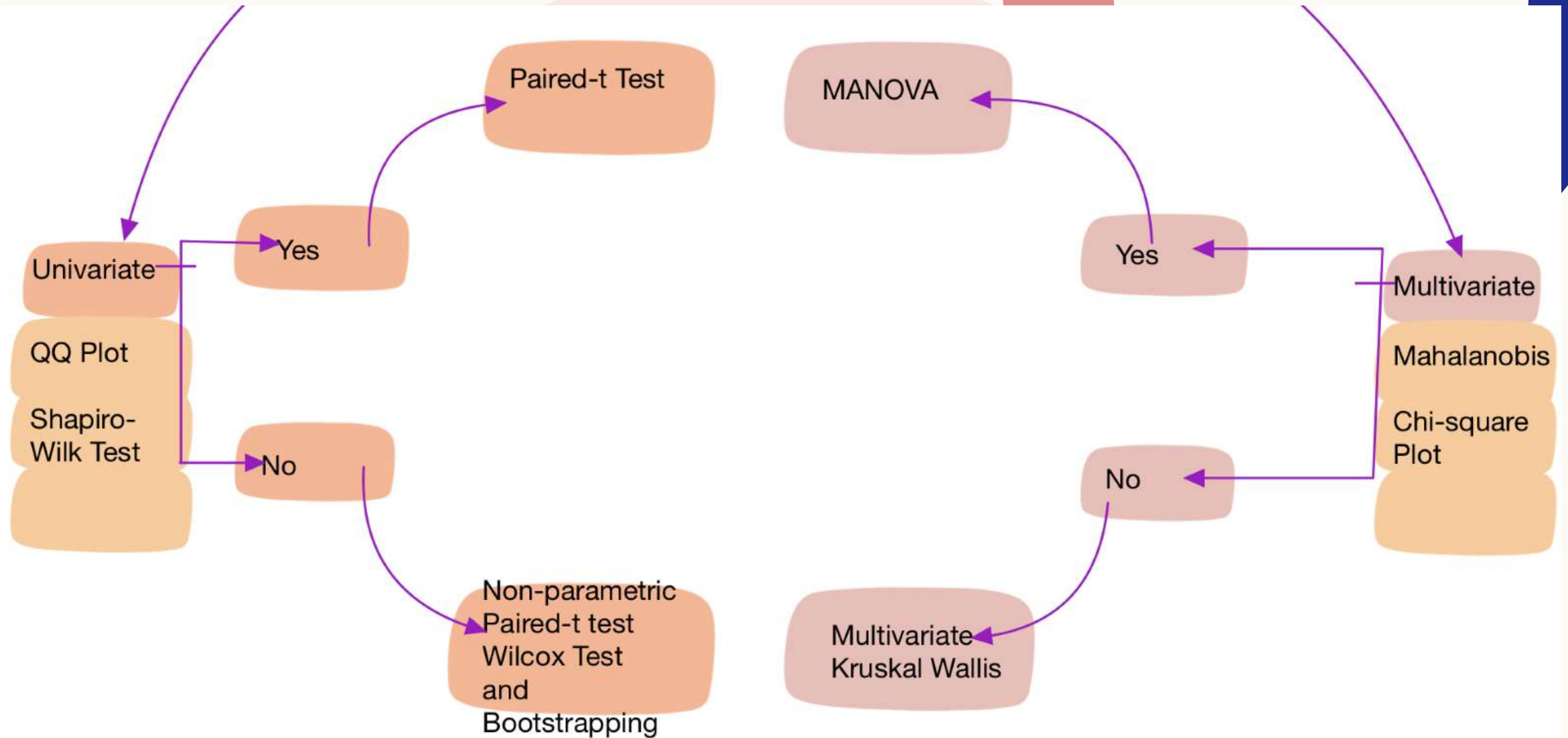
Ratio captures relative changes of duration between TNT and Tea.

e.g. TNT Tea

t1    24    10    diff = 14    ratio = 2.4

t2    12    5    diff = 7    ratio = 2.4

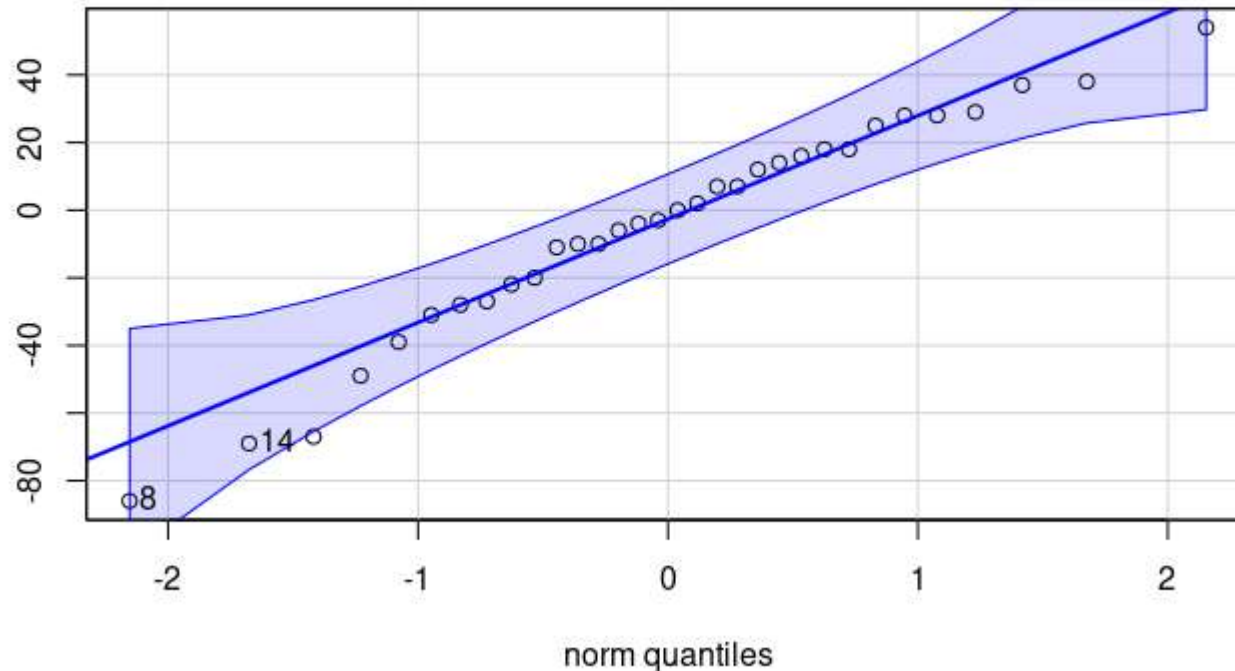
# NORMALITY CHECK



# NORMALITY CHECK FOR UNIVARIATE DATA

Variable: Difference of interaction duration between TNT area and Tea area

## Q-Q Plot



## Shapiro Wilk Test

H0: data follows a normal distribution.

H1: data does not follow a normal distribution.

#code

```
shapiro.test(time_comparison$T1.Duration.Difference  
-time_comparison$T2.Duration.Difference)
```

#R output

```
data: time_comparison$T1.Duration.Difference  
      - time_comparison$T2.Duration.Difference  
W = 0.96345, p-value = 0.3403
```

**# Result:**

Data approximately normally distributed.



# PAIRED T-TEST

Variable: Difference of interaction duration between TNT area and Tea area

- Statistical test used to determine whether there is a significant difference between the means of two related groups.
- Test the means of “TNT v.s Tea interaction duration” between group1 and group2

Group 1: Rats' ages = 10 weeks

Group 2: Rats' ages = ages = 6 months

H0: There is no difference between means of “TNT v.s. Tea interaction duration” group1 and group2

H1: There exists difference between means of “TNT v.s. Tea interaction duration” group1 and group2

## Paired t-test

```
data: time_comparison$T1.Duration.Difference and time_comparison$T2.Duration.Difference
t = -0.7986, df = 31, p-value = 0.4306
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -16.547712  7.235212
sample estimates:
mean difference
 -4.65625
```

- **Result:** There is no significant difference of TNT duration preference between t1 and t2

# BOOTSTRAPPING

Variable: Difference of interaction duration between TNT area and Tea area

## ➤ Advantages :

1. Fix the problem of limited sample size. (Our sample size is relatively small,  $n=32$ )

2. This method is distribution free.

## ➤ (Continuous) Test the means of “TNT v.s Tea interaction duration” between 2 time points

Treat as continuous variables → 95% Confidence interval  $[-19.4, 8.9]$  → “0” is in the CI

Treat as binary variables → 95% Confidence interval  $[0.25, 0.75]$  → “0.5” is in the CI

Binary Variable: The variable is set to 1 if the following condition is satisfied

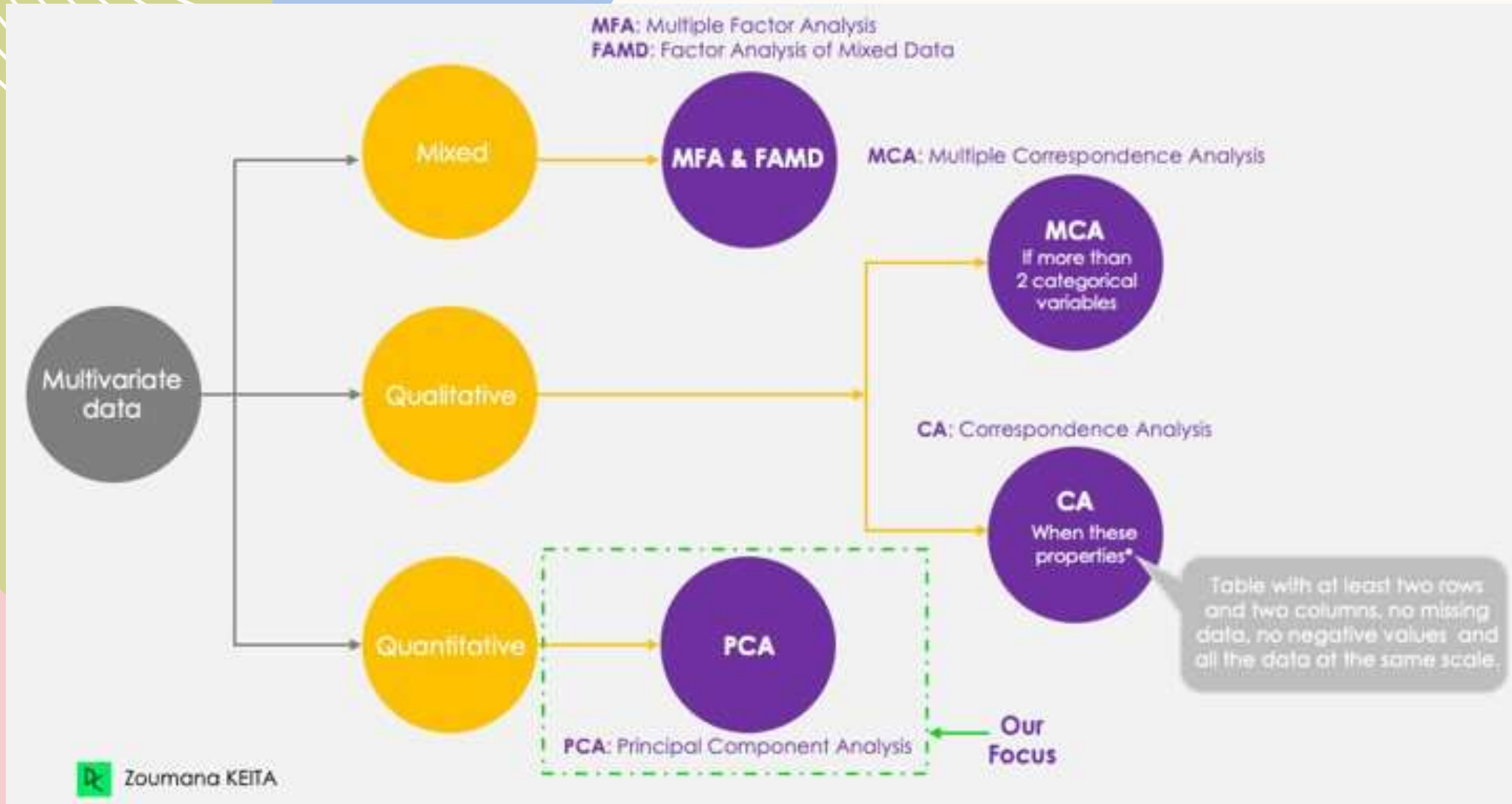
Difference of interaction duration TNT v.s. Tea at T2 > Difference of interaction duration TNT v.s. Tea at T1

## ➤ Results: The difference in interaction duration **stay the same** when age increases.

Bootstrap to PCA: Since only comparing means of the preference indicators and no change respect to time has been observed, we would like to apply linear regression to obtain a holistic view of the relationships between the preference indicators and the objective variables using PCA (linear combinations of durations)

# PRINCIPLE COMPONENT ANALYSIS (PCA)

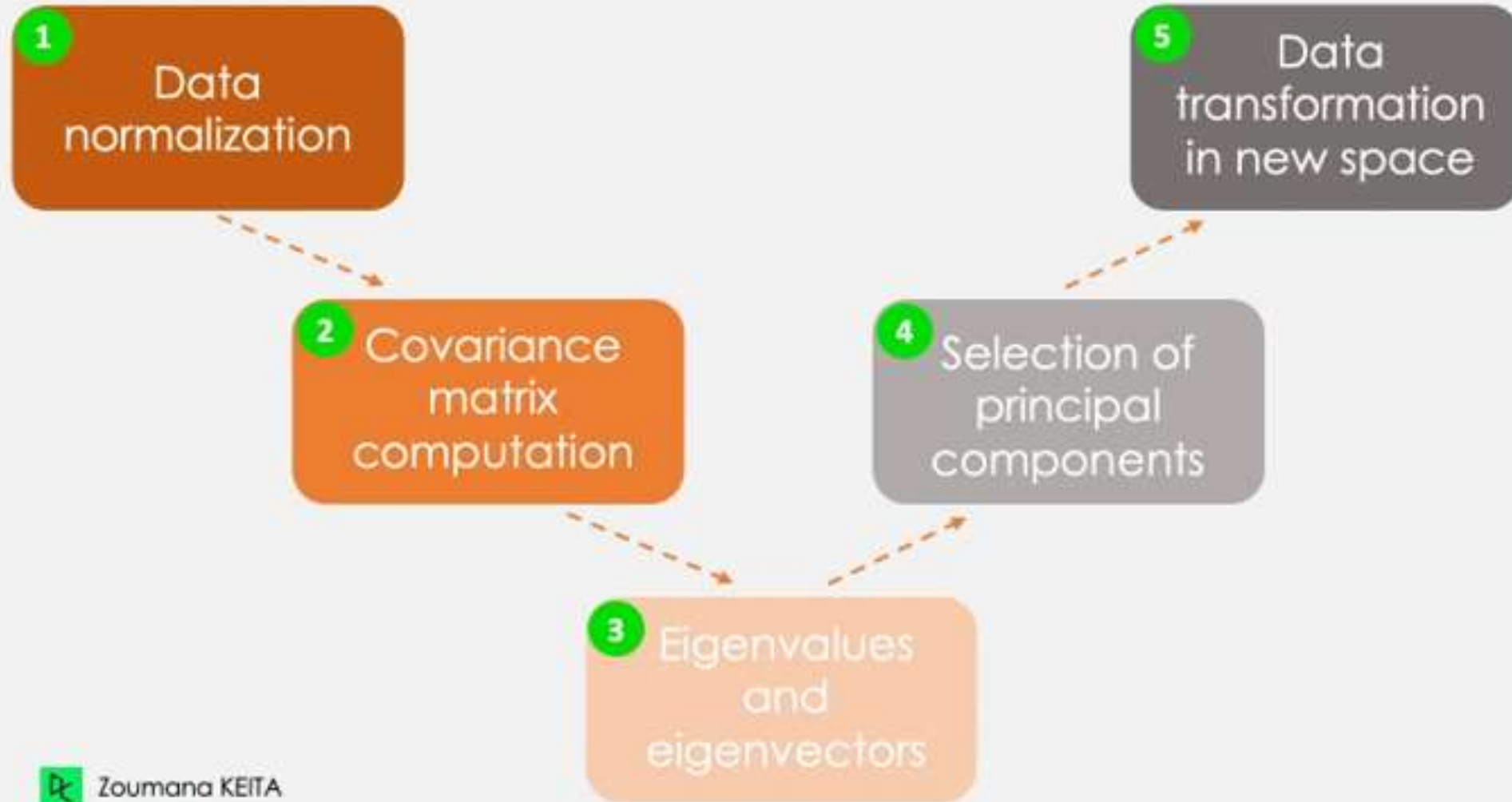
Variable: Ratio of proximal duration and visits between TNT area and Tea area



# PROCESSING PCA METHOD

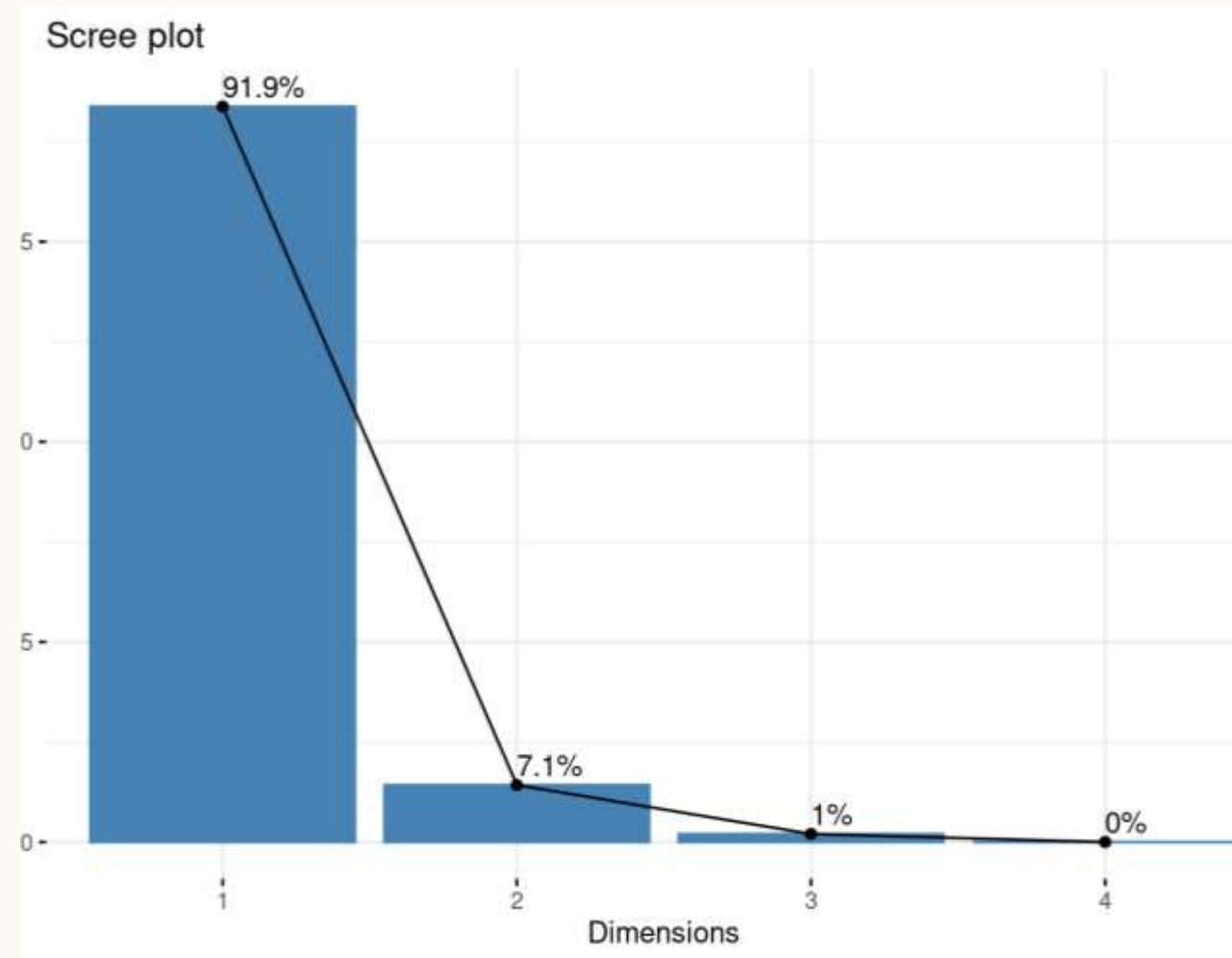
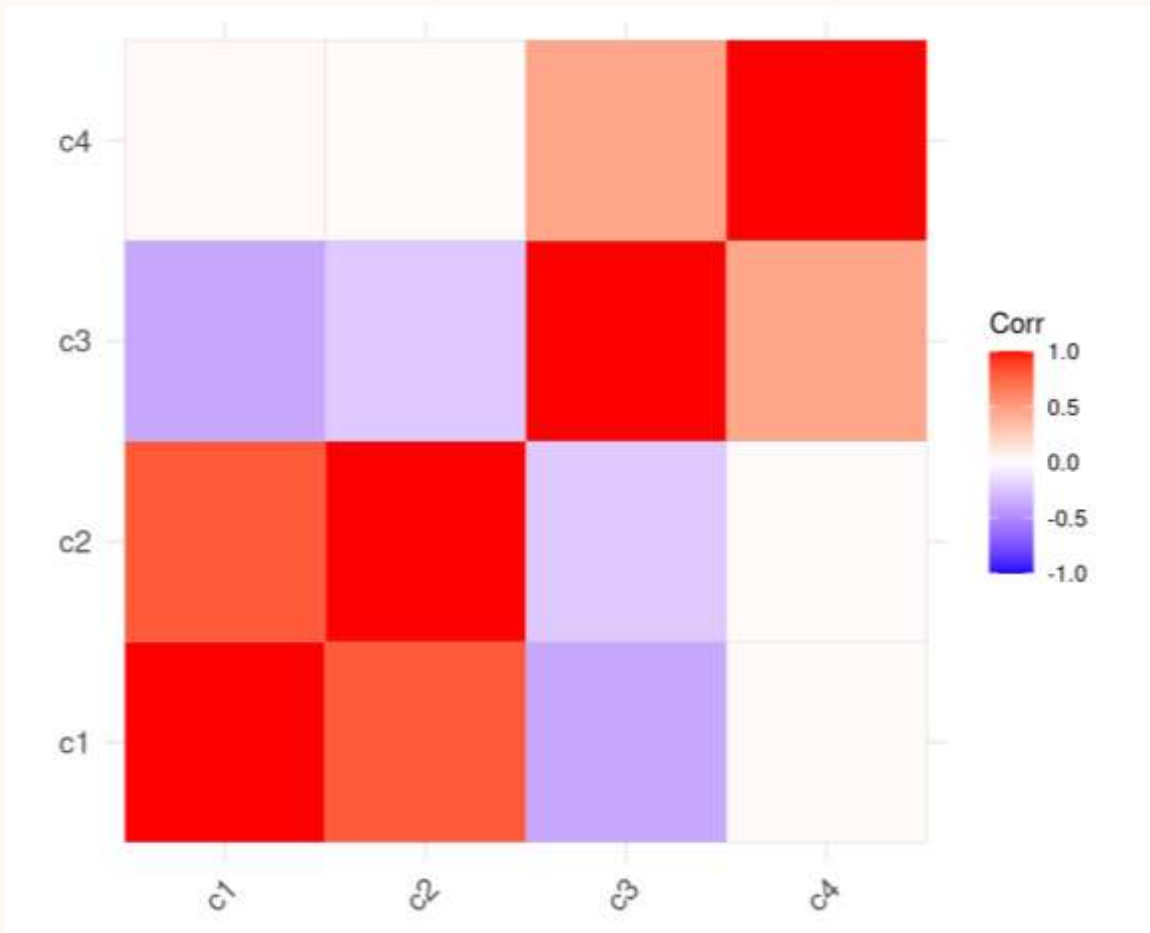
12

Variable: Ratio of proximal duration and visits between TNT area and Tea area



Zoumana KEITA

# Covariance Matrix & Scree Plot



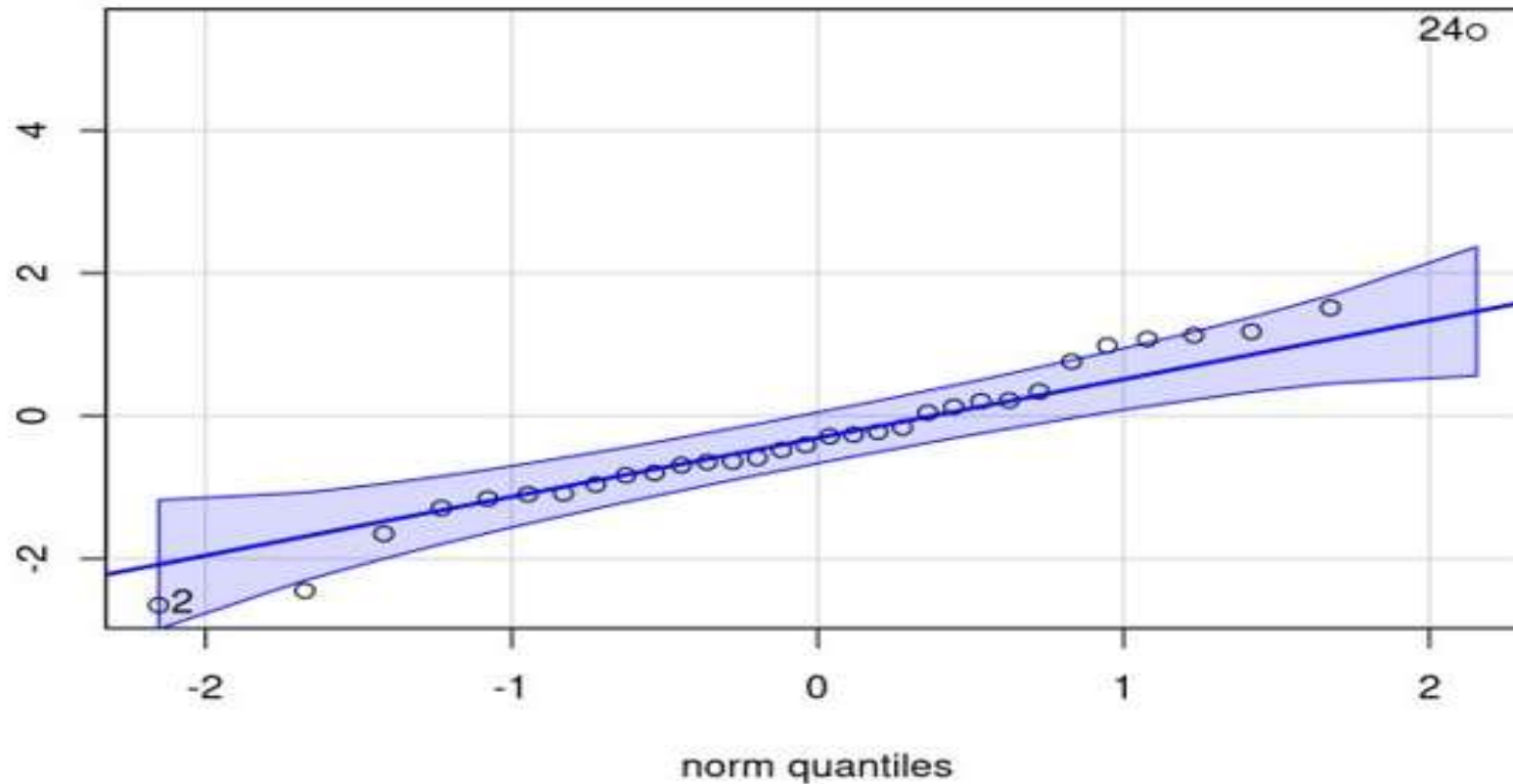
# PRINCIPAL COMPONENT NAMING

## CONTRAST OF T1 AND T2 ACTIVITY

### AND TIME

TNT Proximal Visits at T1	TNT Proximal Duration at T1	TNT Proximal Visits at T2	TNT Proximal Duration at T2
0.5661369	0.5174113	-0.5548703	-0.3223251

# NORMALITY CHECK BEFORE FITTING REGRESSION MODEL



# LINEAR REGRESSION WITH OUTLIERS EXCLUDED DATA

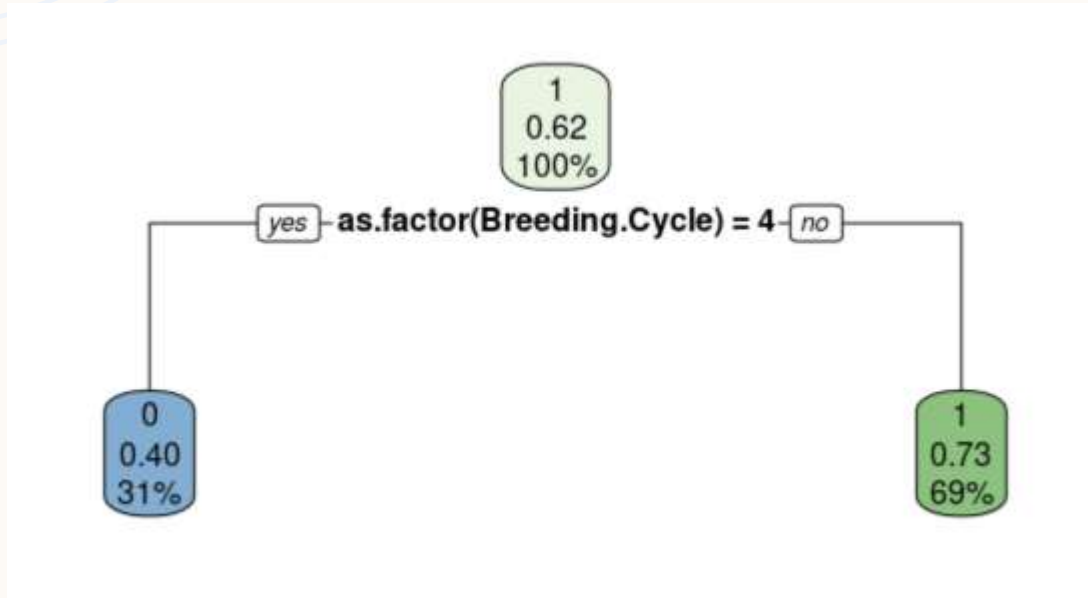
•## Coefficients:

•##		Estimate	Std. Error	t value	Pr(> t )
•##	(Intercept)	-0.27296	0.37973	-0.719	0.4798
•##	as.factor(Sex_binary)1	-0.81712	0.32994	-2.477	0.0214 *
•##	as.factor(Factor.A)2	0.44252	0.47204	0.937	0.3587
•##	as.factor(Factor.A)3	0.29486	0.56115	0.525	0.6045
•##	as.factor(Factor.B)2	0.38363	0.31733	1.209	0.2395
•##	as.factor(Breeding.Cycle)2	-0.42541	0.62412	-0.682	0.5026
•##	as.factor(Breeding.Cycle)3	0.04517	0.47581	0.095	0.9252
•##	as.factor(Breeding.Cycle)4	-0.09794	0.51282	-0.191	0.8503

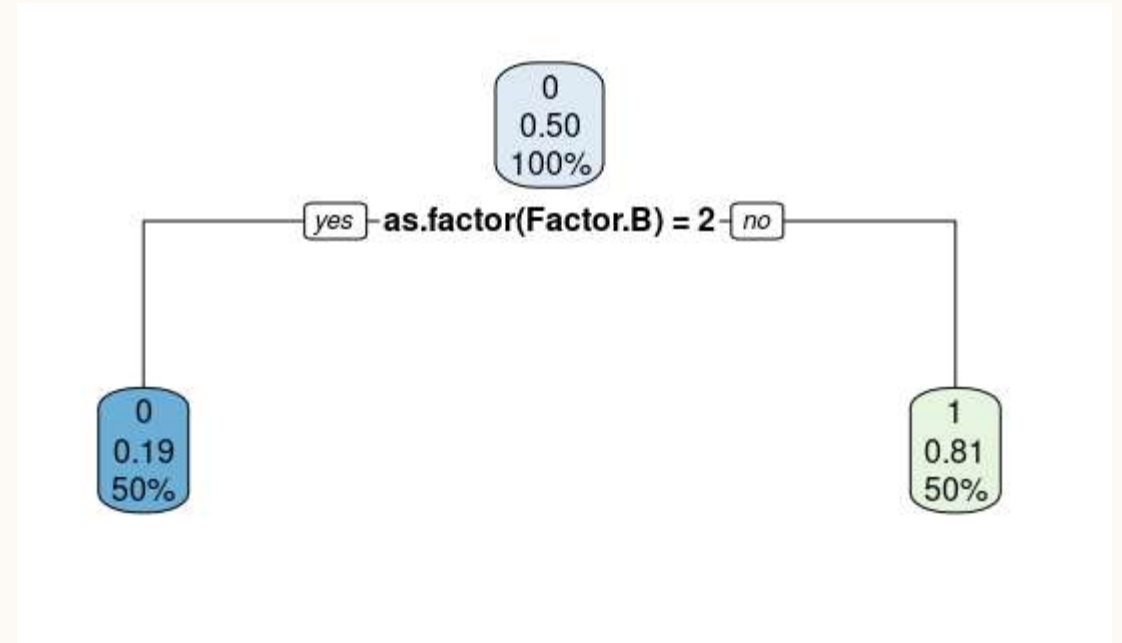


# DECISION TREE

T1



T2



# INTERESTING FINDING FROM DATASET

- **Randomization Check:**
  - Whether TNT.Position affect the first egg that rats interact with
- Look at the behavior variables with respect to time

# Randomization Check:

Whether TNT.Position affect the first egg that rats interact with.

T1

<u>TNT.Position</u>	<u>mean(abs(TNT.Position - First.tea.egg) &lt; 0.5)</u>
1	0.375
2	0.375

T2

<u>TNT.Position</u>	<u>mean(abs(TNT.Position - First.tea.egg) &lt; 0.5)</u>
1	0.6666667
2	0.7058824

Overall from t1 and t2

<u>TNT.Position</u>	Proportion of Visiting TNT First
1	0.5161290
2	0.5454545

The proportion of first interaction with TNT is quite similar

The 2 proportions at each time point are similar to each other, however, another interesting change is observed, the proportion of interacting with TNT first has increased significantly from 0.375 to 0.67. This table can be one indication of rats' growth in preference of TNT as they age.

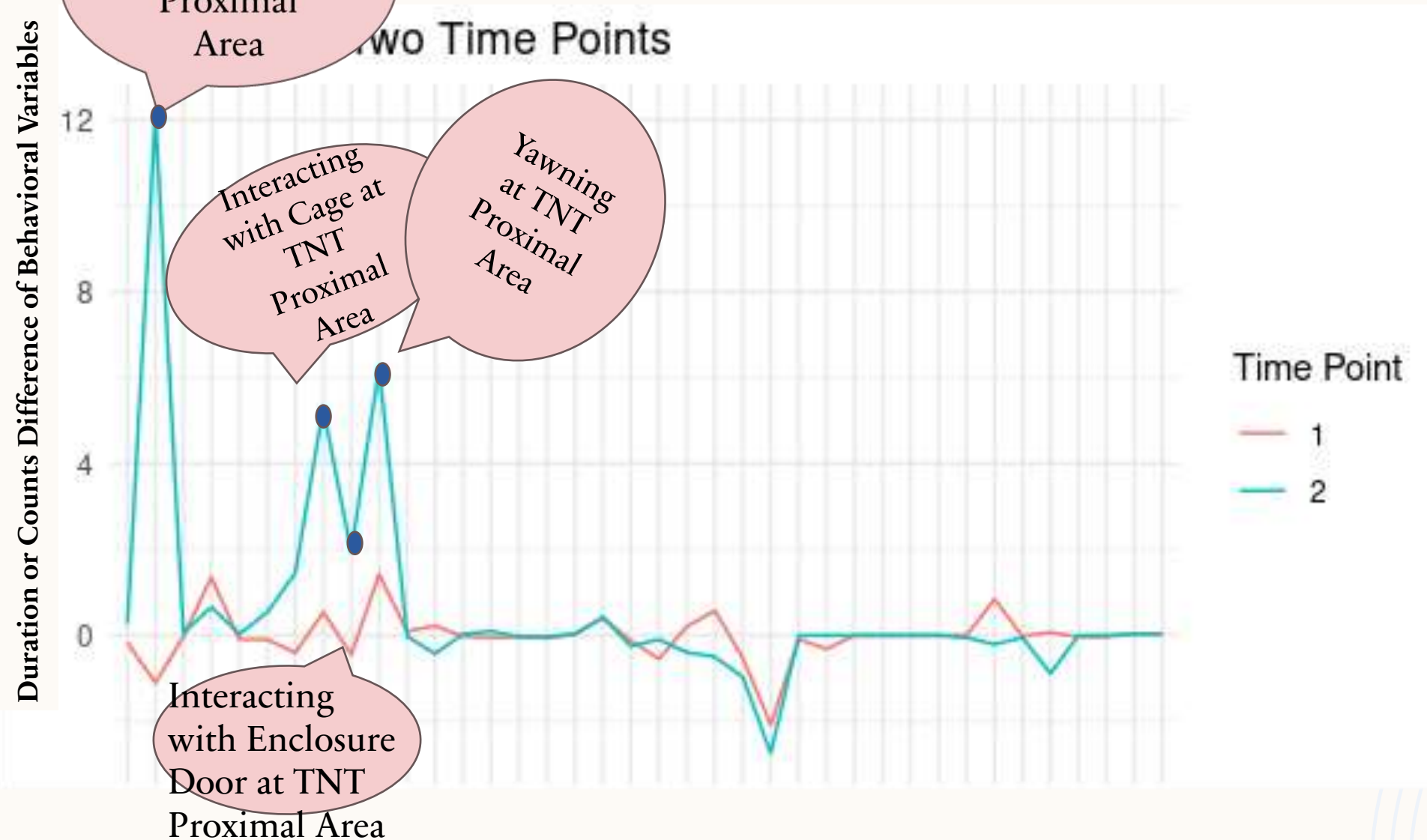
- Look at the behavior variables with respect to time

Duration or Counts Difference of Behavioral Variables

Line Plot of Two Time Points



## ➤ Look at behavior variables with respect to time



# Behavior Variables & P-value

<b>[1] GP.diff</b> 0.38443152495463	<b>[7] CP.diff</b> 0.80132073996468	<b>[13] L.diff</b> 0.42968451449557	<b>[19] Y.Diff</b> 1
<b>[2] RP.diff</b> 0.98488446492794	<b>[8] DP.diff</b> 0.41421617824252	<b>[14] IM.diff</b> 0.37109336952269	<b>[20] U.diff</b> 1
<b>[3] LP.diff</b> 0.79288013824888	<b>[9] YP.diff</b> 1	<b>[15] TER.diff</b> NaN	<b>[21] GP1.diff</b> 0.2639076418163
<b>[4] IMP.diff</b> 0.26520539259150	<b>[10] UP.diff</b> 1	<b>[16] TEL.diff</b> NaN	<b>[22] RP1.diff</b> 0.7495565067870
<b>[5] TERP.diff</b> 0.07023426683851	<b>[11] G.diff</b> 0.81912947086654	<b>[17] C.diff</b> 0.58789016205337	<b>[23] LP1.diff</b> 1
<b>[6] TELP.diff</b> 0.07023426683851	<b>[12] R.diff</b> 0.15668462599590	<b>[18] D.diff</b> 0.86919791107179	<b>[24] IMP1.diff</b> 0.4614509878333

# Continuous

[25] TERP1.diff 0.7223683123262	[29] YP1.diff 1	[33] L1.diff 0.56305732839422	[37] C1.diff 0.59976223745737
[26] TELP1.diff 0.71695146422861	[30] UP1.diff 1	[34] IM1.diff 0.37109336952269	[38] D1.diff 0.48308574282474
[27] CP1.diff 0.74103301732077	[31] G1.diff 0.98860203681852	[35] TER1.diff NaN	[39] Y1.diff 1
[28] DP1.diff 0.42267807417063	[32] R1.diff 0.14259034476783	[36] TEL1.diff NaN	[40] U1.diff 1

None of the p-values is smaller than 0.05, conclude that there is no significant change of behaviors between time.

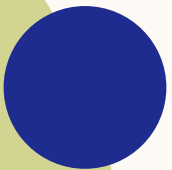
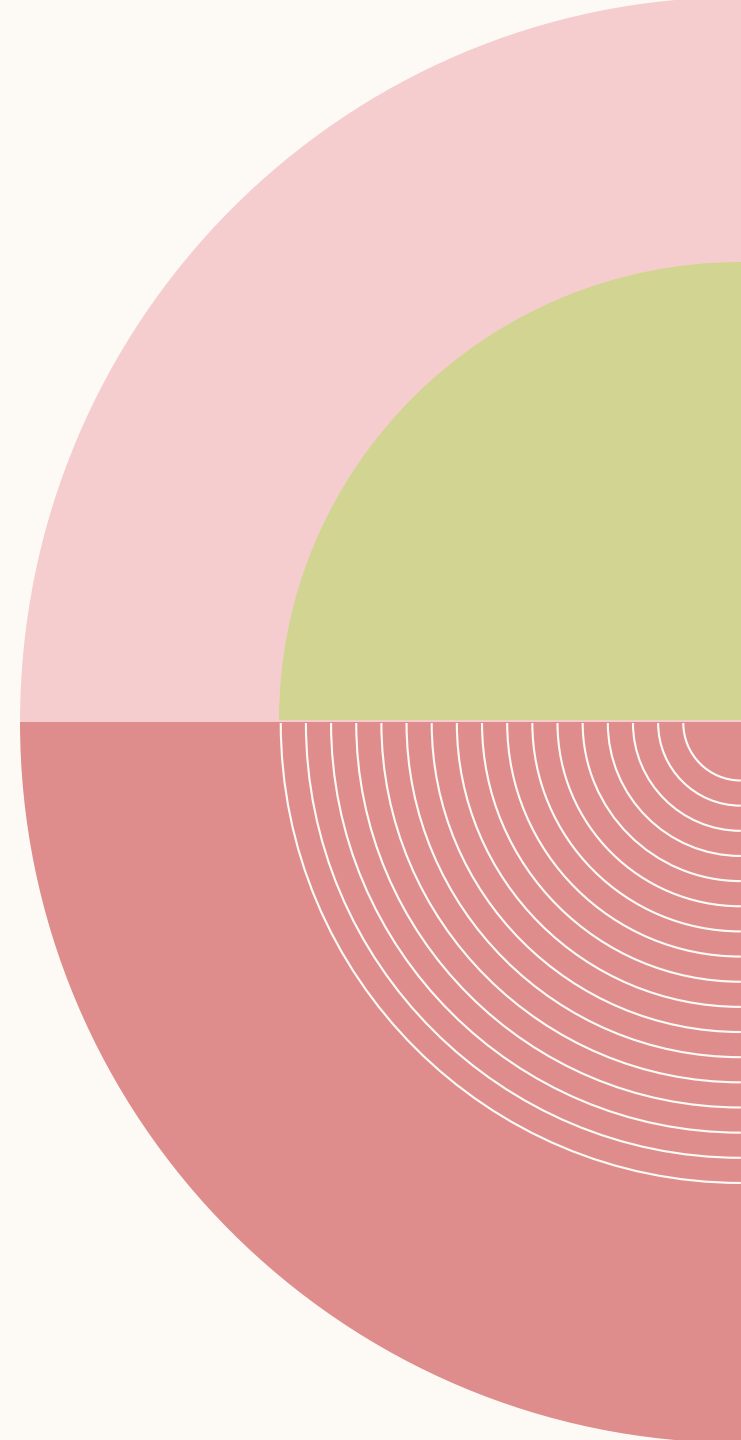
# SUMMARY

Sex

Breeding Cycle

Factor B

Profile Analysis of Behaviors







**THANK YOU**