



**BEB
DIRECT**

A/B Test

Grupo 6

Enia Lahcene, Bruno Rodriguez, Benjamín Mancera

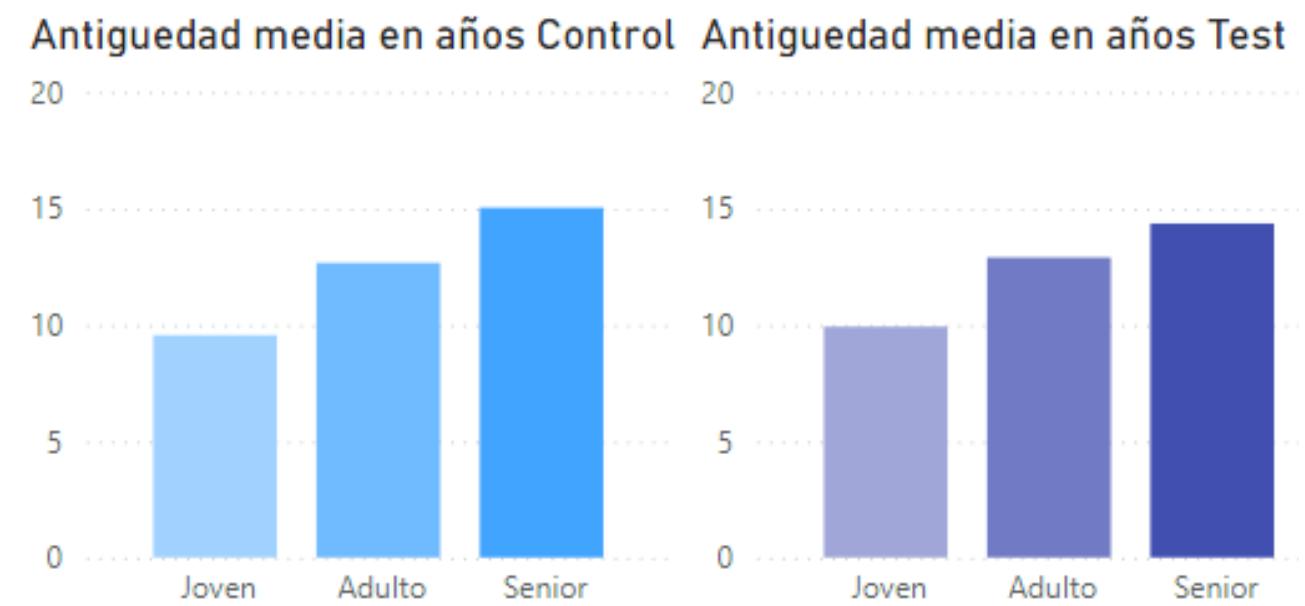
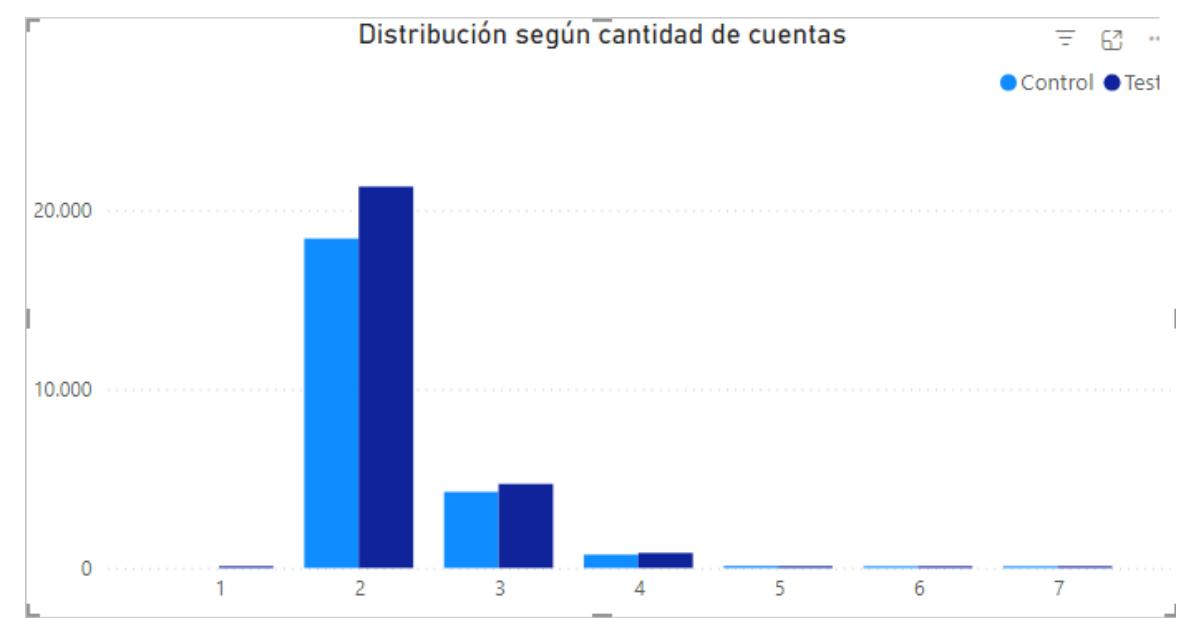
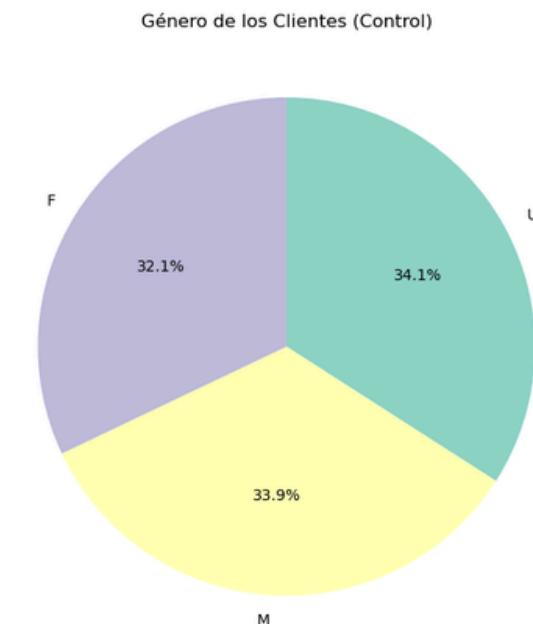
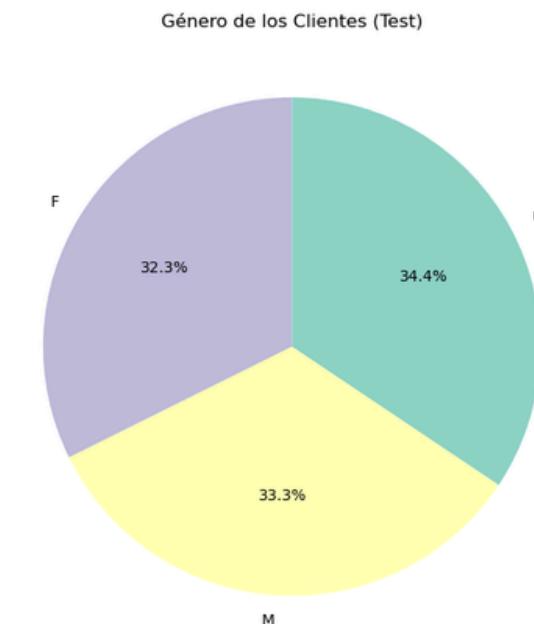
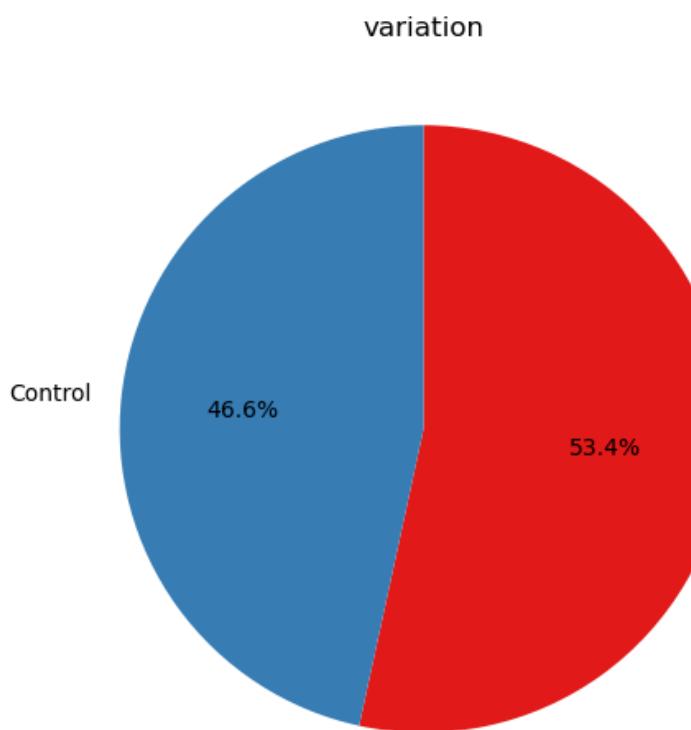
INDEX

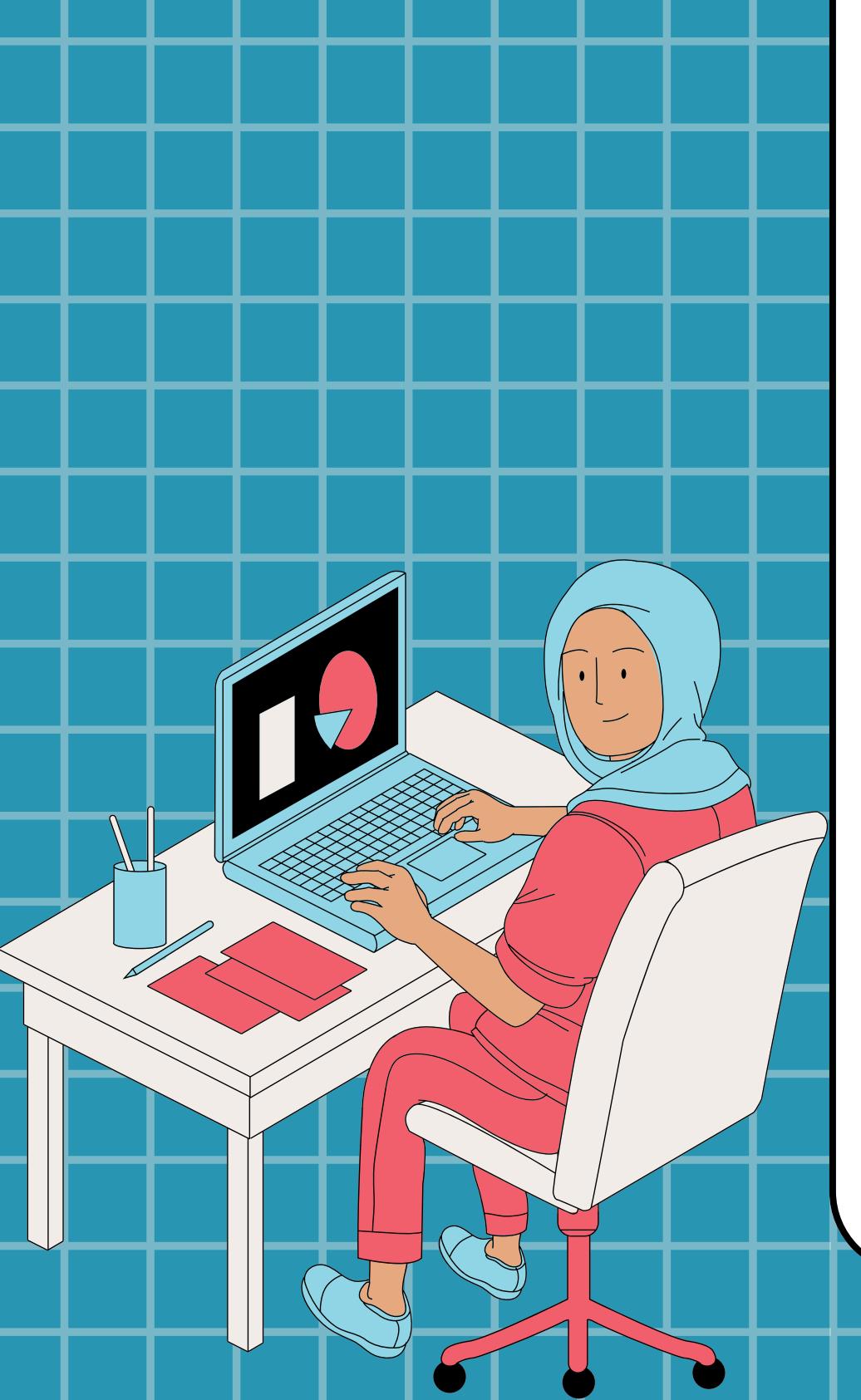
- Introduction:
 - A/B Testing is well done?
 - Data inspection and cleaning
 - Proposed hypothesis
 - EDA (Exploratory Data Analysis)
 - Hypothesis testing
 - Conclusions



A/B TESTING

Comparing both datasets to identify potential bias."

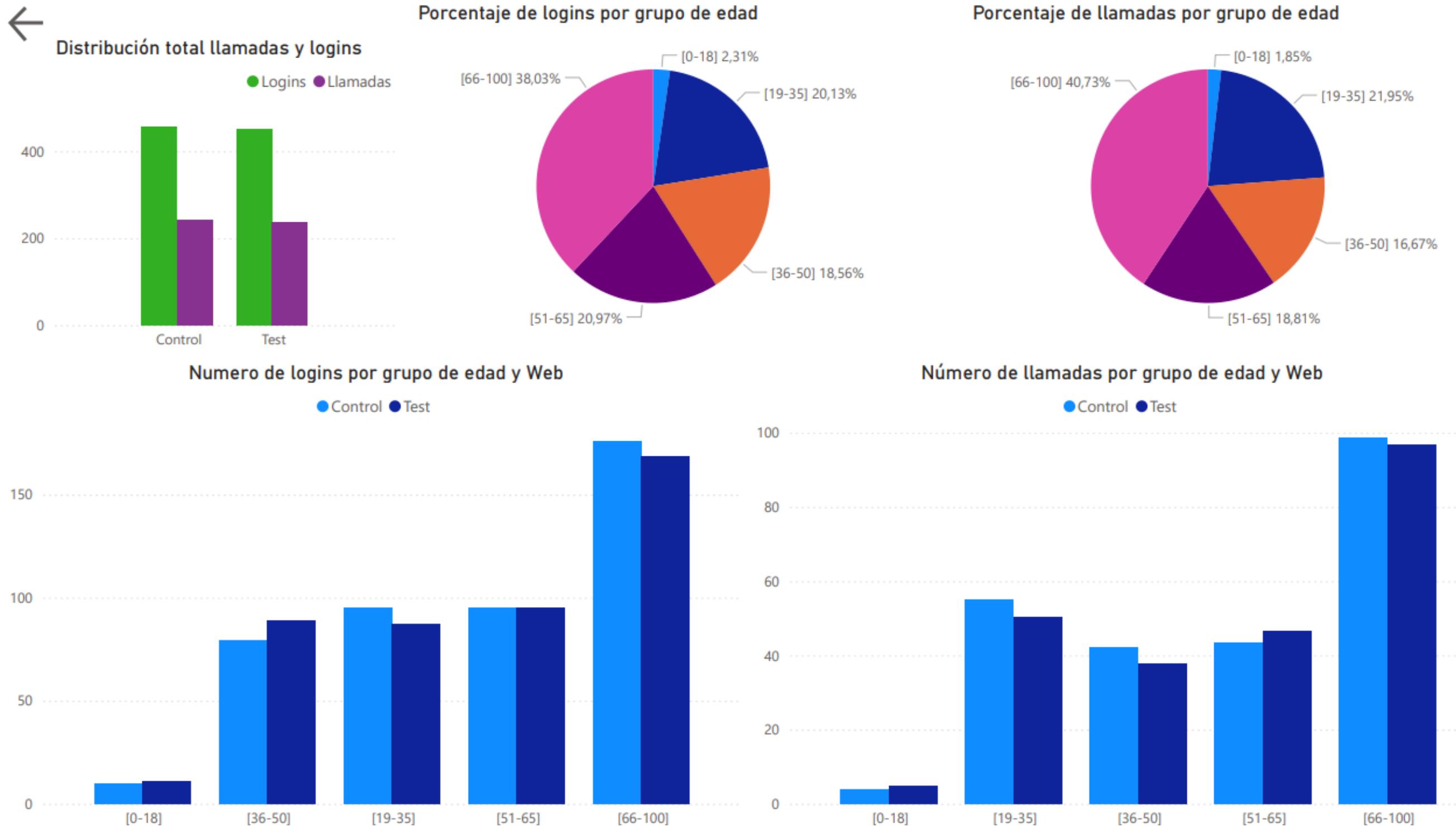




PROPOSED HYPOTHESIS

- The number of calls and logins is significantly lower in the test web.
- The number of completed operations, steps back, and time to completion show better results in the test web.
- The customer balance by age and type of web does not show significant differences.

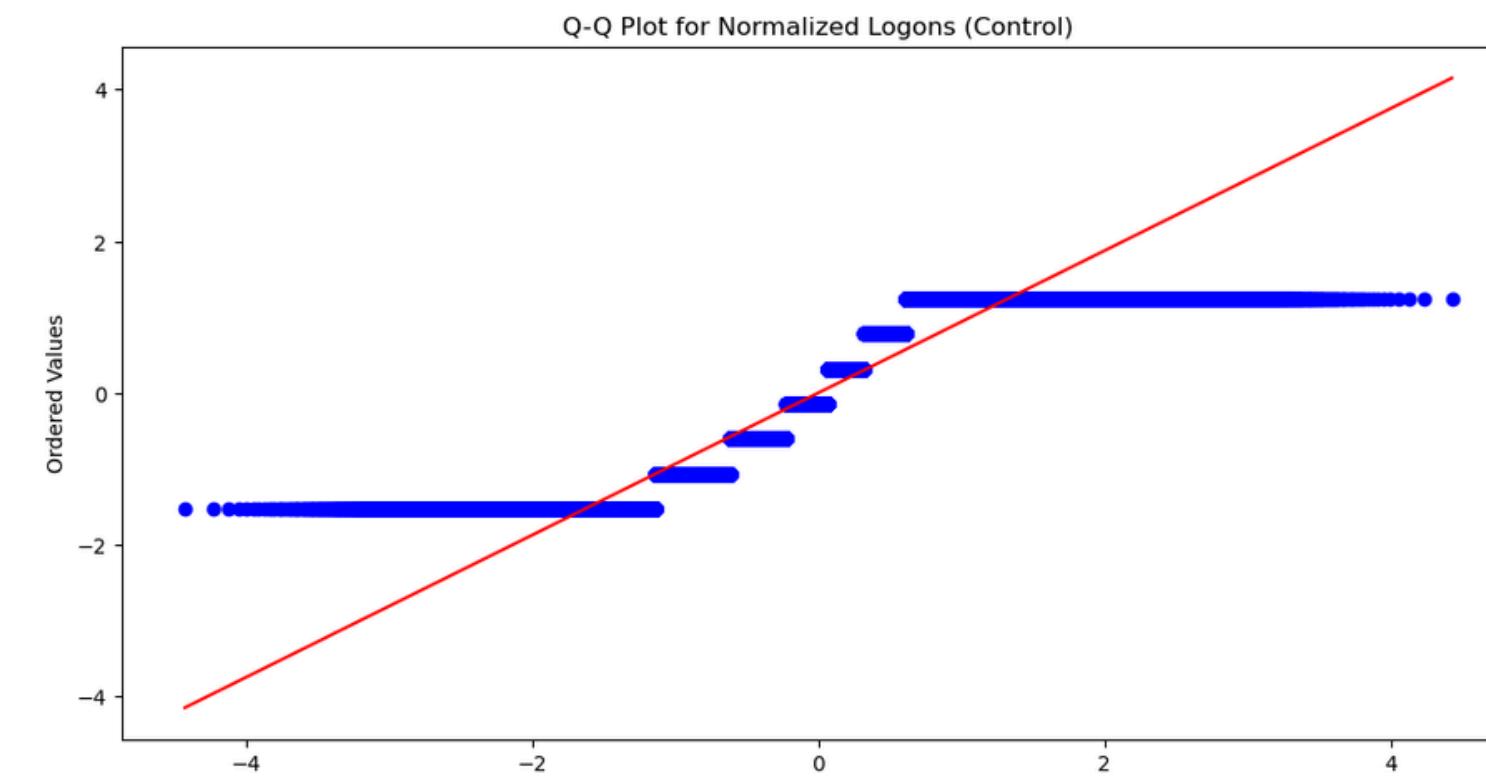
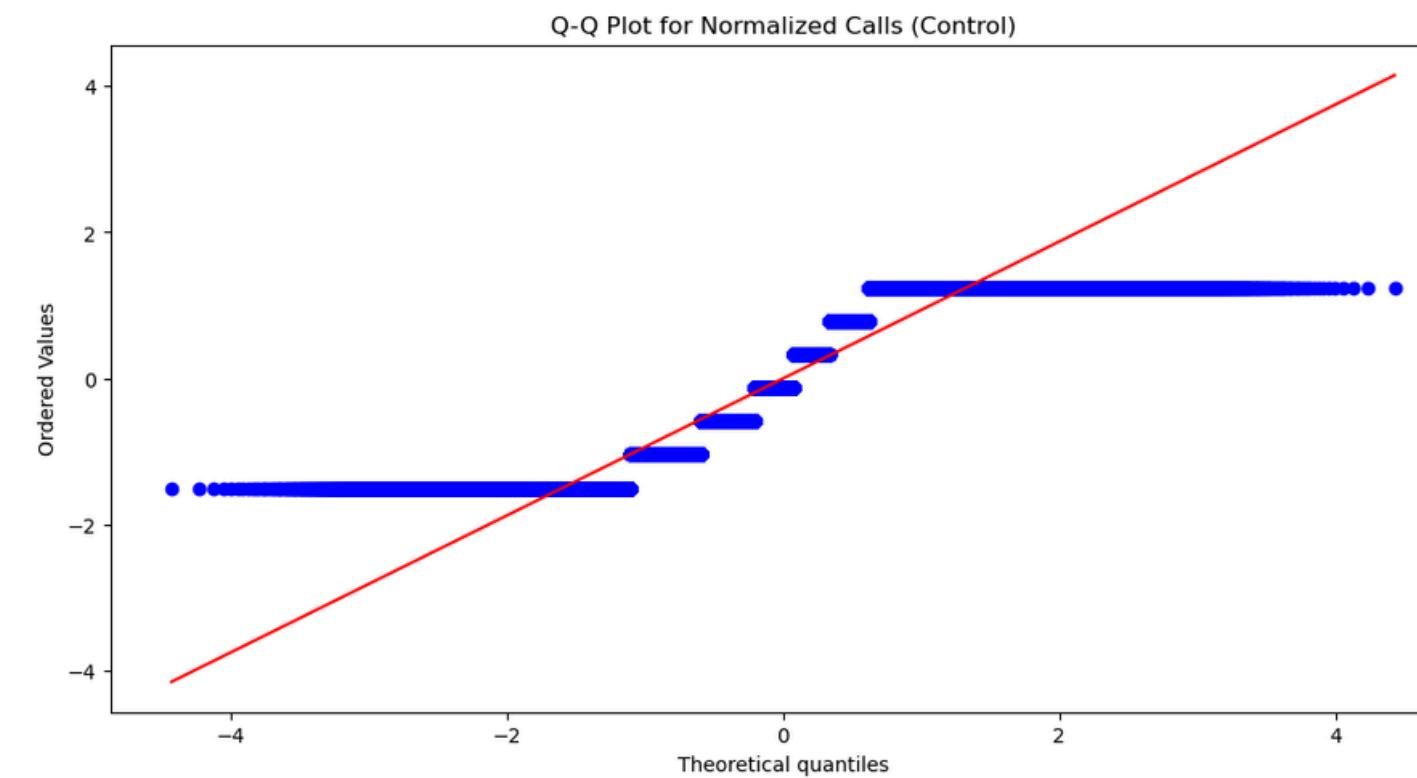
CALLS AND LOGINS



FIRST STEPS

After trying to normalize the data in several ways, Min-Max Saling, Z-Score, Log Transformation, Box-Cox and Quantile Transformation, none of them managed to normalize the data completely.

Therefore we opted for other hypothesis testing statistics such as **t-Student (t-test)**, **Mann-Whitney U**, **Kolmogorov-Smirnov** and **Shapiro-Wilk tests**.



STATISTICAL HYPOTHESIS TEST

Logins

T-test:

- T-statistic: 8.53 P-value: 1.43e-17
- Mean logons in Control: 6.31
- Mean logons in Test: 6.25

Mann-Whitney U test (Control vs Test):

- U-statistic: 12963680581.00
- P-value: 5.04e-17

Kolmogorov-Smirnov test (Control vs Test):

- KS-statistic: 0.02
- P-value: 1.26e-21

The three statistics show that the number of calls is significantly lower in the test group.

Calls

T-test:

- T-statistic: 9.05 P-value: 1.49e-19
- Mean logons in Control: 3.28
- Mean logons in Test: 3.21

Mann-Whitney U test (Control vs Test):

- U-statistic: 12976443204.00
- P-value: 6.65e-19

Kolmogorov-Smirnov test (Control vs Test):

- KS-statistic: 0.02
- P-value: 6.26e-23

The three statistics show that the number of calls is significantly lower in the test group.



HYPOTHESIS TESTING

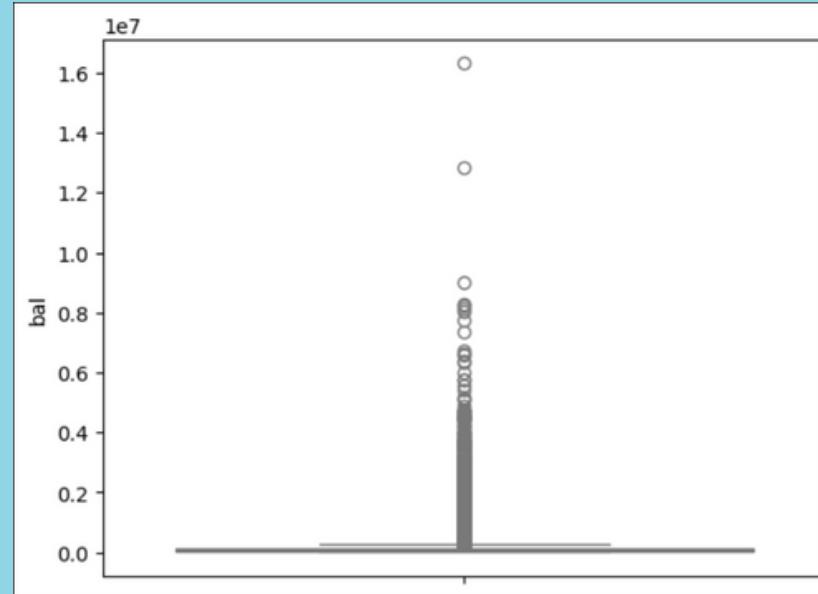
Samples should be well selected, so that their elements look random.

Columns to be analyzed:

- Balance
- Client age
- Client tenure year

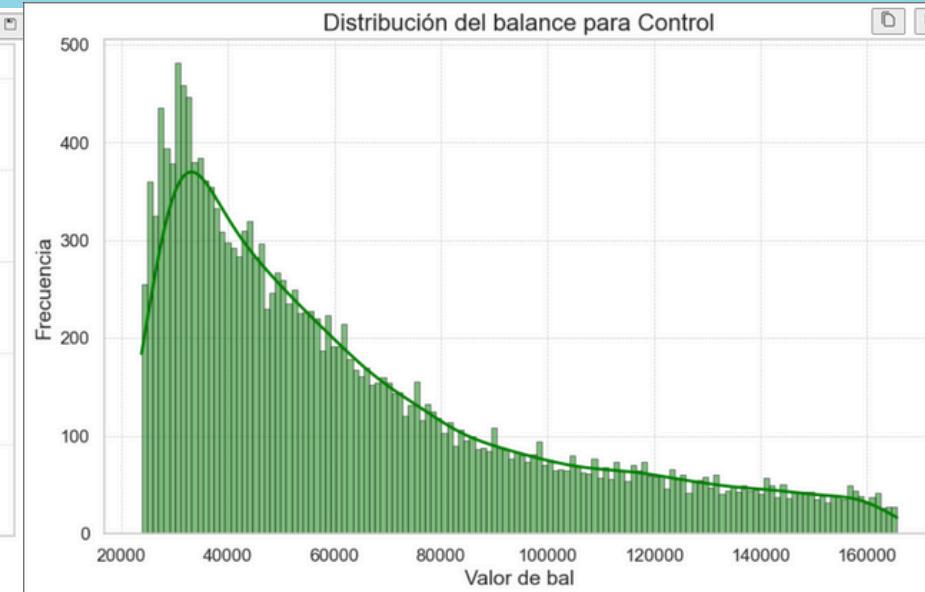
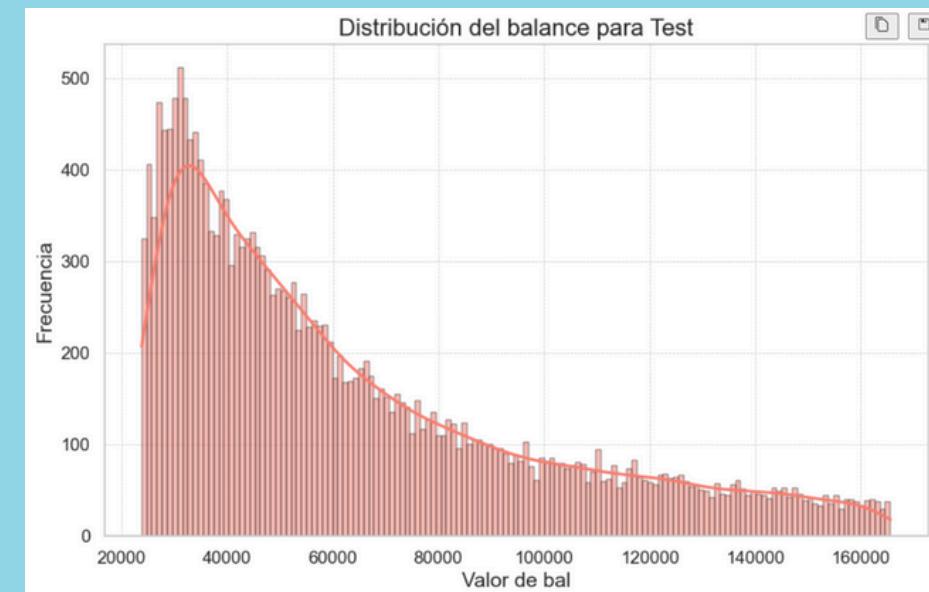
BALANCE

Remove the outliers



- large number of outliers

Visualize the histograms



- distributions are similar

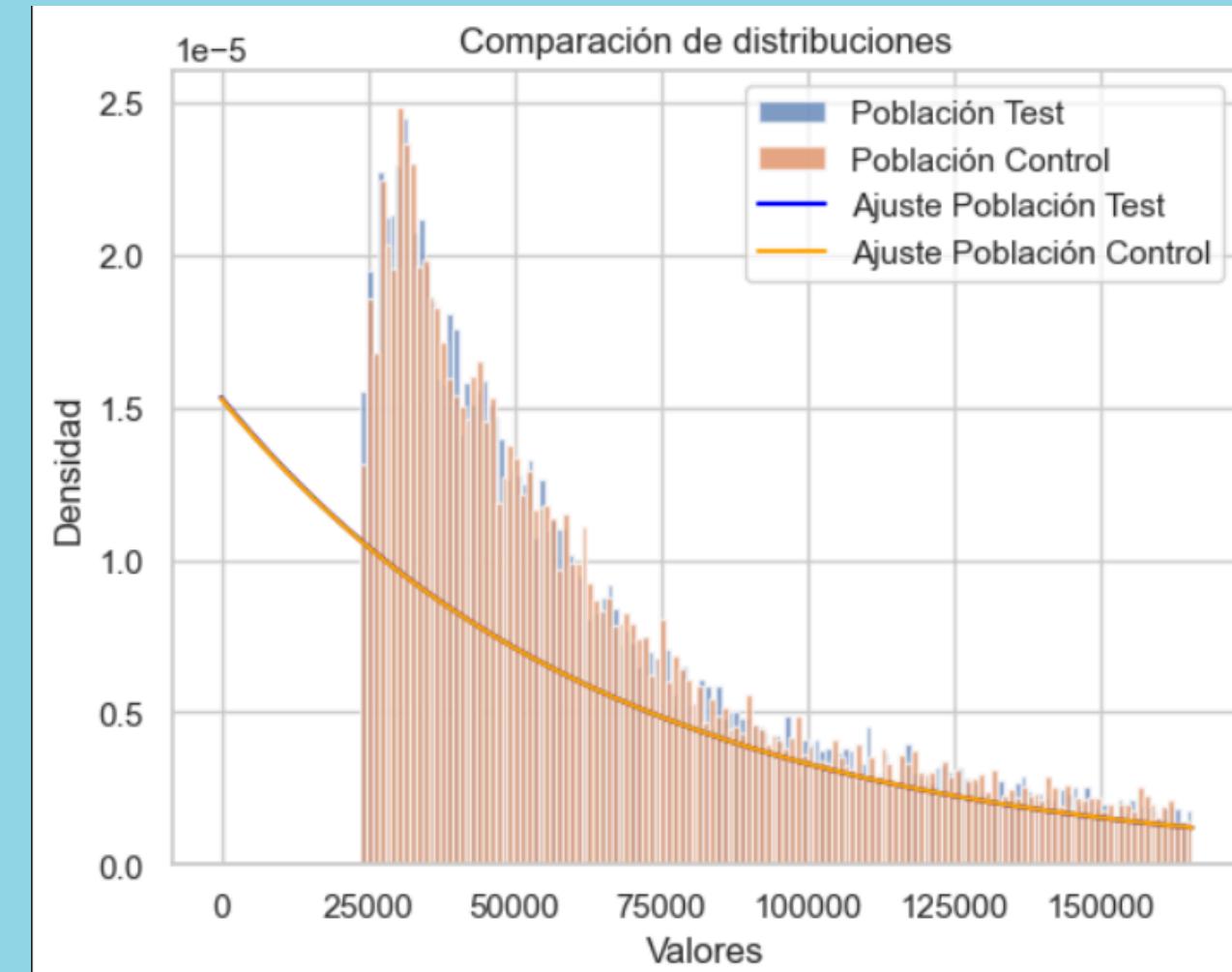
CLIENT AGE

Kolmogorov – Smirnov: 0.16

T-Test: 0.21

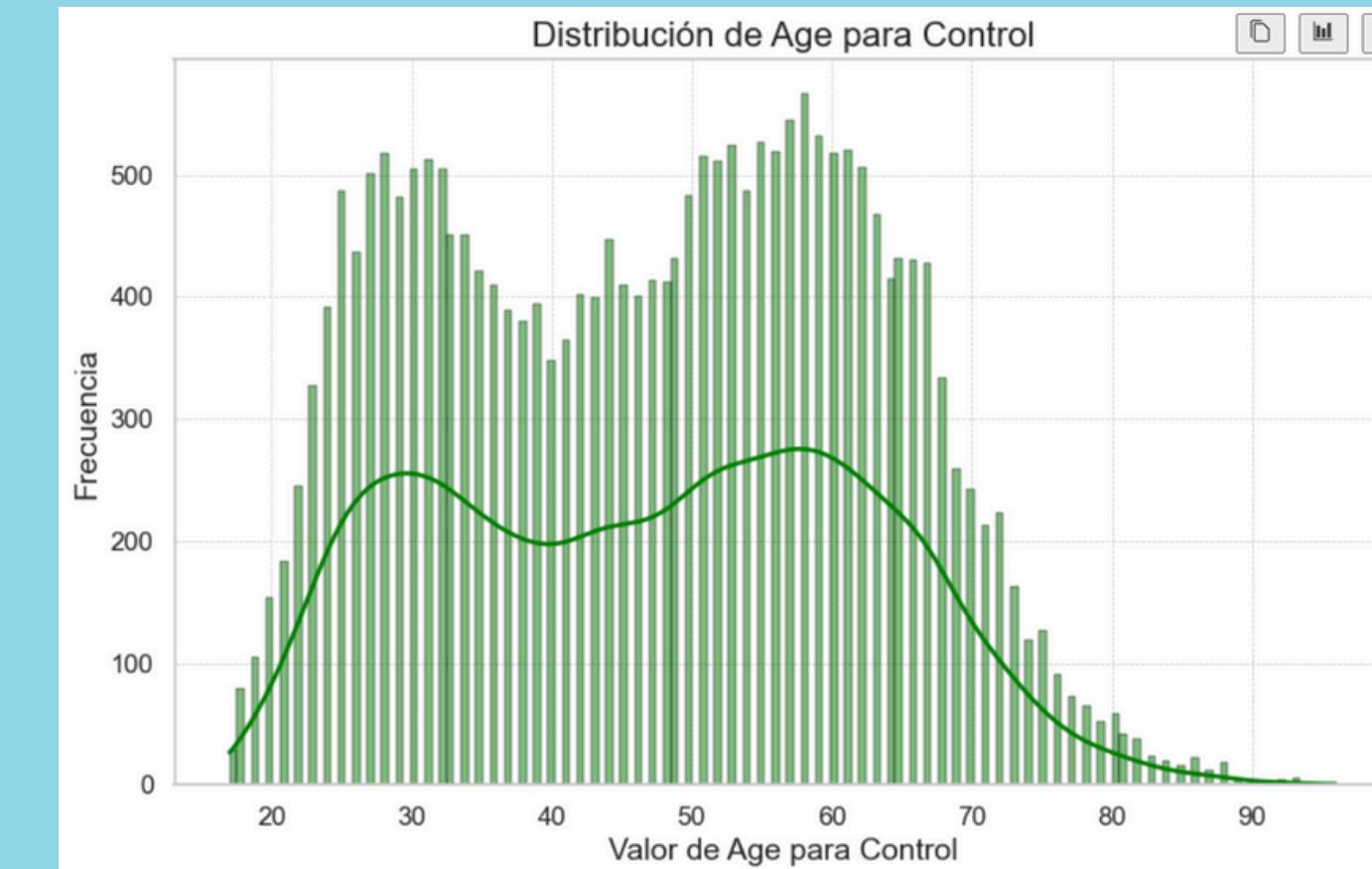
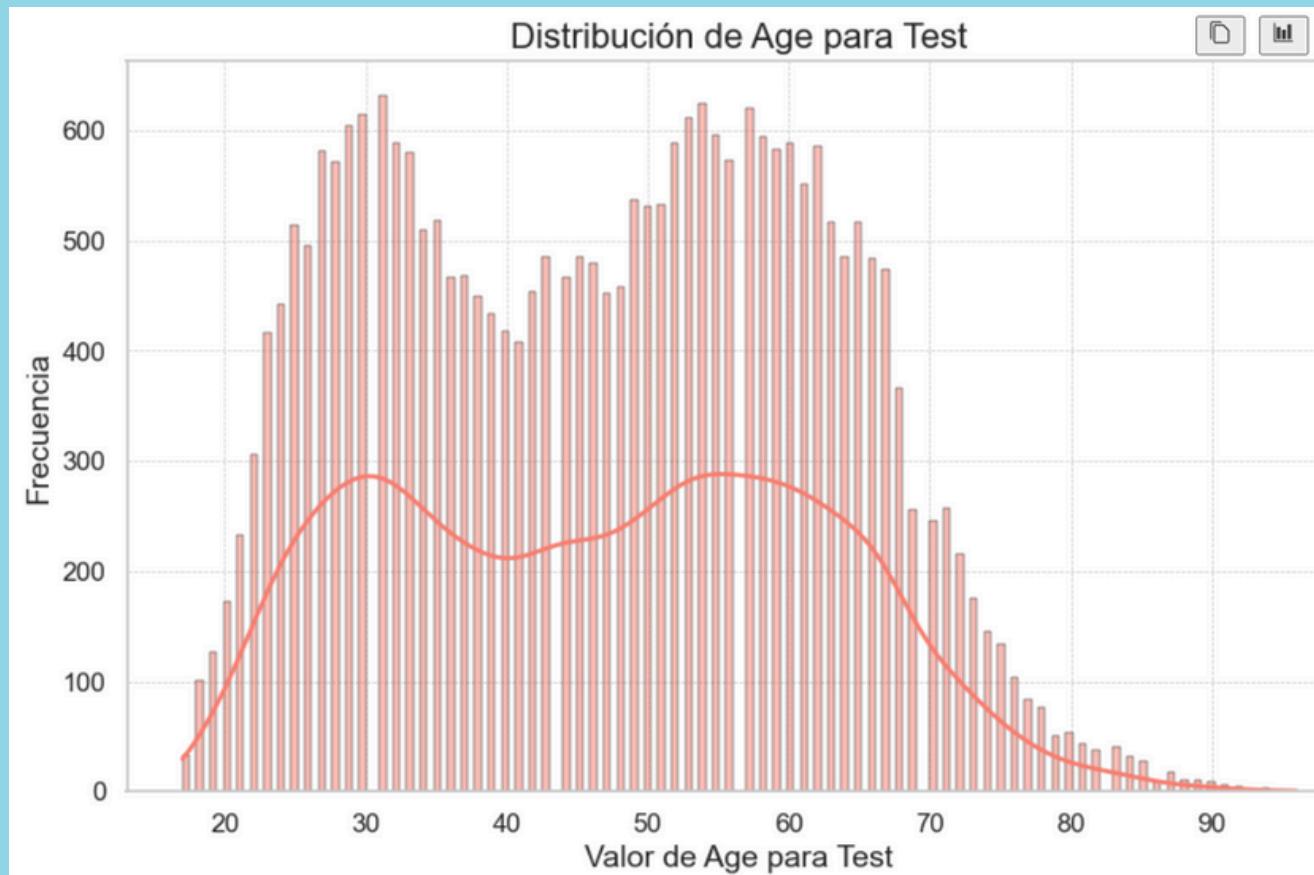
We cannot reject the null hypothesis:
There is insufficient evidence to state
that the distributions are different.

Distribution comparision



CLIENT AGE

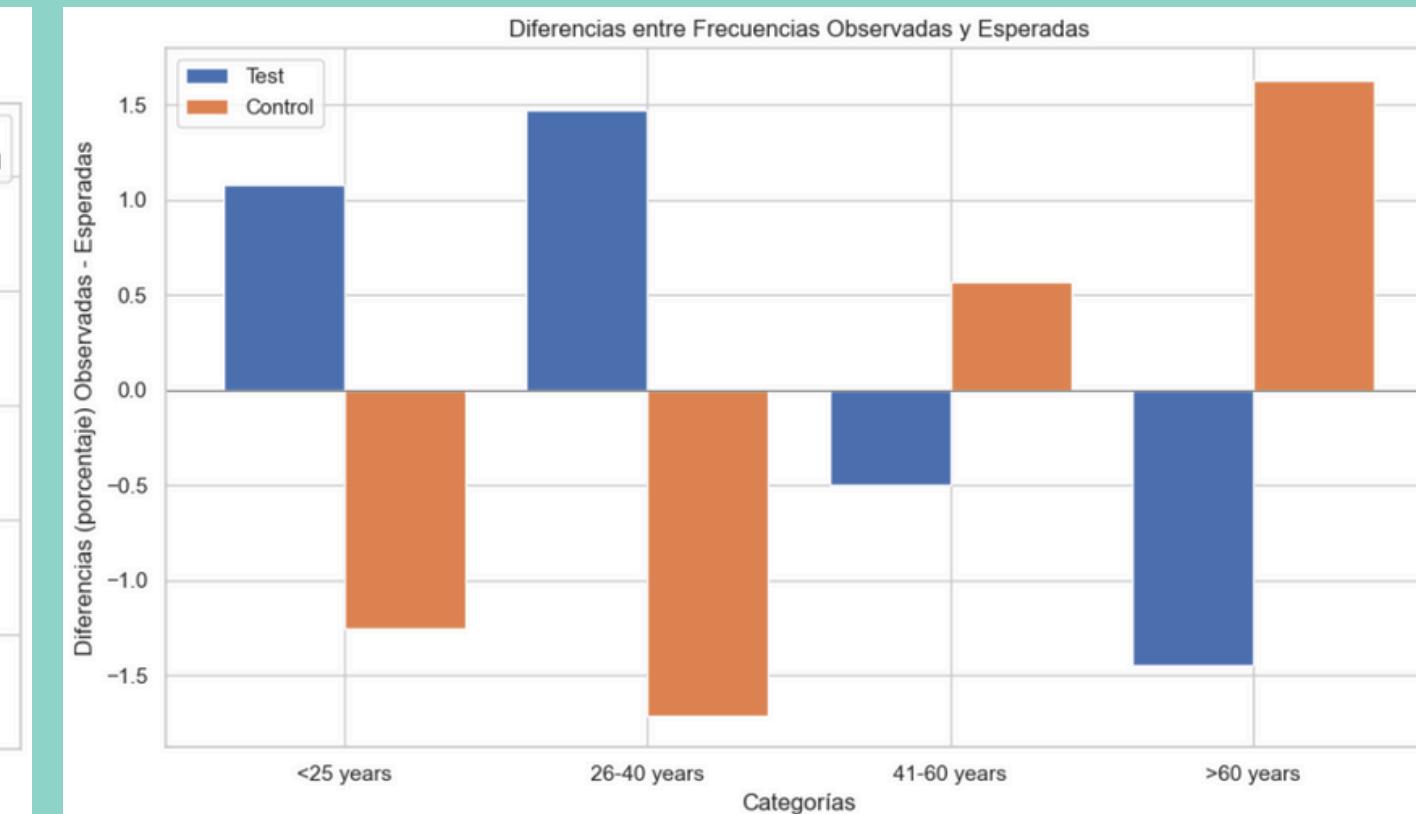
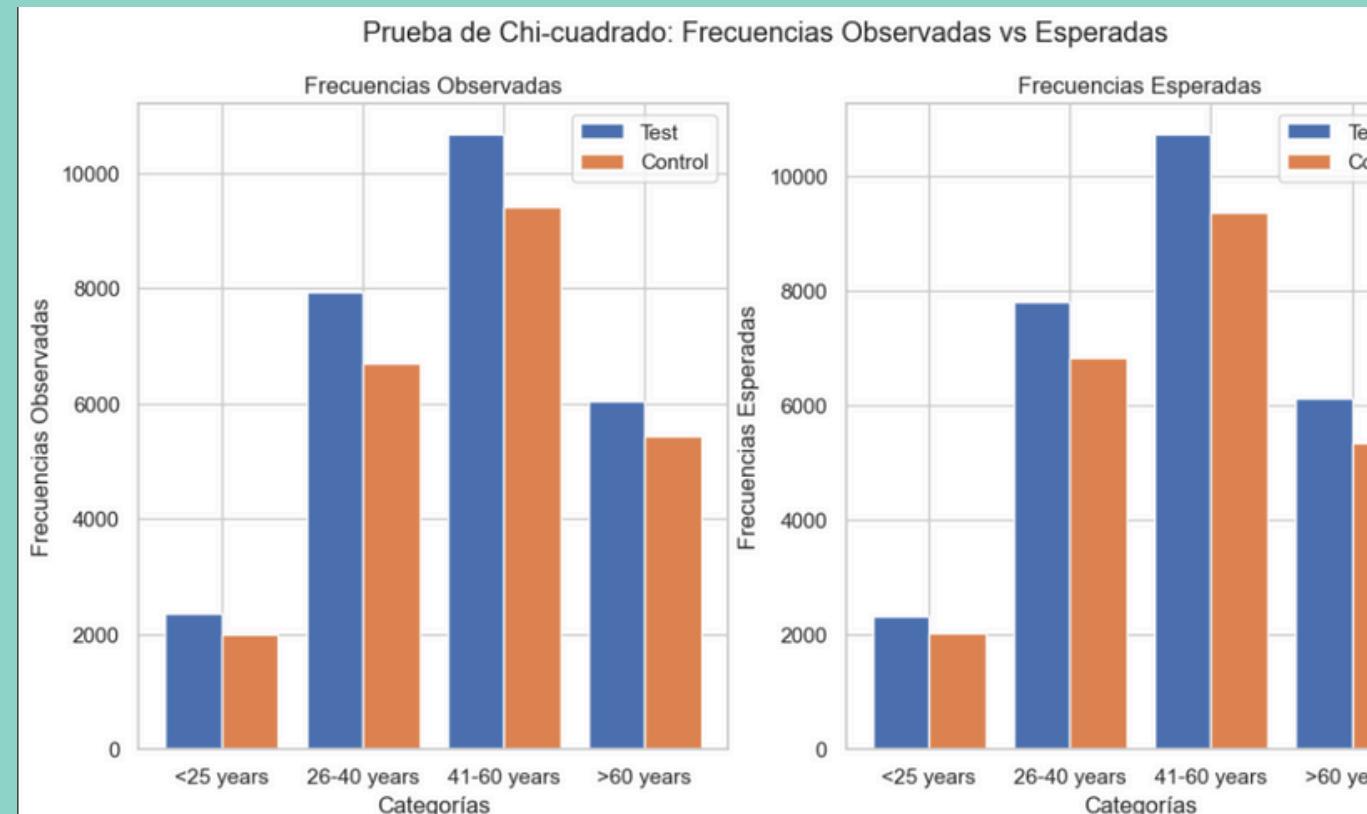
Visualize the histograms



- distributions have two humps and are similar

CLIENT AGE

Visualize the histograms

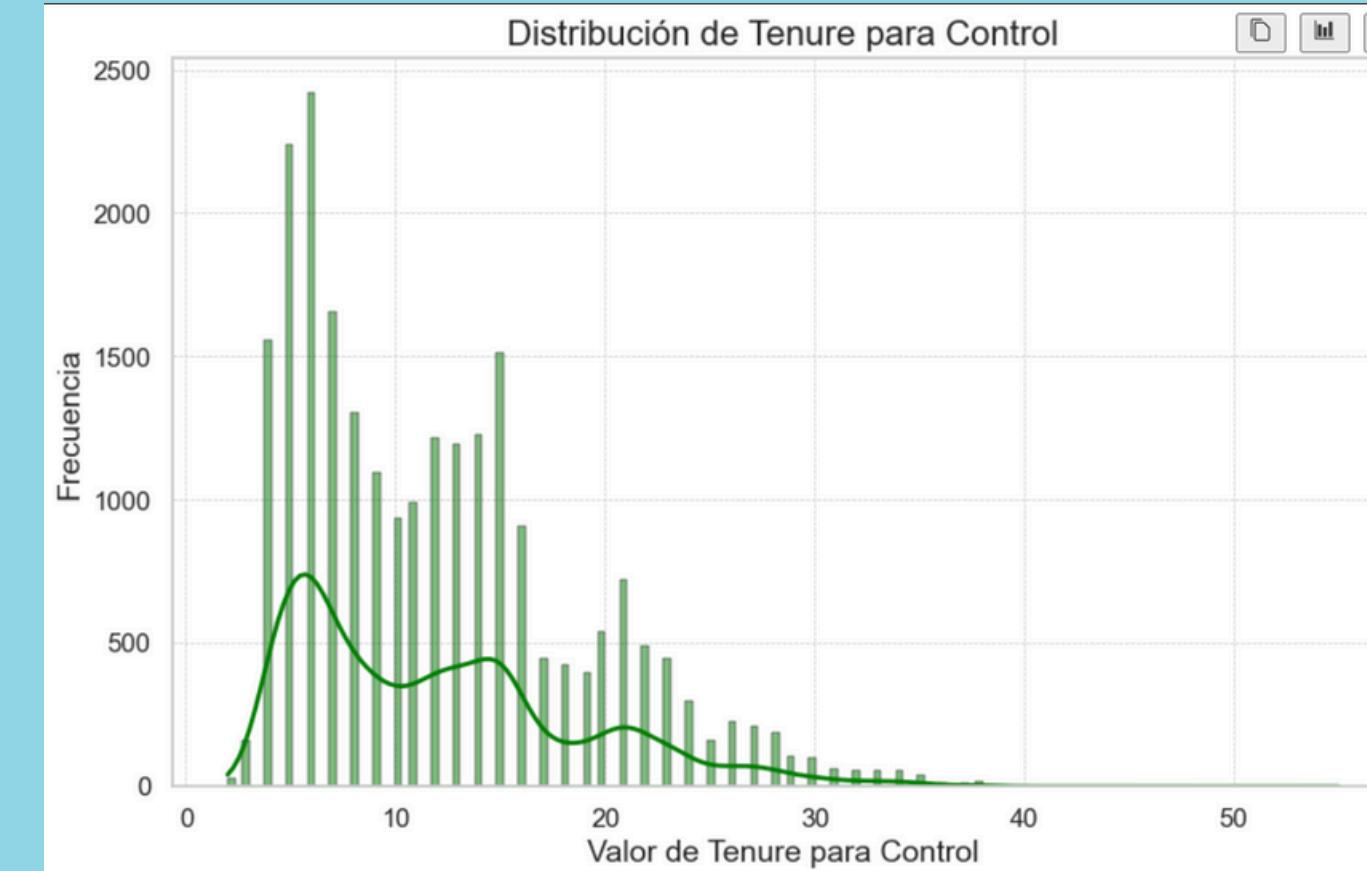
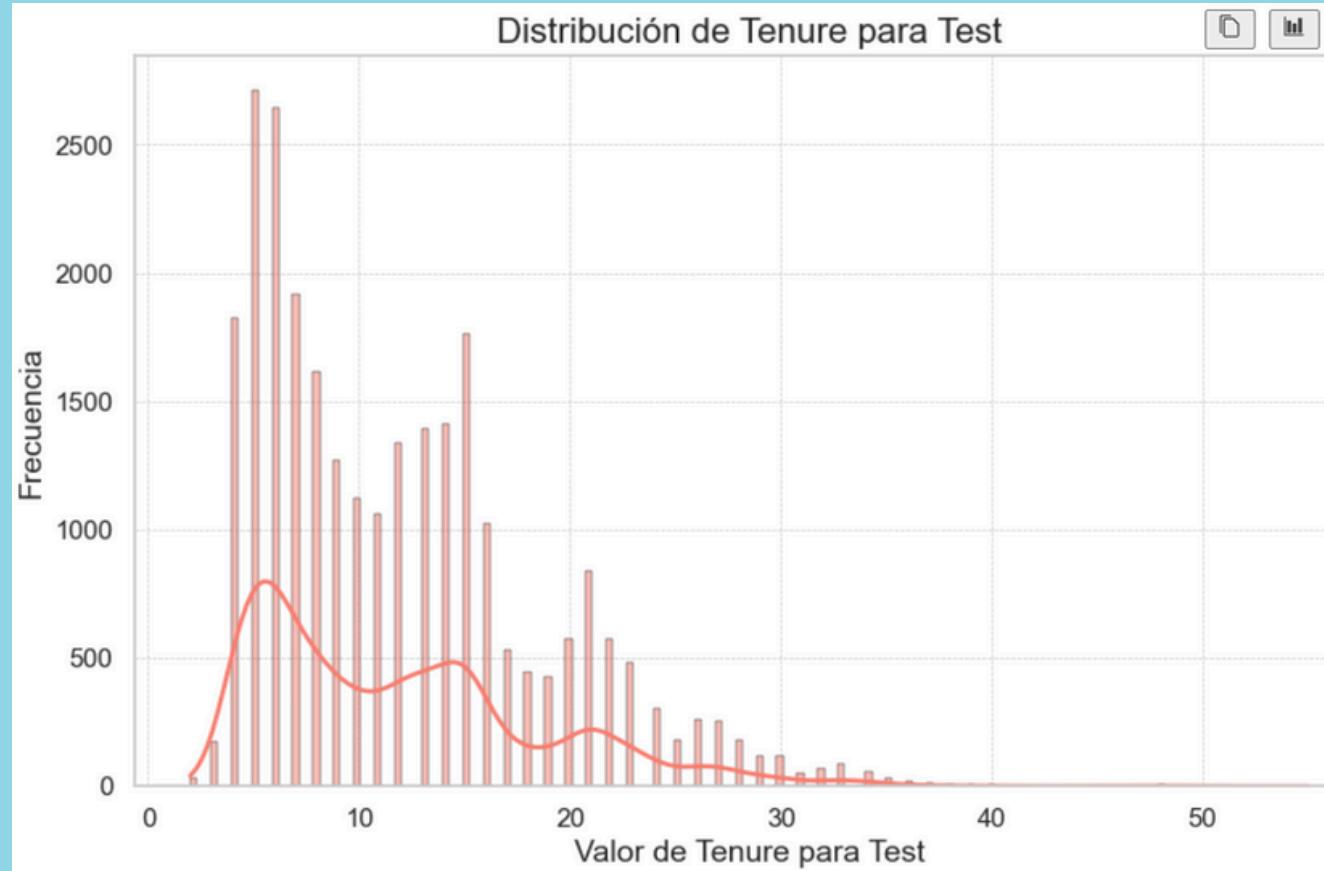


Superimpose and differentiate the frequency.

Chi 2 = 0.057

CLIENT AGE

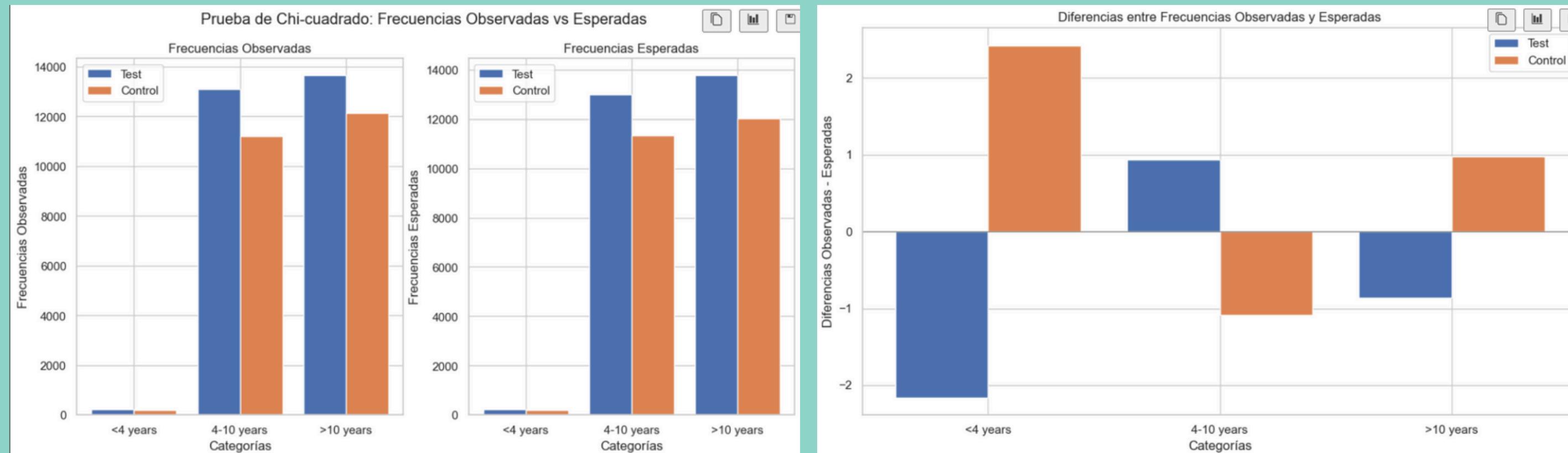
Visualize the histograms



- Distributions are similar

CLIENT AGE

Visualize the histograms



Superimpose and differentiate the frequency.

Chi 2 = 0.09

TOTAL TIME INVESTED

Control web:

- Delta_t mean: 209.16 s
- Delta_t median: 127.00 s
- Delta_t variance: 126168.97 s²
- Delta_t standard deviation: 355.20 s
- Delta_t IQR: 278.00 s

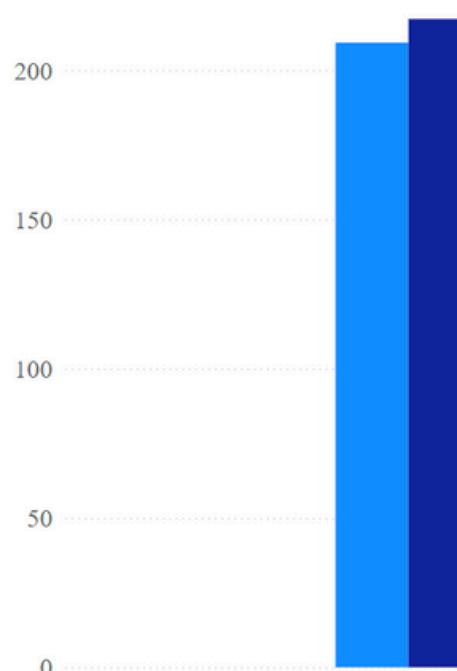


Test web:

- Delta_t mean: 217.11 s
- Delta_t median: 124.00 s
- Delta_t variance: 182854.66 s
- Delta_t standard deviation: 427.62 s
- Delta_t IQR: 248.00 s

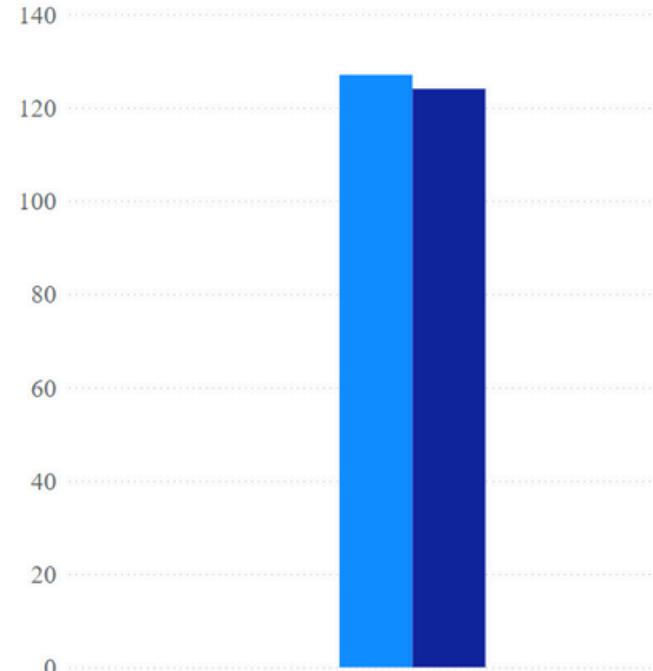
Mean [s]

● Control ● Test



Median [s]

● Control ● Test

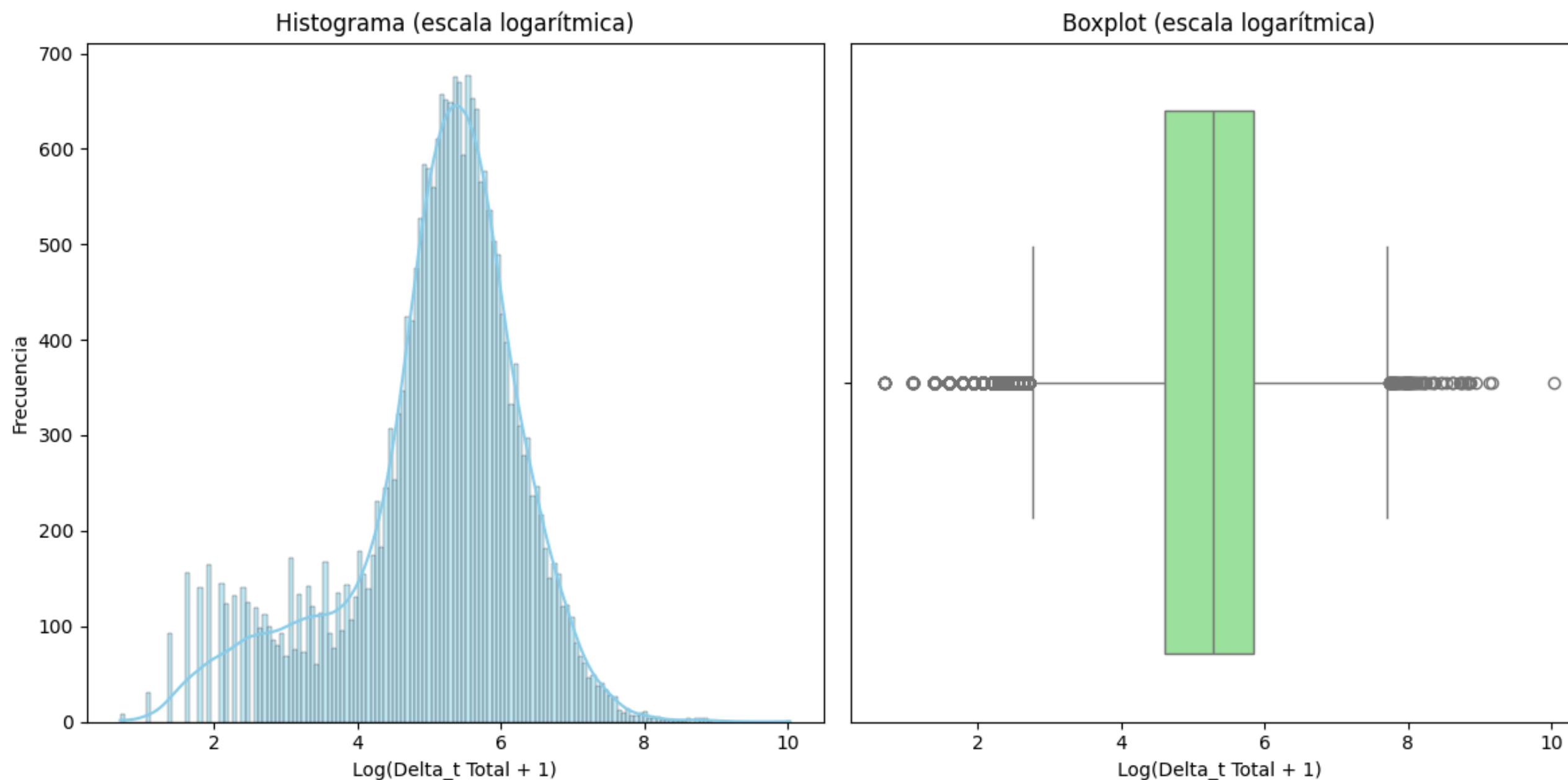


Steps to visualize it in a histogram and in a boxplot:

1. Apply logarithmic transformation to the variable
2. Eliminate null values

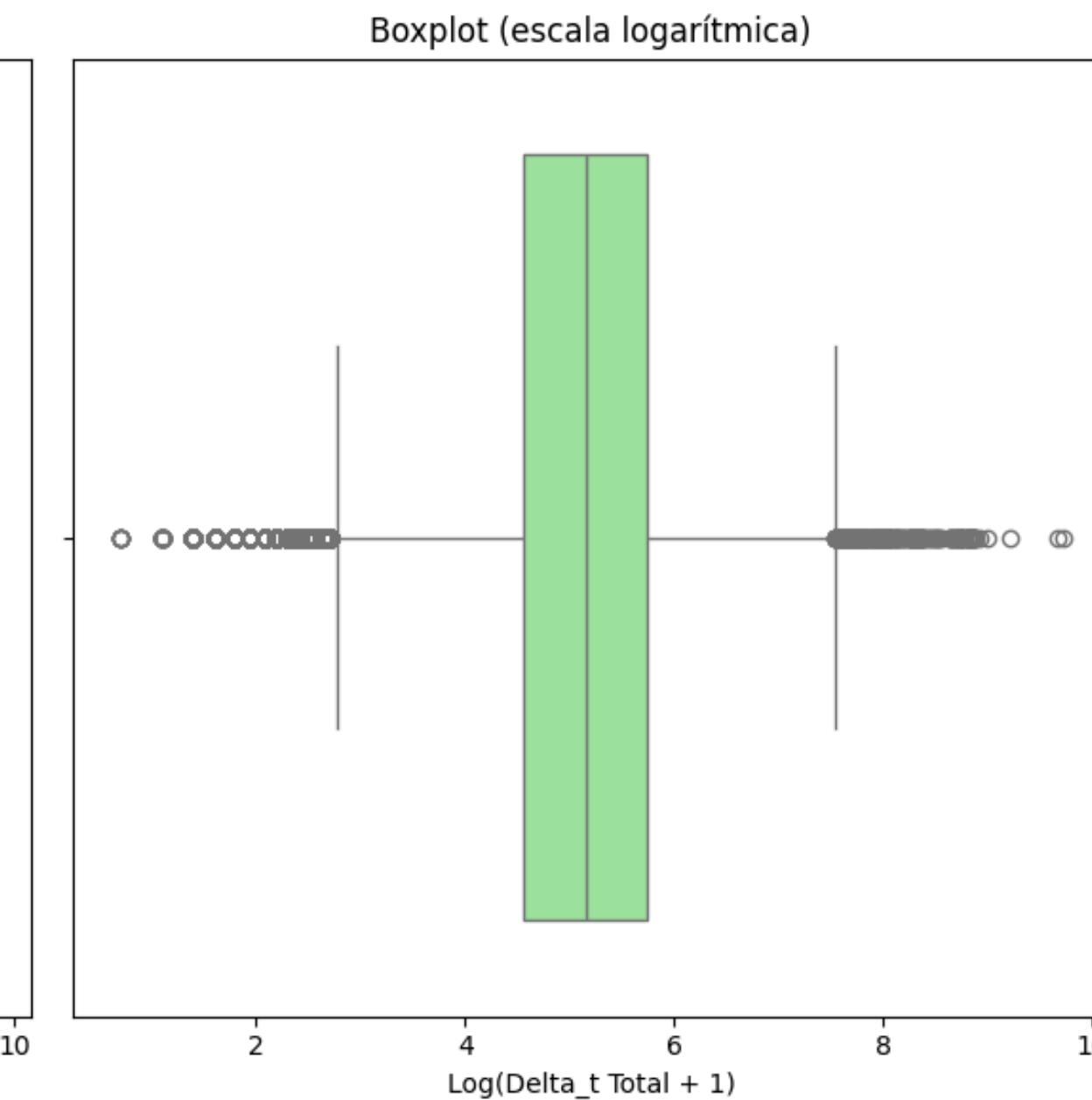
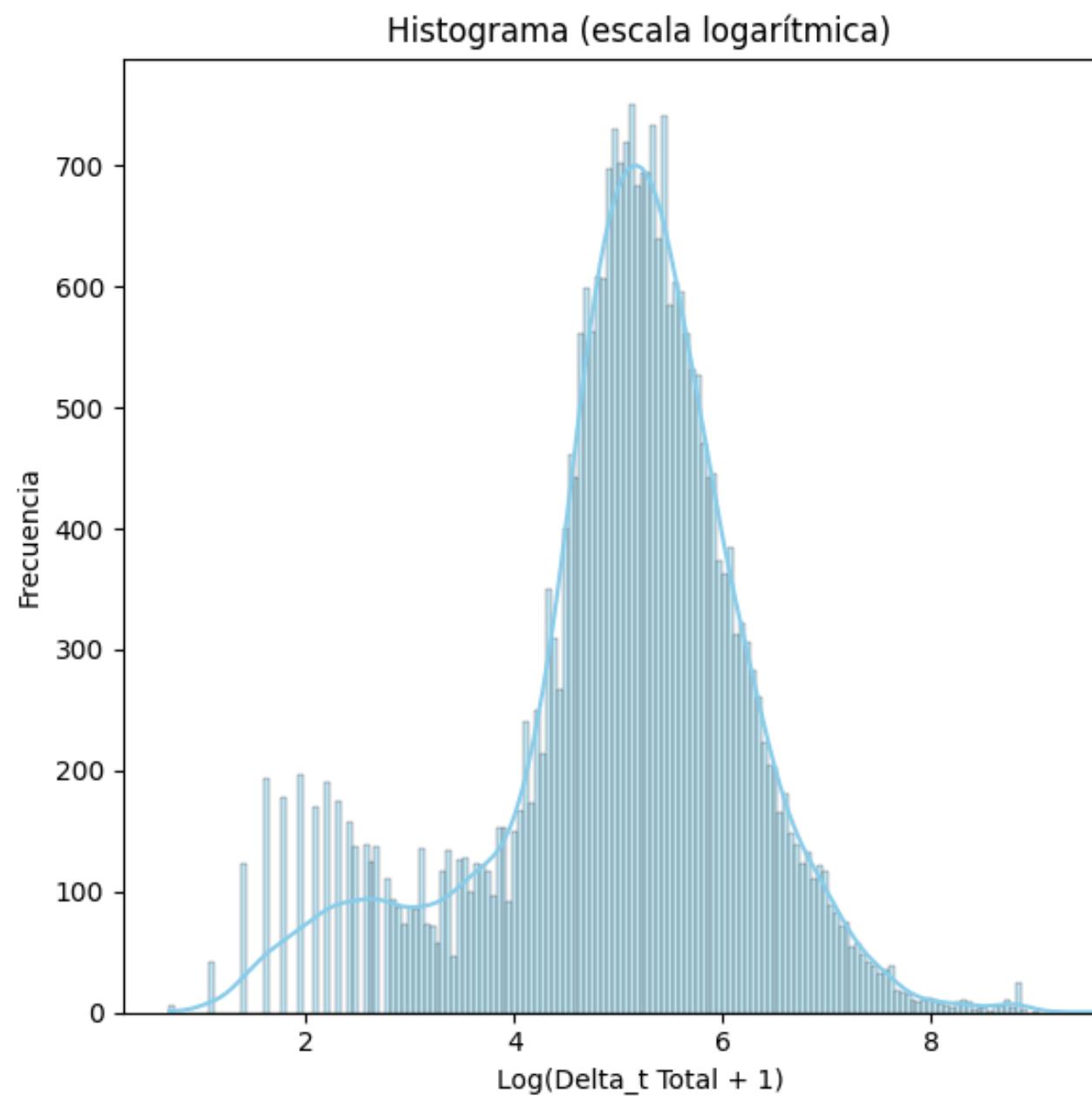
TOTAL TIME INVESTED IN CONTROL WEB.

HISTROGRAM AND BOXPLOT



TOTAL TIME INVESTED IN TEST WEB.

HISTROGRAM AND BOXPLOT



STATISTICAL HYPOTHESIS TEST

Shapiro-Wilk test:

- Control: p-value = 6.22e-67 < 0.05 --> Control's distribution is not normal
- Test: p-value = 1.61e-66 < 0.05 --> Test's distribution is not normal

Mann-Whitney u test (Control vs Test):

- p-value = 7.37e-23 < 0.05 --> There is a significant difference between the distributions.

Kolmogrov-Smirnov test (Control vs Test):

- p-value = 2.36e-36 < 0.05 --> Control and Test have significantly different distributions.

TIME INVESTED PER STEP

Time invested by step

Control Test

300

250

200

150

100

50

0

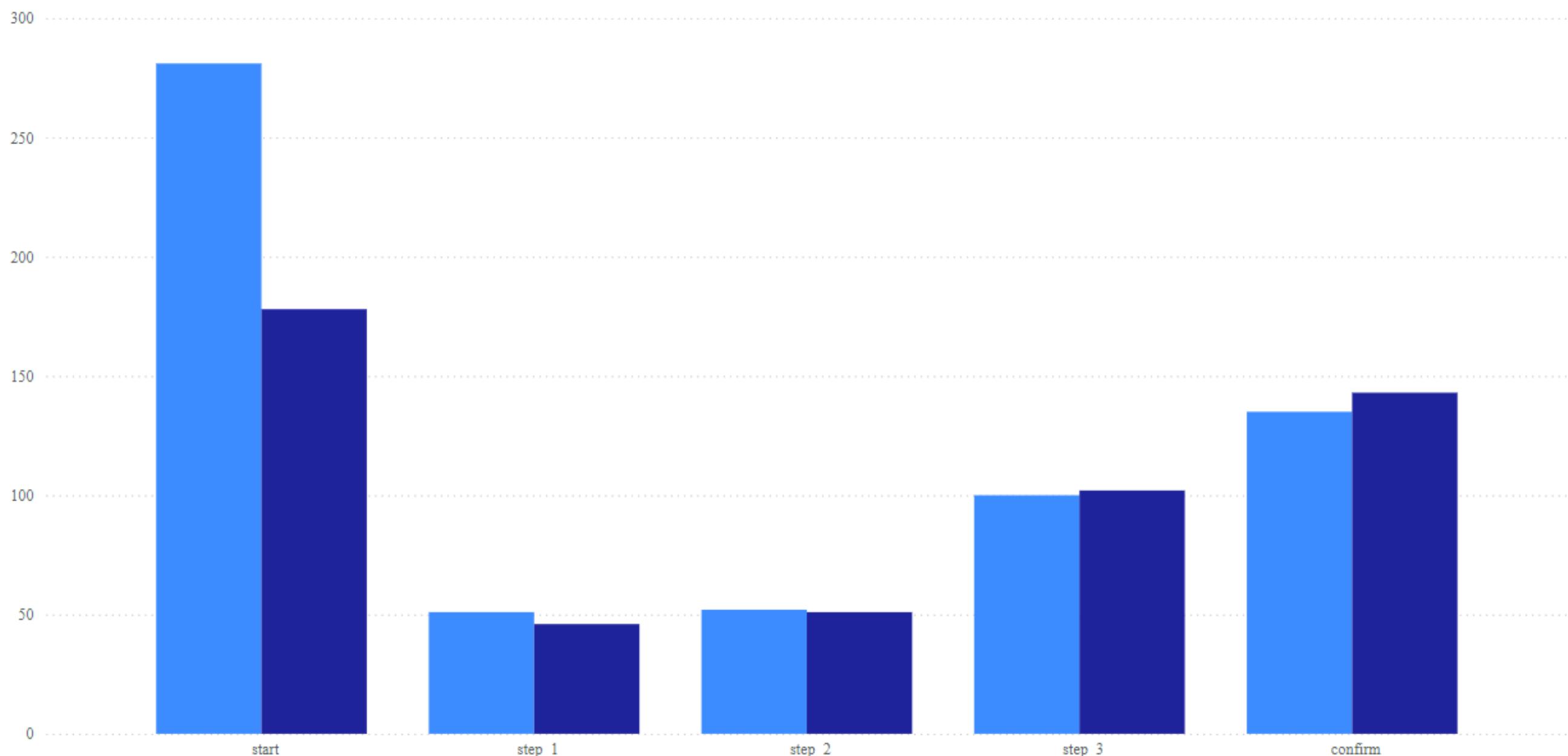
start

step_1

step_2

step_3

confirm



START STEP

Control web:

- Delta_t mean: **281.41 s**
- Delta_t median: **82.0 s**
- Delta_t variance: **519362.61 s²**
- Delta_t standard deviation: **720.66 s**
- Delta_t IQR: **238.00 s**



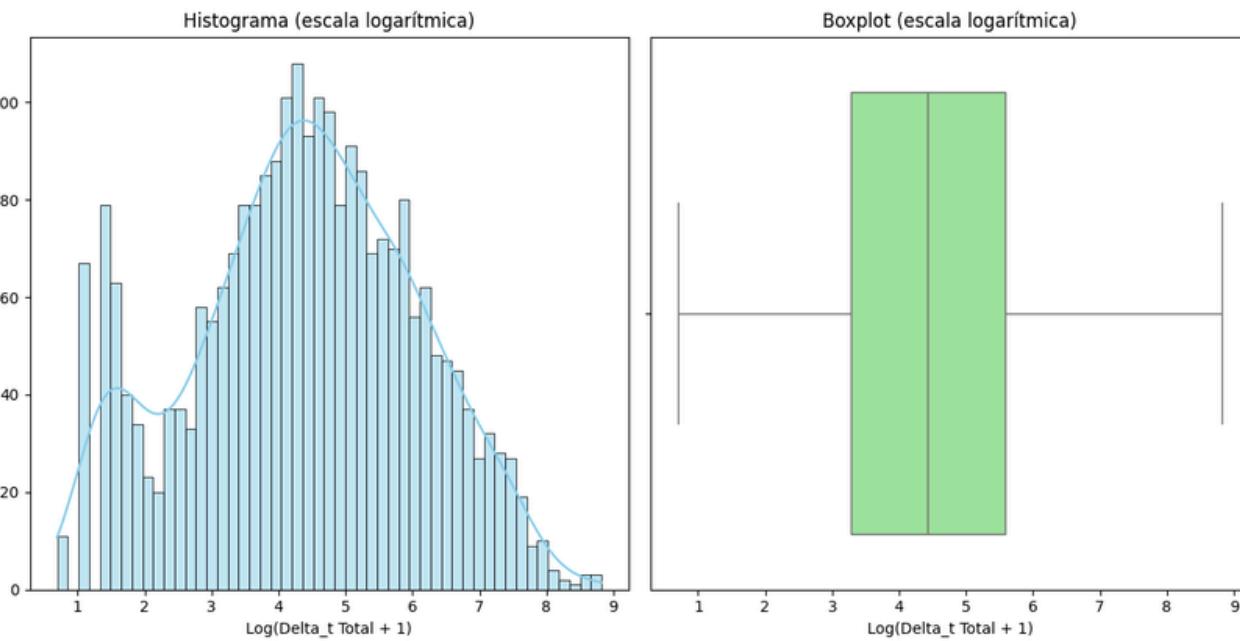
Test web:

- Delta_t mean: **178.65 s**
- Delta_t median: **61.00 s**
- Delta_t variance: **240932.22 s²**
- Delta_t standard deviation: **490.84 s**
- Delta_t IQR: **137.00 s**

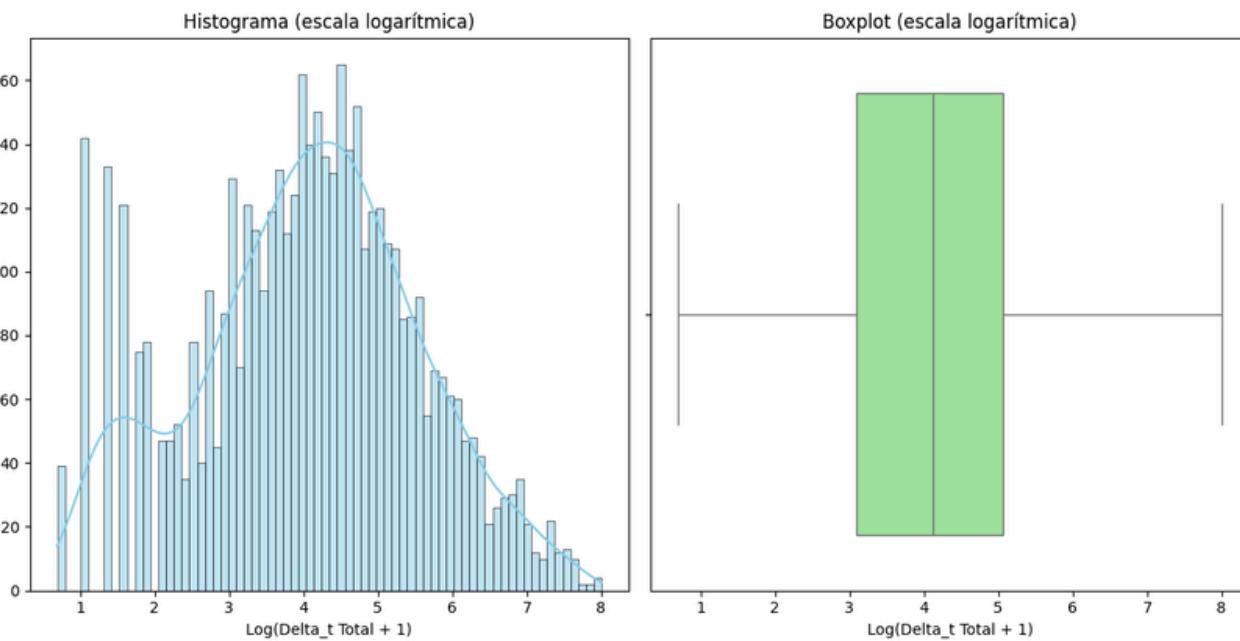
Steps to visualize it in a histogram and in a boxplot:

- Apply logarithmic transformation to the variable
- Eliminate null values
- Eliminate outliers:
 - lower_limit_control = Q1 - 1.5 * IQR
 - upper_limit_control = Q3 + 1.5 * IQR

START STEP



Control Web



Test Web

STEP 1

Control web:

- Delta_t mean: 51.94 s
- Delta_t median: 19.0 s
- Delta_t variance: 18336.34 s²
- Delta_t standard deviation: 135.41 s
- Delta_t IQR: 35.00 s



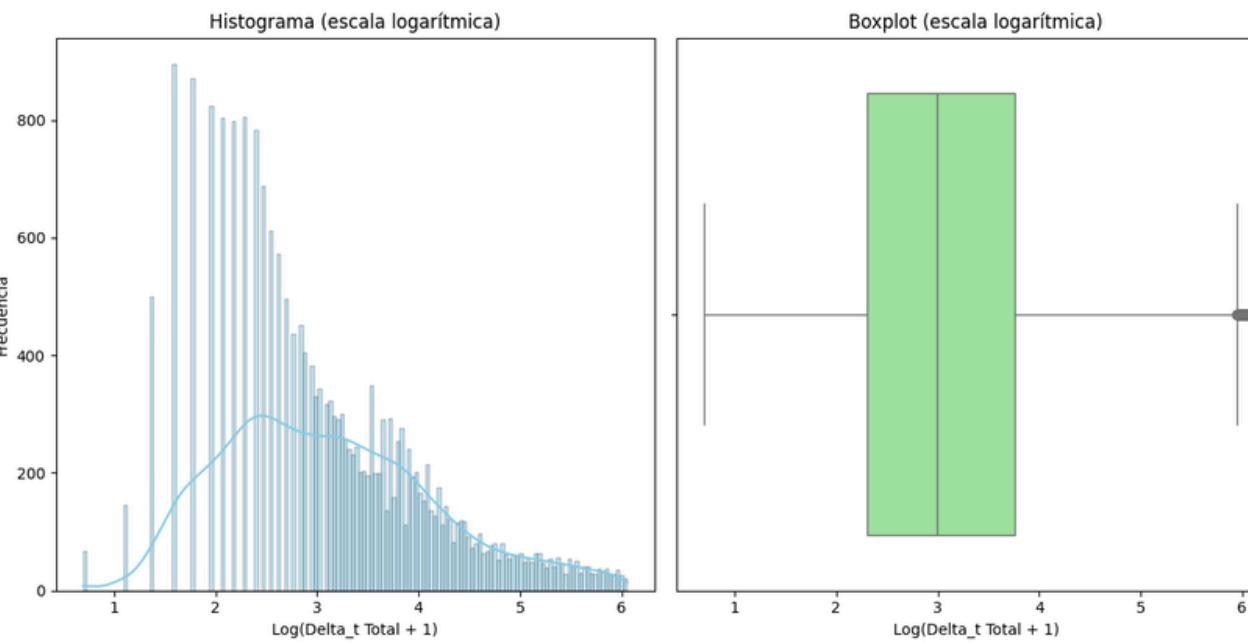
Test web:

- Delta_t mean: 46.75 s
- Delta_t median: 11.0 s
- Delta_t variance: 24307.67 s²
- Delta_t standard deviation: 155.90 s
- Delta_t IQR: 23.00 s

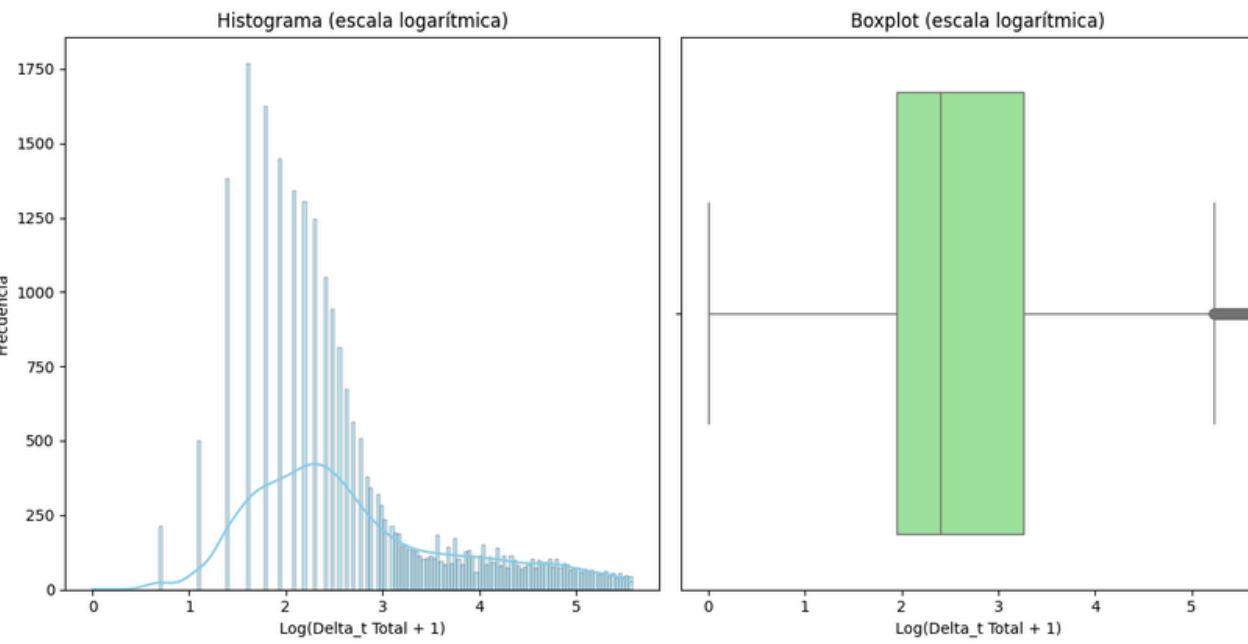
Steps to visualize it in a histogram and in a boxplot:

- Apply logarithmic transformation to the variable
- Eliminate null values
- Eliminate outliers:
 - lower_limit_control = Q1 - 1.5 * IQR
 - upper_limit_control = Q3 + 1.5 * IQR

STEP 1



Control Web



Test Web

STEP 2

Control web:

- Delta_t mean: 52.50 s
- Delta_t median: 22.00 s
- Delta_t variance: 9816.05 s²
- Delta_t standard deviation: 99.07 s
- Delta_t IQR: 45.00 s



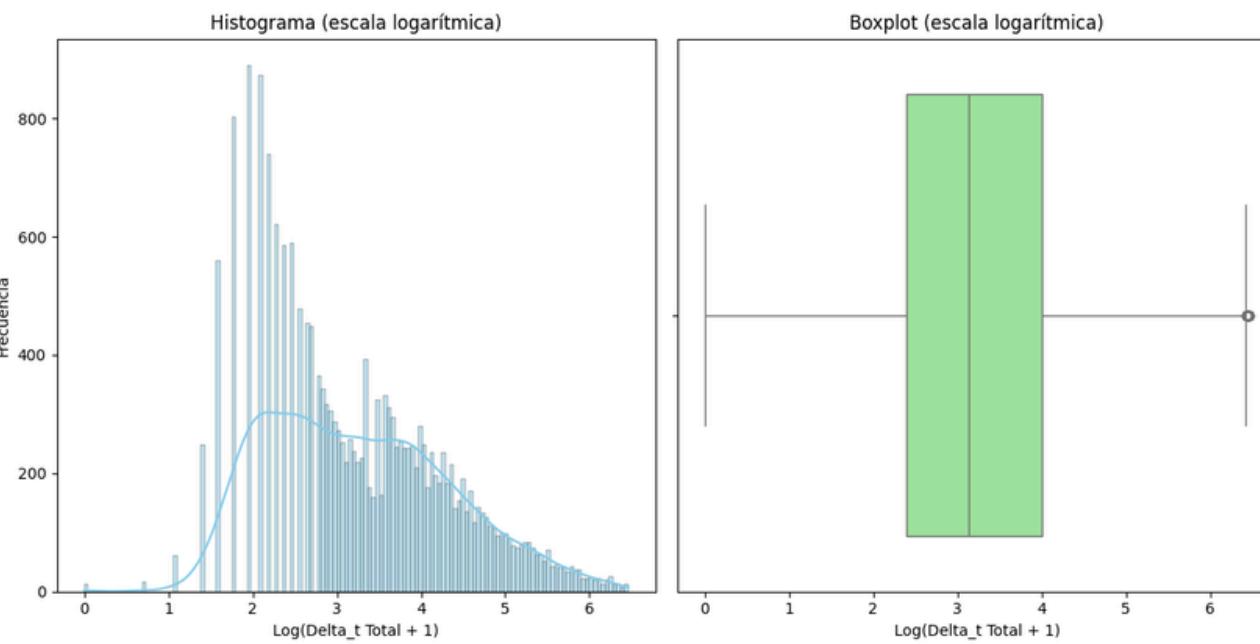
Test web:

- Delta_t mean: 51.64 s
- Delta_t median: 24.0 s
- Delta_t variance: 23474.01 s²
- Delta_t standard deviation: 153.21 s
- Delta_t IQR: 34.00 s

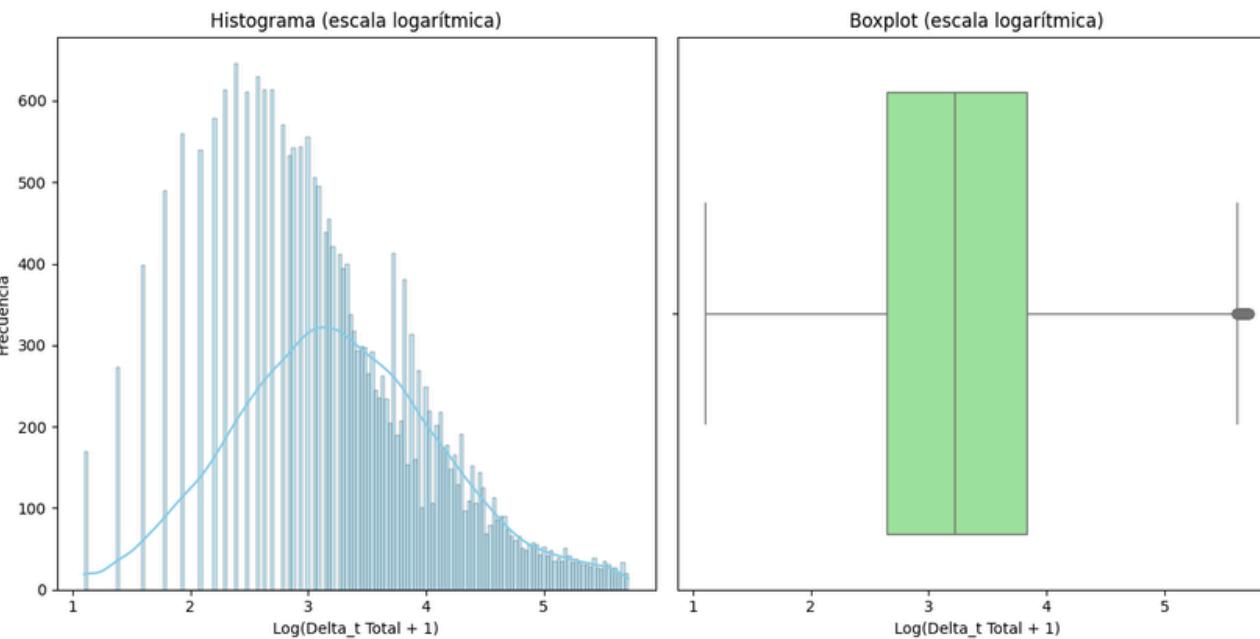
Steps to visualize it in a histogram and in a boxplot:

- Apply logarithmic transformation to the variable
- Eliminate null values
- Eliminate outliers:
 - lower_limit_control = Q1 - 1.5 * IQR
 - upper_limit_control = Q3 + 1.5 * IQR

STEP 2



Control Web



Test Web

STEP 3

Control web:

- Delta_t mean: 100.33 s
- Delta_t median: 69.0 s
- Delta_t variance: 18701.72 s²
- Delta_t standard deviation: 136.75 s
- Delta_t IQR: 78.00 s



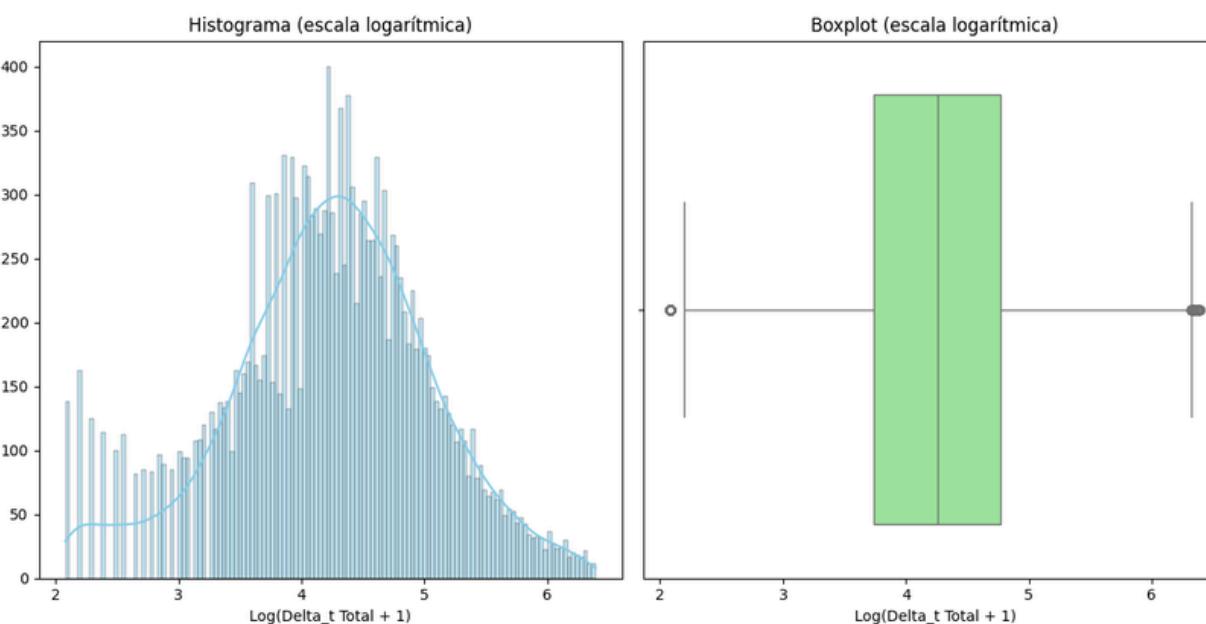
Test web:

- Delta_t mean: 102.60 s
- Delta_t median: 68.00 s
- Delta_t variance: 42011.35 s²
- Delta_t standard deviation: 204.96 s
- Delta_t IQR: 68.00 s

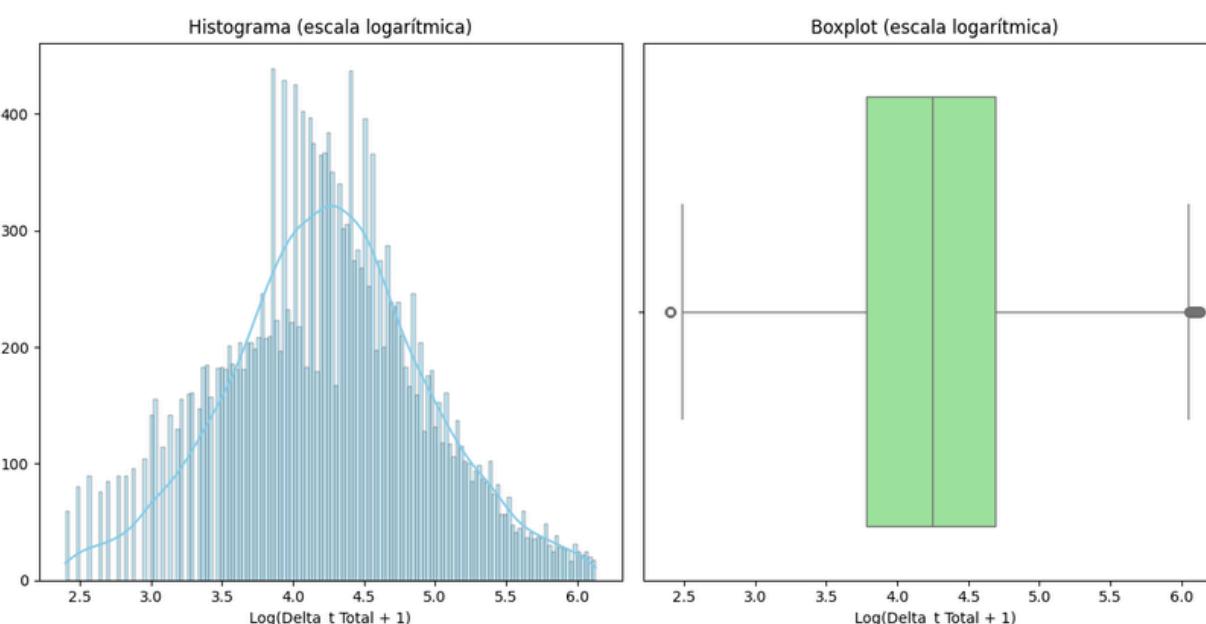
Steps to visualize it in a histogram and in a boxplot:

- Apply logarithmic transformation to the variable
- Eliminate null values
- Eliminate outliers:
 - lower_limit_control = Q1 - 1.5 * IQR
 - upper_limit_control = Q3 + 1.5 * IQR

STEP 3



Control Web



Test Web

CONFIRM

Control web:

- Delta_t mean: 135.94 s
- Delta_t median: 77.00 s
- Delta_t variance: 48996.37 s²
- Delta_t standard deviation: 221.35 s
- Delta_t IQR: 105.00 s

Test web:

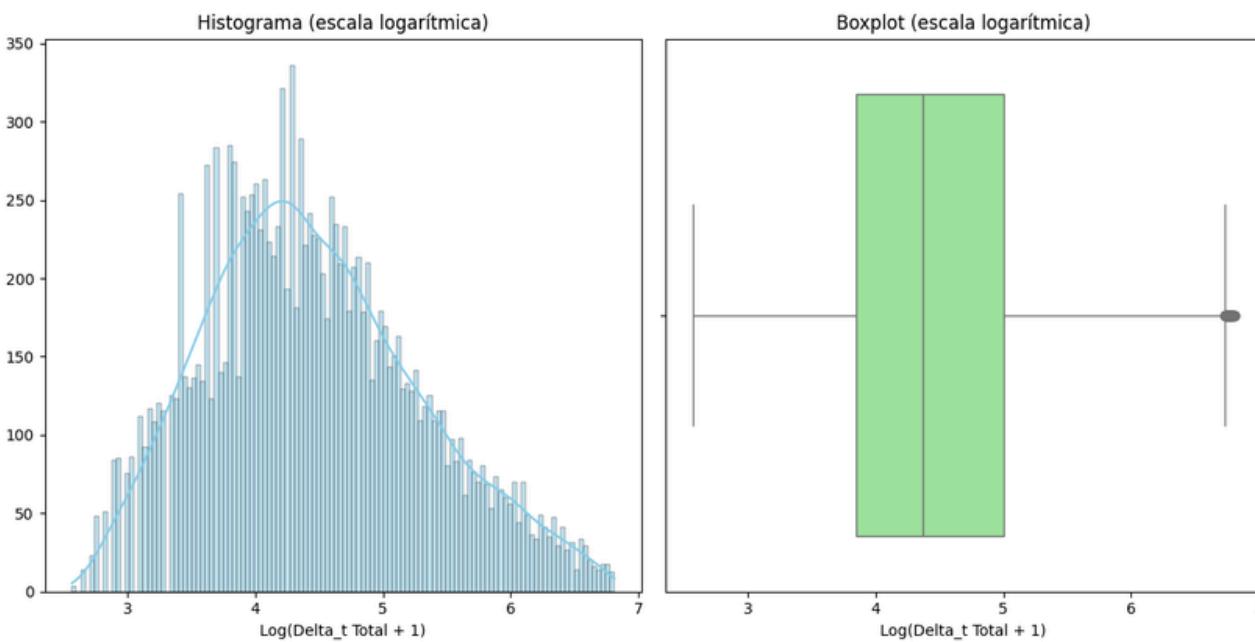
- Delta_t mean: 143.40 s
- Delta_t median: 59.00 s
- Delta_t variance: 128059.88 s²
- Delta_t standard deviation: 357.85 s
- Delta_t IQR: 98.00 s



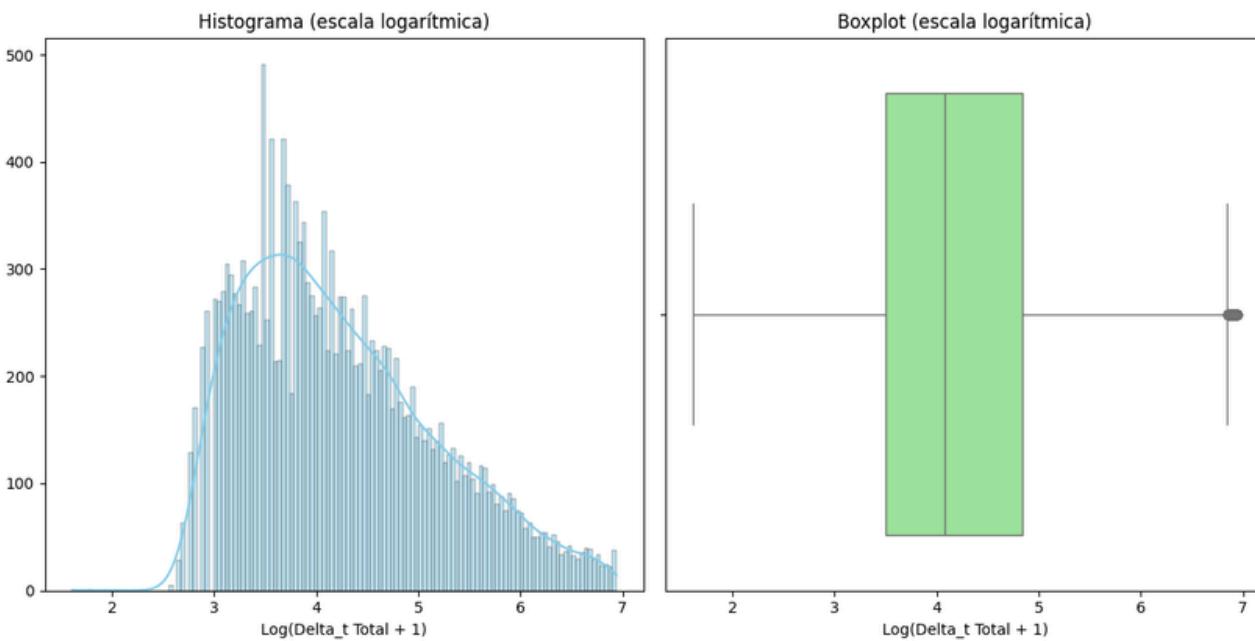
Steps to visualize it in a histogram and in a boxplot:

- Apply logarithmic transformation to the variable
- Eliminate null values
- Eliminate outliers:
 - lower_limit_control = Q1 - 1.5 * IQR
 - upper_limit_control = Q3 + 1.5 * IQR

CONFIRM



Control Web



Test Web

STATISTICAL HYPOTHESIS TEST

Shapiro-Wilk test:

- For all the steps: p-values < 0.05 --> Control and Test distributions are not normal

Mann-Whitney u test (Control vs Test):

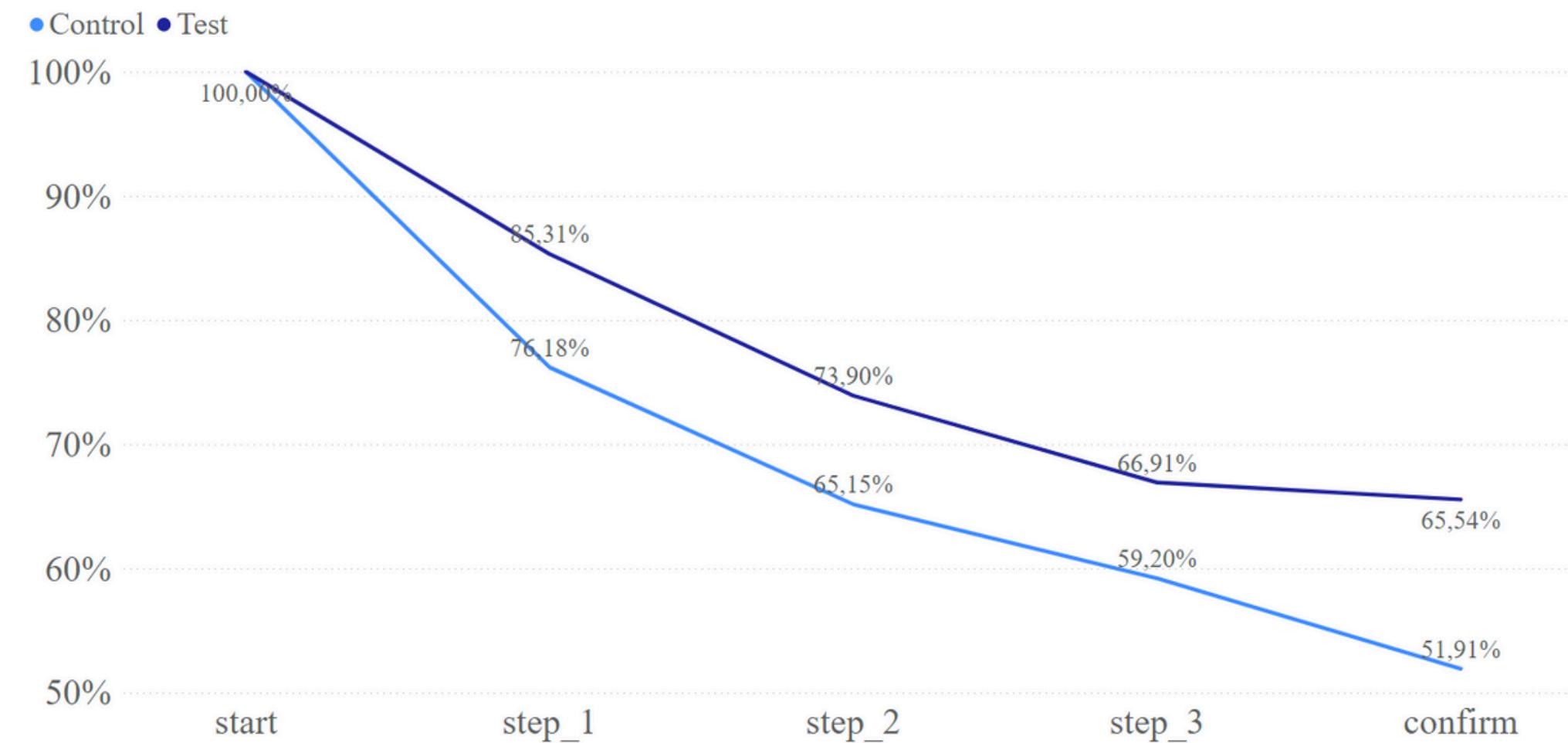
- Step 3: p-value > 0.05 --> There is no significant difference between the distributions.
- Rest of the steps: p-value < 0.05 --> There is a significant difference between the distributions.

Kolmogrov-Smirnov test (Control vs Test):

- For all the steps: p-values < 0.05 --> Control and Test have significantly different distributions.

FREQUENCY OF VISITS PER STEP

Frequency of visits by step (%)



Number of visits

- Control: 31K
- Test: 33K

Number of visits that completed the process:

- Control: 16K
- Test: 22K

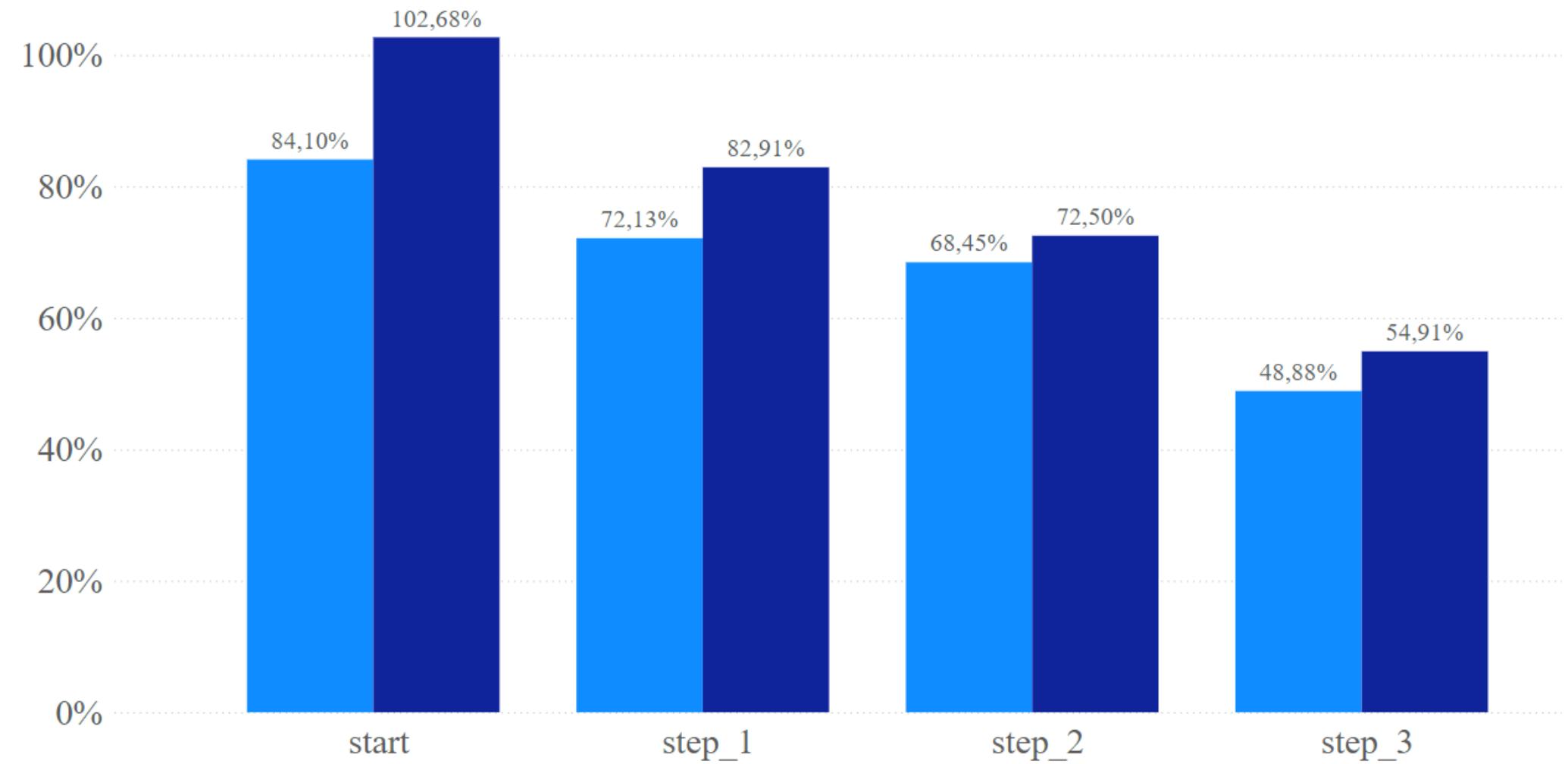
Decay of visits by step

- Control:
 - Step 1: -23.82%
 - Step 2: -11.03%
 - Step 3: -5.95%
 - Confirm: -7.29%
- Test:
 - Step 1: -14.69%
 - Step 2: -11.41%
 - Step 3: -6.99%
 - Confirm: -1.37%

NUMBER OF STEPBACKS PER STEP

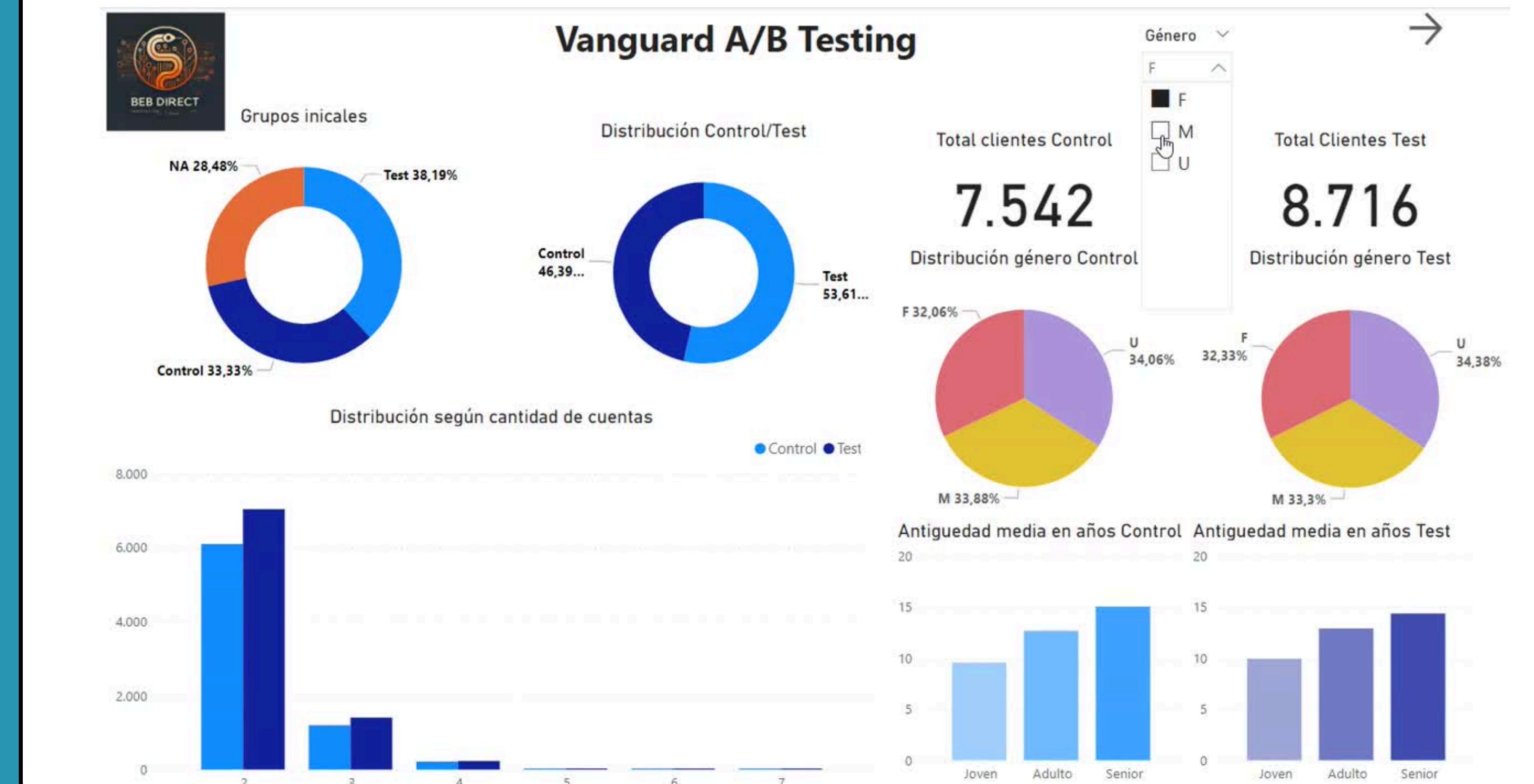
Backtracks (%)

• Control • Test



- **Test:** the number of setbacks from step 1 to start is **greater than 100%** --> more than 1 visit has been **set back more than 2 times**
- The percentage of setbacks is always **higher in test**

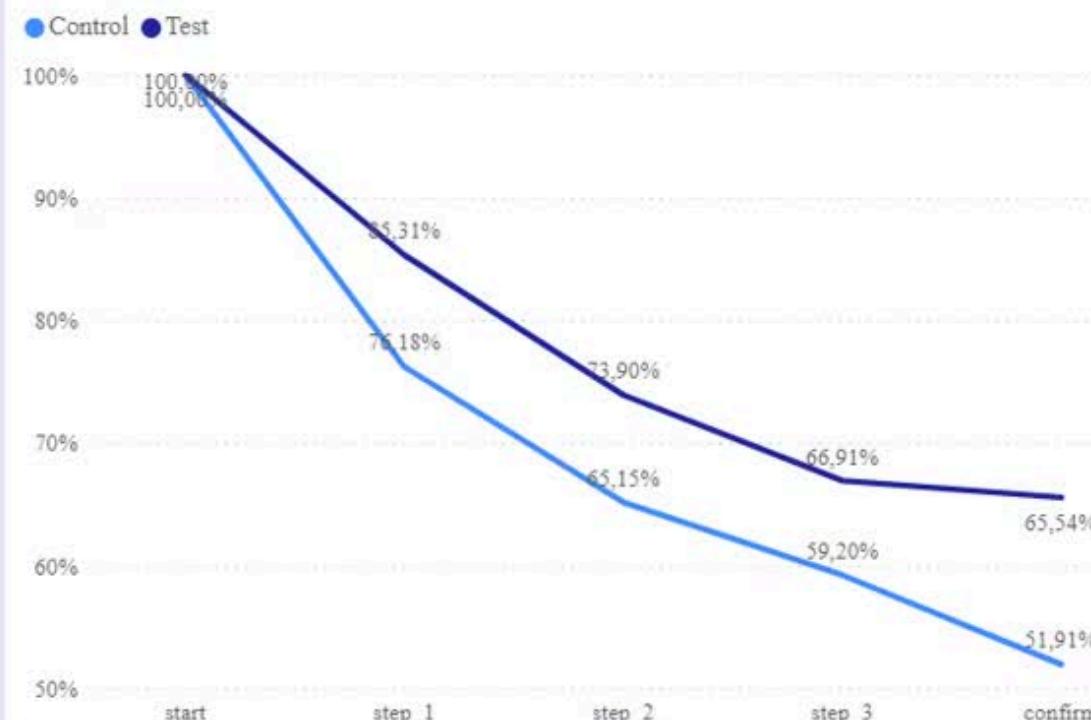
VISUALIZATION OF RESULTS IN POWER BI



VISUALIZATION OF RESULTS IN POWER BI

Test A/B Results

Frequency of visits by step



31K

Nº of visits Web Control

16K

Number of visits that completed the process in Web Control

33K

Nº of visits Web Test

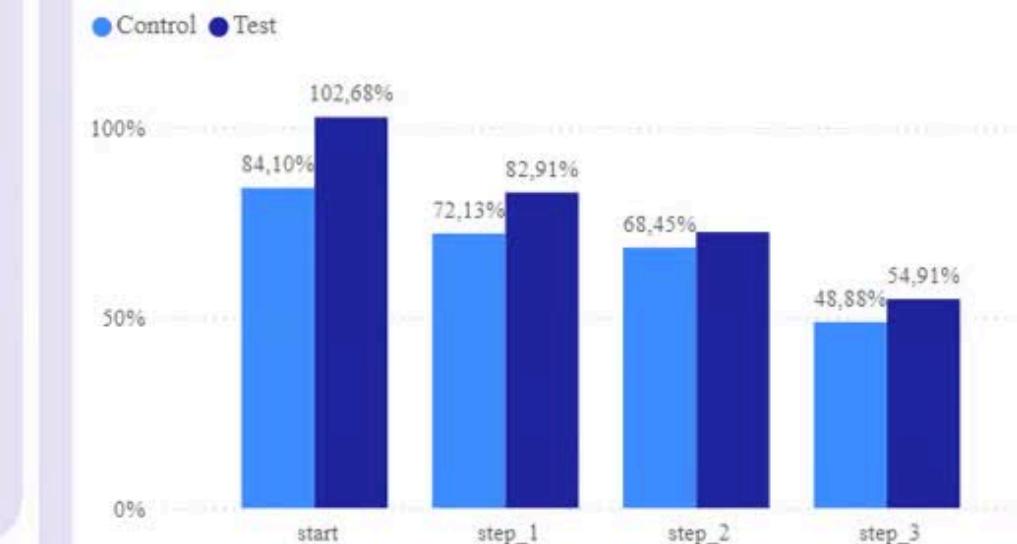
22K

Number of visits that completed the process in Web Test



- steps
- start
 - step_1
 - step_2
 - step_3
 - confirm

Backtracks (%)



More results



CONCLUSIONS ABOUT CALLS AND LOGONS

- It is demonstrated that significantly **fewer calls and logons are made in the test group** compared to the control.
- This may indicate a better user experience or a higher number of process completions without the need for additional calls or logons.



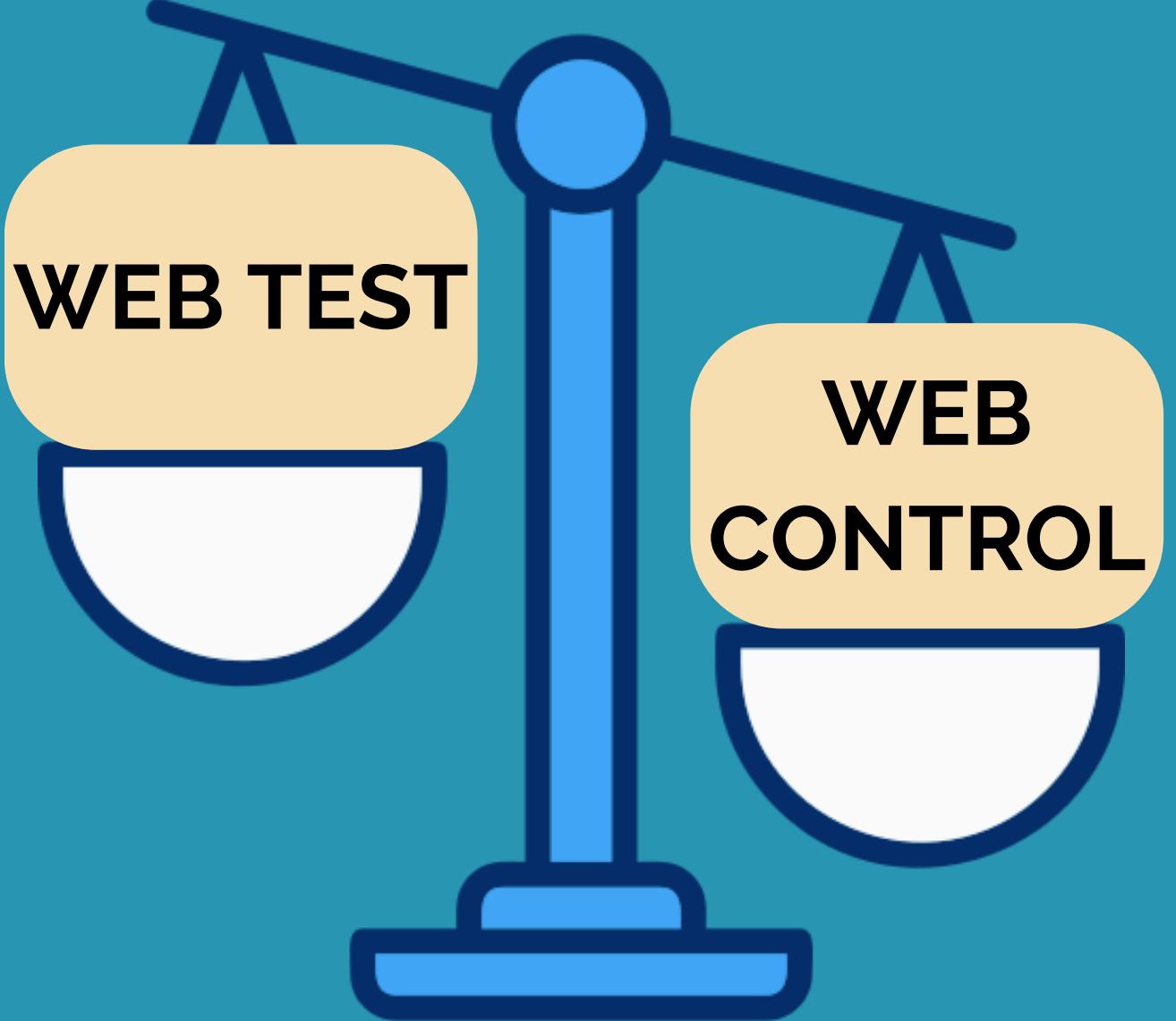
CONCLUSIONS ABOUT STEP DURATION

- The **total time** counting from the start step to the confirm step is on average **similar on both sites**
- The **time spent in the Start step and in Step 1** is **less** in the **Web Test** than in Control --> **Web Test more fluid in the first step**
- In **Step 2 and Step 3** the web is very **similar** in terms of time spent per visit, even in **Step 3 there is no significant difference with one of the hypothesis tests.**
- In the **Confirm step, less time** is spent on the **Control web**, but it could be that the **customer's unfamiliarity with the web test implies that he checks more the last step before confirming the process.**



CONCLUSIONS ABOUT THE PERCENTAGE OF VISIT FREQUENCIES PER STEP

- The percentage of visits that complete the process is higher on the Test site than on the control site.
- The abandonment of the process is much higher in the control web in step 1 and in the confirmation step.
- The abandonment of the process is slightly higher in the test web in steps 2 and 3, it would be a future improvement to make these steps clearer.
- In general the slope of percentage drop is lower in Test than in control.



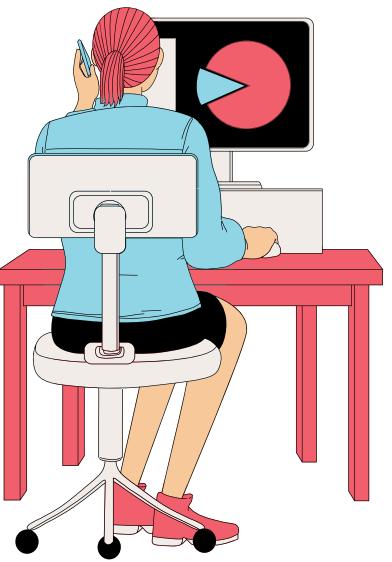
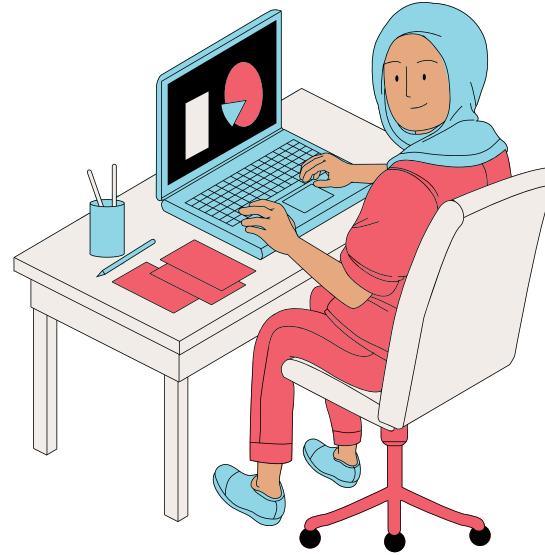
CONCLUSIONS ABOUT PERCENTAGE OF STEPBACKS PER STEP

- The **percentage of regressions to previous steps is higher** in the **Web Test** in **all steps**, but we believe that this is normal for a new website that is not yet known to customers.



**THANKS FOR
LISTENING! ANY
QUESTIONS?**

RESOURCE PAGE



RESOURCE PAGE



**GOOD LUCK &
HAVE FUN!**



Please reach out to your instructor with
any questions about this assignment.