



# E-NEWS EXPRESS PROJECT

STATISTICAL & VISUAL ANALYSIS

# Contents

- Business Problem Synopsis
- Data Overview
- Exploratory Data Analysis (EDA)
- Recommendations Based on EDA
- Key Questions to Answer
- Statistical analysis and Visual Analysis
- Conclusion

# Business Problem Synopsis

- E-news Express is an online news portal
- Have conducted an a/b test to analyze if a new feature will attract more subscribers

## Solution Approach:

- Perform visual analysis and provide recommendations based on the key insights
- Perform statistical analysis and formulate hypothesis
- Establish inferences based on the p-value obtained

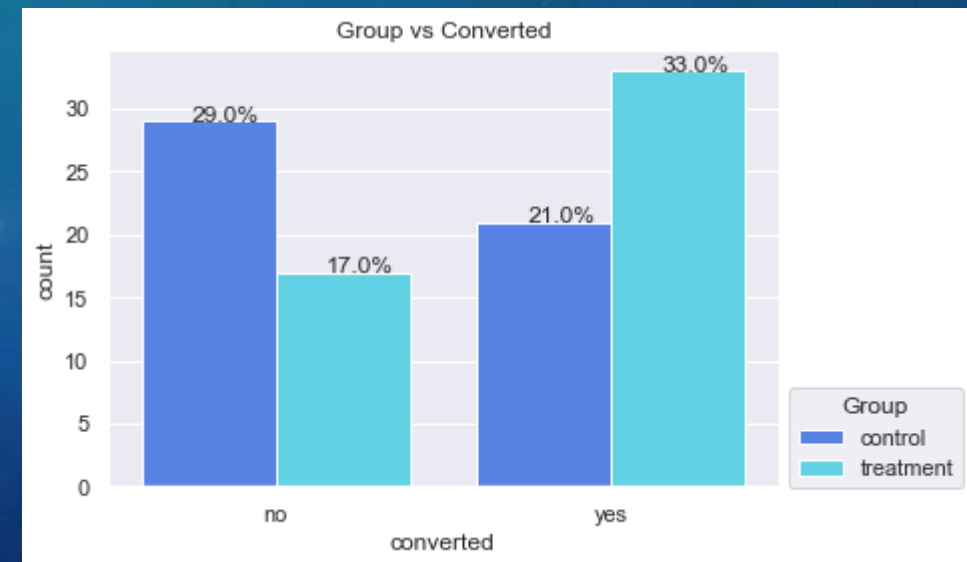
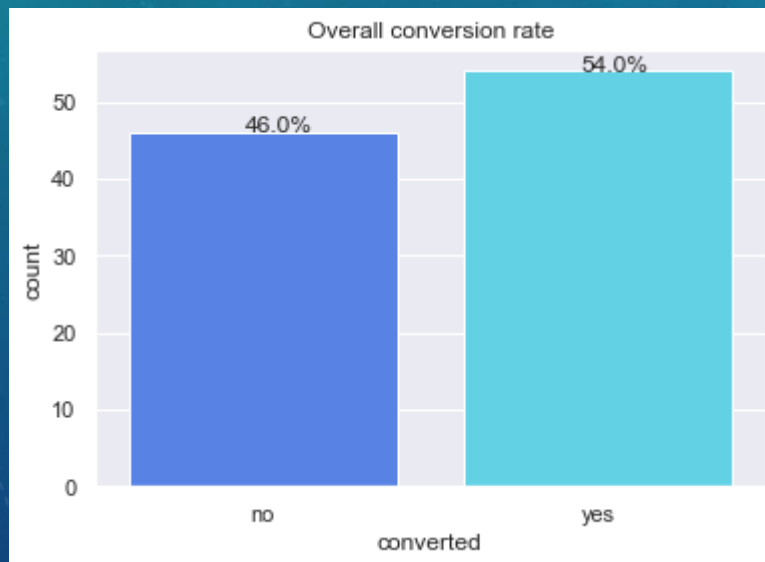
# Data Overview

- Data is based on the a/b test with randomly selected 100 users divided into two equal groups
  - Control group (old landing page)
  - Treatment group (new landing page)
- 100 observations on
  - User ID
  - Group
  - Landing page
  - Time spent on the page
  - Converted
  - Language preferred
- No missing values in the dataset
- *The level of significance ( $\alpha$ ) is 5%*

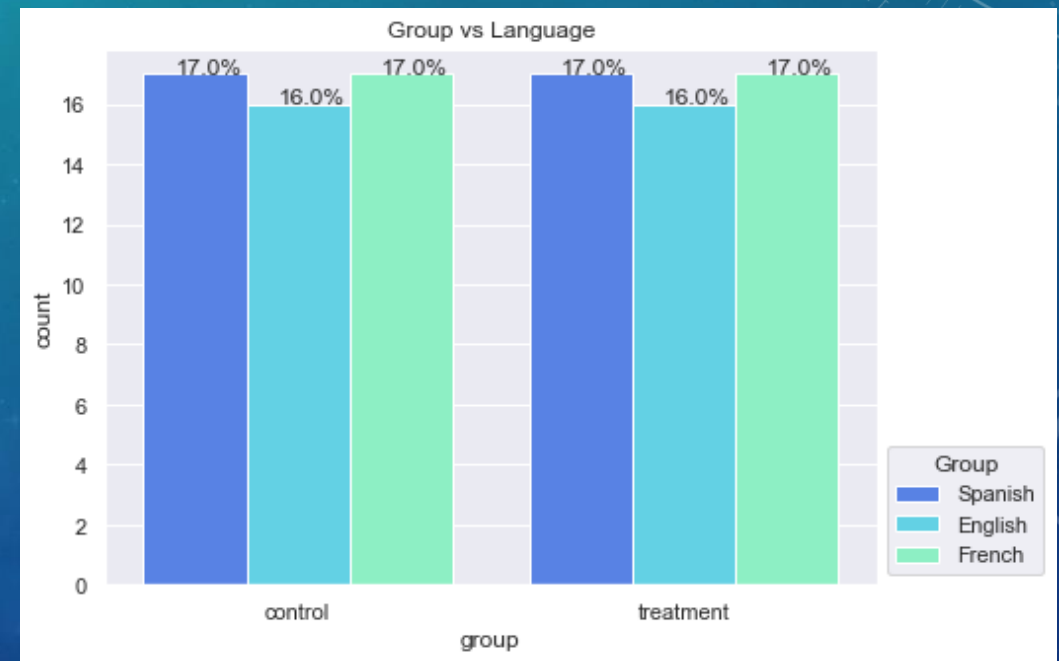
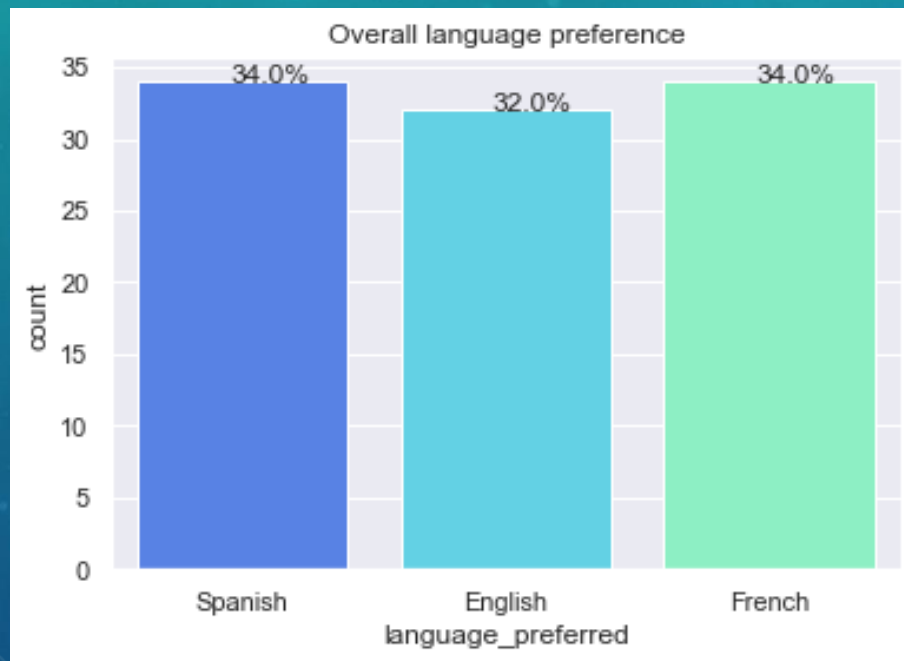


# Exploratory Data Analysis

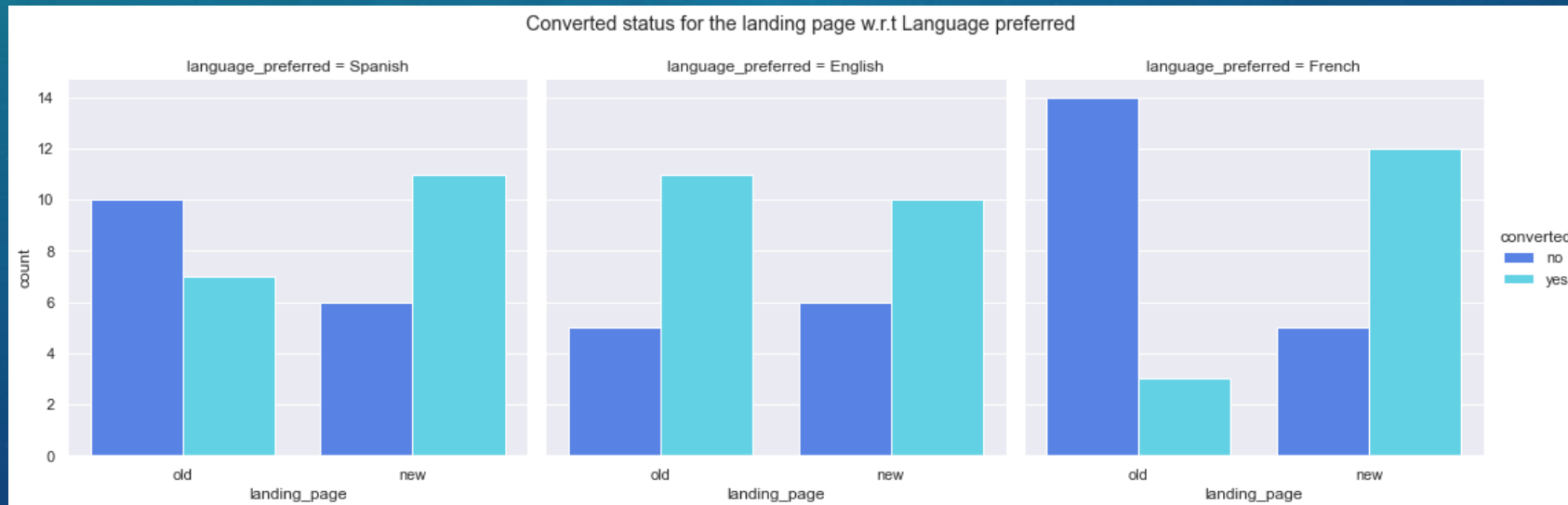
- Among the 54% new subscribers, 33% are from the treatment group
- As the new landing page was served to the treatment group, it can also be inferred that the conversion status is greater for the new landing page



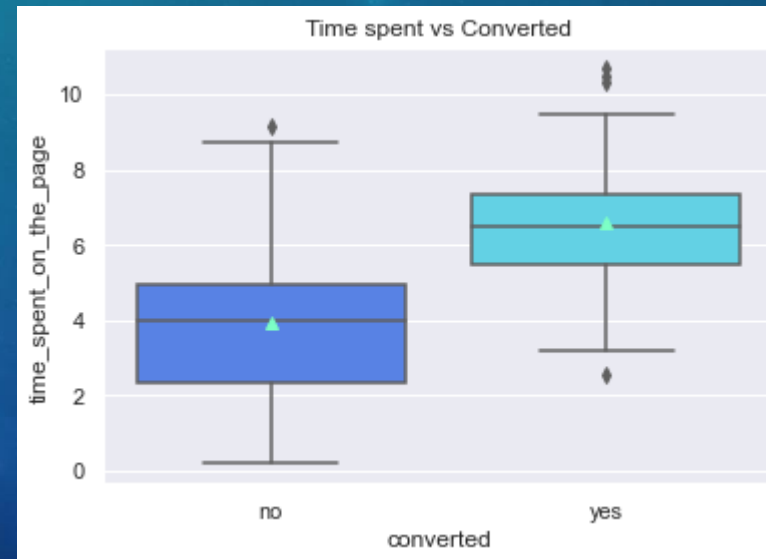
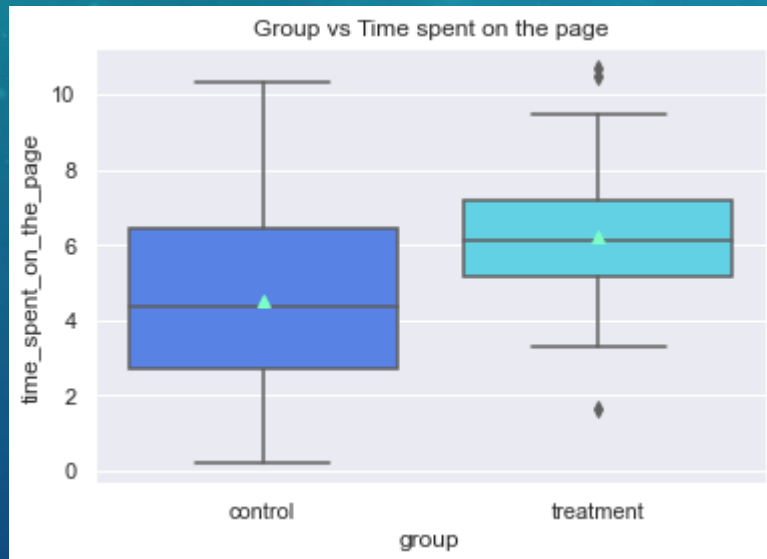
- The total number of users per language is uniform not only on the overall aspect but group-wise as well



- English language has the maximum converted users overall, followed by Spanish
- French language has the least number of converted users overall.
- However, French depicts highest conversion status for the new landing page than the old landing page
- Significant increase in the converted status of the new landing page for both Spanish and French



- The overall time spent on the news portal is greater for the new landing page in comparison
- There is also a high correlation between the time spent on the page by the user and his/her converted status





# Insights

- The user spends an average of 6 minutes on the news portal
- On average more time is spent on the new landing page
- We have subscribers from both new as well as old landing pages
- The users' language preference is equal for both new and old landing pages
- On average, large number of converted users prefer English language
- On average, least number of converted users prefer French
- However, French language depicts a higher conversion ratio for the new landing page
- Conversion status is almost 50-50 for a Spanish language user
- The more time a user spends, he/she is likely to become a subscriber, regardless of the landing page/language preferred

# Recommendation Based on EDA

*Based on the Exploratory Data Analysis, we have enough visual evidence to conclude that the business will improve with the launch of the new landing page.*

# Key Questions to Answer

1. Do the users spend more time on the new landing page than the old landing page?
2. Is the conversion rate (the proportion of users who visit the landing page and get converted) for the new page greater than the conversion rate for the old page?
3. Does the converted status depend on the preferred language?
4. Is the mean time spent on the new page same for the different language users?



*Do the users spend more time on the new landing page than  
the old landing page?*



# Statistical Analysis

- Formulated the hypothesis

$$H_0 : \mu_1 \leq \mu_2$$

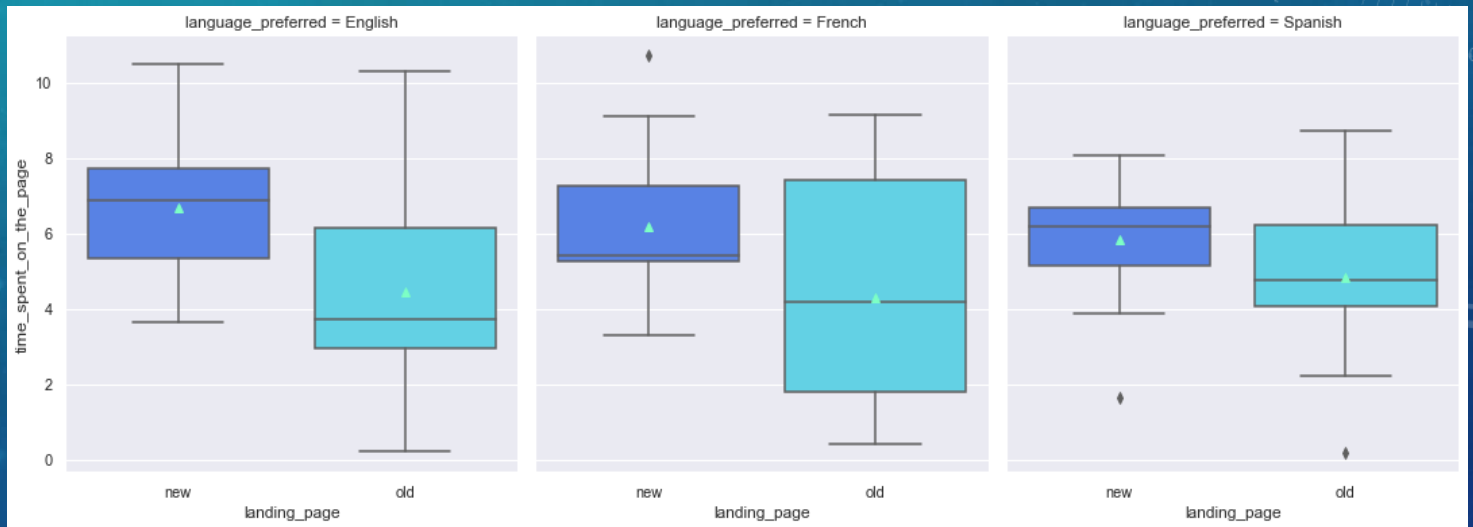
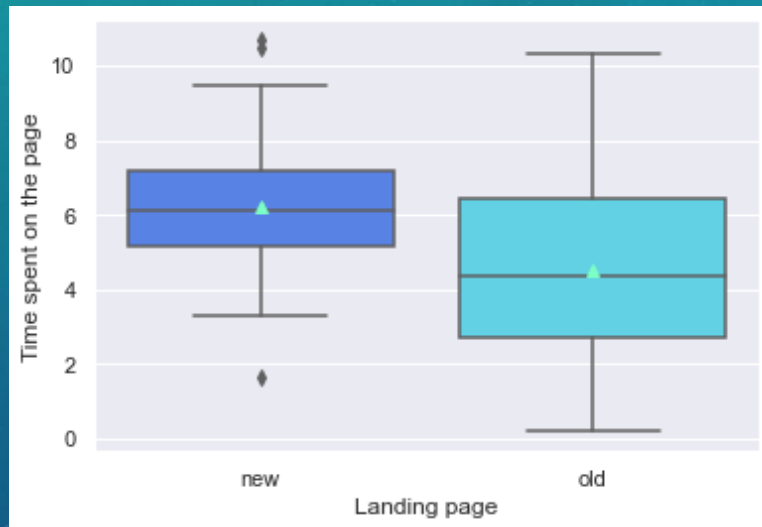
$$H_a : \mu_1 > \mu_2$$

where,  $\mu_1 = \text{time spent on the new landing page}$   
 $\mu_2 = \text{time spent on the old landing page}$

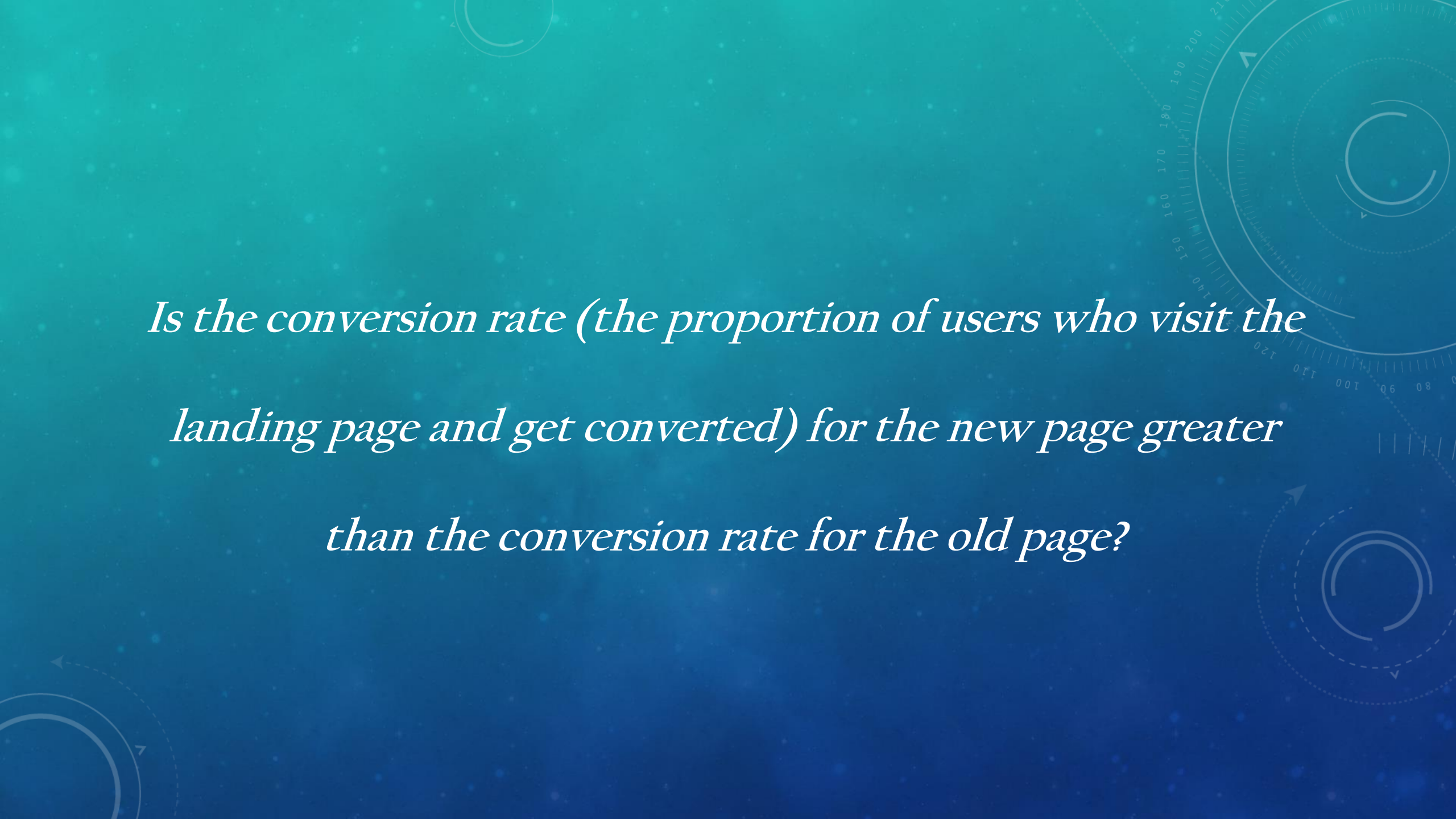
- Checked for the assumptions and then performed the *Two Independent Sample T-test for Equality of Means - Unequal Standard Deviations*
- Obtained a p-value of ( $\sim 0.0001$ ) which is much less than 0.05 and therefore rejected the Null Hypothesis
- Hence, *at 5% significance level we have enough statistical evidence to prove that the mean time spent on the new landing page is greater than the old landing page*

# Visual Analysis

- Mean time spent on the new landing page is around 6 minutes
- Mean time spent on the old landing page is around 5 minutes
- Mean time spent on the new landing page is higher regardless of the languages



*We have enough visual evidence to conclude that the mean time spent on the new landing page is higher compared to the old landing page*



*Is the conversion rate (the proportion of users who visit the landing page and get converted) for the new page greater than the conversion rate for the old page?*

# Statistical Analysis

- Formulated the hypothesis

$$H_0 : p_1 \leq p_2$$

$$H_a : p_1 > p_2$$

*$p_1$  = proportion of users who visit the new landing page and get converted*

*$p_2$  = proportion of users who visit the old landing page and get converted*

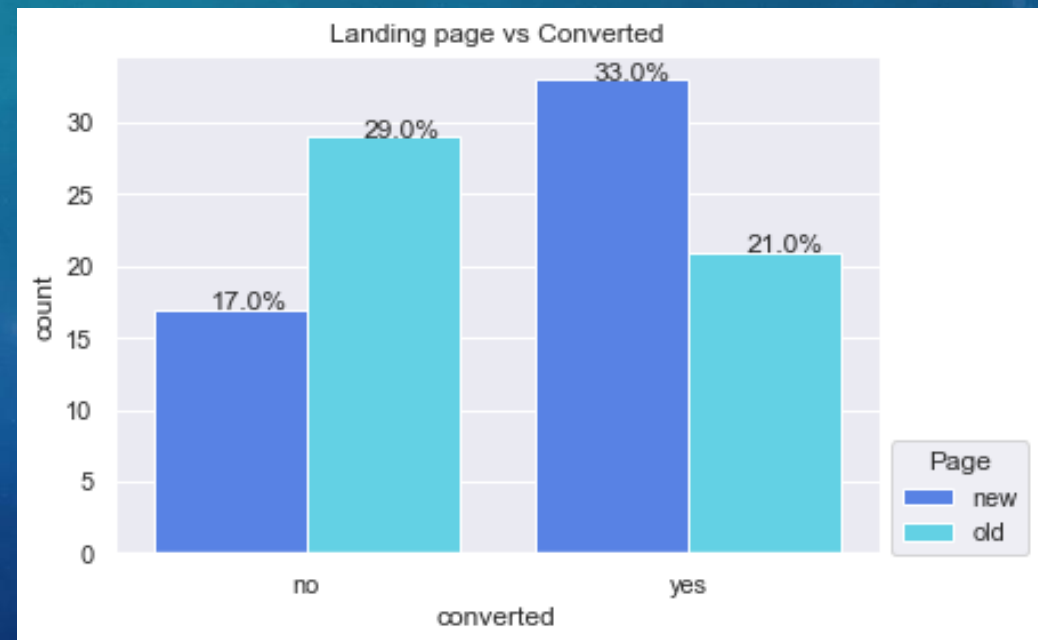
- Checked for the assumptions and then performed the **Two Proportions Z-Test**
- Obtained a p-value of ( $\sim 0.008$ ) which is much less than 0.05 and therefore rejected the Null Hypothesis
- Hence, *at 5% significance level we have enough statistical evidence to prove that the conversion rate of users is greater for the new landing page*



# Visual Analysis

- Out of 50 observations for the new landing page, 33 users have become subscribers to the news portal.
- Out of 50 observations for the old landing page, 21 users have become subscribers.

*We have enough visual evidence to conclude that the conversion rate of users is greater for the new landing page*





*Does the converted status depend on  
the preferred language?*

# Statistical Analysis

- Formulated the hypothesis

Ho : Converted status is independent of language preferred

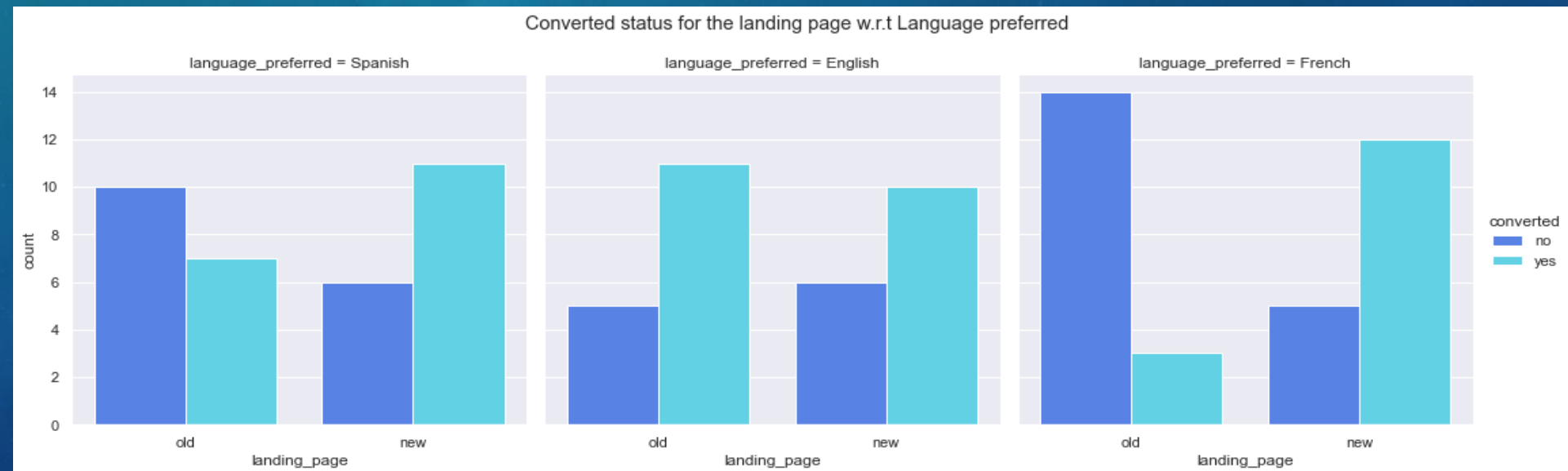
Ha : Converted status depends on the language preferred

- Checked for the assumptions and then performed the *Chi-Square Test for Independence*
- Obtained a p-value of ( $\sim 0.213$ ) which is much greater than 0.05 and therefore failed to reject the Null Hypothesis
- Hence, *at 5% significance level we have enough statistical evidence to prove that the converted status is independent of the language preferred*

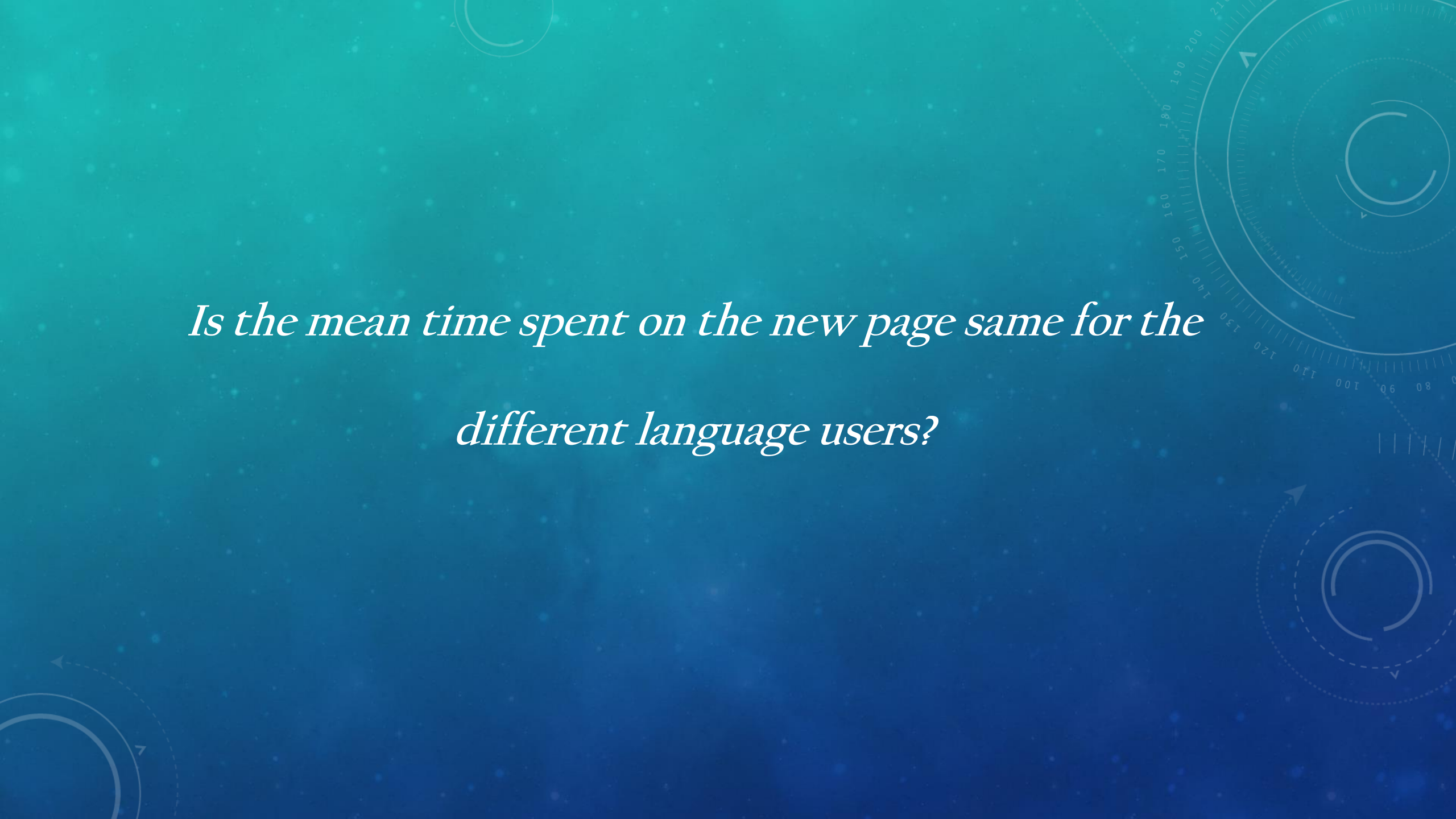
# Visual Analysis

- Overall 18 out of 34 Spanish language users have subscribed to the news portal
- 21 out of 32 English language users have subscribed
- Only 15 out of 34 French language users have subscribed
- However, French language has the highest conversion ratio for the new landing page
- There is little to no correlation between converted status and language preferred

***We have enough visual evidence to conclude that the converted status does not depend on the language preferred***







*Is the mean time spent on the new page same for the  
different language users?*

# Statistical Analysis

- Formulated the hypothesis

Ho : The mean time spent on the landing page w.r.t. the languages are all equal

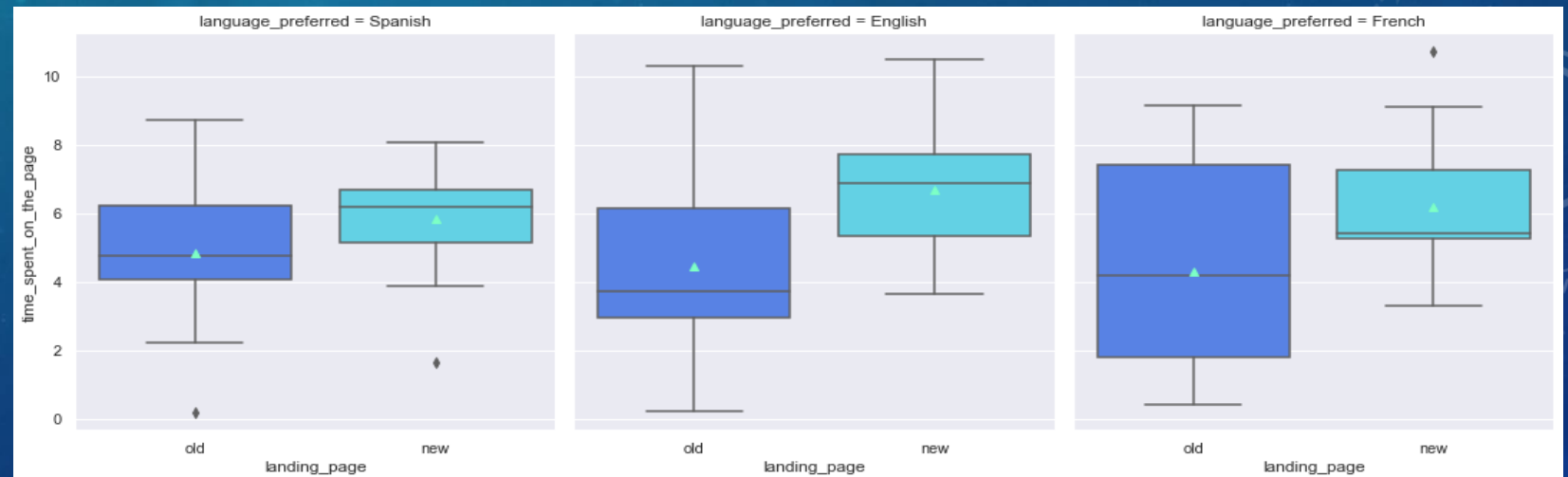
Ha : The mean time spent w.r.t. at least one of the three languages is different

- Tested for the assumptions and then performed the One-way ANOVA F-test
- Obtained a p-value of ( $\sim 0.43$ ) which is much greater than 0.05 and therefore failed to reject the Null Hypothesis
- Hence, *at 5% significance level we have enough statistical evidence to prove that the mean time spent on the new page for the different language users are all equal*

# Visual Analysis

- The mean time spent on the new landing page for all three languages is between 6 to 7 minutes
- The mean time spent on the old landing page for different languages is between 5 to 6 minutes
- The mean time spent for different languages are quite close to one another w.r.t. the landing page

*We have enough visual evidence to conclude that the average time spent for the new page is similar for the different language users*



# Conclusion

From the statistical and visual analysis performed, we can conclude that at a 5% significance level . . .

- The users spend more time on the new landing page than the old landing page
- The conversion rate for the new page is greater than the conversion rate for the old page
- The converted status does not depend on the language preferred
- The mean time spent on the new page is the same for the different language users

*All things considered, there is enough statistical and visual evidence to prove that the business would improve significantly with the launch of the new landing page.*





Thank You