New Landing Page Analysis

Business Statistics - Pam Lozano
October 14, 2023

Objective

Analyzing new landing page for effectiveness of getting new subscribers.

Key Focus Areas:

1. Assess the impact of a new landing page design on user engagement and conversion rates.

2. Determine if the new landing page increases user time spent on the page.

3. Investigate if the new landing page leads to a higher conversion rate.

4. Explore the influence of user language preferences on these outcomes.

Contents/Agenda

- Data Information

- Business Problem Overview

- Solution Approach

- Data Background and Contents

- EDA - Univariate Analysis

- EDA - Bivariate Analysis

- Hypotheses Tested and Results

- Executive Summary

  - Conclusions

  - Recommendations

- Further Analysis

Data Information

The data contains information regarding the interaction of users in both groups with the two versions of the landing page.

Variable	Description
user_id	Unique user of the person visiting the website
group	Whether the user belongs to the first group (control) or the second group (treatment)
landing_page	Whether the landing page is new or old
time_spent_on_the_page	Time (in minutes) spent by the user on the landing page
converted	Whether the user gets converted to a subscriber of the new portal or not
language_preferred	Language chosen the user to view the landing page

Business Problem Overview

- E-news Express, an online new portal, is facing a decline in new monthly subscribers.

- The company believes that the current landing page design may not effectively engage users to subscribe.

- To address this, an A/B test was conducted, comparing a new landing page to the existing page.

- Data on user interactions, time spent, conversions, and language preferences were collected for analysis.

- The goal is to determine if the new landing page design increases user engagement and subscription rates.

Solution Approach

- **Variable Selection:** Identify key variables, including landing_page, time_spent_on_the page, converted, and language_preferred, relevant to the analysis.
- **Data Cleaning:** Address missing values and duplicates in these selected variables to ensure data quality.
- **Exploratory Data Analysis (EDA):** Conduct EDA to explore user behavior on the old and new landing pages. Compare time_spent_on_the_page and conversion rates between the two groups. Analyze language_preferred to understand its impact on conversions.
- **Hypothesis Testing:** Perform statistical tests, such as a t-test for time_spent_on_the_page and a proportions test for conversion rates, to draw valid conclusions.
- **Visualization:** Create visualizations, like histograms and box plots for time spent and stacked bar charts for conversion rates by language, to communicate findings effectively.
- **Interpretation:** Interpret statistical results to determine the effectiveness of the new landing page in terms of user engagement and conversion.
- **Recommendations:** Based on insights, provide actionable recommendations for improving the landing page and increasing conversions.

Data Background and Contents

- There are 100 rows and 6 columns.

- The datatypes of the different columns integer, object, and float.

- There are no missing values.

- The average time spent on the page is 5.37 minutes.

- The median time spent on the page is 5.42 minutes.

- The least amount of time spent on the page is 19 seconds.

- The most amount of time spent on the page is 10.71 minutes.

Exploratory Data Analysis

# Univariate Analysis - Time Spent on the page
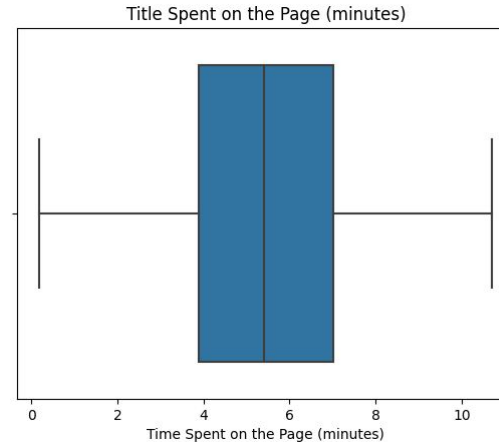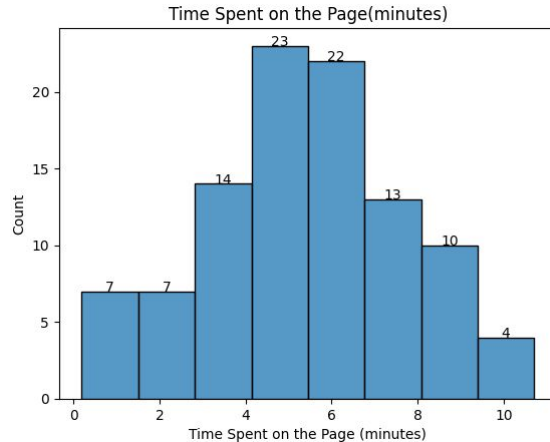
- There are no outliers in this distribution.
- The distribution appears to be evenly distributed.
- The mean time spent on the page is 5.37 minutes, with a standard deviation of 2.38 minutes.
- The median time spent on the page is 5.42 minutes.
- The least amount of time spent on the page is 19 seconds.
- The most amount of time spent on the page is 10.71 minutes.
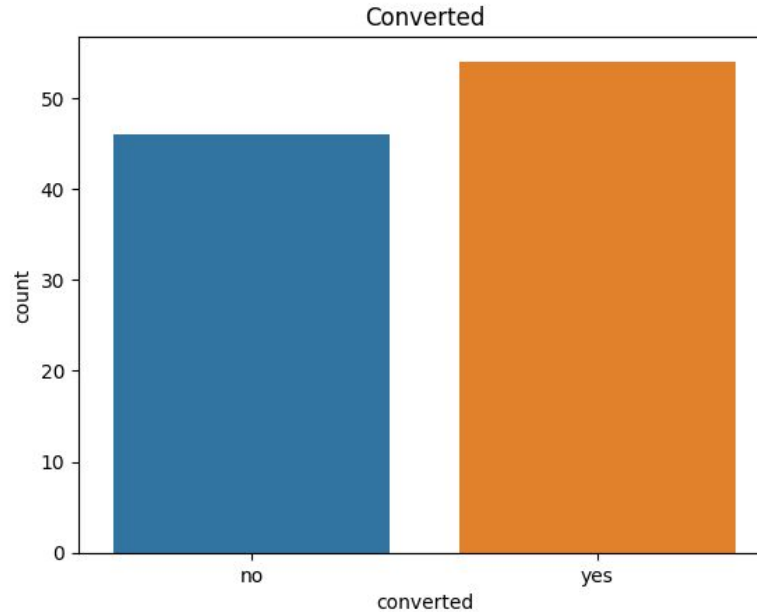
Univariate Analysis - Group
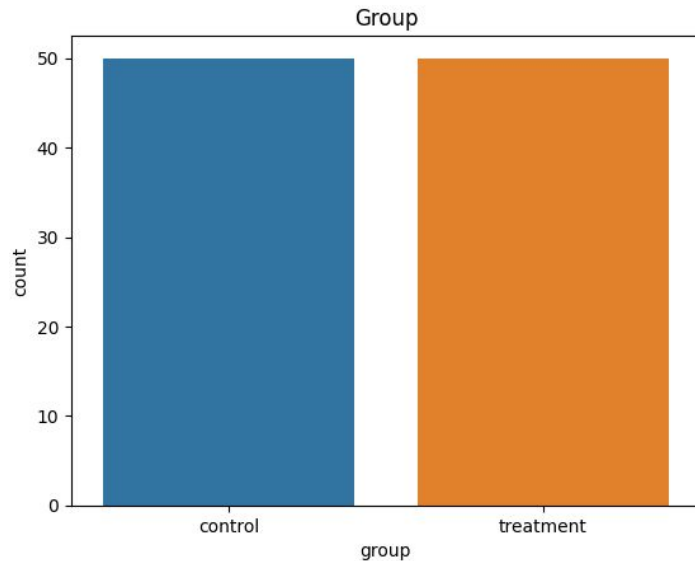
- The group represents the two groups: control and treatment.
- There are an equal number of observations in both groups, with each group having 50 users.
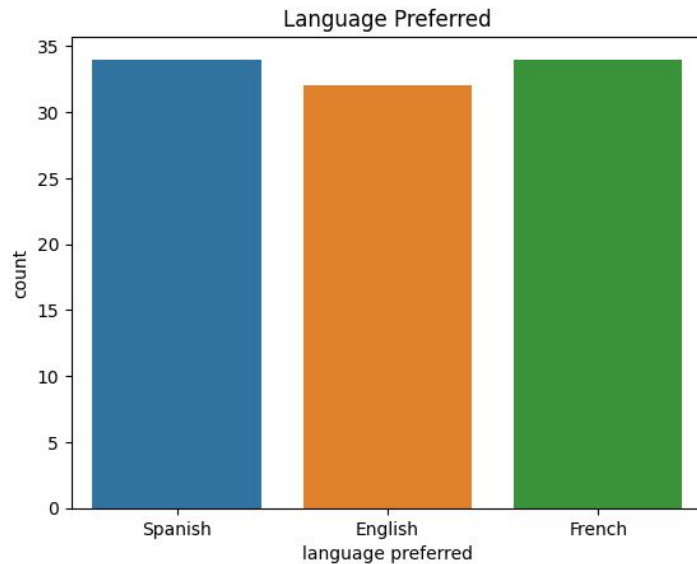
Univariate Analysis - Landing Page
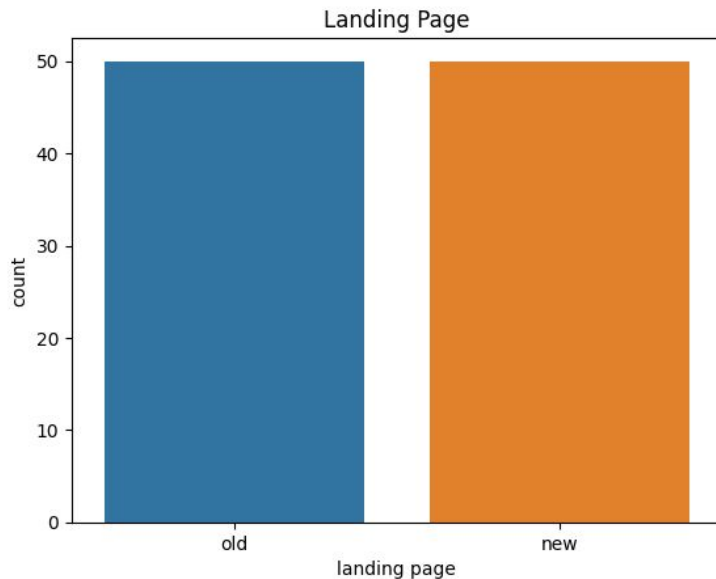
- There is a balanced distribution, with an equal number of users. Each group consists of 50 users to both the old and new landing page.

Univariate Analysis - Converted

- The data reflects that 54 users converted while 46 did not.
- This shows the success rate of the landing pages in converting visitors to subscribers.

Univariate Analysis - Language Preferred

- The data shows that there are three language preferences: Spanish, English, and French with counts of 34, 32, and 34 respectively.
- The distribution of language preferences among users appears to be relatively balanced.

Bivariate Analysis - Landing Page vs Time Spent on the Page


Time Spent on the Page vs Landing Page
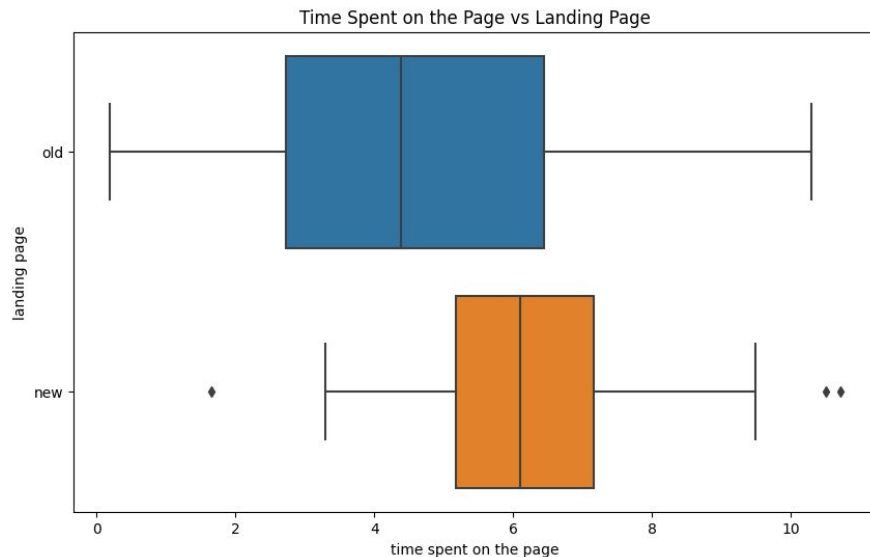
- The new landing page indicates outliers are present in the distribution.
- The distribution of time spent on the new and old landing page seem to differ.
- The median amount of time spent on the new landing page is more than the old page.
- The old landing page reflects the data to be more dispersed than the new landing page.

Bivariate Analysis - Conversion Status vs Time Spent on the Page

Time Spent on the Page vs Conversion Status
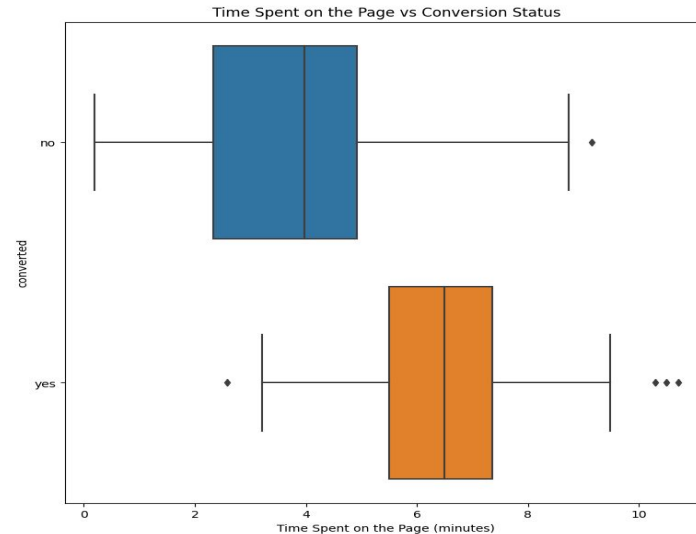
- Both distributions show outliers are present.
- The new landing page impacts more users to convert to subscribers.
- The distributions show a difference between conversion status and the amount of time spent on the page.
- The old landing page appears to be right skewed, whereas the new landing page appears to be symmetric.

Bivariate Analysis - Language Preferred vs Time Spent on the Page


Time Spent on the Page vs Language Preferred
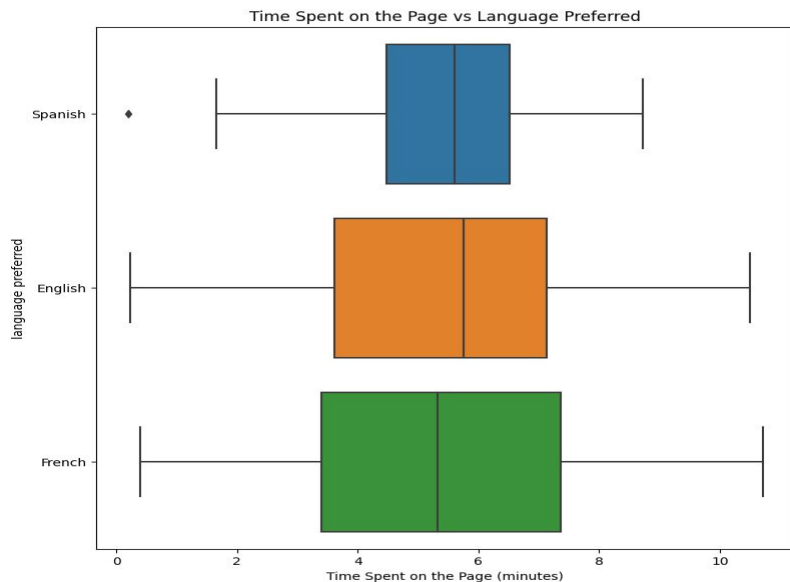
- The landing page translated into Spanish shows some outliers.
- The distribution of languages do not seem to differ among the 3 languages.

Hypotheses Testing

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

- Null hypotheses (H0): The mean time spent on the new landing page is <= the mean time spent on the existing landing page (μ_new <= μ_old) (μ represents the population mean)
- Alternative hypotheses (H1): The mean time spent on the new landing page > the mean time spent on the existing landing page (μ_new > μ_old)

- The test used is a one-tailed t-test to compare the means of two groups (new and old landing pages)

- The p-value = 0.00

- The sample standard deviation of the time spent on the new page is: 1.82

- The sample standard deviation of the time spent on the old page is: 2.58

Hypotheses Tested and Results


Time Spent on New Landing Page
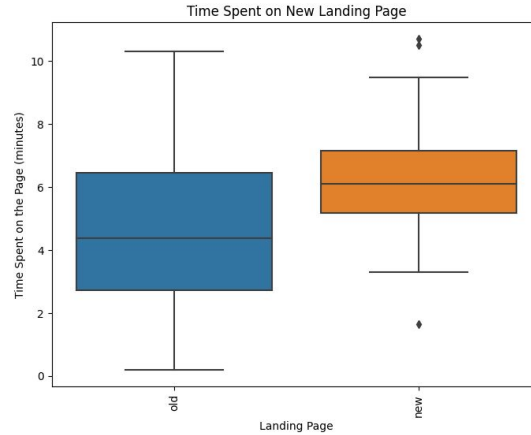
- The p-value 0.00 is less than the level of significance, we then reject the null hypothesis.

- There is enough evidence to reject the null hypothesis. In other words, there is a statistically significant difference in the time spent on the new landing page 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?

- Null Hypotheses (H0): The conversion rate for the new landing page is less than or equal to (<=) the conversion rate for the old landing page (p_new <=p_old)

- Alternative hypotheses (H1): The conversion rate for the new landing page is greater than (>) the conversion rate for the old landing page (p_new > p_old)

- The test used was a two-proportion z-test to compare the proportions of two groups (conversion rates)

- The p-value is 0.01.

- The numbers of users served the new and old pages are 50 and 50 respectively.

Hypotheses Tested and Results


Control (Old) vs Treatment (New)
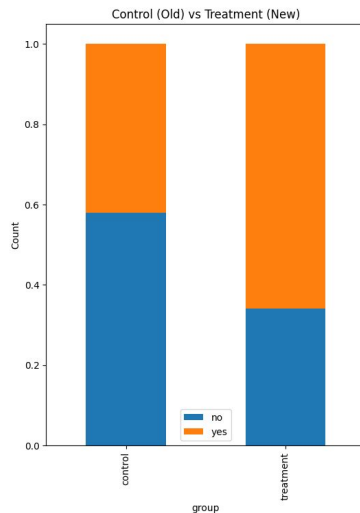
- The p-value 0.01 is less than the level of significance, we reject the null hypothesis.

- This is a statistically significant result and suggests that the new landing page is more effective in terms of converting users compared to the old landing page.

Does the converted status depend on the preferred language?

- Null hypotheses (H0): The converted status is independent for the preferred language.
- Alternative hypotheses (H1): The converted status is dependent for the preferred language.
- A Chi-Square test of independence using a contingency table is suitable when there is categorical data and want to assess whether two categorical variables are independent or related.
- The p-value is 0.21

Hypotheses Tested and Results

Conversion Status vs Preferred Language
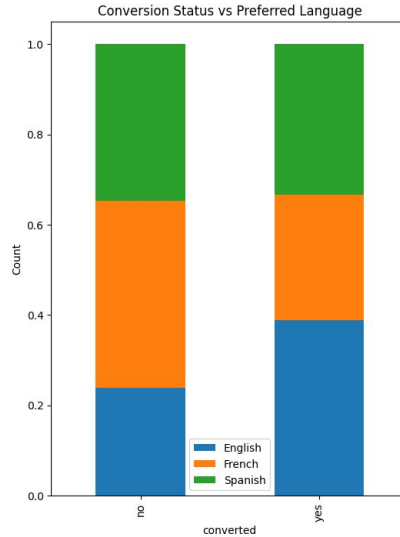
- The p-value 0.21 is greater than the level of significance, we fail to reject the null hypothesis.
- In this context, failing to reject the null hypothesis means that there is no statistically significant association between the preferred language and the conversion status.
- This result suggests that the preferred language choice does not play a significant role in whether a user gets converted or not.

Is the time spend on the new page same for the different language users?

- Null hypotheses (H0): The mean time spent on the new landing page is the same for all language preferences.

  Alternative hypotheses (H1): The mean time spent on the new landing page is not the same for all language preferences.

- The test used was a one-way ANOVA (Analysis of Variance) since there is one categorical independent variable (language) and one continuous dependent variable (time spent on the new page).

- The p-value is 0.43

Hypotheses Tested and Results
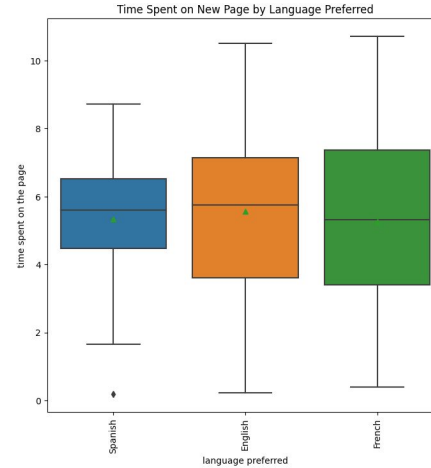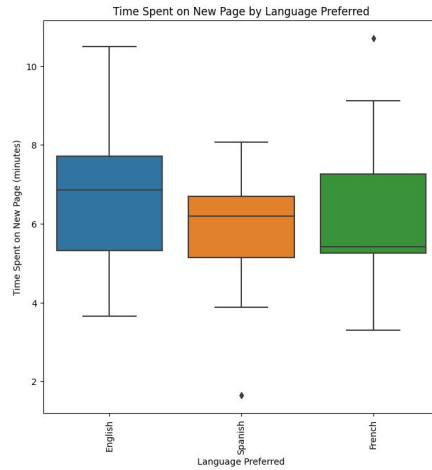
- The p-value is 0.43 is greater than the level of significance, we fail to reject the null hypothesis.

- Since we fail to reject the null hypothesis, the choice of preferred language does not appear to have a significant impact on the time spent on the new page.

Executive Summary

Executive Summary - Conclusions

- **Time Spent on New Page:** Users spend more time on the new landing page than the old one, indicating higher engagement.
- **Conversion Rate:** The conversion rate for the new page is significantly greater than the old page, suggesting it's more effective in acquiring new subscribers.
- **Language Preferences:** The preferred language does not significantly impact conversion rates, meaning the new page is equally effective across languages.
- **Time Spent by Language:** Users across different languages spend a similar amount of time on the new page, implying that the design appeals uniformly to diverse language groups.
- **Effectiveness of New Page:** Overall, the new landing page design is successful in engaging users and increasing conversion rates, aligning with the company's objective to attract more subscribers.
- **Further Investigation:** While the new page is promising, ongoing monitoring and potential refinements should be considered to optimize user engagement continually.

Executive Summary - Recommendations

- **Continual Monitoring:** Maintain regular tracking of user engagement and conversion rates on both landing pages to ensure the new page's effectiveness persists.
- **User feedback:** Collect user feedback to identify specific aspects of the new page that resonate positively and areas that may require improvement.
- **A/B Testing:** Consider further A/B testing with variations of the new page to fine-tune its design and content for maximum user engagement.
- **Multilingual Content:** Invest in creating relevant and engaging content in multiple languages to cater to a diverse audience
- **Mobile Optimization:** Ensure the landing page is mobile-friendly to capture users accessing the portal via smartphones and tablets.
- **Performance Metrics:** Establish clear Key Performance Indicators (KPIs) for user engagement and conversion rates to quantify the impact of page enhancements over time.

These recommendations aim to support the company's goal of increasing subscriber numbers and maintaining a user-friendly experience.