# E-news Express Business Presentation

Objective

Perform the statistical analysis to answer the following questions using the collected data:-

- Explore the dataset and extract insights using Exploratory Data Analysis.
- Do the users spend more time on the new landing page than 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?
- Does the converted status depend on the preferred language?
- Is the mean time spent on the new page same for the different language users?

*Consider a significance level of 0.05 for all tests.

The aim of performing statistical analysis to answer the above questions is to decide wether the new landing page is more effective to attract and bring in new subscribers for the online news portal.The answers of the questions obtained from the statistical analysis performed will help make a business decision.

Data Information
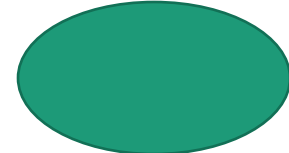
Variable	Description
user_id	This represents the user ID of the customer visiting the website.
group	This represents whether the customer belongs to the first group (control) or the second group (treatment).
landing_page	This represents whether the landing page is new or old.
time_spent_on_the_page	This represents the time(in minutes) spent by the customer on the landing page.
converted	This represents whether the customer converted or not.
language_preferred	This represents the language chosen by the customer to view the landing page.

| Observations	Variables
100	6

Note:
- There are no missing values in the dataset.
- The group, landing page, converted, and language_preferred columns have been converted to category data type leading to a reduced memory usage from 4.8+ KB to 2.6 KB

# Univariate Analysis – Time spent on the page

Time spent on the page

- There are no outliers for time_spent_on_the_page.
- The time_spent_on_the_page distribution shows an almost symmetrical distribution(zero skewness) that is a fairly normal distribution.
- Majority of users(represented by the 2 highest peaks) spend about 5 and 6 minutes on the landing page.

# Univariate Analysis – Group & Landing page
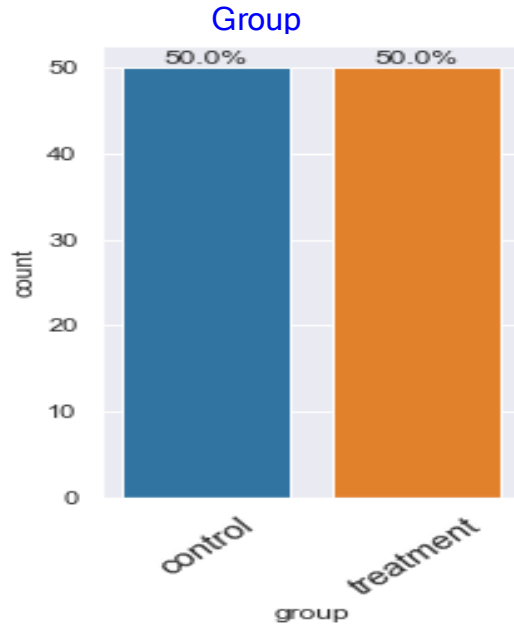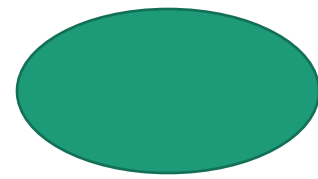
Group

Landing page

- There are 2 unique groups in this dataset -control and treatment.
- The observations are uniformly distributed across the groups.
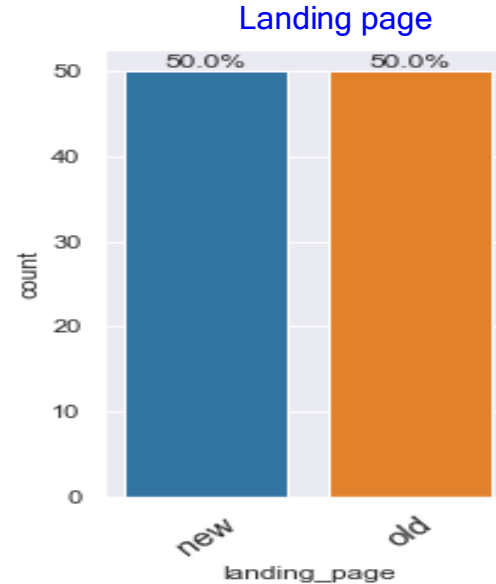
- There are 2 landing_pages - new and old.
- The observations are uniformly distributed across the landing pages.

Univariate Analysis – Converted status & Preferred lang.

Converted status
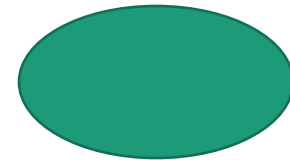

Preferred language
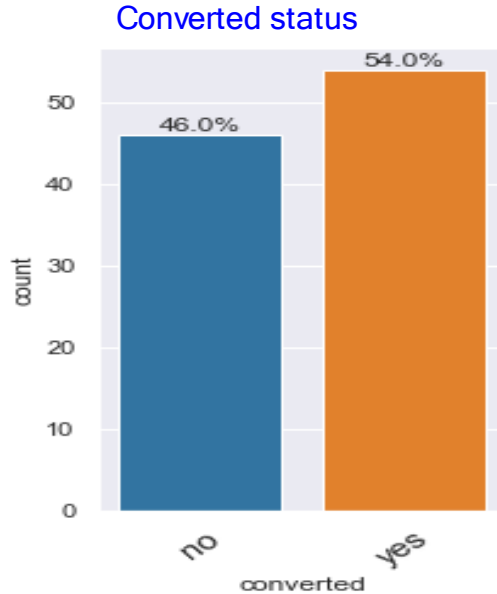
- 54% of the users get converted after visiting the landing page.
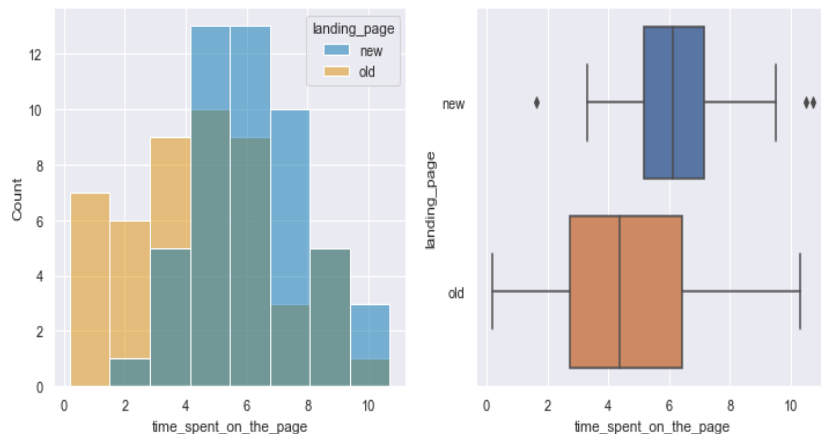- 46% of the users do not get converted after visiting the landing page.
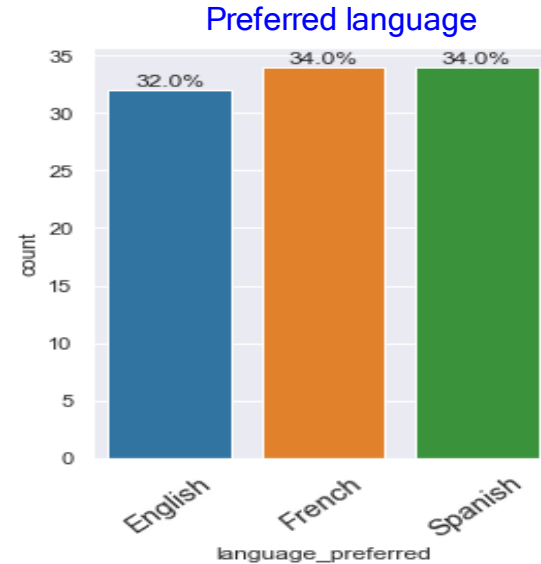
- There are 3 unique preferred languages - English, French, and Spanish.
- 34% of users choose French,34% of users choose Spanish and 32% of users choose English to view the landing page.

Bivariate Analysis

Landing page vs Time spent on the page

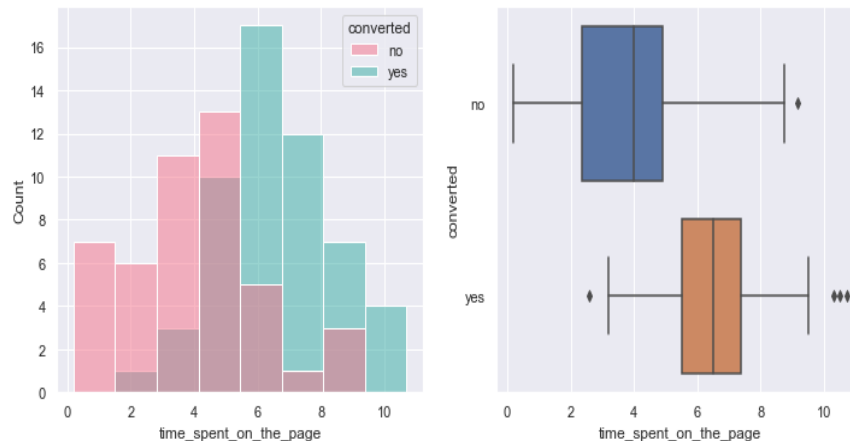## Converted status vs Time spent on the page

- Overall, the time spent on the new page seems to be greater than the time spent on the old page.

- Overall, the users who get converted seem to spend more time on the page.

7

Hypothesis Testing - Time spent across new and old page

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

- H_o: The mean time spent by the users on the new page is equal to the mean time spent by the users on the old page.

- H_a: The mean time spent by the users on the new page is greater than the mean time spent by the users on the old page.

By using the two samples independent t-test, we get the p-value as ~0.0001 that is < 0.05.

Therefore, we reject the null hypothesis that the mean time spent by the users on the new page is equal to the mean time spent by the users on the old page.
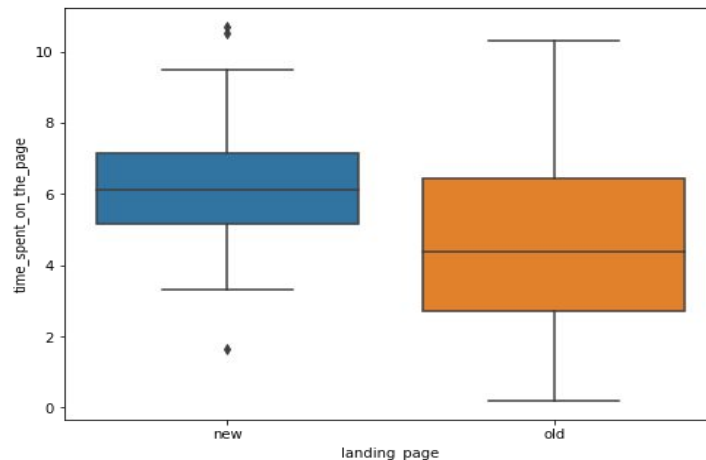
- Visually the difference between the time spent on the new page and old page is apparent. The time spent on the new page is comparatively greater than the time spent on the old page.
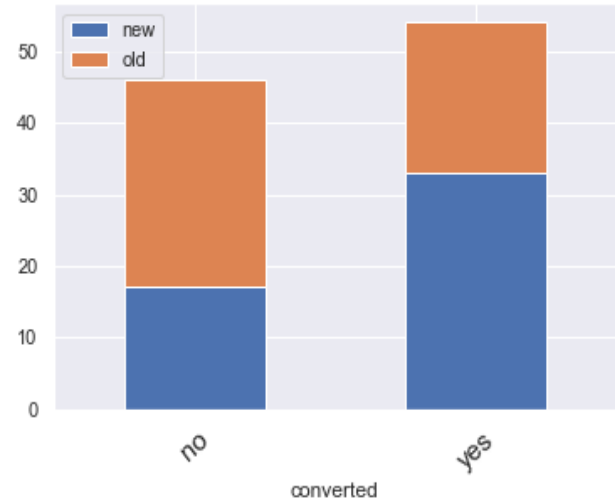
Hypothesis Testing - Conversion rate

Problem: 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?

- Ho: The conversion rate for the new page is equal to the conversion rate f or the old page

- Ha: The conversion rate for the new page is greater than the conversion r ate of the old page.

By using the two sample proportions test, we get the p -value as ~0.0080 that is < 0.05.

Therefore, we reject the null hypothesis that the conversion rate for the new page is equal to the conversion rate for the old page.

- Visually, the number of users who get converted is more for the new page than the old page.
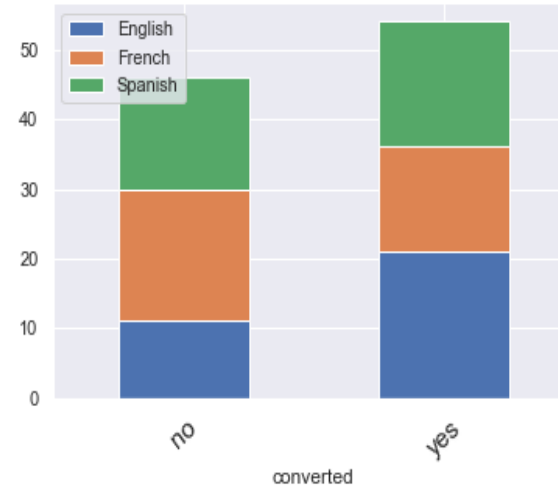
Hypothesis Testing - Converted status across languages

Problem: Is the conversion and preferred language are independent or related?

- H_o: The converted status is independent of the preferred language.

- H_a: The converted status is not independent of the preferred language.

By using the chi-square test of independence, we get the p-value as ˜0.2130 that is > 0.05.

Therefore, we fail to reject the null hypothesis that the converted status is independent of the preferred language.

- The distribution of conversion status across different preferred languages looks fairly similar.

Hypothesis Testing - Time spent on the page

Problem: Is the time spent on the new page same for the different language users?

- H_o: The mean times spent on the new page by English, French, and Spanish users are equal.

- H_a: At least one of the mean times spent on the new page by English, French, and Spanish users is unequal.

By using the one-way ANOVA test, we get the p-value as ~0.4320 that is > 0.05.

Therefore, we fail to reject the null hypothesis that the mean times spent on the new page by English, French, and Spanish users are equal..

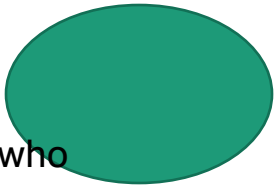Time spent on the new page w.r.t. Language preferred
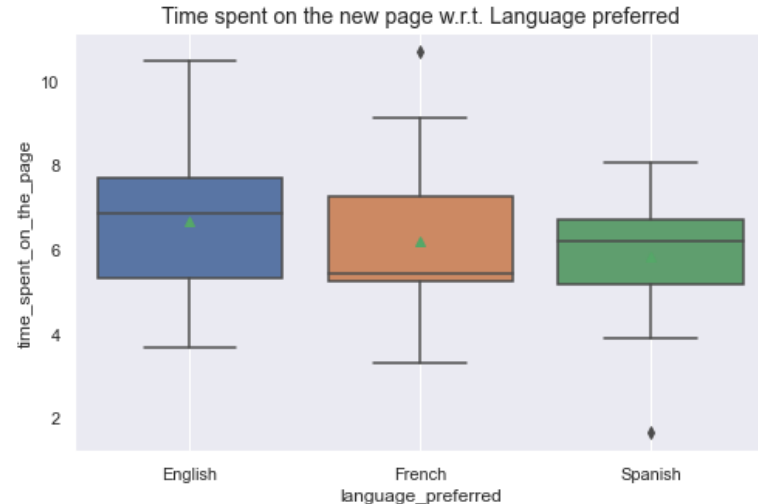
- The time spent on the new page by English, French, and Spanish users does not look drastically different. but, we need to test if this difference is statistically significant or not.
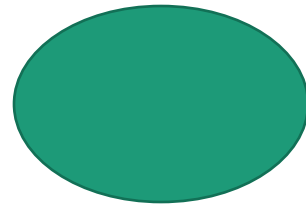
Conclusion

I analyzed a dataset of the response of 100 randomly selected users divided into two groups(samples)-the treatment and the control group.The treatment group is served the new landing page and the control group is served the old landing page. My conclusions from the analysis made are:

- 54.0% of users in the dataset converted to subscribers of the news portal and 46.0% of users in the dataset did not convert to suscribers of the news portal.

- There are 3 unique languages in this dataset that users use to view the landing page. The most chosen language used, is the French and Spanish language.

• In this dataset,the time spent on the page by users ranges from ~0.19(minimum) to ~10.7(maximum) minutes.The average time spent by users on the landing page is ~ 5.38 minutes.

•Visually when we compare the time spent on the page distribution of the two samples(treatment and control group),It seems the users spend more time on the new landing page than the old landing page and hence statistically,we have enough evidence to say that this holds.

- 33.0% of users who viewed the new landing page(treatment group) and 21.0% of users who viewed the old landing page(control group),converted to subscribers of the news portal.Visually,it seems the proportion of users who visit the new landing page and get converted to subscribers of the news portal is more than the proportion of users who visit the old landing page and get converted to subscribers of the news portal. Hence statistically, we have enough evidence to say that the conversion rate for the new page is greater than the conversion rate for the old page.

- From the sample,21.0% of users who chose English language,18.0% of users who choose French language and 15.0% of users who choosed Spanish language to view the landing page, converted to subscribers of the news portal.Even though visually from the sample,it seems that the users who choose English language had the highest converted status ,statistically,**there is not enough evidence to say that the converted status depends on the preferred language.

- When we compare the sample mean time spent on the new landing page for different language users, even though visually it seems as if the mean time spent on the new landing page differs among the three different language users,**there is enough statistical evidence to say that the mean time spent on the new page is same for the different language users.

Business Recommendations

- From the above analysis and conclusions, the new landing page created by the design team of the company which suggest that it might be more captivating as the users spend more time on it. The new landing page is hence more effective to gather new subscribers. If the new landing page is maintained and why not perfected ,it might gather many more subscribers.

- Even though statistically, They have been evidence that converted status is independent of preferred language which visually shows other wise,i think it's better for the company to carry out further investigations on users about this inorder to have a deeper insight and better viewers-news portal experience which might attract many more subscribers.

- I suggest the company should further investigate subscription method and find ways to facilitate subscription. This is because some users though they spent a lot of time on the new landing page, yet they don't subscribe. May be subscription method to the news portal is quite complex or time demanding for some users which might find it quite hard to subscribe even though they enjoy the news portal .