E-news Express Project

Objective

Statistical analysis of business data. Explore the dataset and extract insights from the data.

Analytical tasks-

- 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?

Data Description:

- user_id - This represents the user ID of the person visiting the website.
- group - This represents whether the user 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 user on the landing page.
- converted - This represents whether the user gets converted to a subscriber of the news portal or not.
- language_preferred - This represents the language chosen by the user to view the landing page.

Data Information

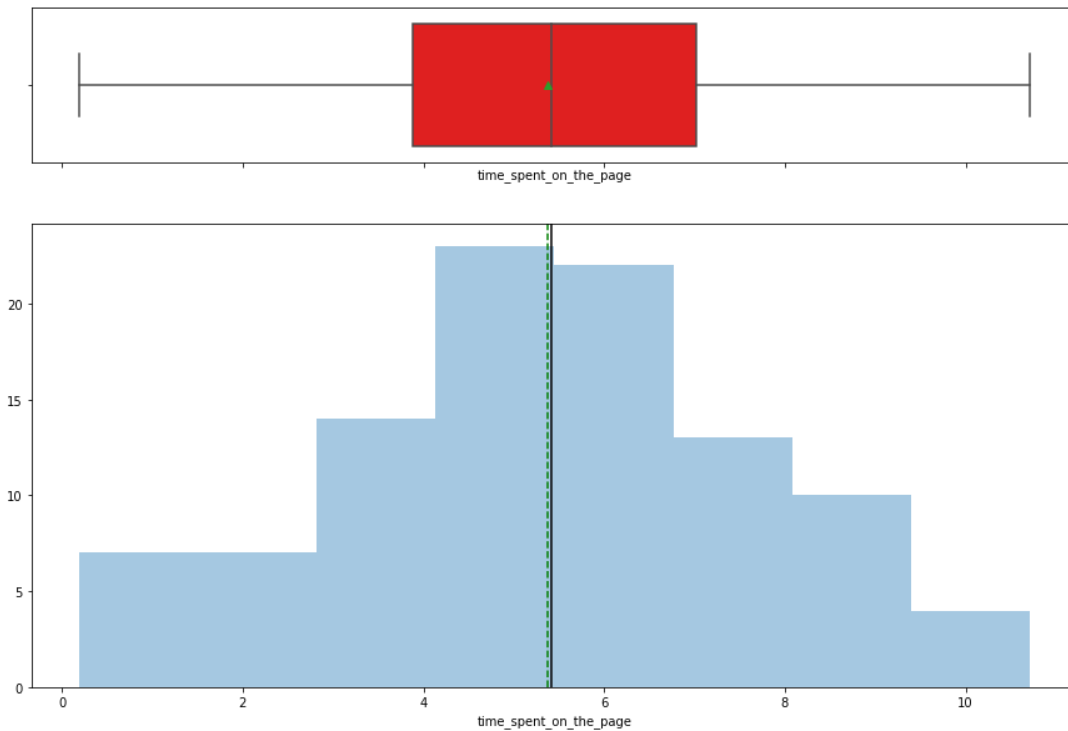| Observations	Variables
100	6

Note:
- group, landing_page, converted and language_preferred variables are categorical variable.
- time_spent_on_the_page is a numerical variable.
- no missing value observed

Data Description

- All the categories in the group and landing_page variables has an equal count which mean that
  - half of the user belongs to the first group (control) and others half to the second group (treatment),
  - it also indicate that half of the landing page is new and others half is old.
- The number of yes observations is higher than no observations which represents that more user gets converted to a subscriber of the news portal
- Equal number of users chosen French and Spanish for preferred language which is higher (34) compared to English (32)

Observation on time spent on the page

- The distribution of time_spent_on_the_page is pretty much normal distribution
- Higher time spent on the page are between 4 to 5 hours
- Median time spent on the pare is almost equivalent to mean
- There is no lot of outliers in this variable.

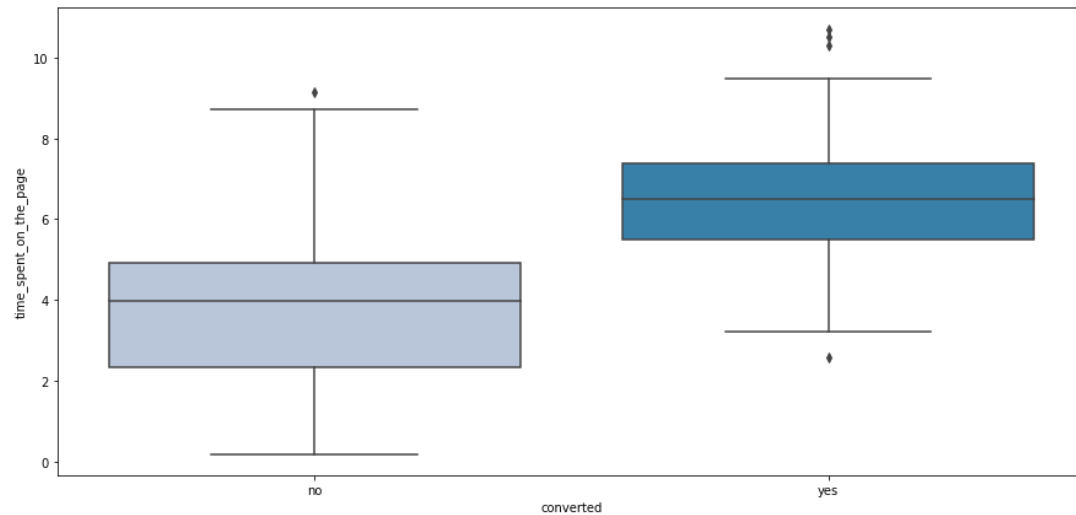# Observation on language preferred
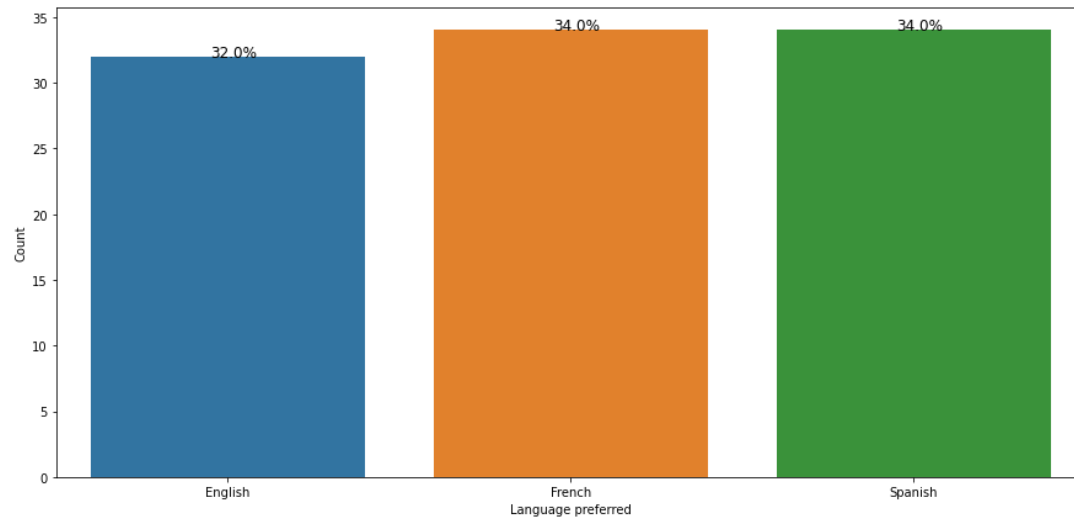
- 32% user chose English language to view the page, whereas 34% user chose French and Spanish
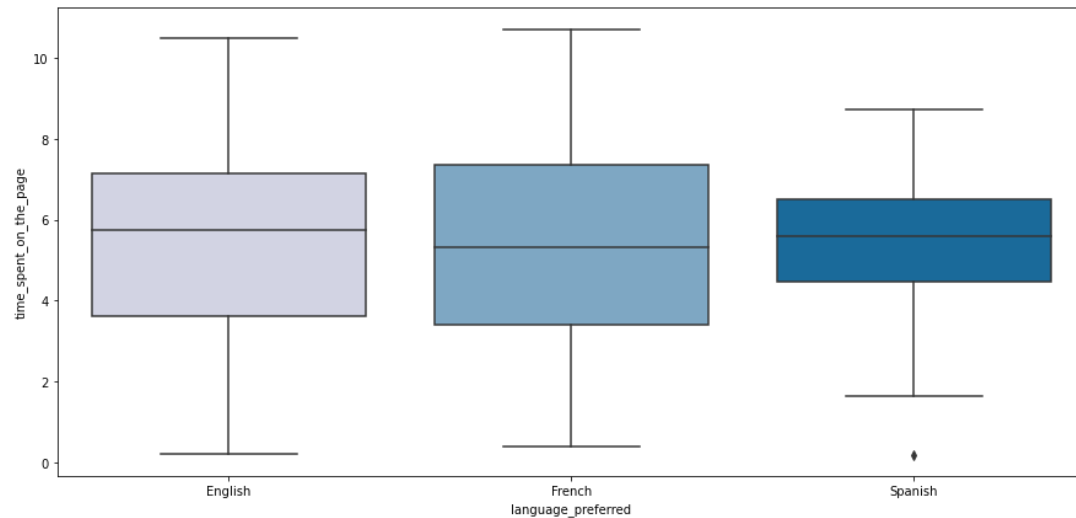
time_spent_on_the_page vs landing_page
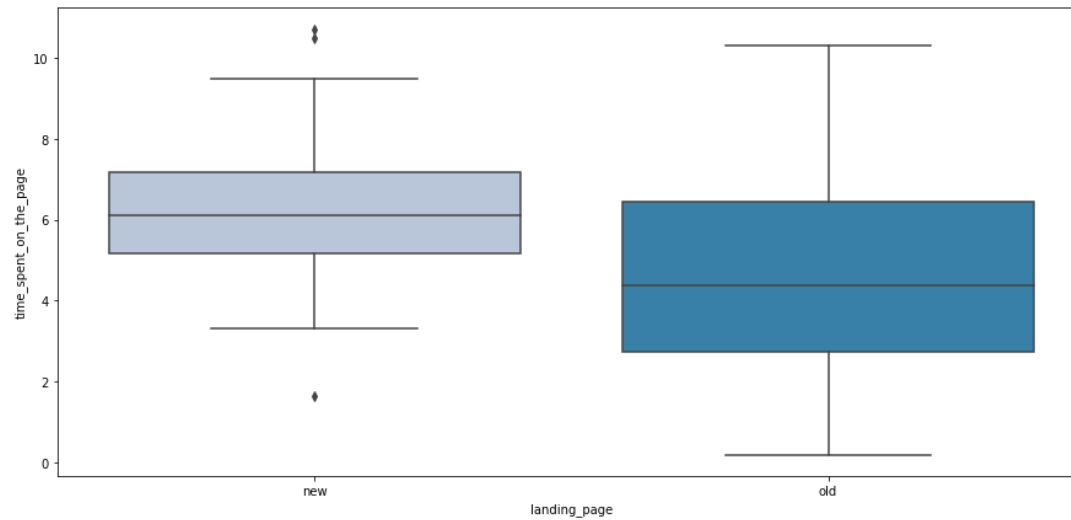
- Users are spending their time more on new landing page as compared to old landing page

# time_spent_on_the_page vs converted

- Users who spent higher time on the page tends to subscribe the news portal compared to who did less

time_spent_on_the_page vs language_preferred

- English and French language preferred users spent wide range of time on the page, but Spanish users are in the range of 5 to 7 (but there are some outliers)

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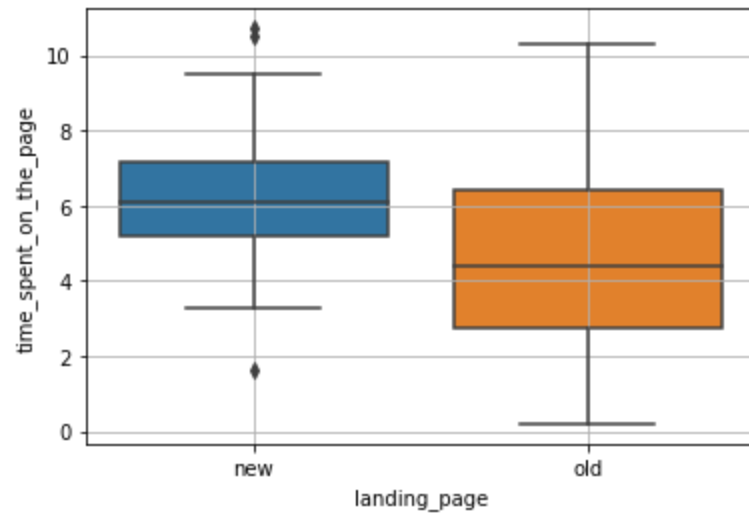

- users in new landing page spent longer time than old landing page

- Hypothesis test at 5% significance level also proved that

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?
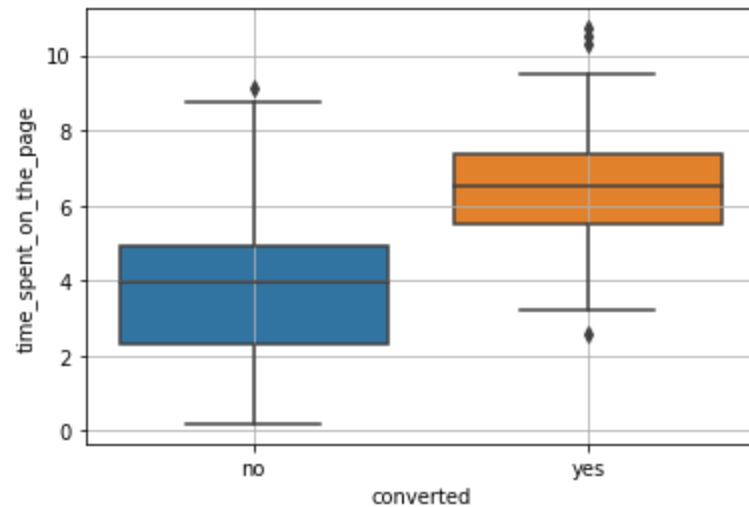
- conversion rate for the new page is greater than the conversion rate of old page.

- Hypothesis test at 5% significance level also proved that

Is the mean time spent on the new page same for the different language users?
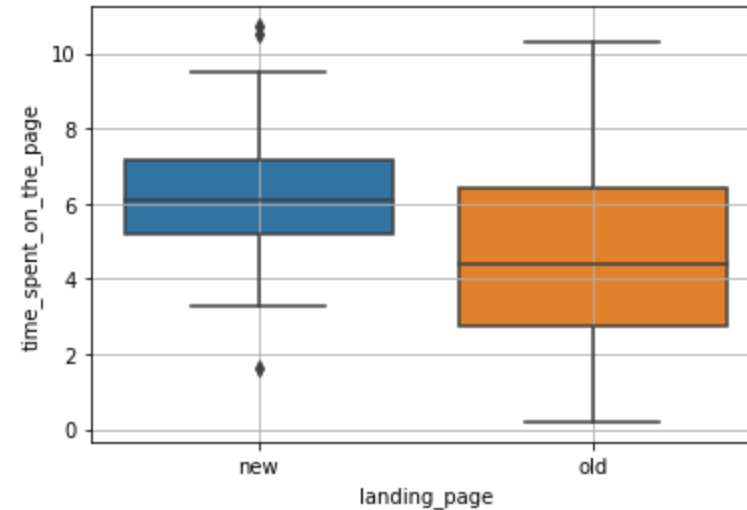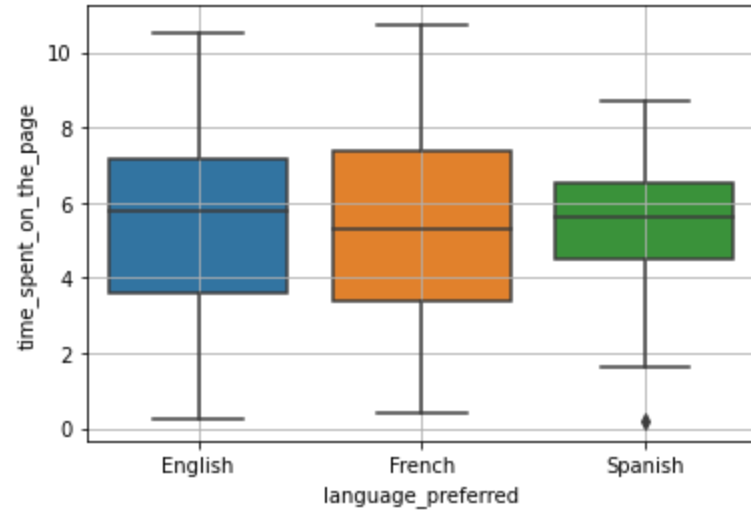
The mean time spent on the new page is not same for the different language users.
Hypothesis test at 5% significance level also proved that

Conclusion and Recommendations

Conclusion

We have been able to conclude that

- All the categories in the group and landing_page variables has an equal count which mean that half of the user belongs to the first group (control) and other half to the second group (treatment), it also indiate that half of the landing page is new and others half is old.

- More user gets converted to a subscriber of the news portal

- Equal number of users chosen French and Spanish for preferred language which is higher (34) compared to English (32)

- Users in new landing page spent longer time than old landing page

- Conversion rate for the new page is greater than the conversion rate of old page.

- new landing page users spent more time compared to old landing page users across the language preferred.

- The mean time spent on the new page is not same for the different language users.

- 54% gets converted to a new subscriber of the news portal.

Conclusion and Recommendations

- 32% user chose English language to view the page, whereas 34% user chose French and Spanish
- Users of control group spent higher time on the page camped to the users of treatment group
- Users are spending their time more on new landing page as compared to old landing page
- Users who spent higher time on the page tends to subscribe the news portal compared to who did less
- The mean time spent on the page by English language preferred user is higher
- The mean time spent on new landing page is higher compared to old landing page across language preference
- The mean time spent on the page by 1st group is higher compared to the 2nd group across language preference
- The mean time spent on the page who converted is higher compared to the who did not convert across language preference
- Converted status is independent on the preferred language
- English language preferred new landing page users spent highest amount of time followed by French language new landing page users and Spanish language new landing page users.

Conclusion and Recommendations

- **Recommendation to business**

- English language preferred users tends to have better conversion rate to subscription, so need to target more English language preferred users.

- Procure more data on converted status to build a predictive model.