

# Business Statistics Case Study

## E-news Express

---

Bhuvaneswari Suriyanarayanan

# Background

An online news portal E-news Express aims to expand its business by acquiring new subscribers. Every visitor to the website takes certain actions based on their interest. The company plans to analyze these interests and wants to determine whether a new feature will be effective or not.

# Objective

We are going to explore the dataset and extract insights from the data using EDA techniques and perform statistical analysis on the business data.

We will be focussing on the below Business problems

- 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?

# Sample Data Information

Variable	Description
user_id	This represents the user id of the person visiting then website
group	This represents whether the user belongs to the the first group(control) or the second group(treatment)
landing_page	This represents whether the landing page is old or new
time_spent_on_the_page	This represents the time(minutes) spent by the user on the landing page
converted	This represents whether the user gets converted or not
language_preferred	This represents the language chosen by the user to view the landing page

Observations	Variables
100	6

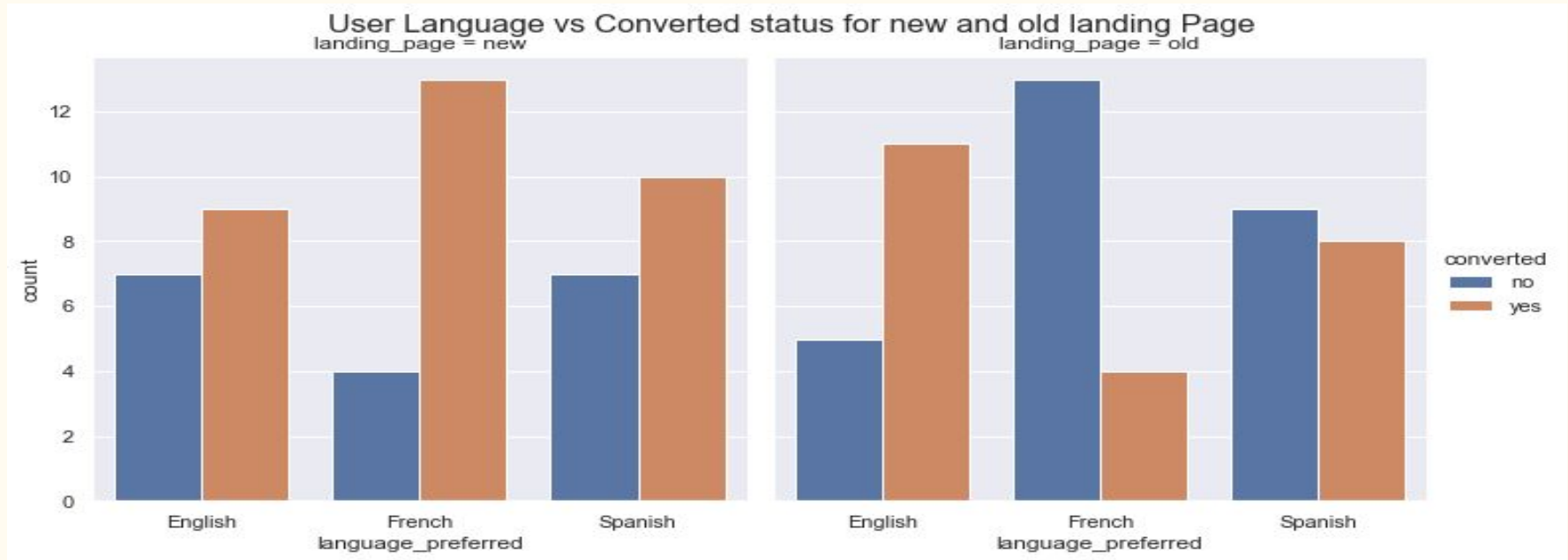
## Note:

- Changed all the objects in to categorical variables.
- Grouped the data of old landing page and new landing page separately for better comparison.

# Assumptions

- This is a simple random sample from population.
- From Central Limit theorem as the sample data is greater than 50, we can assume that the mean will follow a normal distribution.
- We will consider an significance level of 0.05 for all our tests.

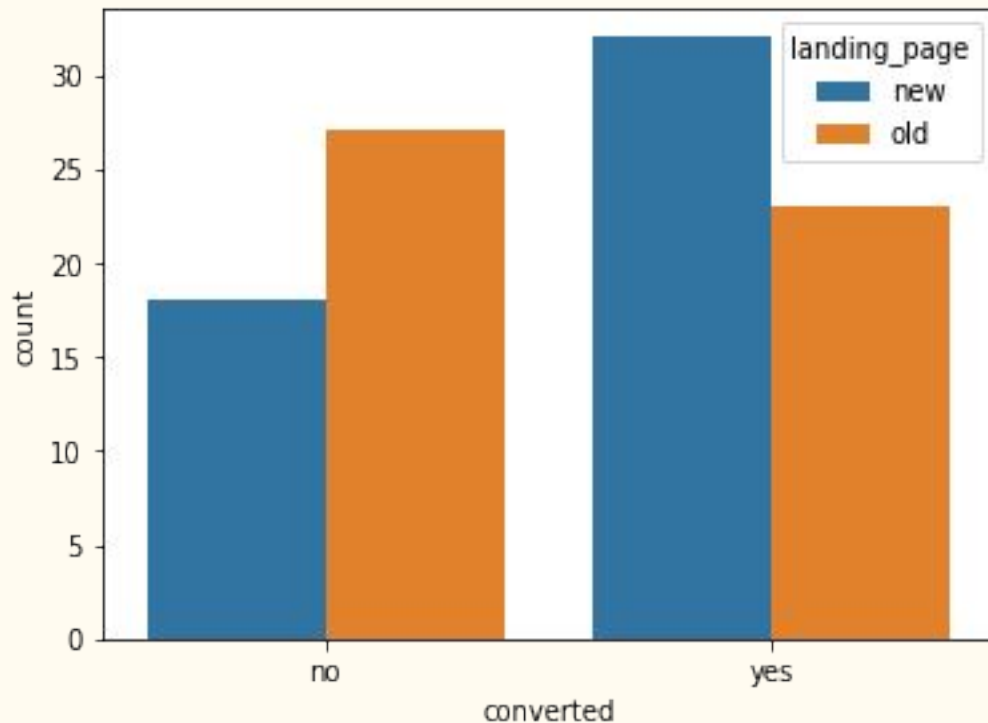
# Exploratory Data Analysis - Univariate Analysis



- ❑ All three language users of the new landing have higher conversion ratio indicating some attractive language specific features especially for French and Spanish that make them subscribe to the new page than the old page. For French, conversion success/failure counts have flipped.
- ❑ The old page seems to be more attractive for English users as they have more subscribers than other two language users.
- ❑ Also when we take a closer look and compare the count of English Language user subscribers in old and new page, the old page has slightly more subscribers than the new page.

# Exploratory Data Analysis - Univariate Analysis

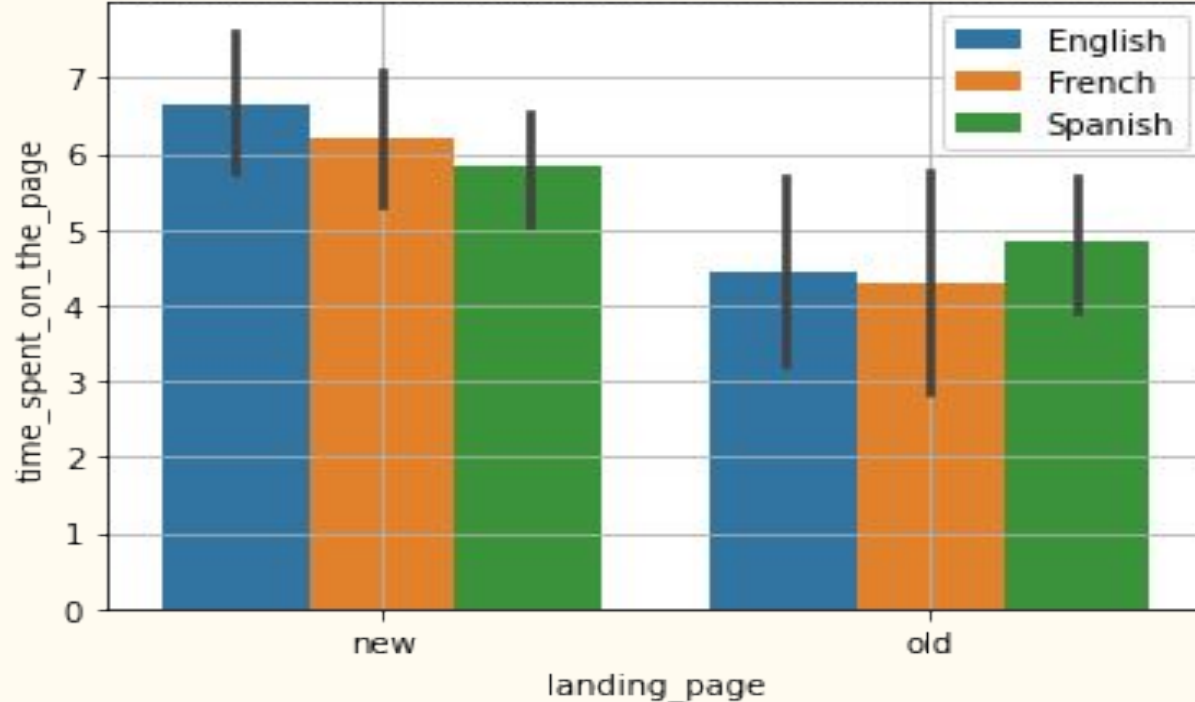
Landing Page vs Converted status



- ❑ The new landing page has more subscribers than old page.

# Exploratory Data Analysis - Bivariate Analysis

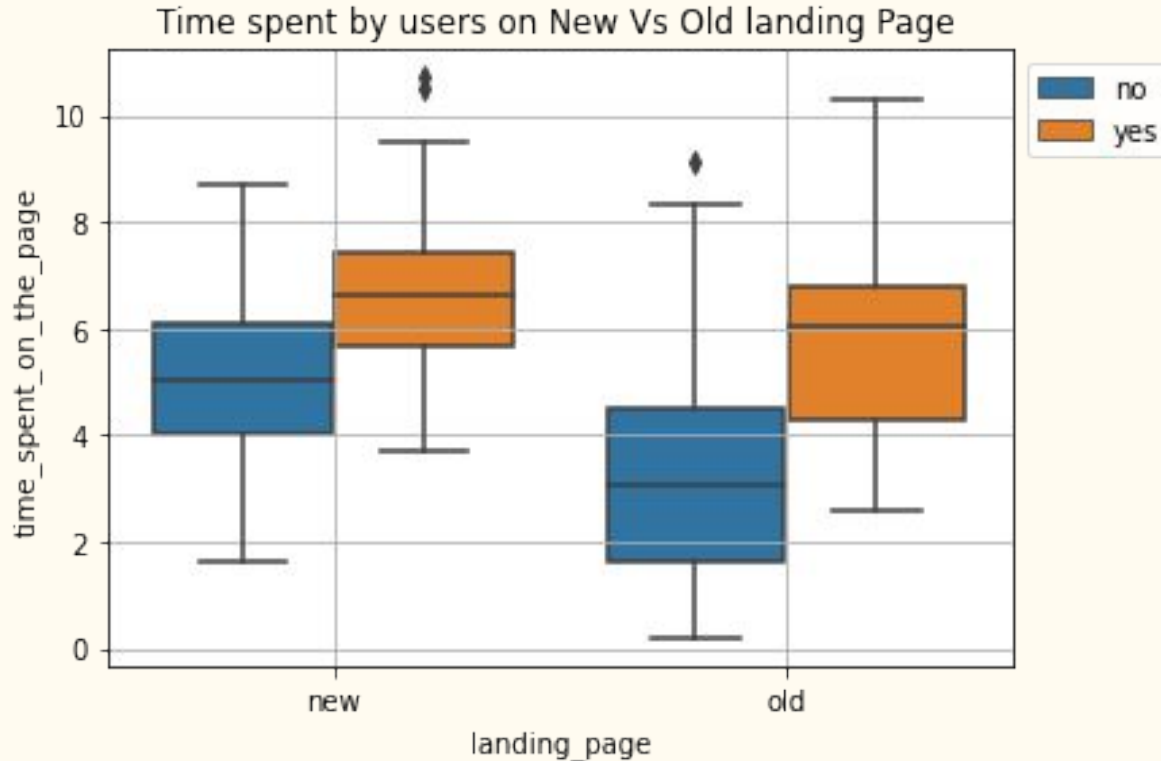
Time spent by different language users on New Vs Old landing Page



- ❑ Irrespective of the different language users, the mean time spent by users on new landing page is more than that of the old landing page.

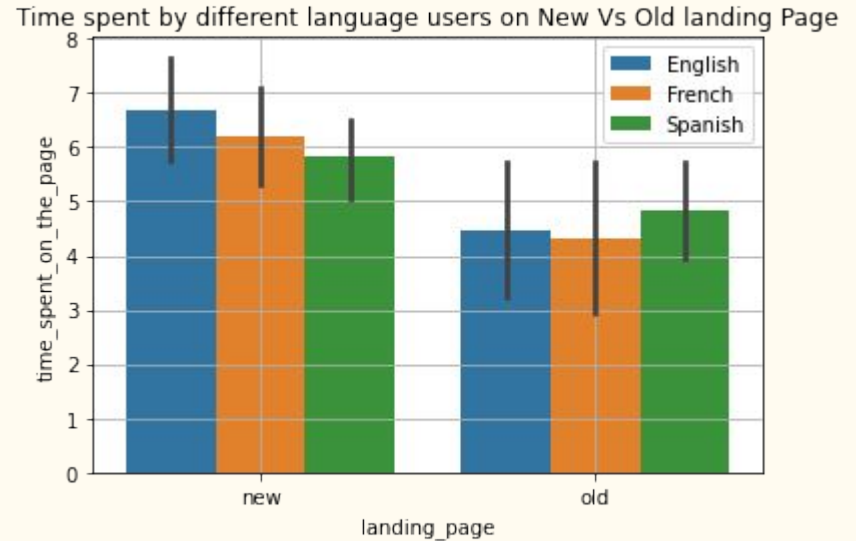
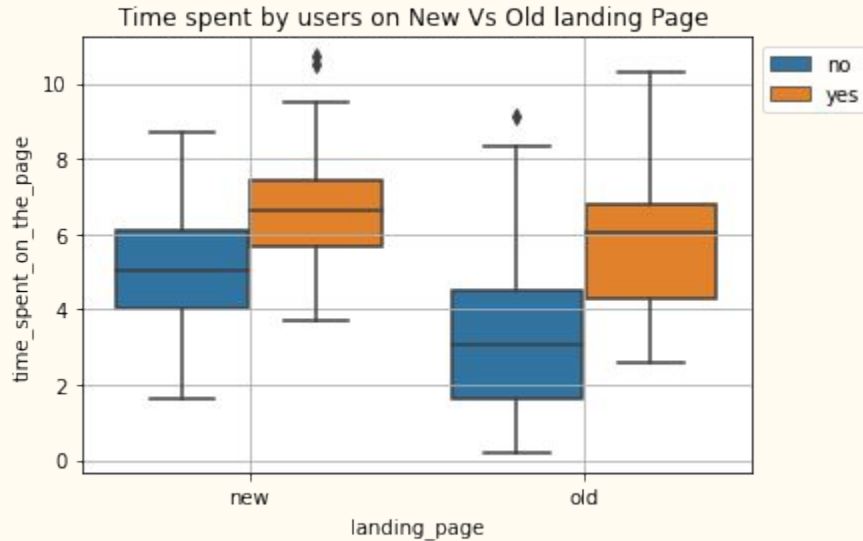


# Exploratory Data Analysis - Bivariate Analysis



- ❑ The mean time spent is higher for new page than old landing page.
- ❑ Users who spend more time on a landing page gets converted.

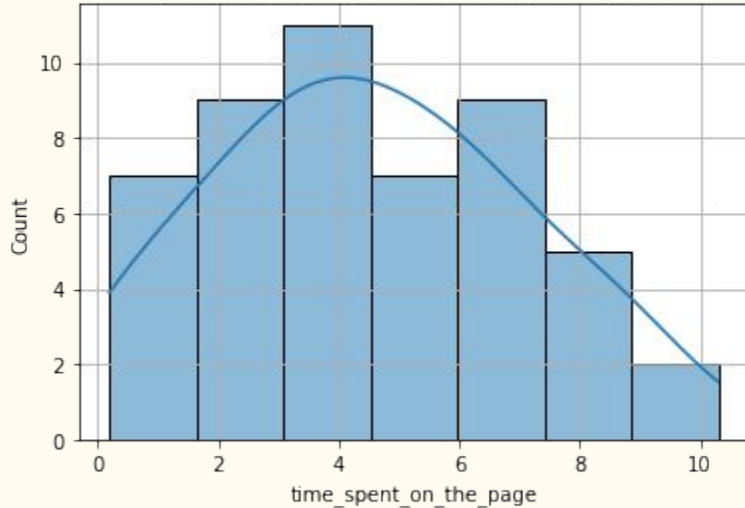
# Exploratory Data Analysis -Do the users spend more time on the new landing page than the old landing page?



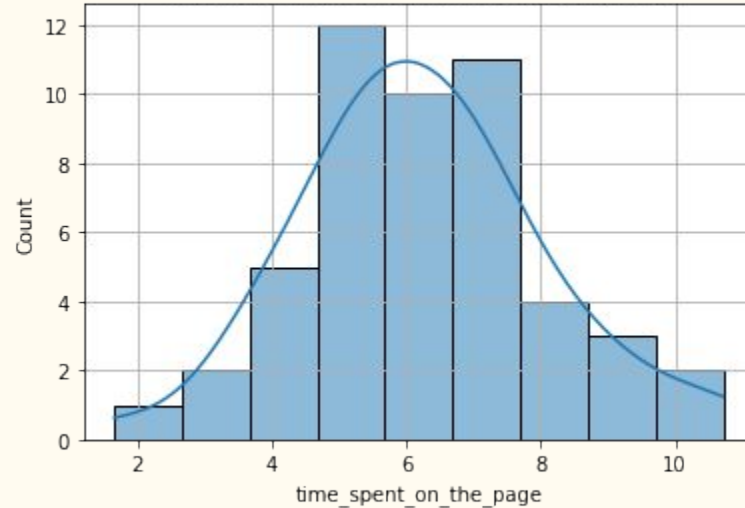
- ❑ The mean time spent is higher for new page than old landing page.
- ❑ Users who spend more time on the landing page are converted users.
- ❑ All the language users spend more time in the new landing page than the old landing page.
- ❑ There is a subtle difference in the mean time spent by the different language users on new and old landing page respectively.

# Statistical Analysis - Do the users spend more time on the new landing page than the old landing page?

Time Spent by users on Old landing page



Time Spent by users on New landing page



- ❑ Mean time spent by the users on new landing page (6.22) > Mean time spent by the users on old landing page (4.53)

# Statistical Analysis - Do the users spend more time on the new landing page than the old landing page?

**Null Hypothesis:**

Let  $\mu_1, \mu_2$  be the mean time spent on the new landing page( $\mu_1$ ) and old landing page( $\mu_2$ ) respectively.

$H_0: \mu_1 = \mu_2$

**Alternate hypothesis:**

$H_a: \mu_1 > \mu_2$

We will test the null hypothesis using two Independent Sample T-test for Equality of Means - Unequal Std Dev

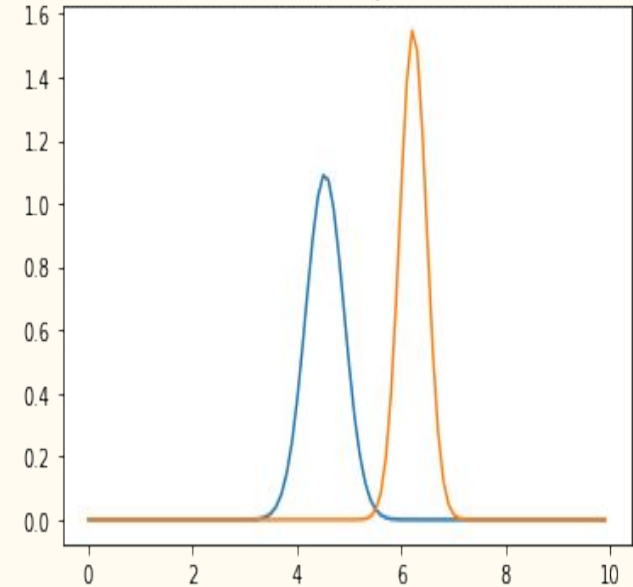
The p-value is 0.0001392381225166549

**Insight:**

As the p-value is much lesser than the level of significance(0.05), we reject the null hypothesis.

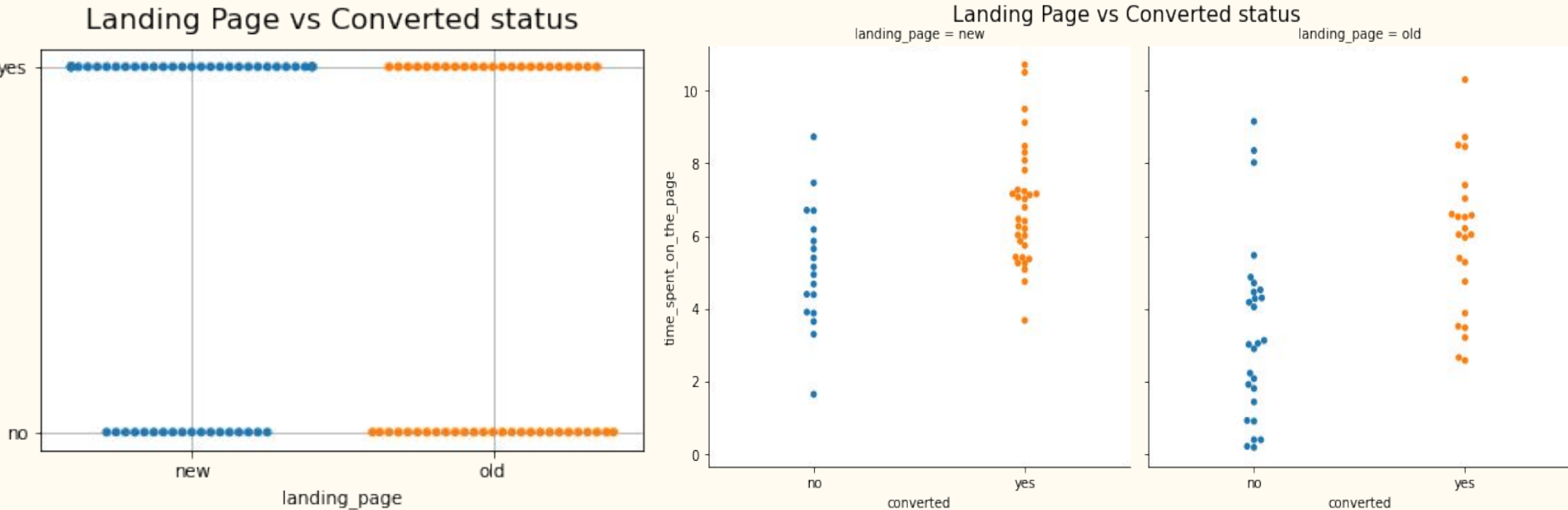
Hence we can conclude that the average time spent on the new landing page is greater than the old landing page at 5% significance level.

Normal distribution for the mean time spent on old and new landing page



No Significant overlap for both the distribution, confirming that their means are significantly different

# Exploratory Data Analysis-Is the conversion rate for the new page greater than the conversion rate for the old page?



- ❑ The above visual plots clearly shows that the conversion rate for the new landing page is more than that of the old landing page.

# Statistical Analysis - Is the conversion rate for the new page greater than the conversion rate for the old page?

**Null Hypothesis:** Let  $p_1, p_2$  be the proportions of converted users in new landing page and old landing page respectively.

**$H_0: p_1 = p_2$**  New landing page conversion rate is equal to old landing page conversion rate

**Alternate hypothesis:**

**$H_a: p_1 > p_2$**

**New landing page conversion rate is greater than the old landing page conversion rate**

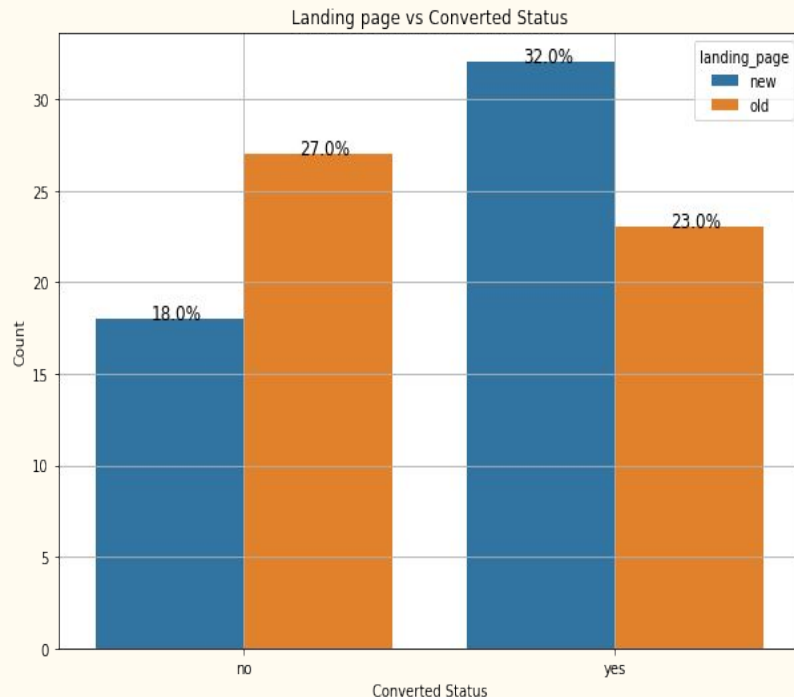
**We will test the null hypothesis using Two proportion Z- test**

**The p-value is 0.035220214636043995**

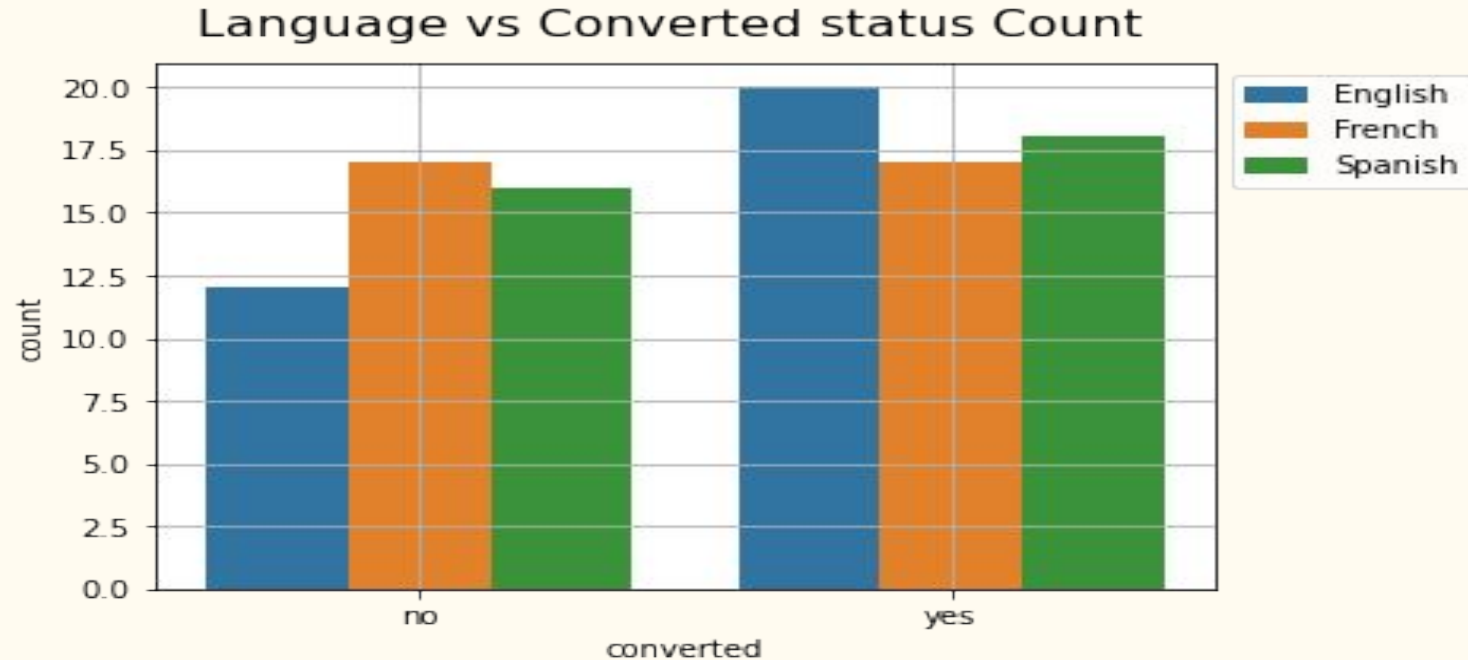
**Insight:**

**As the p-value is lesser than the significance level(0.05), we reject the null hypothesis.**

**Thus, we have enough statistical significance to conclude that the conversion rate of the new landing page is greater than the conversion rate of the old landing page at 5% significance level.**



# Exploratory Data Analysis-Does the converted status depend on the preferred language?



- ❑ The conversion status is slightly greater for English users than the other two languages but it is not very significant.

# Statistical Analysis-Does the converted status depend on the preferred language?

**Null hypothesis:**

*H<sub>0</sub>*: Converted Status is independent of language.

**Alternate hypothesis:**

*H<sub>a</sub>*: Converted Status depends on language.

**We will test the null hypothesis using Chi2 Test for Independence**

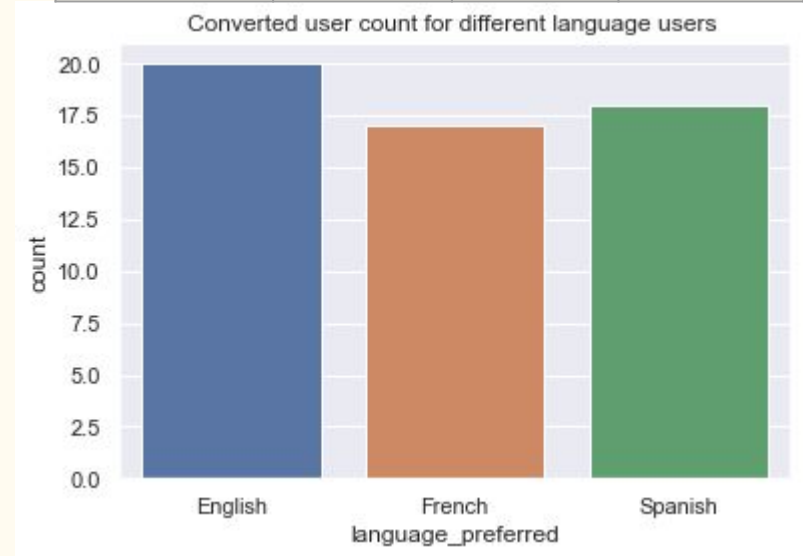
**The p-value is 0.5686624745121615**

**Insight::**

**As the p-value is much greater than the significance level, we fail to reject the null hypothesis.**

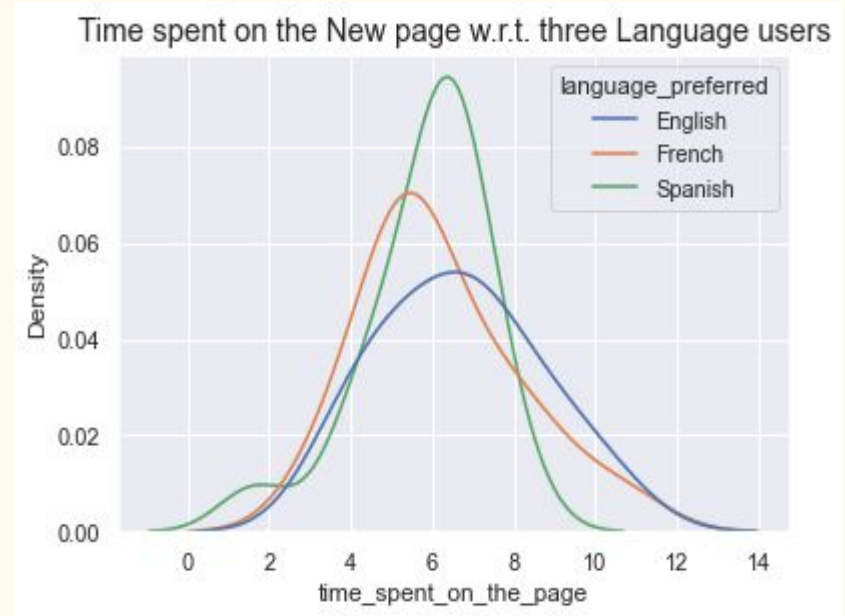
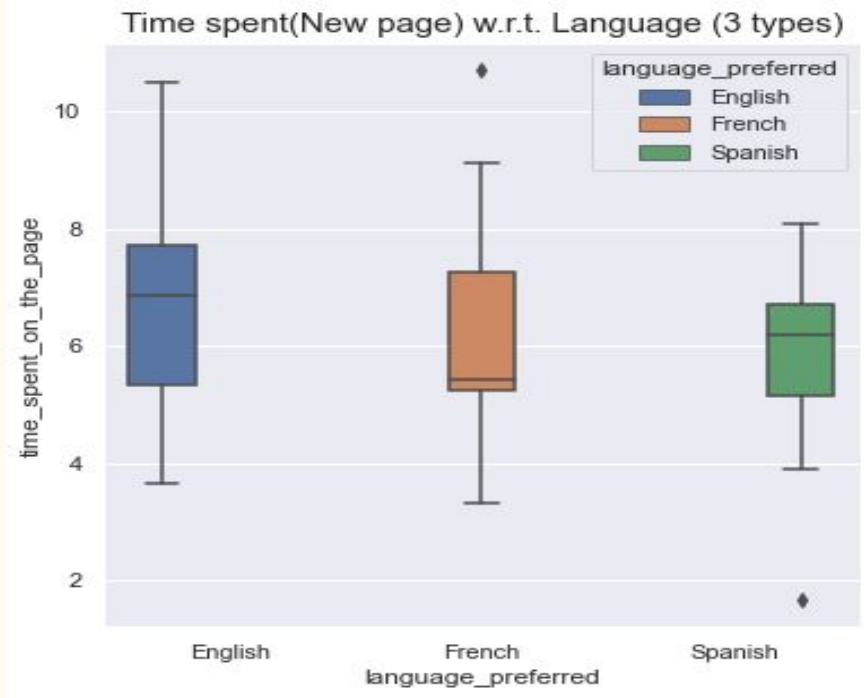
**Hence, there is no enough statistical evidence to conclude that the converted status is dependent on language at 5% significance level.**

Language Preferred	English	French	Spanish
Converted			
No	12	17	16
Yes	20	17	18





# Exploratory Data Analysis - Is the mean time spent on the new page same for different language users?



- There is very subtle difference in the mean time spent on the new landing page for all three language users(English = 6.7 , French = 5.8, Spanish = 6.2)

# Statistical Analysis - Is the mean time spent on the new page same for different language users?

**Null Hypothesis:**

Let  $\mu_1, \mu_2, \mu_3$  be the means of time spent on the new landing page for English( $\mu_1$ ), French( $\mu_2$ ) and Spanish( $\mu_3$ ) language users respectively.

$H_0: \mu_1 = \mu_2 = \mu_3$  (all three language users spend equal time on the new landing page)

**Alternate Hypothesis:**

$H_a$ : The time spent on the new landing page for atleast one language user is different from the rest.

We will test the null hypothesis using one way Anova test

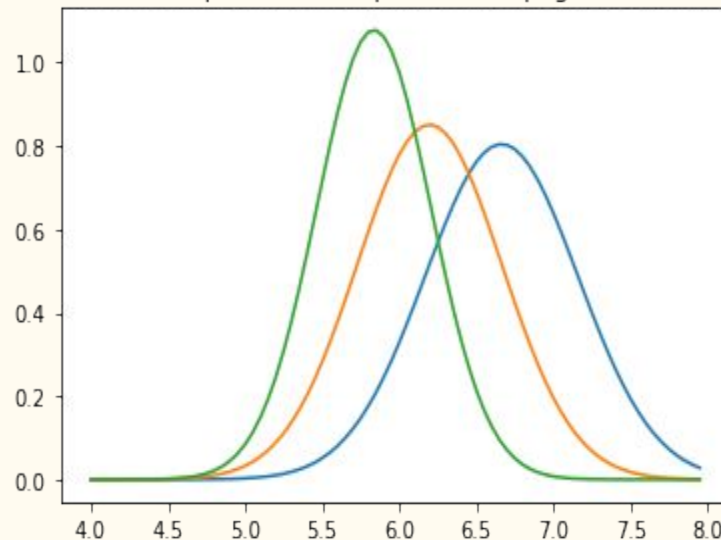
The p-value is 0.43204138694325955

**Insight:**

As the p-value is much greater than the significance level, we fail to reject the null hypothesis.

Hence, we do not have enough statistical evidence to conclude that time spent on new landing page for atleast one language user is different from the rest at 5% significance level

Normal Distribution plot for time spent in new page for three languages



There are significant overlaps for all three distributions for the time spent on the new page by three language users indicating they are not different.

# Conclusion

- ❑ All three language users of the new landing page gets converted indicating some attractive language specific features especially for French(total inverse in subscription status from old to new landing page) and Spanish that make them subscribe to the new page than the old page.
- ❑ The old page seems to be more attractive for English users as they have more subscribers than other two language users.
- ❑ Also when we take a closer look and compare the count of English Language user subscribers in old and new page, the old page has slightly more subscribers than the new page.
- ❑ The new landing page has more subscribers than old page.
- ❑ Irrespective of the different language users, the mean time spent by users on new landing page is more than that of the old landing page.
- ❑ Users who spend more time on a landing page gets converted.
- ❑ The mean time spent by users is higher for new page than old landing page at 5% Significance level
- ❑ The conversion rate of the new landing page is greater than the conversion rate of the old landing page at 5% Significance level.
- ❑ There is no enough statistical evidence to conclude that the converted status is dependent on language at 5% significance level.
- ❑ There is no enough statistical evidence to conclude that time spent on new landing page for atleast one language user is different from the rest.

# Recommendations

- ❑ French users can be the primary target for the New landing page followed by Spanish
- ❑ Business can focus on making the new page equally interesting for English users as well to increase the subscribers as compared to old landing page