

# E-News Express

## Data Science & Business Analytics Business Statistics Project

12-17-22

# Contents / Agenda

- Executive Summary
- Business Problem Overview and Solution Approach
- EDA Results
- Hypotheses Tested and Results

# Executive Summary

- From the 4 tests conducted, we reached the following conclusions:
  - Users spend more time on the new landing page than the existing one.
  - The conversion rate for the new page is also greater than that of the old one.
  - Conversion status does not depend on the preferred language of a user.
  - Mean time spent on the page is the same regardless of preferred language.

# Executive Summary

- Insights gained from EDA:
  - More users in the dataset were converted than not
  - Spanish is the most preferred language from this data
  - Users who are converted spend more time on the website
- Business Recommendations:
  - Implementing the new landing page will surely get the website more subscribers/traction
  - The company could add more options to the preferred languages to further gain more insights on their audience
  - In the future, they can make more changes to the landing page similar to the ones they already made to make it even more appealing to new users, leading them to gain more subscribers

# Business Problem Overview and Solution Approach

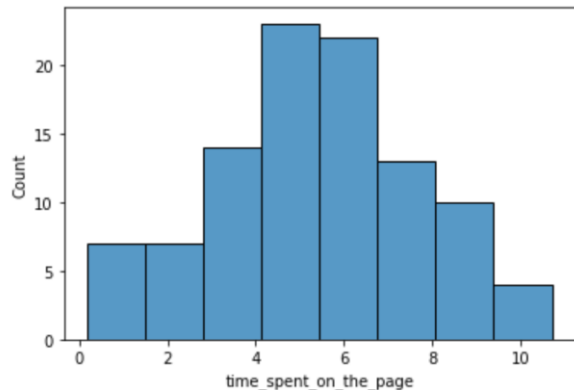
- E-News Express, an online news platform, wants to gain new subscribers to expand their business.
- In order to do this, we will analyze engagement to understand user interests and find how to keep them actively growing. The design team created a new landing page with a different outline & more relevant content. To test the effectiveness of the new landing page in securing new subscribers, we conducted an experiment by randomly selecting 100 users and dividing them into two groups equally. The existing landing page was shown to the first group, the control group, and the new landing page was shown to the second group, the treatment group. Data concerning the interaction of both groups with the older and newer versions of the landing page was collected. We then analyzed the data and performed statistical analyses to find the effectiveness of the new landing page in gaining new subscribers to E-News Express. Here are some questions we answered:
  - Do the users spend more time on the new landing page than on the existing landing page?
  - Is the conversion rate for the new page greater than the conversion rate for the old page?
  - Does the converted status depend on the preferred language?
  - Is the time spent on the new page the same for the different language users?

# EDA Results

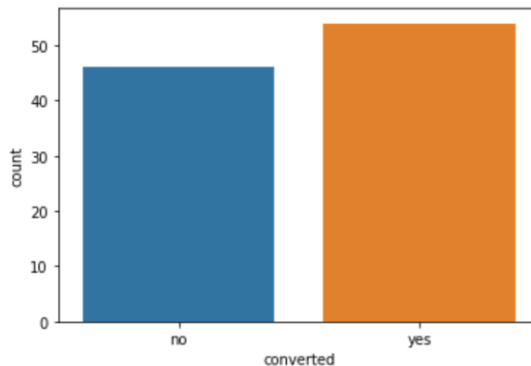
- The shape of the data is 100 rows by 6 columns, meaning there are 100 different users for which 6 measures were taken.
  - Out of these 6 measures, 1 is integer type, 4 objects, and 1 float.
- There are no missing values or duplicated user id's.
- The mean time spent on the page is 5.38 minutes with a standard deviation of 2.38 minutes, the minimum being 0.19 min and the maximum being 10.71 min.

[Link to Appendix slide on data background check](#)

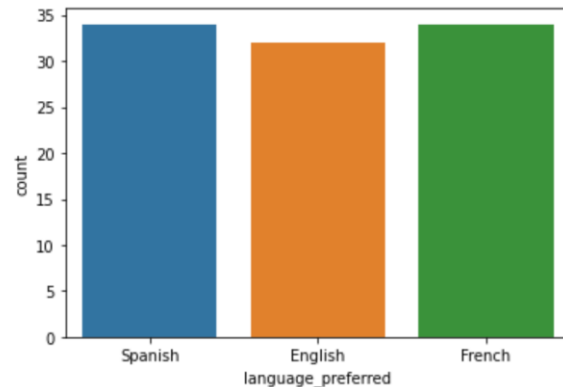
# EDA Results: Univariate Analysis



Here is the distribution of the time spent on the page for both control (existing landing page) and treatment (new landing page) groups.



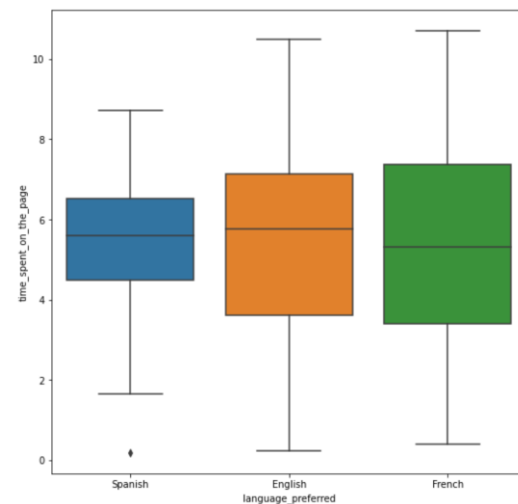
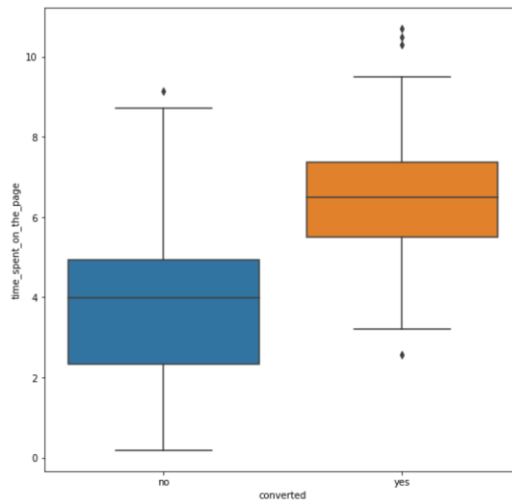
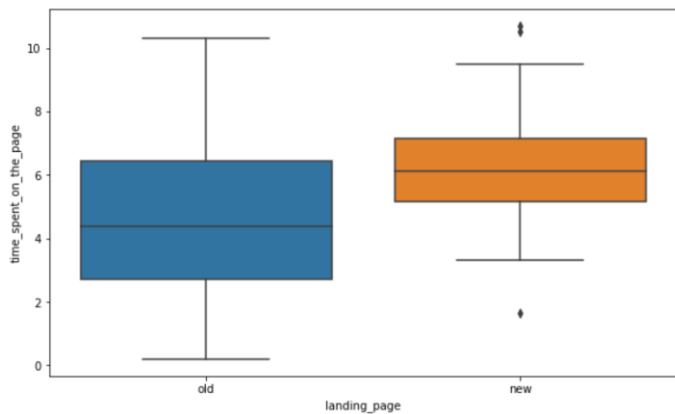
More users were converted than not, meaning they became subscribers to the online news portal.



Here is the distribution of the different languages preferred by users. They are mostly uniform with a few more preferring Spanish and French.

[Link to Appendix slide on data background check](#)

# EDA Results: Bivariate Analysis



This graph shows which landing page a user got vs. the amount of time they spent on it. It seems like users spent more time on the new page.

In this graph we see the time spent on the page by a user and if they were converted to a subscriber or not. If they were, they usually spent more time.

Here we see the preferred language of users vs. how long they spent on the page, fairly similar except for Spanish.

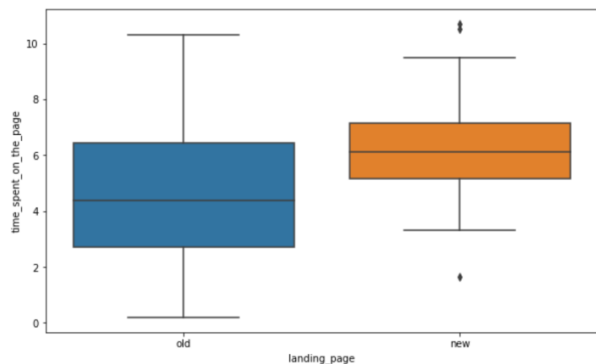
[Link to Appendix slide on data background check](#)



# Hypotheses Tested and Results

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

- Visual analysis: From the graph it looks like more time was spent on the new page, but we do not enough evidence to strongly accept this conclusion yet.



- Hypothesis tested:  $\mu_1$  = mean time spent on the new landing page &  $\mu_2$  = mean time spent on the existing landing page
  - Null Hypothesis:  $\mu_1 = \mu_2$
  - Alternate Hypothesis:  $\mu_1 > \mu_2$

[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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

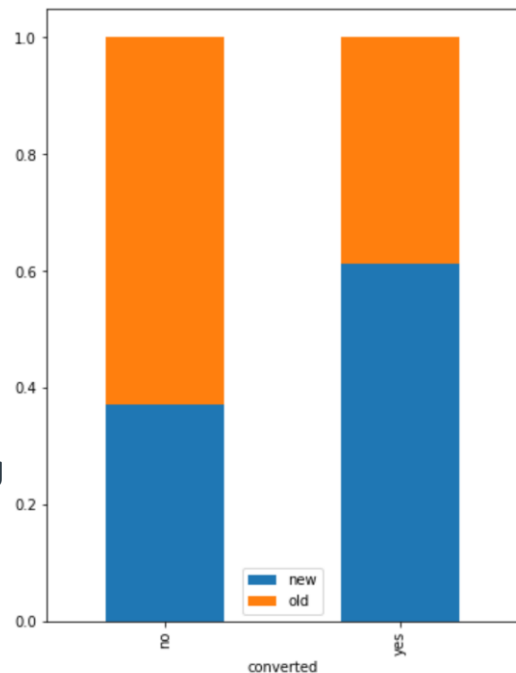
- Test result and inference: We used a 2-sample independent T-test for this hypothesis with a critical value of 0.05
  - The sample standard deviation of the time spent on the new page is: 1.82 and on the old page is: 2.58, so we use the test of unequal variances.
  - The P-value calculated from this test was 0.00013924, which is significantly less than the critical value.
  - We conclude to reject the null hypothesis because we have strong statistical evidence that the mean time spent on the new landing page was greater than that spent on the existing page.

[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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?

- Visual analysis: From the graph to the right, it looks like the conversion rate for the new page is greater than that of the old.
- Hypothesis tested:  $p_1$  = proportion of users visiting the new landing page who get converted &  $p_2$  = proportion of users visiting the old landing page who get converted
  - Null Hypothesis:  $p_1 = p_2$
  - Alternate Hypothesis:  $p_1 > p_2$



[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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?

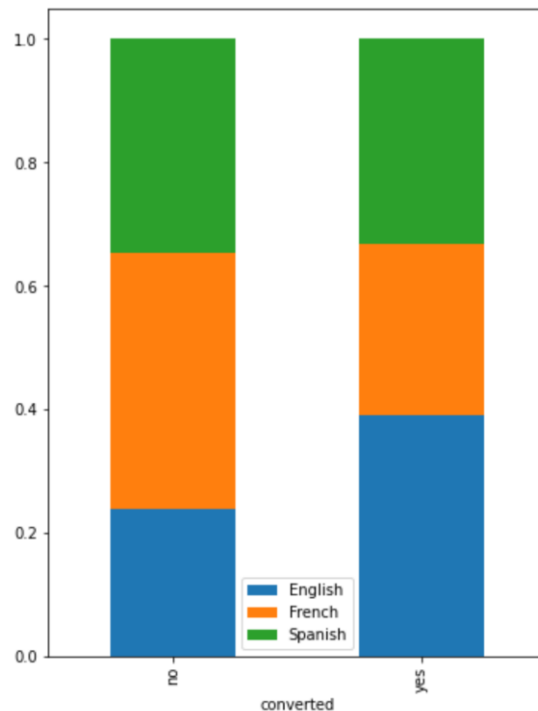
- Test result and inference: We used a 2-proportion Z-test for this hypothesis with a critical value of 0.05 again
  - As we had already decided, the number of users served the old and new page are both 50 for a total of 100 users.
  - The P-value calculated from this test was 0.008026, which is significantly less than the critical value.
  - We conclude to reject the null hypothesis because we have strong statistical evidence that the proportion of users converted after visiting the new landing page was higher than that of users visiting the old page.

[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

Does the converted status depend on the preferred language?

- Visual analysis: From the graph to the right, it looks like there is not a lot of correlation between preferred language and conversion status.
- Hypothesis tested:
  - Null Hypothesis: Conversion status is independent of language preferred
  - Alternate Hypothesis: Conversion status is not independent from language preferred



[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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?

- Test result and inference: We used a  $\chi^2$  test of independence with a critical value of 0.05 again
  - A contingency table, shown below, was created to see the preferred languages in comparison to whether a user was converted or not.

language_preferred	English	French	Spanish
converted			
no	11	19	16
yes	21	15	18

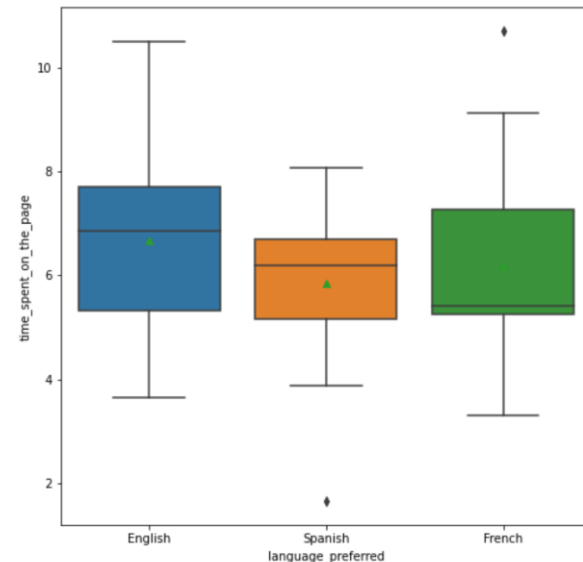
- The P-value calculated from this test was 0.21299, which is not low enough to be statistically significant.
- We conclude to fail to reject the null hypothesis because we do not have enough evidence that conversion status is dependent on preferred language.

[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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

- Visual analysis: From the graph to the right, it looks like there might be a correlation between preferred language and amount of time spent on the page.
- Hypothesis tested:  $u_1$  = mean time spent on page for an English user,  $u_2$  = mean time spent on page for a Spanish user,  $u_3$  = mean time spent on page for a French user
  - Null Hypothesis:  $u_1 = u_2 = u_3$
  - Alternate Hypothesis: At least one of the 3 different language preferred population means of time spent on the page is different



[Link to Appendix slide on details of the test performed](#)

# Hypotheses Tested and Results

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

- Test result and inference: We used a 1-way ANOVA test with a critical value of 0.05 again
  - We inserted data into new groups of mean time spent on page according to the user's preferred language.
  - The P-value calculated from this test was 0.43204, which is not low enough to be statistically significant.
  - We conclude to fail to reject the null hypothesis because we do not have enough evidence that we have enough statistical evidence to say that the mean time spent on the new page differs by preferred language of users.

[Link to Appendix slide on details of the test performed](#)





**Happy Learning !**

