Roman Dzadzic

10/24/2021

Business Presentation

E-news Express

Contents

We will cover the following topics

- Core business idea, objective, and approach
- Brief description of data
- Exploratory Data Analysis (EDA) with graphs and related insights
- Business Insights and Recommendations:
  - Actionable insights based on the statistical analysis results
  - Future exploration and next steps
- Appendix: data dictionary

Business Problem Overview and Solution Approach

Core business idea and problem / objective to tackle

As E-news Express – an online portal – our bread and butter is in acquiring new subscribers to which we upsell other services or advertise to and make revenue through ad dollars. Therefore, our most important conversion metric are users who land on our various landing pages and subscribe to our news and information services.

In order to sustain the business, we conducted a set of A/B tests, and as part of that we have analyzed a set of our landing pages – one NEW (treatment group) and one OLD landing page design (control group – existing landing page). Our idea was to explore how different designs and content may influence our North Star Metric – the landing page conversion rate.

The following few slides will attempt to present you with statistically analyzed information and its results and suggestions.

Data Overview

Actual data example below

	user_id	group	landing_page	time_spent_on_the_page	converted	language_preferred
0	546592	control	old	3.48	no	Spanish
1	546468	treatment	new	7.13	yes	English
2	546462	treatment	new	4.40	no	Spanish
3	546567	control	old	3.02	no	French
4	546459	treatment	new	4.75	yes	Spanish

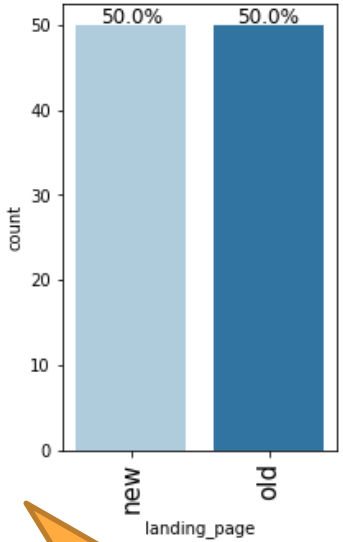
Brief description of data

- The A/B test dataset has 100 rows and 6 columns, where each row corresponds to a user landing on one of two landing page designs along with the resulting conversion being either a ‘yes’ – the user has converted, or a ‘no’ – the user left the page without converting.
- There was no missing data in this set.
- The ‘language_preferred' column has either a label named ‘Spanish’, ‘English’, or ‘French’ for the user language preferences.
- The ‘time_spent_on_the_page' column has numeric value and shows how many minutes the user has spent on the landing page prior to the conversion event.

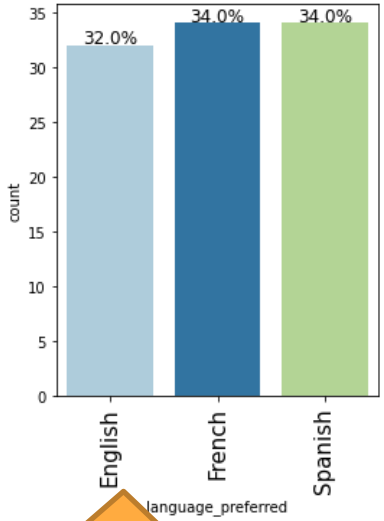
Exploratory Data Analysis (EDA)

- Graphs showing the factors most heavily impacting the target attribute
- Insights from the graphs showing the factors most heavily impacting the target attribute

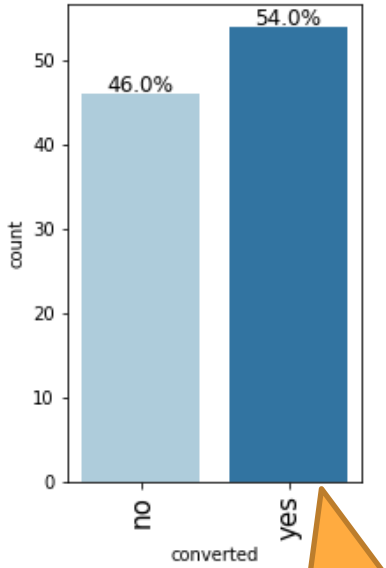
Exploratory Data Analysis (EDA)



As a confirmation of our A/B test split, you can see that 50% of visitors were distributed to the control group, and 50% to the treatment group.

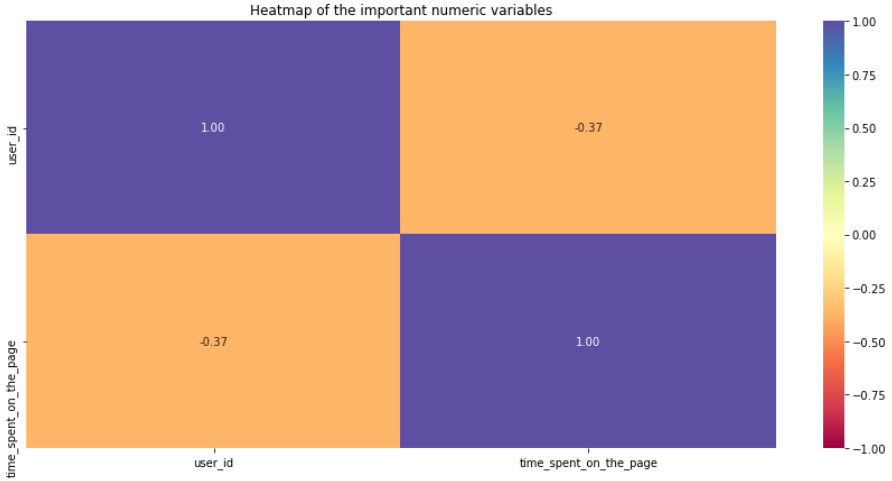


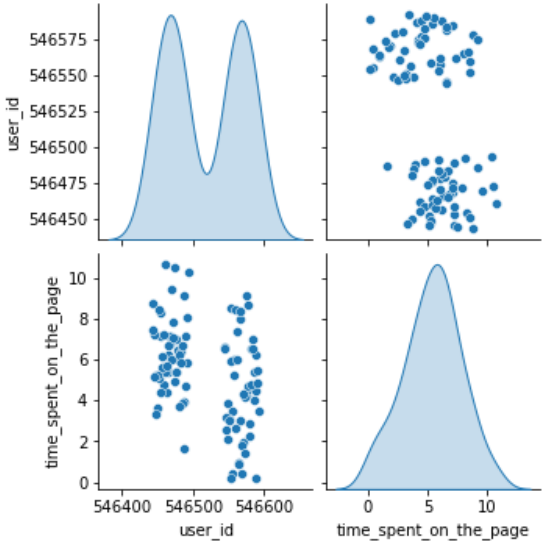
Of all the unique visitors, 32% had English as a language preference, and both French and Spanish languages comprised 34% of visitors each.



Of the total traffic of visitors to AB test, 54% converted.

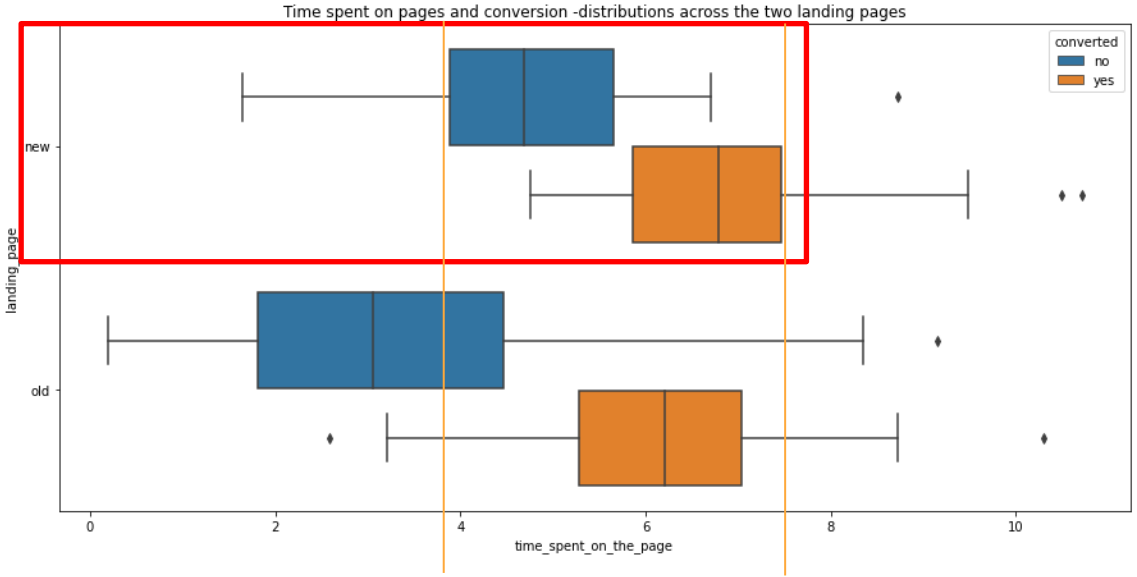
Exploratory Data Analysis (EDA)





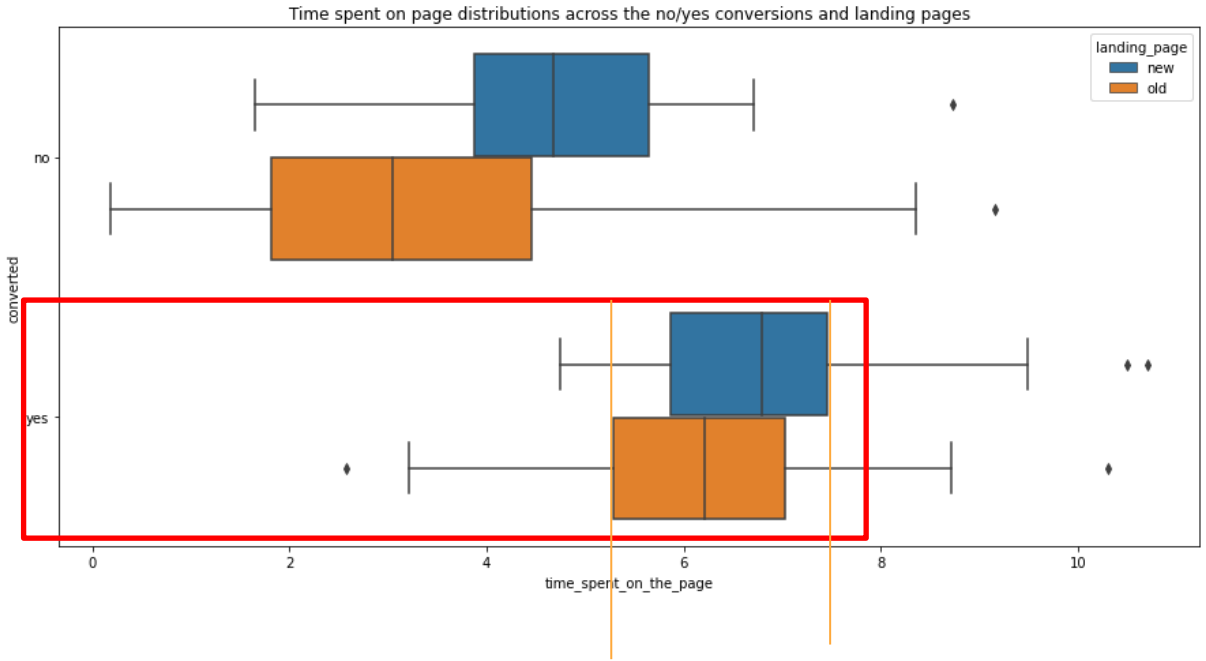
- In the above charts demonstrate a lack of any significant correlation between any of the observed variables.
- User IDs seem to be distributed across two distinct groups of counts with respect to the ID count itself.
- User IDs have twin peaks likely because their visits are distributed across two groups, control and treatment.

Exploratory Data Analysis (EDA)



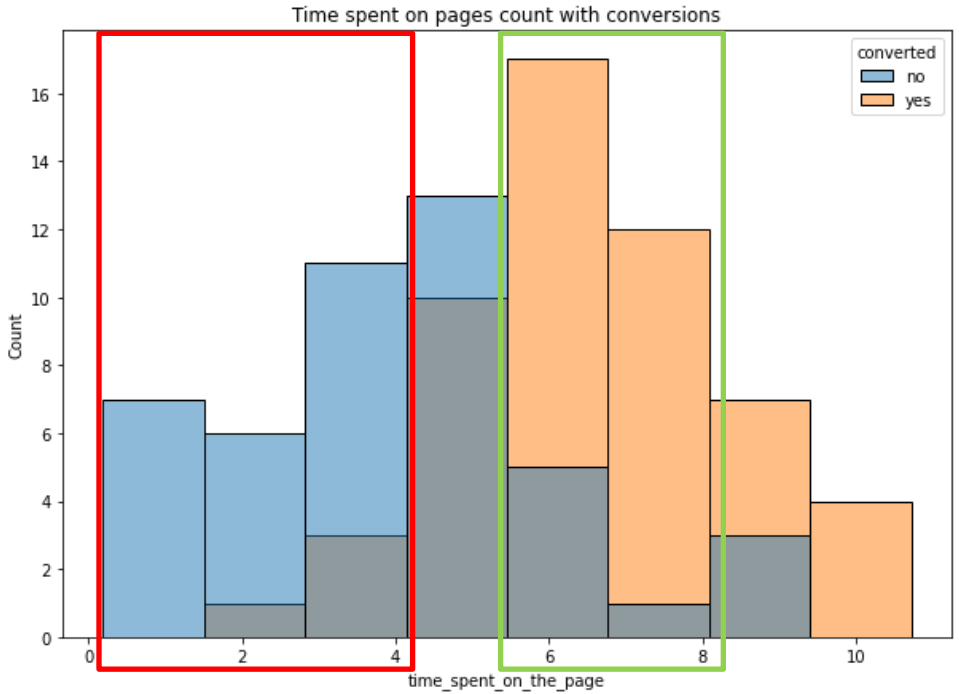
- Based on EDA analysis, both converting and non-converting visitors tend to spend a bit more time on the new page.
- New landing page starts converting at 4.5 min spent on page (or 5.5 min for the 25 percentile of IQR), while the old landing page starts converting as low as 3 min spent on page (or just over 5 min for the 25th percentile of IQR).
- One could say that the old landing page requires somewhat less time spent on page for user visits to result in conversions.
- The old page has bigger distributions of non-converting users in the quartile (IQR) ranges from 1.75 – 4.5 min.

Exploratory Data Analysis (EDA)



- Here we see that for the converted (yes) visitors, both NEW and OLD pages have a relatively speaking similar quartile ranges (IQR) from around 5-7 min for the OLD page, and from 6-7.5 min for the NEW page.
- Unlike the converted users/visitors, we see that the non-converted (no) users have more differing times spent on the landing pages for the OLD and NEW landing pages (IQRs). The OLD landing page 25th and 75th quartile range is from 1.75-4.5 min, while for the NEW page, that IQR range is from 4-5.75min.

Exploratory Data Analysis (EDA)



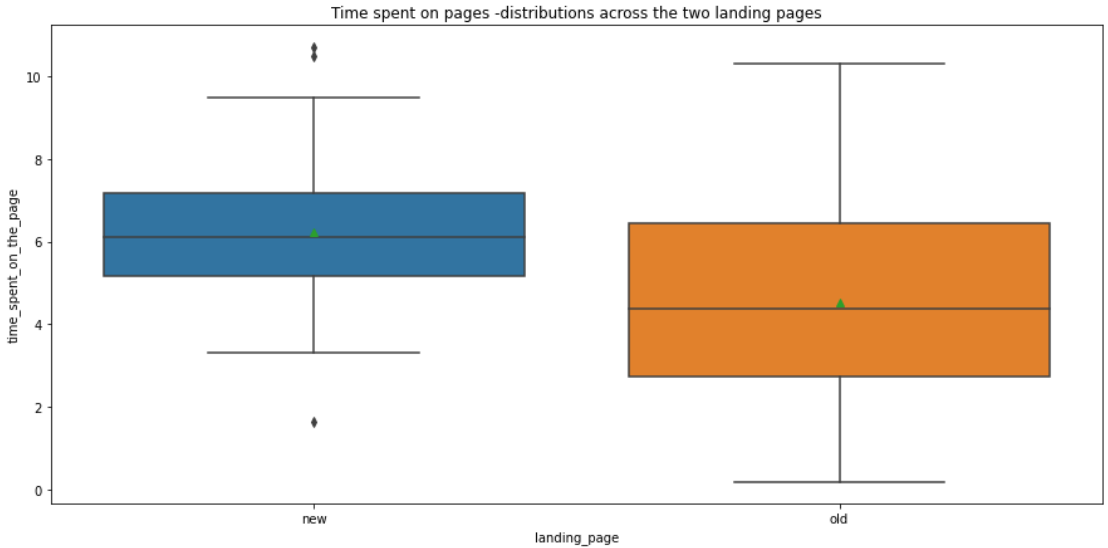
- Here we see that the largest count of converted visitors for both landing pages happens from 5.5 to 8 minutes.
- The largest count for non-converted (no) visitors happens between 0.3 to 4.3 minutes spent on the landing pages.

Business Insights and Recommendations

Conclusions about time spent on pages and conversions

- It looks like users spend overall (cumulatively and statistically speaking) more time on the new version of the landing page vs. the old version.
- There is also statistical significance in the conversion rate of the NEW page, which is likely significantly greater than the conversion rate of the OLD page.
- This could potentially mean several things:
  - The new page might have more content, and so it takes longer time to read it.
  - The content on the new page is more interesting and more related, so the users take more time to read it.
  - The new page has different content from the old content, which could be less or more related, hence more time spent on pages.
  - Or the new page is written in a less clear manner/format that is spread out over the webpage real estate, hence it takes longer time to read it in order to understand it.

Time spent on page distribution across landing pages

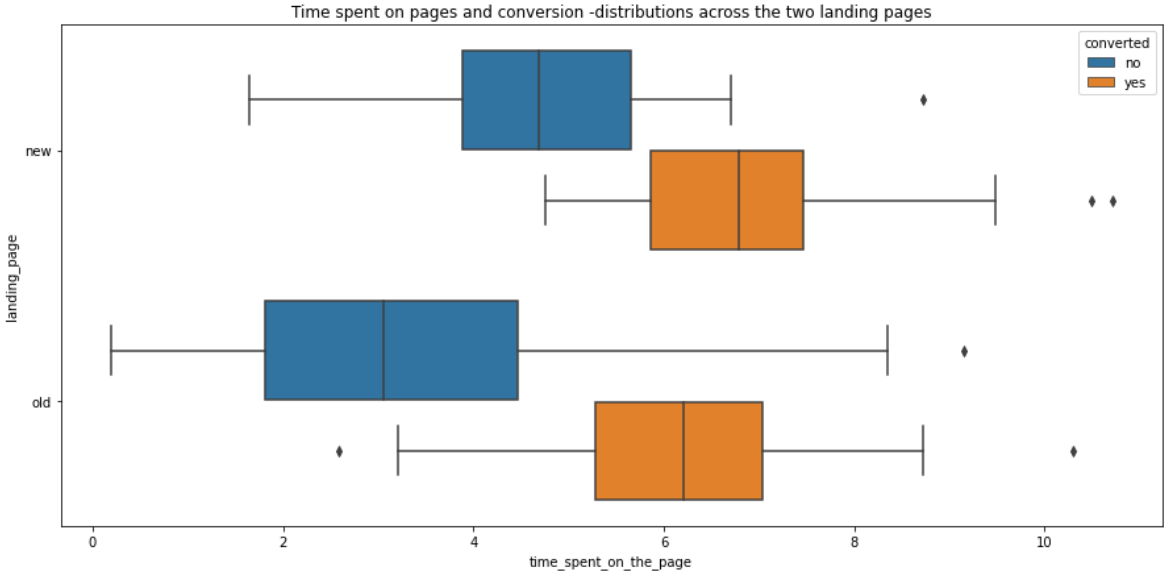


- Less information is visible from the time spent on overall pages.

- While overall pages show that more time is spent on the NEW page (especially looking at interquartile ranges), the OLD page has wider overall range of time spent, and there is a big overlap of times between the two pages.

- Next slide shows more details...

Time spent on page distribution across landing pages



- With this chart, you can see clearer separation between the winning (converting) NEW page and its times vs. the OLD (losing or non-converting) page and its times spent on page.

- The NEW landing page has more time spent on page for both, converting and non-converting users.

- However, the difference between the converting users is smaller between NEW and OLD pages (there is a bigger overlap) vs. the difference between the non-converting users, which seem to clearly spend more time on the NEW vs. OLD page.

Business Insights and Recommendations

Future explorations and next steps related to time spent on page & conversions

- Since the difference in time spent on NEW vs. OLD landing page with converted visitors is smaller than the difference in time spent on NEW vs. OLD landing pages with non-converted visitors, my suggestion is to explore this difference further and conduct more tests as the next steps.
    - Something there should be investigated, because the visual separation between the converted and non-converted users, in terms of the time spent on page, is clearer with respect to the OLD page than the NEW page.
    - With the NEW page, the 5.5 minutes spent on page are critical, because it seems for the most converted visitors if they get over that hump of staying on the page for 5.5 minutes, they tend to convert. The company should do whatever possible (on the NEW page) to keep the users on the page longer than 5.5 minutes.
- Devise A/B tests to further explore the use of and types of content, and potentially reposition the CTA (calls to action) below the fold so that users have to read more and stay more on the page prior to being pushed into action.

Business Insights and Recommendations

Conclusions about the time spent on page, conversions, and language preferences

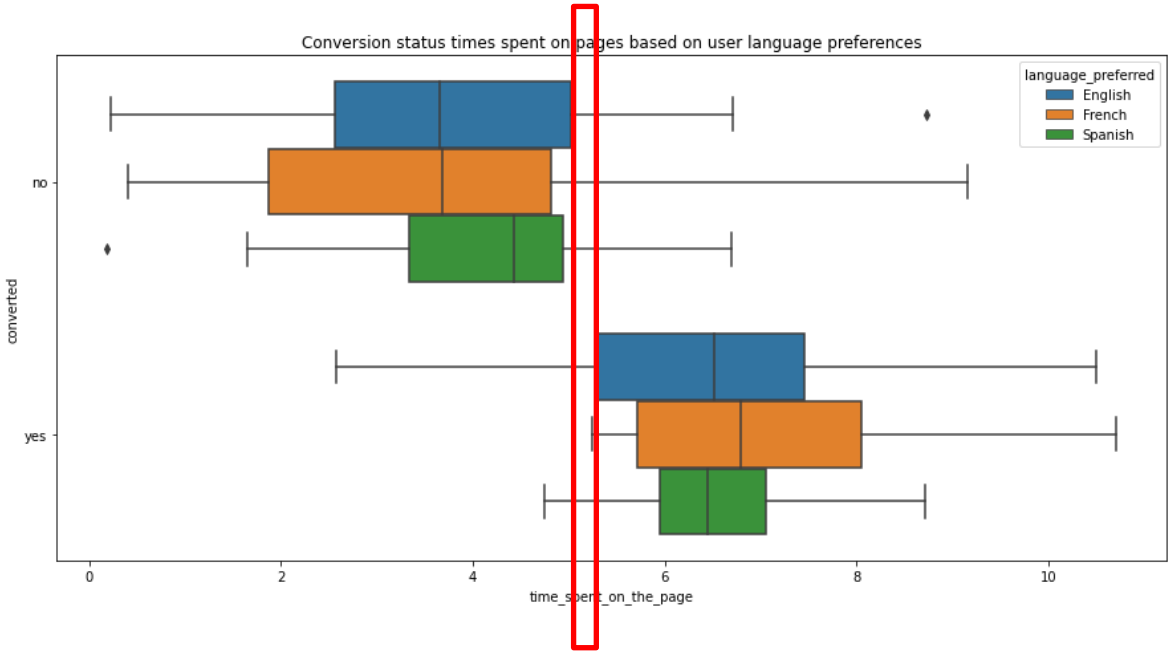
- As with the previous section of this presentation related to times spent on page, this section related to languages also shows that the 5th minute of reading represents a critical “hump” for users of all 3 language preferences. Starting with the 25th percentile of the users for all 3 languages, if they spend more than 5 minutes on the page, they seem more likely to belong to the converted (yes) users.
- English readers start converting the soonest (25th percentile) at around 5 minutes, French at 5.5 minutes, and Spanish at 6 minutes of reading.
- On the other end of the spectrum, French readers who do NOT convert start dropping off the page at 2 minutes, English at 2.5 minutes, and Spanish visitors take the longest to start dropping off, starting at almost 3 minutes of reading.
- Spanish language preferences have the shortest IQRs within both, non-conversion and conversions.
- Overall, all 3 languages show similar patterns with respect to time spent on page for 'yes' conversions, and separately with respect to time spent on page for the 'no' conversions (except French).

Business Insights and Recommendations

Conclusions about the time spent on page, conversions, and language preferences [...*continued from the previous slide*]

- It's important to note, statistically speaking, that we do NOT have enough evidence to say that there are statistically significant differences in the mean times spent on the NEW page for different language users.
- In other words, the mean times spent on the NEW page are independent of the users' language preferences, so the impact of the reader's (user's) language is likely not significant or less significant in terms of how long such user may stay on the page.
- We also do NOT have enough evidence to claim that the converted status depends on the language preferences (overall – for both landing pages combined).
- However, when we filter out and separate the two landing pages for language preferences and conversions, we see that French language preferences performed the worse on the OLD landing page.

Time spent on page (overall – both landing pages) and language preferences



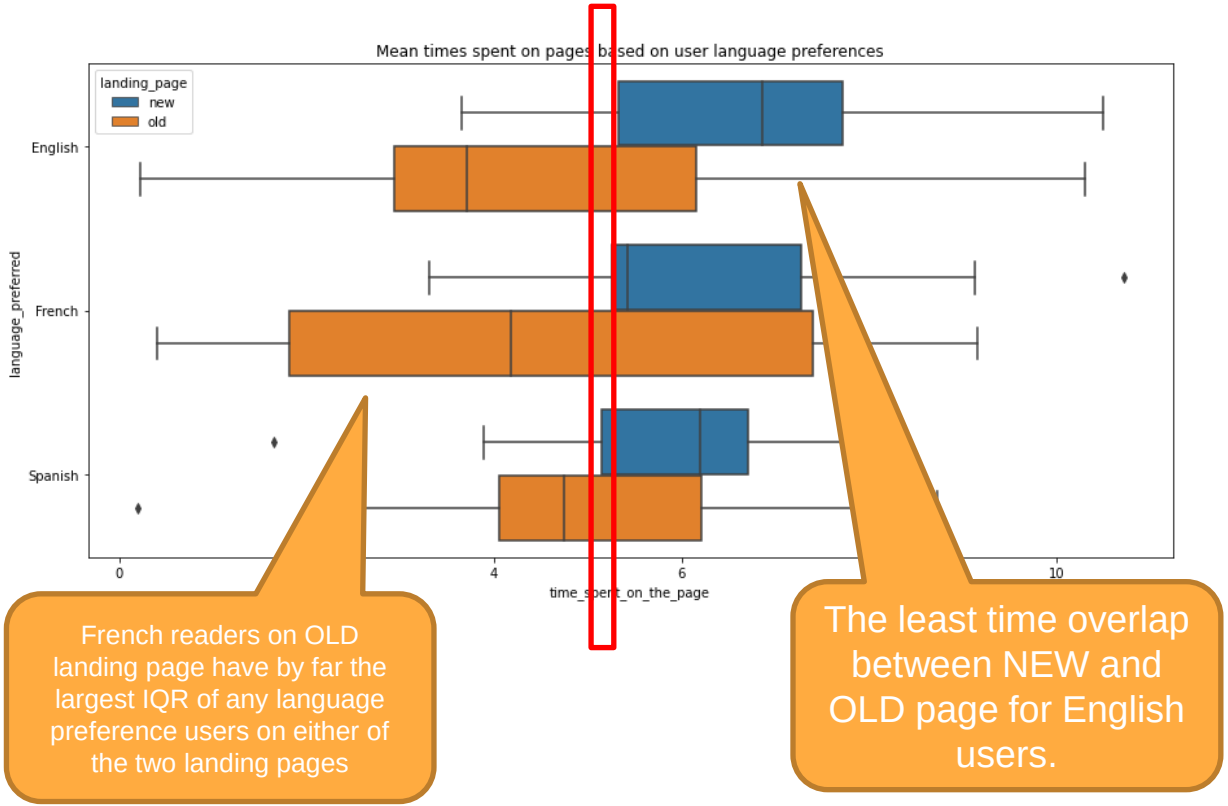
- Red line on the chart shows a clear separation between the non-converting users (upper portion of the chart), and converting users (lower portion) with respect to all 3 language preferences.

- You can see that most non-converting visitors with respect to all 3 language preferences spend less than 5 minutes on the page for the 75th percentile.

- Similarly, for all 3 languages, the 25th percentile of converted visitors and above spend more than 5 minutes on the page.

- The Spanish readers seem to have the shortest ranges overall in both categories, converted and non-converted visitors.

Time spent on page (overall – both landing pages) and language preferences



- On this chart, we are observing times spent on either NEW or OLD pages based on language preferences only (not based on conversions).

- On the NEW landing page (blue color), you can see that starting with the 25th percentile for all 3 language preferences users spend at least 5 minutes on the page.

- However, for different language preferences the 75th percentiles end at various minutes spent on page. English spend more time than French, which spend more time than Spanish.

- For French users only, they spent more time on OLD page than on the NEW page.

Time spent on page and language preferences

French readers on OLD landing page have by far the largest number of non-conversions of any language preference within any landing page.

French readers/visitors experience the best improvement on the NEW landing page. English visitors are the only ones to drop very slightly in conversion rate.

language_preferred	converted	no	yes
English		5	11
French		14	3
Spanish		10	7

OLD landing page

language_preferred	converted	no	yes
English		6	10
French		5	12
Spanish		6	11

NEW landing page

- On this slide, we see two tables with conversion rates for different language preferences. The left table is for the OLD landing page, and the table on the right is for the NEW landing page.
- Note that the French language preference has the least number of conversions and the largest number of non-converted users - almost 5 times as many non-conversions - on the OLD landing page.
- Simultaneously, French visitors experience the best improvement on the NEW landing page.
- Spanish language preference has the 2nd best improvement on the NEW page.

Business Insights and Recommendations

Future explorations and next steps related to time spent on page, conversions, and language preferences

- We would need to further understand how and why times spent on page benefit conversion rates for both NEW and OLD landing pages and based on language preferences.
- The company should invest in the common UX (user experience) practices, and develop user personas and user journeys to understand how the sources of traffic and other pertinent user data (e.g. demographics, psychographics, etc.) influence conversion decisions and navigation to the landing pages and from the landing pages (including language preferences).
- While we can statistically show that the converted status does NOT depend on the preferred language (calculated for BOTH landing pages), we also see statistically speaking that within the OLD landing page the converting status DOES depend on the preferred language.
- And we also intuitively see, based on data, that readers with French language preference had the poorest conversion on the OLD page vs. NEW landing page, with best overall improvement on the NEW landing page.

Business Insights and Recommendations

Recommended A/B tests as part of next steps

1. We could experiment with English language preference, which has dropped by one conversion from OLD to NEW landing page.
2. Since both English and Spanish language preferences both show similar conversion rates on the NEW page, yet Spanish convert on the NEW page by spending much less time, we could hypothesize that by reducing the time spent on page for English visitors, we could further improve their conversion rate.
3. We could also make an assumption and test that hypothesis, that the long times spent on landing pages is negatively correlated with conversion rates as they relate to language preferences and try it out in separate A/B test.

Business Insights and Recommendations

To support such A/B tests with regards to language preferences [...*continued from the previous slide*]

- We could explore hiring writers that are native speakers in the targeted languages and seeing if we can move the subscription conversion needle rates up for all 3 language preferences by strategically placing a more native content structures around CTA (calls to action, e.g. request for email address).
- By creating original content in natively translated languages and AB testing those, we might be able to further improve subscription base for all 3 language preferences.
- We could also explore what specific page elements (buttons, options, links, and other CTAs (calls to action) were selected during visitors' journeys, and how it correlates with the decision to convert or not, and as related to the language preferences.

Business Insights and Recommendations

Explore other data in relations to online analytics and as part of A/B tests [...*continued from the previous slide*]

- The company should also investigate other types of online data and variables resulting from the tests, such as:
  - Other pages visited before/after the landing page (pages per session)
  - Overall session lengths and session by source
  - Bounce rates, common user paths / user flows
  - Users' browser information, screen size and device information
  - On-page viewing, and many other bits of contextual information in order to understand where and how to improve the conversion rates

Appendix

Data dictionary

1. **user_id** - This represents the user ID of the person visiting the website.
2. **group** - This represents whether the user belongs to the first group (control) or the second group (treatment).
3. **landing_page** - This represents whether the landing page is new or old.
4. **time_spent_on_the_page** - This represents the time (in minutes) spent by the user on the landing page.
5. **converted** - This represents whether the user gets converted to a subscriber of the news portal or not.
6. **language_preferred** - This represents the language chosen by the user to view the landing page.

greatlearning
Power Ahead

Happy Learning !