

#### PROJECT | BUSINESS STATISTICS

# DATA-DRIVEN STRATEGIES FOR E-NEWS EXPRESS

SHAISHAV MERCHANT

## AGENDA

•	Executive Summary	03
•	Business Problem Overview & Solution Approach	<u>05</u>
•	Data Overview	<u>08</u>
•	EDA - Univariate Analysis	<u>10</u>
•	EDA – Bivariate Analysis	<u>13</u>
•	Hypothesis Testing	<u>17</u>
•	Conclusion & Recommendations	22

## E-NEWS EXPRESS

## **EXECUTIVE SUMMARY**

Executive summary distills key findings from
E-news Express's landing page evaluation,
spotlighting data-driven insights into user
engagement and conversion efficacy. This
synthesis guides strategic decisions, ensuring that
their digital footprint aligns with user preferences
and market trends.

## **EXECUTIVE SUMMARY**

- WE EMBARKED ON A CRITICAL EVALUATION OF E-NEWS EXPRESS'S REVAMPED LANDING PAGE TO GAUGE ITS IMPACT ON USER INTERACTION AND RETENTION.
- E-NEWS EXPRESS'S COMPREHENSIVE DATASET SPANNED 100 USERS, DELVING INTO NUANCED BEHAVIORAL PATTERNS ACROSS DIFFERENT LANGUAGES.
- A METICULOUSLY BALANCED A/B TEST SETUP PROVIDED A ROBUST FOUNDATION FOR THE ANALYSIS, REINFORCING THE INTEGRITY OF THE FINDINGS.
- I EMPLOYED EXPLORATORY DATA ANALYSIS AND VARIOUS HYPOTHESIS TESTS SUCH AS 2-SAMPLE T-TEST, 2-PROPORTIONS Z-TEST, CHI-SQUARE TEST OF INDEPENDENCE, AND ANOVA (ANALYSIS OF VARIANCE) TO DRAW CONCLUSIONS FOR THE KEY RESEARCH QUESTIONS.
- DATA QUALITY WAS RIGOROUSLY UPHELD. ENSURING THE INSIGHTS DRAWN ARE BOTH RELIABLE AND ACTIONABLE.
- THE OBSERVED 54% CONVERSION RATE IS A TESTAMENT TO THE NEW PAGE'S CAPABILITY TO CAPTIVATE AND CONVERT VISITORS EFFECTIVELY.
- THE INCREASED AVERAGE ENGAGEMENT TIME OF 6.22 MINUTES ON THE NEW PAGE, COMPARED TO 4.53 ON THE OLD, SUGGESTS THE REVAMPED CONTENT SIGNIFICANTLY BOOSTS USER INTEREST.
- THE NEW LANDING PAGE HAS SHOWN A MARKED IMPROVEMENT IN USER CONVERSION RATE, SIGNALING A STEP IN THE RIGHT DIRECTION.
- ENGLISH-LANGUAGE USERS DEMONSTRATED THE HIGHEST CONVERSION RATES, REVEALING LANGUAGE'S SIGNIFICANT ROLE IN USER ENGAGEMENT DYNAMICS.

## **BUSINESS PROBLEM OVERVIEW**& SOLUTION APPROACH

Facing a subscriber decline in a competitive digital news market, E-news Express employs A/B testing to evaluate its webpage designs. Utilizing Python's SciPy and seaborn, we analyze user interactions to understand the impact of page design on time spent and conversion rates, aiming to boost engagement and subscriptions.

## PROBLEM OVERVIEW

#### CONTEXT

E-NEWS EXPRESS, AN ONLINE NEWS PORTAL, SEEKS TO EXPAND ITS SUBSCRIBER BASE AMIDST THE RAPID SHIFT TO ELECTRONIC NEWS. WITH THE DIGITAL LANDSCAPE OFFERING NOVEL WAYS TO ENGAGE AUDIENCES, THE COMPANY FACES THE CHALLENGE OF ADAPTING ITS ONLINE PRESENCE TO ENHANCE USER ENGAGEMENT AND INCREASE SUBSCRIPTIONS.

#### **OBJECTIVE**

- 1. DETERMINE IF THE NEW LANDING PAGE DESIGN SIGNIFICANTLY INCREASES USER ENGAGEMENT COMPARED TO THE EXISTING PAGE.
- 2. ASSESS WHETHER THE NEW PAGE DESIGN POSITIVELY IMPACTS THE CONVERSION RATE OF VISITORS TO SUBSCRIBERS.
- 3. INVESTIGATE IF USER CONVERSION IS INFLUENCED BY THEIR PREFERRED LANGUAGE.
- 4. ANALYZE IF THE TIME SPENT ON THE NEW PAGE VARIES AMONG USERS OF DIFFERENT LANGUAGES.

## SOLUTION APPROACH

IN ADDRESSING THE CHALLENGE OF BOOSTING USER ENGAGEMENT AND SUBSCRIPTIONS FOR E-NEWS EXPRESS, MY APPROACH IS SEGMENTED INTO THREE KEY PHASES. IT IS DESIGNED TO THOROUGHLY ANALYZE USER INTERACTIONS, PROVIDING A FOUNDATION FOR INFORMED ENHANCEMENTS TO THE DIGITAL STRATEGY.

#### 1. EXPLORATORY DATA ANALYSIS (EDA):

• CONDUCT AN IN-DEPTH EXAMINATION OF USER INTERACTIONS, FOCUSING ON IMPORTANT METRICS LIKE TIME SPENT ON THE PAGE, CONVERSION RATES AND PREFERRED LANGUAGE.

#### 2. DATA VISUALIZATION:

EMPLOY TOOLS LIKE SEABORN TO TRANSLATE EDA FINDINGS INTO CLEAR, VISUAL REPRESENTATIONS, HIGHLIGHTING DATA PATTERNS AND INSIGHTS.

#### 3. HYPOTHESIS TESTING:

• USE PYTHON'S SCIPY LIBRARY FOR RIGOROUS HYPOTHESIS TESTS, ASSESSING THE EFFECTIVENESS OF THE NEW LANDING PAGE IN COMPARISON TO THE EXISTING ONE, PARTICULARLY IN TERMS OF USER ENGAGEMENT AND CONVERSION EFFICIENCY.

THIS APPROACH ENSURES A COMPREHENSIVE ANALYSIS, LEADING TO STRATEGIC CONCLUSIONS AND RECOMMENDATIONS FOR E-NEWS EXPRESS.

## DATA OVERVIEW

The analysis utilizes E-news Express's dataset, capturing user interactions with variables like group allocation, landing page version, time spent, conversion status, and language preference.

This data is essential for assessing the effectiveness of the new landing page on user engagement and subscription trends.

## DATA OVERVIEW

COLUMN	TYPE	REMARKS
user_id	int64	Unique identifier for each website visitor
group	object	User's assigned group (control or treatment)
landing_page	object	Version of landing page viewed (old or new)
time_spent_on_the_page	int64	Time in minutes on page
converted	object	Subscription conversion status (yes or no)
language_preferred	object	User's chosen language for page (English, Spanish or French)

RECORDS	COLUMNS	DATA TYPE
100	6	2 (int64, object)

DATA DESC.	MEAN	SD	MIN	50%	MAX
time_spent	5.38	2.38	0.19	5.42	10.71

#### **Dataset Information:**

The dataset analysis will inform how landing page design influences user engagement for E-news Express.

#### Key Insights:

- Dataset covers 100 users, each with 6 variables.
- 'French' (34) and 'Spanish' (34) languages are preferred over 'English' (32).
- A 54% conversion rate may indicate effective user acquisition.
- Equal split in control and treatment groups validates the A/B test setup.
- Average page engagement is 5.38 minutes, with a reasonable variation in user interaction times.
- Clean data with no nulls or duplicates ensures analysis accuracy.

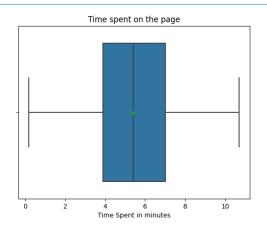
#### Conclusion:

This clean and balanced dataset sheds light on the relationship between page design and user subscription behavior.

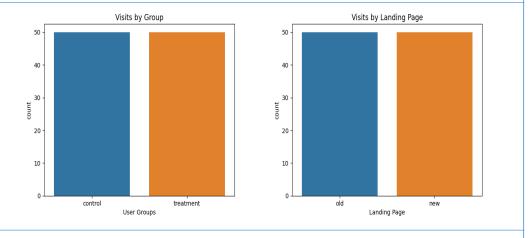
Univariate analysis of E-news Express's user data reveals insights into time spent on the site, user group distribution, conversion rates, and language preferences. The findings suggest a well-designed A/B test environment with balanced language representation, demonstrating the new landing page's potential for higher subscriber conversion.

#### Time Spent on the Page

# 

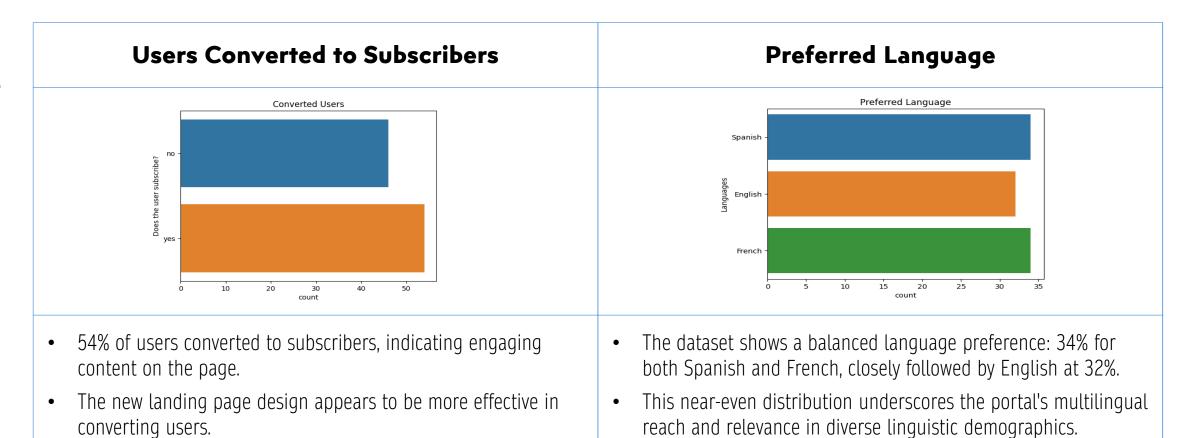


#### **Groups and Landing Page**

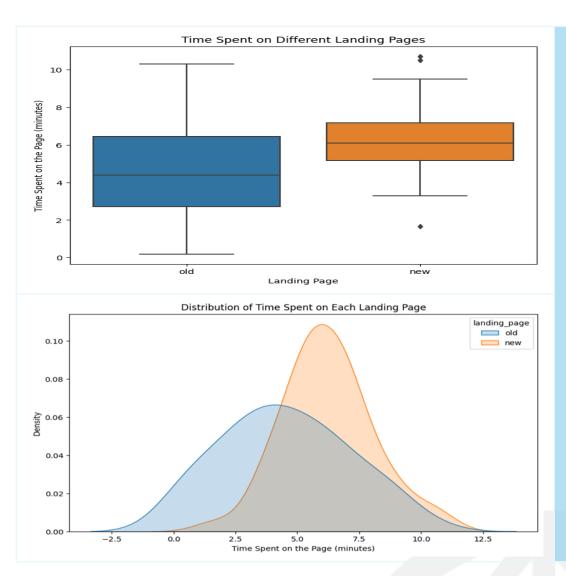


- Engagement ranges from 0.19 to 10.71 minutes, with an average of 5.38 minutes.
- Normal distribution with no outliers depicting consistent user behavior.
- Half of the users stay longer than the median time of 5.42 minutes.
- Most users' engagement falls between 3.88 minutes (Q1) and 7.02 minutes (Q3).

- The 'group' column contains two categories: Control and Treatment.
- Each group has 50 users, guaranteeing a balanced comparison.
- Control group users access the existing landing page, while Treatment group users view the new page.
- This even split is crucial for valid A/B testing results.

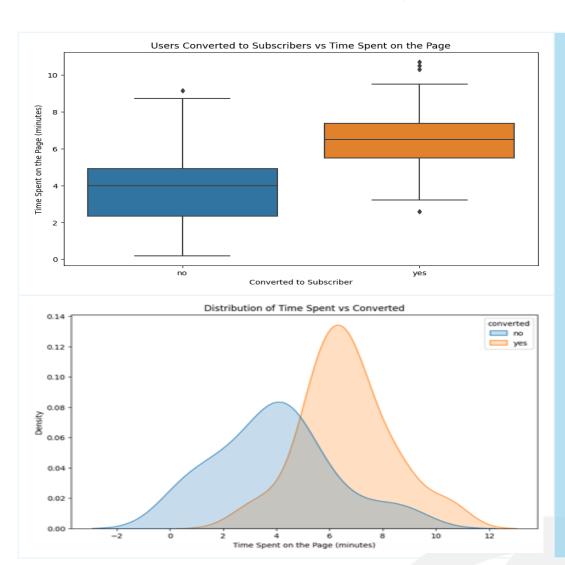


The bivariate analysis is set to delve into the interaction between landing page designs and user engagement. By examining time spent on each page variant, user conversion patterns and preferred language, we expect to uncover actionable insights to guide E-News Express in refining their content delivery, enhancing the user experience.



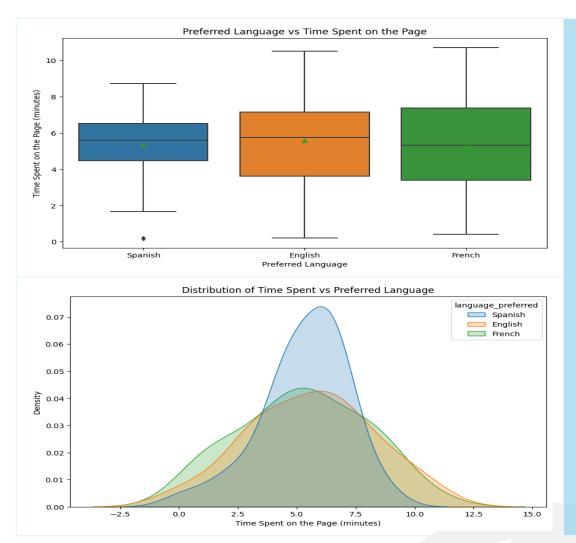
#### Landing Page vs Time Spent on the Page

- Users spend more time on the new landing page than old, averaging 6.22 minutes to 4.53 minutes.
- The median time spent is also higher on the new page (6.10 minutes) than on the old page (4.38 minutes), indicating most users stay longer on the new page.
- The range of time spent shows wider distribution on the old page (0.19 to 10.30 minutes) than on the new page (1.65 to 10.71 minutes).
- The new landing page exhibits more consistent user engagement, with a standard deviation of 1.82 minutes, compared to 2.58 minutes for the old page.
- Conclusion: The new landing page increases user engagement time, indicating improved design efficacy.



#### Converted Users vs Time Spent on the Page

- Non-converters typically stay for 3.92 minutes, whereas subscribers spend a more substantial 6.62 minutes engaging with the page.
- The median visit duration reaffirms this trend: a brief 3.98 minutes for non-converters against an increased involvement of 6.50 minutes for converters.
- The engagement pattern is steadier among converters who generally spend between 2.58 to 10.71 minutes, compared to the wider and more erratic range of non-converters from 0.19 to 9.15 minutes.
- Conclusion: A direct correlation emerges between extended, consistent time spent on the page and the likelihood of conversion, highlighting the importance of engaging content in driving subscriptions.



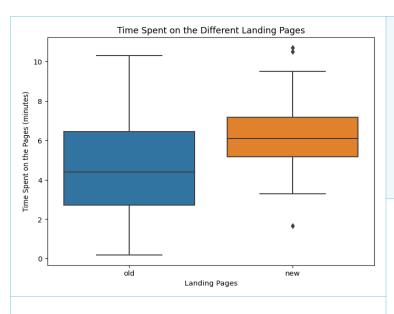
#### Preferred Language vs Time Spent on the Page

- English users average 5.56 minutes, with a broad spread of engagement times.
- French users spend on an average 5.25 minutes, showing varied interaction patterns.
- Spanish users engage more consistently, averaging 5.33 minutes.
- **Conclusion**: Engagement varies across languages, with Spanish users showing the most consistent interaction, while English and French users have wider variations in their time spent.

## HYPOTHESIS TESTING

Hypothesis tests examine if the new landing page design significantly affects user engagement and conversion rates, considering variables like time spent and language preference. This analysis will inform strategic decisions for E-news Express's digital transformation efforts.

## DO THE USERS SPEND MORE TIME ON THE NEW LANDING PAGE THAN THE EXISTING LANDING PAGE?



#### **Observations:**

- Average time: New Page: 6.22 mins; Old: 4.53 mins,
- Median time: New: 6.10 mins; Old: 4.38 mins.
- Time range: New: 1.65 to 10.71 mins; Old: 0.19 to 10.30 mins.
- Standard deviation: **unequal**. New: 1.82 mins; Old: 2.58 mins.
- Both populations are assumed to be normally distributed.
- Samples are drawn from independent populations.

#### 1. Formulate Hypothesis:

**Null Hypothesis:**  $(H_0: \mu_2 \le \mu_1)$  Average time spent on the new page  $(\mu_2)$  is **less than or equal to** the average time spent on the old page  $(\mu_1)$ .

Alternative Hypothesis:  $(H_a: \mu_2 > \mu_1)$  Average time spent on the new page  $(\mu_2)$  is **greater than** the average time spent on the old page  $(\mu_1)$ .

#### 2. Select Appropriate Test:

I'll conduct a **one-tailed two-sample t-test for independent populations** with unknown variances. I've ensured that assumptions of continuity, normality, independence, and random sampling are met.

#### 3. Significance Level ( $\alpha$ ): set to 5% (0.05)

#### 4. Collect and Prepare Data:

Standard deviations are unequal. New page: 1.82 mins; Old page: 2.58 mins.

#### 5. Calculating p-value:

p-value is 0.000139 (rounded to 6 sf)

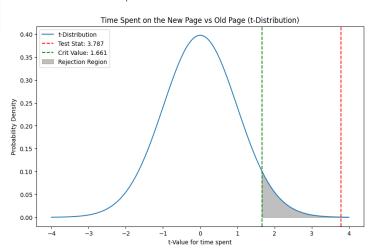
#### 6. Compare p-value to Significance Level ( $\alpha$ ):

p-value (0.000139) is less than  $\alpha$  (0.05).

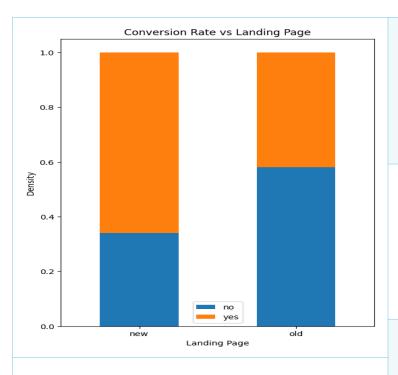
#### 7. Draw Inference:

The p-value of 0.000139 is below the 0.05 threshold which leads us to **reject the null hypothesis**. Therefore, we have sufficient evidence that at the 5% significance level **users** spend more time on the new page than the old page.

Test Statistic: 3.787, Critical Value: 1.661



## IS THE CONVERSION RATE FOR THE NEW PAGE GREATER THAN THE CONVERSION RATE FOR THE OLD PAGE?



#### Observations:

- 33 out of 50 users (66%) who visited the New Landing Page subscribed.
- Contrastingly, the Old Landing page saw a 42% conversion rate (21 out of 50 users).
- The data indicates that the New Landing Page outperforms the Old Landing Page in converting visitors into subscribers.

#### 1. Formulate Hypothesis:

**Null Hypothesis:** ( $H_0$ :  $\rho_2 \le \rho_1$ ) Conversion rate on the new page ( $\rho_2$ ) is **less than or equal to** the old page ( $\rho_1$ ).

Alternative Hypothesis: ( $H_a$ :  $\rho_2 > \rho_1$ ) Conversion rate on the new page ( $\rho_2$ ) is **greater than** on the old page ( $\rho_1$ ).

#### 2. Select Appropriate Test:

A **one-tailed two-proportion z-test** is selected as it compares the proportion of conversions between two independent groups. Test prerequisites: binary outcomes, random sampling, and sufficient size for normal approximation are met.

3. Significance Level ( $\alpha$ ): set to 5% (0.05)

#### 4. Collect and Prepare Data:

Equal sample sizes of 50 for the new and old pages show a higher conversion rate on the new page at 66% (33 users) compared to 42% (21 users) on the old page.

#### 5. Calculating p-value:

p-value is 0.008026 (rounded to 6 sf)

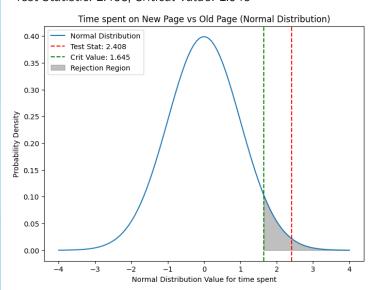
#### 6. Compare p-value to Significance Level ( $\alpha$ ):

p-value (0.008026) is less than  $\alpha$  (0.05).

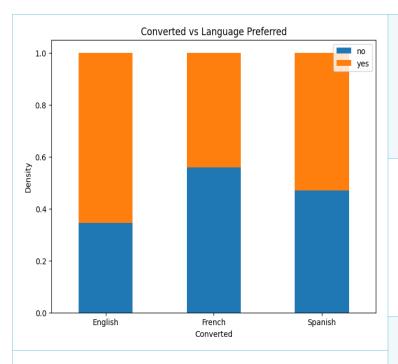
#### 7. Draw Inference:

With a p-value of 0.008026 below the 0.05 significance level, we **reject the null hypothesis** indicating statistically significant evidence that **the New Landing Page has a higher conversion rate than the Old Landing Page**.

Test Statistic: 2.408, Critical Value: 1.645



## DOES THE CONVERTED STATUS DEPEND ON THE PREFERRED LANGUAGE?



#### **Observations:**

- English leads with 66% conversion rate (21 out of 32).
- Followed by Spanish with a 53% rate (18 out of 34).
- And lastly, French has a 44% rate (15 out of 34).
- English language preference appears to have the highest conversion rate.
- To confirm the claim that conversions are dependent on language preference, a detailed hypothesis test would be needed.

#### 1. Formulate Hypothesis:

**Null Hypothesis:** (H<sub>0</sub>) Conversion rate is independent of the preferred language.

Alternative Hypothesis: (H<sub>a</sub>) Conversion rate is dependent on the preferred language.

#### 2. Select Appropriate Test:

Given two categorical variables—converted status and preferred language—and sufficient sample size with all expected frequencies above 5, the **chi-squared test of independence** is suitable for evaluating their dependency. Random sampling further validates its use.

3. Significance Level ( $\alpha$ ): set to 5% (0.05)

#### 4. Collect and Prepare Data:

A contingency table created using the crosstab() to compare conversions across languages. Among the 3 languages, English has the highest conversion rate of 66%.

#### 5. Calculating p-value:

p-value is 0.212989 (rounded to 6 sf)

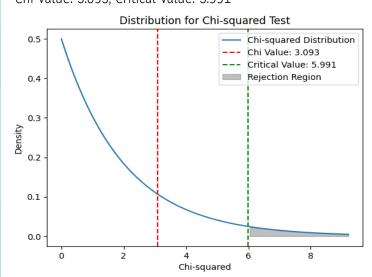
#### 6. Compare p-value to Significance Level ( $\alpha$ ):

p-value (0.212989) is greater than  $\alpha$  (0.05).

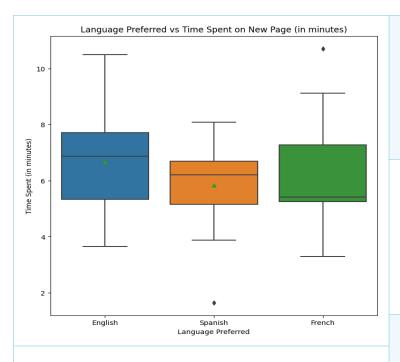
#### 7. Draw Inference:

The observed p-value of 0.212989 exceeds the 0.05 significance level. Hence, we **fail to reject the null hypothesis**. This helps us conclude that there is **insufficient evidence to support the claim** that user conversion rate is dependent on the preferred language.

Chi Value: 3.093. Critical Value: 5.991



## IS THE TIME SPENT ON THE NEW PAGE SAME FOR THE DIFFERENT LANGUAGE USERS?



#### **Observations:**

- English users average 6.66 minutes on the new page.
- French users spend 6.20 minutes on average.
- Spanish users' average time is 5.84 minutes.
- The differences in mean time spent by language preference indicates a potential impact of language on user engagement.
- To confirm this claim, a detailed hypothesis test would be needed.

#### 1. Formulate Hypothesis:

**Null Hypothesis:** ( $H_0$ :  $\mu_e = \mu_s = \mu_f$ ) The mean time spent on the new lading page is the same for all languages.

Alternative Hypothesis: At least one language preference group shows different mean time spent on the new page.

#### 2. Select Appropriate Test:

Given three population means from normally distributed and equal variance populations, a **one-way ANOVA**, supported by Shapiro-Wilk's and Levene's tests for normality and equal variance, is the suitable choice for hypothesis testing.

#### 3. Significance Level ( $\alpha$ ): set to 5% (0.05)

#### 4. Collect and Prepare Data:

Shapiro-Wilk's (p=0.80) and Levene's (p=0.47) tests validate normality and homogeneity of variance, enabling the ANOVA test application. Data was randomly sampled and categorized by language.

#### 5. Calculating p-value (ANOVA):

p-value is 0.432041 (rounded to 6 sf)

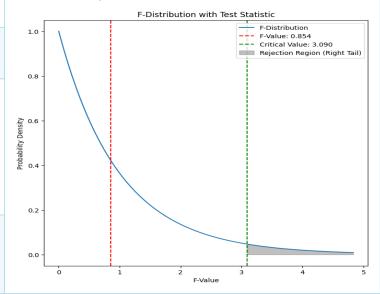
#### 6. Compare p-value to Significance Level ( $\alpha$ ):

p-value (0.432041) is greater than  $\alpha$  (0.05).

#### 7. Draw Inference:

With a p-value of 0.432041 greater than 0.05, we **fail to reject the null hypothesis**. This suggests that there is **no significant effect** of language preference on the time users spend on the new page.

f-Value: 0.854, Critical Value: 3.090



### CONCLUSION & RECOMMENDATIONS

Conclusion and recommendations distill the essence of our findings, offering strategic directives based on solid data analysis. It will pinpoint the actionable steps E-news Express can take to leverage their new landing page, maximizing user engagement and conversion rates in a linguistically diverse user base.

## CONCLUSION

- ANALYTICAL RIGOR CONFIRMED THAT THE NEW LANDING PAGE ENHANCES USER ENGAGEMENT, A VICTORY FOR E-NEWS EXPRESS'S DESIGN AND CONTENT TEAMS.
- HYPOTHESIS TESTING ROBUSTLY SUPPORTED THE SUPERIORITY OF THE NEW LANDING PAGE DESIGN IN FOSTERING USER ENGAGEMENT AND INCREASING CONVERSION LIKELIHOOD.
- ADHERENCE TO STATISTICAL PREREQUISITES, SUCH AS NORMALITY AND EQUAL VARIANCE, SOLIDIFIED THE FOUNDATION FOR HYPOTHESIS TESTS.
- THE SURGE IN CONVERSION RATES FOR THE NEW PAGE UNDERSCORES ITS EFFICACY AND ITS POTENTIAL FOR FURTHER OPTIMIZATION.
- FOR THE HYPOTHESIS ON TIME SPENT ON THE NEW VERSUS OLD LANDING PAGE, OUR ANALYSIS CONFIRMS THAT USERS ARE MORE ENGAGED WITH THE NEW DESIGN. SPENDING A SIGNIFICANTLY LONGER TIME ON AVERAGE.
- THE CONVERSION RATE HYPOTHESIS REVEALS A HIGHER RATE FOR THE NEW PAGE, INDICATING THAT ITS DESIGN RESONATES BETTER WITH VISITORS, LEADING TO MORE SUBSCRIPTIONS.
- REGARDING THE LANGUAGE PREFERENCE, OUR FINDINGS SUGGEST THAT WHILE THERE IS VARIATION IN TIME SPENT, IT DOES NOT SIGNIFICANTLY AFFECT CONVERSION RATES, HIGHLIGHTING THE CONTENT'S UNIVERSAL APPEAL REGARDLESS OF LANGUAGE.
- FINALLY, THE ANALYSIS OF TIME SPENT ACROSS DIFFERENT LANGUAGES SHOWS NO SUBSTANTIAL DIFFERENCE, PROVING THAT THE NEW LANDING PAGE'S DESIGN HAS A CONSISTENTLY POSITIVE IMPACT ACROSS LANGUAGE GROUPS.

## RECOMMENDATIONS

- EMBRACE THE INNOVATIVE NEW LANDING PAGE DESIGN WHOLEHEARTEDLY AS IT HAS PROVEN TO ELEVATE USER EXPERIENCE SIGNIFICANTLY.
- ZERO-IN ON ENHANCING ENGLISH-LANGUAGE CONTENT, SEIZING THE OPPORTUNITY TO AMPLIFY THE STRONGEST CONVERSION RATES.
- CRAFT TARGETED INITIATIVES FOR FRENCH AND SPANISH LANDING PAGES, TAPPING INTO UNCHARTED ENGAGEMENT POTENTIAL TO DRIVE GROWTH.
- CONTINUE THE PRACTICE OF A/B TESTING RIGOROUSLY, FOSTERING A CULTURE OF CONSTANT IMPROVEMENT AND INNOVATION.
- MAINTAIN HIGH STANDARDS FOR DATA EXCELLENCE, ENSURING THAT STRATEGIES ARE INFORMED BY QUALITY-DRIVEN INSIGHTS.
- INVESTIGATE DEEPER INTO LINGUISTIC PREFERENCES, PERSONALIZING CONTENT TO RESONATE WITH THE GLOBAL AUDIENCE'S DIVERSE TAPESTRY.
- LEVERAGE THESE RICH INSIGHTS TO SCULPT CONTENT STRATEGIES AND MARKETING CAMPAIGNS THAT ARE DATA-INFORMED AND USER-CENTRIC.
- COMMIT TO A CYCLE OF REGULAR REVIEW AND REFINEMENT OF ENGAGEMENT METRICS, SUSTAINING THE UPWARD TRAJECTORY OF USER INTERACTION.





## THANK YOU

SHAISHAV MERCHANT
PROJECT – E-NEWS EXPRESS | BUSINESS STATISTICS