***Kanchan Katoghaniya / Utica University / Major Business Analytics / My First Project Report on AB Testing***

**Marketing A/B testing dataset**

Marketing companies want to run successful campaigns, but the market is complex and several options can work. So normally they tun A/B tests, that is a randomized experimentation process wherein two or more versions of a variable (web page, page element, banner, etc.) are shown to different segments of people at the same time to determine which version leaves the maximum impact and drive business metrics.

The companies are interested in answering two questions:

* Would the campaign be successful?
* If the campaign was successful, how much of that success could be attributed to the ads?

With the second question in mind, we normally do an A/B test. The majority of the people will be exposed to ads (the experimental group). And a small portion of people (the control group) would instead see a Public Service Announcement (PSA) (or nothing) in the exact size and place the ad would normally be.

The idea of the dataset is to analyze the groups, find if the ads were successful, how much the company can make from the ads, and if the difference between the groups is statistically significant.

Data dictionary:

* Index: Row index
* user id: User ID (unique)
* test group: If "ad" the person saw the advertisement, if "psa" they only saw the public service announcement
* converted: If a person bought the product, then True, else is False
* total ads: Number of ads seen by person
* most ads day: Day that the person saw the biggest amount of ads
* most ads hour: Hour of day that the person saw the biggest amount of ads

To begin with the AB Testing first datasets was uploaded and number of python library are imported like pandas, numpy, matplotlib, seaborn and warnings and provided python code to read the dataset then after “duplicates” were checked in the “user id” column because user id gives the unique identification to the data of a particular row. And as expected we got the results “No duplicates found in 'user id' column” and then after dropped the column “user id” and “unnamed 0” because that irrelevant to the test of the dataset. Later left with the columns such as 'test group', 'converted', 'total ads', 'most ads day', 'most ads hour'

|  |  |
| --- | --- |
| **test group** | 2 |
| **converted** | 2 |
| **most ads hour** | 24 |
| **most ads day** | 7 |

Checked if the categorical columns have appropriate level

Value counts for column: **test group**

test group

ad 564577

psa 23524

Name: count, dtype: int64

Number of unique levels: 2

Value counts for column: **converted**

converted

False 573258

True 14843

Name: count, dtype: int64

Number of unique levels: 2

Value counts for column: **most ads hour**

most ads hour

13 47655

12 47298

11 46210

14 45648

15 44683

10 38939

16 37567

17 34988

18 32323

9 31004

19 30352

21 29976

20 28923

22 26432

23 20166

8 17627

7 6405

0 5536

2 5333

1 4802

3 2679

6 2068

5 765

4 722

Name: count, dtype: int64

Number of unique levels: 24

Value counts for column: **most ads day**

most ads day

Friday 92608

Monday 87073

Sunday 85391

Thursday 82982

Saturday 81660

Wednesday 80908

Tuesday 77479

Name: count, dtype: int64

Number of unique levels: 7

1. Now begin the Univariate analysis of categorical variable “test group”

Univariate Analysis for 'test group':

|  | **Count** | **Percentage** |
| --- | --- | --- |
| **test group** |  |  |
| **ad** | 564577 | 96.000007 |
| **psa** | 23524 | 3.999993 |

A bar graph with a number of bars

AI-generated content may be incorrect. A blue circle with a number of percentages

AI-generated content may be incorrect.

1. Univariate Analysis for 'converted':

**Count Percentage**

**converted**

**False** 573258 97.476114

**True**  14843 2.523886

A blue circle with a orange triangle and a number of text

AI-generated content may be incorrect.

1. **Univariate Analysis for 'most ads day':**

|  | **Count** | **Percentage** |
| --- | --- | --- |
| **most ads day** |  |  |
| **Friday** | 92608 | 15.746955 |
| **Monday** | 87073 | 14.805790 |
| **Sunday** | 85391 | 14.519785 |
| **Thursday** | 82982 | 14.110161 |
| **Saturday** | 81660 | 13.885370 |
| **Wednesday** | 80908 | 13.757501 |
| **Tuesday** | 77479 | 13.174438 |
|  |  |  |

A graph of blue bars with white text

AI-generated content may be incorrect.

1. Univariate Analysis for 'most ads hour':

|  | **Count** | **Percentage** |
| --- | --- | --- |
| **most ads hour** |  |  |
| **13** | 47655 | 8.103200 |
| **12** | 47298 | 8.042496 |
| **11** | 46210 | 7.857494 |
| **14** | 45648 | 7.761932 |
| **15** | 44683 | 7.597845 |
| **10** | 38939 | 6.621142 |
| **16** | 37567 | 6.387848 |
| **17** | 34988 | 5.949318 |
| **18** | 32323 | 5.496165 |
| **9** | 31004 | 5.271884 |
| **19** | 30352 | 5.161018 |
| **21** | 29976 | 5.097084 |
| **20** | 28923 | 4.918033 |
| **22** | 26432 | 4.494466 |
| **23** | 20166 | 3.429003 |
| **8** | 17627 | 2.997274 |
| **7** | 6405 | 1.089099 |
| **0** | 5536 | 0.941335 |
| **2** | 5333 | 0.906817 |
| **1** | 4802 | 0.816526 |
| **3** | 2679 | 0.455534 |
| **6** | 2068 | 0.351640 |
| **5** | 765 | 0.130080 |
| **4** | 722 | 0.122768 |

A graph of a number of blue bars

AI-generated content may be incorrect.

1. **Univariate Analysis for 'total ads':**

count 588101.000000

mean 24.820876

std 43.715181

min 1.000000

25% 4.000000

50% 13.000000

75% 27.000000

max 2065.000000

Name: total ads, dtype: float64

* The **mean** number of ads is **24.82**, but the **standard deviation is very high (43.72)**, which indicates a **high variability** in the number of ads.
* The **maximum value (2065)** is extremely large compared to the median (13) and 75th percentile (27), suggesting the presence of **strong outliers or a right-skewed distribution**.
* The **75th percentile is 27**, meaning **75% of the data has 27 or fewer ads**.

**The 75th percentile (Q3) for 'total ads' is: 27.00**

Univariate Analysis for 'total ads' values up to the 75th percentile:

count 442253.000000

mean 9.795009

std 7.628273

min 1.000000

25% 3.000000

50% 8.000000

75% 16.000000

max 27.000000

Name: total ads, dtype: float64

A graph of a number of ads

AI-generated content may be incorrect.

When analyzing only the data **up to the 75th percentile**, the **mean drops significantly to 9.80**, and the **standard deviation to 7.63**. This shows that the bulk of the data is **concentrated at lower ad counts**.

* The **median is 8**, and the max is 27 (of Q3).
* This subset is **much more normally distributed** and does **not include extreme outliers**.
* The **overall distribution of total ads is right-skewed**, with a small number of entries having **very high ad counts**.
* Most users have **between 1 and 27 ads**, with the **majority clustered between 3 and 16**.
* Analyzing only up to the 75th percentile gives a **clearer picture of the typical user behavior**, while the full dataset is influenced heavily by **outliers**.

Now begin the bivariate analysis with the target variable “converted”.

**Contingency Table (Counts)**

| **Test Group** | **Not Converted (False)** | **Converted (True)** | **Total** |
| --- | --- | --- | --- |
| **ad** | 550,154 | 14,423 | 564,577 |
| **psa** | 23,104 | 420 | 23,524 |

**Conversion Rates by Test Group**

| **Test Group** | **Conversion Rate (%)** |
| --- | --- |
| **ad** | **2.55%** |
| **psa** | **1.79%** |

**Interpretation**

1. **Conversion Rate Difference**:
   1. The **ad group** has a **higher conversion rate** of **2.55%**, compared to the **psa group's 1.79%**.
   2. This suggests that **ads are more effective** than public service announcements (PSAs) in driving conversions.
2. **Magnitude of Effect**:
   1. The difference between the two groups is **0.76 percentage points**, which is **substantial in high-volume campaigns**.
3. **Practical Insight**:
   1. If the goal is to **maximize conversions**, the **ad group performs better**.
   2. The **psa group may still serve a different purpose** (e.g., awareness, non-commercial goals), but it’s less effective for conversions.

**1. General Trend of most “Ads Hours”**

* **Conversion rates are lowest during early morning hours** (midnight to 6 AM).
* **Conversion rates increase steadily throughout the day**, peaking between **2 PM and 4 PM (14:00–16:00)**.
* After 4 PM, the rates **start to gradually decline** but remain relatively high until 10 PM.

**Top Hours by Conversion Rate**

| **Hour** | **Conversion Rate (%)** |
| --- | --- |
| **16 (4 PM)** | **3.08%** *(Highest)* |
| 15 (3 PM) | 2.97% |
| 14 (2 PM) | 2.81% |
| 13 (1 PM) | 2.47% |
| 12 (12 PM) | 2.38% |

These hours represent the **afternoon peak** in user conversions, suggesting users are **most responsive to ads in the afternoon**.

**Lowest Performing Hours**

| **Hour** | **Conversion Rate (%)** |
| --- | --- |
| **2 (2 AM)** | **0.73%** *(Lowest)* |
| 3 (3 AM) | 1.05% |
| 1 (1 AM) | 1.29% |
| 4 (4 AM) | 1.52% |
| 0 (12 AM) | 1.84% |

Ads shown during **late night and early morning hours** perform **significantly worse**, possibly due to **lower user engagement or intent** at those times.

**Interpretation and Implications**

1. **Afternoon is Prime Time**:
   1. Users shown the most ads between **2 PM and 4 PM** convert at the **highest rates**.
   2. Scheduling **ad-heavy campaigns in the afternoon** could **maximize conversions**.
2. **Low Yield During Night Hours**:
   1. Conversion rates are notably **low from midnight to 6 AM**, indicating **low effectiveness** of ad exposure during these hours.
3. **Data Supports Ad Scheduling**:
   1. These insights can be directly used to **optimize ad scheduling strategies** — prioritize higher budgets and impressions during peak hours and reduce spending during low-converting periods.

A graph of sales

AI-generated content may be incorrect.

**Top Performing Days (Highest Conversion Rates)**

| **Day** | **Conversion Rate (%)** |
| --- | --- |
| **Monday** | **3.28%** *(Highest)* |
| Tuesday | 2.98% |
| Wednesday | 2.49% |
| Sunday | 2.45% |

**Monday** stands out with the **highest conversion rate**, suggesting users may be more responsive at the start of the week possibly due to a **fresh mindset**, **work-week planning**, or **intention-driven behavior**.

**Lowest Performing Days**

| **Day** | **Conversion Rate (%)** |
| --- | --- |
| **Saturday** | **2.11%** *(Lowest)* |
| Thursday | 2.16% |
| Friday | 2.22% |

Weekends and end-of-week days (like **Saturday, Thursday, and Friday**) have **lower conversion rates**, possibly due to **reduced online activity for transactional purposes**, or more **leisure browsing** behavior.

**Interpretation & Insights**

1. **Conversion Peaks Early in the Week**:
   1. **Monday and Tuesday** yield the highest conversion rates. This could be a strategic window for **launching campaigns**, sending emails, or increasing ad visibility.
2. **Decline Toward the Weekend**:
   1. Conversion rates tend to **drop on Thursday through Saturday**, suggesting users are **less likely to act** on ads as the weekends.
3. **Optimize Ad Scheduling by Day**:
   1. **Reallocate budget** to Monday–Wednesday for higher ROI.
   2. Consider **awareness-only content** or lower ad spend on Thursday–Saturday.

A graph of blue bars with white text

AI-generated content may be incorrect.

**Summary Table total ads**

| **Converted** | **Count** | **Mean** | **Std Dev** | **Min** | **25%** | **50% (Median)** | **75%** | **Max** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **False** | 573,258 | 23.29 | 40.86 | 1 | 4 | 13 | 26 | 2065 |
| **True** | 14,843 | 83.89 | 87.46 | 1 | 35 | 64 | 103 |  |

**Key Interpretations**

1. **Average Total Ads Is Much Higher for Converted Users**
   1. Users who **converted** were shown an **average of ~84 ads**, compared to **~23 ads** for those who did **not convert**.
   2. This suggests a **positive correlation between ad exposure and conversion** — people who saw more ads were **more likely to convert**.
2. **Wider Spread Among Converted Users**
   1. The **standard deviation** is significantly higher for converted users (**87.46**) than non-converted (**40.86**), indicating **greater variability** in the number of ads they saw.
   2. Some converted users may have been heavily targeted before acting.

A diagram of a comparison between two blue squares

AI-generated content may be incorrect.

1. **Median and Quartiles Show Clear Separation**

| **Percentile** | **Non-Converted** | **Converted** |
| --- | --- | --- |
| 25% | 4 ads | 35 ads |
| 50% (Median) | 13 ads | 64 ads |
| 75% | 26 ads | 103 ads |

* Even at the **25th percentile**, converted users saw **35 ads**, which is already more than the **75th percentile (26 ads)** for non-converted users.
* This clearly indicates that **higher ad exposure is associated with higher conversion likelihood**.

**4. Outliers Present**

* 1. Both groups have **extreme maximum values** (over 1000 ads), suggesting some users were shown a **very high number of ads**, possibly due to high targeting frequency.

**Summary Insight**

* **Ad exposure strongly influences conversion**: More ads correlate with a higher probability of conversion.
* This supports the idea that **repetition and sustained ad engagement can drive user action**.
* However, the large standard deviation and high maximums also suggest a **diminishing return beyond a point** or the need to **balance exposure to avoid fatigue**.

**Interpretation of Each Chi-Square Test**

**Test Group vs. Converted**

1. **Chi-Square Statistic**: 54.01
2. **Degrees of Freedom**: 1
3. **P-value**: 0.0000
4. **Conclusion**:
   1. Since the **p-value < 0.05**, you **reject the null hypothesis**.
   2. This means there is a **statistically significant association between the test group and conversion status**.
   3. In simpler terms: **the type of group (e.g., ad vs. PSA)** has a real effect on whether users convert.

**Implication**: The choice of campaign (ad vs. PSA) **impacts conversion rates** significantly.

**2. Most Ads Hour vs. Converted**

1. **Chi-Square Statistic**: 430.77
2. **Degrees of Freedom**: 23
3. **P-value**: 0.0000
4. **Conclusion**:
   1. Again, **p < 0.05**, so you **reject the null hypothesis**.
   2. There is a **significant relationship between the hour when most ads were shown and whether the user converted**.

**Implication**: The **time of day when users are most exposed to ads** plays a meaningful role in whether they convert — ad performance is **time-sensitive**.

1. **3. Most Ads Day vs. Converted**
2. **Chi-Square Statistic**: 410.05
3. **Degrees of Freedom**: 6
4. **P-value**: 0.0000
5. **Conclusion**:
   1. With **p < 0.05**, you again **reject the null hypothesis**.
   2. There is a **significant association between the day of highest ad exposure and conversion status**.

**Implication**: **Certain days of the week are more effective** for ad exposure in driving conversions (as seen in your earlier analysis where Monday and Tuesday had the highest conversion rates).

**Overall Summary**

| **Variable** | **Significant Association with Conversion?** | **Practical Meaning** |
| --- | --- | --- |
| **Test Group** | Yes | The type of content (ad vs. PSA) affects conversion success |
| **Most Ads Hour** | Yes | The time-of-day matters — conversions vary by ad exposure hour |
| **Most Ads Day** | Yes | Day of the week also influences conversion likelihood |

These results confirm that **ad timing (hour/day) and content type meaningfully affect conversion behavior**, and this insight can be **used to optimize ad strategies** for better performance.

| **Group** | **Shapiro-Wilk Statistic** | **P-value** | **Normality Conclusion** |
| --- | --- | --- | --- |
| Converted = True | 0.6578 | 0.0000 | Not Normally Distributed |
| Converted = False | 0.4747 | 0.0000 | Not Normally Distributed |

**🔍 Interpretation**

**1. What the Shapiro-Wilk Test Does:**

* This test checks whether your data **follows a normal (bell-shaped) distribution**.
* **Null Hypothesis (H₀)**: Data is normally distributed.
* If the **p-value < 0.05**, we **reject H₀** → data is **not normally distributed**.

**2. Your Results:**

* For **both groups (converted = True and converted = False)**, the **p-value is 0.0000**, which is **far below 0.05**.
* So, for both groups, we **reject the null hypothesis**.

**3. Conclusion:**

* The distribution of **total ads is not normal** in either the converted or non-converted group.
* This is also supported by the **very low Shapiro-Wilk statistics** (much less than 1), indicating **significant deviation** from normality.

A comparison of a graph

AI-generated content may be incorrect.

**Levene’s Test for Equality of Variances** and related statistics:

**Levene's Test Summary**

| **Metric** | **Value** |
| --- | --- |
| **Levene Statistic** | 9121.1970 |
| **P-value** | 0.0000 |
| **Alpha (Significance Level)** | 0.05 |
| **Conclusion** | Reject the null hypothesis |

**🔍 Interpretation of Levene’s Test**

**1. What Levene’s Test Does:**

* It tests whether **two or more groups have equal variances**.
* **Null Hypothesis (H₀)**: The groups have **equal variances**.
* If the **p-value < 0.05**, you **reject H₀**, meaning the variances are **significantly different**.

**2. Your Results:**

* The **p-value is 0.0000**, which is far below the significance level of 0.05.
* Therefore, we **reject the null hypothesis**.
* This means the **variance in total ads is significantly different** between **converted** and **non-converted** users.

**Standard Deviations Show the Magnitude of Difference**

| **Group** | **Standard Deviation of total ads** |
| --- | --- |
| **Converted** | **87.46** |
| **Not Converted** | **40.86** |

* Converted users have a **much higher variability** in the number of ads they saw.
* This indicates that while **some converted users may have been exposed to many ads**, others may have converted after seeing fewer ads — but in general, their exposure varies widely.
* In contrast, **non-converted users had more consistent and lower ad exposure**.

**Implications**

1. **Unequal variances must be accounted for** in any statistical testing:
   * You **should not use a standard t-test** assuming equal variances.
   * Use a **Welch’s t-test** (or non-parametric tests like the **Mann-Whitney U test**, especially since the data is also **not normally distributed**).
2. **Marketing Insight**:
   * The **converted group shows more diverse engagement levels** with ads.
   * This could suggest that **multiple conversion patterns exist**, e.g.:
     + Some users need very few ads to convert (low resistance),
     + Others may require **frequent exposure** (high resistance or high involvement products).