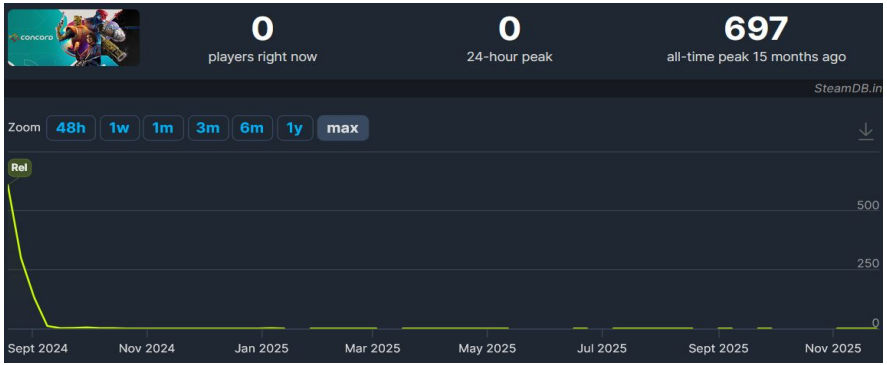
A/B Testing for Mobile Game Retention

BY:
Muhammad Umar Niazi - MSDS25018
Afzaal Kareem - MSDS25062
Numan Hussain - MSDS25067

Player Retention In Gaming - Why Is It Important?

- Retention is the main factor for a game’s health and state.
- No one playing means no revenue generation.
- Player feedback is the source to determining gameplay changes and creative decisions.
- Constant changes and updates required to keep the players in the game.



Why Hypothesis Testing

- To test whether design changes actually improve retention
- To avoid false positives caused by randomness and transforms uncertainty into measurable confidence
- It enables A/B testing at a large scale even with millions of players like most mainstream games do.

About The Dataset

- The dataset includes A/B test results of Cookie Cats to examine what happens when the first gate in the game was moved from level 30 to level 40.
- When a player installed the game, he or she was randomly assigned to either gate_30 or gate_40.
- Ideal for hypothesis testing and suggesting game developers the game design changes.

userid	version	sum_gamerounds	retention_1	retention_7
116	gate_30	3	False	False
337	gate_30	38	True	False
377	gate_40	165	True	False
483	gate_40	1	False	False
488	gate_40	179	True	True
540	gate_40	187	True	True
1066	gate_30	0	False	False
1444	gate_40	2	False	False
1574	gate_40	108	True	True
1587	gate_40	153	True	False

```
[90189 rows x 5 columns]>
```

Hypothesis (Null & Alternative)

Day-1 Retention

- **Null Hypothesis (H_{01}):**
  Day 1 retention remains the same when moving from gate 30 to gate 40 => (retention_1(g30) = retention_1(g40))
- **Alternative Hypothesis (H_{11}):**
  Day 1 retention changes when moving from gate 30 to gate 40 (retention_1(g30) =/= retention_1(g40)).

Day-7 Retention

- **Null Hypothesis (H_{07}):**
  Day 7 retention remains the same when moving from gate 30 to gate 40 => (retention_7(g30) = retention_7(g40))
- **Alternative Hypothesis (H_{17}):**
  Day 7 retention does not remain the same when moving from gate 30 to gate 40 => (retention_7(g30) =/= retention_7(g40))

Libraries Used

`OS:` Used for interacting with the operating system (e.g., file paths, directory management).

`Pandas:` For data manipulation and analysis: reading CSVs, cleaning data, grouping, creating summary tables, pivoting, and displaying results.

`Numpy:` Provides numerical operations and arrays; used for calculations such as sums, differences, proportions, and standard errors

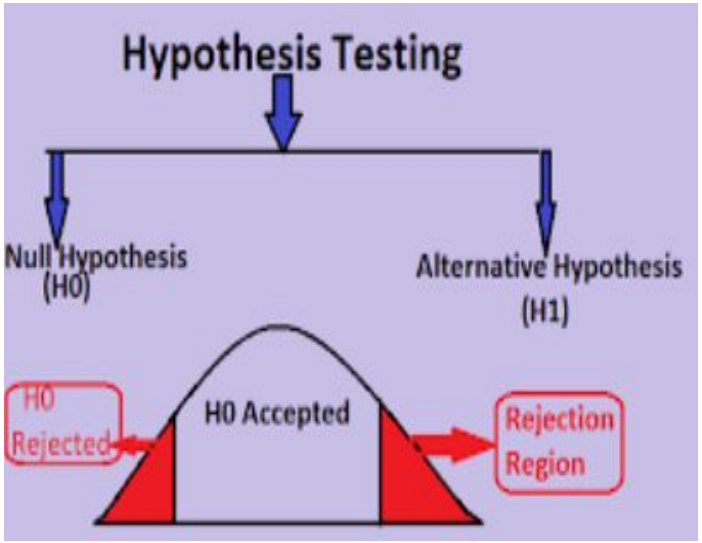
`Scipy.stats:` Statistical functions, e.g., `stats.norm.ppf` for z-critical values and confidence intervals.

`Statsmodels - proportions_ztest:` Performs the two-proportion z-test to compare retention rates between Gate_30 and Gate_40.

`Matplotlib.pyplot:` Visualization library used for creating graphs and charts

Methods Used

- Data Collection and Preprocessing
- Grouping and Summary Statistics
- Null And Alternative Hypothesis
- Z-Test
- Computing CI
- Visualization
- Interpretation and Reporting

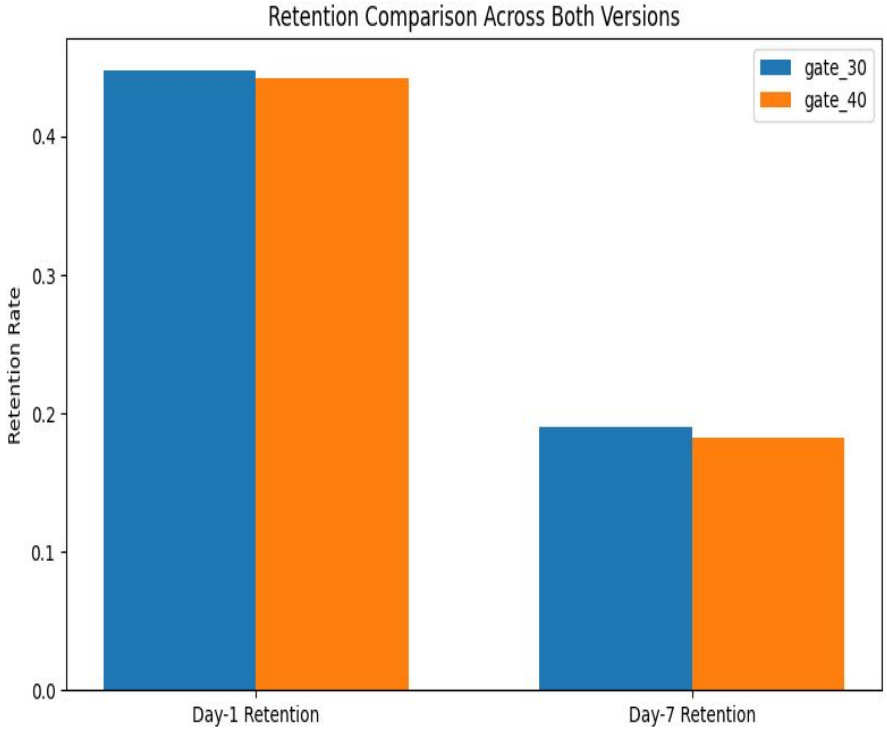


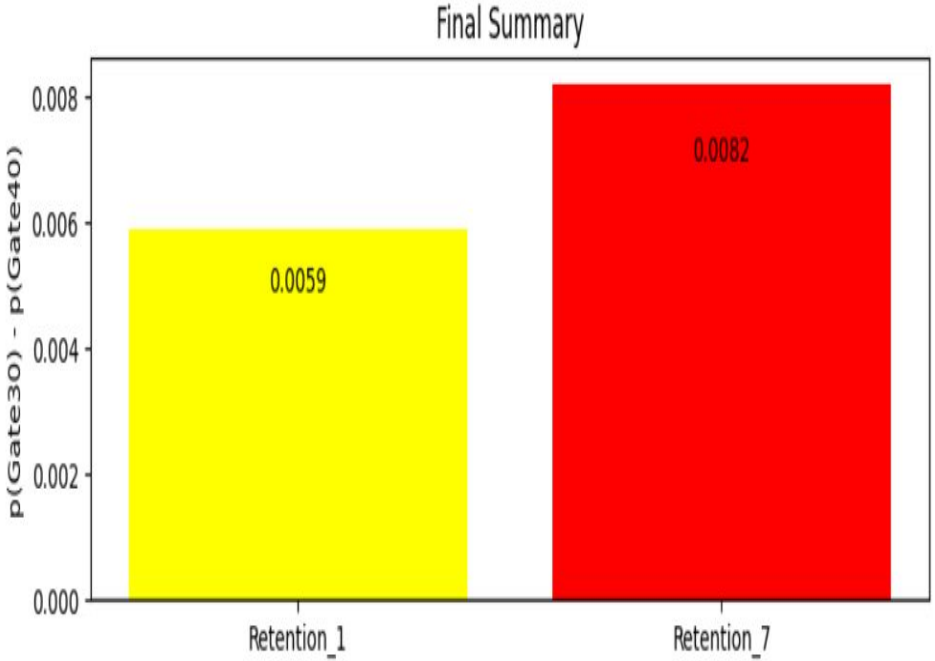
Important Values

- **Z-Statistic:**
  - A large positive z-value → Gate_30 performed better than Gate_40.
  - A z close to zero → both gates performed similarly, difference likely due to randomness.
  - A negative z-value (not the case here) → Gate_40 would have been better.
- **Significance level (α):**
  - threshold we choose before testing to decide how much risk of being wrong we are willing to accept.
  - Chosen value 0.05 or 5%
- **P-Value:**
  - The p-value measures the probability that an observed difference between Gate_30 and Gate_40 retention occurred by chance.

Metric	Gate_30	Gate_40	Difference (p30-p40)	P-Value
Retention_1	0.448188	0.442283	0.005905	0.074410
Retention_7	0.190201	0.182000	0.008201	0.001554

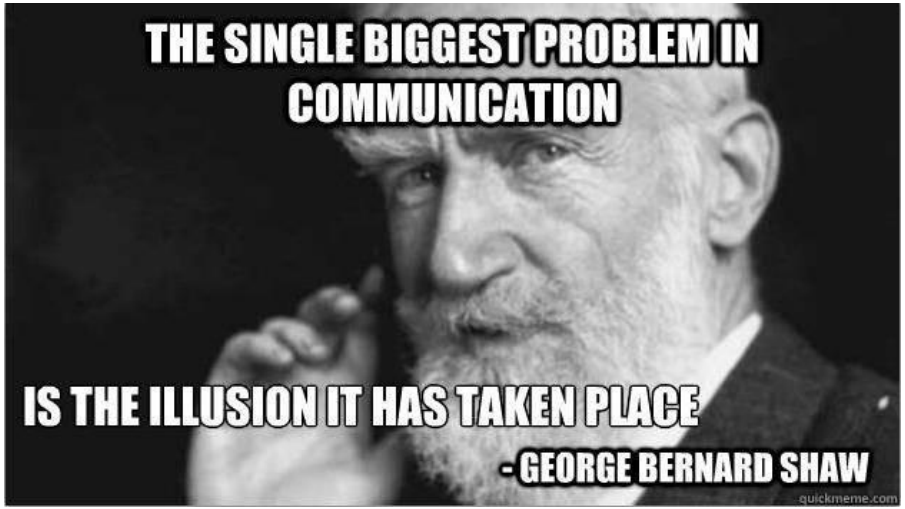
Output Graphs





Suggestion To Game Devs

- Keep the progression gate at level 30 rather than moving it to level 40, as Day-7 retention is slightly higher with Gate-30.
- to experiment with a later gate adding incentives such as rewards, boosters, or mini-goals can maintain player engagement.
- Monitor retention carefully for different player segments and over longer periods to ensure the change doesn't negatively impact overall engagement.
- Ensure A/B testing is performed for the newer version too.



Conclusion

- Hypothesis testing allowed us to avoid intuition based decisions and prevented incorrect assumptions that modifying the game experience would automatically lead to better engagement.
- It supported data-driven decision-making, ensuring that changes implemented by the development team were based on solid analytical proof.
- The use of hypothesis testing made it possible to communicate results clearly and confidently to the game developers.
- Output values and differences provided objective proof that could be trusted in decision-making discussions.

THANK YOU

Any Questions?