

Investigating Mana Value Popularity

Robert D. Clark

Western Governors University

Table of Contents

A. Project Highlights	2
B. Project Execution	3
C. Data Collection Process	5
C.1 Advantages and Limitations of Data Set.....	5
D. Data Extraction and Preparation	6
E. Data Analysis Process	6
E.1 Data Analysis Methods.....	6
E.2 Advantages and Limitations of Tools and Techniques	6
E.3 Application of Analytical Methods	7
F. Data Analysis Results	7
F.1 Statistical Significance.....	7
F.2 Practical Significance	8
F.3 Overall Success.....	8
G. Conclusion	8
G.1 Summary of Conclusions.....	8
G.2 Effective Storytelling	9
G.3 Recommended Courses of Action.....	9
H. Panopto Presentation.....	9
References.....	10
Appendix A	11
My Code	11
Appendix B.....	12
Date Source.....	12
My data source comes from Scryfall and is the Oracle Cards JSON. Download it from here: https://scryfall.com/docs/api/bulk-data	12

Project Plan

A. Project Highlights

A.1 Research Question

Which MtG card mana value is the most popular among Modern players?

A.2 Project Scope

The project scope included downloading the data to my local computer, loading it into python, cleaning the data as the analysis needs, separating the data by card rarity, calculating the price averages for each rarity, then calculating price averages for each mana value within each rarity, displaying the resulting averages with bar charts, calculating correlation metrics and Z-scores, interpreting the metrics and bar charts, and finally making a recommendation as to which mana value will be the most popular. This matches my scope defined in my Task 2 submission. I stayed within the scope throughout the whole project.

A.3 Solution Overview

I used Python, including a few libraries, to perform the project. I used the Agile methodology to plan the project, the Z-test methodology to get the results to the research question, and a correlation methodology to verify the results were accurate.

Project Execution

B. Project Execution

B.1 Project Plan

The overall goals and objectives I planned to achieve were:

- Goal 1: Find the most popular card mana value.
 - Objective 1.1: Find the average card price for each mana value.
 - Deliverable 1.1.1: A bar chart for each card rarity each showing the average card price for each mana value. Also with a way to compare the averages to a standard (threshold line.)

I achieved every single one of these goals exactly as they were planned in Task 2. The next part explains the plan in more detail.

B.2 Planning Methodology

I used the Agile project planning methodology. Which meant taking an iterative approach to the project. Any aspect of the project may have been subject to change, except the research

question. Any phase could have been returned to after it's completed if a future phase requires it. The phases I planned to and followed are data collection, data cleaning, analysis, visualization and metric evaluation, and finally results and recommendations.

Data Collection: In this phase, I downloaded and loaded the data I needed into my environment.

Data Cleaning: The data cleaning phase is when I determined and extracted a sample from the data based on common parameters, as well as removed columns that were unnecessary. I also separated the data into multiple dataframes, each filtered on a different card rarity.

Analysis: In the analysis phase, I grouped each dataframe by mana value and calculated the average card price for each mana value. I also calculated the overall average card price per card rarity to use as a threshold. Also in this phase, I calculated the Z-score for every price value in each card rarity dataframe, grouped them by mana value and averaged them; similar to the way I did price.

Visualization and Metric Evaluation: In this phase, I visualized the average data with thresholded bar charts. I made four different bar charts, one for each card rarity, with each bar related to a specific mana value and the height being the average price for that mana value. Each chart also had a threshold line showing the total average for all cards of that rarity. That's to make it easier to read how much more or less popular a specific mana value makes a card. Also included in this phase was metric evaluation, in which I calculated the correlation coefficient and p-value for mana value on card price for each card rarity. This put a number to how much and in what way does mana value affect card price. At the end, I averaged the four average Z-scores per mana value (separated by card rarity) together to get a single Z-score for each mana value. This shows how much a specific mana value increases or decreases the card price.

Results and Recommendations: This final phase is where I used the results from the charts and metrics to understand which mana value is most popular and made my recommendation to Wizards of the Coast.

There was no difference in the planned phases from Task 2 than what was executed.

B.3 Project Timeline and Milestones

While the milestones did not change, the duration for one of them was longer than planned.

Milestone or Deliverable	Duration (Hours)	Start Date	End Date
Find the average card price per mana value (Goal 1)	4h	10/17/2024	10/17/2024
Make bar charts for each card rarity showing average card price per mana value. (Deliverable 1)	4h	10/17/2024	10/17/2024

I was able to keep the project within the full day I planned, but making the bar charts with the threshold lines (row 2) took 2 hours more than planned in Task 2.

Methodology

C. Data Collection Process

The data collection process was as simple as downloading a JSON from <https://scryfall.com/docs/api/bulk-data> and loading it into my Python environment with pandas. Which is exactly what I planned in Task 2. I had no obstacles during this process. It is a very simple process, after all. No data governance issues arose either because the data is publicly available. There were no precautions or regulations I needed to take or work around. Smooth sailing.

C.1 Advantages and Limitations of Data Set

An advantage of the dataset is how specific the file has been designed. The specific file I downloaded is designed to be only one row for each card ever printed. For example, I didn't have to deal with any duplicate rows or multiple rows that refer to the same card but are slightly different; nor have to arbitrarily decide which of these rows to keep. It made the cleaning process non-existent.

A disadvantage of the dataset is it's just a snapshot that can be downloaded. Scryfall updates the card prices once a day, but you have to download a new file to get those updated

prices. So, in the case of my project, after about a month, the prices may have changed enough for the results to be misleading or inaccurate. The dataset would have to be manually updated each day.

D. Data Extraction and Preparation

I saved the JSON to a pandas dataframe, then I selected from that only cards that were legal in the Modern format, had an english printing, had price data, were of a card format that didn't skew the mana value, had a rarity value, and had a mana value between 0 and 8. I kept Modern cards because that's what I specified in the business question, and the cards I removed were because there were so few of them that the averages would be skewed and mess with the analysis. I removed less than 5% of the cards and I consider them outliers. Then I removed all the columns I didn't need for the analysis. The columns I kept were card id, name, mana value, rarity, and price. I didn't end up using the card id or name.

E. Data Analysis Process

E.1 Data Analysis Methods

The analytical method I used is a comparative aggregation. I averaged the card prices per mana value and compared to the overall card average. Sometimes called a Z-score. This analytical method is appropriate to prove or disprove my hypothesis because it shows how and by how much each mana value increases or decreases the card price. This was used to determine the actual most popular mana value to prove or disprove my hypothesis.

E.2 Advantages and Limitations of Tools and Techniques

An advantage of using Python for my project is the versatility of it. With the large library of possible modules, almost any purpose for Python has been made easier with one. I used pandas to seamlessly load in, manipulate, and use the data throughout the whole project. However, a disadvantage of python is also the versatility of it. Because it can be used for almost anything, there's nothing in particular it's really good at. It's not very efficient at what it does and

there may be some parts of code that could be one-liners in other languages, but not with Python. Python modules can also be fairly difficult to discover.

E.3 Application of Analytical Methods

I used a comparative aggregation by finding the average card price per each mana value and displaying them with a bar chart with a threshold line set at the overall card average. I also measured the exact number of standard deviations above or below the overall average each mana value was, which is called a Z-score. I used a stats python module and the zscore method to find it. I verified the Z-score was an accurate depiction of that mana value's popularity by finding the correlation coefficient for mana value and card price. If mana value and card price weren't correlated at all, the Z-scores produced were just a coincidence. I go into more detail in my submitted Panopto video.

Results

F. Data Analysis Results

F.1 Statistical Significance

The null hypothesis for the correlation test is that there is no correlation between mana value and card price. The correlation test returns a correlation coefficient and a p-value. The correlation coefficient measures the linear correlation between two variables. It can measure between -1 and +1, where positive means a direct correlation between the two variables, and a negative coefficient means inversely correlated variables. The p-value is essentially the probability the correlation coefficient is a coincidence. For determining statistical significance, if the alpha value is ≥ 0.05 and the p-value is ≤ 0.1 , the correlation is statistically significant and the null hypothesis is false. From the results returned from the correlation test between mana value and card price, I can conclude that there is a correlation between mana value and price and the null hypothesis is false.

The null hypothesis for the Z-test is that no mana value is more popular than any other mana value. A Z-test returns a Z-score, which is classified as the number of standard deviations

away from the total average a specific value is. If at least one Z-score greater than or equal to -0.05 , but less than or equal to $+0.05$, then the null hypothesis is false. A Z-score outside the same range is statistically significant. From the results returned from the Z-tests, I can conclude that there is a mana value that is more popular than the others and the null hypothesis is false.

F.2 Practical Significance

The practical significance of finding the most popular mana value is that Wizards of the Coast can use the information to design cards that are more likely to be popular and make more sales. The analysis can also be used to design cards that synergize with the most popular mana value.

Because one mana value has been proven to be more popular than the others, Wizards can use that information to design more cards with that mana value and subsequently more popular cards.

F.3 Overall Success

The criteria for success I set in Task 2 are: that card price averages are properly calculated, as well as the grouped by mana value amounts. The bar charts display the data properly and legibly. The correlation coefficient and p-value are within a reasonable range for each value (e.g. -1 to $+1$). And the Z-scores are calculated in a way that accurately shows how popular or unpopular each mana value is, taking into account anything else that may affect popularity.

My project reached all of these criteria for success. I deem my project a success.

G. Conclusion

G.1 Summary of Conclusions

The conclusion from my project is:
The most popular mana value among Modern players is 0.

Other information I learned from my project is that mana value is very related to card price. Because mana value is not set or changed based on card price, it is mana value that directly affects card price, not the other way around.

G.2 Effective Storytelling

I used several bar charts to display the relationship between each mana value and card price. It was an effective way to show that information because it made it simple to see how much higher an average was from the total average by using actual bar height to show it. It allows you to visually compare the two values instead of in your head.

I used tables to show the Z-scores and correlation coefficients because they are simply values that convey the information. There's no comparison that needs to be made, you just need to see the number, which is harder to read if in a bar chart.

G.3 Recommended Courses of Action

The two recommendations I make are to Wizards of the Coast. I recommend designing more cards with a mana value of 0 because they are more likely to be popular. I also recommend using your own knowledge on what makes cards synergize with other cards and design cards that synergize with existing 0 mana value cards. Cards that synergize with popular cards are more likely to be popular, too.

H. Panopto Presentation

I've provided a link to my Panopto video [here](#) and in my submission. It includes a summary of my project and goes through the code I wrote and how it works.

References

Your Playmat. (2023, February 22). *Are Magic Cards Worth Anything? MTG Cards Value*.

Retrieved from Your Playmat: <https://www.yourplaymat.com/blogs/tcg-players-blog/are-magiccards-worth-anything>

Mark Rosewater. (2003, April 21). *DESIGN 101*. Retrieved from Wizards: [https://](https://magic.wizards.com/en/news/making-magic/design-101-2003-04-21-0)

magic.wizards.com/en/news/making-magic/design-101-2003-04-21-0

Alex Wiltshire. (2019, November 6). *How Magic: The Gathering card sets are designed*.

Retrieved from RockPaperShotgun: <https://www.rockpapershotgun.com/how-magic-the-gathering-card-sets-are-designed>

Appendix A

My Code

I've submitted my code in an ipynb file with my submission.

Appendix B

Data Source

My data source comes from Scryfall and is the Oracle Cards JSON. Download it from here:

<https://scryfall.com/docs/api/bulk-data>