The Beatles*

What are their top 5 popular songs?

Hyunje Park, Charlie Zhang

November 21, 2024

This paper analyzes which songs from the 1960s rock-band The Beatles had the highest rating in the United States, using the popularity rating extracted from the Spotify API. It was discovered that the top 5 songs in popularity were Here Comes the Sun, Let It Be, Blackbird, Yesterday, In My Life.

Table of contents

1 Introduction

The Beatles, is one of the most popular English rock band in the world, dominating the music industry in the 1960s. This paper analyzes which songs from the band `The Beatles` had

*Code and data are available at: https://github.com/davidpxrk/beatles-top-5-songs

the highest popularity rating, using the Spotify API to collect information on their song's popularity. The method of this web scrapping is further mentioned in Data (Section 2).

The paper is structured as follows: Data (Section 2) introduces the dataset; Results (Section 3) presents key findings; Discussion (Section 4) explores observations and limitations; and Conclusion (Section 5) summarizes the analysis.

2 Data

This report analyzes the Canadian Grocery Price dataset, collected by Jacob Filipp and hosted by Project Hammer (Filipp 2024). Of the three tables in the dataset, only the product table was used, focusing on the `id` (product identifier) and `vendor` (name of the vendor) variables.

First, the original dataset was downloaded from `Project Hammer` (Filipp 2024). Using `SQL` (ISO/IEC Joint Technical Committee 1987), the `id` and `vendor` variables were selected, and a table was created that containing the number of unique product IDs (`id`) for each vendor. This table was then exported as a `.csv` file. The exported `.csv` file was then analyzed in `R Programming Language` (R Core Team 2023) to visualize the relationship between vendors and their unique product counts, using the `ggplot2` package (Wickham 2016).

3 Results

?@fig-difference shows the unique product counts for each Canadian grocery vendors. Overall, Loblaws, Metro and Walmart reigned at the top with over 20,000+ products, while international-focused chains such as Galleria or T&T had the least amount of unique products.

	name	popularity
1	Here Comes The Sun - Remastered 2009	82
2	Let It Be - Remastered 2009	77
3	Blackbird - Remastered 2009	76
4	Yesterday - Remastered 2009	76
5	In My Life - Remastered 2009	76

4 Discussion

?@fig-difference showed that big companies such as Metro and Walmart far passed Galleria and T&T in product variety, showing a competitive challenge for smaller vendors.

4.1 Correlation vs Causation

Larger vendors, like Walmart (the world's largest company by revenue) offer more unique items, while smaller vendors such as Galleria and T&T have a smaller revenue, which could explain the disparity in product variety.

4.2 Missing Data

The `upc` (universal product code) and `sku` (inventory identifier) variables were excluded from the analysis as they were missing for all products.

4.3 Source of bias

This raises the question of why vendors like Galleria and T&T lag in product variety. As Asian-fusion grocery chains, they rely on costly imports, which could limit their inventory. However, without data on product types, commenting on the efficiency and competition on these vendors is impossible.

5 Conclusion

In summary, this paper analyzed Canadian grocery vendors' product variety, finding Metro and Walmart lead with over 20,000 unique items, while chains like Galleria and T&T, focused on Asian products, have fewer than 10,000.

References

Filipp, Jacob. 2024. *Project Hammer*. https://jacobfilipp.com/hammer/.

ISO/IEC Joint Technical Committee. 1987. *SQL: Structured Query Language*. https://www.iso.org/standard/63555.html.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.