

# Vinyl ★ Profitability Explorer

A tool built to optimize  
record store profit and  
efficiency.




# Project Outline




The tool built in this project will create a value matrix, where record sellers can evaluate Discogs data and interpret a live visual metric of demand. It's easy to look up a record's price, but it is very difficult to make an evaluation relative to demand. The web application built in this project enables the user to locate undervalued records and stock records that are likely to net profit.

## *Project Workflow*

- 
- Data Collection from Discogs
  - Data Cleaning in Jupyter Notebook
  - Feature Engineering
  - High & Low Value Tier System
  - Offer Price Valuation Tool
  - R Shiny Application

The goal of this project is to configure an application capable of dictating vinyl profitability.





# Data Source & Web-Scrape


Discogs is the largest online database for physical media purchases. They have cataloged over 750 million records.


# Discogs Marketplace



### Search results


Showing 2 of 2




[Alice In Chains - MTV Unplugged](#) 

Columbia, Music On Vinyl, Sony Music Entertainment, Sony Music Entertainment, Music On Vinyl B.V., Music On Vinyl B.V., Brooklyn Academy Of Music, Sony/CBS, Haarlem, Record Industry, Precision Lacquer – MOVLP138  
Vinyl, LP, Album, Reissue, Repress  
2010, Europe

[List item](#)





[Alice In Chains - MTV Unplugged](#) 

Music On Vinyl, Sony/CBS, Haarlem, Record Industry, Precision Lacquer, Columbia – MOVLP138  
Vinyl, LP, Album, Limited Edition, Numbered, Reissue, Repress  
2010, Europe



[List item](#)

### List item




[Alice In Chains](#) – [MTV Unplugged](#) 

Columbia, Music On Vinyl – MOVLP138  
2xLP, Album, RE, RP, 180  
2010, Europe

Rated: 4.78  
 12545 have  
 4591 want

Last sold: May 18, 2023 for £136.00 [View complete sales history](#)

Highest	Lowest	Average	Median
£240.00 (NM or M-)	£22.11 (M)	£129.65	£145.99

[29 for sale from £108.00](#) 

Media condition

Sleeve condition

optional [How do I grade my items?](#)

Item condition comments optional

0/500

★ **What?**

Online database  
for physical media

★ **How?**

Community  
marketplace with  
prices and demand

# Web-scrape Data Collection



**Marketplace** All Items Items I Want Purchases Cart Buyer Settings

Search Marketplace

**You Selected:** **Shop The Beatles - Please Please Me (All Versions)**



Release: The Beatles - Please Please Me (All Versions)

**Ships From**  
United Kingdom 719  
United States 708  
Germany 317  
Italy 295  
Japan 292  
Show more...

**Currency**  
EUR (€) 1,453  
USD (\$) 1,082  
GBP (£) 734  
JPY (¥) 138

1 – 25 of 3,629 < Prev Next >

Sort Listed Newest Show 25

Condition, Artist, Title, Label	Seller	Price	
 Rated: 4.56 5178 have 711 want	<b>The Beatles - Please Please Me (LP, Album, RE, RM, 180)</b> Label: Parlophone, Parlophone #: 0094638241614, PCS 3042 Media Condition: Mint (M) Sleeve Condition: Mint (M) Sealed and with hype sticker! <a href="#">View Release Page</a>	<b>Good Things Gmunden</b> ★★★★★ 100.0%, 1,647 ratings Ships From: Austria	<b>€34.90</b> <del>€43.88</del> shipping €47.90 total <a href="#">Add to Cart</a> <a href="#">Details</a>
 3 have 89 want	<b>The Beatles - Please Please Me (LP, Album, RE, Isr)</b> Label: Portrait Cat#: PCS 3042 Media Condition: Very Good Plus (VG+) Sleeve Condition: Very Good Plus (VG+) Old name sticker on cover, please ask for pictures <a href="#">View Release Page</a>	<b>israeli-records</b> ★★★★★ 99.6%, 242 ratings Ships From: Israel	<b>\$125.00</b> +\$14.00 shipping about €127.88 total <a href="#">Add to Cart</a> <a href="#">Make Offer / Details</a>



# Data Feature Engineering

Which records are popular and how much are they worth given marketplace demand and rating?

# Data Variables



## **RATING**

community vote  
assigned to the  
vinyl record



## **WANT**

Number of  
community  
members that  
desire record



## **PRICE**

representation of  
marketplace  
value



## **HAVE**

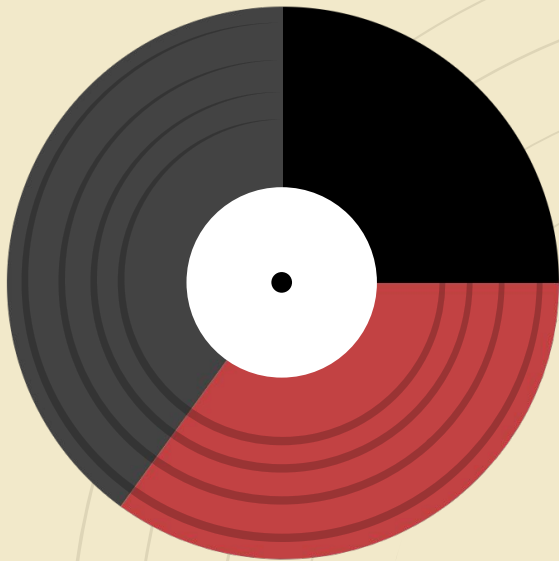
Number of  
community  
members that  
own record

# Data Cleaning

album	price	rating	have	want	tier	popularity
Nilsson* - Knillsson	15.0	4.02	1340	215	Mid-Tier (Popular)	0.1382636655948550
Rockin' Horse - Rockin' Horse	10.0	4.26	272	212	Mid-Tier (Popular)	0.4380165289256200
The Beatles - Abbey Road	30.0	4.64	484	1053	Top Rated	0.6851008458035130
Rod Stewart - Foot Loose & Fancy Free	3.99	3.7	7628	110	Mid-Tier (Popular)	0.01421555957611790
Bob Weir - Heaven Help The Fool	6.0	3.9	1770	92	Mid-Tier (Popular)	0.04940923737916220
Keith Urban - Ripcord	20.0	4.23	816	127	Mid-Tier (Popular)	0.13467656415694600
Aerosmith - Toys In The Attic	8.0	4.02	22835	2603	Mid-Tier (Popular)	0.1023272269832530
The J. Geils Band - Bloodshot	5.0	3.5	143	38	Niche/Underground	0.20994475138121500
Phil Ochs - There But For Fortune	39.5	5.0	15	24	Niche/Underground	0.6153846153846150
Mott The Hoople - The Hoople	10.0	3.89	2561	144	Mid-Tier (Popular)	0.05323475046210720
Electric Light Orchestra - A New World Record	7.0	4.11	12772	694	Mid-Tier (Popular)	0.05153720481211940
Grand Funk Railroad - Caught In The Act	8.0	4.0	194	40	Niche/Underground	0.17094017094017100
The Rolling Stones - Sweet Summer Sun - Hyde Park Live	176.98	4.43	57	63	Mid-Tier (Popular)	0.525
Pete Townshend - All The Best Cowboys Have Chinese Eyes	15.0	3.84	930	27	Niche/Underground	0.02821316614420060
Led Zeppelin - In Through The Out Door	15.0	4.09	1887	422	Mid-Tier (Popular)	0.18276310090948500



# Feature Engineering



## Popularity

Want / Want + Have count



## Rating

Community given vote



## Tier System

Model for profitability

In future versions of the app, rating will be replaced with a unique score based on inventory and sale velocity.

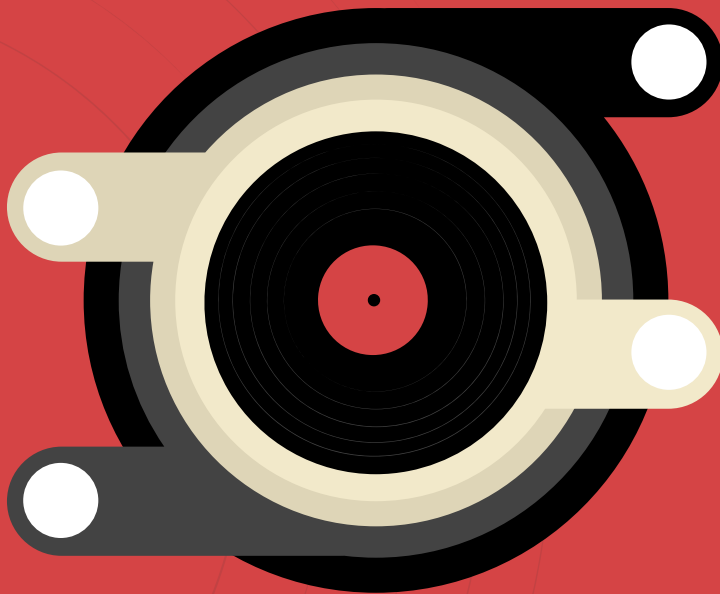
# Tier System

## Top Rated

High Ratings  
High Demand

## Mid-Tier

Medium Ratings  
High Demand



## Niche / Underground

Medium Ratings  
Low Demand


## Low Priority

Low Ratings  
Low Demand



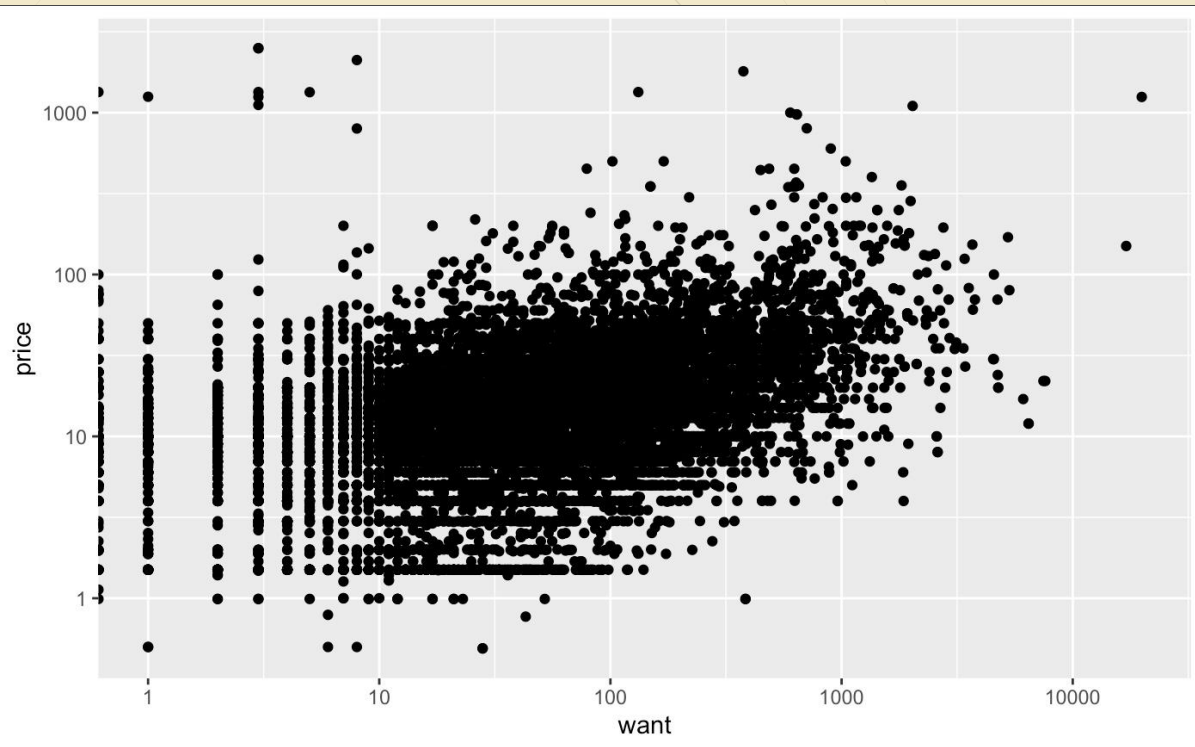
# Data Questions



- Do records with a higher demand also have a higher prices?
  - Which records are high in demand but low in cost?
  - What are the most popular records in a specific price range?
- 



# Exploratory Analysis



✱ **Positive**  
Want and price  
are correlated  
events

✱ **Weak**  
Statistically  
significant  
estimated  
correlation

0.2559768

# R Shiny Application

Popularity is calculated using discogs community votes. Large values indicate more demand relative to supply. This may indicate better resale value.

