# Capstone Project Presentation

## Tim Williams

# Method

## Unleash the Scrapers

---

**amazon**.co.uk prime

Electronics & Photo | digital slr

Deliver to Tim
Chertsey KT16 0QN

Shop by Department | Tim's Amazon | Today's Deals

Hello, Tim Your Account | Your Prime | Your Lists | Basket 0

Electronics | Best Sellers | Deals | Phones & Accessories | TVs & Home Cinema | Camera & Photo | Audio & HiFi | Computers & Accessories

1-24 of over 40,000 results for Electronics & Photo : "digital slr"

Sort by Relevance

### Show results for

Any Category
**Electronics & Photo**
Digital SLR Cameras
Digital Cameras
Point & Shoot Digital Cameras
Digital Camera Accessories
Compact System Cameras
SLRs
See more

### Refine by

**Subscription Option** (What's this?)
☐ Subscribe & Save

**Delivery Option**
☐ ✓prime

**Delivery Day**
☐ Get It Today

**Brand**
☐ Nikon
☐ Canon
☐ Sony
☐ Olympus
☐ Pentax
☐ Fujifilm
☐ Sigma
☐ Opteka
☐ Kaavie
☐ Lowepro
☐ DBROTH
☐ Tamron
☐ Hama
☐ webseller
See more

**Colour**

**Camcorder Display Size**
☐ Up to 1.9"
☐ 2 - 2.9"
☐ 3" & more

**Camera Optical Zoom**
☐ Up to 2.9x
☐ 3 - 5.9x
☐ 6 - 9.9x
☐ 10 - 19.9x
☐ 20 - 29.9x
☐ 30x & more

**Camera Sensor Resolution**
☐ Up to 7.9 MP
☐ 8 - 9.9 MP
☐ 10 - 11.9 MP
☐ 12 - 15.9 MP
☐ 16 - 19.9 MP

SPONSORED BY EVECASE
**Professional gear backpack for cameras and laptops**
Shop now ›

Evecase Shell DSLR Camera / 15.6-inch Laptop Water Resistant Backpac...
★★★★★ 5
✓prime

Ad feedback

Sponsored ⓘ
**Pixel Vertax D15 Battery Grip Power Pack for Nikon Digital SLR Camera D7100 D7200(Replacemen... MB-D15)**
by Pixel
£72.00 ✓prime
Only 10 left in stock - order soon.
★★★★★ ▾ 2

Sponsored ⓘ
**Tepoinn Vintage Waterproof Canvas SLR DSLR Digital Camera Bag Case for Canon Digital Camera Gadget... Bag**
by Tepoinn
£28.99 £69.99 ✓prime
Some colours are Prime eligible
★★★★½ ▾ 12

Sponsored ⓘ
**PIXEL Hot Shoe Mount Flash GN60 1-500Hz M/MULT Flash Modes Wireless Speedlite Flash Come wi... Digital SLR Camera.**
by PIXE
£58.99 ✓prime
Only 3 left in stock - order soon.

Sponsored ⓘ
**Pixel Vertax E11 Battery Grip For Canon EOS Digital SLR 5D Mark III 5DIII 5D3 5DS 5DSR Cameras (Rep... Canon BG-E11)**
by Pixel
£76.00 ✓prime
Only 15 left in stock - order soon.
★★★★★ ▾ 3

Amazon's Choice
More options available
**Canon EOS 1300D DSLR Camera with EF-S18-55 DC III F3.5-5.6 Lens - Black**
by Canon
£319.00 £329.99 ✓prime
Get it by Tomorrow, Jun 5
More buying choices
£233.00 (17 used & new offers)
★★★★½ ▾ 229

More options available
**Nikon D3400 + AF-P 18-55VR Digital SLR Camera & Lens Kit - Black**
by Nikon
£359.98 £499.99
Only 14 left in stock - order soon.
Eligible for FREE UK Delivery
★★★★☆ ▾ 120

---

a fraud as the product was not originally made to be sold in EU/EEA and does not meet EU/EEA standards. I was never made aware of this. I have suffered the consequences of "parallel-trade".

Sort by: Top | Filter by: All re... | All st... | All f... | Keywo | Search

### great quality at a reasonable cost
By Carles R. on 22 August 2017
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

Excellent camera for people that want to take good quality photos and are starting the world of photography. It has WIFI connectivity which means you can send images directly to your smartphone. Overall, very good quality at a reasonable cost.

20 people found this helpful

Helpful | Comment | Report abuse

### My product came without a serial number nor guarantee. ...
By Jaime Grau Lluch on 22 April 2018
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

My product came without a serial number nor guarantee. The Canon Counterfeit Department have certified that the Camera I bought is genuine but the labels have been falsified. This is a fraud as the product was not originally made to be sold in EU/EEA and does not meet EU/EEA standards. I was never made aware of this. I have suffered the consequences of "parallel-trade".

17 people found this helpful

Helpful | 1 Comment | Report abuse

### Five Stars
By Shirley Cross on 12 June 2017
Style: EF-S 18-55 IS II F3.5-5.6 Lens | Verified Purchase

Easy to use instructions, like the wifi option, easy to look at on my iPad.

15 people found this helpful

Helpful | Comment | Report abuse

### Good camera!!
By Emanuela on 18 September 2017
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

Fantastic camera, lightweight, easy to use, the display is big and clear.
I bought it for my mother and she is very happy with the clarity and quality of the pictures.

7 people found this helpful

Helpful | Comment | Report abuse

### Strange restrictions...
By forrest on 7 May 2018
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

I bought this directly from Amazon.
Basically a good camera. I have had cheaper cameras in the past, and the learning curve is steep, but it does the job. What lets this camera down is the fact that "linking" with media is incredibly poor. Even Windows 10 does not pick it up when you plug it in, and Linux is not supported in any shape or form. If you have time, you can apparently download software from the Canon website, but why you cannot just transfer pictures is bizarre. (All my former cameras - cheaper and simpler - just plugged in.) You can install an app for Android and Apple called Canon Connect, and this allows you to view pictures, though strangely you cannot do the same with the movies you shoot. These are all rather odd restrictions for a relative advanced camera. The most effective way of transferring images and movies to your pc is simply to yank out the DS card and use that.

3 people found this helpful

Helpful | Comment | Report abuse

### Five Stars
By wire21 on 6 November 2017
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

bought as a present, recipient loves it - photo quality is amazing

6 people found this helpful

Helpful | Comment | Report abuse

### Great value for money
By Mr Brian D Boyce on 4 April 2018
Style: EF-S 18-55 DC III F3.5-5.6 Lens | Verified Purchase

Canon EOS 1300D DSLR Camera with EF-S18-55 DC III F3.5-5.6 Lens - BlackA very good spec for the price. Using as a second camera with standard lens and my other camera fitted with a telephoto zoom/macro to save me have to swop lenses on the same body. The Canon 1300d was very easy to set up straight out of the box.Lightweight, but feels solid and well built, the HD movie feature is a bonus and produces excellant results. Have only had the camera for a few days but the initial results are very good and I am sure the camera will serve me well. No indication how many shots from the battery so will probably invest in a second battery in due course.

---

**Questions? Get fast answers from reviewers**

What do you want to know about Canon EOS 1300D DSLR Camera with EF-S18-55 DC III F3.5-5.6 Lens - Black?

See all 292 answered questions | Ask

### Customers also viewed these items

Nikon D3400 + AF-P 18-55VR Digital SLR Camera & Lens Kit - Black
by Nikon
£359.98
120

Canon EOS 750D Digital SLR Camera with 18-55 mm Lens
by Canon
£489.88
133

Canon EOS 2000D DSLR Camera and EF-S 18-55 mm f/3.5-5.6 IS II Lens - Black
by Canon
£469.99
1

Canon EOS 200D Digital SLR Camera with EF-S 18-55 mm f/4-5.6 IS STM Lens - Black
by Canon
£513.40
34

Canon EOS 1300D Kit incl. EF-S 18-55mm DC-III lens
by Canon
£326.99

Need customer service? Click here

# What did I Scrape?
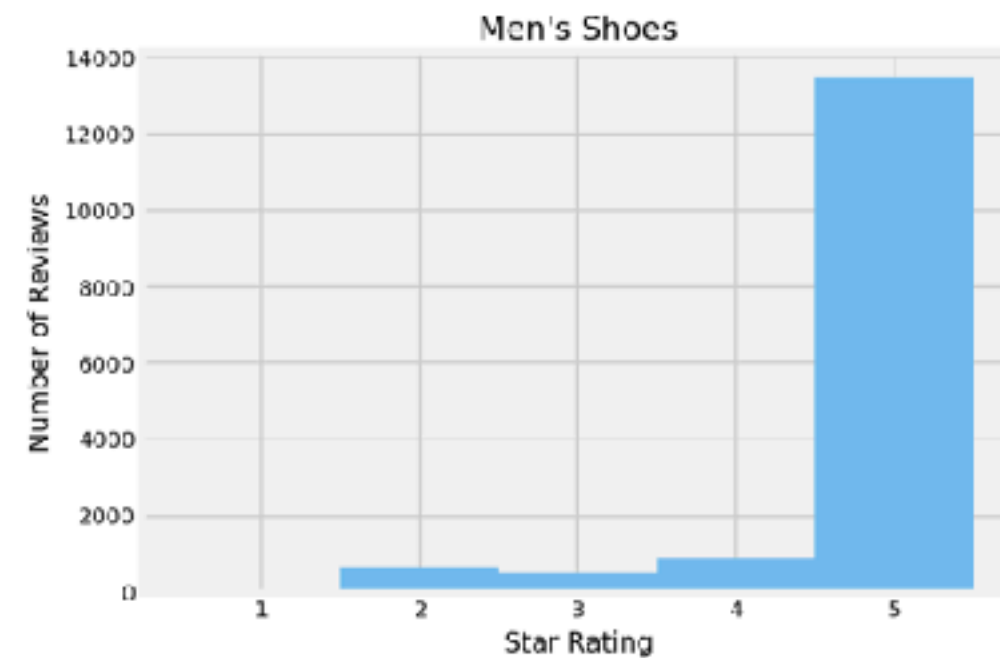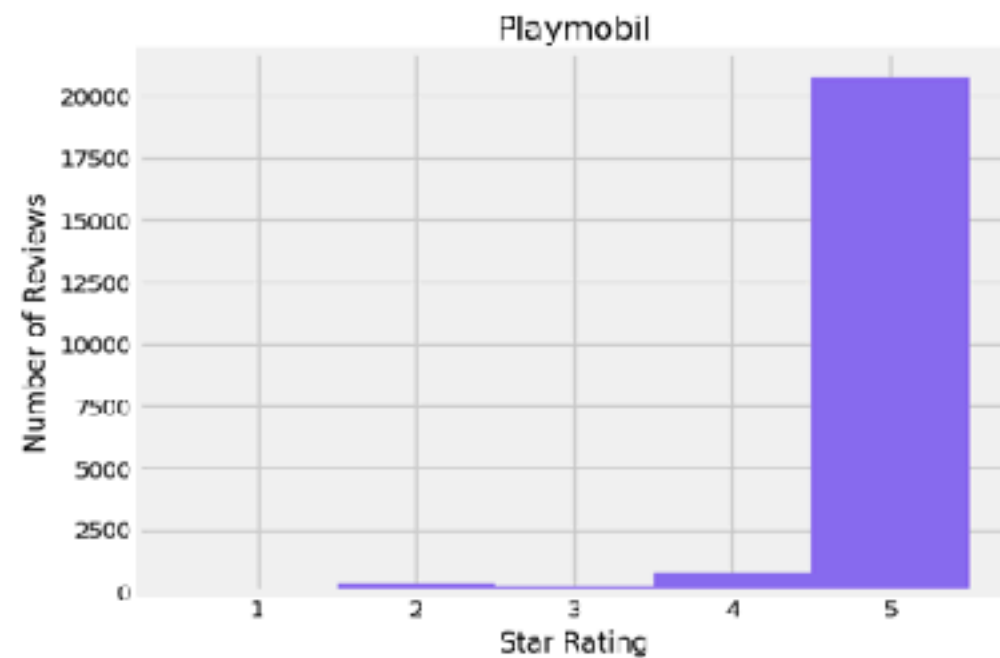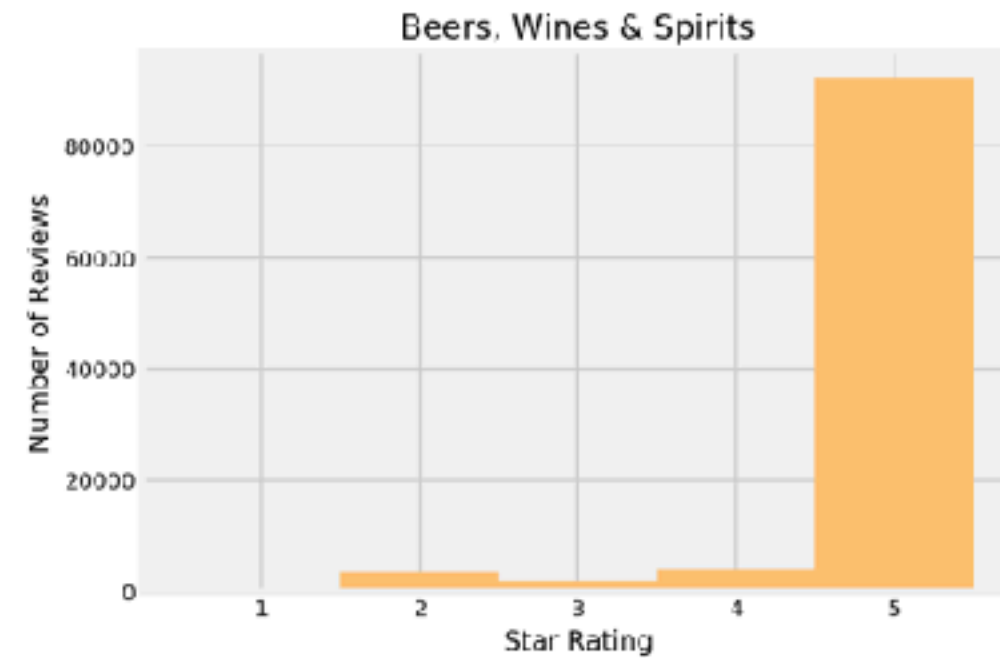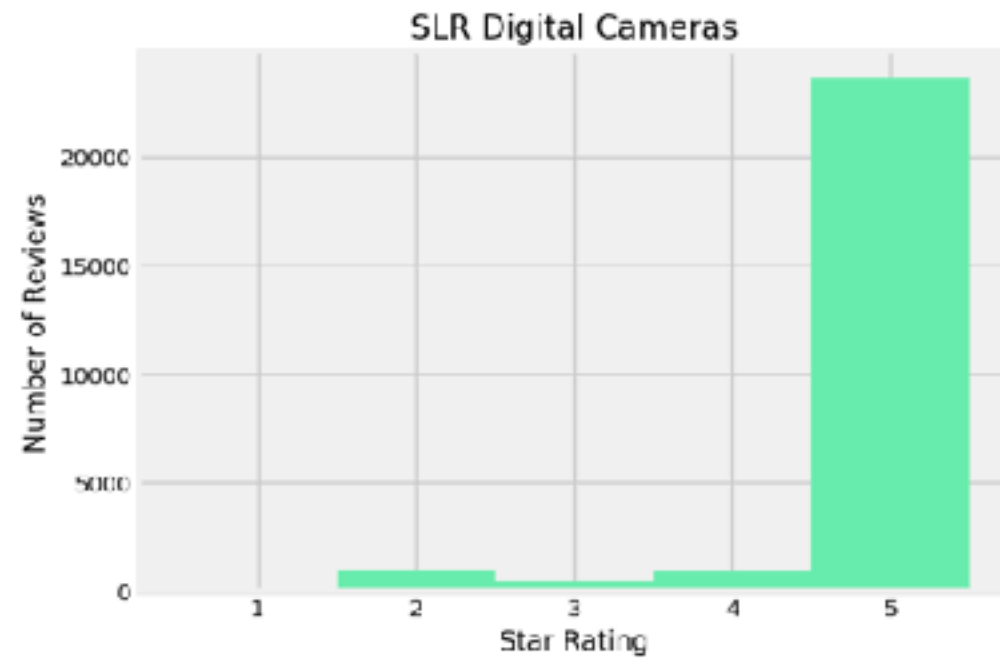
## The Four Categories

1. Digital SLR Cameras
2. Playmobil Toys
3. Men's Shoes
4. Beers, Wines & Spirits

## The Review Information

1. Product ASIN
2. Star Rating
3. Author ID
4. Page Index
5. Date
6. Review Text
7. Review Title
8. Comment ID
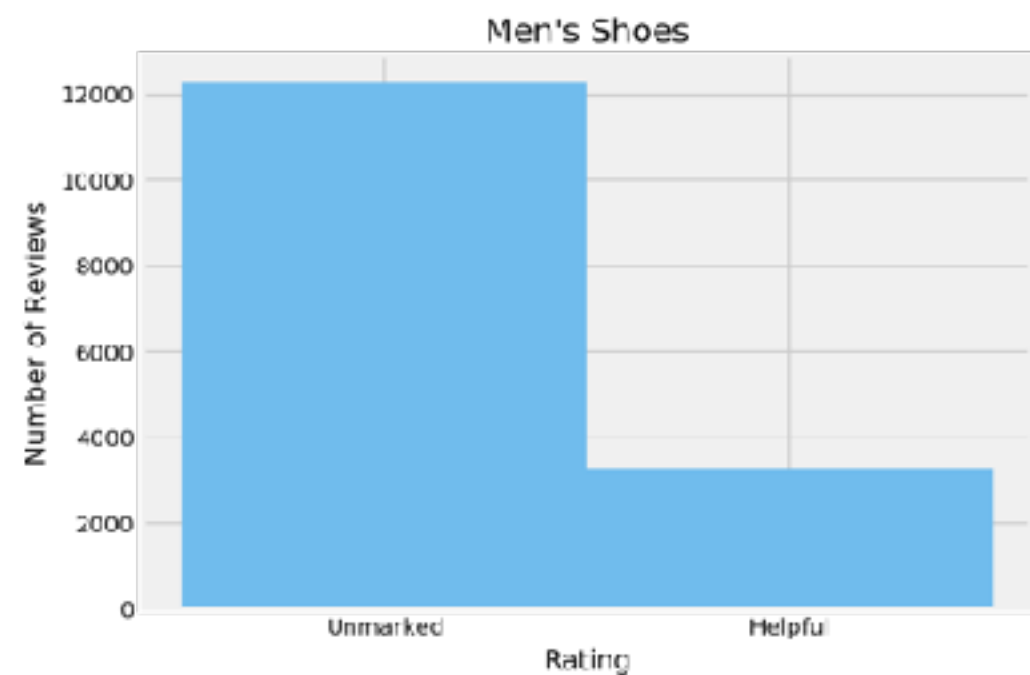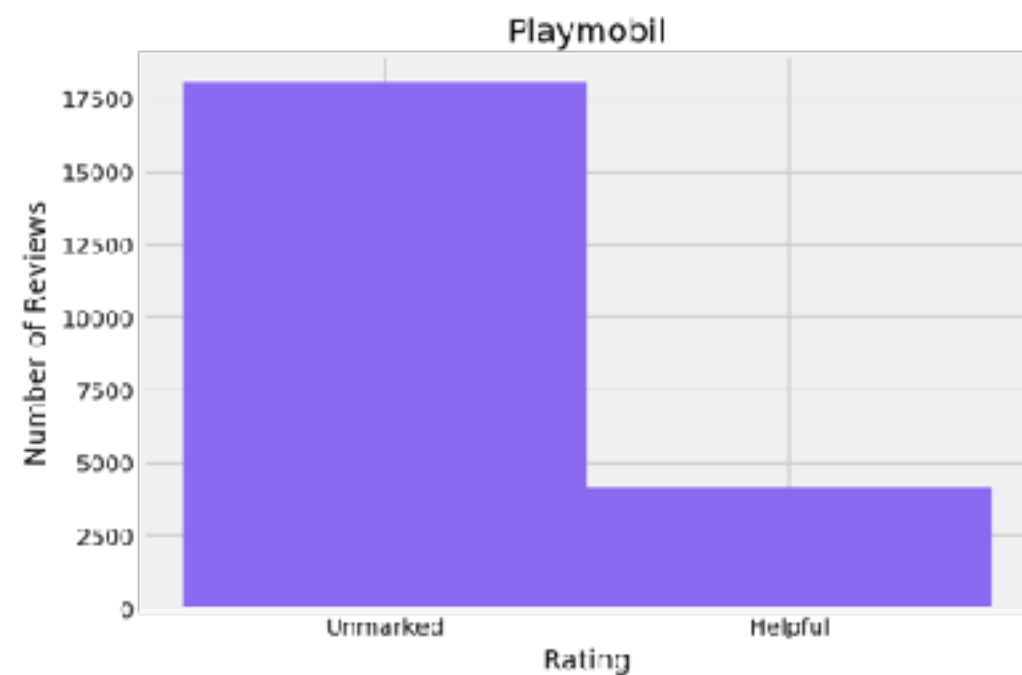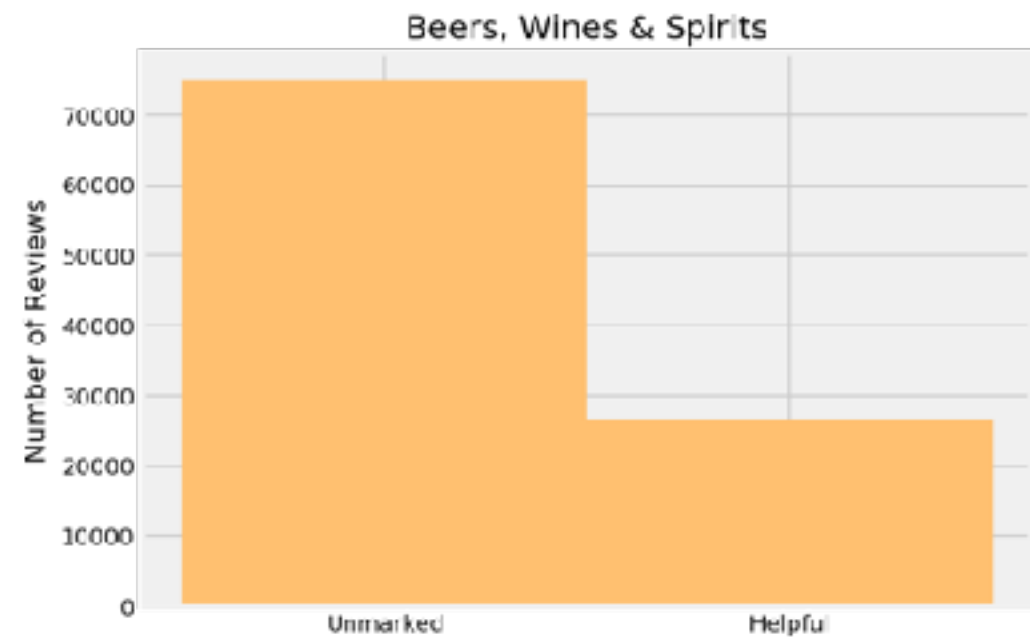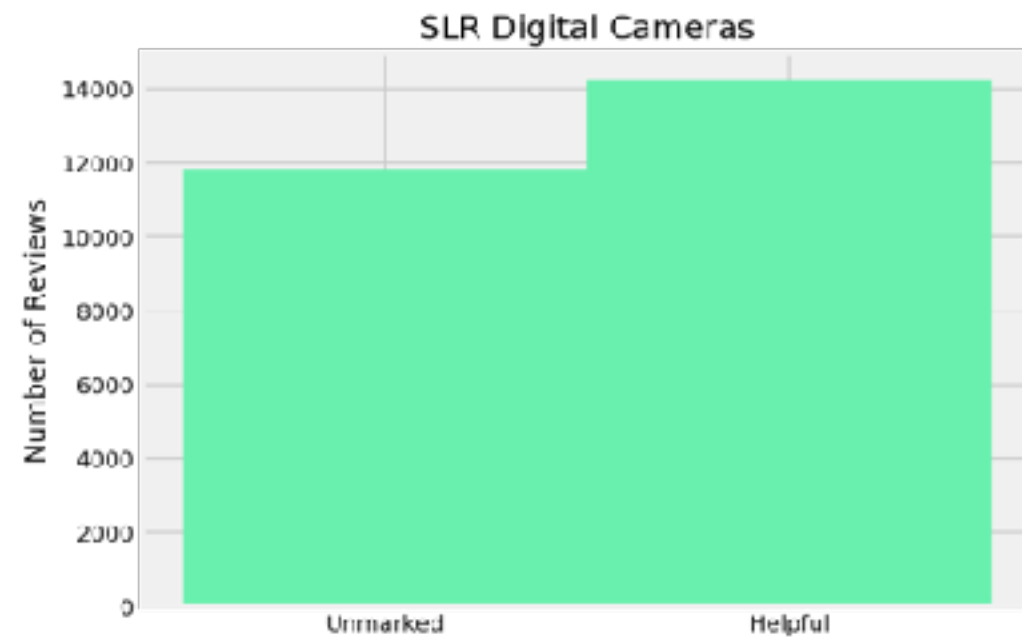9. Count of 'Helpful' Votes

# What Did I End Up With?
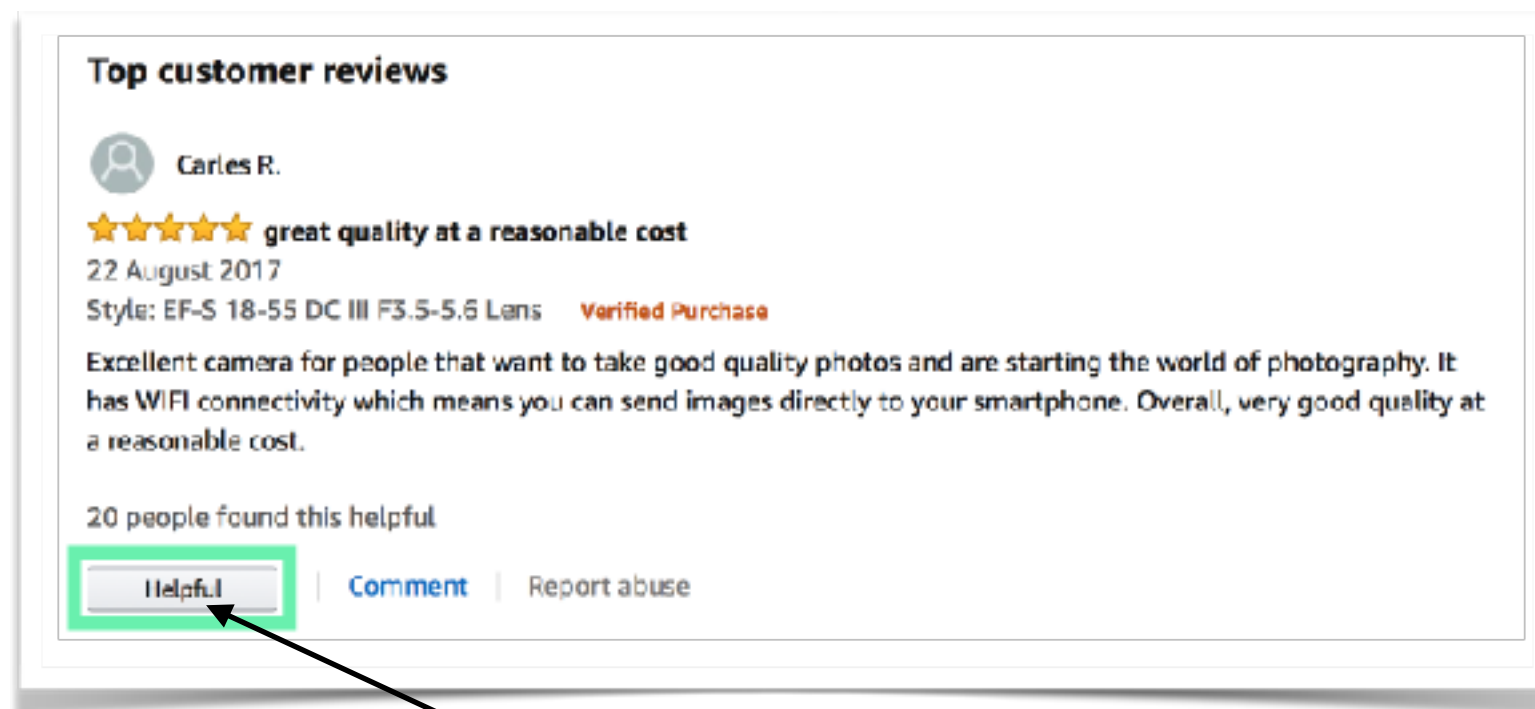


Histograms Showing Star Rating by Number of Reviews

# What Did I End Up Using?



Number of Reviews Marked as Helpful

# Amazon Customer Reviews

## What Was I Looking to Solve?



Top customer reviews

Carles R.

⭐⭐⭐⭐⭐ great quality at a reasonable cost

22 August 2017

Style: EF-S 18-55 DC III F3.5-5.6 Lens    Verified Purchase

Excellent camera for people that want to take good quality photos and are starting the world of photography. It has WIFI connectivity which means you can send images directly to your smartphone. Overall, very good quality at a reasonable cost.

20 people found this helpful

Helpful | Comment | Report abuse

**Can I predict whether another customer would find a review helpful?**

# LDA Analysis

## Topic Identification within the Reviews through statistical inference

Words that are more frequently found together in documents are more likely to belong to the same topics. Through repeated scanning through the text, the statistical properties are established and topics will show themselves.

## Expectation

Purchased for Size 9(uk) 43(eur), its about at least 1 size too large. The way it is advertised, it is supposed to fit quite snug so that it does not come off that easily and does not extend beyond your toes as to not trip you when you are walking around. But the advertised size is definitely too large to do either of what is advertised.

I bought this for my 2 year old granddaughter: she is now nearly four and she still loves it. The pieces are a little fiddly, and I was disappointed that the figures could only lie-down on the swing, but it didn't concern my granddaughter. We have since added to the park by buying additional figures and a camping set. It relates to her level of development and interests. Some people criticise such toys because they claim they leave nothing to the child's imagination, but I disagree: my granddaughter has developed her language skills, imagination, and social skills in playing with this. Both my children had Playmobil (and we still have many sets after 30 years) so I knew its value.

This is such a lovely wine; the flavour is dry but isn't harsh - I'd recommend this to people who don't normally like rosé if that makes sense! I find that a lot of rosé is too sweet or has an overly acidic aftertaste, this has a nice clean finish similar to a medium-dry white.

# LDA - Reality

**Topic 1:** 'taste', 'like', 'bottle', 'drink', 'gin', 'whisky', 'best', 'smooth', 'flavour', 'rum', 'really', 'vodka', 'wine', 'tried', 'buy'

**Looks like BWS**

**Topic 2:** 'camera', 'use', 'shoe', 'lens', 'good', 'size', 'like', 'quality', 'fit', 'canon', 'bought', 'picture', 'im', 'really', 'easy'

**Mix of Shoes & Cameras**

**Topic 3:** 'great', 'good', 'price', 'excellent', 'product', 'delivery', 'quality', 'arrived', 'item', 'time', 'service', 'expected', 'wine', 'fast', 'good price'

**Relating to Delivery/Service**

**Topic 4:** 'love', 'nice', 'gift', 'bought', 'lovely', 'present', 'loved', 'christmas', 'son', 'really', 'great', 'pleased', 'birthday', 'happy', 'grandson'
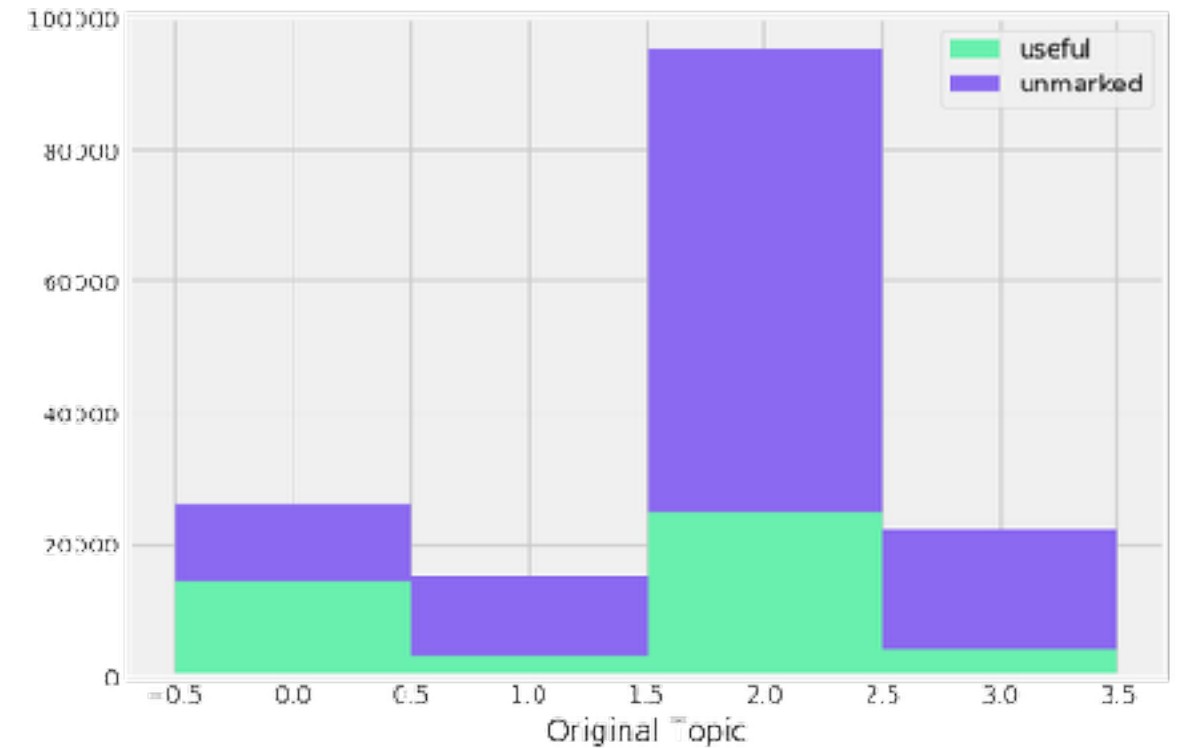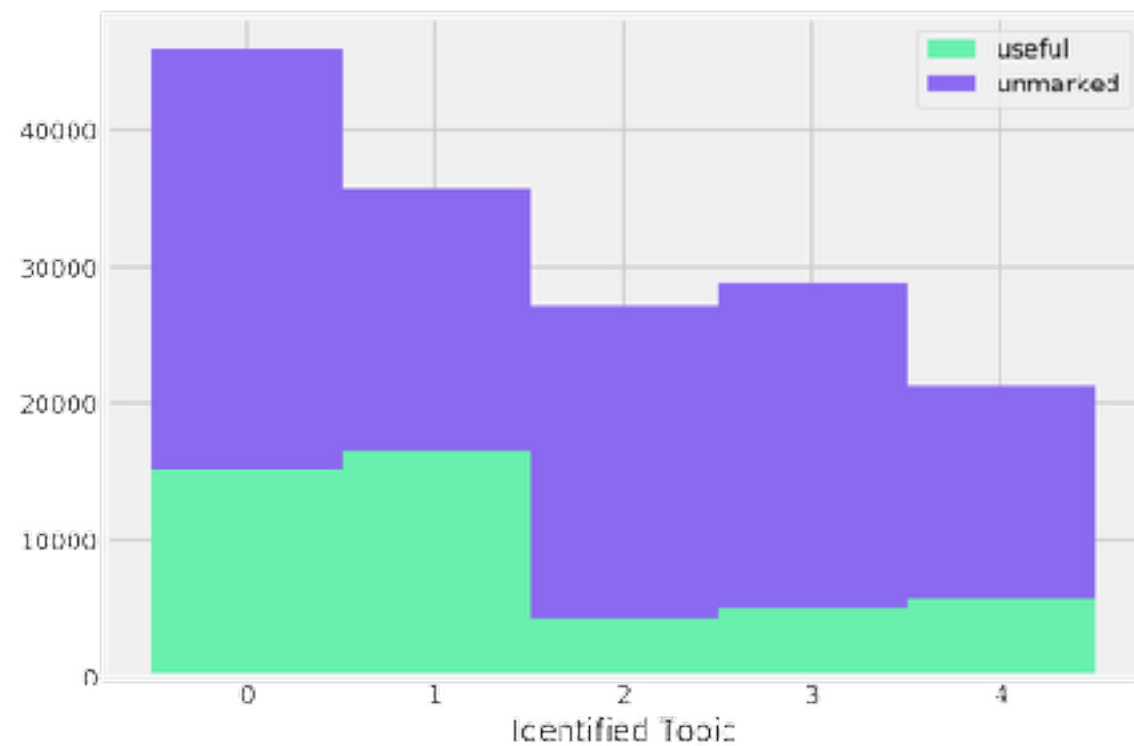
**Gift Giving/ Receiving**

**Topic 5:** 'year', 'playmobil', 'old', 'great', 'little', 'value', 'toy', 'set', 'good', 'year old', 'play', 'lot', 'money', 'day', 'bit'

**Looks Like Playmobil**

# Visual Inspection of Topic Utility



Histograms of Number of Reviews by Original and LDA Identified Categories

# Sentiment Analysis

## Analysis of Comment Content to Discover Author's Feeling Towards Topic

Words and sentences are compared to a pre scored set of words, to establish how positive or negative a review is

## Expectation

I expected that the most useful reviews would be those that were within a central bound of neutrality - either not too positive, or too negative as these would indicate an emotional response rather than objective response.

# Sentiment Analysis - Reality

**Most Positive Comments**

Excellent
Super!
Thanks
Amazing Value
Excellent! I loved!

**Most Neutral Comments**

Product is exactly as described.
Does the basic job
As expected
Cheap and does the job
Does what it says on the tin.

**Most Negative Comments**

Disappointed
Fake
Horrid
Awful
TERRIBLE

**Great at identifying the underlying sentiment in the review**

# Results of Classification

| | Predicted Unmarked | Predicted Helpful |
|---|---|---|
| **Actual Unmarked** | 28118 | 0 |
| **Actual Helpful** | 11642 | 0 |

**The model did not predict any reviews as useful**

# Parts of Speech Analysis

**Identification of the Makeup of Each Review by Review Structure**

By Identifying the parts of speech (noun, verb, adjective etc.) that go into making up a review, I could then generate the proportion of each part of speech that a review was made up from

## Expectation

I expected that If I could establish a structure that most typically modelled a useful review, then predictions could be made on new reviews to asses whether they matched the structure of a useful review.

# Parts of Speech - Reality
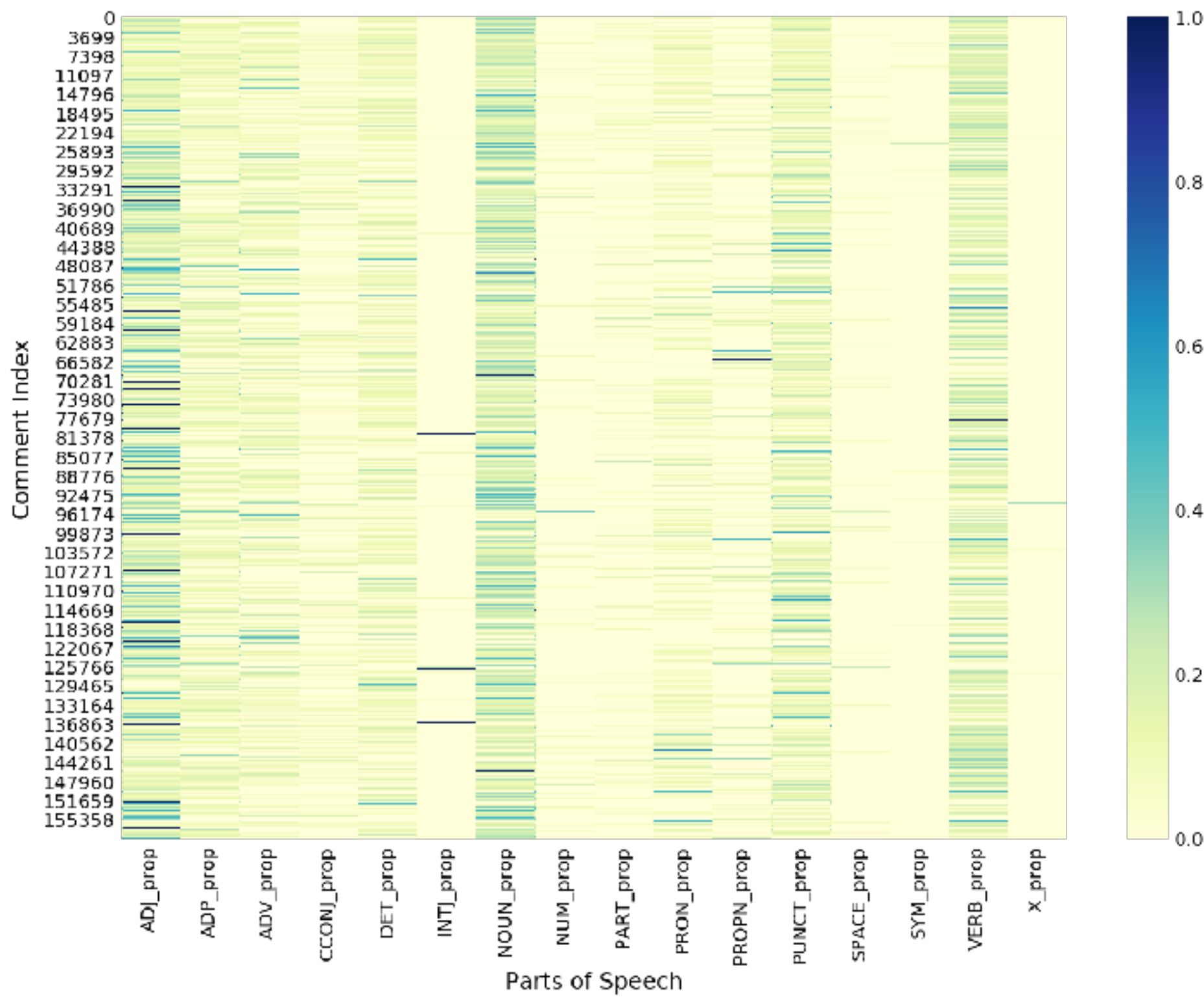


It is clear that Adjectives form the greatest part of reviews. Are these more descriptive reviews useful?

# Results of Classification

| | Predicted Unmarked | Predicted Helpful |
|---|---|---|
| **Actual Unmarked** | 27842 | 301 |
| **Actual Helpful** | 11521 | 88 |

**The model is an improvement over sentiment, but it still is not a good classifier**

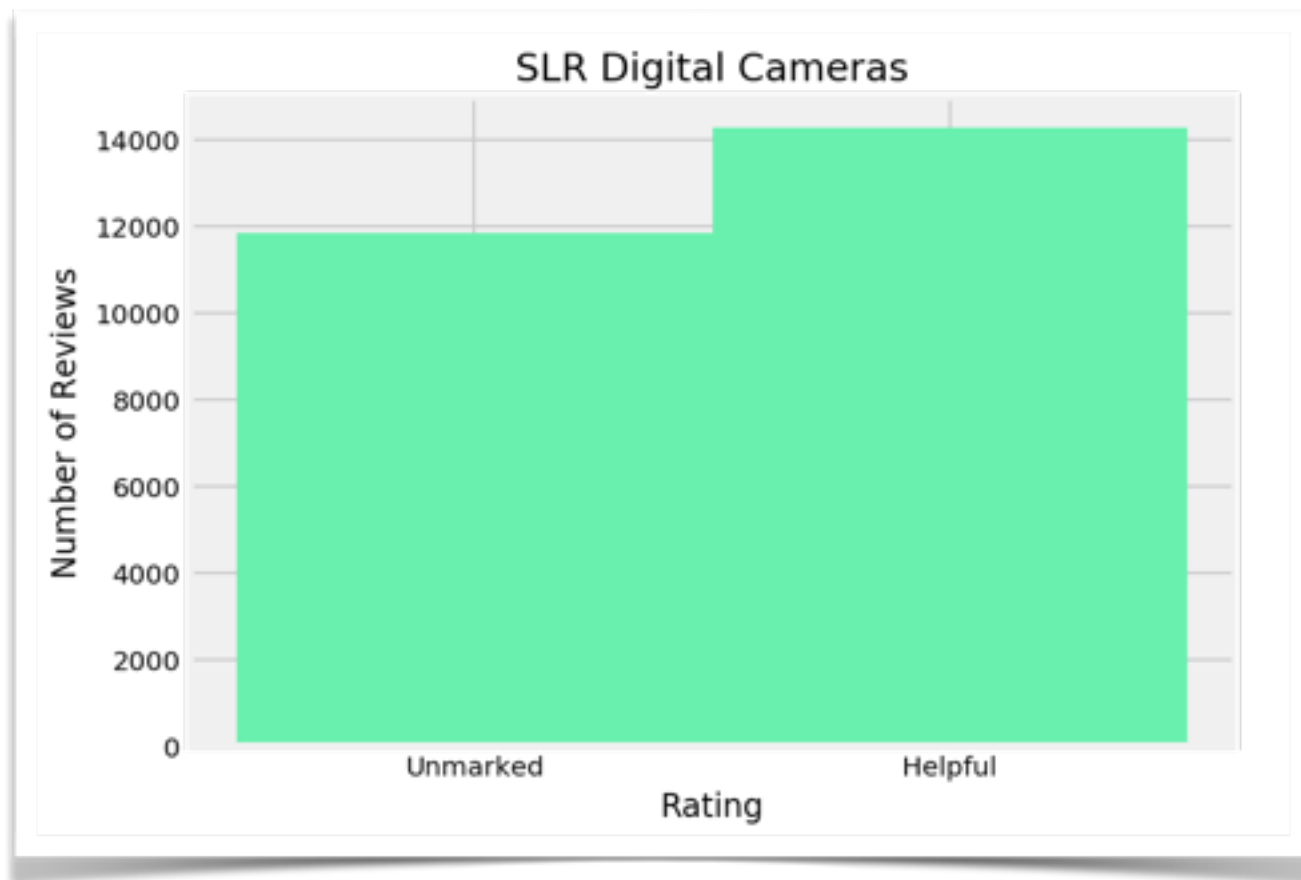| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.71 | 0.99 | 0.82 | 28122 |
| 1 | 0.25 | 0.01 | 0.02 | 11638 |
| avg / total | 0.57 | 0.70 | 0.59 | 39760 |

**We want good recall with our model - we don't mind showing users unhelpful reviews, but don't want to not show useful ones**

This model's recall of 0.01 out of 1 can be improved on

# The Separation

# SLR Digital Cameras



**Objective**

Now I was making multiple models, it was all about building the least computationally intensive model for the job at hand, because a model for each topic running in real time could become more costly than the value derived from such a model
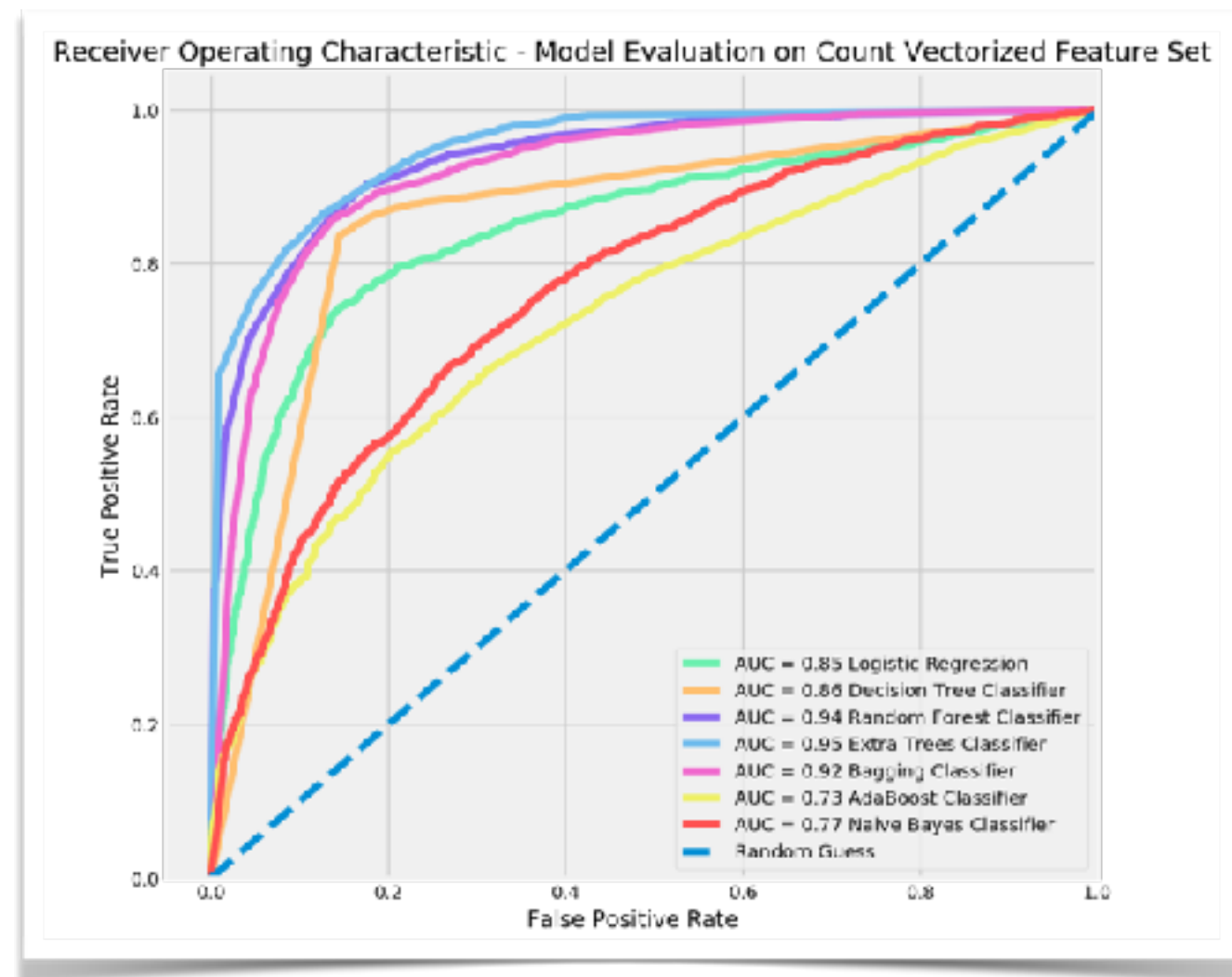
**This is the most even split of classes, and therefore was my best bet at building a good classifying model**

# Simple Counting

**I looked at simple counts of the most frequent words used in the reviews corpus per review**

**Then fitting a classifying Algorithm**

In this case, an Extra Trees Classifier proved to be my best estimating model



Receiver Operating Characteristic - Model Evaluation on Count Vectorized Feature Set

AUC = 0.85 Logistic Regression
AUC = 0.86 Decision Tree Classifier
AUC = 0.94 Random Forest Classifier
AUC = 0.95 Extra Trees Classifier
AUC = 0.92 Bagging Classifier
AUC = 0.73 AdaBoost Classifier
AUC = 0.77 Naive Bayes Classifier
Random Guess

# Better Counting

**I Then Looked at the Term Frequency - <mark>Inverse Document Frequency</mark>**

With count vectorised matrices of words, each word is considered as important as each other within the individual review, but with the inverse document frequency, we are able to reduce the importance of words that appear in multiple reviews, effectively filtering for the most 'important' words per review

**..And fit a classifying Algorithm**
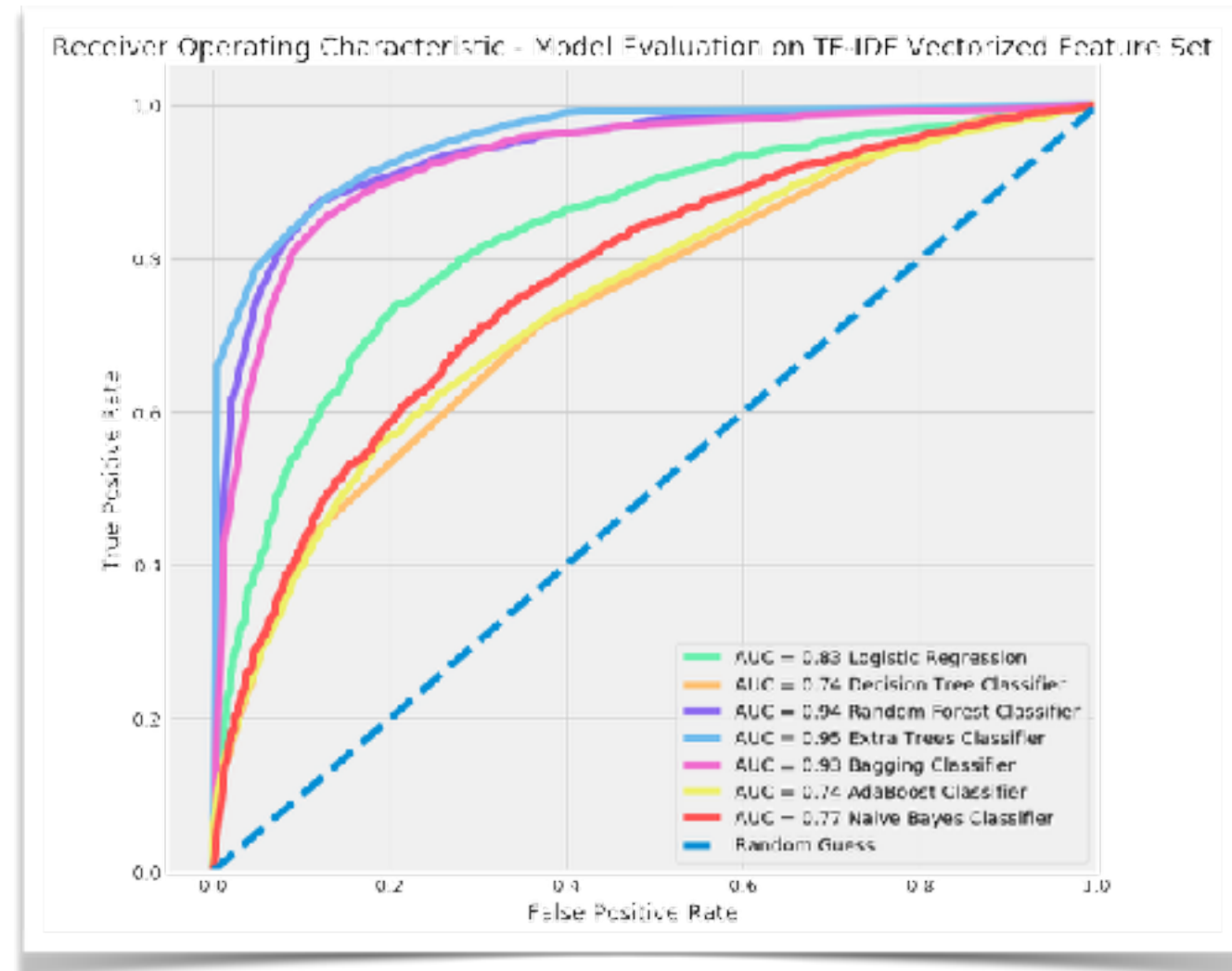


In this case an Extra Trees Classifier proved to be my best estimating model too

# Picking Between the Two



Precision-Recall curve: Average Precision

TF-IDF Extra Trees Classifier 0.84
Count Vectorised Extra Trees Classifier 0.61
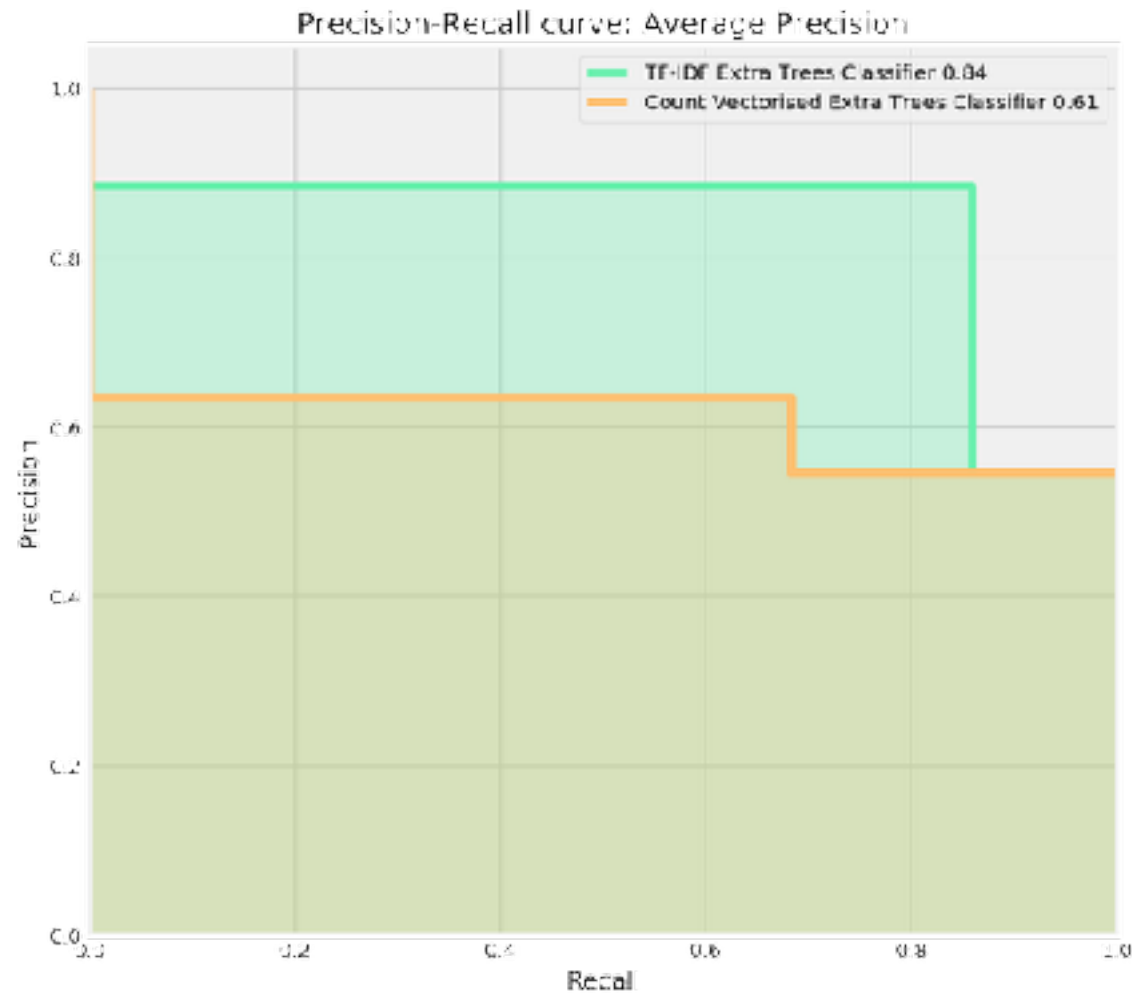
**As with most graphs, we want: UP and to the RIGHT**

**The TF-IDF matrix of words gives us much greater precision - it will better indicate when the review is useful**

# What Makes a Useful SLR Review?

**The good thing about classifiers using word matrices is that we can look at the most important words/phrases that make useful reviews**

From the important features, it looks like the useful reviews have a narrative style, highlight positives and mention downsides

| Word/Phrase | Relative Importance |
|---|---|
| going strong | 0.009352 |
| got nikon | 0.008304 |
| came time | 0.007653 |
| camera work | 0.004414 |
| workhorse | 0.004188 |
| produces excellent | 0.003741 |
| love camera | 0.003592 |
| june | 0.003584 |
| value price | 0.003544 |
| excellent buy | 0.003431 |
| photography camera | 0.003297 |
| died | 0.003125 |
| hang | 0.003011 |
| upgraded nikon | 0.002983 |
| bought body | 0.002956 |
| downside | 0.002866 |
| lens 18 55mm | 0.002862 |
| early | 0.00279 |
| pushed | 0.002786 |
| camera worth penny | 0.002716 |

# Website Implementation



**ismyreviewuseful.com**

# Are There any pitfalls?

Sure, this model is based on only the words that go into the review
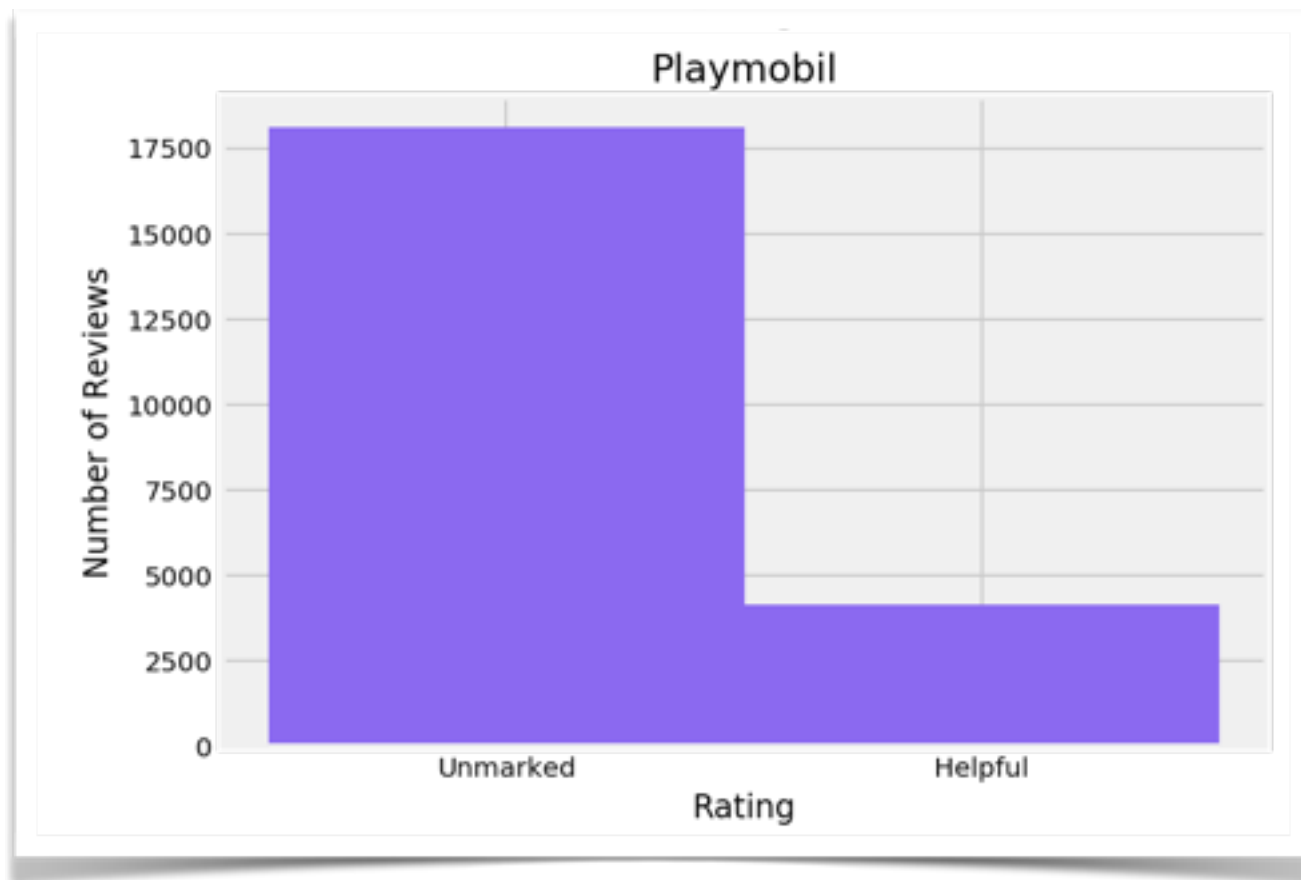
Q. What if I just typed random words?
A. it would make no sense, but could still be regarded as useful

# Playmobil

# Playmobil



The Biggest Problem here:

Those using the product are not the ones reviewing it. Reviews are likely to be about the perceived enjoyment of a toy, rather than experienced enjoyment.

**This is the most uneven split of classes, and therefore could be tough to find a good model**

# What even is a useful review?

In exploring the reviews in this category, I was unable to find any very good models. Upon closer inspection, i could see why..

Using a cosine difference between tf-idf vectorised reviews, I could see reviews that were similar, but not necessarily considered as useful as each other

## 2 Useful Reviews

1 Great,  but beware of prying young eyes. It comes as is seen with an address label stuck on the front, so I had to grab it and hide it under my jumper before it was noticed!

2 This is a great toy. Very happy four year old. Unfortunately the postage label was stuck directly on the box so don't order as a surprise.

## 1 Useful Review

1 Very cross!!!! This arrived with no outer box, so daughter saw what it was when it was delivered!!!

2 Arrived unboxed so my child, who's present it was saw it, also parts missing. Shocking.

**Which one was useful?**

# What's Next?

Keep chipping away with the other categories

Look for more balanced categories within Amazon

Encourage you to mark useful reviews on Amazon

Encourage Amazon to include a 'not helpful' button again

~~Change the world~~

# Thank You
# Vielen Dank
# Merci Beaucoup