



Scrapy Project: Microsoft App Store

08/08/2018

BY JENSEN HU

Table of Contents

- ▶ Introduction
- ▶ Scrapy Flow Diagram
- ▶ What was scraped?
- ▶ Summary of Results

Table of Contents

- ▶ Introduction
- ▶ Scrapy Flow Diagram
- ▶ What was scraped?
- ▶ Summary of Results

Introduction

- ▶ Objective: scrape the Microsoft App store to explore the contents, ratings, reviews, overall app community.

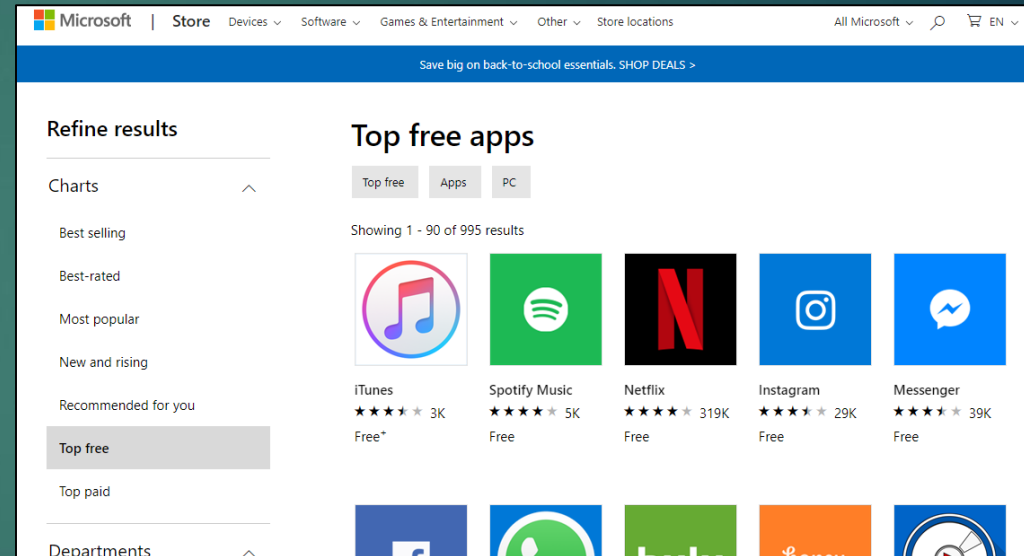
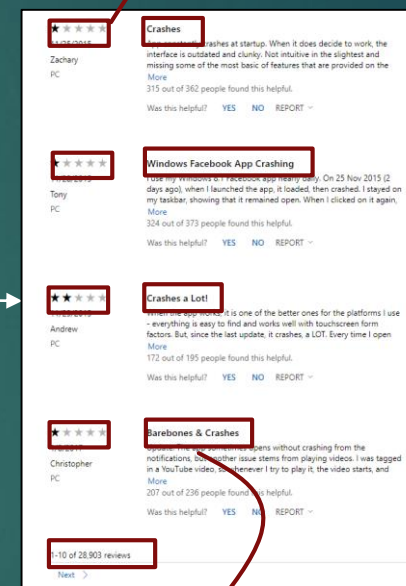
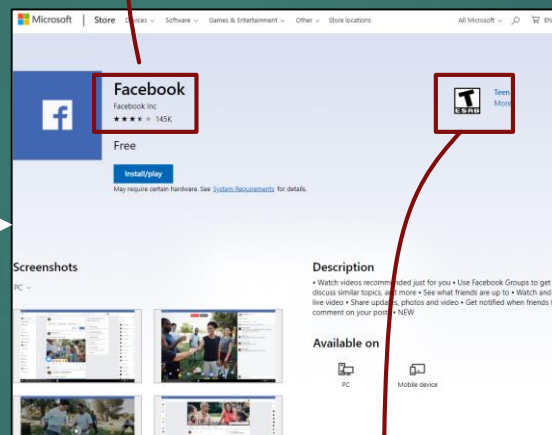
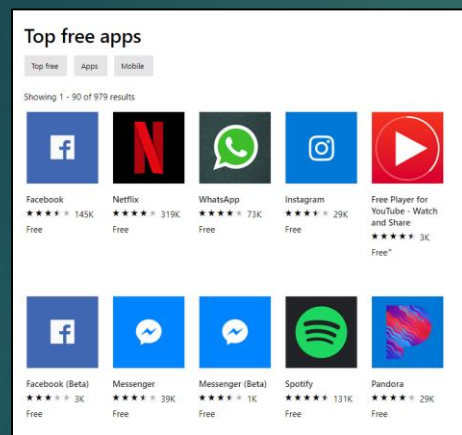


Table of Contents

- ▶ Introduction
- ▶ Scrapy Flow Diagram
- ▶ What was scraped?
- ▶ Summary of Results

Scrapy Flow Diagram



Individual
Comment
Rating

ESRB Rating

Comment
Header

Table of Contents

- ▶ Introduction
- ▶ Scrapy Flow Diagram
- ▶ What was scraped?
- ▶ Summary of Results

What was Scraped?

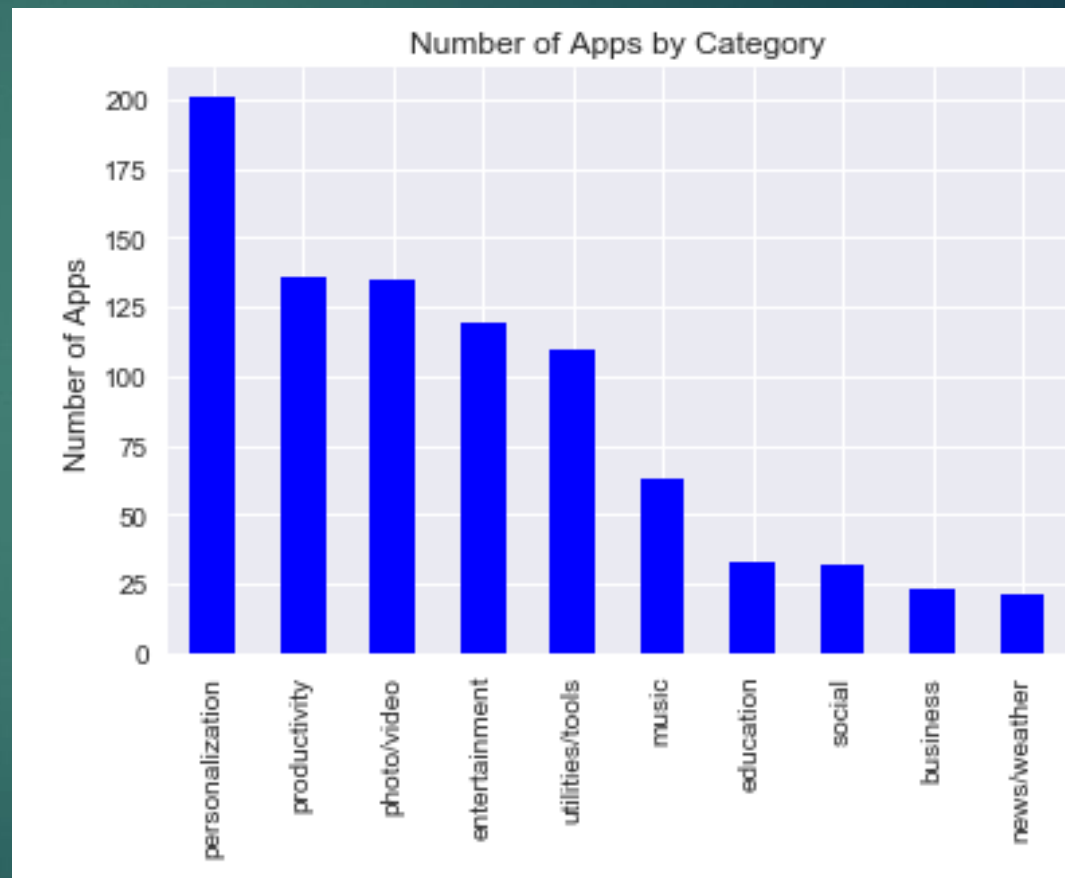
A	B	C	D	E	F	G	H
App_Name	Category	Comment_header	Comment_rating	ESRB_rating	Num_Rated	Num_Reviewed	Rating
iTunes	music	Don't bother...	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music		1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	iPod won't work with it	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	MICROSOFT STORE IS TRASH	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Love it	5	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Does no work	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Outdated and Laggy	3	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Does Not Install Apple Mobile Software and Others	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Doesn't work	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
iTunes	music	Terribad	1	Rating Board ESRB. Rated For ages 13 and up.	3758	1-10 of 1,451 reviews	3.5
AdBlock	productivity	Check the Acceptable Ads filter first! Great adblock.	5	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Generally Works Great!	5	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	A working adblocker for Edge	4	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Extension keeps shutting off	1	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Doesn't always work, can't access options *UPDATE 8/4*	5	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Problems... SOLVED?... Nope	1	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Disables Video	3	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Not useful as an ad blocker	1	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Wonderful!! Add free experience!	5	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
AdBlock	productivity	Worked fine for a couple of weeks, then stopped working	2	Rating Board ESRB. Rated For all ages.	1112	1-10 of 688 reviews	3.5
World National Parks	personalization	Great	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5
World National Parks	personalization	Beautiful scenes!	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5
World National Parks	personalization	Stunning	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5
World National Parks	personalization	ðŸ’	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5
World National Parks	personalization	The USA Parks, Canada come I rest World National Parks	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5
World National Parks	personalization	this app butifull	5	Rating Board ESRB. Rated For all ages.	74	1-10 of 24 reviews	4.5

Table of Contents

- ▶ Introduction
- ▶ Scrapy Flow Diagram
- ▶ What was scraped?
- ▶ Summary of Results

Summary of Results

- Scraped 873 apps
- Categories with > 20 apps



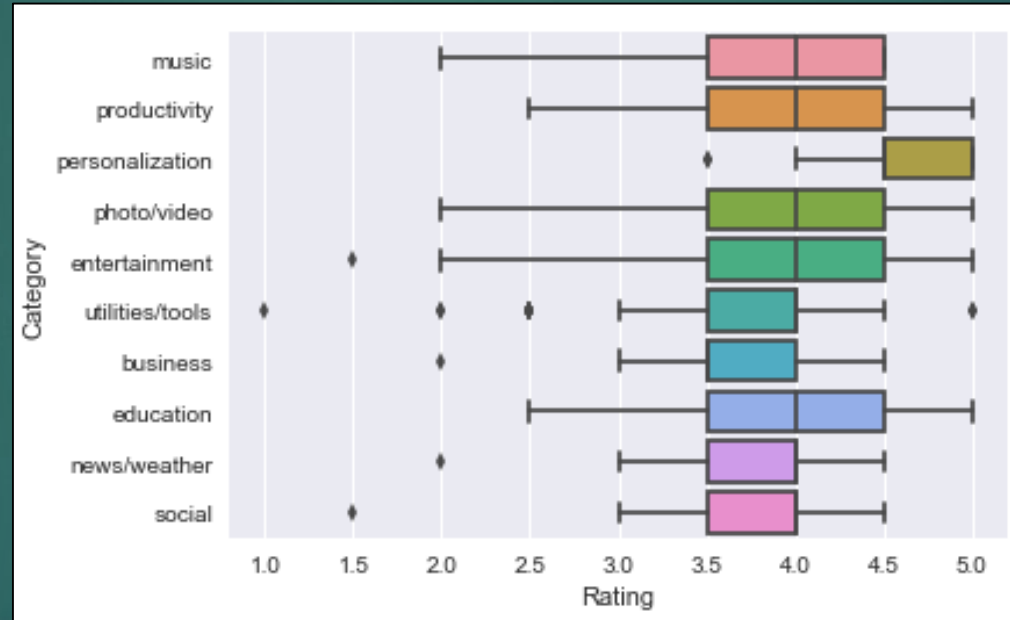
Summary of Results: Personalization Apps

App_Name	Category	ESRB_rating	Num_Rated	Num_Reviewed	Rating	u
World National Parks	personalization	Rating Board ESRB. Rated For all ages.	74	24	4.5	
Islands in the Sun	personalization	Rating Board ESRB. Rated For all ages.	55	21	4.5	
Remix 3D Winterscapes	personalization	Rating Board ESRB. Rated For all ages.	44	26	4.5	
Twilight by Chad Copeland	personalization	Rating Board ESRB. Rated For all ages.	21	10	5.0	
Colors of India	personalization	Rating Board ESRB. Rated For all ages.	24	10	5.0	

Summary of Results: News/Weather Apps

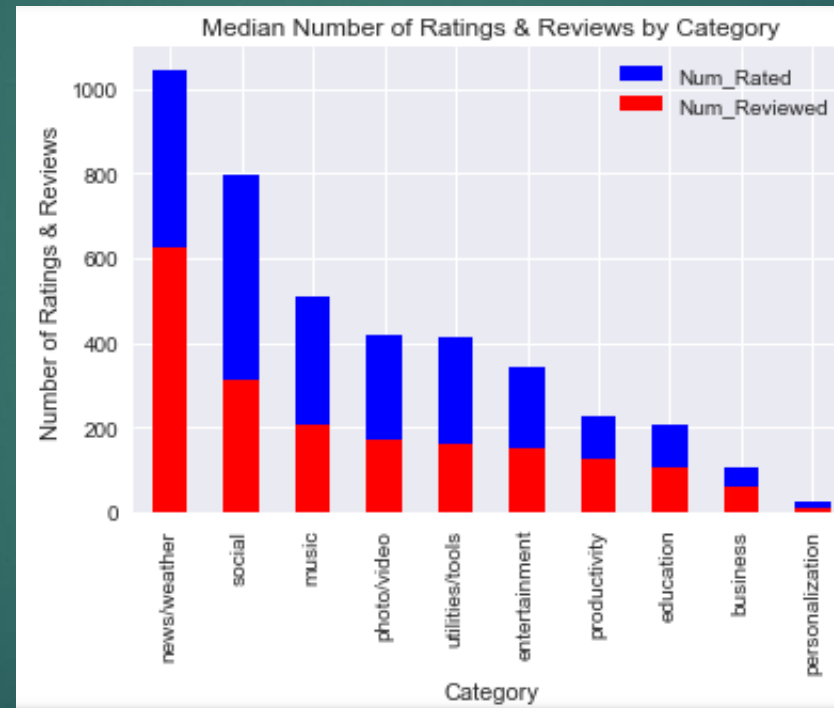
App_Name	Category	ESRB_rating	Num_Rated	Num_Reviewed
The Weather Network	news/weather	Rating Board ESRB. Rated For all ages.	796	370
Windows Central	news/weather	Rating Board ESRB. Rated For all ages.	1044	689
Sense Clock	news/weather	Rating Board ESRB. Rated For all ages.	95	73
Police Radio Scanner	news/weather	Rating Board ESRB. Rated For all ages.	2129	853
Newsflow	news/weather	Rating Board ESRB. Rated For ages 10 and up.	509	305
CBS Local	news/weather	Rating Board ESRB. Rated For ages 10 and up.	388	225
NPR One	news/weather	Rating Board ESRB. Rated For ages 13 and up.	1060	657

Summary of Results: Ratings by Category



Summary of Results:

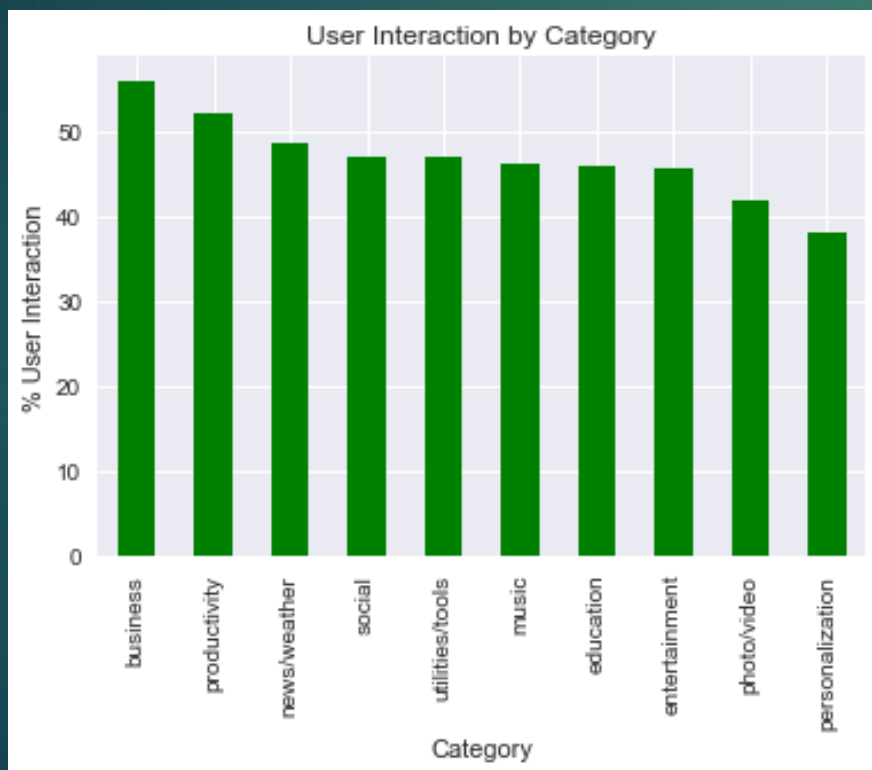
Number of Ratings & Reviews



Summary of Results: User Interaction

App_Name	Category	ESRB_rating	Num_Rated	Num_Reviewed	Rating	user_interaction
iTunes	music	Rating Board ESRB. Rated For ages 13 and up.	3758	1451	3.5	38.610963
AdBlock	productivity	Rating Board ESRB. Rated For all ages.	1112	688	3.5	61.870504

User interaction: Number of Reviews / Number of Ratings



Business WordCloud



Personalization WordCloud





Questions?