



ReDI School of
Digital Integration

Data Analytics 2023
Assignment 2
Analysis of IKEA app reviews
from Play Store

Introduction - Why this?

IKEA Android App

- From a big Swedish brand
- Interesting historical data
- Close to a real work scenario



Methodology

Data cleaning

- Fix misaligned due to bad CSV import
- Remove duplicated
- Handle missing values

(e.g. missing app version, replaced with est. version based on date)

- Convert data to proper types

(e.g. date strings into Excel dates)

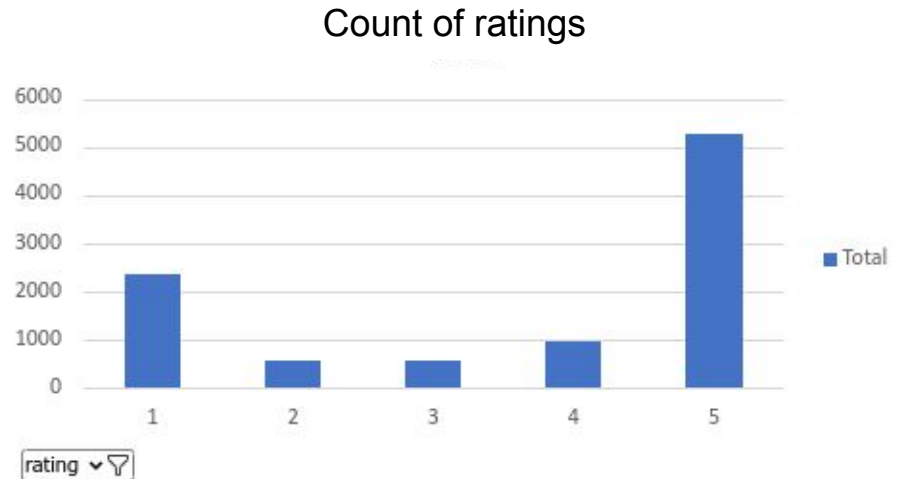
The screenshot shows an Excel spreadsheet with a data table. The formula bar contains a complex LAMBDA function. The task pane on the right shows several tabs, including 'ikea_dataset_raw', 'base_cleaning', 'fix_misalignment', 'remove_duplicates', 'type_cleaning', 'functions', 'ikea_dataset_tab', 'remove_duplicates_tab', and 'cleaned_tab'.

id	index	user_fullName	metadata_avatarURL	rating	review	dateTime	foundUseful	answer	appVersion	answer_fullName
003d6091-dd50-45e3-9fc5-934c9e4c32de	153	Nathan DUBOUFF	https://play-lh.google...	5	I love th	2023-08-28T21:03:	0	#N/A	3491	Inter IKEA System
0055e1a8-4737-47f5-89f6-b5f974345f04	89	J F	https://play-lh.google...	5	Very qui	2023-09-08T20:58:	0	#N/A	3500	Inter IKEA System
009106bd-2e8c-418b-b8c1-03c9cd60d6c1	1016	Sana Shaikh	https://play-lh.google...	5	Nice	2023-02-15T14:59:	0	#N/A	3321	Inter IKEA System
00b73251-9ec6-4d24-ac5d-146ab93f77cb	38	Jenny Tang	https://play-lh.google...	5	Great a	2023-09-19T07:44:	0	#N/A	3500	Inter IKEA System
00eafb19-848e-4e77-85eb-46c845529342	511	cellathescot	https://play-lh.google...	5	Helpful	2023-06-19T15:46:	13	#N/A	#N/A	Inter IKEA System
011c99f-84a2-45f1-8044-7ecd28e563fc	694	cumulus cloud	https://play-lh.google...	5	Great, a	2023-04-30T18:09:	0	#N/A	3401	Inter IKEA System
01763d37-da67-4b94-b254-28deb42baf9b	179	Vatsala Raj	https://play-lh.google...	5	Love ike	2023-08-24T18:19:	0	#N/A	3470	Inter IKEA System
0192b055-9cd2-4cfe-97ad-3fde3f14c7f	55	slawomir chwaast	https://play-lh.google...	5	One of r	2023-09-16T18:45:	0	#N/A	3500	null
01b51a67-8ca6-4be9-bf5a-0be7d684c302	46	Bernadette Tarlo	https://play-lh.google...	5	Fun!!	2023-09-18T02:03:	0	#N/A	3500	Inter IKEA System
01f8439b-44c1-4969-a1a0-0c0ee02b537	1074	Pass TheMash	https://play-lh.google...	2	E d eed	2023-02-08T01:08:	0	#N/A	3080	null
01f98097-539c-474d-b995-026d123243e3	503	Margaret L	https://play-lh.google...	5	Searcha	2023-06-21T12:51:	0	#N/A	3440	Inter IKEA System
0212b30b-8051-4e4c-bfa3-c0a0fec4788f	207	Abby AB	https://play-lh.google...	1	Amn isla	2023-08-19T05:16:	0	#N/A	#N/A	null
023e75cf-f52e-462f-b1f3-f4f5194b6df9	844	Sanoj Mathew	https://play-lh.google...	5	Good. E	2023-03-23T21:01:	0	#N/A	3380	Inter IKEA System
024701c0-e45-4539-94ee-494043a4ff0b	854	Iosie Spectator	https://play-lh.google...	5	Good w	2023-03-21T06:56:	1	#N/A	3371	Inter IKEA System
02560a1f-5a2f-4a3d-aa0b-27f6f3edf99	670	akhil patil	https://play-lh.google...	2	Why no	2023-05-06T15:06:	3	#N/A	3401	null
02cb165e-2532-4f90-b0d2-cc8ec68aabc4	664	Iody Buttl	https://play-lh.google...	5	Love it!	2023-05-07T09:22:	0	#N/A	#N/A	Inter IKEA System
02d32e0d-94de-4b3b-a4d2-f72231977d7d	339	Pitchaiha Naik B	https://play-lh.google...	5	Seen in	2023-07-26T15:58:	0	#N/A	3470	null
0309c209-8d31-4d43-8e53-2b310c681e39	487	Prathyusha Yada	https://play-lh.google...	5	Amazing	2023-06-24T11:09:	0	#N/A	3440	Inter IKEA System

Methodology

Initial Analysis

- Obtain statistics for interesting fields
- Look for patterns over time
 - full history
 - cyclic: by month / time of day
- Create a correlation matrix across different fields



Methodology

Initial Analysis

- Obtain statistics for interesting fields
- **Look for patterns over time**
 - **full history**
 - cyclic: by month / time of day
- Create a correlation matrix across different fields

Reviews over year, full history



* partial data for 2023, still ongoing

Methodology

Initial Analysis

- Obtain statistics for interesting fields
- **Look for patterns over time**
 - full history
 - **cyclic: by month / time of day**
- Create a correlation matrix across different fields



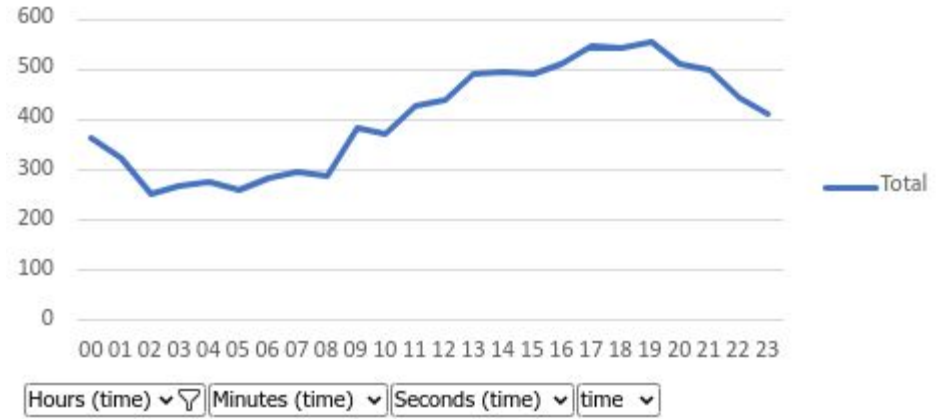
* partial data for 2023, still ongoing

Methodology

Initial Analysis

- Obtain statistics for interesting fields
- **Look for patterns over time**
 - full history
 - **cyclic: by month / time of day**
- Create a correlation matrix across different fields

Reviews by hour



Answers by hour



Methodology

Initial Analysis

- Obtain statistics for interesting fields
- Look for patterns over time
 - full history
 - cyclic: by month / time of day
- **Create a correlation matrix across different fields**

Correlation Matrix

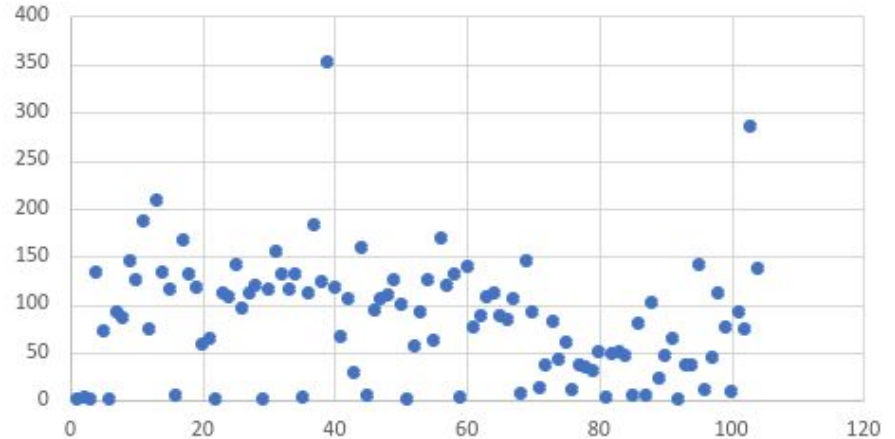
	0 rating	date_review	time_review	foundUseful	app_version	score_review
rating	1	0.07	0.01	-0.14	0.09	0.63
date_review	0.07	1	-0.03	-0.02	0.94	0.03
time_review	0.01	-0.03	1	-0.02	-0.03	0.01
foundUseful	-0.14	-0.02	-0.02	1	-0.02	-0.21
app_version	0.09	0.94	-0.03	-0.02	1	0.05
score_review	0.63	0.03	0.01	-0.21	0.05	1

Methodology

Review score (sentiment analysis) by app version

- No clear direction emerged from available data
- Reproduces a possible work setting scenario

App version vs review amount



Methodology

Review score (sentiment analysis) by app version

- Hypothesis:
 - **Null:** there's no significant correlation between app version and review score, $\mu_0 = 0$
 - **Alternative:** the app version correlates positively with the review score, $\mu_0 > 0$

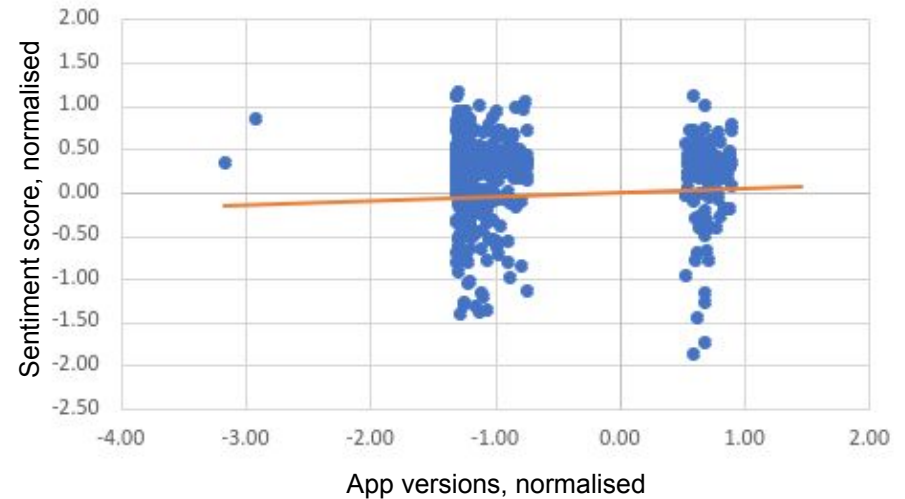


Results

Review score (sentiment analysis) by app version

- Regression t-test with 95% confidence
 - p value: -1.65, reject H0
 - confidence: 95%
 - intercept: 0, slope: 0.05
- Result: Significant

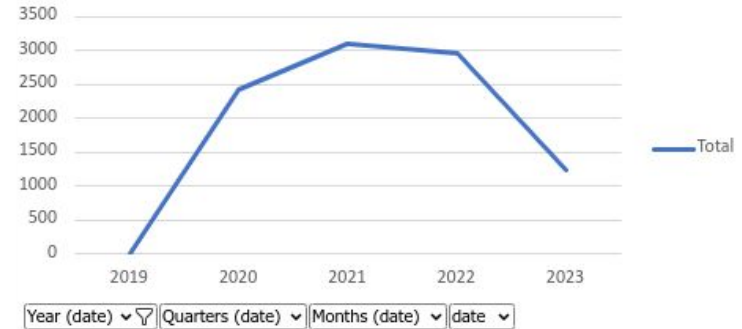
Regression analysis
App version vs review sentiment score



Conclusion and summary

- Data: IKEA app reviews
- No clear pattern emerged
- Analysis on review score (sentiment analysis) by app version
- There is a small but significant increase of the review score by app version

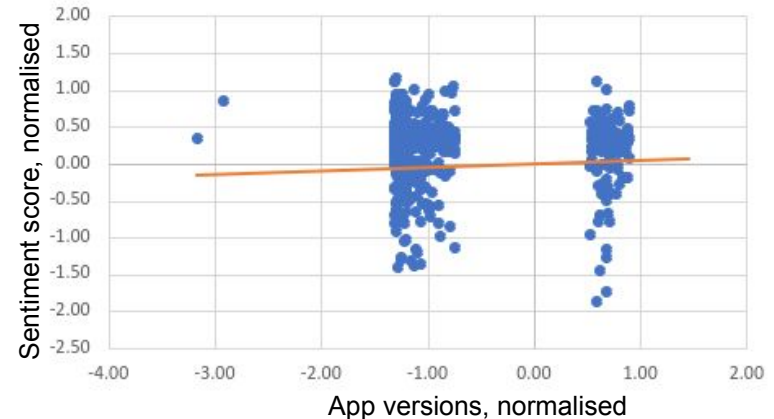
Reviews over year, full history



* partial data for 2023, still ongoing

Regression analysis

App version vs review sentiment score



Future directions

- Investigate what led to more positive user feedback in latest app versions.
- Apply sentiment analysis for topic discovery to evaluate how main review topics change over time.



Limitations

- Additional factors beyond version updates may also influence user sentiment towards the app.
- Due to time limitations, it was not possible to appropriately assess the homogeneity of the sample over time, a key factor when applying regression analysis.





ReDI School of Digital Integration

Data Analytics 2023

Assignment 2

Analysis on IKEA app reviews from Play Store

Thank you for your attention!

Questions?