



Are you a robot?

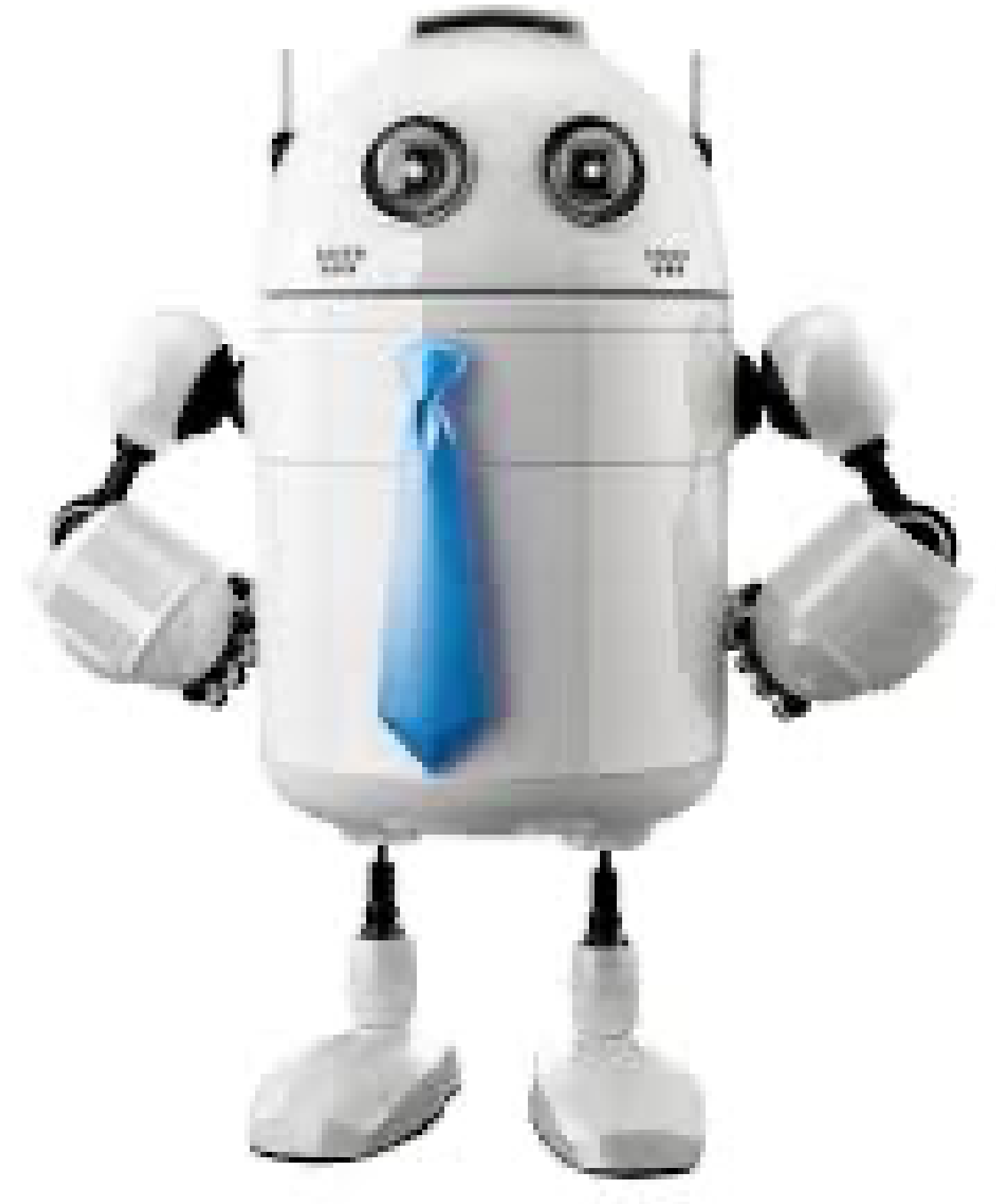


reCAPTCHA

[Privacy](#) - [Terms](#)

A Look Into Amazon Music Reviews

Hackathon Submission By Toliy A., Nidhi A., Rosania P.



Team Captcha



Rosania Pinto

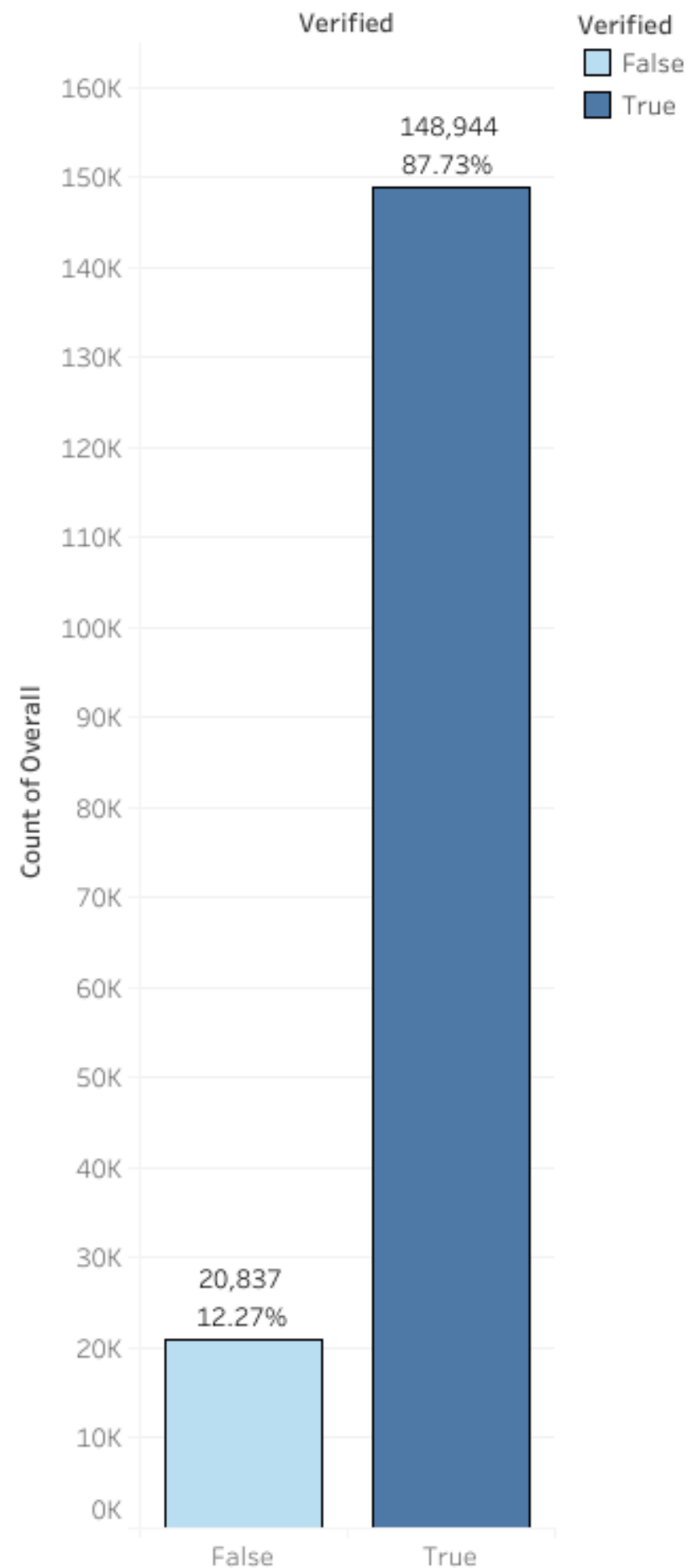


Toliy Anchakov



Nidhi Arora

Verified vs non-verified
purchases



Overview of digital music reviews

- 169,781 reviews
- 16,566 unique user ID's
- Average of 10.24 reviews per user ID
- 87.73% of reviews come from verified purchases
- **We are looking at the 12.27% of reviews with no verified purchases**

Criteria of Identifying Potential Bots



- Non verified purchase review
- Overall rating 5 stars
- Review text is 5 or less words

All reviews posted on
same day



Potential User Bot
Reviewer ID
AZ7IXJZ79RMCC

	overall	vote	verified	reviewTime	reviewerID	asin	style	reviewerName	reviewText	summary	unixReviewTime
7121	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B000VZWMZ0	{'Format': 'MP3 Music'}	Anessa	love	Five Stars	1462406400
24312	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B00137QMFG	NaN	Anessa	love	Five Stars	1462406400
25269	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B00137SPH4	{'Format': 'MP3 Music'}	Anessa	love	Five Stars	1462406400
30449	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B000VZWMZ0	{'Format': 'MP3 Music'}	Anessa	love	Five Stars	1462406400
47640	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B00137QMFG	NaN	Anessa	love	Five Stars	1462406400
48597	5	NaN	False	05 5, 2016	A1VLJIGQNNCTL5	B00137SPH4	{'Format': 'MP3 Music'}	Anessa	love	Five Stars	1462406400

Multiple products



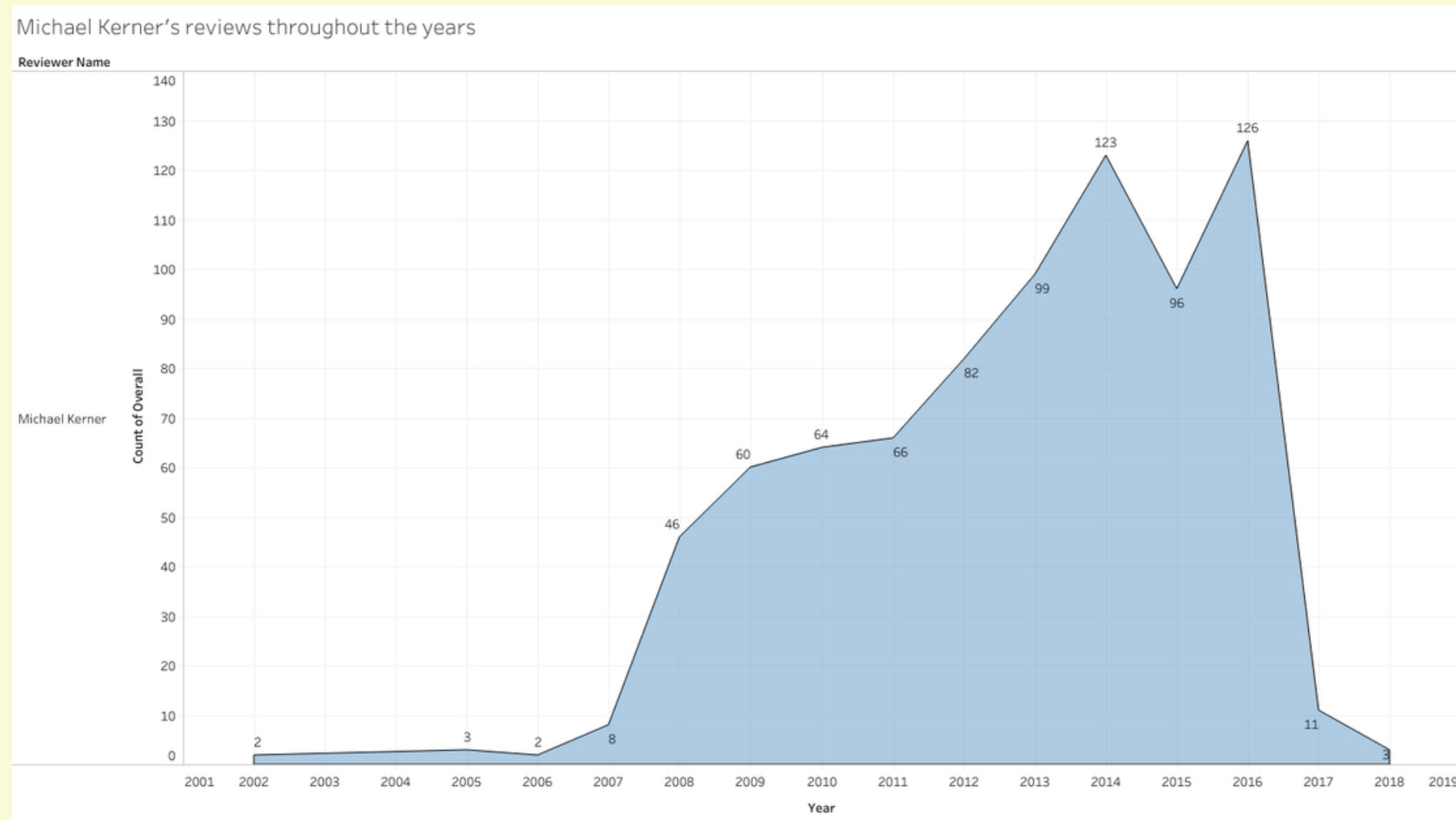
Same comment
for different products



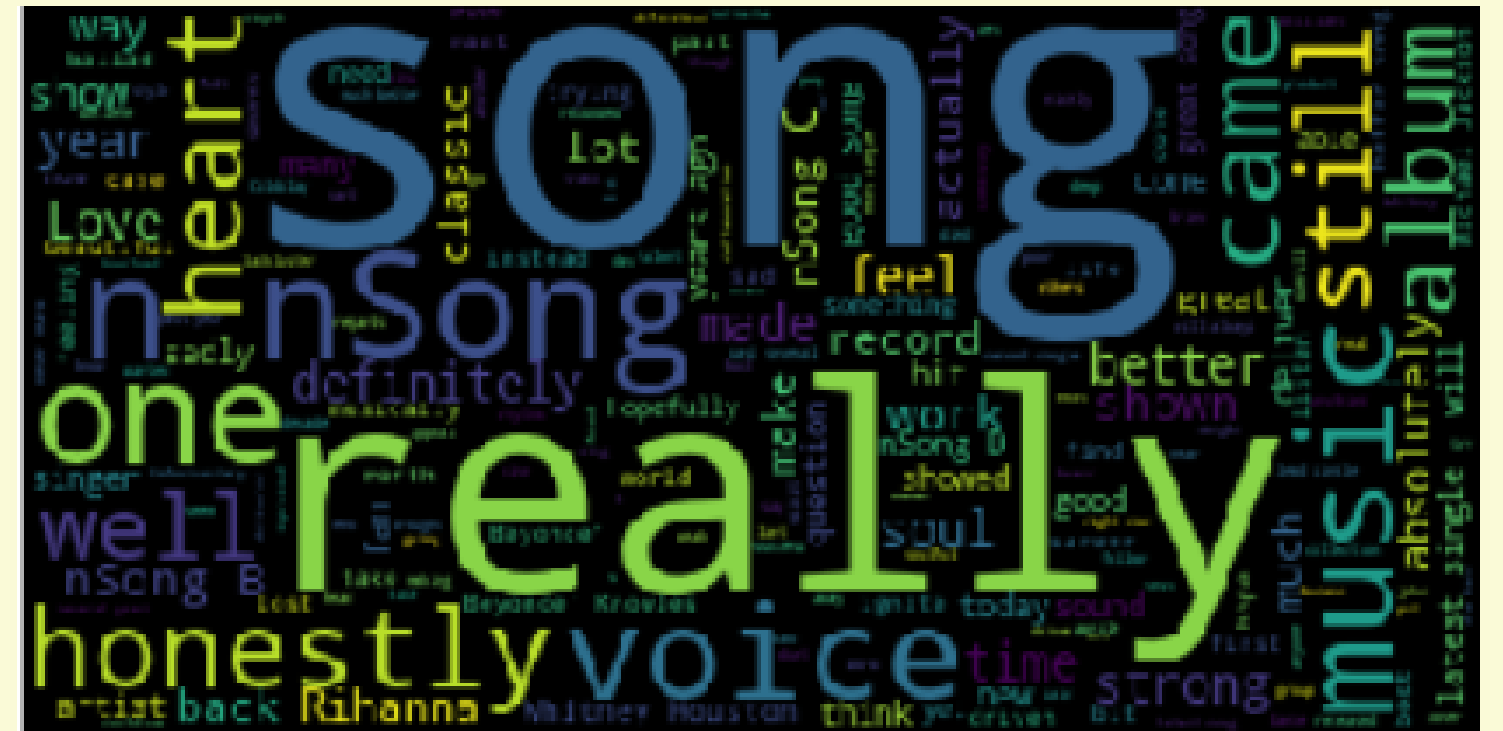
Other anomalies

User with the most reviews

Michael Kerner



- From his total of 791 reviews, 766 are non-verified purchases
- 134,000 words used in total for all his reviews
- Average of about 170 words per review
- Word Cloud of some of his most used words
- Has been leaving reviews since 2002
- Has peaked in 2016 with 126 reviews
- As of 2018, Mr. Kerner's activity has died down to a very suspiciously low level. Maybe Amazon should check in on him...



Conclusions and Next Steps

Project highlights

Identifying Bots and Spam:

- All unverified purchases
- Lack of word diversity in reviews
- Same time and date of reviews

Identifying Anomalies:

- Having a large number of reviews does not correlate with a user being a bot.
- Unverified purchases are also not enough to determine whether reviews are by a bot.
- Diverse posting times/dates, and diverse word use lets us determine who is not a bot.

Next Steps

Finding More Suspects:

- Are there more subversive methods used by bots and spammers?
- Can we use NLP (Natural Language Processing) to identify additional bots?
- Are there bots with verified purchases?

Lingering Questions:

- How do bots affect Amazon's bottom line?
- How do they affect the bottom line of musicians?
- Last but not least, where in the world is Michael Kerner?

Thank you, let's keep in touch!



Rosania Pinto

Email: RosaniaPinto@gmail.com
LinkedIn: linkedin.com/in/rosaniapinto/
GitHub: github.com/RosaniaPinto



Toliy Anchakov

Email: Toliy.Anchakov@gmail.com
LinkedIn: linkedin.com/in/toliy-anchakov/
GitHub: github.com/Tokey94



Nidhi Arora

Email: narora1183@gmail.com
LinkedIn: linkedin.com/in/narora1183
GitHub: github.com/narora1183

