

# THE GROUP CONVO

Analyzing the Facebook Group Conversation from July 1st 2016 to July 12th 2018.

## WHO'S COUNTED

Rogges Anandarajah, Saad Muhammed, Ali Javed, Kishan Baskaran, Aushvin Vasanthakumaran, Ali Nawed, Thannoj Thavalingam, Janusshan Paramasivam, Hamza Ahsan, Harris Ismael

Sometimes: Diahanna, Vinoth, Joey Bot, and Farhin



## MILDLY INTERESTING FACT

Farhin's only message to the group: Farhin Chowdhury has left the group.

## VERBOSITY

At an average of 49.8 characters per message, Kishan is the most verbose of us all.

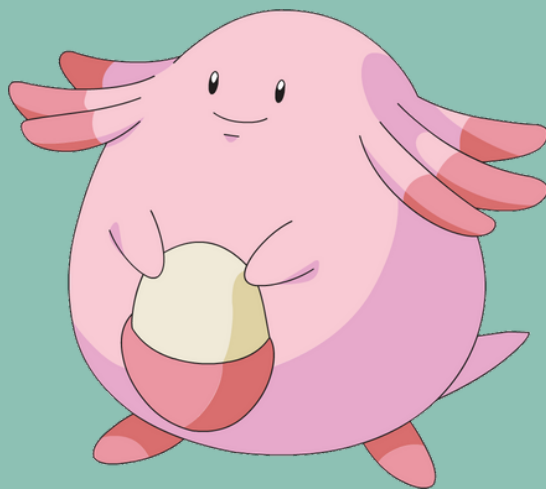


## MR. NEGATIVE

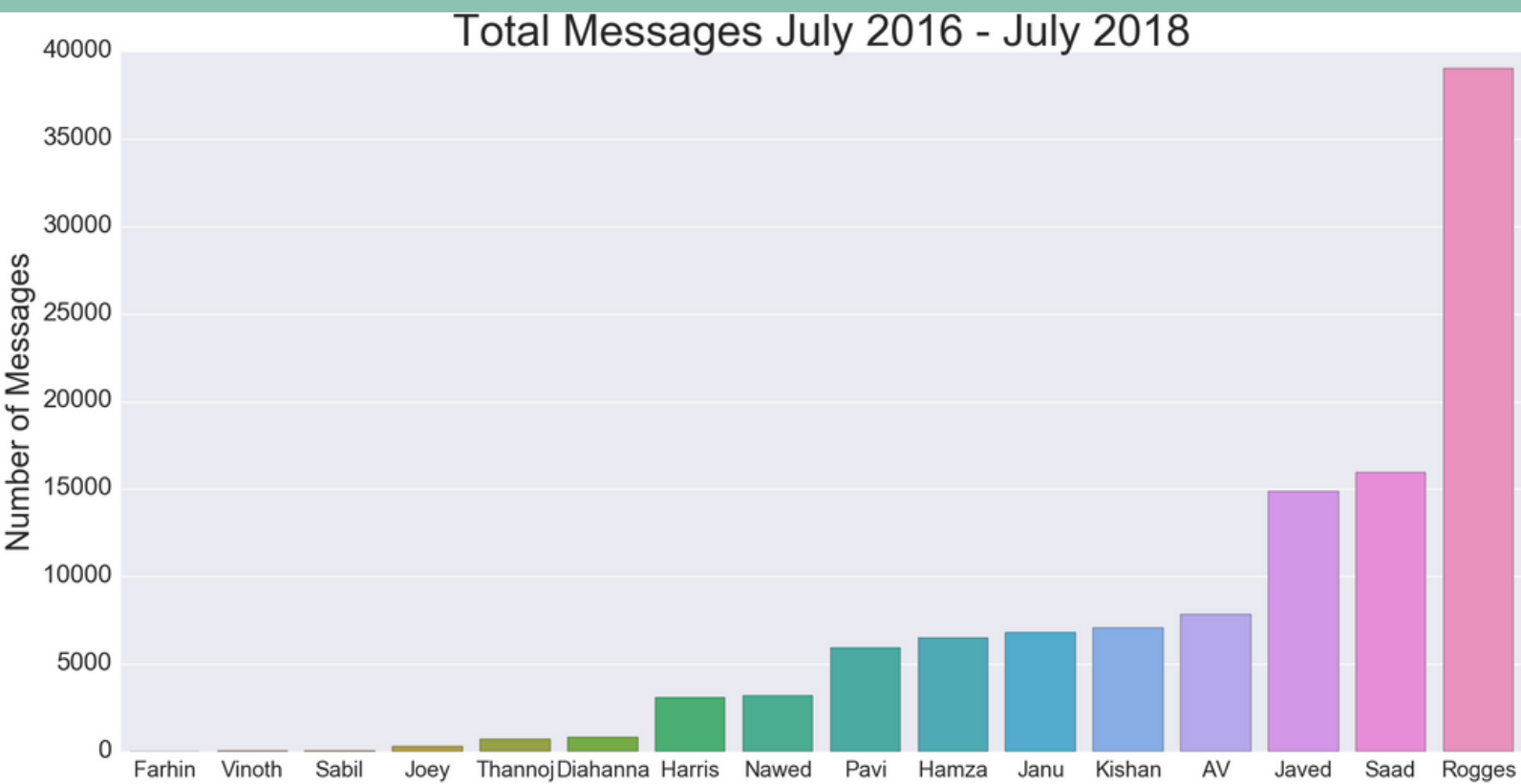
With the only negative average on comment polarity, Janusshan is the most toxic member of the group.

## NOT THANNOJ

The highest, positive average on comment polarity belongs to one Harris Ismael. That's right kids, Thannoj brings the best kind of toxicity, relative.



# Quantity of Messages



**ABOVE**

A simple bar chart of the amount of messages each member contributed

Rogges, at number 1, has more messages than the 2nd, 3rd, and 4th most messages combined.



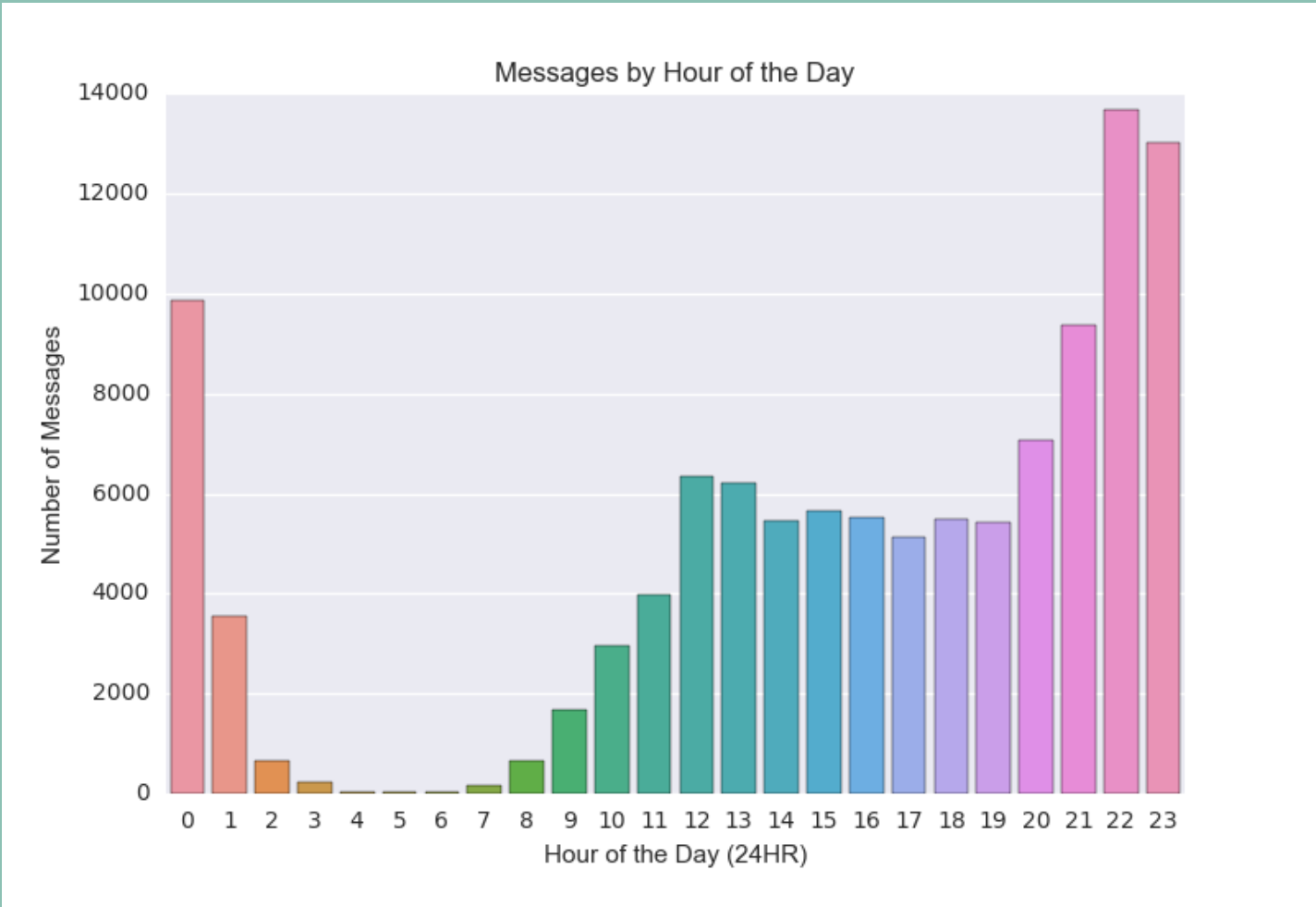
Javed, Saad, and Rogges combined account for 62.2% of all messages.

A raw data table with the precise message totals per member for interested parties.

**BELOW**

	Number of Messages
Rogges Anandarajah	39040
Saad Daas	15956
Ali Javed	14864
Aushvin Vasanthakumaran	7862
Kishan Baskaran	7088
Janusshan Paramasivam	6792
Hamza Ahsan	6514
Pavi Subenderan	5947
Ali Nawed	3184
Harris Ismael	3087
Diahanna Rose	831
Thannoj Thavalingam	710
Joey Bot	307
Sabil Ahmed	71
Vinoth Maruthalingam	26
Farhin Chowdhury	1

# When to Message



## ABOVE

A bar chart distribution with the number of messages by hour of day. 0 HR is Midnight, and 23 HR is 11PM.

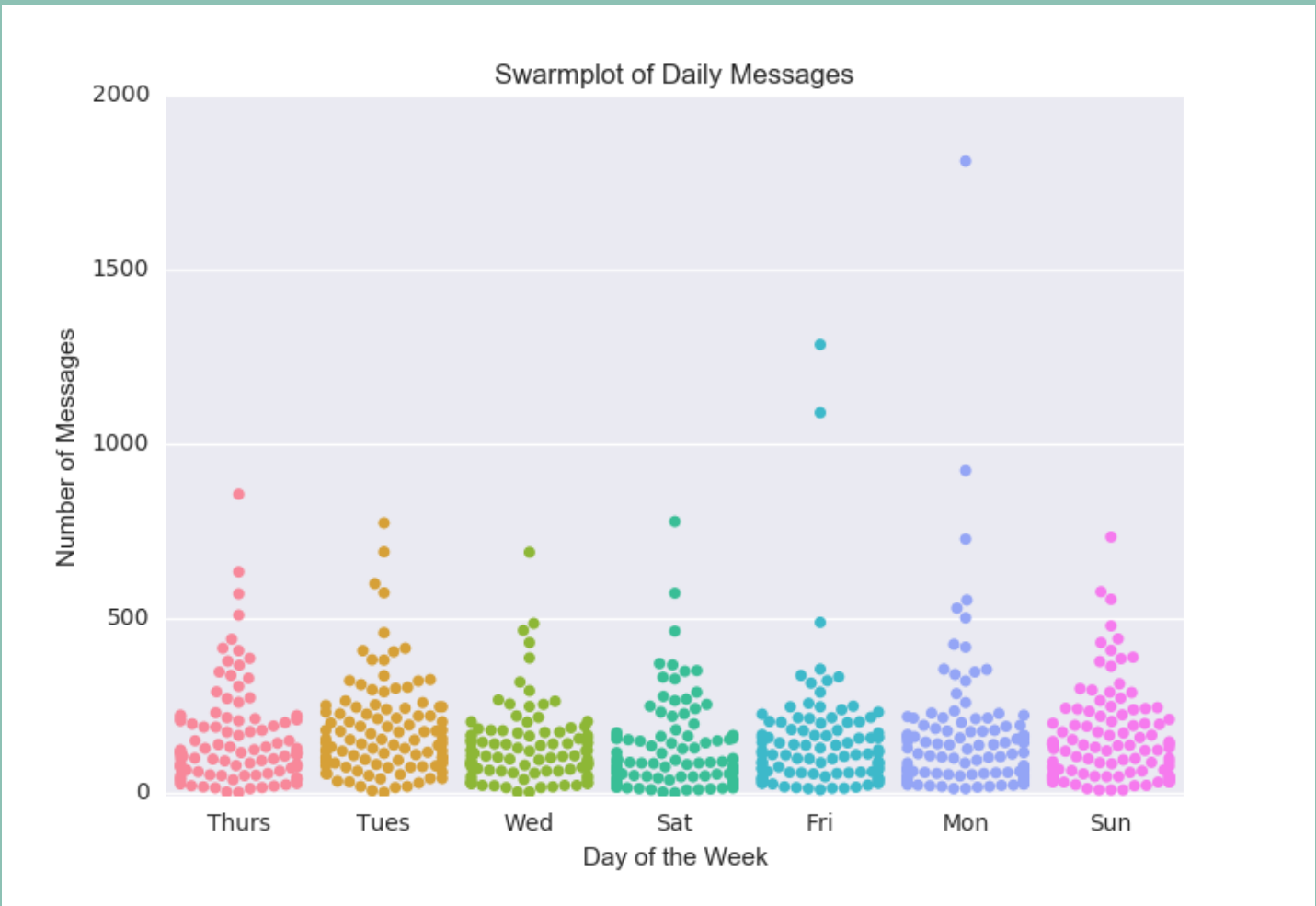
Don't message between 2AM and 11AM. Activity is depressed at these times.



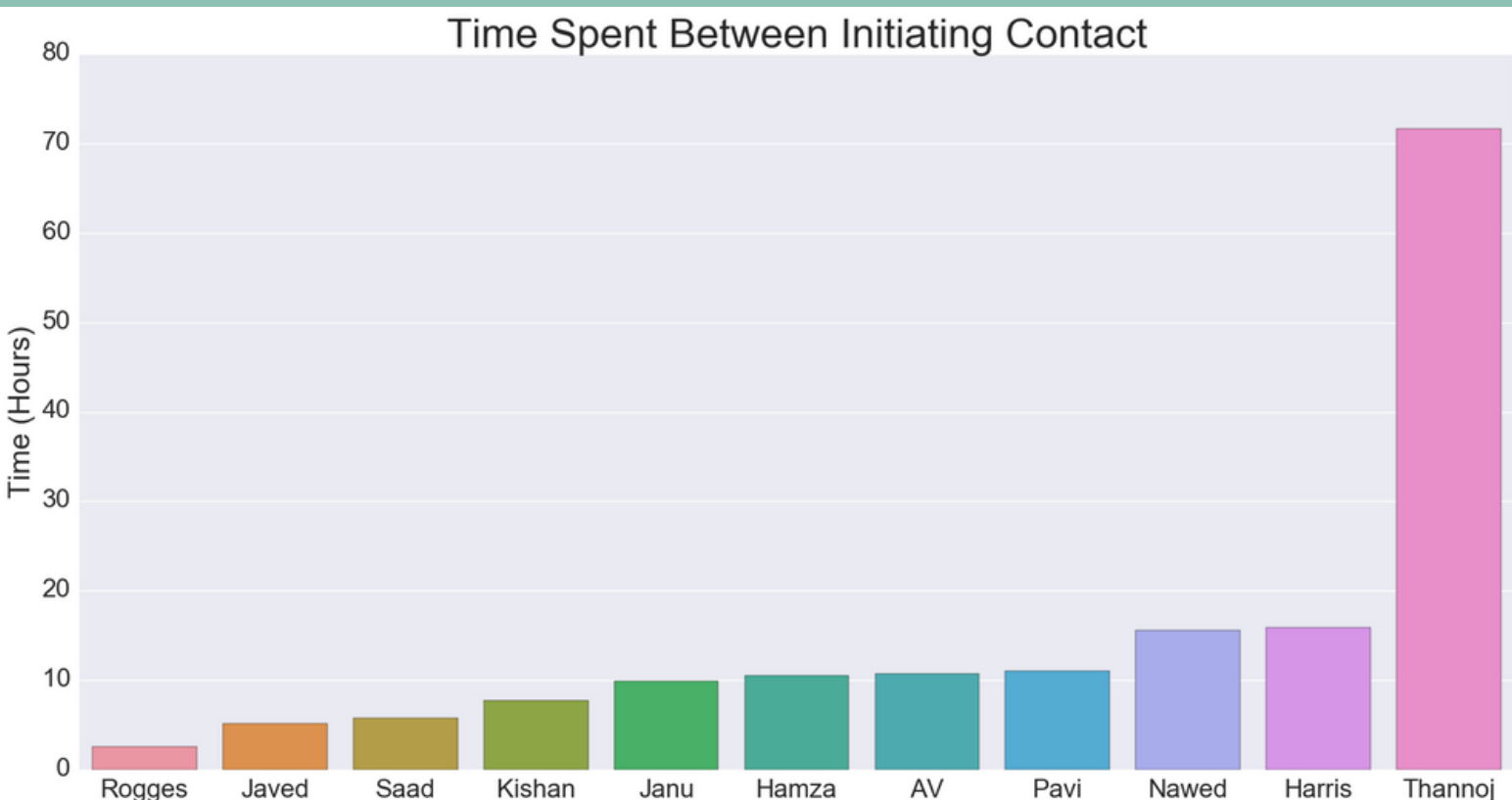
Despite decent activity, Saturdays and Wednesdays have 21% and 15% lower daily message volume respectively.

Swarmplots plot every single value by category. E.g. For Sunday below, the number of messages on every Sunday is plotted.

## BELOW



# Who to Message



## ABOVE

Measures how long it takes for a group member to initiate contact (excludes participation in conversations by other members).

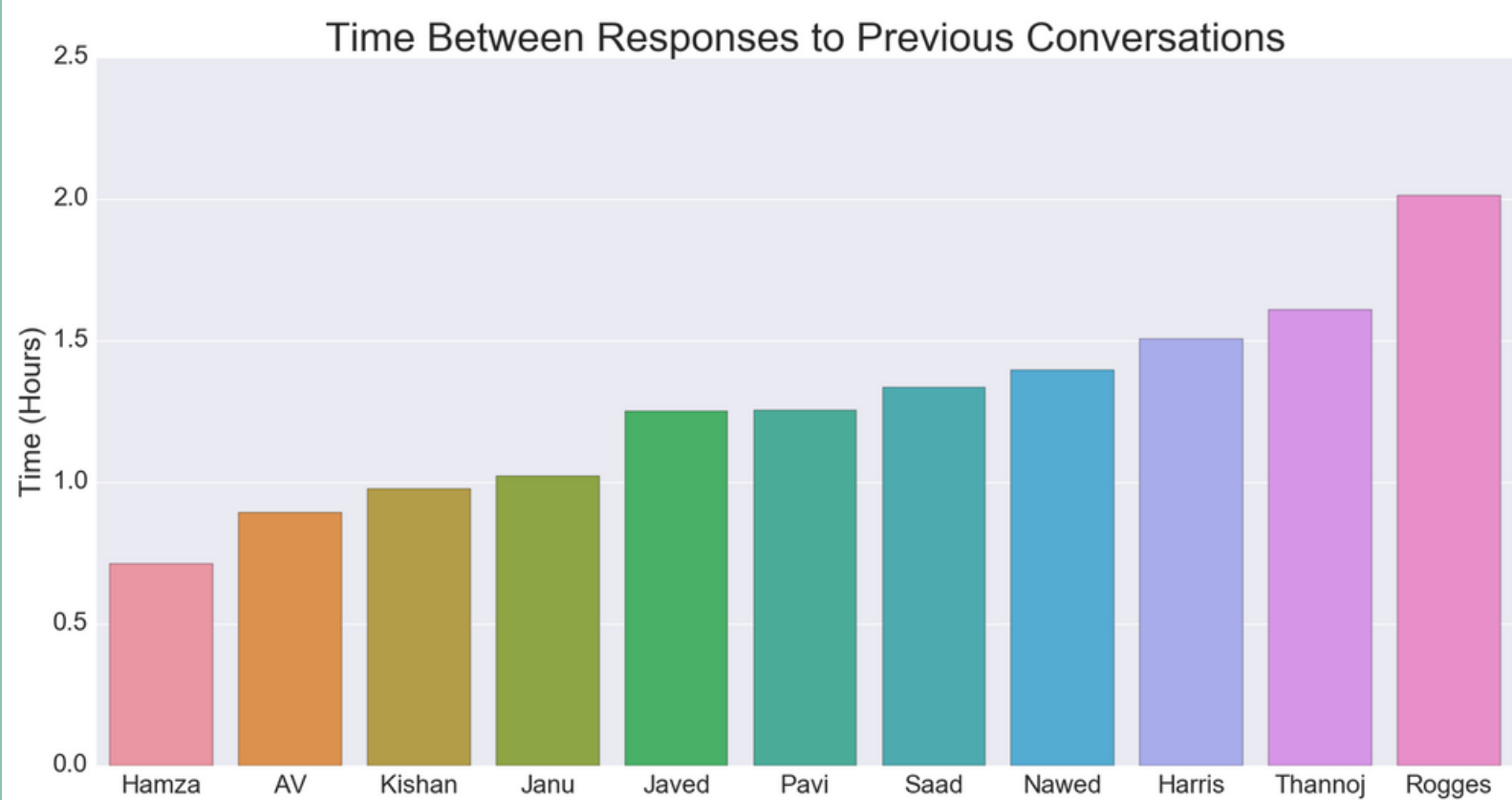
If a conversation hasn't taken place in a while; message, or wait for: Rogges, Javed, or Saad. They'll likely be on in: 2hrs, 5hrs, or 5.5hrs respectively.



If a conversation just ended a few minutes ago; message, or wait for: Hamza, AV, or Kishan. They usually follow up within an hour.

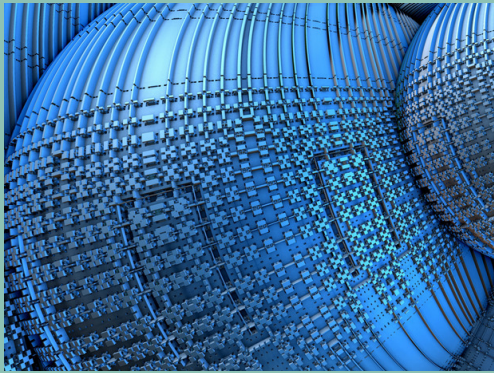
Measures follow up to a concluded conversation. Conversation must've concluded 3 mins or more in the past to qualify.

## BELOW





# Authors / Concluding Remarks



## Data Acquisition

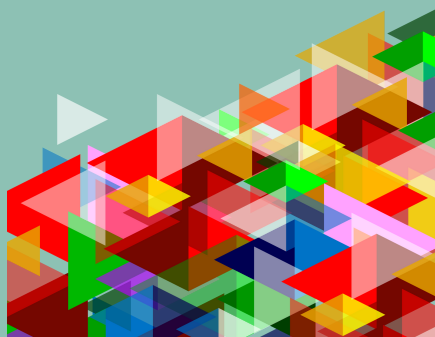
Pavi Subenderan



## Coding / Data Wrangling

Rogges Anandarajah

Pavi Subenderan



## Graphs and Infographic

Rogges Anandarajah

### Tools

Python 3.5.x on an Anaconda Distro

Jupyter Notebooks

Canva.com - Infographic



### Packages

JSON - Raw data

Pandas - Basic dataframe object

Numpy - Arrays and operations

Re - Regular expressions

NLTK - Word stemming and tokenizing, other NLP purposes

PyContractions - Expanding contractions in speech

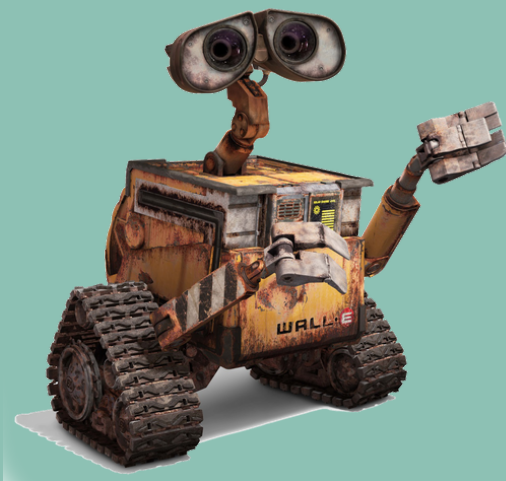
Matplotlib - Graphics engine

Seaborn - Graphics package



## Potential Future Direction

Training a Chat-bot on a single user's messages to try and recreate their speech.



## Potential Future Direction

Using supervised and unsupervised Machine Learning models for Topic Extraction. Allowing us to classify messages by topic, e.g. Sports, Music

