Overview

Our Client

Business Problem & Solution Overview

Data Collection

Data Analysis

Valuable Insights

Challenges & Limitations

Our Client

RESTROOM ASSOCIATION (SINGAPORE)

Non-Profit Organization, Works closely with NEA and MOH

Advocates for clean toilets in Singapore

Launched the Happy Toilet Programme in 2003

Our Client

Happy Toilet Programme

- Star-grading initiative
- Toilet owners can apply to be assessed by RAS
- Awarded with a certification with their star rating, one year validity

Business Problem

Part of the mission of Restroom Association Singapore is to 'Lobby the cleaning industry to meet the right standards'

With around 1150 toilets to look out for, it can be **challenging to know which toilet to focus lobbying attention on**.

Business Problem

Toilet owners registered under the **Happy Toilets Programme** schedule an audit by Restroom Association. The results of these audits are valid for **a year** afterwards.

However, there is a **likelihood that the toilets won't remain clean** after the audit is completed.

Business Solution

Our solution analyzes social media mentions of various toilets in Singapore, and integrates these sentiments and comments into the current Happy Toilet Programme.

This will allow for a more **dynamic review** of the toilets, as well as for concrete comments on the toilets to be reflected so that RAS can take action according to the feedback.
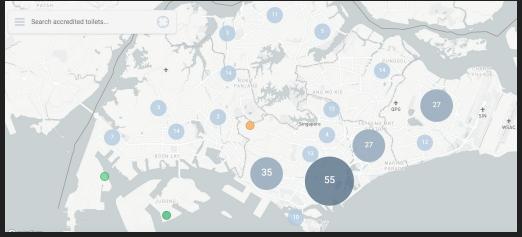
Business Solution

"(It will) definitely (be) useful! It would be even better if it could blend into our Happy Toilet Programme"
- Mr Loh, RAS

About our Client

Business Problem & Solution

Data Collection

Data Analysis

Valuable Insights

Challenges & Limitations

Project Objectives

1. What are the **sentiments on toilets** in Singapore on Social Media for a specified period of time?

2. Where are the **clusters of dirty/clean toilets**?

3. What are some **concerns/trends** of toilets in Singapore?

4. What are some **factors** that makes using the toilet a **pleasant experience**?

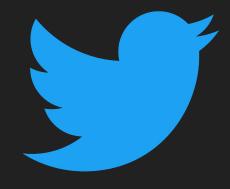5. Which toilets should be chosen for **"surprise" audit checks**?
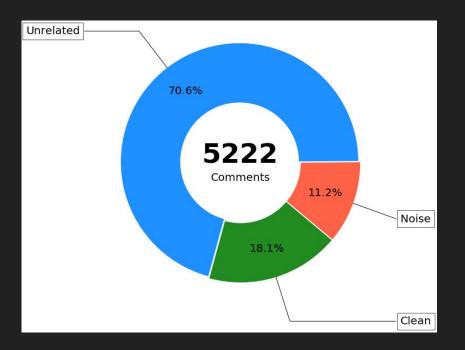
Data Collection

Data Collection - Reddit

- Collected using their API
- Search terms used "toilet" on Singapore Subreddit
- After Data Cleaning:
  - **11.2%** of data are noise (advertisements, direct quotes to a prior comment, or bots),
  - **70.6%** of data are unrelated (comments that do not contain keywords such as "toilets", "urinal", "cubicle", "flush", "poop", "urine", and more).
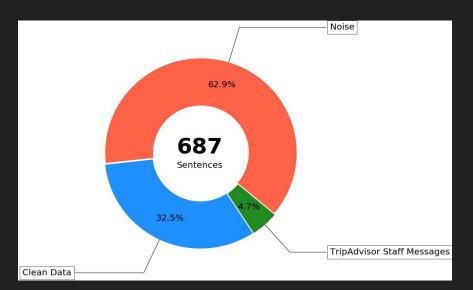 - Only **18.1%** of data is clean and relevant.

Data Collection - TripAdvisor

- Collected using Selenium
- Search terms used "toilet" on Singapore
- After Data Cleaning:
  - **62.9%** of data ar noise (advertisements, greetings and salutations, URL),
  - **4.7%** of data are TripAdvisor Staff Messages
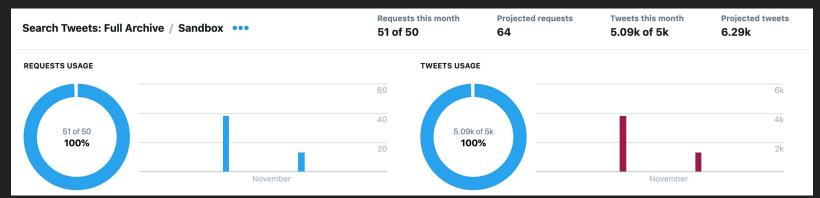 - Only **32.5%** of data is clean and relevant.

Data Collection - Twitter

- Collected using the sandbox version of Twitter Premium API
- Keyword: "Singapore Toilet"
- Time Span: 2017 - 2019
- Python package - "TwitterAPI" is used to connect to the endpoints

Search Tweets: Full Archive / Sandbox •••

Requests this month	Projected requests	Tweets this month	Projected tweets
51 of 50	**64**	**5.09k of 5k**	**6.29k**

REQUESTS USAGE

51 of 50
100%

60
40
20
November

TWEETS USAGE

5.09k of 5k
100%

6k
4k
2k
November

Twitter Developers Dashboard

About our Client | Business Problem & Solution | Data Collection | Data Analysis | Valuable Insights | Challenges & Limitations
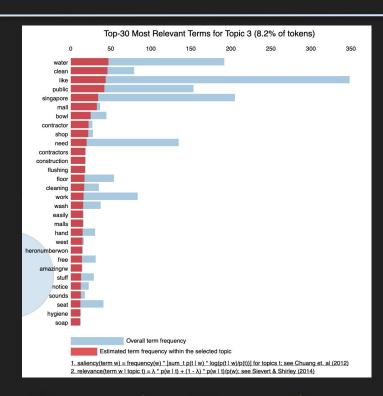
Data Analysis - Reddit Topic Modelling

Top-30 Most Relevant Terms for Topic 3 (8.2% of tokens)

Relevant Comments under Topic "Water" and "Clean"

"Actually, i don't rate precisely because i just washed my hands and i don't want to touch that screen because others don't wash their hands after pooing and peeing…" (opinion on touch-screen toilet review boards)

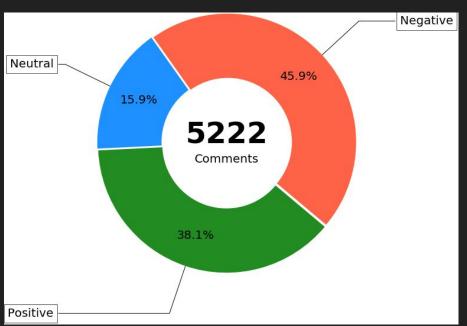"Usually, the MRT ones tend to be dirtier than the average mall toilet. Not sure who is the most common culprit."

Data Analysis - Reddit Sentiment

5222 Comments

- Negative: 45.9%
- Positive: 38.1%
- Neutral: 15.9%

| Strong Active Sentences Give the Best Accuracy	
"Yep. I always rate the highest whenever i'm at the airport. Their toilets are so clean"	Positive, +0.37
"If my legs ever touch a public toilet seat, I'm cutting is off. "	Negative, -0.21
"Disgusting behaviour"	Negative, -1.0
"Menstrual blood all over the toilet seat. My friend discovered it at work some more, so it was definitely the work of an adult woman seriously wtf."	Negative, -0.11
"CCK MRT toilet. The bowl was up to the brim with liquid shit. Wondered how they cleaned that one."	Negative, -0.12
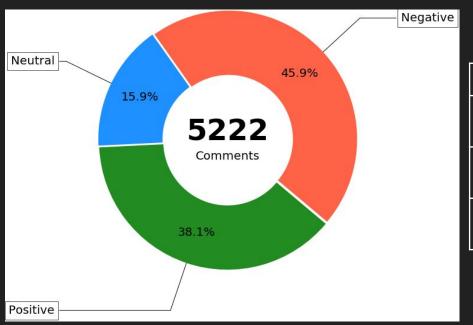
Data Analysis - Reddit Sentiment

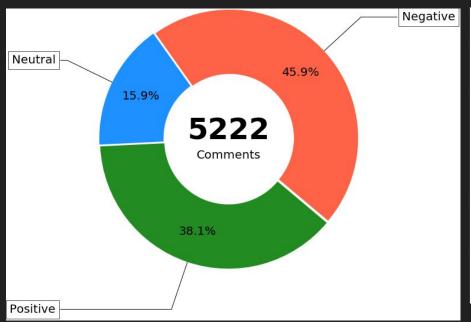| Susceptible to False Positive due a Lack of Context	
"Same in this toilet. No brushes or anything provided to clean.	Positive, +0.17
"well you are using PUBLIC toilet.. don't expect much."	Positive, +0.06
"where can i find a 6 star toilet ??"	Negative, -0.03

Data Analysis - Reddit Sentiment

Donut chart showing sentiment distribution of 5222 Comments: Negative 45.9%, Positive 38.1%, Neutral 15.9%.

| Weak to Sarcasm	
"Cleaner than your intentions."	Neutral, +0.0
"Spreading singapore culture to the natives. Good job. If Peeping Tom was an Olympic sport we'd own the entire podium,then celebrate by filming the female athlete in their own showers."	Positive, +0.23
Location detection relies on spelling accuracy of commenter	
"Went in no toilet paper, realised no flush. Bras Bersah."	Negative, -0.03

Data Analysis - TripAdvisor Sentiment

Neutral — 96.9%

223 Sentences

2.7%

0.4%

Negative
Positive
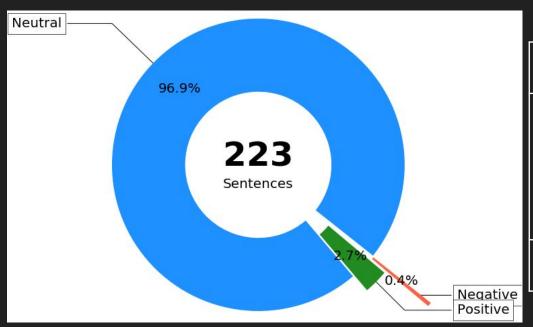
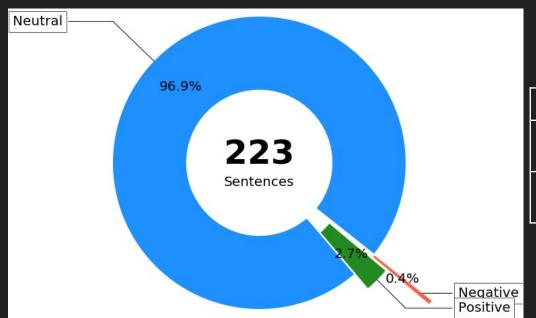| Positive Sentiment Comments do provide good feedback	
"Yes, it's a good idea to get used to toilet paper. All toilets here are dry toilets. Having said that, for most toilets, the last cubicle, normally, a squat toilet, will have a hand spray or tap. Alternatively you may want to bring a disposable water bottle with water to the toilet. Hope this helps."	Positive, +2.0
"the ones at newton circus are very'zen' and lovely:)"	Positive, +3.0

About our Client | Business Problem & Solution | Data Collection | Data Analysis | Valuable Insights | Challenges & Limitations

Data Analysis - TripAdvisor Sentiment

Neutral — 96.9%

223 Sentences

2.7%

0.4%

Negative
Positive

False Positives are aplenty due to sarcasm:	
"Anyway, I haven't seen a 'dry toilet' in ages."	Neutral, +0.0
"LOL WTF IS THIS SHIT"	Positive, +0.03

About our Client   Business Problem & Solution   Data Collection   Data Analysis   Valuable Insights   Challenges & Limitations
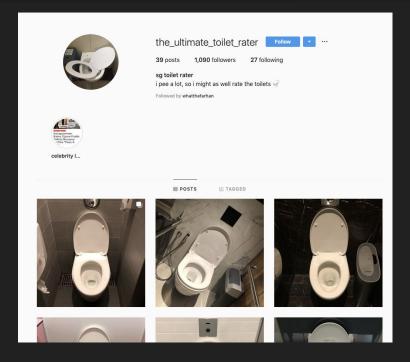
Analysis from Instagram Data Set

Analysis from Instagram Data Set

- Found some potential hashtags to scrape on Instagram
  - #toiletsg
  - #sgtoilet

Analysis from Instagram Data Set

Data Analysis - Twitter Wordmap

Posts on Twitter are more about legal issues:

"Malaysian man jailed, fined after he secretly filmed men in toilet cubicles"

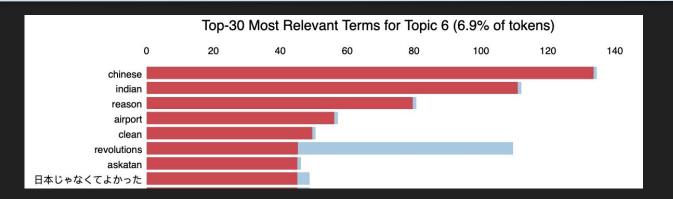"IT consultant jailed for filming men urinating in poly toilet"

Data Analysis - Twitter Topic Modelling

Top-30 Most Relevant Terms for Topic 6 (6.9% of tokens)

chinese	
indian	
reason	
airport	
clean	
revolutions	
askatan	
日本じゃなくてよかった	

Tourists Enjoyed the toilets at the Changi Airport:

> "Singapore's epic toilet flush!
> #Jewel #Singapore @ Jewel Changi Airport"

News Article about Chinese and Indian toilet:

> "Chinese and Indian toilet revolutions look to
> Singapore's bottom line"

About our Client | Business Problem & Solution | Data Collection | Data Analysis | Valuable Insights | Challenges & Limitations
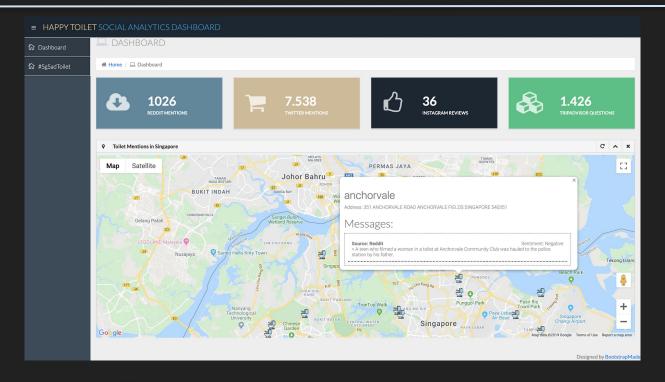
Valuable Insights - Integrated Public Toilet Rating System

Valuable Insights - Social Media Campaign

HELLO THERE!

Please leave a Tweet/Instagram post to let me know how I've made you feel! Hashtag #SgHappyToilet if you felt happy or #SgSadToilet if you didn't like me

With your feedback, we can take action and make every toilet a Happy Toilet!

HAPPY TOILET

We recommend a campaign that encourages users of the toilet to leave a tweet or Instagram post about their sentiments with a specified hashtag

This will help us to collect more relevant data in order to provide better insights on public sentiments in the future

Valuable Insights - RAS Toilet Grading Rubric

We suggest adding ply of tissue paper to the grading rubrics

Valuable Insights - Toilet Grading Rubric

The availability of bidets really differentiates a high-end
toilet, from the standard ones.
Hence, we suggest adding this in the rubric as well.

Challenges & Limitations

Max scrape 5k per month per application

Locked down for data retrieval

API requires $$

Lack of wordlist for location Misspellings may affect results

Thank you for listening!