



Sentiment Analysis for Canada Post Reviews

Google Reviews & Reddit & TrustPilot

ABHISHEK SHARDA

ANTRIKSHA GOSWAMI

DAVID AGUDELO

MAROINE BEN LTAIEF

YUE MA

Problem Statement

Problem statement

The problem we aim to address in this data science project is to perform sentiment analysis on customer reviews for Canada Post. By analyzing a diverse set of reviews from Google, Reddit, and Trust Pilot, we seek to gain valuable insights into the perceptions of customers towards each courier service.

Problem Statement

Scope

- 1.Data Collection:** Collected sample dataset of customer reviews from Google reviews, Reddit, and Trust Pilot.
- 2.Sentiment Analysis:** Utilized NLP techniques and LDA and SIA, and classified the sentiment of each review as positive, negative.
- 3.Cross-Platform Comparison:** By comparing sentiment scores, identified variations in customer feedback and assessed the impact of platform-specific factors on customer perceptions.
- 4.Thematic Insights:** Went beyond sentiment and extracted common themes and topics mentioned in positive and negative reviews for each company (service, convenience, damages and waiting times).
- 5.Actionable Recommendations:** Based on findings, provided actionable recommendations to Canada Post, for enhancing customer satisfaction and optimize their courier services.

Data Collection & Understanding

DATA SOURCE

- Customers' reviews about Canada Post were scraped from 3 major sources, Google reviews, Reddit and Trust Pilot.
- To cross-compare and help Canada Post identify their differences, customers' reviews data towards UPS and Purolator were also scraped.

DATA UNDERSTANDING

- Index and reviews
- Google Reviews contain other variables such as user names, date, number of reviews and locations.

	User Name	Number of Reviews by User	RfnDt 2	Time When reviewed	Reviews
0	Anita E	32 reviews	Local Guide	4 months ago	Great customer service, never had an issue here. Barbara is super helpful and I'm always confident my packages will be dispatched efficiently here. Thank you!
1	Dakkar (Rabanne)	11 reviews	NaN	a month ago	The woman that works there in the afternoons is rude and has an attitude if you ask for a receipt. I wish they'd replace her
2	Joan Beverley	22 reviews	NaN	4 months ago	Great service fast and the staff always polite. Barbara thank you for your professionalism.

Tech Used For "Scraping"



reddit "API"

```
import praw

user_agent = "Scraper 1.0 by/u/Cool-Dog-3907"
reddit = praw.Reddit(
    client_id="3WCxCK3_hxLOIMcpKzc2RA",
    client_secret="Bd1B0lwj-jXkUvaZ4pLG6o_RQ3rxw",
    user_agent=user_agent
)

Version 7.7.0 of praw is outdated. Version 7.7.1 was released Tuesday July 11, 2023.

reviews = set()
for submission in reddit.subreddit('canadapost').hot(limit=None):
    reviews.add(submission.title)
print(len(reviews))

995

df = pd.DataFrame(reviews)
df = pd.DataFrame(reviews, columns=["reviews"])
df.to_csv('Reviews_reddit_Canada_Post.csv', header=False, encoding="utf-8", index=False)
df.head()
```

	reviews
0	Anyone else get opened mail in their mailboxes?
1	What does this even mean?!
2	Stamp Question
3	Taxi cab delivering mail bags
4	Help with shipping a package to America full o...



INSTANT DATA SCRAPER

- It's an automated tool to scrape data on any website.
- Scraped google reviews for Canada Post, Purolator and UPS.
- Targeted different store locations across major cities like Toronto and Ottawa.

href	Mapped href	A503be	dehyf	OWOW	Jlu8Td	QV3IV
local/review/rp/report?https://www.google.com/maps/contrib/10339704	15 reviews 12 photos	3 years ago	4	Like		Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/10146795	9 reviews 1 photo	6 years ago	2	Like	The same day my flight was, was the day my pe	
local/review/rp/report?https://www.google.com/maps/contrib/10819916	9 reviews	3 years ago	1	Like		
local/review/rp/report?https://www.google.com/maps/contrib/10450007	183 reviews 197 photos	a year ago	1	Like	Can't believe I'm reviewing a post office, but serv	Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/11021421	1 review	a year ago	0	Like	Very courteous and helpful staff. Kudos to Mr K	
local/review/rp/report?https://www.google.com/maps/contrib/11115347	17 reviews 5 photos	11 months ago	0	Like	this place needs an audit, standing in line forev	Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/10392965	7 reviews	10 months ago	0	Like	Came first thing in the morning. Not open even l	
local/review/rp/report?https://www.google.com/maps/contrib/10056764	5 reviews	2 years ago	2	Like	Honestly, very rude person. Not helpful at all. Sc	
local/review/rp/report?https://www.google.com/maps/contrib/11564043	11 reviews 3 photos	4 years ago	2	Like	Rude and snooty clerk. Kept on correcting me b	Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/10780061	435 reviews 24 photos	6 years ago	1	Like	Decent post office. Best location and service.	Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/11821522	465 reviews 16 photos	3 years ago	1	Like	The staff is always friendly and very helpful	Local Guide
local/review/rp/report?https://www.google.com/maps/contrib/10999666	14 reviews 2 photos	6 years ago	3	Like	not open sunday	
local/review/rp/report?https://www.google.com/maps/contrib/11455801	50 reviews 61 photos	7 years ago	1	Like	canada post 5/5	Local Guide



Selenium WebDriver

Tech Used For “Scraping”

```
"""
Created on Sun Jun 11 12:54:58 2023

Scraping and saving the data from trustpilot to a csv

@author: Maroine
"""

# Import Library
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.chrome.options import Options
import pandas as pd

# Set Chrome options to run in headless mode
chrome_options = Options()
chrome_options.add_argument("--headless")

# Chrome Path
driver = webdriver.Chrome(options = chrome_options)
df=pd.DataFrame()

# Get Text
for i in range(2):
    # Web URL
    url=f'https://ca.trustpilot.com/review/www.canadapost.ca?page={i+1}&sort=recency'
    driver.get(url)

    element = driver.find_elements(By.TAG_NAME,'p')

    temp=[]

    for i in range (len(element)):
        temp.append([element[i].text])

    aux=pd.DataFrame(temp)
    df=df.append(aux,ignore_index=True)

df.to_csv('reviews.csv',index=False)

# Close the window
driver.close()
```

0	1.2
1	NaN
2	Visit this website
3	5,389 total
4	5-star
5	2%
6	4-star
7	<1%
8	3-star
9	<1%
10	2-star
11	2%
12	1-star
13	95%
14	Filter
15	Most relevant
16	Signed up for MyMail just now. I had to use a different e-mail address because the recovery process for my long-forgotten e-Post username and password crashed. Created new profile under a different e-mail only to discover MyMail is not available for my address. COMPLETE CANADA POST APP-TESTING FAIL!!!
17	Date of experience: June 08, 2023
18	Sent an iPhone expresspost to Philippines from Ontario. Came back 14 days later with sticker 🚫 dangerous goods. I told the guy at the counter it was a used iPhone in original box, and he does not know it is classed as dangerous goods? If he does not know, how should I know? 🙄
19	Date of experience: June 08, 2023

Library

LIBRARIES USED IN THIS PROJECT

- **Pandas**
- **Seaborn**
- **matplotlib.pyplot**
- **Re** for removing special characters in reviews
- **nltk**
- **Stopwords** (nltk.corpus) for identifying stopwords
- **SentimentIntensityAnalyzer** (nltk.sentiment) for performing sentiment analysis on text
- **TfidfVectorizer** (sklearn.feature_extraction.text) for converting text documents into numerical feature vectors
- **LatentDirichletAllocation** (sklearn.decomposition) for Topic Modelling

Data Cleaning & Feature Engineering

DATA CLEANING

- Removed the rows that contains NULL values

```
df = pd.read_csv("CanadaPostGoogleReviews.csv")
df.head(3)
```

	User Name	Number of Reviews by User	RtnDt 2	Time When reviewed	Reviews
0	Anita E	32 reviews	Local Guide	4 months ago	Great customer service, never had an issue here. Barbara is super helpful and I'm always confident my packages will be dispatched efficiently here. Thank you!
1	Dakkar (Rabanne)	11 reviews	NaN	a month ago	The woman that works there in the afternoons is rude and has an attitude if you ask for a receipt. I wish they'd replace her
2	Joan Beverley	22 reviews	NaN	4 months ago	Great service fast and the staff alway polite. Barbara thank you for your professionalism.



```
# Load the dataset
df = pd.read_csv("CanadaPostGoogleReviews.csv")
df["Source"] = "Google reviews"
df = df[["Reviews", "Source"]]
df = df.dropna()
df.head(3)
```

	Reviews	Source
0	Great customer service, never had an issue here. Barbara is super helpful and I'm always confident my packages will be dispatched efficiently here. Thank you!	Google reviews
1	The woman that works there in the afternoons is rude and has an attitude if you ask for a receipt. I wish they'd replace her	Google reviews
2	Great service fast and the staff alway polite. Barbara thank you for your professionalism.	Google reviews

FEATURE ENGINEERING

- Removed variables like 'user name', 'number of reviews', 'time of review' and 'locations';
- Added variable 'Source' (data source) before joining all the datasets;
- Lowercased text and removed special characters;
- Created new variable called 'Sentiment' and applied TF-IDF vectorization into this variable

SENTIMENT ANALYSIS

- `SentimentIntensityAnalyzer()` to create an instance;
- `lambda x: sia.polarity_scores(x)['compound']` to return sentiment score;
- `vectorizer.fit_transform()` to convert the raw text reviews into numerical vectors;
- `LatentDirichletAllocation()` to perform topic modelling. In this project, we created 3 topics.

	Reviews	Source	Sentiment	Topic
0	great customer service never had an issue here barbara is super helpful and im always confident my packages will be dispatched efficiently here thank you	Google reviews	0.9595	1
1	the woman that works there in the afternoons is rude and has an attitude if you ask for a receipt i wish theyd replace her	Google reviews	-0.0772	2
2	great service fast and the staff alway polite barbara thank you for your professionalism	Google reviews	0.7650	1
3	horrible horrible horrible service constantly they have lost my packages on a number of occasions once they even gave me someone elses package when going to pick up my own i didnt notice until i was already back home and opening it	Google reviews	-0.9100	1
4	conveniently located in college squares rexall i had no idea it was there until i searched for a post office in the area a newer sign would draw more attention to it the service was oood i did have to wait quite a while like most	Google reviews	0.5390	0

Sentiment Analysis

COMPLAINT & CUSTOMER SERVICE ANALYSIS

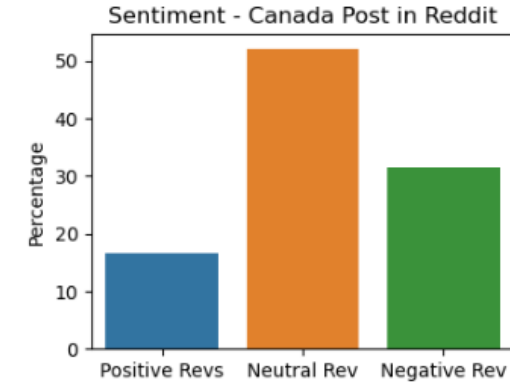
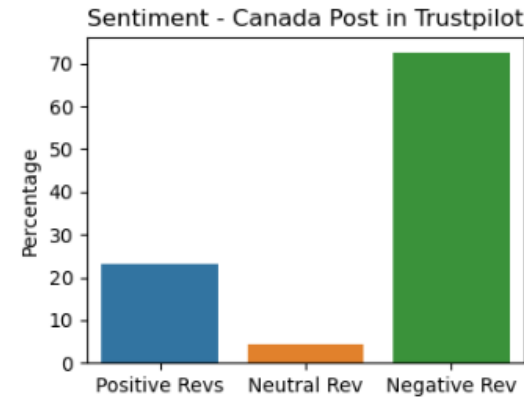
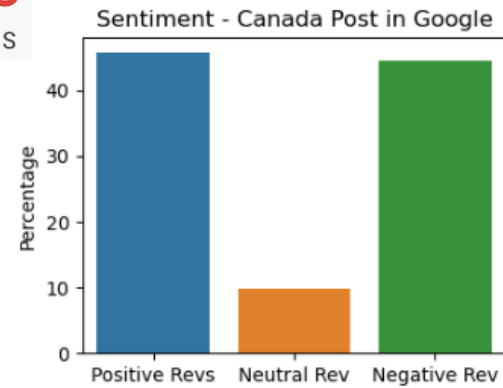
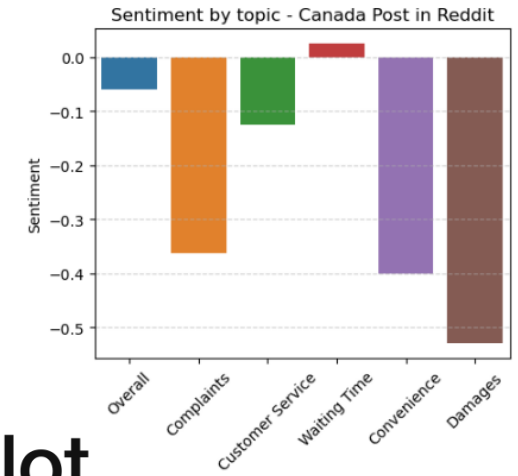
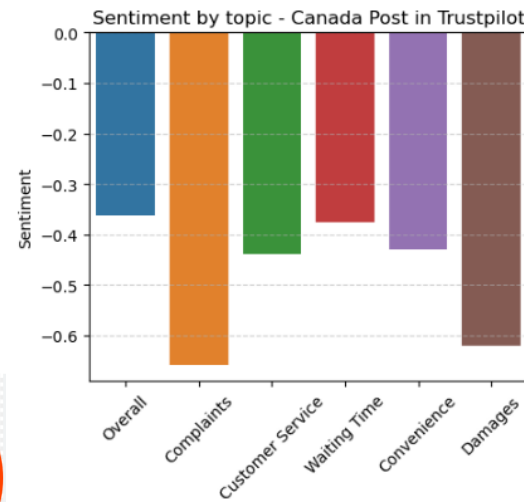
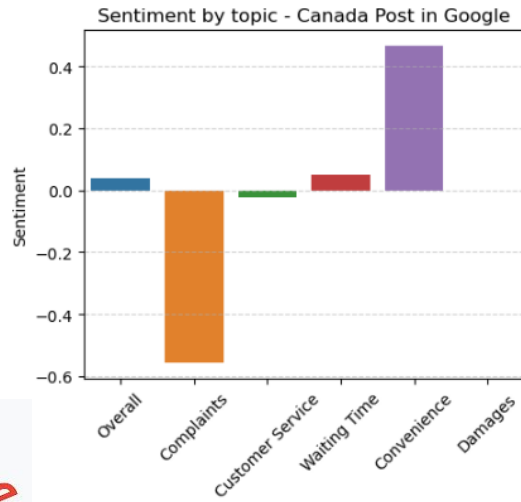
Complaint analysis

- The negative_review is defined with `df['Sentiment'] < 0`;

Customer services analysis

- the customer_service_reviews is defined as ones with 'service' in the reviews;
- So are waiting_time_reviews, convenience_reviews, damages_reviews.

Complaint & Customer
Service Analysis |



Data Visualization



Conclusion

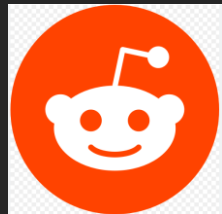


CONCLUSIONS & CROSS-COMPARISON



Overall Sentiment	0.0410	0.3378	-0.2737. Predominantly negative sentiment among customers
Complaints	-0.5581. Predominantly negative.	-0.5917 Negative sentiment. Customers expressing complaints are dissatisfied with the service.	-0.6707. Strongly negative sentiment in reviews that mention complaints.
Customer Service Reviews	-0.0246. Some dissatisfaction with the customer service experience	0.4636. Positive sentiment. Customers seem to have positive experiences with UPS's customer service.	-0.3080. Negative sentiment. Customers are dissatisfied with the customer service experience.
Waiting Time Reviews	0.0498. Slightly positive sentiment	0.1481. Slightly positive sentiment. Customers may have a generally positive experience with waiting times.	-0.3106. Negative sentiment. Customers are expressing dissatisfaction with waiting times.
Convenience Reviews	0.4655. Significantly positive sentiment. Customers seem to appreciate the convenience provided by the service.	-0.0337. Slightly negative sentiment. Some customers may find UPS's services less convenient.	-0.2694. Negative sentiment. Customers may not find the service convenient.
Damages Reviews	Nan. Due to missing or insufficient data.	0.0900. Slightly positive sentiment.	-0.4456. Negative sentiment.

Conclusion



CONCLUSIONS & CROSS-COMPARISON



Overall Sentiment	-0.0598. Slightly negative sentiment among users.	-0.0386. Slightly negative sentiment among users	-0.0267. Slightly negative sentiment among users..
Complaints	-0.3638. Negative sentiment in reviews that mention complaints	-0.3252. Negative sentiment in reviews that mention complaints.	-0.4150. Negative sentiment in reviews that mention complaints
Customer Service Reviews	-0.1255. Users are dissatisfied with the customer service experience.	-0.0971. Negative sentiment. Users seem to be dissatisfied with the customer service experience.	-0.3182. Negative sentiment.
Waiting Time Reviews	0.0243. Slightly positive. Users may have a more positive experience with waiting times.	-0.1054. Negative sentiment. Users may be dissatisfied with waiting times.	nan
Convenience Reviews	-0.4005. Negative sentiment. Users may not find the service convenient.	nan.	nan
Damages Reviews	-0.5301. Negative sentiment. Users are dissatisfied with experiences related to damages	-0.3779, Negative sentiment. Users are dissatisfied with experiences related to damages.	nan

Conclusion



CONCLUSIONS & CROSS-COMPARISON



Overall Sentiment	-0.3621 Predominantly negative sentiment among customers.	-0.3709. Predominantly negative sentiment among customers.	0.1256 Slightly positive sentiment among customers.
Complaints	-0.6580. Strongly negative sentiment.	-0.6644. Strongly negative sentiment..	-0.6722. Strongly negative sentiment.
Customer Service Reviews	-0.4380. Negative sentiment.	-0.4350. Negative sentiment.	-0.3086. Negative sentiment
Waiting Time Reviews	-0.3750. Negative sentiment.	-0.4182. Negative sentiment.	-0.3624. Negative sentiment.
Convenience Reviews	-0.4293. Negative sentiment.	-0.4317. Negative sentiment.	-0.3888. Negative sentiment.
Damages Reviews	-0.6213. Strongly negative sentiment.	-0.6668. Strongly negative sentiment.	-0.3339. Negative sentiment..

Recommendations & Further Research

For the sample dataset selected for this study

- According to GoogleReviews, Canada Post does significantly better than the other two competitors do in 'conveniences', which indicates Canada Post covers wider population and provides easier services than its peers.
- Based on Reddits, Canada Post is slightly better than other competitors in 'waiting time', which also makes sense because Canada Post has more service partners that helps reduce the customer traffics.
- According to GoogleReviews, Canada Post would need to improve 'customer services' as this seems to be an area it is facing a lot of challenges from UPS.
- Trust Pilot seems to be full of complains and negative sentiments for all the companies however gives an overall positive sentiments for Purolator. This needs to be further looked into and investigated.

The background features a dense, abstract pattern of numerous thin, wavy lines in various colors including red, blue, green, yellow, and purple. These lines flow across the frame, creating a sense of movement and depth. The overall composition is set against a dark gray background.

THANK YOU
