



# UNIVERSITÀ DI TRENTO

Dipartimento di Ingegneria e Scienza dell'Informazione

Corso di Laurea in Informatica

ELABORATO FINALE

## THE EMOTIONAL IMPACT OF THE COVID-19

*Studying the emotional impact of the Covid-19 pandemic using social media*

Supervisore  
Alberto Montresor

Laureando  
Simone Alghisi

Anno accademico 2020/2021

# Ringraziamenti

*...thanks to...*

# Contents

<b>Sommario</b>	<b>2</b>
<b>1 Introduction</b>	<b>2</b>
1.1 Context and motivations . . . . .	2
1.2 Project description . . . . .	2
<b>2 Data collection</b>	<b>3</b>
2.0.1 Tweets . . . . .	3
2.0.2 Analyzed period and languages . . . . .	4
2.1 Cras in aliquam quam, et . . . . .	4
2.1.1 Sed pulvinar placerat enim, a . . . . .	4
2.2 Vivamus hendrerit imperdiet ex. Vivamus . . . . .	4
<b>3 Conclusioni</b>	<b>4</b>
<b>Bibliografia</b>	<b>5</b>
<b>A Titolo primo allegato</b>	<b>7</b>
A.1 Titolo . . . . .	7
A.1.1 Sottotitolo . . . . .	7
<b>B Titolo secondo allegato</b>	<b>8</b>
B.1 Titolo . . . . .	8
B.1.1 Sottotitolo . . . . .	8

# Sommario

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

Sommario è un breve riassunto del lavoro svolto dove si descrive l'obiettivo, l'oggetto della tesi, le metodologie e le tecniche usate, i dati elaborati e la spiegazione delle conclusioni alle quali siete arrivati.

Il sommario dell'elaborato consiste al massimo di 3 pagine e deve contenere le seguenti informazioni:

- contesto e motivazioni
- breve riassunto del problema affrontato
- tecniche utilizzate e/o sviluppate
- risultati raggiunti, sottolineando il contributo personale del laureando/a

## 1 Introduction

The COVID-19 pandemic is having a huge impact on our lives, that goes beyond the direct effects of the virus. Besides the fear of infection, lockdown measures adopted by many countries are limiting the possibility to move, work, have contact with others, and are creating a situation of economic crisis and generalized uncertainty about the future. The psychological effects of this unprecedented situation need to be studied.

### 1.1 Context and motivations

During this year, everyone daily life changed significantly and we had to adapt to restrictive measures in order to stop the disease: whether we liked it or not. This research proposed by Eurecat really caught my eye: the possibility to study how people perceived all of this situation, and better understand which measures were more welcomed than others, was really fascinating and, above all, may be useful in the case of some other unfortunate event.

### 1.2 Project description

The project consisted in an **analysis of emotions as emerging from Twitter messages** during the pandemic.

**Lexicon-based sentiment analysis tools** have been employed to characterize emotions associated with content on a large scale. Moreover, users have been divided into two different groups w.r.t. their gender, to study the different emotional response of males and females, and also based on their location.

This could allow us to contrast the emotional reaction with the evolution of contagions and deaths, and with the different lockdown and de-escalation stages, in different areas.

## 2 Data collection

The dataset used for the project is the **echen102/COVID-19-TweetIDs** GitHub repository[1]. The repository contains an ongoing collection of tweets IDs, starting on the 28th of January 2020, from specified accounts and also real-time tweets that mention specific keywords.

Number of files	<b>10 402</b>
Number of identified languages	<b>65</b>
Number of tweets	<b>1 055 843 481</b>
Number of unique tweets (no retweets)	<b>323 504 667</b>
Dataset compressed size	<b>865 GB</b>
Dataset estimated uncompressed size	<b>6.252 TB</b>

Table 2.1: Dataset general statistics

language	ISO	unique tweets	retweets	total	percentage
English	en	195 645 826	473 950 322	669 596 148	63.41%
Spanish	es	35 533 886	111 464 189	146 998 075	13.92%
Portuguese	pt	15 459 760	29 912 427	45 372 187	4.30%
French	fr	9 547 251	23 635 273	33 182 524	3.14%
Undefined	und	20 560 392	8 590 707	29 151 099	2.76%
Indonesian	in	9 029 012	16 479 537	25 508 549	2.41%
German	de	8 091 516	11 447 554	19 539 070	1.85%
Japanese	ja	3 228 542	10 220 609	13 449 151	1.27%
Italian	it	5 256 748	7 173 234	12 429 982	1.18%
Turkish	tr	3 347 597	6 698 252	10 045 849	0.95%

Table 2.2: Top 10 languages with the most tweets

### 2.0.1 Tweets

To comply with Twitter's Term of Service, tweets cannot be released publicly: the repository is in fact a collection of tweets IDs. The original tweets can be retrieved, or hydrated, using the Python library Twarc with a Twitter Developer Account. Given an id, Twarc simply uses the token of the associated developer account to contact the API, and returns the corresponding tweet as a json object.

The original structure of the tweets was changed, in order to consider only the relevant fields:

```
{
  "id": 1307025659294674945,
  "full_text": "Here's an article that highlights the updates...",
  "lang": "en",
  "created_at": "Fri Sep 18 18:36:15 +0000 2020",
  "retweet_count": 11,
  "favorite_count": 70,
  "user": {
```

```

    "id": 2244994945,
    "id_str": "2244994945",
    "screen_name": "TwitterDev",
    "name": "Twitter Dev",
    "description": "The voice of the #TwitterDev team and your official...",
    "location": "127.0.0.1",
    "followers_count": 513958,
    "statuses_count": 3635,
    "default_profile_image": false,
    "profile_image_url_https": "https://pbs.twimg.com/profile_images
      \1283786620521652229\1E0DkLTh_normal.jpg"
  }
}

```

## 2.0.2 Analyzed period and languages

We have decided to consider the period from January 2020 to March 2021

## 2.1 Cras in aliquam quam, et

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

### 2.1.1 Sed pulvinar placerat enim, a

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

## 2.2 Vivamus hendrerit imperdiet ex. Vivamus

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

# 3 Conclusioni

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu

purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis.  
Phasellus gravida tellus velit, non eleifend justo lobortis eget.

# Bibliography

- [1] Emily Chen, Kristina Lerman, and Emilio Ferrara. Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set. *JMIR Public Health and Surveillance*, 6(2):e19273, 2020.



# Allegato A      Titolo primo allegato

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

## A.1      Titolo

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

### A.1.1      Sottotitolo

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

# Allegato B      Titolo secondo allegato

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

## B.1      Titolo

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

### B.1.1      Sottotitolo

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.