

SØREN KOLBYE JENSEN  
Bsc. Ha(IT)  
#43124133123  
December 1, 2017

Contents

1 Summary 1

2 Introduction 2

2.1 Case introduction . . . . . 2

3 Problem Formulation and Reserach questions 3

3.1 Something . . . . . 3

4 Methodology 3

4.1 Data Acquisition and Dataset Description . . . . . 3

# 1 Summary

Summary section

## 2 Introduction

Big Data has changed how data is viewed in society, businesses have invested heavily in infrastructure that can improve data collection. It is now possible, to use Big Data to analyze the behaviour of Social Media channels such as Twitter, Facebook and Instagram.

These principles are called "Datascience and Datamining". According to (INSERT BOOK), datascience and datamining can be defined as follows:

Data science is the act of using fundamental principles, to guide extraction of knowledge from datasets.

Datamining on the other hand, deals with extracting knowledge from data using technologies, which incorporates the principles of data science.

This research paper, aims to use the fundamental principles of Data science, in conjunction with datamining technologies. To analyze the behaviour of U.S president Donald J. Trump. and how this behaviour affected his political campaign, as well as how his behaviour affected the American financial markets.

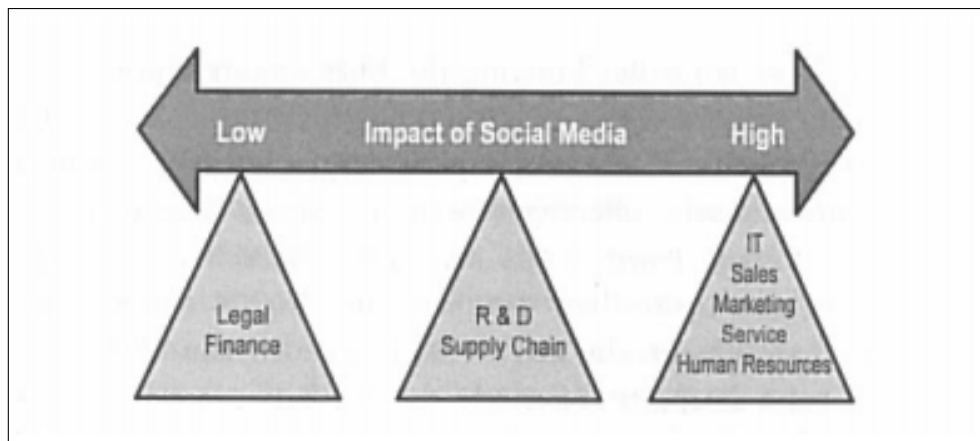


Figure 2: real, local caption for reference

### 2.1 Case introduction

Donald John Trump, born in 1946 in Queens New York, is the CEO of the Trump Organization and is well known as a real-estate broker in the United States. However, on June 16th, 2015, Trump officially announced his candidacy as the 20th US President". This political campaign was heavily involved with Social Media, as Trump was actively fighting big news channels, such as CNN, NBC etc. labelling them as "fake news". Thus, Trump figured out a way to distribute highly controversial statements and gain political support by his use of Social Media Channels such as Twitter.

### 3 Problem Formulation and Reserach questions

The goal of this research paper is to analyze trumps behaviour on Twitter and how he used Twitter to combat "fake news". A visual analysis on his behaviour between DATE and DATE will be conducted to reveal the patterns in his behaviour in this timeframe. This leads to the reserach question:

What was trump's behaviour on social media, prior to his announcement of his presidency. And how did it change, during and after his political campaign.

Furthermore, Trump's political campaign was heavily focused on combating illegal immigration. One of the primary targets of Trump's statement was Mexico. This leads to the research question:

How did Trump's political statements on Twitter affect the American and Meixcan currency.

Testing git

#### 3.1 Something

## 4 Methodology

To answer the research questions two different types of analysis is presented

Furthermore, the CRISP framework has been used in a modified fashion, in order to better illustrate the context of this research paper. The model serves to illustrate how the data was processed throughout the proejct in order to yield a final result/answer to the resarch questions. The model is iterative and serves to explore the Twitter data, so a better understanding of the data can be reached. This process has been conciously iterated throughout the project.

"Business understanding" is in this context the overall problem formulation and research questions that we wish answered.

Table 1: Distribution of Number of Injured Non-Motorists by Injury Severity Level

Injury Severity	Pedestrian		Bicyclist		All Non-Motorists	
Possible injury	1700	67.70%	502	59.40%	2202	65.60%
Non-Incapacitating injury	523	20.80%	259	30.70%	782	23.30%
Incapacitating injury	250	10.00%	84	9.90%	334	9.90%
Fatal injury	39	1.50%	0	0.00%	39	1.20%
<b>Total</b>	<b>2512</b>	<b>100.00%</b>	<b>845</b>	<b>100.00%</b>	<b>3357</b>	<b>100.00%</b>

#### 4.1 Data Acquisition and Dataset Description

The datasets in this research paper has been collected from Trump's Twitter page, ranging from 2009 to 2017.

The first datasets consists of Twitter data on Trump's primary twitter channel.

The second dataset consists of Historical data on the American and Mexican currency between Trump's announcement to 14/07/2017

05/04/2009 07/14/2017