Discuss how data analytics can be used to influence people.

In 2014, the findings from the first emotional contagion[1] experiment carried out by an online social network, Facebook, on a pool of 689,003 users were published[2], signifying a disquieting turning point for many: a realisation of the powerful capabilities companies' use of data analytics can have, in this case with regards to the manipulation of arguably one of the most fundamental aspects of human behaviour - our emotions. Through the analysis and categorisation of content as either positive or negative[3], and then by altering the volume of such content presented on the News Feeds of the two groups, the study demonstrated that a clear correlation could be seen between the "emotional valence" of the content viewed and that of the users' own posts.

Even 9 years later, this study is one of the most striking and statistically-backed examples of how data analytics can be used to influence people. However, its impact is not solely confined to affecting people's emotions and actions but also their perspectives, habits, relationships and more. Often misinterpreted as the crystal ball of technology conglomerates, data analytics is most aptly defined as a process that allows one to observe trends and draw conclusions about the information data sets contain, which can then be used to change plans of action moving forward. Consequently, its broad use as a tool lends itself to its versatility in different sectors, ranging from business to education, which shall be

[1] 'a phenomenon of an automatic adoption of an emotional state of another person'
Hatfield, E., Cacioppo, J.T. and Rapson, R.L. 1993
[2] Kramer, A.D., Guillory, J.E. and Hancock, J.T. (2014)
[3] 'Posts were determined to be positive or negative if they contained at least one positive or negative word, as defined by Linguistic Inquiry and Word Count software (LIWC2007)'

discussed. Yet, all its uses are underpinned by the cardinal idea that it is effective in enabling these entities to influence their users, customers, population: people.

Application in business:

In business, big data analytics can allow for a greater degree of precise control over consumer markets. Digital activity has permeated daily life, creating markets for the data based on these information-seeking and socialising behaviours in the process, or rather, markets for shaping and keeping the attention of the consumers[4]; access to this is advantageous to not only media companies, but any company that wishes to gain an edge over competitors. The use of descriptive analytics on data mined from potentially millions of real-time customer interactions with digital touchpoints, such as targeted advertisements and navigation through the business' website, combined with, but not limited to, historical spending or viewing habits and reviews, can be used to obtain behavioural insights[5] into customers on an individual basis. For instance, Google Analytics is able to provide a more customer-centric measurement of interactions, using multiple identity spaces as well as marketer-provided User IDs; this is unlike previous measurements, which were fragmented by device or platform[6]. Additionally, the use of predictive data analytics can then allow organisations to predict future interests and provide recommendations based on this past consumer behaviour, not only from the specific user, but also from other pools of users with similar interactions.

[4] Schroeder, R. (2018)
[5] Holmlund, M. et al. (2020)
[6] Srinivasan, V. (2020)

One of the most notable examples of this can be seen in Netflix's recommender system, Cinematch, which uses multiple algorithms that rely on statistical data analytics and machine-learning techniques including "both supervised (classification, regression) and unsupervised approaches (dimensionality reduction through clustering or compression, e.g., through topic models)"[7]. The system thus ensures that in a diverse pool of members, each will be recommended only titles that should compel them, without the need for "additional input or an explicit expression of their intents or goals"[8], based on their viewing habits and the behaviour of similar users - thus influencing them to watch more and spend more time on the platform.

Application in journalism:

"The biggest impact of big data will be that data-driven decisions are poised to augment or overrule human judgement." [9]

The omnipresence of editorial analytics in modern-day newsrooms is undeniable; dominant metrics platforms such as Chartbeat and Ophan are a staple in the offices of many prominent news companies such as The Guardian and BBC News. These platforms provide clients with dashboards displaying real-time data analytics of audience attention from as broad as concurrent visits to as granular as the entry rate of specific pages, in order to inform the way editors decide to structure homepages and fine-tune headlines.[10] Hence, the use of editorial

[7] Gomez-Uribe, C.A. and Hunt, N. (2015)
[8] Lamkhede, S.D. and Kofler, C. (2021)
[9] K. Cukier and V. Mayer-Schönberger, (2013)
[10] Cherubini, F. and Nielsen, R.K. (2016)

analytics enables these companies to structure content in a format optimal for influencing readership retention and for building a loyal and returning audience, which they can consequently monetise in a plethora of ways - one example being through paywalls. However, in the continual search for the increased revenues that these data analytics can represent, the development of a form of participative gatekeeping is an increasing possibility: the endless feedback loop centred on engagement metrics holds the potential to "overrule human judgement" and influence journalists themselves, to the extent of the abdication of their duty to inform audiences about the most significant public issues of today[11]. This is due to the incentivisation of click-worthy content that is sensationalist or incendiary, be it celebrity stories, misogynistic articles or conspiracy theories.

Application in education:

Educational data mining and learning analytics are new disciplines that have emerged from the implementation of data analytics in education, all deriving from the aim to enhance the academic experiences of students and teachers. Data analytics techniques can enrich the learning process of individual students by providing real-time feedback, which can take the form of adaptive learning paths that tailor learning environments, making them more compatible with students' acquired skills and knowledge base, learning style or metacognitive abilities[12]. By applying the data and process mining tools traditionally used in business process management to students' learning flows throughout the course duration, information gleaned from the learning trajectories assembled can be utilised to provide

[11] Petre, C. (2015)
[12] Smedt, J.D. et al. (2017)

tutors with monitoring systems. These can detect complex topics for students, suggest optimum courses of action to take, or even help to construct predictive models with the ability to assess the probability of students performing a certain way in exams. For example, social network analysis can be used to track contributors in student groups, using forum discussions and other media data. In a similar vein, pattern mining from exercise solutions with multiple stages can be used to find successful resolution strategies.[13] Hence, the influence data analytics can have on students is the enhancement of their learning experience, which manifests in the greater retention of information learnt and a greater depth of understanding.

Application in politics:

Predictive data analytics lies at the crux of political consulting firms' work, as they leverage ever-expanding datasets to create accurate predictive models about the inclinations of the voter pool, their expected behaviours and reactions to campaign outreach activities[14]; the aim of data analytics here is to develop micro-targeted communication efforts in order to influence members of the public, and convert hesitancy to votes, maximising the chances of victory for the campaign. To do so, such firms first purchase a broad range of information from consumer data vendors, such as estimated years of education or mortgage information. They combine this data with that which individuals have provided themselves, for example, through previous donation or volunteering history or receptiveness to phone calls. Therefore, they are able to develop three categories of "predictive scores" for each voter in

[13] Ibid
[14] Nickerson, David W., and Todd Rogers. (2014)

the database: responsiveness scores, behaviour scores and support scores. Ultimately, these predictive scores allow for the forecasting of political behaviours and engagement, and more efficient attempts to influence mainly those who are least likely to vote for the party.

A prominent example of this was seen in the creation of "Narwhal" during the 2012 campaign to re-elect President Obama: it was a program that allowed for the centralisation of all the data collected from these information sources, and campaign applications, into one interface. With applications now such as Dreamcatcher, "a tool developed to 'micro target' voters based on sentiments within text", and Dashboard, "the 'virtual field office' application that helped volunteers communicate and collaborate"[15], a successful campaign was achieved, demonstrating the power of data analytics in political campaigns at a time when previous campaigns struggled to assimilate their various sources of data in order to analyse it with such efficiency.

More recently, the Cambridge Analytica scandal in 2018 revealed the mass scale of targeted manipulation possible with the use of data analytics by businesses with access to large volumes of detailed data. Facebook obtained information about 87 million individuals: their psychographic profiling, obtained through surveys[16], and digital activity, obtained through likes and other exploitable traits. This data was shared with Cambridge Analytica, enabling them to direct targeted political messaging at specific users and most likely influence some voters' choices in the 2016 American presidential election campaign of Donald Trump[17].

[15] Gallagher, Sean (2012)
[16] Hern, A. (2018)
[17] Ghosh, D. and Scott, B. (2018)

Application in healthcare:

Precision medicine is a healthcare approach that has been revolutionised by advancements in genomic data analytics. The development of expansive biologic databases, such as the human genome sequence combined with clinical data derived from the vast number of electronic health records, provides a source of big data with which machine-learning algorithms can be used to classify patients methodically into phenotypic groups[18]. In doing so, it is possible to use statistical analysis techniques to identify clinically disease-contributing gene variants, enabling "a quicker diagnosis for patients with a rare disease" as well as "matching people to the most effective medications and interventions"[19] personalised to them. A vital example of this in practice is evidenced in The NHS Genomic Medicine Service's work with Genomics England, which has laid the foundations for such use of genomics in routine clinical care, through the ground-breaking 100,000 Genomes Project they delivered[20]. Hence, data analytics can also be employed in healthcare to influence patients' lives through the development of personalised treatments, which can improve their quality of life significantly.

Ultimately, although the use of data analytics to influence people is often associated with business, it can have applications in multiple spheres of contemporary human life. However, it is essential to consider that while this continually expanding area of data science is attractive, there is a growing need for some form of ethical oversight to prevent it from being

[18] He, K., Ge, D. and He, M. (2017)
[19] NHS (undated)
[20] NHS (2022)

used to achieve misguided intentions - and undermine the very reasons for which it was developed.

Bibliography

1. Hatfield, E., Cacioppo, J.T. and Rapson, R.L. (1993) “Emotional contagion,” *Current Directions in Psychological Science*, 2(3), pp. 96–100. Available at: https://doi.org/10.1111/1467-8721.ep10770953.
2. Kramer, A.D., Guillory, J.E. and Hancock, J.T. (2014) “Experimental evidence of massive-scale emotional contagion through social networks,” *Proceedings of the National Academy of Sciences*, 111(24), pp. 8788–8790. Available at: https://doi.org/10.1073/pnas.1320040111.
3. Ibid
4. Schroeder, R. (2018). “Big data: shaping knowledge, shaping everyday life”. In *Social Theory after the Internet: Media, Technology, and Globalization*, pp. 126–148. UCL Press. Available at: https://doi.org/10.2307/j.ctt20krxdr.9
5. Holmlund, M. *et al.* (2020) “Customer experience management in the age of big data analytics: A strategic framework,” *Journal of Business Research*, 116, pp. 356–365. Available at: https://doi.org/10.1016/j.jbusres.2020.01.022.
6. Srinivasan, V. (2020) “Introducing the new google analytics”. *Google*. Available at: https://blog.google/products/marketingplatform/analytics/new_google_analytics/ (Accessed: January 14, 2023).
7. Gomez-Uribe, C.A. and Hunt, N. (2015) “The Netflix Recommender System,” A*CM Transactions on Management Information Systems*, 6(4), pp. 1–19. Available at: https://doi.org/10.1145/2843948.
8. Lamkhede, S.D. and Kofler, C. (2021) “Recommendations and results organization in netflix search,” *Fifteenth ACM Conference on Recommender Systems* [Preprint]. Available at: https://doi.org/10.1145/3460231.3474602.
9. K. Cukier and V. Mayer-Schönberger, (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, pp. 141. Houghton Mifflin Harcourt
10. Cherubini, F. and Nielsen, R.K. (2016) “Editorial analytics: How news media are developing and using audience data and metrics,” *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2739328.
11. Petre, C. (2015) “The Traffic Factories: Metrics at Chartbeat, Gawker Media, and The New York Times,” *Columbia University*. Available at: https://doi.org/10.7916/D80293W1.
12. Smedt, J.D. *et al.* (2017) *Data Analytics applications in education*. CRC Press.
13. Ibid
14. Nickerson, David W., and Todd Rogers. (2014). "Political Campaigns and Big Data." *Journal of Economic Perspectives*, 28 (2): 51-74. Available at: http://dx.doi.org/10.1257/jep.28.2.51

15. Sean Gallagher - Nov 14, 2012 4:15 pm U.T.C. (2012) *Built to win: Deep Inside Obama's campaign tech*, *Ars Technica*. Available at: https://arstechnica.com/information-technology/2012/11/built-to-win-deep-inside-obamas-campaign-tech/ (Accessed: January 15, 2023).
16. Hern, A. (2018) *Cambridge Analytica: How did it turn clicks into votes?*, *The Guardian*. Guardian News and Media. Available at: https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie (Accessed: January 15, 2023).
17. Ghosh, D. and Scott, B. (2018) *New facebook scandal shows how political ads Manipulate you*, *Time*. Time. Available at: https://time.com/5197255/facebook-cambridge-analytica-donald-trump-ads-data/ (Accessed: January 15, 2023).
18. He, K., Ge, D. and He, M. (2017) "Big Data Analytics for genomic medicine," *International Journal of Molecular Sciences*, 18(2), p. 412. Available at: https://doi.org/10.3390/ijms18020412.
19. *NHS Genomic Medicine Service* (no date) *NHS choices*. NHS. Available at: https://www.england.nhs.uk/genomics/ (Accessed: January 15, 2023).
20. *Accelerating genomic medicine in the NHS* (2022) *NHS choices*. NHS. Available at: https://www.england.nhs.uk/long-read/accelerating-genomic-medicine-in-the-nhs/ (Accessed: January 15, 2023).

Word count: 1646 (excluding titles, footnotes and references).