**Big Data and (In)equality**

***Listening***

**Listen to the podcast 1 (from 1’42’’ to 5’15’’) and complete the following statements:**

1. Chris Bud works

……………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………

1. The journalist wants to know

……………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………

1. What mostly struck Chris Budd was

……………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………

1. The main difference Chris Budd makes between the use of Big Data in the past and now is that

……………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………

**Big Data and (In)equality**

***Reading and Grammar***

1. **In each case below decide whether the word requires an ‘s’ at the end or not.**

We live in the **information(s)** age. Most of what we do is hugely influenced by our access to massive **amount(s)** of data — whether this is through the Internet, on our computers, or on our **mobile(s)** phones. The buzz word to describe this deluge of information is *Big Data*. In 2012 the UK government identified Big Data as one of the eight great technologies of the future. So what does the challenge of Big Data entail and how can we meet it?

Perhaps the leading source of **current(s)** Big Data is the Internet. According to a recent estimate, about 1021 bytes (a zettabyte) of **information(s)** are added to the Internet every year, much of which is graphical in content. The *internet penetration* in the UK is over 80%, and in all but a few **country(ies)** it is over 20%.

A major source of this **data(s)** is the ever growing content on social media **website(s)**. For example, Facebook was launched in 2004. It now has 2 **billion(s)** registered users (about a quarter of the world's population!) of which 1.5 billion are active. Around 2.5 billion **piece(s)** of content (around 500 terabytes of information) are added every day to Facebook, with most of this data stored as pictures. The search engine Google is estimated to seek information from around 15 exabytes (1015 bytes) of data (which it searches using a clever mathematical algorithm).

Another source of Big Data are mobile and **smart(s)** phones. There are now more mobile phones than **people(s)** in the world, with the potential for 25,000,000,000,000,000,000 possible simultaneous conversations. The forthcoming plans for a 5G network will offer data **rate(s)** at 1 gigabyte per second simultaneously to **ten(s)** of workers on the same office floor. Another fast-approaching **technology(ies)** are sensors that can provide constant monitoring of, say, our state of health (with significant ethical implications). The 5G network will support several **hundred(s)** of thousands simultaneous connections for massive sensor deployments. Indeed, the future is rapidly approaching: soon our devices will simply communicate with each other (for example the cooker talking to the dishwasher and also to the supermarket every time a meal is prepared) with little or no human interference — it's called the *Internet of things*.

Significant amounts of data, of significant **interest(s)** to the social sciences, also come from the way that we use our devices and the information this gives about our **lifestyle(s)**. Again there are significant ethical issues here. Every time we make a purchase on Amazon, use our bank on-line, switch on an electrical **device(s),** or simply use a mobile phone or write an email, we are creating data which contains information that can in principle be analysed. For example, our shopping **habit(s)** can be determined, or our location tracked and recorded. Mathematics can be used at all stages of this, but we must never lose sight of the moral dimension in so doing.

**Big Data and (In)equality**

***Reading***

# Can Big Data Analytics Help Bridge The Economic Inequality In India?

******

**Match the paragraphs (A to D) with the subtitles in the box below:**

|  |
| --- |
| 1. **Predictive Analytics For Farming** 2. **Providing Density And Location Coordinates To Fishermen** 3. **Big Data For Telemedicine** 4. **Reducing Dropout Rate At Schools With Microsoft** |

**Introduction**

Can Big data help alleviate economic inequality in India? According to a report by the Johannesburg-based company New World Wealth, India is the second-most unequal country across the globe, with millionaires controlling 54 percent of its wealth.

With a total individual wealth of $5,600 billion, it’s among the 10 richest countries in the world — but still the average Indian is relatively poor and lives below the baseline. Data from Credit Suisse on India indicates that the richest 1 percent own 53 percent of the country’s wealth. While at the end of the pyramid, the poorer half jostles for a mere 4.1 percent of national wealth. Severe economic disparity has been one of the common refrains in India’s growth story and experts are wondering as to how it could pose a threat to the country’s future growth.

***A………………………………………………………………….***

How can data science techniques be used to reduce the income inequality? [Vijay Nadadur](https://twitter.com/vjkant?lang=en), computer scientist and CEO at [Stride.ai](https://stride.ai/) mentioned at a [forum](https://www.quora.com/How-can-big-data-help-rural-people-in-India) about how big data can help tackle rural India’s problems and help the country grow sustainably. Sharing his thoughts, Nadadur mentioned how big data can play a major role in transforming healthcare for rural India with a sensor-based data collection framework. This could indicate various factors — even warnings about seasonal diseases. These indicators can then pave the way for need-based preventive healthcare to the citizens. Also, big data can help in promoting telemedicine, at least at a primary healthcare level.

**B………………………………………………………………………….......**

Nadadur also cited another use case where analytics helped in providing information related to the density of fish along with the location coordinates to fisherman. This could be a big booster to the lives of fishermen, increasing their efficiency, and in turn, profits.

**C……………………………………………………………………………...**

In 2016 Microsoft collaborated with International Crops Research Institute for Semi-Arid Tropics (ICRISAT) and Andhra Pradesh government, to develop a new mobile application for farmers. The idea behind developing the Sowing App and providing Personalised Village Advisory Dashboard was to provide powerful cloud-based predictive analytics so that farmer felt empowered with information and insights which were required to take farming decisions. This app helped farmers reduce crop failures and increase yield. The app and dashboard were already loaded with in-depth data about rainfall of over the last 45 years and 10 years of groundnut sowing progress data. This led to better income opportunity for the farmers. Smart sensors are already being deployed for precision farming that help to get out of drought-related information to farmers.

**D………………………………………………………………………………**

In a bid to improve Andhra Pradesh’s education ecosystem, Microsoft is working with the state Government on a machine learning-based model to analyse and predict dropouts and take preventive action. Andhra Pradesh government uses Azure Machine learning to predict which students are likely to drop out of school across its 10,000 schools. Officials have created more than 6,00,000 predictions using Azure Machine Learning, revolutionising how Indian local governments increase student retention.

**Read the article and match the following definitions with words from the text:**

•To ease, to reduce (Introduction)

•A minimum, or starting point (Introduction)

•To make one's way by pushing or shoving (Introduction)

•To grapple with, to try to deal with (a task, a difficulty, etc.); to try to solve (a problem) (p. A)

•to facilitate or lead on to a result or object in view (p. A, three words)

•That has the character, quality, or function of forecasting the future (p.C)

•production; that which is produced, produce; *esp.* amount of produce. (p. C)

•To perform the action of scattering or depositing seed on or in the ground so that it may grow. (p.C)

•The condition or quality of being dry; dryness, aridity, lack of moisture. (p. C)

•an attempt to win or secure something (p. D)

•A person who withdraws from a course of study or from society (p. D)

**Big Data and (In)equality**

***Listening***

**Listen to the podcast 2 (from 0’00’’ to 2’45’’) and decide whether the following questions are true or false:**

1. Algorithms cannot be used to make sense of the world.
2. Kathy O’Neil has written a book on Big Data and inequality.
3. She thinks that Big Data is fundamentally racist.
4. She denounces the system of criminal justice algorithms as being fundamentally racist.
5. Many of these algorithms take into account personal characteristics like age, sex, geography and family background.

**Listen to the podcast 2 (from 0’00’’ to 2’45’’) again and complete the following sentences:**

1. Algorithms are more and more being used in

……………………………………………………………………………………………………………………………………………………………………………………

1. Kathy O’Neil is

……………………………………………………………………………………………………………………………………………………………………………………

1. O’Neil thinks that western societies are

……………………………………………………………………………………………………………………………………………………………………………………

1. The machine learning algorithms that we train

……………………………………………………………………………………………………………………………………………………………………………………

1. The “recidivism risk” algorithms are trying to

……………………………………………………………………………………………………………………………………………………………………………………