Hello everyone, my name is AmanDeep Singh and I am a student at University of Applied Science in Utrecht pursuing a minor in Big Data & Design.

as of I know we are all living in a digital age where our personal information is constantly at risk right?... So today I'll dive into a recent popular topic of DATA BREACHES: a Data Story of Trust, ethics, and the responsibility of protecting our digital lives.

Data breaches have a larger impact than just numbers; they compromise trust between users and companies. It raises key concerns about like ethical responsibility and the right to privacy. BUT let’s look into this using the 'List of Top Data Breaches' (2004 - 2021)" dataset from Kaggle.

My data story starts with cleaning and preparing the data. It's like preparing the stage for the main act, making sure that all missing or duplicate entries were carefully checked and removed. I also ensured that the data types were consistent,

years had to be in a precise chronological integer format, not just random strings/ object, so that the timeline of data breaches could be clearly defined. After all these cleaning and preparing the data step, I was quite confident that the data story graphs I was about to show you will be accurate.

Now… let take look at the first visualization about Users affected by data breaches. Have you noticed the rising figures in 2013 and 2019? Behind these rising numbers ARE USER, communities whose trust has been compromised by the huge exposure of user’s personal information leaks.

NOW A GOOD RESEARCH QUESTION THAT COMES TO MY MIND IS …

1. Who were the users affected and through what means did their data get breached?

I will answer these questions with my upcoming graphs. As shown in the second scatterplot graph, the size of data breaches in some large businesses is significant. This graph shows a very different picture where the depth of data breaches are defined in terms of frequency and variety rather than the number of users as shown above. This visual also depicts a reoccurring increasing data breach pattern over time, indicating the need for stricter and more disciplined security measures.

BUT To fully understand a cyber security failure, I will examine the methods of how the data gets breached. As you can see this graph, the bars on the graph reflect the unique data breach methods and are color coded for clarity. It’s clear from the visual that despite the technological progresses, vulnerabilities still persist, and it frequently resulting in catastrophic outcomes. But it is also apparent from the graph that how complex today’s cybersecurity threats have evolved over the Years.

But let’s take a break and reflect on what matters most, the people the ethics, the principles. These data breaches was more than an email and phone number; this was about people and the invasion of their personal world - life without their consent. Ethically speaking, I believe one has to rise beyond the social norms and legal expectations with which these companies have failed to rise above. It's not just breaking rules; it's failing to protect basic rights to which people should be entitled.

Economically: The data is valuable—maybe even tempting to exploit for personal profit. But at what cost? To me, there is an ethically sound way through which data should be handled. It does not involve compromising privacy for financial gain.

Technologically: We are more connected than ever, and it is beautiful, but that comes with responsibility. The technology that unites us should not be the one that brings risks upon us.

All these are about ethical obligations, about ensuring that our shared humanity does not come at the cost of our privacy.

And that’s why I will conclude my data story for now and let you remind that the data breaches we have seen for example at Yahoo, Facebook, Microsoft

represent more than just security weaknesses. They represent a larger issue of user trust.

The bottom line is that the companies with our data really need good care. They have to be responsible because, when they screw it up, that's not a technical glitch; it's a betrayal of the user's trust.

There could be many reasons: at times, security is weak, and at other times, one shares, due to ignorance or by chance, what was meant to be kept secret. This shows that even reputable large companies find it difficult to keep our data safe.

So, what is the most important conclusion from this data story?

Companies must raise the bar in how they treat our personal data. They really have to get their act together and reflect on what relevance it really has to safeguard privacy. It is more about doing the right thing and making sure they take care of the trust people place in them. In fact, today, it is almost a daily occurrence that data breaches happen.

It’s actually a wake-up call for businesses to step up and value the data they have been entrusted with as much as they would value the trust of a friend for example. BUT it’s also a good reminder for us to ask for and fight for a digital world in which our privacy is not only expected but should ALSO BE RESPECTED And with that said thank you for listening and evaluating my data story! Thank you again!