Eurostat: Stats in a Wrap

Collaboration with universities and statistical innovation

Release date: 20 October 2025

SPEAKERS
Sofia Rodrigues (Statistics Portugal), Aleksandra Bujnowska (Eurostat), Peter Lugtig (Utrecht university), Jonathan Elliott (host)

Jonathan Elliott
Stats in a Wrap, the podcast series from Eurostat.

Jonathan Elliott
Hello and welcome to Stats in a Wrap, the podcast all about statistics from Eurostat, the statistical office of the European Union. I'm Jonathan Elliott, your host for this episode, and I invite you to join me on an adventure into the numbers, the data and all the points on the map that tell the stories of the 27 Member States and their 450 million inhabitants in all their dazzling diversity.

And no, we won't be talking about medians, means or modes, nor Bayesian networks, nor stochastic block models or even heteroscedasticity, not even that. I'm sorry, this podcast is all about you and me and what the numbers tell us about the Union of European countries that we live in, how they help planners to plan and governments to govern, and yes, voters to vote. You see, data is delicious, and we're here to wrap it up in our wrap cafe to nourish body, mind and soul.

Our statisticians love the art and science of turning the raw numbers into a magic lens, an X-ray machine, to see things as they really are. So come on in, and let's get started. If you've ever filled in a survey, ticked a box online, or wondered where governments get all their numbers about jobs, health or how people spend their time, today's episode is for you.

Behind the scenes, Europe's statistical offices are working hard to make data gathering smarter, faster and more effective, and they're not doing it alone. Eurostat has teamed up with innovators in universities and research institutes across Europe, and together, they're rethinking the painstaking process of data gathering by using smartphones.

And they're also working out safe ways to zoom in to minute details in statistics without risking our privacy. At the heart of all this inventiveness is collaboration. Academics bring fresh ideas, national statistical institutes bring real world needs, and Eurostat helps connect it all across the borders.

The goal is to keep official stats trusted, relevant and timely in a world of information and misinformation coming at us from all directions. To guide us through this fascinating topic of innovation in stats, we have 3 guests, and it is my pleasure to introduce them. Sofia Rodrigues, Head of Unit at Statistics Portugal – Sofia, welcome.

Sofia Rodrigues
Hi, Jonathan. It's really a pleasure to be here, especially with Peter and Aleksandra. Thank you so much for having me.

Jonathan Elliott
Thank you for joining us. And next to Sofia is Peter Lugtig, professor at Utrecht University and senior researcher at Statistics Netherlands. Peter, welcome.

Peter Lugtig
Thanks for having me.

Peter Lugtig
And our third guest is Aleksandra Bujnowska, statistical officer at Eurostat. Aleksandra, welcome.

Aleksandra Bujnowska
Hello. Thanks for inviting me.

Jonathan Elliott
Great! So, let's dive into the world of innovation, academia and the statistics that help shape our democracies. Why is innovation so vital in official statistics today? Just give us your take on innovation and its importance.

Sofia Rodrigues
Innovation, it's not really an option now. No, I think it's imperative to innovate because there are so many challenges in this so fast changing world. We are flooded with information, with data. We have to respond to urgent needs. We see that during the Covid time pandemic and then in the Ukrainian war that we need information. We need relevant information and consistent information faster and more disaggregated. We need to innovate in order to change how things are done.

Jonathan Elliott
Peter, let me come to you now: one of your jobs is inventing new ways of conducting surveys and making sure they're highly relevant. The importance of relevance here can't be overstated either. The practice of statistics really has to keep up in a world that is constantly changing, doesn't it?

Peter Lugtig
You're absolutely right, I think. Society is changing. So, if you think about the last 10 years or so, we've had ongoing digital change, and now, in the past few years, a transition to AI models. This means that we may need new statistics for policymakers, but it also means that the way we collect data is changing, and this is one reason why I think we need to both continuously adapt and change and innovate.

At the same time, we also want to keep the people in our surveys that are maybe not fully up to date with all the new AI models, right? So, we also want to keep some of the old time series of statistics that we've been collecting for years and years. We want to keep doing that as well. So, we both want to change and innovate, but we also want to keep the old.

Person on Street 1
It is just because how it is used, the information that they got from us. If it's for the purposes to improve maybe the product that they use, and it's always nice, but the downside could go that it's going to be used against you, basically, certain area of information.

Person on street 2
I am sharing my personal data now, just by being here, and it has always been an issue with personal data. And we should be careful with personal data and with modern technologies and AI tools. Things are going so fast and it's much more efficient, so you can collect much more personal data and, like, in a much shorter time, and you can do much more correlation.

Jonathan Elliott
And we're not going to talk too much about trust in this podcast, because, in a sense, every podcast we do on stats is about trust, it's so important in official statistics; but anxieties about trust often come or are accentuated when data about people gets very detailed and specific, when it's granular, zoomed in to small sets or areas, and there's a lot of innovation in this area.

Aleksandra, can you just explain for us how innovation, as well as giving us more detail, can also help with managing sensitive data and making it available safely, which balances privacy with usability.

Aleksandra Bujnowska
Innovations are very important in helping us to better and better meet the needs of the researchers, so to prepare micro-data that are useful for their research, but also, on the other hand, to protect the data and to be sure that we find adequate balance between the risk of disclosure, risk of identification, and the quality of the data.

There is a lot of new methods, a lot of discussion about facilities that can be used, safe facilities allowing researchers to do their analysis according to their needs, but at the same time making sure that this data are safe and well protected.

Jonathan Elliott
We've talked about micro-data without really telling our listeners what it is. We should do that quickly before we talk any more about it.

Aleksandra Bujnowska
Micro-data - these are sets of records where each record represents the person or household or business entity. This is a kind of a raw material for all statisticians. This is the basis for doing statistics. This is the basis also for doing research. The point is that micro-data – as they represent persons, business entities, respondents, in a general way – they are normally confidential. They used to be only available for statistical offices.

This is relatively new, that researchers can also be granted access to this type of data, under several conditions, under legal requirements and also technical requirements that we apply on the service. It's not for everyone to get access to this type of data.

Jonathan Elliott
This is much smaller groups of numbers, more zoomed in, more detail, and as a statistician and a researcher, you're closer to the subject, but you're not actually getting raw data as it's gathered on the ground. It's still carefully processed to protect the people that the data is about. What is that process exactly?

Aleksandra Bujnowska
Yes, of course, the micro-data the researchers receive, these are not the same micro-data as the statisticians use to produce statistics. The researchers get access to the data that are anonymised. So, the first thing is that never the micro-data contain any, as we call them, direct identifiers - so, names, addresses, anything that identifies anyone directly.

Person on street 3
I think today, when we are collecting the data, it's, it's faster than in the past, because now we have everything online, and you can make it from the whole world and not by a textbook and a letter, and you have to write something down. Now you can type it in, and in 5 seconds, it's all over the world.

Jonathan Elliott
Peter, you and your university colleagues might be thought of as typical users of microdata. How has it helped in your work? How has it changed the way that you and other academic researchers work?

Peter Lugtig
Micro-data are incredibly valuable for both researchers, but also, I think, for policymakers. So, if you take a general topic like understanding poverty in Europe, for example, as a general topic, there are lots of different types of studies that you can actually do with micro-data. So, for example, you can look at spatial variation in poverty. So what areas are where is more poverty? You could also look at differences in, for example, household situations.

So, what types of households are more likely to be poor. And then – potentially, at least – you could look into the combination of these two things, right? So, in what areas are what particular types of households poor? And this is – both for sociologists and for social science researchers generally – very interesting to understand; but then ultimately, I think, also for policymakers to then use this information in some way to, for example, set up some targeted policy interventions.

But here we run exactly into this data protection issues again, right? If you of course, know where someone lives, and you know their exact household situation, you, of course, run the risk that you maybe identify who this actual person is, and we want to prevent that at all cost.

Jonathan Elliott
Yeah, a subject for another podcast another day, I think − always a rich topic, confidentiality and trust, particularly in official statistics. Peter, you are in the very fortunate position of having 2 roles which complement each other. You're both an academic at the University of Utrecht, and you work with Statistics Netherlands as a researcher as well. Just talk us through how these roles complement each other.

Peter Lugtig
What I think is very interesting about collaboration between academia and official statistics is that in official statistics, there are a lot of data that are being collected, and there are a lot of problems or issues or potential innovations. And then at universities, I think we're typically very interested in these kinds of innovations, but these are not the kinds of institutes that collect vast amounts of data, universities themselves, right?

So, the collaboration, I think people at the university generally have an interest in problems, and they have time to study them, and then at the national statistical offices, they have a lot of problems or issues, and they have a lot of data. And I think, yeah, trying to, that's what I at least try to do, is try to kind of bridge these two.

Jonathan Elliott
Yes, a classic criticism when it comes to innovation in AI is that the innovators often are not the people who have to use the new tech, but you're different. You see the problems that the NSI's have, and you bring those back into the University of Utrecht and say: ‘Look, guys, this is what these institutes are having to deal with every day’.

This is what the messy real world really looks like. It must be very useful to be able to see how stats are used in a national institution, for example, and then design solutions you know are going to work in practice.

Peter Lugtig
Yes, I mean, ideally, it's kind of like a cycle, and you end up with solutions and there's new problems and new solutions. In practice, of course, it's sometimes just very complicated. So, I think in official statistics, some of the problems that we nowadays have in data collection, so, for example, combining data: there's lots of data out there, but somehow we have to combine all this data and make statistics that make sense to us, to policymakers, but also are consistent over time, for example.

And these are really huge challenges, I think, and there's no ready-made answer, I think, that universities have for this. The only thing that we can try to do is try to contribute to a solution. But I'm afraid we don't have a definite answer to all these problems straight away.

Jonathan Elliott
You've talked about universities and their relationship with national statistical institutes. What about Eurostat? How does that fit into this collaboration?

Peter Lugtig
There's a lot of collaboration at the national level, but I think at the European level, there are definitely initiatives that help facilitate collaboration. So, one of these is, for example, the European Master in Official Statistics. So, this is a network, really, of 32 universities across Europe, and each of these universities collaborates with one or more of these national statistical institutes in doing research, but mostly in training the future generation of statisticians who are now studying at the university.

This programme has been around for 10 years and has been really successful in delivering a new cohort of graduates who know about new techniques, know about data science, and are then ready to work in official statistics and develop new methodologies, for example.

Jonathan Elliott
One of the innovations that is particularly interesting is so-called smart surveys. Tell us why surveys need to be modernized and how they can be made better using things like smartphones and tech.

Peter Lugtig
So, some surveys that are conducted across Europe are done in the same way in every country, so that we get comparable statistics across European countries. Some of these surveys have always been in the form of a diary.

So, one example of that is that people keep track of their household spending for a certain amount of time, so they basically write down everything that they buy in the supermarket or other stores, and these kinds of data are then used for calculating inflation statistics, so, how much do prices actually go up and what kind of product category where a price increase is higher?

So, these data have been collected using diaries for a very long time, but this means that households have to keep these diaries. They have to keep track of all their spending for quite a long time, and this is quite cumbersome for households.

Jonathan Elliott
What was the solution?

Peter Lugtig
One innovation project that I've been working on together with other national statistical institutes, and this was also coordinated by Eurostat as well, a project to innovate here and to see whether we could actually use people's smartphones to do this in a better way, in an easier way. So, the idea that we piloted was that people basically take pictures with their smartphone of everything that they spent.

So, when you go to the supermarket, you, you get a receipt, and the receipt basically has all the things that you bought at the supermarket. So, the idea is that if people take a picture of these receipts, that receipt already contains a lot of information, and usually if people go to the supermarket, right, it's not one thing that they buy.

So, this, this saves them a lot of time in not having to do these diaries and not having to fill in all the products that they buy. So, we tested this, we piloted this, we developed software for this and infrastructure with the idea, of course, that in the future, this may then be rolled out to more countries in Europe.

Person on street 4
Artificial intelligence is, like, it's like a tool that can make everything, like, faster and more efficient than humans.

Person on street 5
But I also think the scariest thing about AI is just people using it without using their brains, you know, not, not…choosing AI over writing a paper for themselves and stuff like that.

Jonathan Elliott
Also, the need to modernise training and education. Sofia, how do you get your workforce ready for a world in which data science is going to be as important as literacy? It's going to be one of those fundamental functional things that nobody leaves school without. What hard skills now do you think your colleagues are having to grapple with that perhaps they didn't anticipate 10 years ago, and how are you preparing them for this new world?

Sofia Rodrigues
If we are talking about administrative data, data has a lot of mistakes and errors, and we have to use special techniques, machine learning techniques, in order to detect anomalies, in order to impute missing data. For doing that, we need to train people in that kind of new…of new ways of doing things. Data scientists are needed that understand different dimensions, dimensions of the problems.

So, we are investing in training courses designed specifically for our offices, that were performed by the one of the most significant universities in Portugal. And we are talking about artificial intelligence for official statistics, in which you can use real cases.

We are talking about data privacy and security as well, that was performed also by a university. And we also have, just like Peter, we have our own resident researchers that they really helped us transforming the culture and bringing expertise inside the house.

Jonathan Elliott
Yes, I mean that holism is critical here, because you were telling me that there is a need for cross-disciplinary, interdisciplinary, communication, soft skills, all sorts of things which statisticians, historically have not been noted to have a reputation for. They like to live in silos and concentrate on their areas of expertise. But that doesn't work anymore, does it?

Sofia Rodrigues
That's because we have too much time occupied by producing our mandatory official statistics, and that's why academia bring us this new way of doing things as well. Of course, the pressure to meet new needs that came from this changing society plays an important role, but then we have these academia liaisons to help us to think outside the box and produce other things with all the data that we have inside and integrate large and complex databases.

Jonathan Elliott
Now, every podcast has to have its little corner with AI and machine learning, and we are no different in the wrap cafe: where, Peter, do you think AI fits into official statistics? What are its biggest use cases? And how do you see AI being most useful in official stats?

Peter Lugtig
Well, we're only really at the start of figuring out how AI is going to be useful. I think one of the uses that I see is in data collection, so trying to, for example, process data in an easier, faster and perhaps better way. Another thing that I see is potentially in the presentation of statistics. So, for example, producing numbers, tables, statistics for policymakers, right?

If now policymakers have a question for, for a certain statistic, they have to go through a lengthy process to somehow get those statistics, and there's a lot of manual steps that have to…have to be

taken. And I think AI at least has the potential to partly automate this so that, again, we can maybe produce more targeted statistics for more targeted questions.

Jonathan Elliott
Absolutely - Sofia, isn't Portugal actively piloting AI applications at the moment? Can you just tell us a little bit about that?

Sofia Rodrigues
Yeah, we've got several examples, but I think the most suitable is this output checking that we must do whenever a researcher uses our data in the safe centre, because, well, Aleksandra already explained, we have different ways of make data available for researchers. The most sensitive one can only be accessed by a safe centre, and up to now, it's only physical. We get the researcher inside the room. He accesses all the data that we have available, of course, non-identified and anonymised.

But then we have to check as well the outputs that are coming from that work that the researcher is doing inside the safe centre. And the output checking is very manual and time consuming, and we are investing in the development of a system when we integrate large language models and prompt engineering techniques in order to predict that if the request of that researcher is secure. So, it will help us to reduce the time.

Jonathan Elliott
Fantastic. Thank you very much. Now we're running horribly out of time. I've got to move very quickly to our final questions, which are all about future directions and aspirations, and it is: what you would love to see in innovation, either in people, in tech, in investment, anything really.

You can have whatever it is you like, magic wand time for everybody. Sofia, let me come to you: in 10 years' time, or even in five years' time, what would be a dream situation for you if you could, sort of like, make your dream come true?

Sofia Rodrigues
I really believe that in 10 years' time, statistical offices can produce data fast enough to solve the needs of the society and the policymakers, and that can actually solve people's problems.

Jonathan Elliott
Peter, I'm going to ask you the similar question: what developments are you hoping for?

Peter Lugtig
I think there's really the potential that we can produce better statistics and also do that faster. The reason for that is that we get more and more data from different sources. So nowadays it's very common that you have all kinds of different data sources on, say, a topic like poverty. And the question is then more: how do we actually combine all these different sources to get statistics that give us a quick answer, but also a right answer to the question that we have?

Jonathan Elliott
Aleksandra - in 5- or 10-years' time, what would be a dream situation for you and your micro-data team at Eurostat?

Aleksandra Bujnowska
What researchers would love would be to have more sources of data available, and I think we are going in this direction, to try to provide new micro-data sets, new domains for research.

Jonathan Elliott
Fantastic. Well, we've run out of time, and that was really interesting. Thank you very much for that, Aleksandra. It's been a fascinating discussion, and it only remains for me to say a big thank you to our guests today: Sofia Rodrigues, Director at Statistics Portugal's.

Sofia Rodrigues
My pleasure. It was really my pleasure to be here. Thank you very much.

Jonathan Elliott
Peter Lugtig at Utrecht University, and also Statistics Netherlands - Peter, thank you so much for your contributions today.

Peter Lugtig
Yeah, thanks for having this conversation.

Jonathan Elliott
And finally, Aleksandra at Eurostat, thank you so much for sharing with us the wonderful and mysterious world of micro-data and the incredible power that it gives researchers. Thank you so much for joining us today.

Aleksandra Bujnowska
Thank you, Jonathan, thank you for listening to me.

Jonathan Elliott
It was wonderful. Thank you very much, everybody. If you've enjoyed Stats in a Wrap, don't forget to follow us on the streaming platforms, where the show can be found on Spotify, Apple and YouTube, and share our adventures with friends and colleagues.

And if you'd like to know more about the topics covered in this podcast and practically every aspect of statistics in the EU, you should make your way to Eurostat's amazing educational resource Statistics Explained, it's really worth checking out. And of course, come to the wrap cafe for the next edition, in which we'll be finding out what the numbers are telling us about young people and the labour market.

If you're about to start your first job, or you can remember what it was like if you're mulling the strange world that beckons beyond education, or trying to get the young people in your life onto the first rung of the ladder of their careers, this will tell you what the stats say about this sometimes painful, sometimes exhilarating, often baffling transition. Join us then, but for now, goodbye.