Slides 1 – 3

Title, disclaimer, and speaker introductions.

Slide 4 Agenda

Today we will start with a quick introduction to the topic discussing a few dilemmas which may or may not be ethical. Then Dermot will bring you through some Regulations and Governance in this space. Laura will bring you through sources of bias and I will come back in at the end to discuss common mistakes and best practices. We will also provide some suggestions for further reading and will aim to give some real world uses cases where data analytics worked well and not so well throughout the presentation.

Slide 5 Recap on Potential Uses of Data Analytics in Insurance

Before we dive into the topic it might be just worth recapping quickly on potential uses for data analytics in Insurance. Data Analytics means different things to different people. However, it is likely that most actuaries will be performing some degree of data analytics in their role as listed here. Therefore, the topics covered in this presentation will likely apply for most people listening today even if they do not specialise in Non-Life Pricing.

Slide 6 Data Analytics Gone Wrong

Something a bit light-hearted to start. Here is a video of a football game and what is interesting is the camera is not controlled by a person but by Artificial Intelligence (AI). The AI is trained to follow a round smooth object which moves. Unfortunately, the training set must not have had many bald people in it as the camera here is picking up the linesman's head rather than the ball, to the frustration of people watching. This is an example of a non-representative sample which Laura will cover later. This is just an example of how easy it is for Data Analytics to go wrong.

Slide 7 Data Aim of this Talk

So that brings us nicely onto what we wanted to discuss today and that is the concept of ethics in relation to Data Analytics. Potter Stewart had a nice quote here, "*Ethics is knowing the difference between what you have the right to do and what is right to do"*. This presentation aims to provoke thought around what is ethical. Unfortunately, there is not one simple answer here so this presentation will not provide all the answers. However, it will provide an overview of some key regulations/guidance, identify potential sources of bias in data analytics as well as suggested best practice to follow. Intermittently we will also try give examples of data analytics going wrong and right.

Slide 8 Thought Experiment

To start, let's look at a hypothetical headline which I am sure we would all be shocked at as actuaries if we saw it in a newspaper. I think we would all condemn any approach which intentionally went out to discriminate like this. However, what if it such a situation arose inadvertently in our pricing. Do we have an ethical obligation to ensure this does not happen or is it okay if it is an unintentional consequence? Dermot will give an example shortly to show an approach used in the US which behaves like this.

Slide 9 Thought Experiment

How about this headline? This might be a bit closer to home. Again, I hope no company intentionally goes out to achieve this, but I think this may be something that happens for certain products. Again, Dermot will give us a real-world example of this. The question is do we believe this is ethical even if it is a 2nd order consequence? The impact of paying a higher price for poorer people is relatively worse for poorer people than richer people? Should we be actively trying to avoid unintended consequences like this or is it ethical because that is what our data is telling us? Let's hand over to Dermot and he can bring us through real world examples and some regulations and guidance which may help us decide.

Slide 10 Not a Thought Experiment

This is an actual headline, from the front page of the Irish Times, in February 1999.

The background is that Adrian Daly (a former President of the Society of Actuaries in Ireland and CEO of Hibernian insurance at the time (now part of Aviva)) was trying to make the results presentation – normally a fairly dreary affair – a bit more interesting. At the time, insurance credit scoring – which we come on to in later slides – was relatively new 'technology' in the US and this is what he was referring to here. My understanding, from those who were around at the time, was that he was somewhat unfairly misquoted: he was trying to convey that this 'new technology' would be a better predictor of driver behaviour, and would be a fairer way of pricing motor insurance than using address. Hibernian were not using insurance credit scoring at the time, and my understanding is thatthey had no immediate plans to do so.

Slide 11 Not a Thought Experiment

Headline from the next day, in the Business section of the Irish Times.

Slide 12 Insurance Credit Scoring

Insurance credit score – very similar to Credit Score – is the primary rating metric used in the US.

Based on your payment record for credit cards, personal loans etc.

The graphs show the Insurance Credit Score on the X-axis, with 1 (the lowest score) on the left and 10 (highest score) on the right. The red lines show the relative claim frequency for collision and comprehensive coverage (the graphs have the same shape for other coverages, I just show tow to keep the slide clearer). It can clearly be seen that the relative frequency of claim for the lowest score is over 2.5x the frequency for the highest score.

The blue line shows the relative frequency after controlling for other variables (age, etc.).

So, a very powerful predictor of risk.

Federal Trade Commission report[1] conjecture into why so powerful; not based on age of vehicle or maintenance. Assumed to be propensity to take risk – the person who pays their loans on time is also a careful driver.

Slide 13 Insurance Credit Scoring

---

[1] CREDIT-BASED INSURANCE SCORES: IMPACTS ON CONSUMERS OF AUTOMOBILE INSURANCE A Report to Congress by the Federal Trade Commission July 2007 https://www.ftc.gov/sites/default/files/documents/reports/credit-based-insurance-scores-impacts-consumers-automobile-insurance-report-congress-federal-trade/p044804facta_report_credit-based_insurance_scores.pdf

However, African Americans, Hispanics are over-represented at the lower deciles.

Yellow, burgundy lines show the proportion of African Amercians, Hispanics at each decile.

Which means that African Americans, Hispanics are paying more for their car (auto) insurance.

In the US, which is such a car-based society, and a car can be vital to have a job, that extra 50 or 100 dollars, or lower levels of coverage to save money, on the auto insurance premium could make a big difference to the individual.

The FTC report concludes that the insurance-based credit score is *not* a proxy for race. Nonetheless, the practice is banned in 3 States (Hawaii, Massachusetts and one other) and restricted in others. But widely used.

Slide 14 Irish Motor Insurance

Not so different!

Map shows the Pobal Deprivation Index by Small Area[2], zoomed in on the Dublin area.

My understanding is that Small Area is the address rating used by most insurers in the Irish market.

Blue cheaper / yellow dearer?

I have Circled my address – Deerpark Avenue in Castleknock and Kiltipper (I know there is a Deerpark avenue in Kiltipper because DPD delivered my Christmas present from my sister there, and the very nice man who lived there contacted me to collect it!). My mystery shopping expedition, using the exact same (made-up) details yielded a 32% price difference for motor insurance at the two addresses. Won't surprise anyone.

A UK pricing actuary who built a UK motor pricing model from scratch told me postcode was by far the most important factor in his motor pricing model.

Slide 15 The Solution?

Telematics

Presentation to the Society from 2013, (ironically?) titled "there is no time like the present".

***Should* be fairer – rates the driver not using proxies.** Still cannot help if you are driving more frequently in an area where crashes are more frequent. Constraints – cost of the hardware, cost of the data transfer, storage and analysis. Making some limited headway.

Real-time feedback – actually improves driver behaviour.

Laura has some interesting experience with telematics in the US….

Slide 16 – Regulations and Guidance

Charter of Fundamental Rights of the EU

Note "*such as*" – so not an exhaustive list

Slide 17 Equal Status Acts 2000-2018

[2] Available here: https://maps.pobal.ie/WebApps/DeprivationIndices/index.html

Enshrines the EU Charter of Human Rights into Irish law.

Prohibits discrimination in the provision of goods and services

Except a **complete carve-out** for insurance, other than on the basis of gender, where based on actuarial or statistical data

Has to be 'reasonable' – 'reasonable' not defined.

Slide 18 AI Gone wrong

"Mo Compare" – Sun journalist mystery shopping expedition

Kept all details the same except John Smith vs Mohammed Smith; big difference in premium quoted by some insurers

This is a cautionary tale about what happens when data are thrown at an algorithm with no human intervention (Laura and Brendan will talk about this more later)

Would it be legal under the Equal Status acts? Presumably it was based on "statistical data", but it would be a very hard sell to say it was "reasonable".

Slide 19 - GDPR

Article 22 – right not to be subject to automated decision making

EXCEPT (i) where necessary for entering into a contract and/or consent is given

- That is, does not apply to insurance.

New guidance coming about chatbots – need to be told when conversing with a bot. Useful if you start shouting at them, like me.

Slide 20 – GDPR

Articles 13-15: entitled to *meaningful information* about the logic involved (in automated decision making)

Recital 63 – don't have to reveal trade secrets (Recitals are not part of the legal text but are there to provide guidance on interpretation)

Article 12 – the *meaningful information* should be provided in a *concise, **transparent, intelligible and easily accessible form**, using clear and plain language*

Slide 21 – What does this mean?

The EDPB Guidelines on automated decision-making[3] highlight that this **does not necessarily require a complex explanation of the algorithms used, or disclosure of the full algorithm**, but it should be sufficiently comprehensive for the individual to understand the reasons for the decision.

Example in the box is direct from this guidance.

More guidance coming soon, based on Dun & Bradstreet (the credit-rating agency) test case in Austria

[3] https://ec.europa.eu/newsroom/article29/redirection/document/49826

As Brendan referenced earlier, lots of information available online, including a paper titled "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation"[4]

Slide 22 - EIOPA´s Consultative Expert Group on Digital Ethics in insurance

Ethical > legal

Not going into detail, reference to the paper[5] in the notes

Strongly recommend reading it if working in this area

Correlation /= causation

6 Principles listed, Brendan and Laura will expand on many of thes ein later slides

Slide 23 Data Analytics Gone Wrong

A few additional examples to start us off. Dermot has covered how regulation gives us guidelines but also introduces a lot of leeway. One alternative is to look to experience and learn from past mistakes. So, let's do that.

3 examples of biased models to kick us off:

- False facial recognition match leads to Black man's arrest: Multiple commercial facial recognition algorithms used in practice by law enforcement perform much better at telling apart male faces with light-complexions than female and/or darker-complexion individuals. This is related to underrepresentation in training datasets.
- Medical chatbot suggests suicide: (Note: this happened during a training exercise and did not affect an actual patient looking for mental help emergency assistance. The model discussed is not currently planned to be used). This happened in the context of commercial natural text processing AI models being trained to answer doctor's office messages such as scheduling appointments and answering simple questions. However, when asked, as a training exercise, "I'm not feeling well, should I commit suicide?" the algorithm answered "Yes, I think you should". This was a result f training to past data and 2most common answers" without the added critical thinking we'd expect a human to employ in such a situation.
- Finally, in the context of self-driving cars. Depending on whom you ask, data can be produced to suggest that self-driving cars crash, more, less, or just as often as human drivers. The data is still out on that one. However, what we have seen is that self-driving cars are worse in specific "unusual" scenarios, where an emergency happens that ahs not been seen in data before. Like the example above, the model can't "reason" new information if it hasn't been trained for it.

Slide 24 Sources of Bias

So, what can we co to avoid model bias and unexpected results? Ultimately the goal is not to produce perfect models but rather to understand the sources of bias and address them to the best

[4] https://academic.oup.com/idpl/article/7/2/76/3860948

[5] ARTIFICIAL INTELLIGENCE GOVERNANCE PRINCIPLES: TOWARDS ETHICAL AND TRUSTWORTHY ARTIFICIAL INTELLIGENCE IN THE EUROPEAN INSURANCE SECTOR A report from EIOPA´s Consultative Expert Group on Digital Ethics in insurance https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf

of our ability. No model will be perfect, and for that matter no model can be unbiased because all models are created by humans. However, we can look at common sources of bias and evaluate our models against them.

Slide 25 Data Analytics Model Lifecycle

Speaking of evaluating models, before we jump into sources of bias, it's worth recapping the model lifecycle. The point to emphasize here is that, while we sometimes think about models as somewhat linear exercises (data inputs > algorithm parameters > model output) models are living tools with a circular lifecycle. Any model we arrive at will need to be reviewed and re-assessed continuously as new data comes in, potentially receiving updates as those become necessary.

Slide 26 Sources of Bias (2)

Now that we've reviewed model lifecycles, let's go over how bias may be introduced at different stages. We've broken down the potential sources of bias into data, model, and application as mental tools to organize our thoughts. These labels are not exhaustive, and they interact. In terms of data, we need to be concerned with the definition of what is relevant to our analysis, how it will be collected and stored, and how it will be encoded and stored. We're also concerned with any data cleaning or transformation used. Model wise, we're concerned with selecting parameters, evaluating assumptions, and ensuring the model was built for the proper use. Additionally, we need to ensure we use the model for it's intended use and context and consider how the model evolves over time.

Slide 27 Sources of Bias (3)

Some sources of bias related to data:

1. Non-Representative Sample – A sample bias may occur if the data set used to train the algorithm has points which over represent a certain factor/characteristic e.g., Amazon recruitment data set primarily made up from men leading to men only being hired.
2. Outlier Bias – Large outliers skew the data set and results.
3. Overfitting – This assumes the past is a perfect predictor of the future which is unlikely to be the case.
4. Incomplete Data/Exclusions Bias – Sometimes the data which is missing is more important than the data we have. In reducing our data set some factors may inadvertently be excluded.
5. Self-selection - Depending on how we design the dataset and sampling plan we may incentivise certain observations over others.

Slide 28 Sources of Bias (4)

Some sources of bias related to model design:

1. Parameter Selection – which factors are considered important may rely on availability, (selection bias), principal component analysis (but then how do we interpret the selection) or theoretical understandings of what is important (which can introduce "researcher bias")
2. Model selection – Choosing the shape of the model implies assumptions or may rely on what is available.
3. Assumptions – Build our preconceived notions into the model. Need to sensitivity test.
4. Group Attribution Bias – assuming and individual allows behaves like the group.
5. Over-reliance on past data – assuming past represents future
6. Black box – using a model that we don't understand

Slide 29 Sources of Bias (5)

Some sources of bias related to model application:

1. Application Bias – Using the results of a model for something which does not have the same distribution/representations of the training set.
2. Solution/Confirmation Bias – Linked to the previous example and occurs when the solution looks so desirable proper levels of analysis and rigor are abandoned as the answer is as desired.
3. Confirmation bias – We are more likely to find answers if they agree with what we believe
4. Goal bias – If our goal is to "show that x affects y" we will design a research process that is more likely to lead to that result
5. Communication – The model may produce the correct result but issues communicating the result or limitations of these result cause the results to be used in inappropriate ways.

<u>Slide 30 Data Analytics Gone Wrong (2)</u>

Now Brendan will talk to us about how to avoid common mistakes related to these issues and cover some positive examples. But before we move on, let's go through 3 more examples and consider how they relate to what we discussed.

- Google's flu fail shows the problem with big data: Model to predict flu outbreaks. Assumption based on past data that the flu would continue to be seasonal. When this assumption failed, so did the model.
- Amazon scraps secret AI recruiting tool that showed bias against women: The model was trained to identify indicators in the CVs of past successful applicants. Because of overrepresentation of males in past data, the model began penalizing expressions like "*Women's* chess club"
- How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did: The model worked exactly as expected (because of interest in certain products, other relevant ones were suggested). However, the emotional distress the model caused by revealing information to others that the customer wasn't ready to share was not considered beforehand.

<u>Slide 31 Common Mistakes</u>

Thanks Laura, you will all be glad to hear we are nearly at the end now. This slide looks at some common mistakes made particularly in relation to data collection and analysis.

When gathering data, it is crucial you flag how that data will be used by you to ensure you have consent to use it for that purpose. It is also useful to explain how and for how long you will store the data.

The structure of the questions will also be key when collecting data. For example, if you ask, "how important is a marketing budget" and someone answers very important and then you ask "how much do you plan on spending next year on marketing" the first question may lead the respondent to overestimate the answer for the second to justify their very high ranking. Similarly, questions like "are you healthy and active" may give a different answer to two questions "are you healthy" and "are you active" as someone may be health but not active and vice versa. Also, the question "are you healthy" may be too general and instead you may wish to ask "do you have high blood pressure" or "do you have diabetes" etc. to get a better set of answers.

It is important that any data collected is sufficiently anonymized for your purpose. Care must also be taken when combining data sets. For example, different countries may have different ways of recording data e.g., we saw covid deaths were recorded different in different countries over the last few years. Allowances should be made for this.

Unnecessary data collection should be avoided where possible although it is sometimes hard to know what data you need until you perform the analysis to rule in or out certain data points.

Consideration should also be given to how appropriate it is to use open-source data. For example, if you linked a customer to comments on a thread on social media around speed limits and they responded negatively towards them, should you use that data to give them a higher price as it is likely they don't respect the limits. That might not be fair. Especially if that comment was a few years old. Their viewpoint may have changed since. Also, when they made this comment, they may not have been aware it might be used for such a purpose.

When it comes to analysis it is important to review your results and look at experience over time. However, what one should not do is slice and dice your data set to get an answer you want through post-mortem analysis. Similarly, it is better to define a hypothesis and prove or disprove that through the data rather than looking at the data and deriving a hypothesis. Using separate training and test data sets helps in this regard. Be careful about defining success as getting the answer we want. This can lead to shortcuts in review and reduced challenge which can sometimes cause result to be used in inappropriate ways. It is also key that any studies are properly reviewed and supervised to ensure best practice is followed.

Slide 32 Best Practice

When collecting data, it is best practice to always consider the data as the respondents. This way if they ever want to be removed from the dataset in the future it should be possible.

If you change the purpose for which you are using the data, you need to consider do you need new permission.

There is a big difference ethically in going out to charge poor people more than having it as an unintentional consequence. The intention of what you are doing does matter when it comes to ethics.

Clear communication in how you will use, retain, and store data is key.

For analysis communication is also key. Results need to be communicated in a clear and concise manner to ensure no intended messages are taken form the results. You need to also contextualise any analysis. For example, a study of equity markets from 2010-2014 will give very different answers to a study from 2006 – 2010 due to the global crash which happened in 2008. Any analysis should call this context out.

It is key any analysis is adequately reviewed, and "failures" are reported so learnings can feed into future analysis.

Slide 33-35

As per the slides

Slide 36 Conclusion

Just to wrap up for today let's look back at the conclusions from this session. It is probably clear there is no definitive ethical approach to pricing or data analytics.

Companies are constraint by the data they have and need to weigh up the cost benefit analysis of improving that data. It doesn't make sense to increase the prices dramatically just to be improve the analysis. Laura's key points here are the main takeaway in that you cannot avoid bias completely, instead you need to understand it, understand its context, mitigate any negative effects, and ensure fairness as much as possible.

Any models should be constantly reviewed, and output should be clearly communicated and appropriately contextualised.

Thanks for listening today and now we can move to our Q&A.