# MODELING OF DATA FROM IMPERIAL COLLEGE COVID BEHAVIOUR TRACKER TO IDENTIFY MENTAL HEALTH RISKS OF PANDEMIC

January 23 , 2022

James Pheby

# Summary

- Following warnings from the WHO and the UN that the pandemic has increased risk factors for suicide, the UK government asked our consultancy firm to immediately identify groups most at risk of mental health problems. These will be targeted in upcoming awareness campaigns

- To do this, we analysed data from the Imperial College London/YouGov Covid-19 Behaviour Tracker Data Hub -- a global survey of hundreds of thousands of people conducted throughout the pandemic

- We then applied various techniques to model whether a respondent would report  that they had felt depressed during the previous week, given their responses to the other questions.  The model was then analysed to find the factors most correlated with mental health problems.

# Main recommendations

- The government should work with job centres and student services at universities to raise awareness of symptoms and who to contact for help
- It  should run television, Facebook and Instagram campaigns targeting young people, particularly young women
- It should also launch a campaign, once it is safe, to encourage people to get out of the house and socialise. This could include financial incentives to eat out, or join clubs
- Inform Italian counterparts of the alarming rates of mental health issues reported by those living in the country

# Outline

- Business Problem
- Data Collection, Analysis and Modelling
- Results
- Conclusion
- Next Steps

# Business Problem

- To identify the groups most at risk from mental health problems in the pandemic for an upcoming UK government awareness campaign
- This involves creating a strong predictive model, and also inferring from the workings of the model which factors most increase the odds that someone will suffer mental health issues
- The National Health Service (NHS) should be able to use our insights to flag up any high-risk individuals, without having to directly quiz them about this highly-sensitive issue
- Any findings of international interest should be shared with the relevant governments

# Data and Methods

## Data

- **Data drawn from the Imperial College London/YouGov Covid-19 Behaviour Tracker Data Hub**

- **The dataset comprises results of a 450+ question survey completed by hundreds of thousands of people during the pandemic from 29 countries around the world**

# Data

- **We used the 11 countries that provided the most comprehensive responses. These were: Australia, Canada, Denmark, France, Germany, Italy, Norway, Spain, the UK and the US**

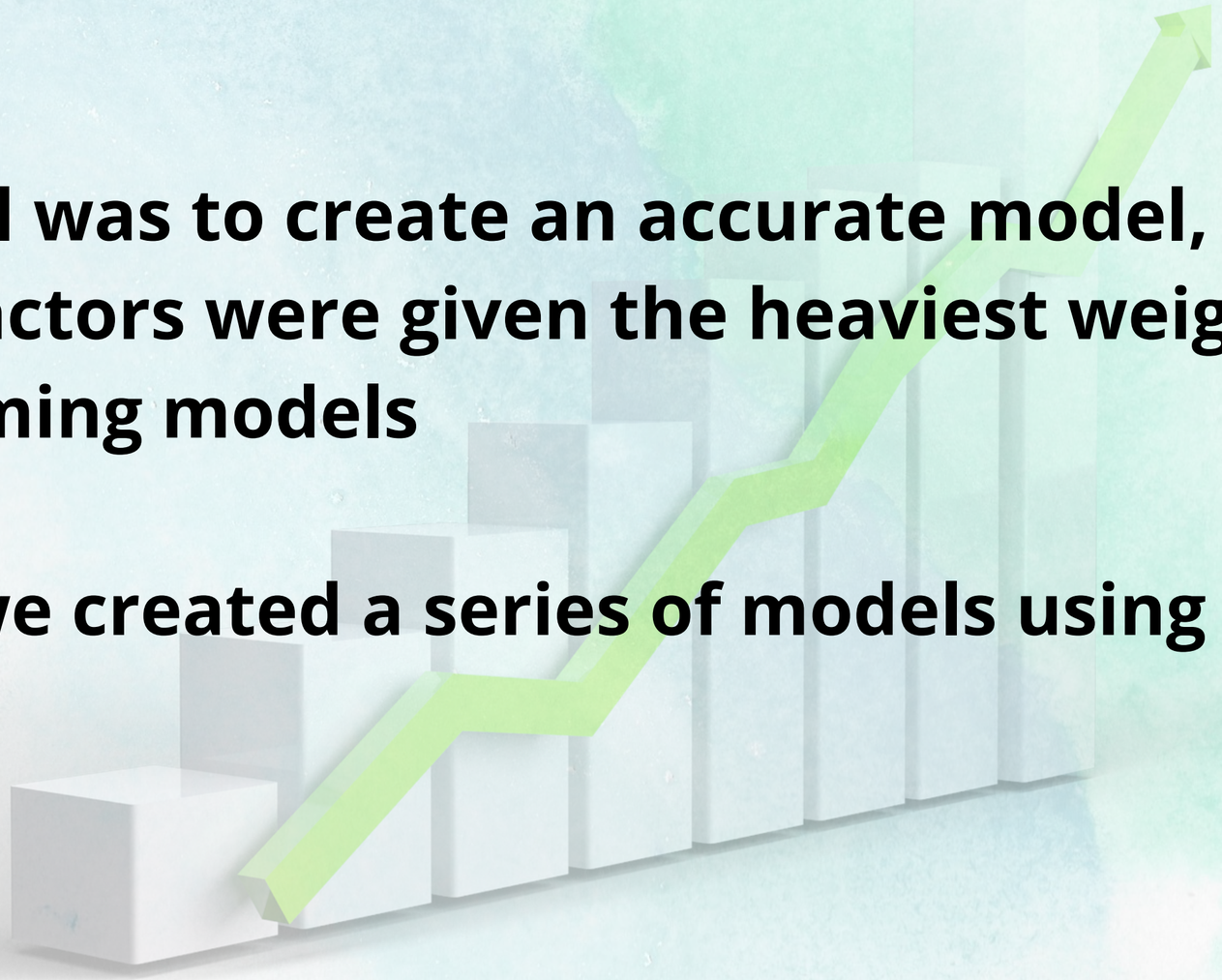
- **We also dropped questions that were repetitious or infrequently answered**

- **This left us with 128,117 records and 44 questions with which to create our model**

# Methods

- **The end goal was to create an accurate model, and to find out which factors were given the heaviest weighting by the best-performing models**

- **To do this, we created a series of models using different techniques**



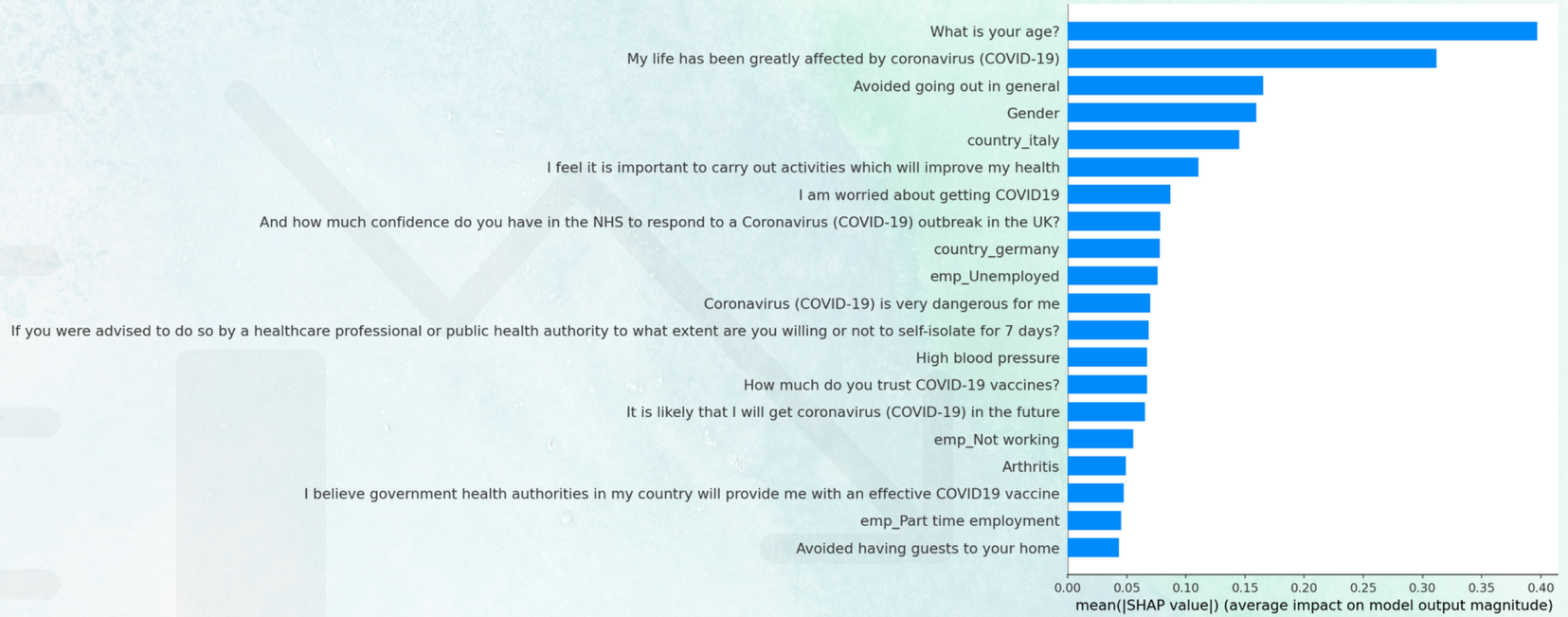
# Methods

- **Each model was put to the test with previously unseen data, and its performance evaluated in terms of how accurately it predicted whether a respondent would answer 'yes' or 'no' to the question of whether they had recently suffered from depression or hopelessness**

- **Each model was tuned to maximise its performance, and the best two models, MODEL A and MODEL B, were selected. These were then analysed to find the most important factors across both**

# Results

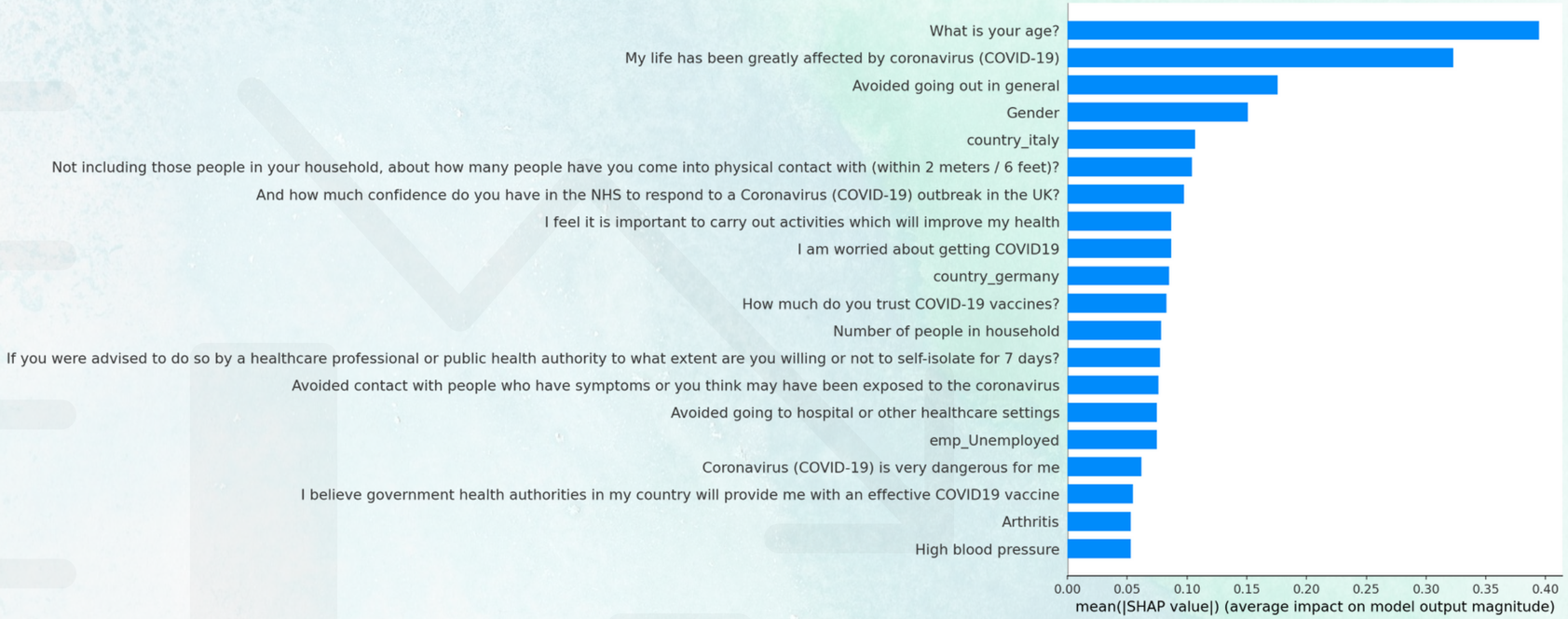
## Importance of Features in MODEL A




- **Being personally affected by the pandemic and avoiding going out were strongly correlated with how the respondent answered the question 'have you felt down or depressed in the past week'**
- **Age and gender were also key factors**
- **Being a student or unemployed had a strong bearing on reporting mental health issues**
- **Living in Italy was a strong predictor of feeling down or depressed**

# Results

## Importance of Features in MODEL B



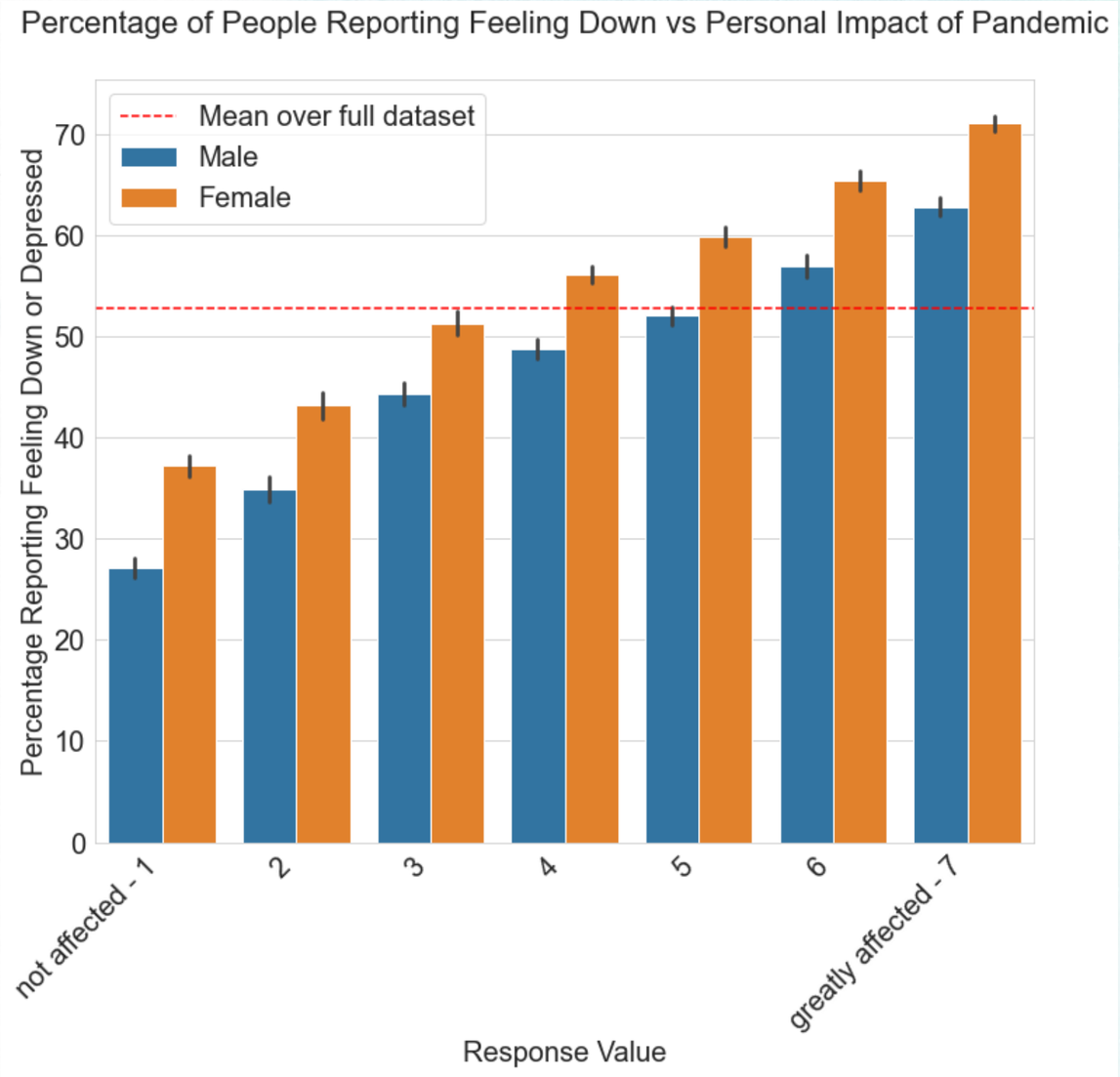

- **Despite Using different techniques, MODEL B produced almost identical results to MODEL A in terms of its most important features**

# Results

## Importance of features in MODEL B by individual record

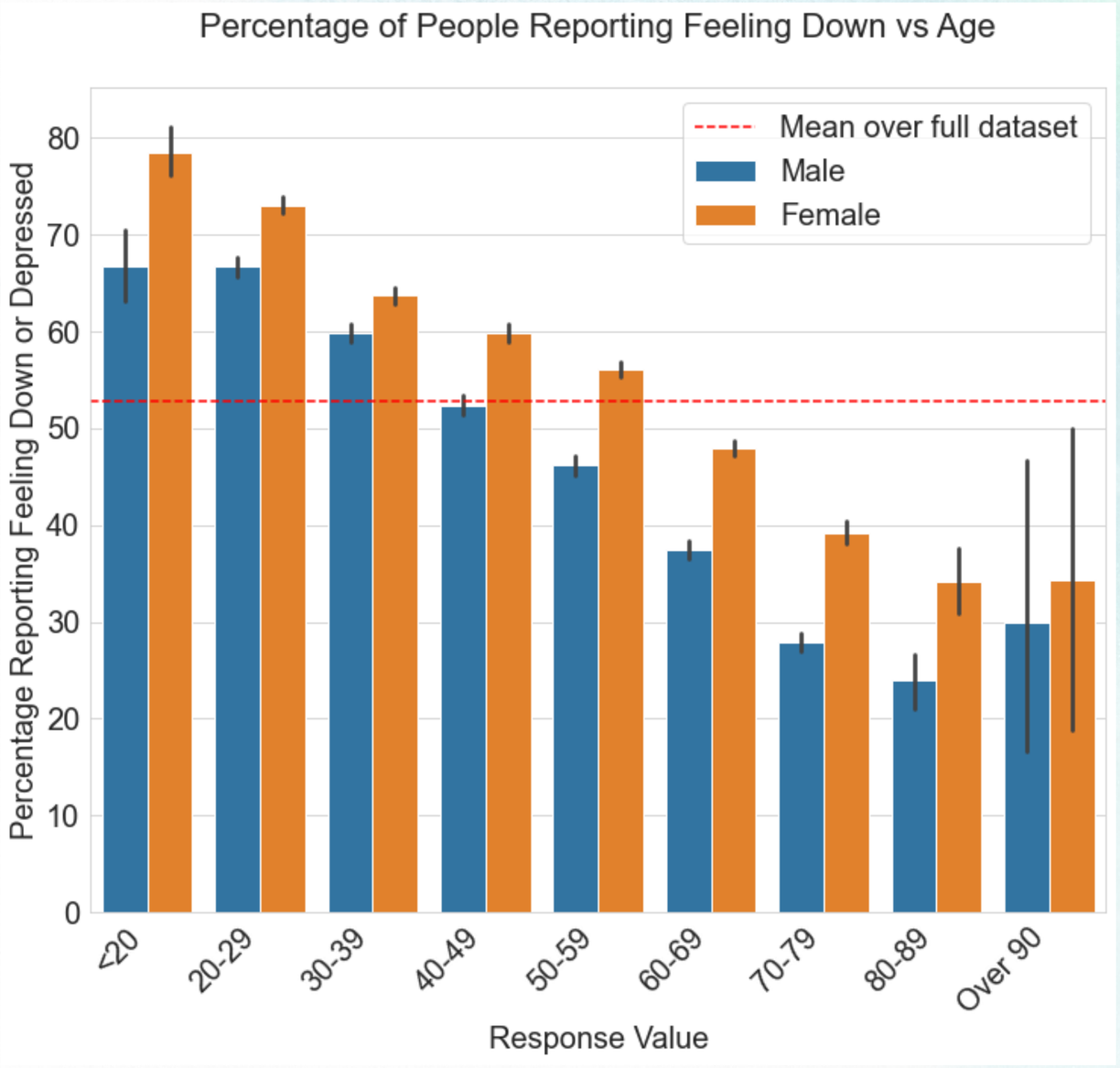

- **Being young and being female were both strongly correlated with reporting mental health issues**
- **People who avoided leaving the house and met few people outside the home were more at risk**
- **Being unemployed or a student raised the chances of reporting feeling down**
- **Previous conditions such as arthritis and asthma and were also correlated with being down or depressed**

# Results



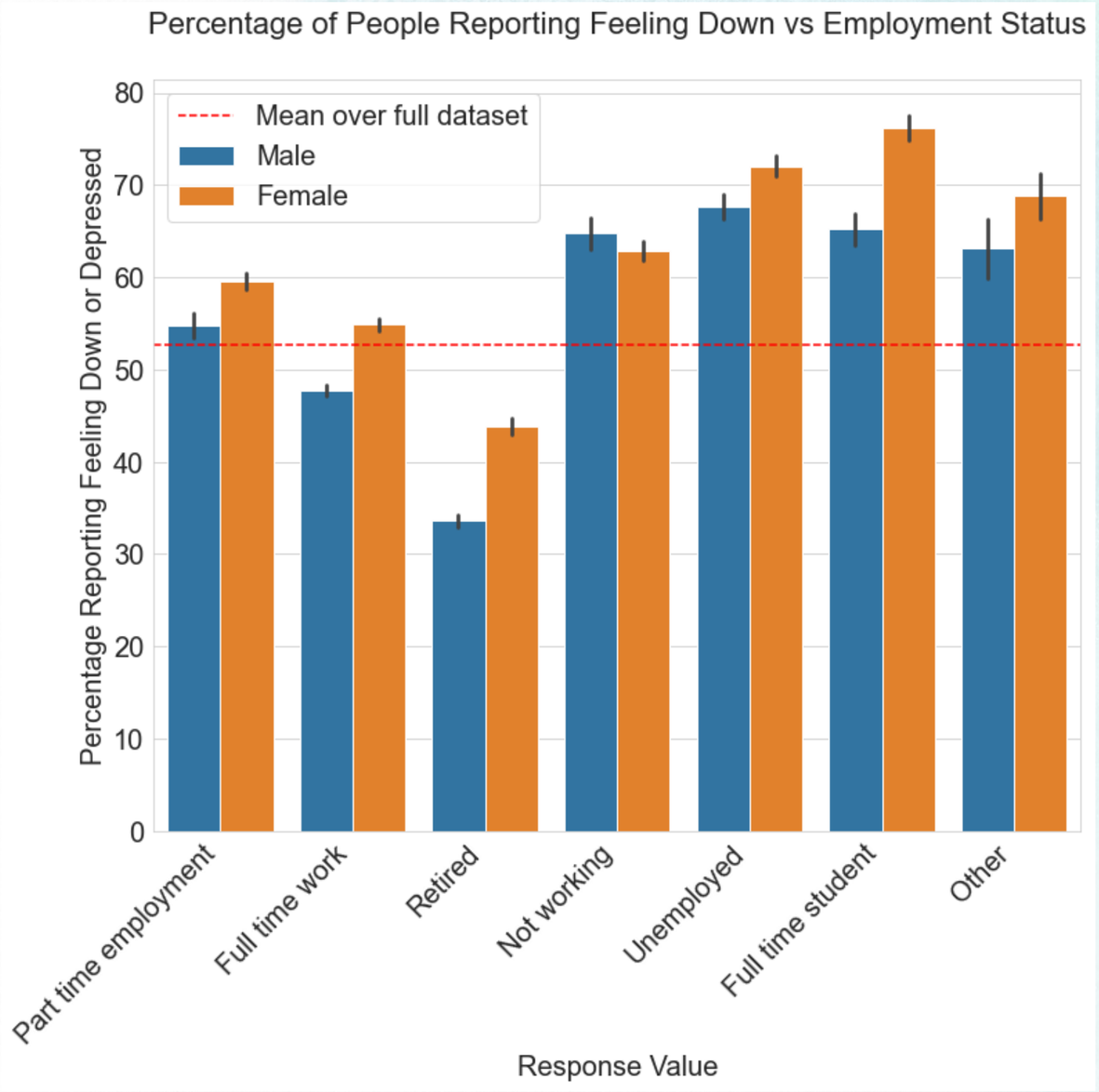

- **People whose life was most affected by the pandemic were more likley to report mental health problems**
- **Taking all other factors into account, MODEL A found that every one unit increase in the response value raised the odds of experiencing mental health symptoms by 5.8%**
- **This suggests the pandemic was a major factor for those feeling down or depressed**

# Results




- **The proportion of people reporting feeling down or depressed decreased with age, except in the case of the over 90s**
- **MODEL A found that every extra year of life decreased the odds of reporting depression by  0.7%**
- **Females were more likely than males to report mental health struggles in every age bracket, a trend that will be seen across every feature**
- **MODEL A found that being female increased  the odds of reporting depression by 30%**

# Results

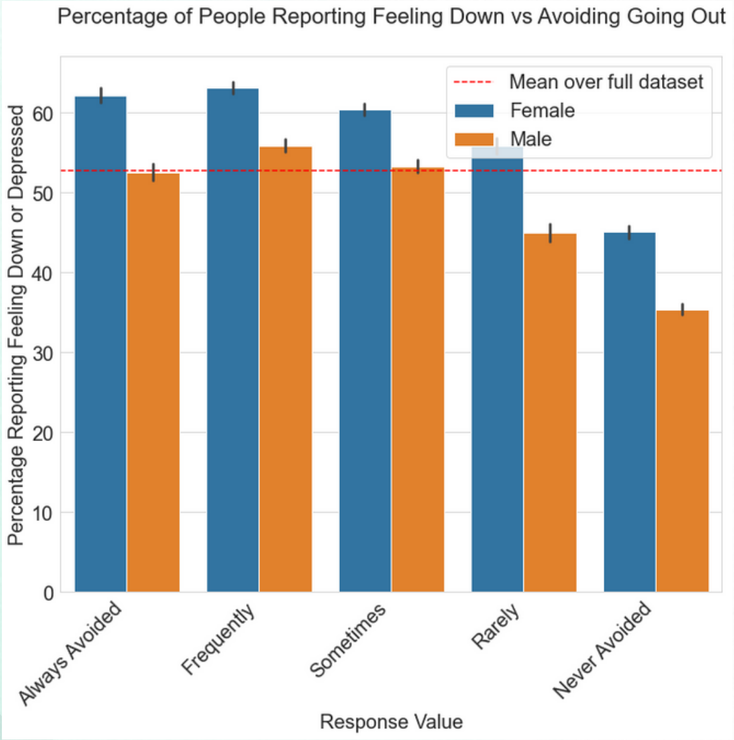


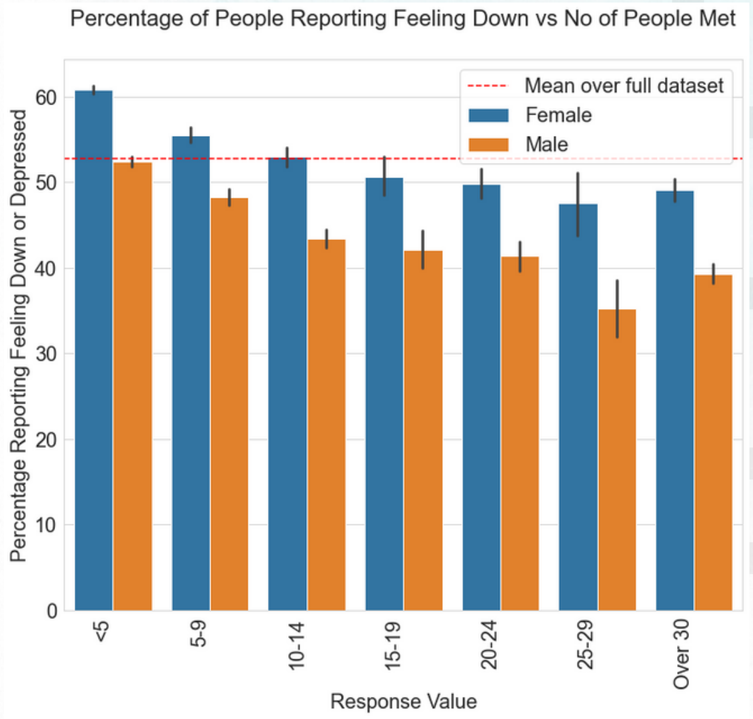

- **Students and people out of work were much more likely to report mental health issues**

- **MODEL A found that being unemployed increased the odds of reporting depression by 45%, while being a student increased the odds by 17%**

# Results

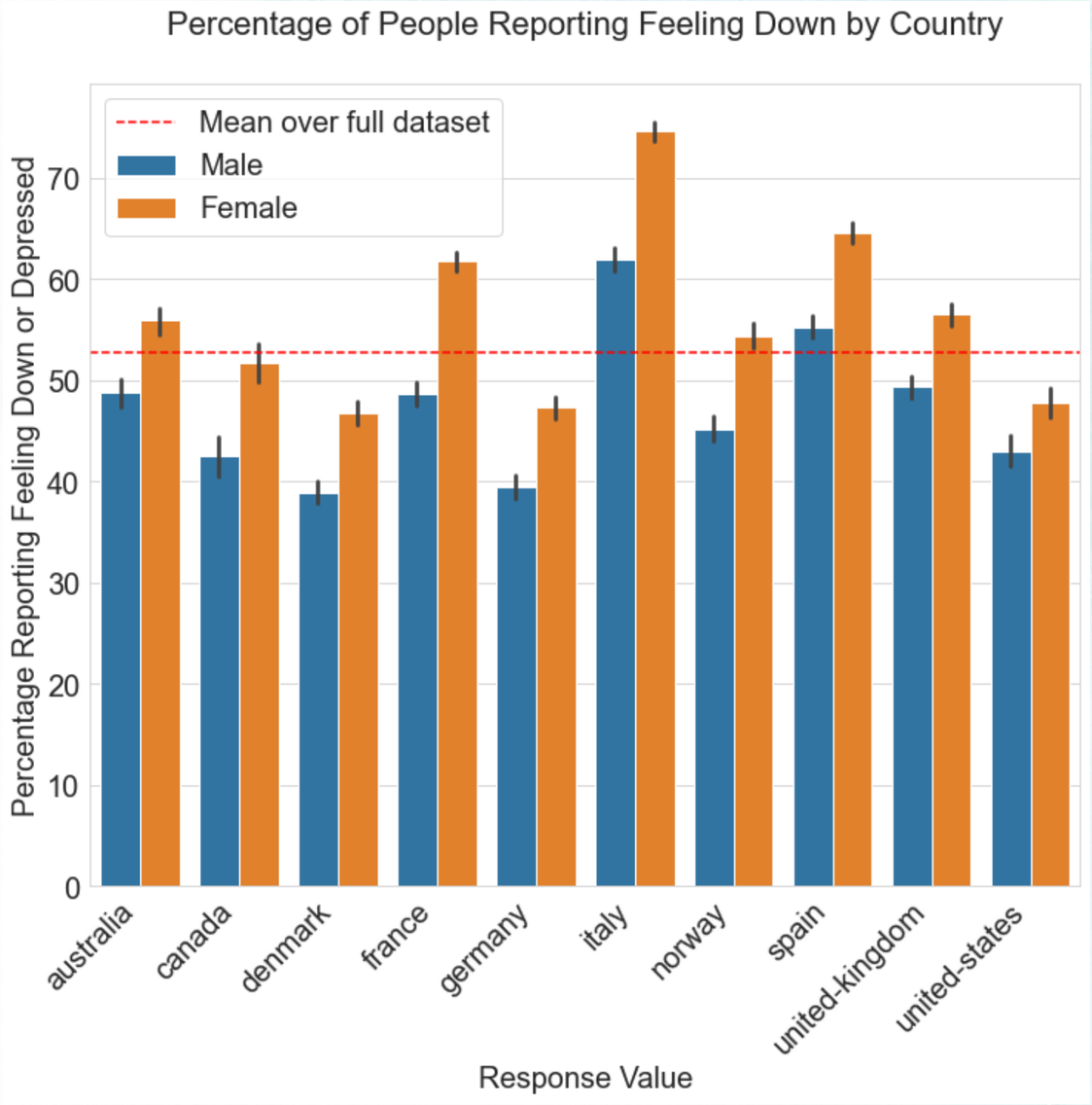
- **People who avoided going out were more likely to report mental health issues. MODEL A found every unit increase in willingness to go out reduced the odds of reporting depression by 5.6%**




- **People who met few people from outside their household were most likely to report feeling depressed. MODEL A found every extra person met reduced the odds by 0.04%**

# Results




- **People living in Italy were particularly prone to feeling depressed**
- **MODEL A found that living in Italy raised the odds of reporting depression by 53%**
- **Around three quarters of women living in Italy complained of feeling down or depressed in the previous week**

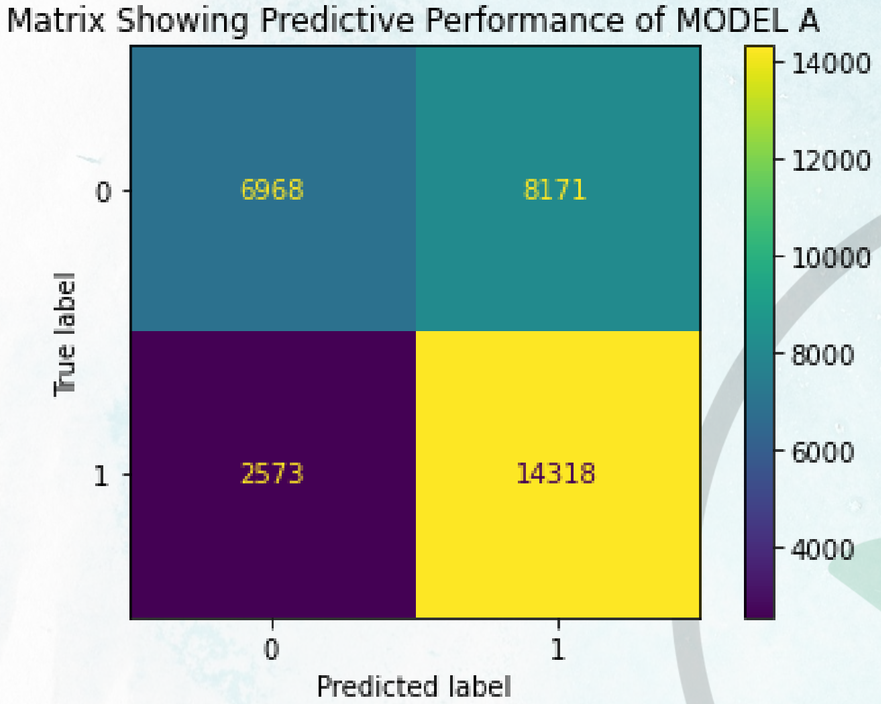
# How reliable are the models?

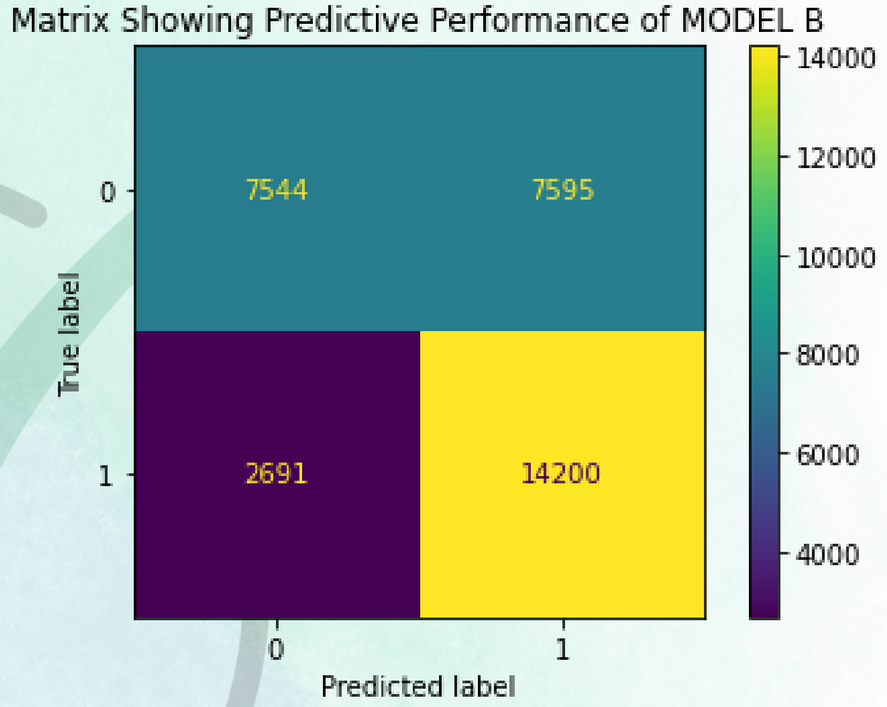
- The models were both around 69% accurate, meaning they correctly predicted whether someone would report feeling down or depressed around 7 times out of 10

- In reality, we are more concerned about catching those who are suffering mental issues than incorrectly  diagnosing those who aren't. Reaching out to someone who isn't experiencing symptoms is less problematic than missing someone who is

- By adjusting the parameters of both models, we were able to catch around 85% of those who were suffering symptoms, but we had to sacrifice some overall accuracy. We have to tolerate misdiagnosing 50%, or more, of those who don't have symptoms, until we have the data to make a stronger model

# How reliable are the models?




- **MODEL A, on unseen data, caught 14,318 of 16,891 cases of people reporting mental health problems, a rate of 85%**
- **It correctly identified 6,968 out of 15,139 who did not suffer from issues, a rate of 46%**
- **The overall accuracy was 66%**

- **MODEL B on unseen data, caught 14,200 of 16,891 cases of people reporting mental health problems, a rate of 84%**
- **It correctly identified 7,544 out of 15,139 who did not suffer from issues, a rate of 50%**
- **The overall accuracy was 68%**

# Conclusion

- **Students and unemployed people are highly vulnerable. The government should work with job centres and student services at major universities to raise awareness of symptoms to look out for, and who to contact for help**

- **Young women are also at risk. The government should run television, Facebook and Instagram campaigns targeting young people, but particularly young females**

# Conclusion

- **People who leave the house less often and don't meet many people outside of their home are more likey to report feeling down**

- **Once the pandemic has subsided, the government should find ways to incentivise getting out and about and meeting people. Perhaps by subsidising eating out or joining a club**

- **Living in Italy was a strong predictor of mental health problems. This finding should be shared with the Italian authorities handling the response to the pandemic**

# Next Steps

- **Tailor a new questionnaire focussed on the key factors identified in the models. Each factor needs to be explored in more detail to produce a more granular model**

- **Once we have a high-performing model, we can then produce a concise survey for the NHS to offer those who use its services. This can be used to flag up individuals at high risk of suffering mental health problems without having to directly question them about this highly-sensitive issue**

- **Although we found that there was a strong link between those whose lives had been affected most by the virus and mental health problems, we need to analyse pre-pandemic data to identify pre-existing predictive factors. This will help us more accurately model whether the mental health of certain groups has been disproportionately affected by the pandemic**

# Next Steps

- **There were sparse responses to questions about working from home, which made them unusable. We should conduct our own survey on this issue, as it is potentially an important factor that has only arisen during the last two years**

- **There appears to be evidence from the models that pre-existing conditions such as arthritis, HIV and asthma have a significant effect but there were too few positive cases in the dataset to make a definitive statement. We should encourage charities dealing with those diseases to conduct surveys. The responses would help us make a more robust model**

# Next Steps

- **The responses on vaccine uptake and hesitancy were also a little inconsistent as different countries were at different stages in their vaccination programs over the duration of the survey. Now countries are on more of an even footing, we should conduct a survey focussed on issues surrounding access and attitudes to vaccines and mental health**

- **We need to further analyse existing data on which groups are more reluctant to admit feeling depressed. Our data is anonymised, but it is still likely that some of those suffering from problems did not report it. This will obviously limit the effectiveness of our model**

**Email:** jpheby93@gmail.com

**GitHub:** https://github.com/jimp93

**Medium:** https://medium.com/@jpheby93/