

Our first question asks about how gender and age correlate with (and so presumably influence) the prevalence and severity of the disease in patients. We begin by simply tabulating the number of male cases against the number of female cases (c.f. 1A). We see about an 8% surplus of additional female cases. This seems significant but is by itself not very informative. To get a better sense of things, we break down the numbers further by age group (c.f. 1B). Once we do this, we see a mixed picture. In certain age groups, female cases seem significantly more prevalent, whereas the opposite is true for other age groups. In order to better evaluate the significance of the discrepancies we observe we compute the ratio of male cases to female cases (c.f. 1C). Here we observe that there is an overwhelming bias towards female cases in ages above 80, a smaller bias in the same direction for ages 40-60, and a smaller bias still in the opposite direction for ages below 40.

There are many possible hypotheses that could explain the gender-age related discrepancies we have observed. For instance, it is well-known that women have, in general, longer life spans. Hence the overwhelming bias towards female cases in the elderly is ostensibly explained by the simple fact that most people who survive to that age are women. An alternative explanation is that older women are perhaps more likely to require assistance with everyday tasks in their later years, and so are perhaps more likely to reside in nursing homes, where the brunt of the outbreak was so sharply felt earlier this year.

To eliminate the effect of the gender-imbalance in old-age homes, we restrict ourselves to the cases which are caused by ‘proximity’ (c.f. 1D and 1E). When we do this, we see that the discrepancy is significantly diminished but is not gone. It therefore stands to reason that both of our explanations are necessary to explain the effect.

We also tabulate the recovery rates for both men and women across the ages (c.f. 1G). Here the pattern is clear and undeniable. Women recover more often than men across the board, in every age range.

For our second question, we examine how age and gender correlate with the method of contracting the disease. Tabulating age against method of contraction (c.f. 2A, 2B and 2C). We see that older folks contract the disease almost exclusively during outbreaks (presumably the ones in nursing homes), younger folks contract the disease mostly through proximity (presumably due to lack of precautions during outings), and that travel is an important but not primary cause of transmission of the disease for middle-aged persons (presumably due to the high rates of business travel for persons of such ages).

When we do the same thing for gender (c.f. 2D, 2E and 2F), we see that men tend to contract the disease at higher proportions from travel and proximity as opposed to outbreaks. We attribute this to the risk-seeking behaviour that males are known for.

For our third question, we examine how economic activity (c.f. 3B) and class (c.f. 3A) correlate with the case numbers. The results are mixed. On the one hand, we do see something of a positive correlation. The most economi-

cally active regions do seem to have the higher case numbers, but it's not clear cut. Furthermore, we could not adjust for population and population density. Hence it is quite possible that the patterns we observe are better explained by population size and density.