1. One possible reason is that many COVID-19 deaths have not been officially recorded or were attributed to other causes. Another possible reason is that some countries may not have the resources or capacity to accurately track and report COVID-19 deaths.

2. To investigate the prevalence of "long COVID," I propose using a combination of surveys and electronic health record (EHR) data. The surveys would be administered to individuals who have recovered from COVID-19, asking about any lingering symptoms they may be experiencing. The EHR data would be used to identify individuals who have been diagnosed with COVID-19 and track their health outcomes over time. By analyzing both sources of data, we can identify patterns and risk factors associated with long COVID, such as demographics, comorbidities, and severity of initial infection. This approach could help identify target populations that may be more susceptible to long COVID and inform the development of interventions to mitigate its impact.

3. Smoking damages the cilia that line the airways, reducing their ability to clear mucus and debris, which can lead to inflammation and narrowing of the air passages. Additionally, smoking causes damage to the walls of the alveoli, reducing their elasticity and ability to exchange gases effectively.

4. i. Biological systems express rhythms for two reasons: environmental adaptation, and internal coordination.
   ii. One example of a biological rhythm is the circadian rhythm, which regulates the sleep-wake cycle in many organisms.
   iii. If illustrating the circadian rhythm, the x-axis would represent time (24-hour clock), and the y-axis would represent a biological output, such as hormone levels or activity.
   iv. One way to test whether your measurements are rhythmic is to use statistical methods, such as autocorrelation analysis, periodogram analysis, or Fourier analysis, to detect significant periodicity in the data.

5. DNA carries genetic information and is made up of nucleotides, while proteins are functional molecules that perform various roles in the cell and are made up of amino acids.

6. One advantage of using baboon data is that their physiology and anatomy are more similar to humans than rodents. One advantage of using mice data is that they are more readily available and easier to manipulate genetically.

7. The surgically implanted cuff around the base of the aorta will have a lower amplitude of change in pressure measurements moment to moment because it measures blood pressure directly from the aorta, which is the body's main artery, and therefore experiences less fluctuation in pressure compared to the wrist, which is further from the heart and experiences more variation.

8. Three different things that might change in a neuron and affect behavior are:

   1. The strength of synaptic connections between neurons, which can be altered through processes such as long-term potentiation and depression.
   2. The expression of neurotransmitters, which can be regulated through gene expression and can influence the transmission of signals between neurons.
   3. The activity of ion channels, which can affect the excitability of neurons and influence their firing patterns.

9. One broad category of biomedical problem that data science might be able to help with is the development of personalized medicine. One major hurdle to achieving that ambition is the lack of high-quality data and privacy concerns surrounding patient data.

10. One reason it is hard to detect Parkinson's Disease earlier is that the symptoms are not specific to the disease and can overlap with other conditions. Additionally, the onset of symptoms can be gradual and easily overlooked, making it difficult to diagnose until the disease has progressed significantly.