Compulsory task 1

State whether the **mean**, **median**, or **mode** would be useful in the following scenarios:

- You are doing population statistics. You are asked to give an estimate of the *typical* income of a single person in the country. There is one snag: wealth distribution is out of whack, and 10% of the population holds 70% of the nation's wealth.

  The income distribution in this case is highly skewed, meaning that a small percentage of the population (10%) hold most of the wealth, while the majority earn much less.

  Using the median (middle value when sorted) would be the best approach to give a sense of a typical person income because the median is less sensitive to extreme values.

  The mean would overestimate what a typical person earns, because it includes outliers from the wealthiest individuals.

- You are running a restaurant, and you are reviewing your menu. You have a list of all orders over the last six months. You are trying to find out which item you should keep based on what customers seem to like the most.

  In this case using the mode (most frequently ordered dish) will directly tell us what customer prefers from our menu.

- You have been buying electricity once a month for the first six months of the year. You are trying to budget your electricity for the rest of the year and therefore need to estimate how much you will spend for the remainder of the year.

  Electricity consumption is normally stable over time unless there are major seasonal changes. In this case the mean (average cost) is useful for predicting future expenses – especially given that we have a good sample size of data (half a year) – because it smooths out small fluctuations, helping us to budget electricity for the rest of the year

- You work in healthcare insurance. You are asked to provide an estimate of the typical amount of money spent on healthcare. This is taking into account the fact that there are a few people who spend a large amount of money on medical healthcare due to major issues.

  In this case the mean would be inflated by the few people requiring expensive treatments. Therefore, the median (middle value) is a better measure of what most people spend on healthcare.

State whether you would use **variance** or **standard deviation** to inform the following decisions:

- You are choosing a new Internet provider. You find two providers with the same **mean** speed, but you want to have a more stable connection. You get a list of all reported speeds over the last month and are trying to find the provider that doesn't move too much from the mean value.

  The standard deviation measures the average deviation from the mean in internet reported speeds. Hence, a lower standard deviation means more stable speeds.

- You are going on holiday to Mauritius. You need to find a shuttle from the airport to your hotel, but you are worried about being overcharged or undercharged (being undercharged might mean that you get unreliable transport). You get a list of all available shuttle service prices and need to find out which services, if any, are overcharging or undercharging.

  In this case, while variance gives an idea of the spread of prices, standard deviation is more interpretable because it is in the same units (currency). A higher standard deviation indicates greater variability in prices, which helps identify services that overcharge or undercharge relative to the average.