

## Modeling

- Column B has the valuation of the companies
- Please choose a standard framework and run a model based on the data – Using the 19 factors to forecast whether the company is likely to be a “unicorn” (i.e. have a valuation  $\geq \$1\text{bn}$ )
- The output of the model should be a % - % likelihood that this company has a valuation  $\geq \$1\text{bn}$ .
- If the valuation of a company in Column B is n.a., you can assume it is  $< \$1\text{bn}$
- Please choose 5 most important/relevant factors out of the 19 factors – i.e. **the final model should only have 5 factors.**
- For the factors, if a factor is not available for a particular company, it is shown as 0 so you might have to clean up 0 with NA
- Which framework did you choose and why
- 5 factors selected and why
- Final model and relevant significance measurement (e.g.  $R^2$ , p-value, precision, fl)
- Summary takeaways (e.g. is the model statistically significant? What were your top constraints? How would you ideally like to improve it?)