1. Turbit primarily deals with time-series data generated by sensors for temperature, pressure and other quantities. Which modern time-series forecasting approach would you choose to model such data? Describe its advantages and disadvantages. How would you implement it using tensorflow?

    - I would strongly consider using Recurrent Neural Networks (RNNs) or models based on RNNs, such as Long Short-Term Memory models (LSTMs).
        - Advantages:
            - The recurrent nature of such models lends itself to memorize both short-term and long-term dependencies simultaneously.
            - If thousands of sensors are analysed simultaneously, such models are more likely to correctly infer any high-dimensional patterns, i.e. patterns that can't be explained by the output of several sensors, but can only be explained by a large fraction of the available sensors.
        - Disadvantages:
            - Such models require much more training data and training resources (compute and time) than traditional statistical models.
            - They are much less interpretable than traditional statistical models.
            - If the dimensionality of the problem (e.g. number of sensors) isn't high or one doesn't expect strong interactions between all/most of the sensor outputs, then such models might be overkill.
    - I don't' have any practical experience with tensorflow but they have built-in classes for both LSTM layers and RNN layers, so one could quite easily build a first prototype, and the fitting/predicting pipeline is the same as in sklearn.

2. We often deal with data gaps. What approaches can you use to deal with missing data in the context of time-series? What are the advantages and disadvantages?
    - Removing periods of time with missing data from training set
        - Advantages:
            - Doesn't require making assumptions on the most reasonable way to fill missing data, potentially preventing bias.
        - Disadvantages:
            - If missing data is spread out across the majority of the dataset, then this might shrink the training dataset to a degree that negatively affects model performance.
            - If the data "gets lost" in a systematic way that is relevant to the prediction, then removing it will lead to biased model predictions. For instance, if a sensor malfunctions and doesn't register data when its temperature is too high, but temperature also predicts wind turbine power, then removing datapoints where the sensor doesn't register data could lead to a biased prediction of wind turbine power.
    - Interpolating between existing values (e.g. linearly)
        - Advantages:

        - Preserves the trend of the data
    - Disadvantages:
        - Requires an assumption on the form of interpolation, which might be wrong if e.g. the data is discrete or known to suddenly change in a stepwise manner.
- Predict the missing values using the other (non-missing) features or previous data points
    - Advantages:
        - Utilizes the correlations between features to provide a potentially more informed guess of missing values than interpolation
    - Disadvantages:
        - Can increase chance of overfitting, as the training data is "used twice" – once in the imputation of missing values, and once in the training.

3. When reporting turbine anomalies, customers usually want to know the underlying reason for the unusual data points. Imagine that a thermometer in the gearbox of a turbine suddenly started reporting temperatures ten degrees higher than expected. List possible causes for such a pattern and how Turbit could distinguish between them?

    - Accident that actually led to an increase in temperature. Detection method: investigate other sensors (temperature, vibration, power) to look for extreme anomalies.
    - Sensor malfunction. Detection method: look for a lack of anomalies in other sensors, or a repetition of this sudden change in the past for this particular sensor.
    - Environmental factors, e.g. sudden heat wave. Detection method: inspect and use external data, such as weather forecasts.