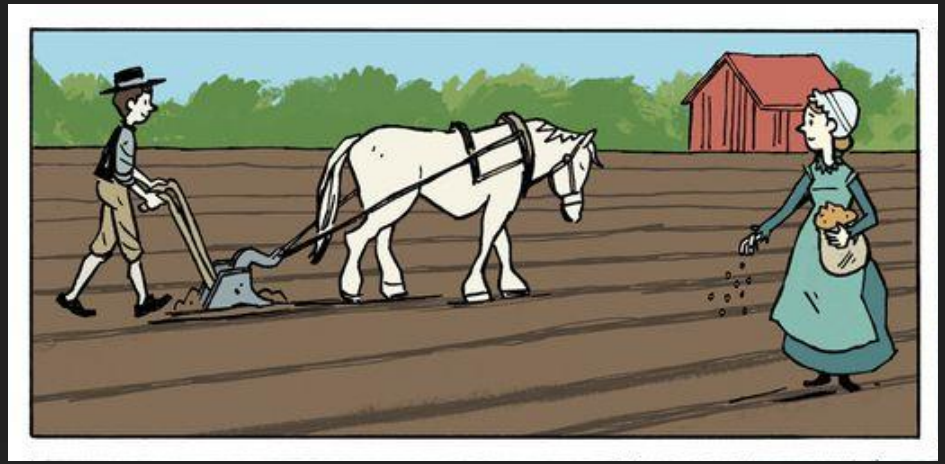
/ Time data

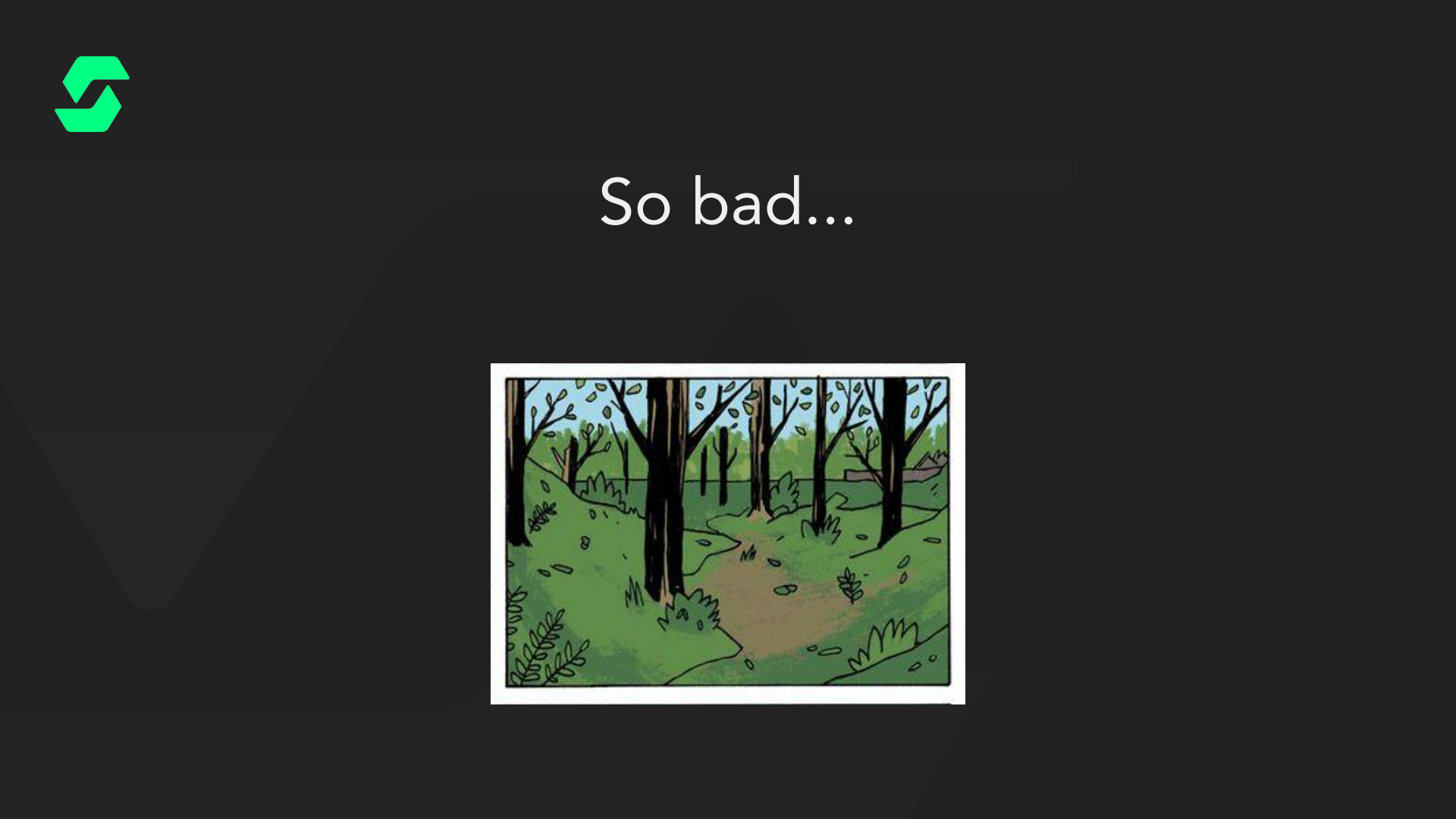
Once upon a time there was a farming couple who lived happily from their harvest.



They left the farm because they did not see the drought and high temperatures coming.

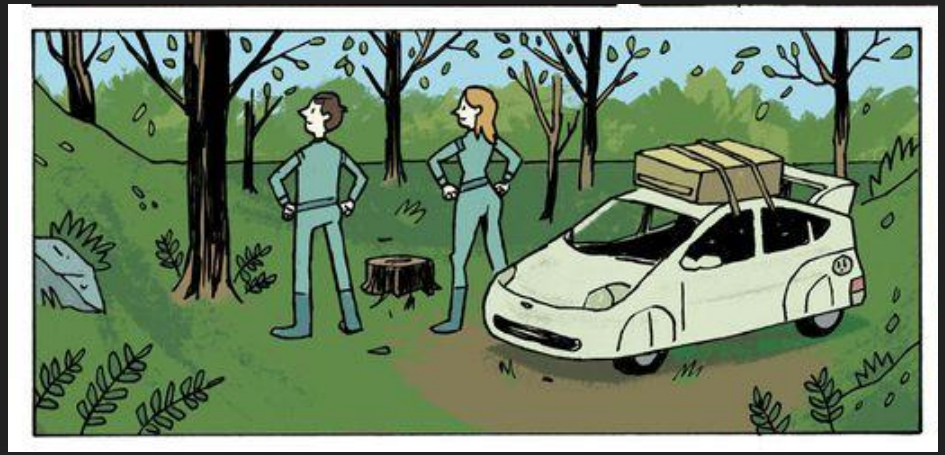
Years went by and the whole farm went to waste

Very bad although there were good seasons with stable temperatures.

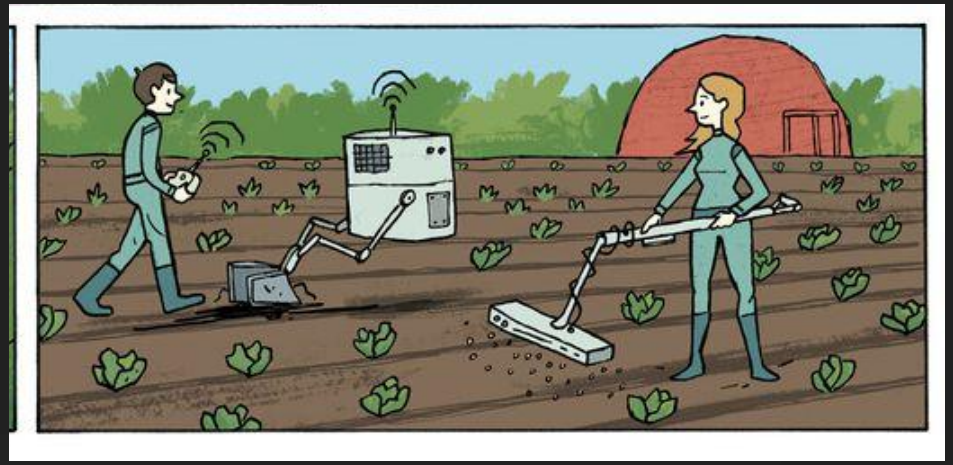


So bad...

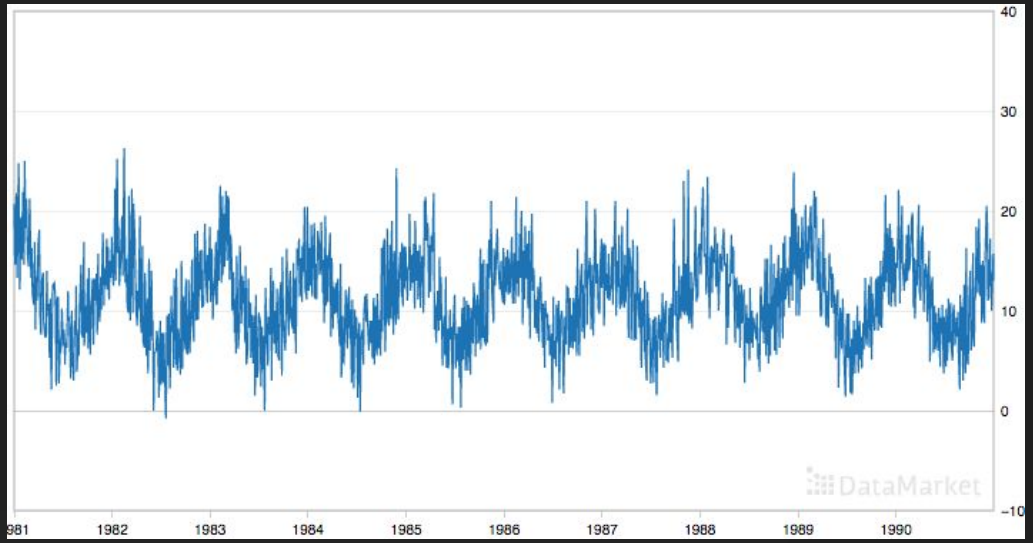
Until one day, the farmers' grandchildren came along and wanted to continue their grandparents' life.



Modernized the farm with robots and installed thermometers for predictive temperature modeling



Here is the temperature collected from the robot



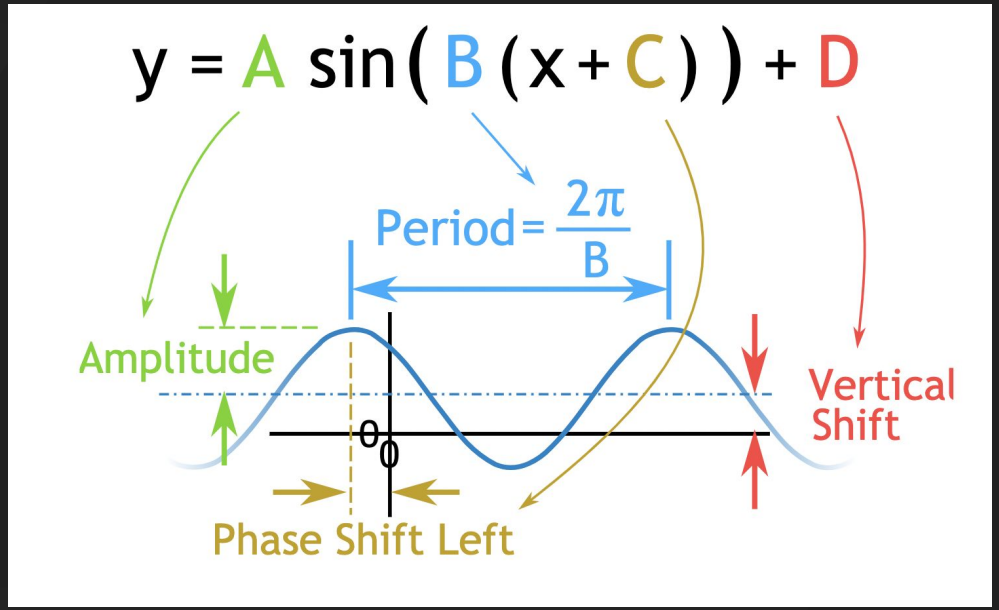
Your task is to help the new farmers by predicting the coming temperature

1. Extract year, month and day_of_month
2. Get lags features → Pandas shift()
    a. The temperature of day before, and 2 days ago
3. Rolling Window Statistics → Pandas rolling()
    a. Mean, min & max of 7 previous days
4. Try new features
5. Select the year 1990 as a validation set
6. Try some some ML models to predict 1990
    a. You can use the RMSE as the metric

Ideas of simple ML models (baselines)

- Use the previous year
- Average of all the prev. Years
- Average by months of the previous year.
- Smooth signal / moving average
  - Try different windows sizes and types (uniform, gaussian, triangular, hanning)
  - Use the smoothed previous year for prediction
- Polynomial
  - 1) Generate polynomials (day of the year squared or cubed)
  - 2) Fit a linear model
- [OPTIONAL] Use the Fourier Transform (FFT)
  - 1) It gives us the frequency, but it give us also the phase
  - 2) Build a general wave fn. With frequency, phase, amplitud & vertical shift.

Wave function




Strive
SCHOOL