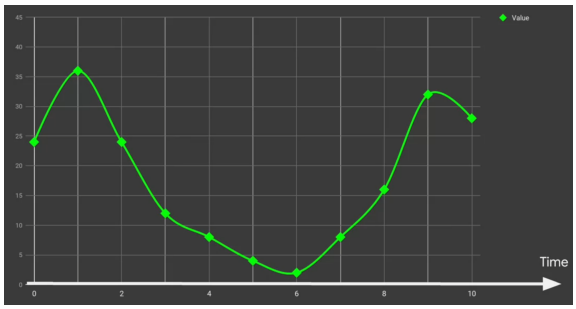
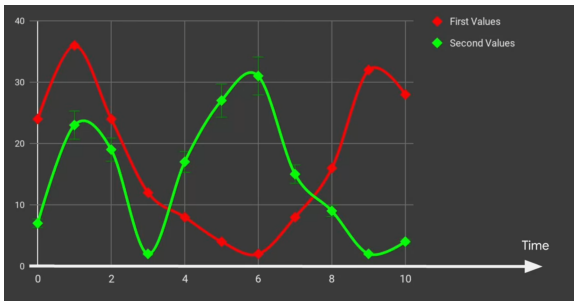
Week 1

- Time series are everywhere, some examples are:
  - Stock prices
  - Weather forcast
  - Historical trends (Moore's Law)
- Time series is typically defined as an order sequence of values equally spaced over time
- Single value at each time step is called univariate
- Time series that have multiple values at each time step is called multivariate



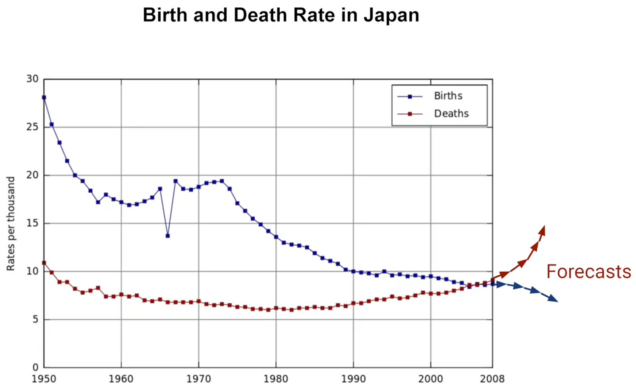
Univariate



Multivariate

ML and Time series

- Examples to apply ML to time series:
  - Birth and death rate in Japan



- Project back to the past to see how the data progressed
  - This processed is called **imputation**
- Analyze sound waves for speech recognition

Common Patterns in Time Series

- There are upward upward facing trend
- Another concept is seasonality, it follows a very distinct pattern
- Some time series can have trend and seasonality
- There's also values that are a complete set of random values producing what's called white **noise**
- Another concept is auto-correlated time series
  - It correlates a delayed copy of itself often called a lag
  - Spikes on graphs are called innovations which cannot be predicted using past values
- Time series in real life most often has:
  - Trend
  - Seasonality
  - Autocorrelation
  - Noise
- **Non-Stationary Time Series** when trends are not always the same

Plotting Different Graphs

- The code below plots all the graphs discussed earlier

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/22d00a4c-53f9-48a4-849d-893b4f19f0ca/plotting_different_time_series.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/261e8af9-4287-4d1a-9be5-fdf3479506eb/plotting_different_time_series.py

Train Time series

- To train a time series, typically, you split the data into 3 parts
  - Training Period
  - Validation Period
  - Test Period
- This is called **fixed partitioning**
- Another way is called Roll-Forward Partitioning
- This method uses the splits the data in days or weeks and predict the data for the next day or week
  - Basically doing fixed partitioning multiple times

Metrics

- Metrics evaluates how wrong the prediction is therefore measuring performance
- Here are a few metrics
  - Mean Square Error

```
mse = np.square(errors).mean()
```

  - root mean squared error

```
rmse = np.sqrt(mse)
```

- Mean absolute error

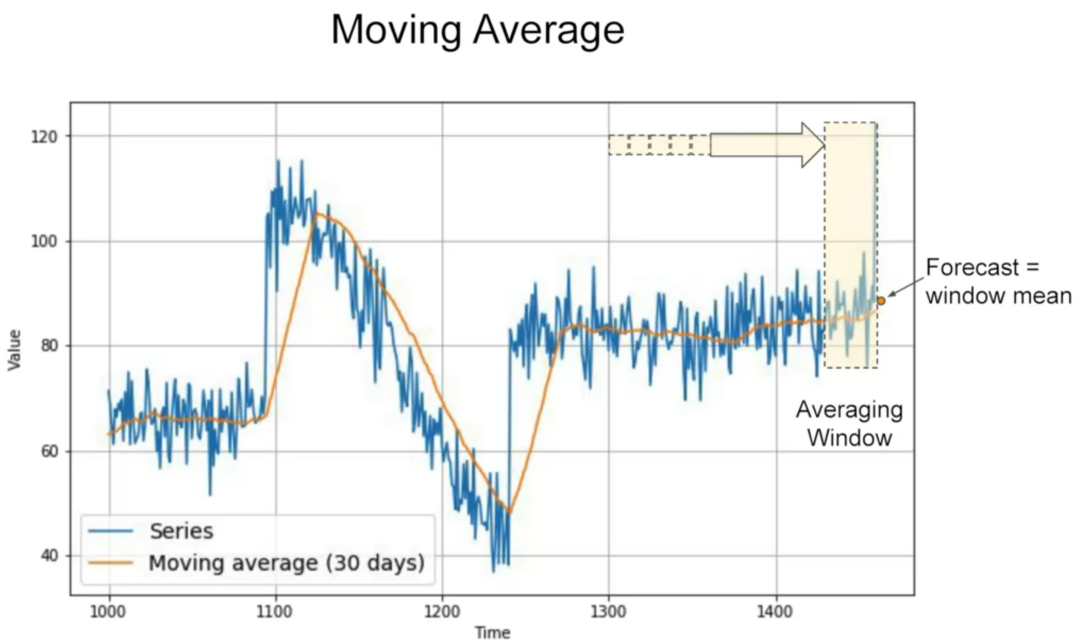
```
mae = np.abs(errors).mean()
```

- mean absolute percentage error

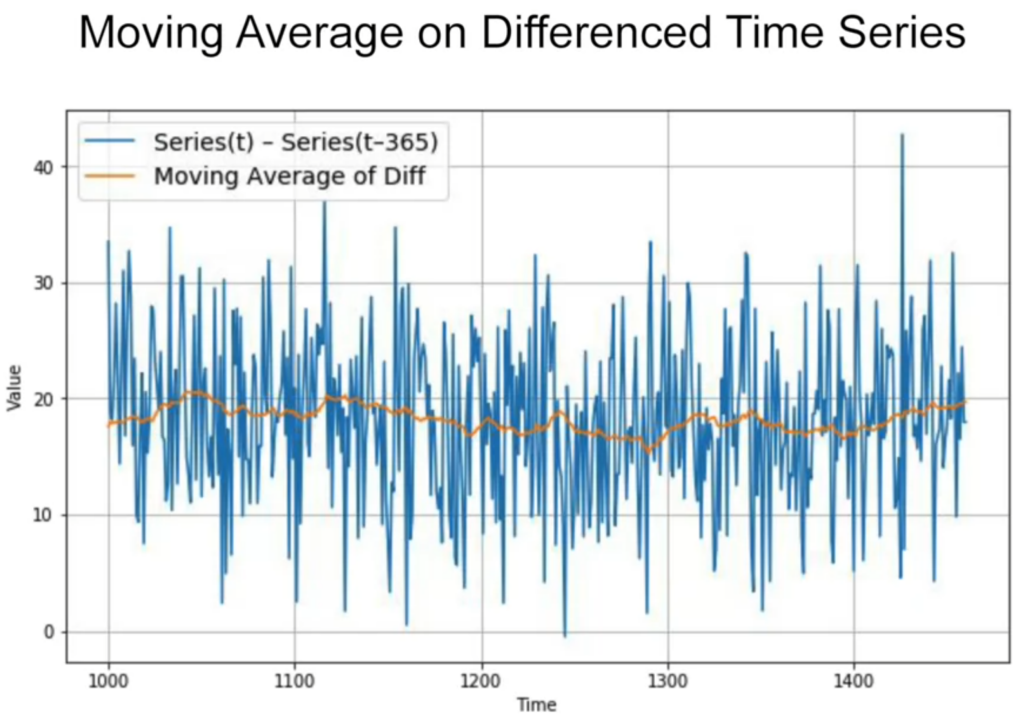
```
mape = np.abs(errors / x_valid).mean()
```

Moving Average and differencing

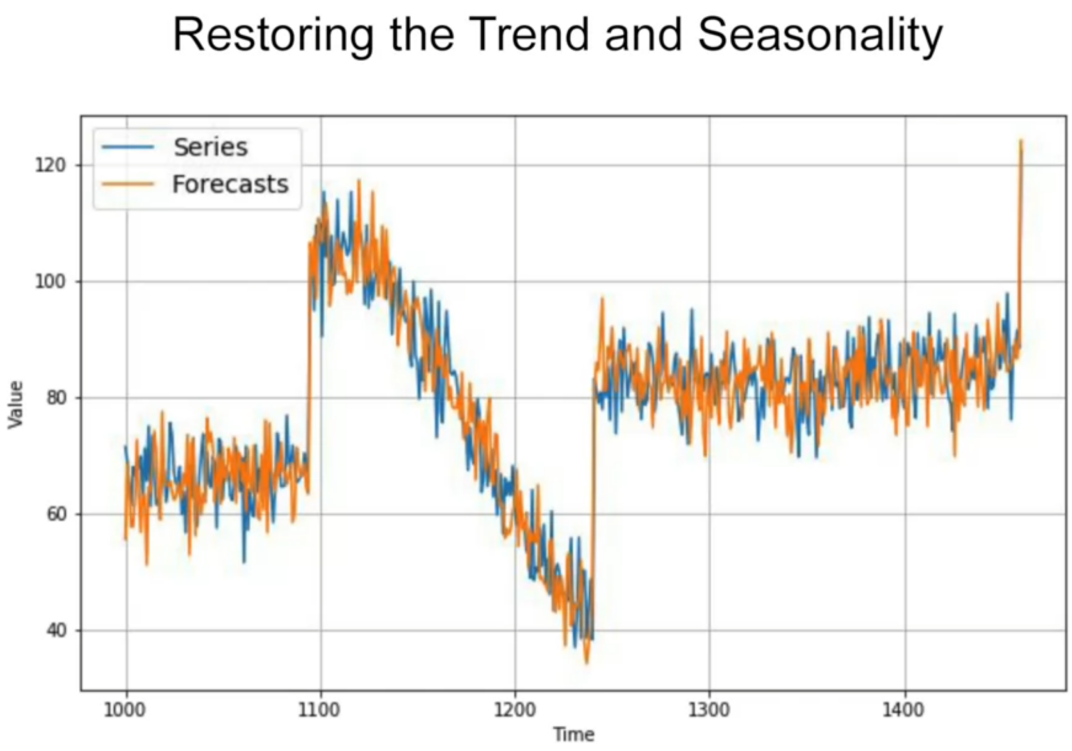
- A common forecasting method is the moving average
  - In the graph below, the yellow line is a plot of the average of the blue values over an fixed period called a fixed window
  - Does not anticipate trend of seasonality
  - Depending on the time, could be worse than naive forecast



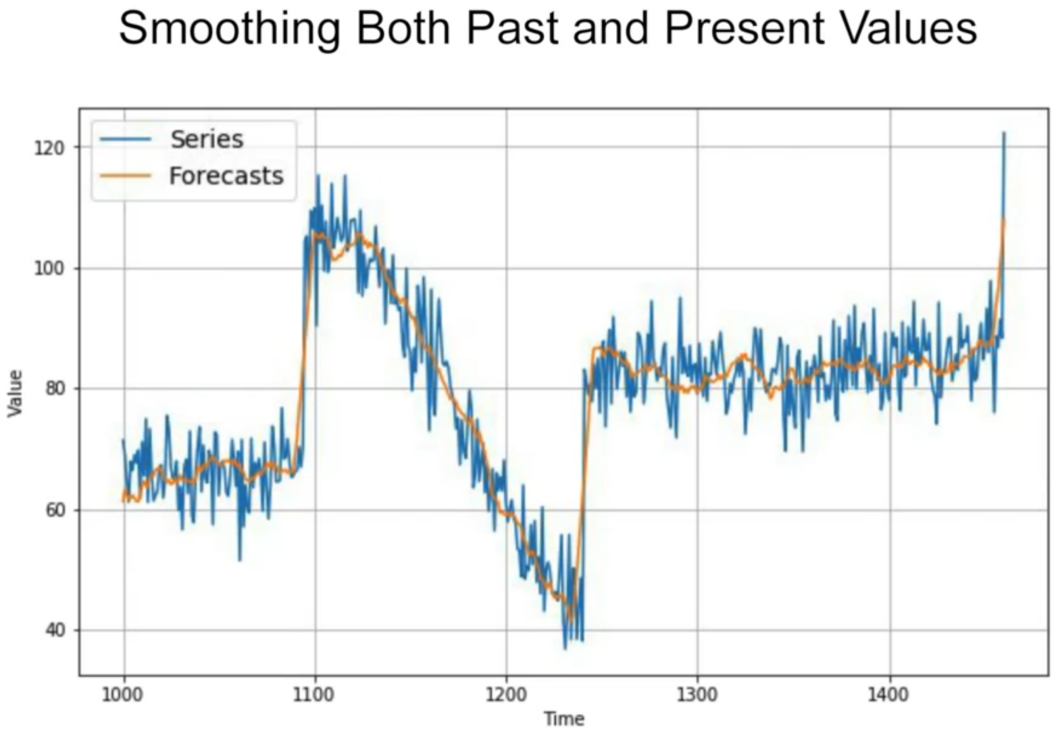
- An way to solve the issue with moving average, use the method differencing
- Differencing study the difference between the time instead of the time itself
- The graph below is the difference time series



- To move it back to original time series, just add t-365



- To remove the noise, use a moving average on the past data too



Forecasting

- Code for forecasting in TensorFlow

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/ae28d60a-c944-4e81-9d26-3013f70bf11a/forecasting.py

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/e86d6059-76cb-407f-90ef-9390ce18c40b/forecasting.ipynb

Exercise

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/7f6e251d-7291-48b8-9680-254fff9508c3/exercise_1_w1.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/dbb62064-cde7-4f4f-afb1-1292e5c3c117/exercise_1_w1.py

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/0cc04c05-d519-47dc-945d-bcf69a9dca3d/time_series_quiz_1.pdf