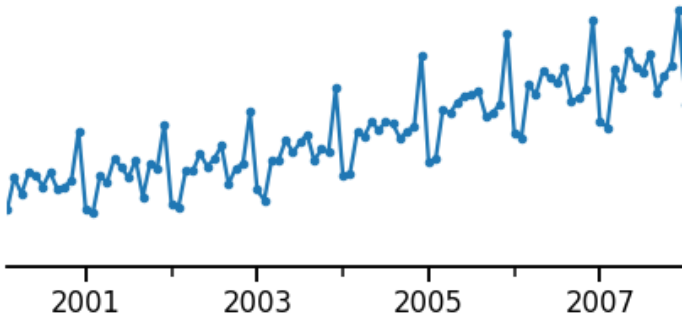
Datasets, Features and Target

Extracting Features for Forecasting

Single Time Series

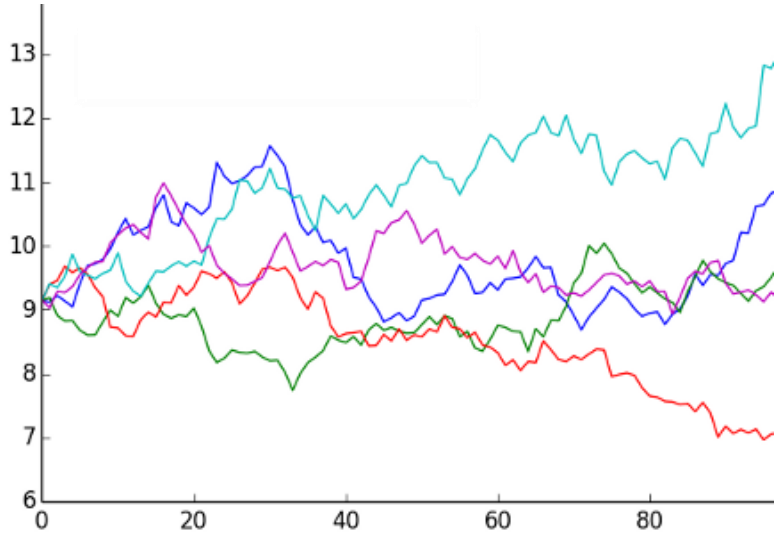


Time	Sales
30/03/20	200
31/03/20	220
01/04/20	230
02/04/20	235

In most basic time series examples we are shown a single column time series (plus time index)

- Sales (one channel), energy consumption (one household)
- **Real life datasets are more challenging**

Multiple Time Series



Time	Sales UK	Spain	Germany	Italy
30/03/20	200	100	330	120
31/03/20	220	120	300	135
01/04/20	230	150	335	133
02/04/20	235	175	340	200

Often, we want to forecast more than 1 time series, simultaneously.

- Sales across multiple channels, or in different countries.
- Energy demand per city.

Datasets ➔ variables + targets

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Often, we have more data accompanying the target series ➔ predictor variables.

- Additional information about the time and situation in which the events developed.

Predicting future values

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

In time series forecasting, we want to predict future events, based on past data.

Predicting future values - Features

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day
4	3
4	4

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

For some features, we know the values in the future.

Predicting future values - Features

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day	Temp	Rain	inflation
4	3	?	?	?
4	4	?	?	?

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

For some features, we do NOT know the values in the future. We need proxies.

Predicting future values - Features

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day	Temp	Rain	inflation	Ads
4	3	?	?	?	1
4	4	?	?	?	1

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

Through additional features, we can simulate plausible scenarios.

- What would happen if we launch an ads campaign?

Predicting future values with predictors

Predictor variables

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Target variables

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day	Temp	Rain	inflation	Ads
4	3	?	?	?	1
4	4	?	?	?	1

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

In some cases, we can predict the target based solely on predictor variables

- Energy production = f(temperature, sun, wind, #solar panels, #wind turbines)

Predicting future values with past data

Predictor variables

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Target variables

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day	Temp	Rain	inflation	Ads
4	3	?	?	?	1
4	4	?	?	?	1

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

In some cases, we can predict the target based on past records of the target itself

- Sales, temperature, stock prices

Predicting future values – mixed models

Predictor variables

Month	Day	Temp	Rain	inflation	Ads
3	30	15	50	0.2	0
3	31	16	10	0.2	0
4	1	17	0	0.19	1
4	2	19	5	0.17	0

Target variables

Sales UK	Spain	Germany	Italy
200	100	330	120
220	120	300	135
230	150	335	133
235	175	340	200

Month	Day	Temp	Rain	inflation	Ads
4	3	?	?	?	1
4	4	?	?	?	1

Sales UK	Spain	Germany	Italy
?	?	?	?
?	?	?	?

In some cases, we use both features and target to predict future events

- Mixed models, panel data, dynamic regression

Summary

Datasets can be complex: predictors + target variables.

Values of predictors can be known or unknown at the time of forecast.

We can extract features from predictors and target series for forecasting.